

# Spectrum™ Technology Platform

Version 12.0

Administration Guide - AMI



# Table of Contents

## 1 - Getting Started

---

|   |    |
|---|----|
| Configuring a New System                        | 5  |
| Configuring a New Spectrum™ Technology Platform |    |
| AMI Instance                                    | 6  |
| Accessing Management Console                    | 7  |
| Starting and Stopping the Server                | 22 |
| Installing the Client Tools                     | 24 |
| Network Ports                                   | 24 |

## 2 - Security

---

|   |    |
|---|----|
| Security Model                                    | 28 |
| Users   | 29 |
| Roles   | 34 |
| Access Control                                    | 45 |
| Security for the Location Intelligence Module     | 48 |
| Limiting Server Directory Access                  | 53 |
| Configuring HTTPS Communication                   | 55 |
| Web Service Authentication                        | 58 |
| Using LDAP or Active Directory for Authentication | 63 |

## 3 - Data Sources

---

|                              |     |
|------------------------------|-----|
| Introduction to Data Sources | 71  |
| Connection Types             | 75  |
| Deleting a Connection        | 124 |

## 4 - Spectrum Databases

---

|  |     |
|--|-----|
| Introduction to Spectrum Databases       | 126 |
| Adding a Spectrum Database               | 126 |
| Database Pool Size and Runtime Instances | 126 |

|                              |     |
|------------------------------|-----|
| Deleting a Spectrum Database | 129 |
|------------------------------|-----|

## 5 - Services

---

|                       |     |
|-----------------------|-----|
| Spectrum Services     | 131 |
| External Web Services | 139 |

## 6 - Flows

---

|                                       |     |
|---------------------------------------|-----|
| Configuring Flow Defaults             | 153 |
| Scheduling Flows                      | 163 |
| Viewing Flow Status and History       | 166 |
| Triggering a Flow with a Control File | 168 |
| Command Line Execution                | 172 |

## 7 - Performance

---

|                              |     |
|------------------------------|-----|
| Performance Tuning Checklist | 191 |
| Monitoring Performance       | 209 |

## 8 - Monitoring

---

|   |     |
|---|-----|
| Email Notification                            | 215 |
| Audit Log                                     | 218 |
| System Log                                    | 220 |
| Logging the Record that Caused a Flow to Fail | 222 |
| Transaction Limit Warnings                    | 222 |
| Viewing Version Information                   | 223 |
| Viewing and Exporting License Information     | 223 |

## 9 - Backup and Restore

---

|                            |     |
|----------------------------|-----|
| Scheduling Backups         | 226 |
| Creating a Backup Manually | 230 |

|                    |     |
|--------------------|-----|
| Restoring a Server | 231 |
|--------------------|-----|

## 10 - Administration Utility

---

|   |     |
|---|-----|
| Getting Started with the Administration Utility | 234 |
| Using a Script with the Administration Utility  | 235 |
| Data Sources                                    | 236 |
| Dataflows                                       | 245 |
| Enterprise Routing Module                       | 251 |
| Folders   | 270 |
| Jobs  | 273 |
| Location Intelligence Module                    | 282 |
| Match Rules                                     | 288 |
| Notification                                    | 291 |
| Open Parser Cultures                            | 295 |
| Open Parser Domains                             | 296 |
| Performance Monitor                             | 297 |
| Permissions                                     | 301 |
| Process Flows                                   | 301 |
| Remote Servers                                  | 310 |
| Roles   | 314 |
| Search Indexes                                  | 318 |
| Services  | 322 |
| Spectrum Databases                              | 326 |
| System  | 362 |
| Tables  | 370 |
| Tokens  | 373 |
| User Accounts                                   | 375 |

## 11 - Clustering

---

|  |     |
|--|-----|
| Clustered Architecture                                     | 381 |
| Using Enterprise Designer with a Cluster                   | 383 |
| Starting a Cluster   | 384 |
| Stopping a Cluster   | 385 |
| Removing a Node from a Cluster                             | 385 |
| Managing a Cluster for the Location Intelligence<br>Module | 386 |

## 12 - About Spectrum™ Technology Platform

---

|  |     |
|--|-----|
| What Is Spectrum™ Technology Platform?     | 394 |
| Enterprise Data Management Architecture    | 395 |
| Spectrum™ Technology Platform Architecture | 399 |
| Modules and Components                     | 404 |

# 1 - Getting Started

## In this section

---

|  |    |
|--|----|
| Configuring a New System                                     | 5  |
| Configuring a New Spectrum™ Technology Platform AMI Instance | 6  |
| Accessing Management Console                                 | 7  |
| Starting and Stopping the Server                             | 22 |
| Installing the Client Tools                                  | 24 |
| Network Ports  | 24 |

## Configuring a New System

When you first install Spectrum™ Technology Platform there are few things you should do to ensure that your system has a basic level of security as well as access to the data you want to process through Spectrum™ Technology Platform.

1. Change the password for the admin user.

**Important:** You should change the admin password immediately after installing Spectrum™ Technology Platform to prevent unauthorized administrative access to your system.

- a) In a web browser go to this URL:

`http://server.port/managementconsole`

Where *server* is the server name or IP address of your Spectrum™ Technology Platform server and *port* is the HTTP port used by Spectrum™ Technology Platform. By default, the HTTP port is 8080.

- b) Log in using the default administrative credentials:

User name: admin

Password: admin

- c) Go to **System > Security**.

- d) Check the box next to the **admin** account then click the Edit button .

- e) In the **New password** field, enter a new password. Enter it again in the **Confirm password** field.

- f) Click **Save**.

2. Create users and roles as needed.

For more information, see [Adding a User](#) on page 30.

3. Specify which folders on the Spectrum™ Technology Platform server you want to allow users to access.

For more information, see [Limiting Server Directory Access](#) on page 53.

4. Decide if you want to allow Basic authentication for web service requests to your Spectrum™ Technology Platform server, or if you want to require token authentication. If you want to require token authentication, disable Basic authentication. For more information, see [Disabling Basic Authentication for Web Services](#) on page 60.

5. Define database resources if applicable.

To determine if you need to define database resources, go to **Resources > Spectrum Databases**. If you do not see the **Spectrum Databases** menu then you do not need to define database resources.

6. Define the databases, file servers, and other data sources that you want to access from Spectrum™ Technology Platform. To define data sources, go to **Resources > Data Sources**.
7. Configure scheduled backups of your Spectrum™ Technology Platform server so that you can restore your server in the event of a severe system failure or other disaster. For more information, see [Scheduling Backups](#) on page 226.

## Configuring a New Spectrum™ Technology Platform AMI Instance

After launching an instance using a Spectrum™ Technology Platform AMI, you must set the administrative password, configure basic security settings, and define connections to your data.

1. Connect to the new instance using Remote Desktop.
2. Set the password for the Spectrum™ Technology Platform "admin" account. This account is the administrator account for Spectrum™ Technology Platform.
  - a) On the instance, open Internet Explorer and go to:  
`http://localhost:8080/spectrum-ami`
  - b) Complete the form to set the password of the "admin" account.

**Note:** You are allowed three attempts to change the password. After three failed attempts Spectrum™ Technology Platform becomes locked. If this happens you must terminate the instance and create a new one. There is no way to unlock Spectrum™ Technology Platform once it is locked.

3. Install the license key.
  - a)
  - b)
4. Open Management Console by opening a web browser and going to `http://localhost:8080/managementconsole`.
5. Use Management Console to create users and roles.  
For more information, see [Adding a User](#) on page 30.
6. Use Management Console to specify which folders on the Spectrum™ Technology Platform server you want to allow users to access.  
For more information, see [Limiting Server Directory Access](#) on page 53.
7. Use Management Console connect to the databases, file servers, and other data sources that you want to access from Spectrum™ Technology Platform.

## Accessing Management Console

Management Console is the tool for administering Spectrum™ Technology Platform. Using Management Console, you can perform such tasks as:

- Manage users and security settings
- Define connections to data sources such as databases or web services
- Specify default settings for services
- Schedule job execution

To access Management Console:

1. In a web browser go to this URL:

`http://server:port/managementconsole`

Where *server* is the server name or IP address of your Spectrum™ Technology Platform server and *port* is the HTTP port used by Spectrum™ Technology Platform. By default, the HTTP port is 8080.

2. Enter a valid user name and password.

The administrative user name is "admin" and it has a default password of "admin".

**Important:** You should change the admin password immediately after installing Spectrum™ Technology Platform to prevent unauthorized administrative access to your system.

## Setting the Language and Region

The Management Console can be viewed in English, French, Japanese, Portuguese, and Spanish. By default, you will see Management Console in the display language set in your web browser's language setting. If you want to use the Management Console in a different language than the language set by your browser, follow these steps.

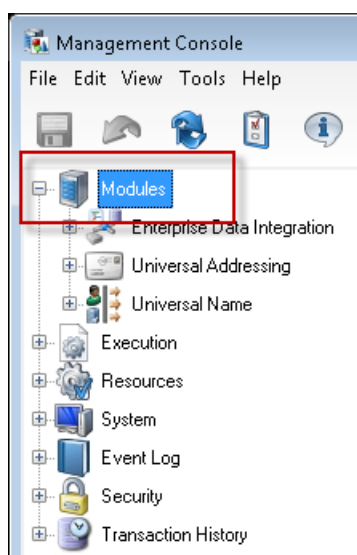
1. Log in to Management Console.
2. Click the user menu in the top right corner.
3. Select **Profile**.
4. In the **Language** field, choose the language you want.
5. In the **Country** field, choose your region. This setting controls the format to use when displaying dates, and times.
6. Click **Save**.

## Comparison of Windows and Web Versions of Management Console

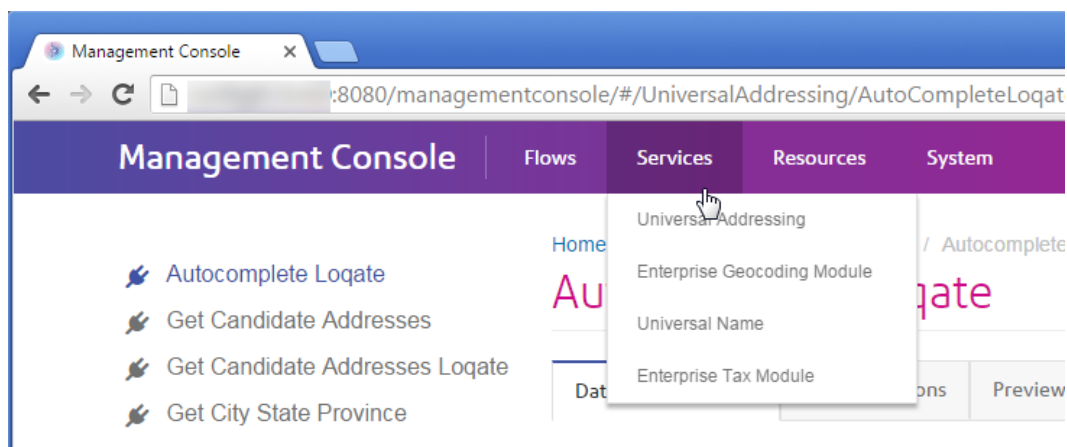
Management Console is a tool for administering Spectrum™ Technology Platform. Beginning in version 11.0, Management Console is no longer available as a Windows client. If you are familiar with the Windows version of Management Console from past releases, this information will help you get oriented in the web browser version.

### Modules

In the Windows client, the Modules section contained settings for service defaults and module-specific tools such as database resources.

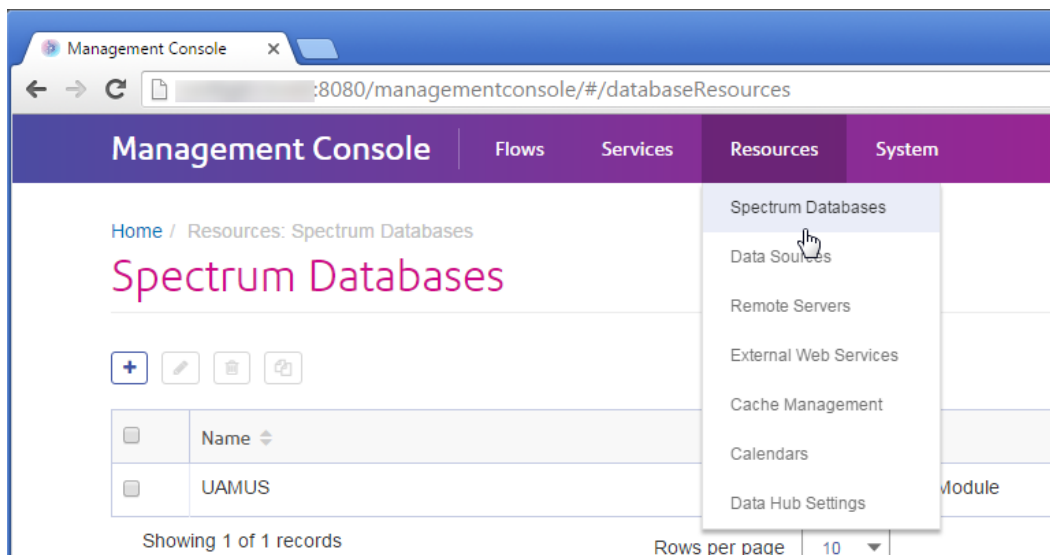


The web version of Management Console provides access to modules in two places. Module services are available under the **Services** menu:



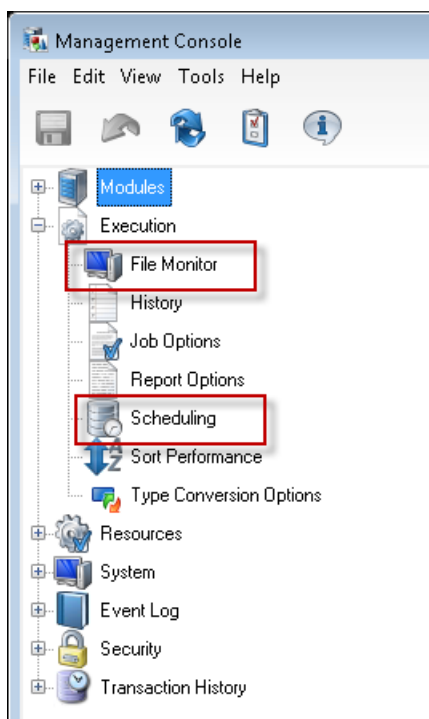
Database resources and other module tools are available under **Resources** menu:





### Execution - File Monitor and Scheduling

In the Windows client, File Monitor and Scheduling features are separate sections under Execution:



In the browser version of Management Console, these settings are available together under **Flows > Schedule**. This provides you with a complete picture of your jobs and process flow execution in one place.

Management Console

Home / Flows: Schedules

## Schedules

Filter

| Schedule Name       | User  | Trigger | Next Run | Last Run        |
|---------------------|-------|---------|----------|-----------------|
| AnotherScheduledRun | admin | ⌚       | -        | -               |
| ExampleSchedule     | admin | ⌚       | -        | 11/5/15 3:18 PM |

### Execution - History

In the Windows client, the History section listed the jobs and process flows that have run on your system. Execution history displays information such as when dataflows and process flows ran, whether they were successful, and the number of records processed.

Management Console

File Edit View Tools Help

- Modules
- Execution
  - File Monitor
  - History
  - Job Options
  - Report Options
  - Scheduling
  - Sort Performance
  - Type Conversion Options
- Resources
- System
- Event Log
- Security
- Transaction History

In the browser version of Management Console you can view execution history under **Flows > History**:

Management Console

Home / Flows: History

## History

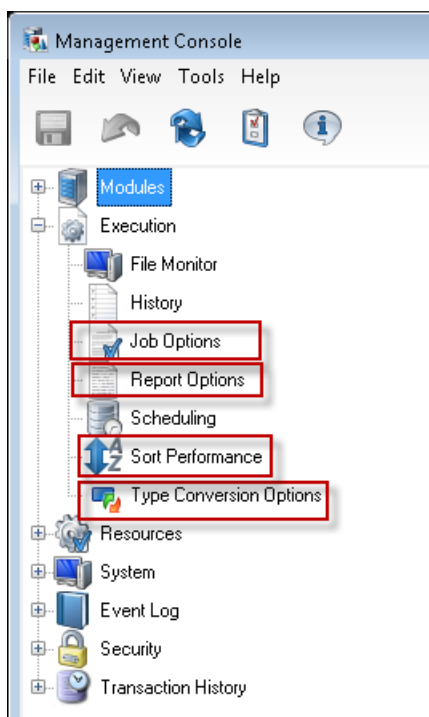
Flows Transactions

Started on 11/6/15

| ID | Name       | User   | Flow Type | Start Time      | End Time        | Duration | Status |
|----|------------|--------|-----------|-----------------|-----------------|----------|--------|
| 5  | ExampleJob | admin  | Job       | 11/6/15 8:46 AM | 11/6/15 8:46 AM | 00:00    | ⚙️     |
| 4  | ExampleJob | admin  | Job       | 11/6/15 8:44 AM | 11/6/15 8:45 AM | 00:16    | 🚫      |
| 3  | ExampleJob | system | Job       | 11/6/15 8:43 AM | 11/6/15 8:43 AM | 00:00    | ✅      |

### Execution - Options

In the Windows version of Management Console, the options that control the handling of malformed records, reports, sort performance, and data type conversion are contained in four different sections:



In the browser version of Management Console, these options are available together under **Flows > Defaults**.

Management Console | Flows | Services | Resources | System

Home / Flows: Defaults

## Defaults

Data Type Conversions | Sort Performance | Malformed Records | Reports

Use these settings to control the default data type conversion behavior for your system. ?

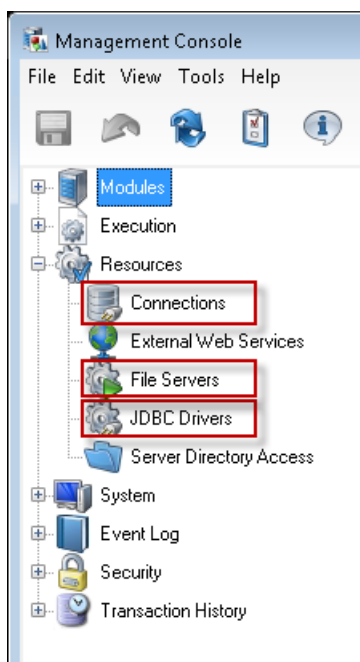
**String conversion format**

Locale

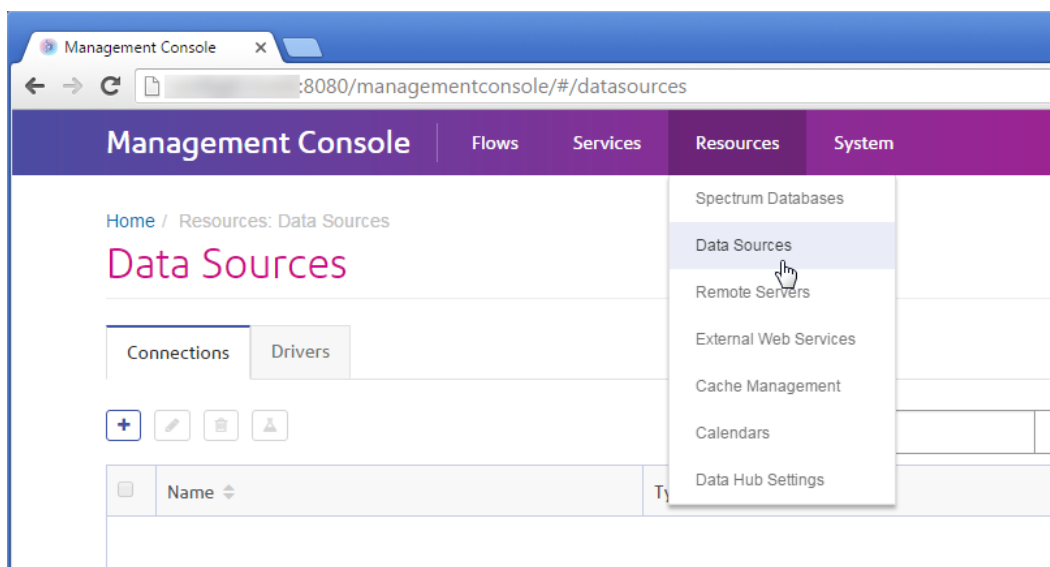
English (United States)

### *Resources - Connections, File Servers, and JDBC Drivers*

In the Windows version of Management Console, the settings for database connections, file server connections, and JDBC drivers were in separate sections under Resources:

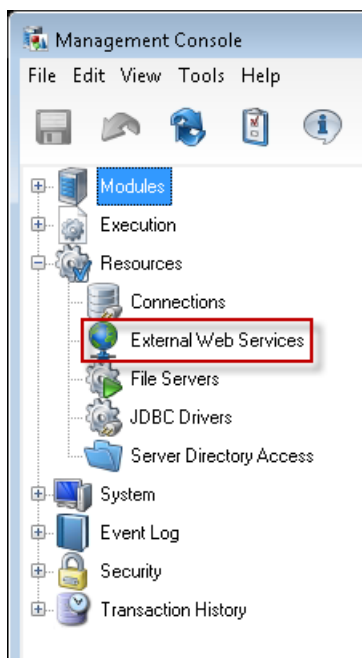


In the browser version of Management Console, these settings are available together under **Resources > Data Sources**.

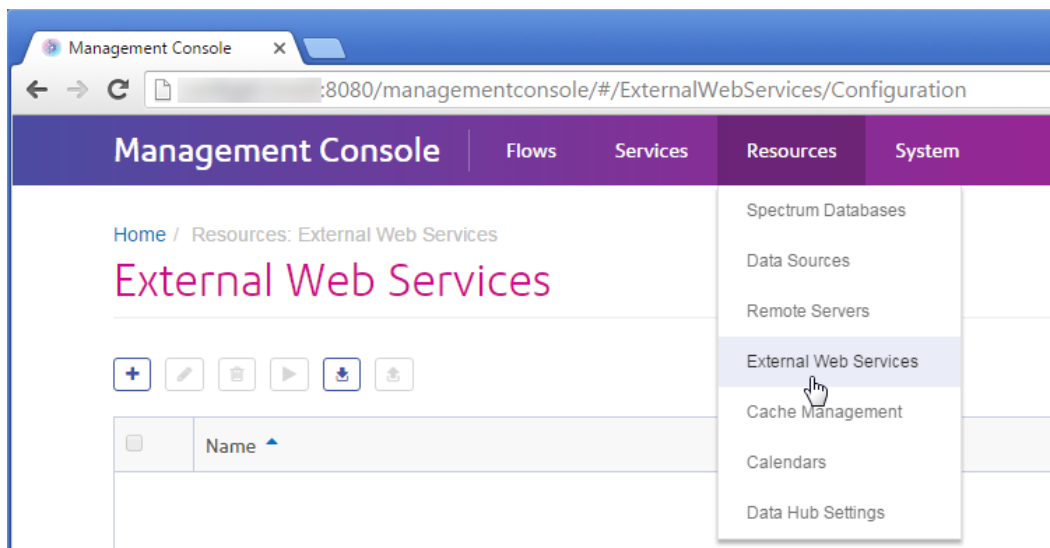


### Resources - External Web Services

In the Windows version of Management Console, the settings for external web services are available under the Resources node:

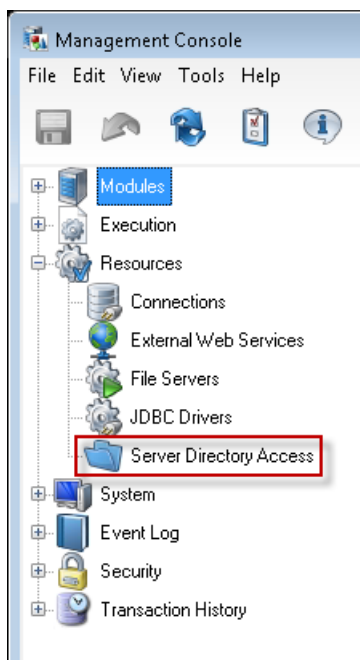


In the browser version of Management Console, external web services are available under **Resources > External Web Services**:

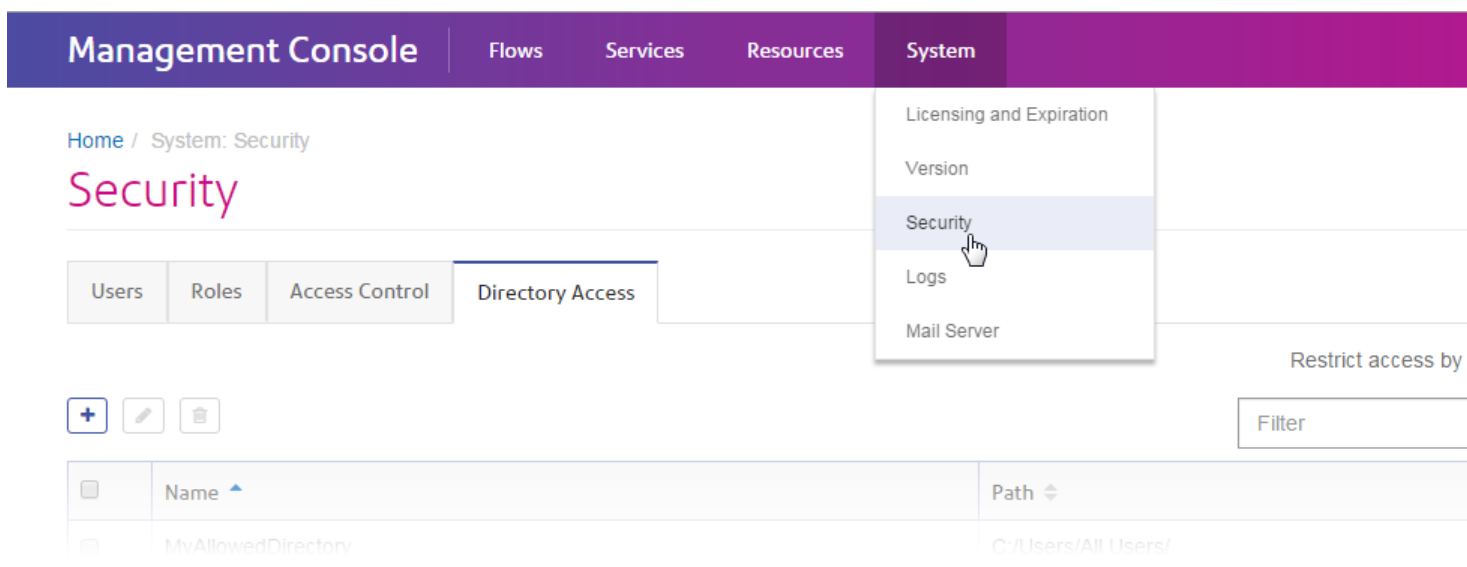


### Resources - Server Directory Access

In the Windows version of Management Console, the settings that restrict access to directories on the server are located under the Resources node:

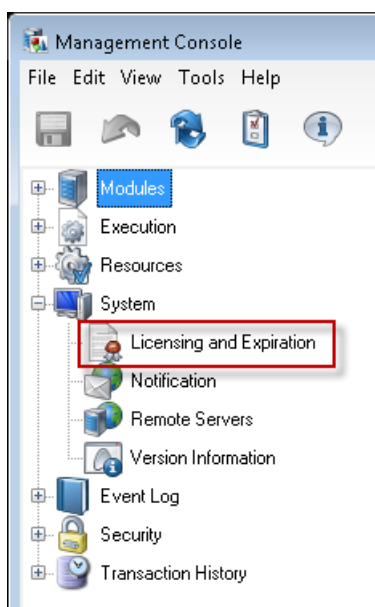


In the browser version of Management Console, these settings are under **System > Security**.

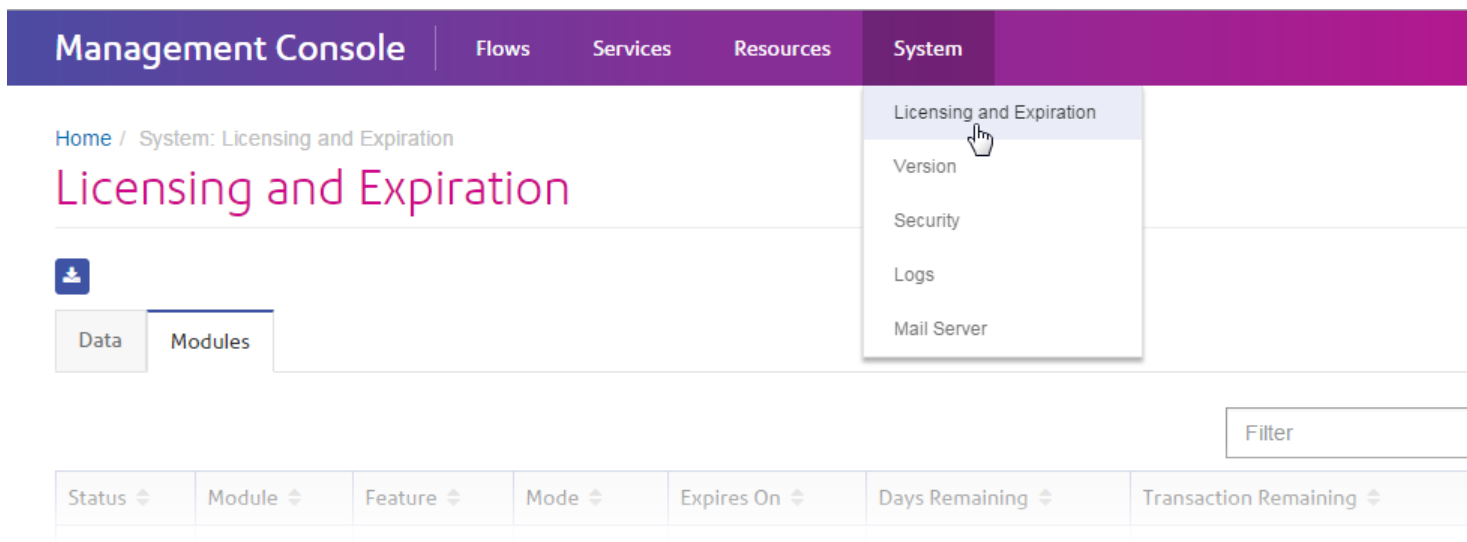


### System - Licensing and Expiration

In the Windows version of Management Console, information about your licenses and the expiration of licenses is available in the Licensing and Expiration section under System:

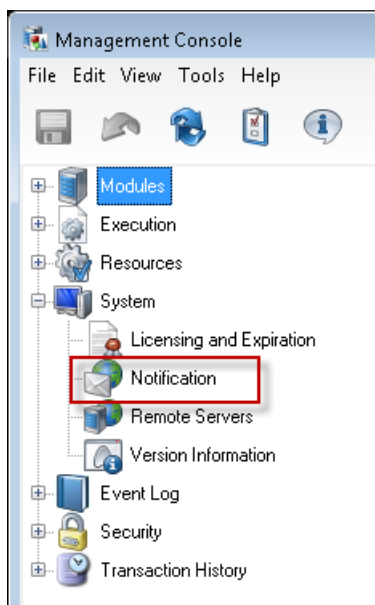


In the browser version of Management Console, this information is available under **System > Licensing and Expiration**.



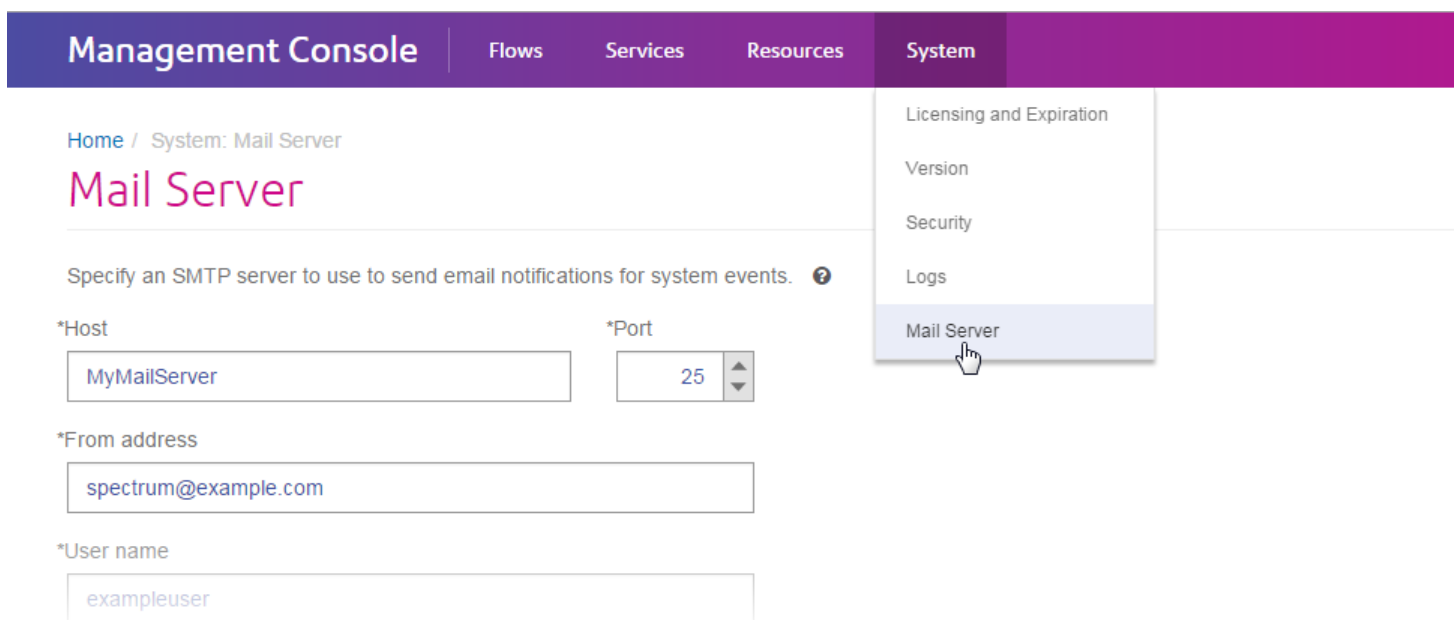
### System - Notification

In the Windows version of Management Console, the settings that control email notification are under System:



In the browser version of Management Console, these settings are in two different places. The settings for configuring an SMTP server are under **System > Mail Server**.





Management Console | Flows | Services | Resources | System

Home / System: Mail Server

## Mail Server

Specify an SMTP server to use to send email notifications for system events. ?

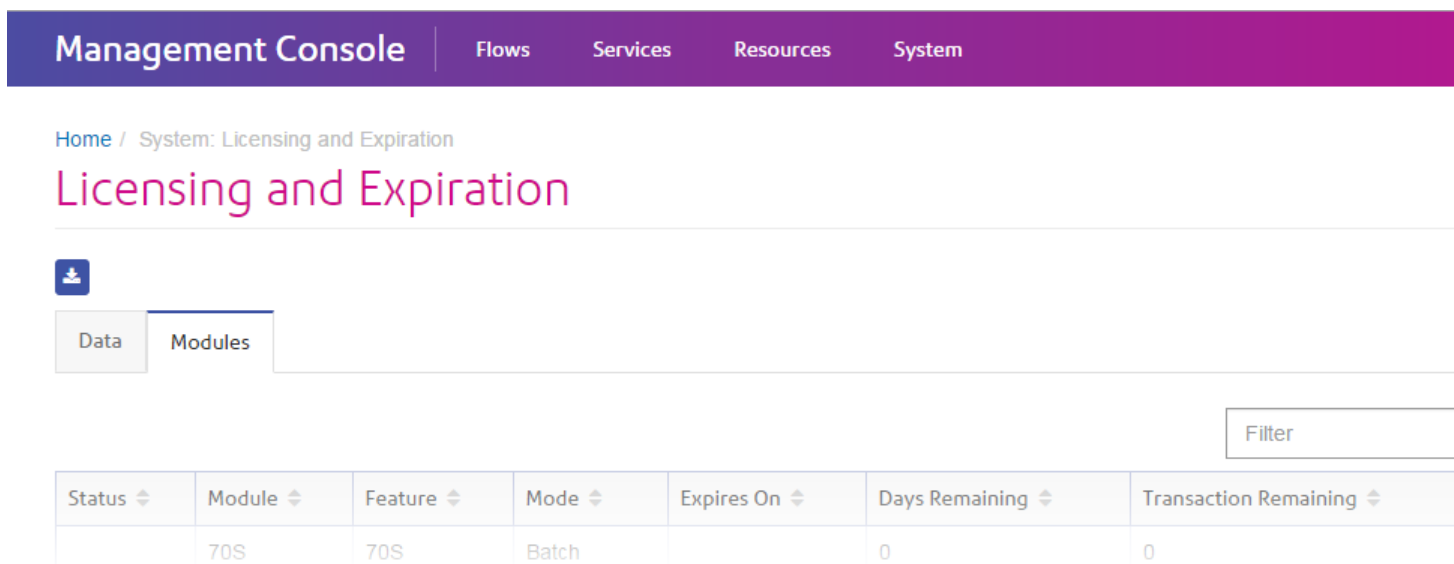
\*Host: MyMailServer      \*Port: 25

\*From address: spectrum@example.com

\*User name: exampleuser

System menu items: Licensing and Expiration, Version, Security, Logs, Mail Server (selected)

The settings for specifying which expiration notifications you want to receive and the recipients of notifications are under **System > Licensing and Expiration > Configure Notification**.



Management Console | Flows | Services | Resources | System

Home / System: Licensing and Expiration

## Licensing and Expiration

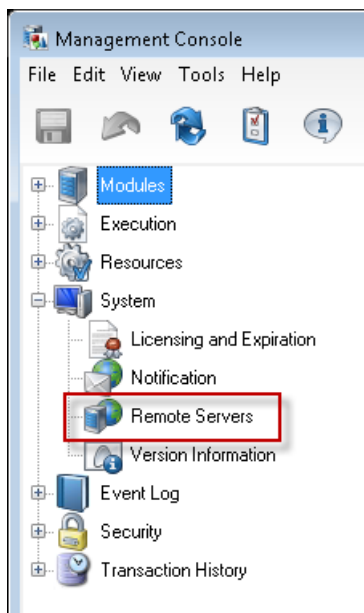
Data | Modules

Filter

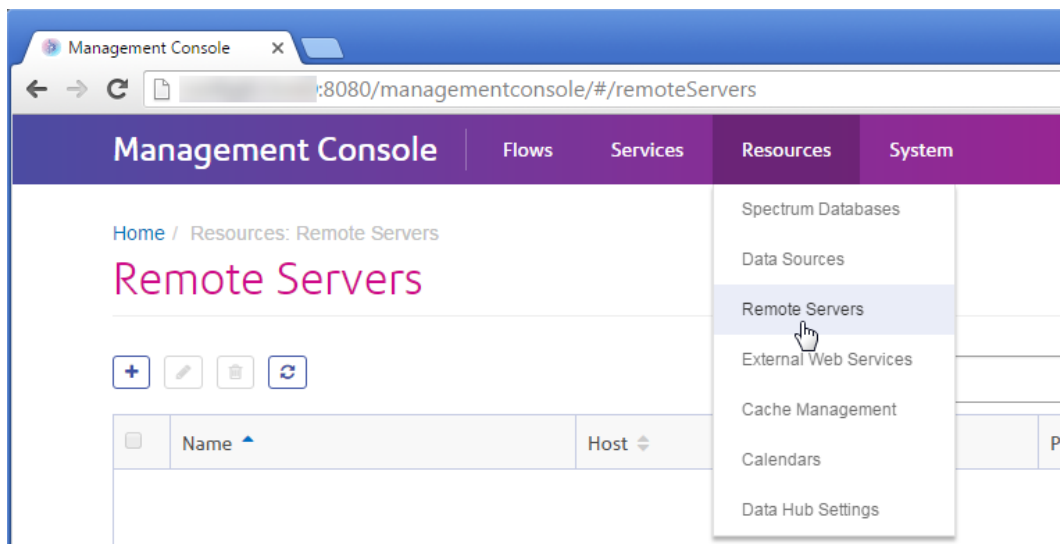
| Status | Module | Feature | Mode  | Expires On | Days Remaining | Transaction Remaining |
|--------|--------|---------|-------|------------|----------------|-----------------------|
|        | 70S    | 70S     | Batch |            | 0              | 0                     |

### System - Remote Servers

In the Windows version of Management Console, the settings for remote servers are under the System section:

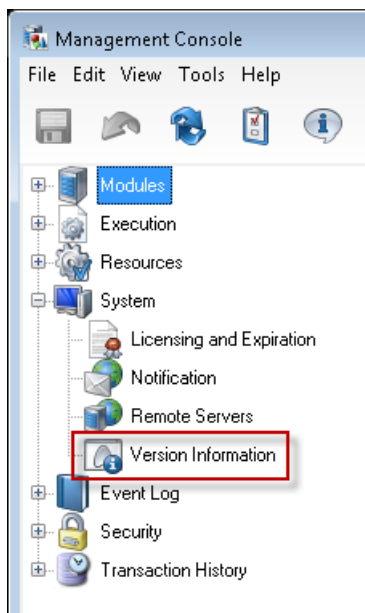


In the browser version of Management Console, these settings can be found under **Resources > Remote Servers**:



### *System - Version Information*

In the Windows version of Management Console, version information is located under the System node:



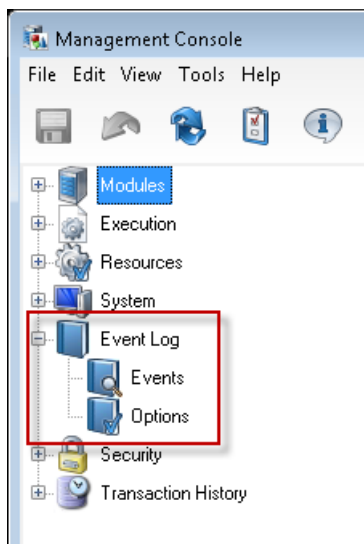
In the browser version of Management Console, version information is under **System > Version**.

 A screenshot of the browser version of Management Console. The top navigation bar is purple and contains 'Management Console', 'Flows', 'Services', 'Resources', and 'System'. The 'System' tab is active. Below the navigation bar, the breadcrumb 'Home / System: Version' is shown, followed by the title 'Version'. A dropdown menu is open under the 'System' tab, listing 'Licensing and Expiration', 'Version' (with a mouse cursor), 'Security', 'Logs', and 'Mail Server'. Below the menu, there are three tabs: 'System', 'Services', and 'Components'. The 'System' tab is selected, displaying a table of system information:
 

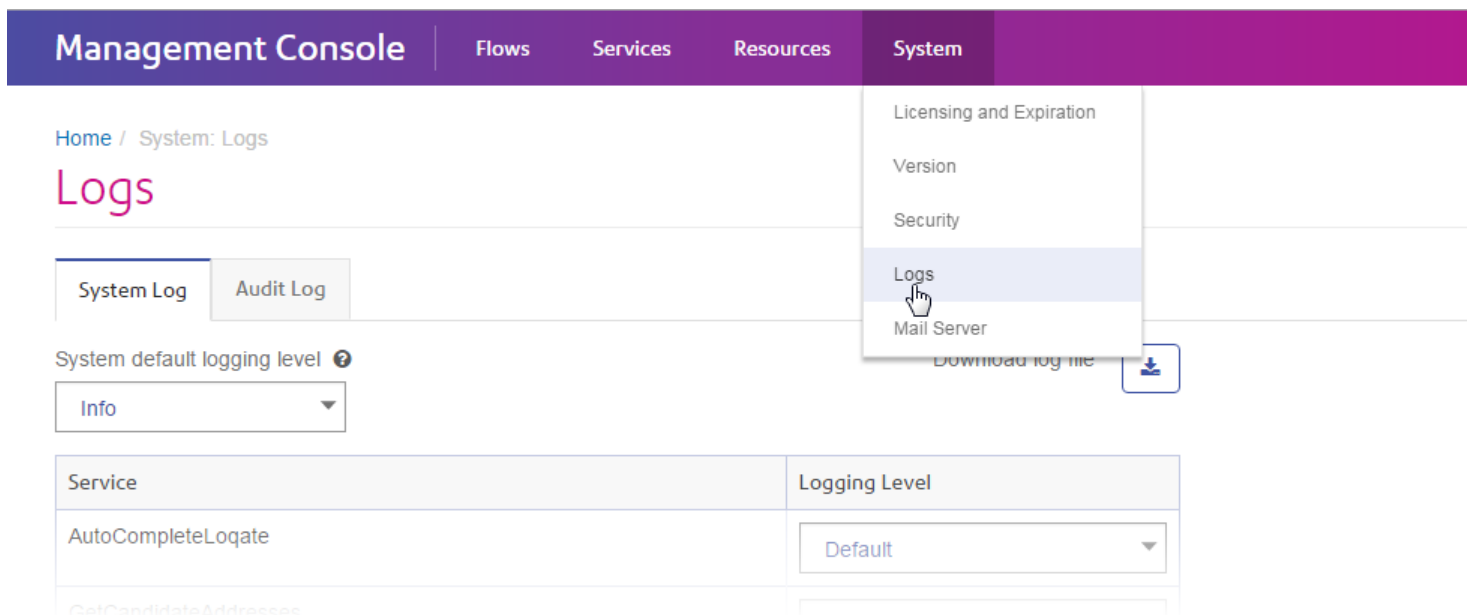
|              |                     |
|--------------|---------------------|
| Server:      | 0-SNAPSHOT          |
| Java:        | 1.8.0_45            |
| Java Vendor: | Oracle Corporation  |
| OS:          | 6.2                 |
| OS Name:     | Windows Server 2012 |

### Event Log

In the Windows version of Management Console you can configure the logging level for each service and view the event log under the Event Log node:

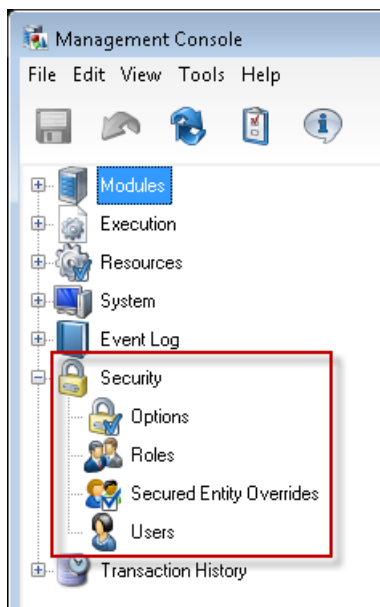


In the browser version of Management Console, you can configure the logging level of services under **System > Logs**. You can also download the system log (also known as the wrapper log) and view the audit log.



### Security

In the Windows version of Management Console, the settings for users, roles, and secured entity overrides are in the Security node:



In the browser version of Management Console, the Options, Roles, and Users settings are located under **System > Security**.

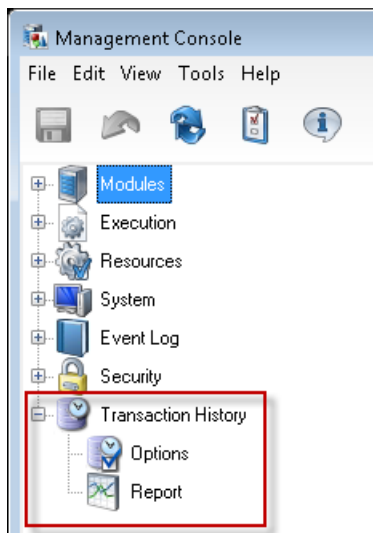
The screenshot shows the browser version of the Management Console. The top navigation bar includes 'Management Console', 'Flows', 'Services', 'Resources', and 'System'. The 'System' menu is open, showing options: Licensing and Expiration, Version, Security (highlighted with a mouse cursor), Logs, and Mail Server. Below the navigation bar, the breadcrumb is 'Home / System: Security'. The main heading is 'Security'. There are four tabs: 'Users' (selected), 'Roles', 'Access Control', and 'Directory Access'. Below the tabs are three icons: a plus sign, a pencil, and a trash can. A 'Filter' input field is on the right. A table lists users and their roles:

| <input type="checkbox"/> | Name ▲ | Roles ⇅  |
|--------------------------|--------|--|
| <input type="checkbox"/> | admin  | admin, designer, integrator, spatial-admin, spatial-user, user |
| <input type="checkbox"/> | amy    | designer   |
| <input type="checkbox"/> | jack   | designer   |
| <input type="checkbox"/> | john   | integrator   |

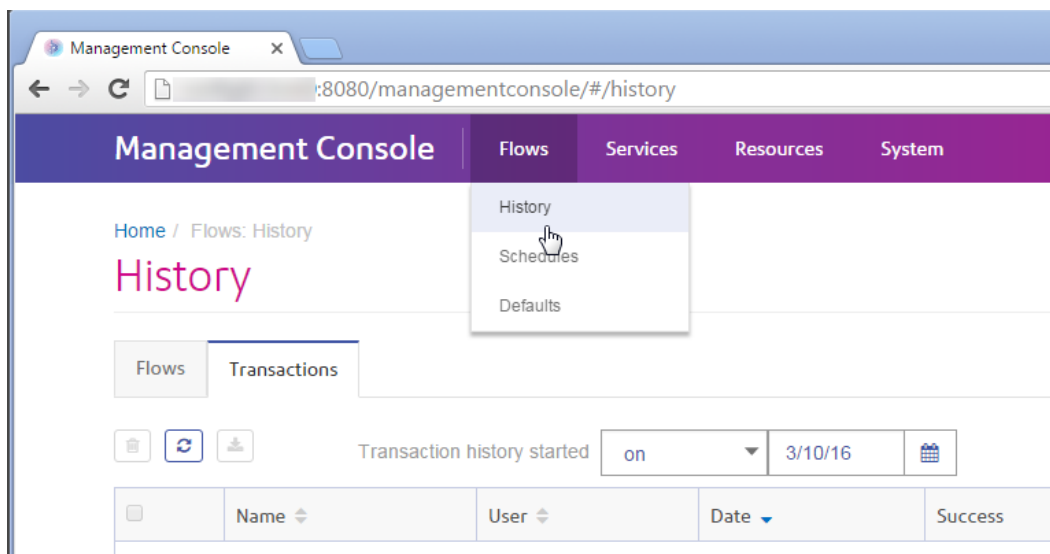
**Note:** In the Windows client the permission that grants users the ability to change an entity is called "Modify" but in the web browser version it's called "Edit".

## Transaction History

In the Windows version of Management Console, the transaction report and its configuration settings are available under Transaction History:



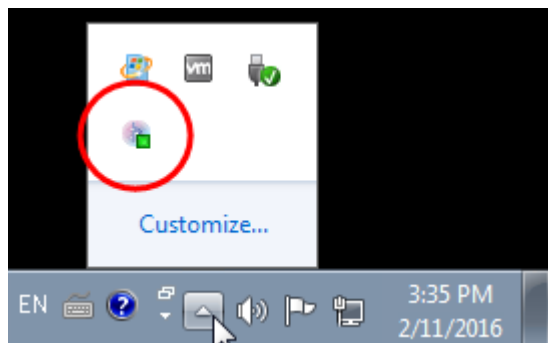
In the browser version of Management Console, transaction history is under **Flows** > **History**.



## Starting and Stopping the Server

You may need to stop and start the Spectrum™ Technology Platform server to perform maintenance such as installing a product update.

- On Windows, Spectrum™ Technology Platform is set to start automatically when Windows starts up. To ensure that Spectrum™ Technology Platform has started, right click on the Spectrum™ Technology Platform icon in the notification area. If the icon is green then the server has started.



To stop Spectrum™ Technology Platform, right-click the icon and select **Stop Spectrum™**.

- To start the server on Unix or Linux:
  - a) Change the working directory to the `bin` directory where Spectrum™ Technology Platform is installed.

For example:

```
cd /usr/g1/tst/server/bin
```

- b) Source the setup file.

For example:

```
./setup
```

- c) Start Spectrum™ Technology Platform.

- To start Spectrum™ Technology Platform in the background, type the following command:

```
./server.start
```

- To start Spectrum™ Technology Platform in the foreground, type the following command:

```
./server.start console
```

- To stop Spectrum™ Technology Platform on Unix or Linux, type the following command:

```
./server.stop
```

**Note:** Java uses `/var/tmp` as its temporary directory by default. If there is not enough space in this directory, the Spectrum™ Technology Platform server may not start.

## Installing the Client Tools

The Spectrum™ Technology Platform client tools are applications that you use to administer your server and design and run flows. The client tools are:

- Enterprise Designer** Enterprise Designer is a Windows application that you use to create, modify, and run flows.
- Job Executor** Job Executor is a command line tool that you can use to run a job from a command line or script. The job must have been previously created and saved on Spectrum™ Technology Platform using Enterprise Designer.
- Process Flow Executor** Process Flow Executor is a command line tool that you can use to run a process flow from a command line or script. The process flow must have been previously created and saved on Spectrum™ Technology Platform using Enterprise Designer.
- Administration Utility** The Administration Utility provides command line access to administrative functions. You can use it in a script, allowing you to automate certain administrative tasks. You can also use it interactively.

**Note:** Before installing, be sure to read the release notes. The release notes contains important compatibility information as well as release-specific installation notes.

To install the client tools:

1. Open a web browser and go to the Spectrum™ Technology Platform Welcome Page at:

`http://<servername>:<port>`

For example, if you installed Spectrum™ Technology Platform on a computer named "myspectrumplatform" and it is using the default HTTP port 8080, you would go to:

`http://myspectrumplatform:8080`

2. Click **Platform Client Tools**.
3. Download the client tool you want to install.

## Network Ports

The Spectrum™ Technology Platform server uses several network ports for communication. Network port conflicts can result in module components failing to start. One indication that a component has failed to start is if it does not appear in Management Console. To troubleshoot the problem, look in the Spectrum™ Technology Platform wrapper log. This log shows which port is causing the problem. You can find the Spectrum™ Technology Platform wrapper log in:



```
server\app\repository\logs\wrapper.log
```

You can change the ports by modifying the properties in this file and restarting the server:

```
server\app\conf\spectrum-container.properties
```

**Note:** In a clustered environment you must modify the `spectrum-container.properties` file on *each node* in the cluster.

| Port | Description   |
|------|---|
| 5001 | <p>This port is used by the Spectrum™ Technology Platform configuration database.</p> <p>To use a different port in a non-clustered environment, modify this property:</p> <pre>spectrum.repository.server.coordinator.port</pre> <p>To use a different port in a clustered environment:</p> <ul style="list-style-type: none"> <li>Specify the port you want instead of 5001 in <code>spectrum.repository.server.coordinator.port</code>.</li> <li>Specify the seed nodes for the configuration database in <code>spectrum.repository.server.cluster.seeds</code>.</li> </ul>  |
| 5701 | <p>This port is used by Hazelcast for managing distributed processing between Spectrum™ Technology Platform servers in a cluster.</p> <p>To use a different port in a non-clustered environment, modify this property:</p> <pre>spectrum.hazelcast.port</pre> <p>To use a different port in a clustered environment:</p> <ul style="list-style-type: none"> <li>Specify the port you want to use instead of 5701 in <code>spectrum.hazelcast.port</code>.</li> <li>Include the Hazelcast port number after each IP address specified in <code>spectrum.cluster.seeds</code>. For example, if <code>spectrum.hazelcast.port</code> is set to 5702 and the IP address of a seed node is 1.2.3.4.5, you would specify <code>1.2.3.4.5:5702</code> in <code>spectrum.cluster.seeds</code>.</li> </ul> |
| 6362 | <p>This port is used if you enable backups of the Spectrum™ Technology Platform configuration database. To use a different port, modify this property:</p> <pre>spectrum.backup.http.port</pre>   |
| 7474 | <p>This port is used by the Spectrum™ Technology Platform configuration database. To use a different port, modify this property:</p> <pre>spectrum.repository.server.connector.http.port</pre>  |

| Port        | Description   |
|-------------|---|
| 7687        | <p>This port is used by the Spectrum™ Technology Platform configuration database. To use a different port, modify this property:</p> <pre>spectrum.repository.server.connector.bolt.port</pre>  |
| 8080        | <p>The port used for communication between the server and Enterprise Designer and Management Console. This port is also used by web services. To use a different port, modify this property:</p> <pre>spectrum.http.port</pre>  |
| 9200        | <p>This port is used by the search index engine used by the Advanced Matching Module. To use a different port, modify this property:</p> <pre>spectrum.index.http.port</pre>  |
| 9300        | <p>This port is used by the search index engine used by the Advanced Matching Module. To use a different port in a non-clustered environment, modify this property:</p> <pre>spectrum.index.tcp.port</pre> <p>To use a different port in a clustered environment:</p> <ul style="list-style-type: none"> <li>• Specify the port you want to use instead of 9300 in <code>spectrum.index.tcp.port</code> and after the colon in <code>spectrum.index.client.addresses</code>.</li> <li>• Specify the seed nodes for the search index in <code>spectrum.index.server.cluster.seeds</code>.</li> </ul> |
| 10119       | <p>This port is used for API calls made to services. To use a different port, modify this property:</p> <pre>spectrum.socketgateway.port</pre>  |
| 15431-15432 | <p>These ports are used by the Machine Learning Module.</p>   |
| 32750       | <p>This port is used for JDBC connections model stores which are created in Metadata Insights. To use a different port, modify this property:</p> <pre>spectrum.metadata.jdbc.port</pre>  |
| 32751       | <p>This port is used for ODBC connections model stores which are created in Metadata Insights. To use a different port, modify this property:</p> <pre>spectrum.metadata.odbc.port</pre>  |

# 2 - Security

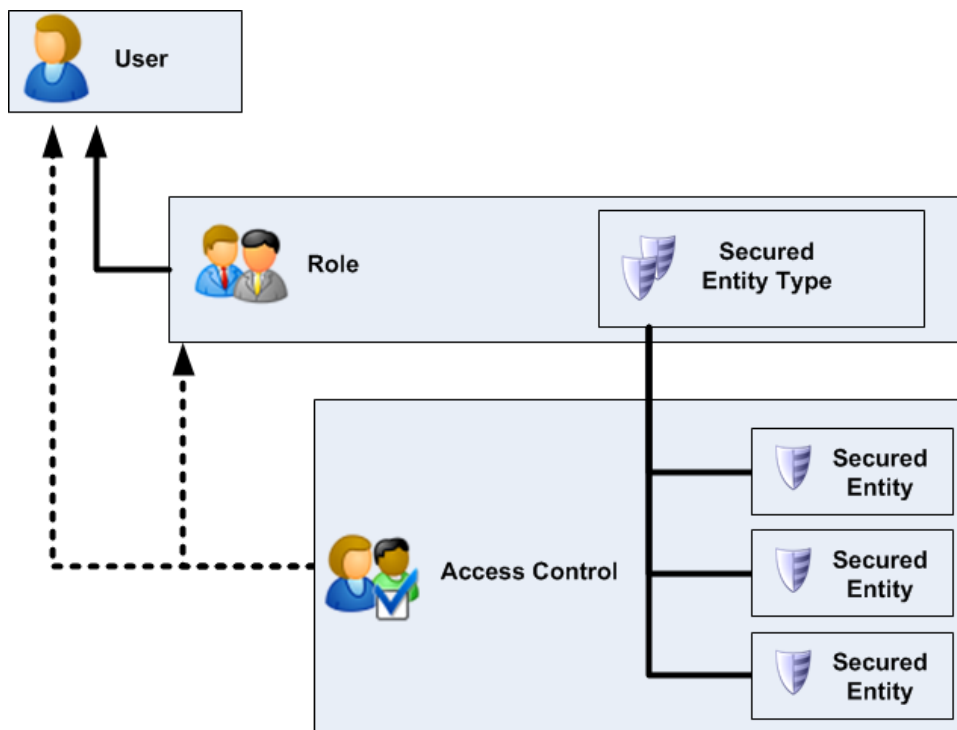
## In this section

---

|   |    |
|---|----|
| Security Model                                    | 28 |
| Users   | 29 |
| Roles   | 34 |
| Access Control                                    | 45 |
| Security for the Location Intelligence Module     | 48 |
| Limiting Server Directory Access                  | 53 |
| Configuring HTTPS Communication                   | 55 |
| Web Service Authentication                        | 58 |
| Using LDAP or Active Directory for Authentication | 63 |

## Security Model

Spectrum™ Technology Platform uses a role-based security model to control access to the system. The following diagram illustrates the key concepts in the Spectrum™ Technology Platform security model:



A *user* is an account assigned to an individual person which the person uses to authenticate to Spectrum™ Technology Platform, either to one of the client tools such as Enterprise Designer or Management Console, or when calling a service through web services or the API.

A user has one or more roles assigned to it. A *role* is a collection of permissions that grant or deny access to different parts of the system. Roles typically reflect the kinds of interactions that a particular type of user has with the system. For example, you may have one role for dataflow designers which grants access to create and modify dataflows, and another role for people who only need to process data through existing dataflows.

A role grants permissions to secured entity types. A *secured entity type* is a category of items to which you want to grant or deny access. For example, there is a secured entity type called "Dataflows" which controls the default permissions for all dataflows on the system.

If you need to fine-tune access you can optionally override the settings in the role or user by configuring access control. Access control settings work in conjunction with roles to define the permissions for a user. Roles define the permissions for categories of entities, such as all dataflows or all database resources, and access control settings define the permissions for specific entities,

called *secured entities*. Examples of secured entities include specific jobs or specific database connections. Defining access control settings is optional. If you do not define access control settings, the permissions defined in the role will control the user's permissions.

Access control settings work in conjunction with roles to define the permissions for a user. Roles define the permissions for categories of entities, such as all dataflows or all database resources, and access control settings define the permissions for specific entities, called *secured entities*. Examples of secured entities include specific jobs or specific database connections. For example, you may have a role that has granted the Modify permission to the secured entity type "Dataflows", but you may want to prevent users from modifying one specific dataflow. You could accomplish this by using access control to remove the Modify permission for the specific dataflow you do not want modified. You can specify access control settings for users and roles. Access control settings for a user override that specific user's permissions as granted by the user's roles. Access control settings for roles apply to all users who have that role.

### Related concepts

[Roles](#) on page 34

[Users](#) on page 29

### Related tasks

[Disabling Role-Based Security](#) on page 37

[Adding a User](#) on page 30

[Creating a Role](#) on page 35

[Configuring Access Control](#) on page 45

## Users

Spectrum™ Technology Platform user accounts control the types of actions users can perform on the system. User accounts are required to:

- Use tools like Management Console, Enterprise Designer, Metadata Insights, and command-line tools
- Run jobs on a schedule
- Run jobs from the command line
- Access services through web services or the API

There is an administrative account called **admin** that comes with the system. This account has full access. The initial password is "admin".


**Important:** You should change the admin password immediately after installing Spectrum™ Technology Platform to prevent unauthorized administrative access to your system.

You can create as many user accounts as you need.

**Related concepts**[Security Model](#) on page 28[Roles](#) on page 34[Security for the Location Intelligence Module](#) on page 48**Related tasks**[Configuring Access Control](#) on page 45

## Adding a User

This procedure describes how to create a Spectrum™ Technology Platform user account and assign a role to the account.

1. Open Management Console.
2. Go to **System > Security**.
3. Click the Add button .
4. Leave the **Enabled** switch set to **On** if you want this user account to be available for use.
5. Enter the user name in the **User name** field.

**Note:** User names can only contain ASCII characters. User names are case sensitive.

6. Enter the user's email address in the **Email address** field. The email address is used by some modules to send notifications to users.
7. Enter a description of the user in the **Description** field.
8. Enter and confirm the user's password.
9. Select the roles you want to give to this user.

You may create your own roles or use one of the default roles. The default roles are:

|                   |   |
|-------------------|---|
| <b>admin</b>      | This role has full access to all parts of the system.   |
| <b>designer</b>   | This role is for users that create dataflows and process flows in Enterprise Designer. It provides the ability to design and run dataflows.   |
| <b>integrator</b> | This role is for users who need to process data through Spectrum™ Technology Platform but do not need to create or modify dataflows. It allows the user to access services through web services and the API, and to run jobs. |
| <b>user</b>       | This is the default role. It provides no access to the system. Users who have this role will only gain access to the system if you grant permission through secured entity overrides.   |


For information about creating roles, see [Creating a Role](#) on page 35.

10. Click **Save**.

**Related concepts**[Security Model](#) on page 28**Related tasks**[Disabling Role-Based Security](#) on page 37[Creating a Role](#) on page 35[Configuring Access Control](#) on page 45

## Changing a Password

This procedure describes how to change a user's password.

1. Open the Management Console.
2. Go to **System > Security**.
3. Select a user then click the Edit button .
4. Click **Change password**.
5. Enter the new password and enter it a second time to confirm it.
6. Click **Save**.

## Setting a Minimum Password Length

The minimum password length is enforced when creating or changing a password. Existing passwords that are shorter than the minimum length will continue to be valid.

1. Open a web browser and go to `http://server:port/jmx-console`  
Where:  
*server* is the IP address or hostname of your Spectrum™ Technology Platform server.  
*port* is the HTTP port used by Spectrum™ Technology Platform. The default is 8080.
2. Log in using the admin account.
3. Under "Domain: com.pb.spectrum.platform.config", click **com.pb.spectrum.platform.config:manager=AccountConfigurationManager**.
4. In the **updatePasswordPolicy** operation, set the **enableAdvanceControl** option to **True**.
5. In the **minLength** field, enter the minimum password length.
6. Click **Invoke**.
7. Click **Return to MBean View** to go back to the Account Configuration Manager screen.

## Changing Your Email Address


The email address associated with your user account is used by some modules to send you notifications. To change your email address, follow these steps.

1. Log in to Management Console.
2. Click the user menu in the top right corner.
3. Select **Profile**.
4. In the **Email** field, enter your new email address.
5. Click **Save**.

## Disabling a User Account

You can disable a user account so that it cannot be used to gain access to Spectrum™ Technology Platform. Any jobs that run on a schedule using a disabled user account will not run.

**Note:** The user account "admin" cannot be disabled.

1. Open the Management Console.
2. Go to **System > Security**.
3. Check the box next to the user you want to modify then click the Edit button .
4. Switch the **Enabled** switch to **Off**.
5. Click **Save**.

The user account is now disabled and cannot be used to gain access to Spectrum™ Technology Platform.

### Related tasks

[Deleting a User](#) on page 32

### Related reference


[token userrevoke](#) on page 374

## Deleting a User

This procedure describes how to permanently delete a Spectrum™ Technology Platform user account.

**Tip:** User accounts can also be disabled, which prevents the account from being used to access the system without deleting the account.



1. Open the Management Console.
2. Go to **System > Security**.
3. Check the box next to the user you want to delete then click the Delete button .

**Note:** The user account "admin" cannot be deleted.

#### Related tasks

[Disabling a User Account](#) on page 32

#### Related reference

[token userrevoke](#) on page 374

## User Account Locking

As a security precaution, user accounts are disabled after five unsuccessful authentication attempts in row. This includes unsuccessful authentication attempts to Enterprise Designer, Management Console, web services, and the Client API.

As an administrator, you can re-enable a user account by logging into Management Console, editing the user, and switching the **Enabled** switch to **On**. User accounts can also be re-enabled using the Administration Utility. Users do not have the ability to unlock their own accounts.

**Note:** If you are using LDAP or Active Directory for authentication, the account locking rules of these services will apply. Your LDAP or Active Directory rules may allow more or fewer unsuccessful login attempts than Spectrum™ Technology Platform.

### Unlocking the admin Account

User accounts are locked after several unsuccessful login attempts. Most user accounts can be unlocked through Management Console, but the admin account cannot. Instead, you must run a script on the server to unlock the admin account.

1. Log in to the server running Spectrum™ Technology Platform.

If you are running Spectrum™ Technology Platform in a cluster, log in to any of the nodes. You only need to run the unlock script on one of the nodes.

2. Open a command prompt and go to the *Spectrum Folder\server\bin* folder.
3. (Unix and Linux only) Run the following command:

```
. ./setup
```

4. Run the enableadmin script by typing the following command:

On Windows:

```
enableadmin.bat -h HostAndPort -p AdminPassword [-s]
```

On Unix/Linux:

```
./enableadmin.sh -h HostAndPort -p AdminPassword [-s]
```

Where:

|                             |  |
|-----------------------------|--|
| <b><i>HostAndPort</i></b>   | The hostname and HTTP port used by Spectrum™ Technology Platform. For example, <code>spectrumserver:8080</code> .  |
| <b><i>AdminPassword</i></b> | The password for the admin account. If you do not know the admin account password and the admin account is locked, contact Pitney Bowes Technical Support. |
| <b>-s</b>                   | Specify <code>-s</code> if Spectrum™ Technology Platform is configured to use HTTPS.   |

## Automatic Logout Due to Inactivity

Users of Enterprise Designer and web clients such as Management Console, the Relationship Analysis Client, Business Steward Portal, and others are automatically logged out after 30 minutes of inactivity.

## Roles

A *role* is a collection of permissions that grant or deny access to different parts of the system. Roles typically reflect the kinds of interactions that a particular type of user has with the system. For example, you may have one role for dataflow designers which grants access to create and modify dataflows, and another role for people who only need to process data through existing dataflows.

Spectrum™ Technology Platform comes with these roles already defined:

|                   |   |
|-------------------|---|
| <b>admin</b>      | This role has full access to all parts of the system.   |
| <b>designer</b>   | This role is for users that create dataflows and process flows in Enterprise Designer. It provides the ability to design and run dataflows.   |
| <b>integrator</b> | This role is for users who need to process data through Spectrum™ Technology Platform but do not need to create or modify dataflows. It allows the user to access services through web services and the API, and to run jobs. |
| <b>user</b>       | This is the default role. It provides no access to the system. Users who have this role will only gain access to the system if you grant permission through secured entity overrides.   |

**Note:** See [Security for the Location Intelligence Module](#) on page 48 for information about the predefined roles for the Location Intelligence Module.

To view the permissions granted to each of these roles, open Management Console, go to **Security** and click **Roles**. Then select the role you want to view and click **View**.

**Tip:** You cannot modify the predefined roles. However, you can create new roles using the predefined roles as a starting point.

#### Related concepts

[Security Model](#) on page 28

[Users](#) on page 29


[Security for the Location Intelligence Module](#) on page 48


#### Related tasks

[Configuring Access Control](#) on page 45

## Creating a Role

A role is a collection of permissions that you assign to a user. If the predefined roles that come with Spectrum™ Technology Platform do not fit your organization's needs, you can create your own roles.

1. Open Management Console.
2. Go to **System > Security**.
3. Click **Roles**.
4. Click the Add button .

**Tip:** If you want to create a role that's similar to an existing role, you can make a copy of the existing role by checking the box next to the role you want to copy then clicking the Copy button . Then, edit the new role and continue with the following steps.

5. In the **Role name** field, enter the name you want to give to this role. The name can be anything you choose.
6. Optional: Since the list of secured entity types can be long, you may want to display only a certain group of secured entity types. This can be useful if you want to apply the same permissions to all entities in a group. For example, if you want to remove the Modify permission from all database resources, you could filter to show just the Database Resources group. To display and modify only one group:
  - a) Check the **Enable group filtering** box.
  - b) Click the funnel icon in the header of the **Group** column and select the group you want to display.
  - c) Check or clear the box in the column header of the permission you want to apply.

- d) To return to the full list of secured entity types, click the filter icon and select **(All)** then clear the **Enable group filtering** box.
7. Select the permissions you want to grant for each entity type. The permissions are:
- View** Allows the user to view entities contained by the entity type. For example, if you allow the View permission for the JDBC Connection entity type, users with this role would be able to view database connections in Management Console.
  - Modify** Allows the user to modify entities contained by the entity type. For example, if you allow the Modify permission for the JDBC Connection entity type, users with this role would be able to modify database connections in Management Console.
  - Create** Allows the user to create entities that fall into this entity type's category. For example, if you allow the Create permission for the JDBC Connection entity type, users with this role would be able to create new database connections in Management Console.
  - Delete** Allows the user to delete entities contained by the entity type. For example, if you allow the Delete permission for the JDBC Connection entity type, users with this role would be able to delete database connections in Management Console.
  - Execute** Allows the user to initiate processing of jobs, services, and process flows. For example, if you allow the Execute permission for the Job entity type, users with this role would be able to run batch jobs. If you allow the Execute permission for the Service entity type, users with this role would be able to access services running on Spectrum™ Technology Platform through the API or web services.
8. Click **Save**.

The role is now available to be assigned to a user.

#### Related concepts

[Security Model](#) on page 28

#### Related tasks

[Disabling Role-Based Security](#) on page 37

[Adding a User](#) on page 30


[Configuring Access Control](#) on page 45

## Deleting a Role

A role may be deleted if it is no longer assigned to any users.

**Note:** The following roles cannot be deleted: admin, user, designer, and integrator.

1. Open Management Console.
2. Go to **System > Security**.

3. On the **Users** tab, make sure the role you want to delete is not assigned to any users. You cannot delete a role if it assigned to a user.
4. Click **Roles**.
5. Check the box next to the role you want to delete then click the Delete button .

## Disabling Role-Based Security

Role-based security is enabled by default. This means that the security restrictions assigned to users through roles are enforced. If you want to disable role-based security, the security restrictions assigned to users will not be enforced and all users will be able to access all parts of the system. Note that a valid user account is always required to access services even if you disable role-based security.

This procedure describes how to disable role-based security.

**Warning:** If you follow this procedure all users will have full access to your Spectrum™ Technology Platform system.

1. Open the Management Console.
2. Go to **System > Security**.
3. Switch the **Limit access by role** switch to **Off**.

### Related concepts

[Security Model](#) on page 28

### Related tasks

[Adding a User](#) on page 30

[Creating a Role](#) on page 35

[Configuring Access Control](#) on page 45

[Disabling Authentication for Web Services](#) on page 60

## Secured Entity Types - Advanced Matching Module

An entity type is a category of items to which you want to grant or deny access. The following entity types control access to parts of the Advanced Matching Module.

**Match Rules Management** Controls access to the match rules in the Interflow Match stage, Intraflow Match stage, Transactional Match stage, and Match Rules Management Tool in Enterprise Designer.

**Search Index Management** Controls access to search indexes in Write to Search Index, Candidate Finder, and the Search Index Management Tool in Enterprise Designer.

## Secured Entity Types - Business Steward Module

An entity type is a category of items to which you want to grant or deny access. The following entity type controls access to parts of the Business Steward Module.

**Exceptions** Controls the ability to modify or delete exception records assigned to other users in the Read Exceptions stage and the Exception Editor. Also controls the ability to review and manage exception record activity in the Exception Manager and enables access to Performance data in the Business Steward Portal.

Users with Edit permissions for the Business Steward Portal group must also have Security-Users View permissions for the Platform group in order to assign or reassign exception records.

In addition to creating secured entity types, you can also create security entity overrides that specify restrictions for specific dataflows or even specific stages. These overrides supersede user- and role-based secured entity type settings in that they make the permissions more restrictive.

In the Business Steward Module, you can create overrides for dataflows and stages as follows:

- To view information on the Business Steward Portal Performance page
- To modify data on the Business Steward Portal Dashboard and Exception Editor
- To reassign exceptions and modify status on the Business Steward Portal Exceptions Manager
- To delete dataflows on the Business Steward Portal Exceptions Manager

For example, if user JohnSmith has Edit permissions based on the secured entity type, but there is a specific dataflow that his manager doesn't want anyone to alter, his manager can define an access control setting that restricts John from modifying that particular dataflow. He is still able to edit other dataflows, but not the one for which there is an access control setting that restricts him. Also, he will still be able to view and update exception records that are assigned to him from that dataflow in the Exception Editor.

## Secured Entity Types - Data Federation Module

An entity type is a category of items to which you want to grant or deny access. For example, the entity type called "Baseview" controls permissions for all Baseviews in the Data Federation Module.

**Note:** The Spectrum Data Federation entity types apply to only the Data Federation Module.

The Data Federation entity types are:

|                                     |   |
|-------------------------------------|---|
| <b>Baseview</b>                     | Controls access to all Baseviews in the Data Federation Module.                           |
| <b>Datasource - Amazon DynamoDB</b> | Controls access to all Amazon DynamoDB connections created in the Data Federation Module. |

|  |  |
|--|--|
| <b>Datasource - Amazon SimpleDB</b>    | Controls access to all Amazon SimpleDB connections created in the Data Federation Module.                              |
| <b>Datasource - Apache Cassandra</b>   | Controls access to all Apache Cassandra connections created in the Data Federation Module.                             |
| <b>Datasource - Flat File</b>          | Controls access to all Flat File connections, whether Fixed Width or Delimited, created in the Data Federation Module. |
| <b>Datasource - Marketo</b>            | Controls access to all Marketo connections created in the Data Federation Module.                                      |
| <b>Datasource - MS Dynamics CRM</b>    | Controls access to all MS Dynamics CRM Online connections created in the Data Federation Module.                       |
| <b>Datasource - NetSuite</b>           | Controls access to all NetSuite connections created in the Data Federation Module.                                     |
| <b>Datasource - Salesforce</b>         | Controls access to all Salesforce connections created in the Data Federation Module.                                   |
| <b>Datasource - SAP</b>                | Controls access to all SAP NetWeaver connections created in the Data Federation Module.                                |
| <b>Datasource - Siebel</b>             | Controls access to all Siebel connections created in the Data Federation Module.                                       |
| <b>Datasource - Splunk</b>             | Controls access to all Splunk connections created in the Data Federation Module.                                       |
| <b>Datasource - SuccessFactors</b>     | Controls access to all SuccessFactors connections created in the Data Federation Module.                               |
| <b>Datasource - SugarCRM On-Demand</b> | Controls access to all SugarCRM On-Demand connections created in the Data Federation Module.                           |
| <b>Metaview</b>                        | Controls access to all Metaviews created in the Data Federation Module.  |
| <b>Virtual Data Source</b>             | Controls access to all the Virtual Data Sources created in the Data Federation Module.                                 |

**Note:** In the Data Federation Module, for any role, the 'Execute' right can be provided to access the Virtual Data Source entity type only.

## Secured Entity Types - Data Hub Module

An entity type is a category of items to which you want to grant or deny access. The following entity types control access to parts of the Data Hub Module.

- Algorithms** Controls the ability to execute algorithms in the Relationship Analysis Client.
- Model Admin** Controls the ability for users to perform the following actions using Data Hub stages and the Relationship Analysis Client:
  - Read model data: entities, relationships, and their properties

- Create models and model data: entities, relationships, and their properties
- Modify model data: entities, relationships, and their properties
- Delete models and model data: entities, relationships, and their properties

For example, if you had a model comprised of insurance data, you might have doctors and patients as entities, with relationships such as "visited" or "filed a claim for," and properties that depict the dates for those visits or claims. Using this example, the Model Admin entity type would allow you to grant permissions for the following actions:

- Read data for doctors and patients and their visit or claim properties
- Create doctor and patient entities, link them together with visit or claim relationships, and include any properties such as addresses, claim IDs, dates, and so on.
- Modify doctor and patient entities, their relationships, and their properties such as addresses, claim IDs, dates, and so on.
- Delete doctors and patients from the model

**Model Metadata** Controls the ability for users to perform the following actions using Data Hub stages and the Relationship Analysis Client:

- Read entity types and relationship types
- Create entity types, relationship types, and their properties
- Modify entity type properties and relationship type properties
- Delete entities and relationships

**Note:** This permission includes clearing a model in the Write to Hub stage.

**Monitor Admin** Controls the ability to create monitors in the Relationship Analysis Client that detect changes to a model's entities or relationships.

**Theme Admin** Controls the ability to define themes for models in the Relationship Analysis Client.

## Secured Entity Types - Data Normalization Module

An entity type is a category of items to which you want to grant or deny access. The following entity types control access to parts of the Data Normalization Module.

**Advanced Transformer Tables** Controls access to the Advanced Transformer tables in the Table Management tool in Enterprise Designer.

**Open Parser Cultures** Controls access to cultures in the Open Parser Domain Editor tool in Enterprise Designer.

**Open Parser Domains** Controls access to domains in the Open Parser Domain Editor tool in Enterprise Designer.



|                               |  |
|-------------------------------|--|
| <b>Open Parser Tables</b>     | Controls access to the Open Parser tables in the Table Management tool in Enterprise Designer.     |
| <b>Standardization Tables</b> | Controls access to the Standardization tables in the Table Management tool in Enterprise Designer. |

## Secured Entity Types - Database Resources

An entity type is a category of items to which you want to grant or deny access. Database resources are available depending on which modules you have installed, for example:

- **Centrus Database Resources**
- **Enterprise Routing**
- **Spatial Database Resources**

## Secured Entity Types - Enterprise Data Integration Module

An entity type is a category of items to which you want to grant or deny access. The following entity types control access to parts of the Enterprise Data Integration Module.

|              |   |
|--------------|---|
| <b>Cache</b> | Controls access to caches used by the Write to Cache and Query Cache stages, and the Cache Management tool in Management Console. |
|--------------|---|

## Secured Entity Types - External Web Services

An entity type is a category of items to which you want to grant or deny access.

There is only one secured entity type in the External Web Services category: **Connection**. This secured entity type controls access to external web services in Management Console and Enterprise Designer.

## Secured Entity Types - Location Intelligence Module

An entity type is a category of items to which you want to grant or deny access. The Location Intelligence Module has the following module-specific entity types:

|                        |  |
|------------------------|--|
| <b>Named Resources</b> | Controls permissions to all named resources in the Location Intelligence Module. Users of Location Intelligence Module services must have at least read permissions for the resources they use as well as for any dependent resources. When a named resource is created (using any tool, including Spatial Manager, the Administration Utility, the Named Resource Service, and WebDAV), a new |
|------------------------|--|

LocationIntelligence.Named Resource secured entity is automatically created for the named resource.

**Dataset.DML** Controls permissions to datasets used in the Location Intelligence Module that are associated with named tables. When a named table is created or uploaded (using any tool, including Spatial Manager, the Administration Utility, the Named Resource Service, and WebDAV), a new LocationIntelligence.Dataset secured entity is automatically created for the associated dataset of that named table. A user must have View permissions on a named table *and* Create/Modify/Delete permissions on the dataset in order to perform DML operations on writable (JDBC-based) tables. DML operations include insert, update, and delete operations performed using the Write Spatial Data stage or the Feature Service.

## Secured Entity Types - Metadata Insights

An entity type is a category of items to which you want to grant or deny access. The secured entity types for Metadata Insights control access to the Metadata Insights web application's modeling, profiling, lineage, and impact analysis features.

**Lineage & Impact Analysis** Controls access to the **Lineage & Impact Analysis** view in Metadata Insights. Since Lineage & Impact Analysis only provides the ability to view information, the only permission available is **View**.

**Note:** Giving users the **View** permission for **Lineage & Impact Analysis** will allow them to view logical models, physical models, and data stores even if they do not have the **View** permission for these secured entity types.

**Logical Model** Controls access to logical models in the **Modeling** section of Metadata Insights.

**Model Store** Controls access to model stores in the **Modeling** section of Metadata Insights.

**Physical Model** Controls access to physical models in the **Modeling** section of Metadata Insights.

## Secured Entity Types - Platform

An entity type is a category of items to which you want to grant or deny access. For example, there is an entity type called "Dataflows" which controls permissions for all dataflows on the system. Platform entity types apply to all Spectrum™ Technology Platform installations, as compared to module-specific entity types that apply only if you have installed particular modules. The platform-level entity types are:

**Audit Log** Controls access to the **System > Logs > Audit Log** area in Management Console.

|   |  |
|---|--|
| <b>Dataflows</b>                            | Controls access to all dataflow types (jobs, services, and subflows) in Enterprise Designer.<br><br><b>Note:</b> If a user does not have the Edit permission, the user will only see the exposed version and the last saved version in the <b>Versions</b> pane in Enterprise Designer.  |
| <b>Dataflows - Expose</b>                   | Controls the ability to make dataflows available for execution.<br><br><b>Note:</b> In order to expose the latest saved version of the dataflow (the version that's always at the top of the <b>Versions</b> pane in Enterprise Designer) the user must have the Edit permission for the <b>Dataflows</b> secured entity type in addition to the Edit permission for the <b>Dataflows - Expose</b> secured entity type. This is because the latest saved version must first be saved as a version before it can be exposed, which requires the Edit permission for the dataflow. |
| <b>Flow Defaults - Data Type Conversion</b> | Controls access to the <b>Flows &gt; Defaults &gt; Data Type Conversions</b> area in Management Console. All users have View access to data type conversion options. You cannot remove View access.  |
| <b>Flow Defaults - Malformed Records</b>    | Controls access to the <b>Flows &gt; Defaults &gt; Malformed Records</b> area in Management Console. All users have View access to malformed record options. You cannot remove View access.  |
| <b>Flow Defaults - Reports</b>              | Controls access to the <b>Flows &gt; Defaults &gt; Reports</b> area in Management Console. All users have View access to report options. You cannot remove View access.  |
| <b>Flow Defaults - Sort Performance</b>     | Controls access to the <b>Flows &gt; Defaults &gt; Sort Performance</b> area in Management Console. All users have View access to sort performance options. You cannot remove View access.   |
| <b>Flow History - Jobs</b>                  | Controls access to view job execution history in Enterprise Designer and Management Console.   |
| <b>Flow History - Process Flows</b>         | Controls access to process flow execution history in Management Console and Enterprise Designer.   |
| <b>Flow History - Transactions</b>          | Controls access to the <b>Flows &gt; History &gt; Transactions</b> area in Management Console.   |
| <b>Flow Scheduling</b>                      | Controls access to the <b>Flow &gt; Schedules</b> area in Management Console.  |
| <b>Jobs</b>                                 | Controls the ability to execute jobs in Enterprise Designer, Management Console, job executor, and the Administration Utility.   |
| <b>Notification - License Expiration</b>    | Controls access to configure license expiration notification emails in Management Console.   |
| <b>Notification - SMTP Settings</b>         | Controls access to the <b>System &gt; Mail Server</b> area in Management Console.  |

|  |   |
|--|---|
| <b>Process Flows</b>                       | Controls access to process flows in Enterprise Designer.<br><br><b>Note:</b> If a user does not have the Edit permission, the user will only see the exposed version and the last saved version in the <b>Versions</b> pane in Enterprise Designer.   |
| <b>Process Flows - Expose</b>              | Controls the ability in Enterprise Designer to make process flows available for execution.<br><br><b>Note:</b> In order to expose the latest saved version of the process flow (the version that's always at the top of the <b>Versions</b> pane in Enterprise Designer) the user must have the Edit permission for the <b>Process Flows</b> secured entity type in addition to the Edit permission for the <b>Process Flows - Expose</b> secured entity type. This is because the latest saved version must first be saved as a version before it can be exposed, which requires the Edit permission for the dataflow. |
| <b>Resources - Database Connections</b>    | Controls the ability to configure database connections in Management Console.   |
| <b>Resources - External Web Services</b>   | Controls access to managing external web services in Management Console.  |
| <b>Resources - File Server Connections</b> | Controls the ability to configure file servers in Management Console.   |
| <b>Resources - JDBC Drivers</b>            | Controls the ability to configure JDBC drivers in Management Console.   |
| <b>Resources - Remote Server</b>           | Controls access to the <b>Resources &gt; Remote Servers</b> area in Management Console.   |
| <b>Security - Access Control</b>           | Controls access to access control settings in the <b>System &gt; Security &gt; Access Control</b> area in Management Console.   |
| <b>Security - Access Token</b>             | Controls the ability to view users' tokens and delete tokens. A token facilitates authentication between a client and the server. Read permission allows you to see a list of the active tokens, each of which represent an active session. The Delete permission allows you to delete users' tokens, which ends their session.   |
| <b>Security - Directory Access</b>         | Controls the ability to enable or disable restrictions on server directory resources using the <b>System &gt; Security &gt; Directory Access</b> area in Management Console.  |
| <b>Security - Directory paths</b>          | Controls the ability to configure server directory resources the <b>System &gt; Security &gt; Directory Access</b> area in Management Console.  |
| <b>Security - Options</b>                  | Controls access to the ability to turn security on and off in the <b>System &gt; Security &gt; Roles</b> area in Management Console.  |
| <b>Security - Roles</b>                    | Controls access to role configuration in the <b>System &gt; Security &gt; Roles</b> area in Management Console.   |

|                                     |  |
|-------------------------------------|--|
| <b>Security - Directory paths</b>   | Controls the ability to configure server directory resources the <b>System &gt; Security &gt; Directory Access</b> area in Management Console. |
| <b>Security - Users</b>             | Controls access for managing user accounts in the <b>System &gt; Security &gt; Users</b> area in Management Console.                           |
| <b>Services</b>                     | Controls the ability to execute services through the API and web services.   |
| <b>Stages</b>                       | Controls whether exposed subflows are available as a stage in dataflows in Enterprise Designer.  |
| <b>System - Licensing</b>           | Controls access to the license information displayed in the <b>System &gt; Licensing and Expiration</b> area in Management Console.            |
| <b>System - Version Information</b> | Controls access to the <b>System &gt; Version</b> area in Management Console.  |
| <b>System Log</b>                   | Controls access to the system log in Management Console.   |

## Access Control

Access control settings work in conjunction with roles to define the permissions for a user. Roles define the permissions for categories of entities, such as all dataflows or all database resources, and access control settings define the permissions for specific entities, called *secured entities*. Examples of secured entities include specific jobs or specific database connections. For example, you may have a role that has granted the Modify permission to the secured entity type "Dataflows", but you may want to prevent users from modifying one specific dataflow. You could accomplish this by using access control to remove the Modify permission for the specific dataflow you do not want modified. You can specify access control settings for users and roles. Access control settings for a user override that specific user's permissions as granted by the user's roles. Access control settings for roles apply to all users who have that role.

## Configuring Access Control



Access control settings work in conjunction with roles to define the permissions for a user. Roles define the permissions for categories of entities, such as all dataflows or all database resources. Access control settings define the permissions for specific entities, such as a specific job or a specific database connection.

In order to configure access controls you must have View and Modify permissions to these secured entity types:

- Security - Access Control
- Security - Roles

- Security - Users

To configure access control:

1. In Management Console, go to **System > Security**.
2. Click the **Access Control** tab.
3. Click the Add button .
4. Do one of the following:
  - If you want to specify access controls for a role, click **Role**. The access control permissions you specify will affect all users who have the role you choose.
  - If you want to specify access controls for a single user, click **User**. The access control permissions you specify will only affect the user you choose.
5. Select the role or user for which you want to define access controls.
6. Click the Add button .
7. Select the secured entity type that contains the secured entity you want. For example, if you want to configure access control for a dataflow, choose **Platform.Dataflows**.
8. Choose the secured entity you want to configure access controls for, then click the **>>** button to add it to the **Selected Entities** list.
9. Click **Add**.
 

The secured entities you chose are displayed. The check boxes indicate the permissions in effect for the selected role or user.
10. Specify the permissions that you want to grant for each secured entity. Each secured entity can have one of the following permissions:

- The permission is inherited from the role.

---

- The permission is inherited from the role and cannot be overridden.

---

- The permission is granted, overriding the permission specified in the user or role.

---

- The permission is denied, overriding the permission specified in the user or role.

---

### Access Control Example

The following shows access control settings for the role RetentionDepartmentDesigner.

Home / System: Security / Add Access Control

## Add Access Control

Save Cancel

Role User

RetentionDepartmentDesigner

+ -

| Platform.Dataflow        |             | Create                   | View                                | Modify                              | Delete                   | Execute |
|--------------------------|-------------|--------------------------|-------------------------------------|-------------------------------------|--------------------------|---------|
| <input type="checkbox"/> |             | <input type="checkbox"/> | <input checked="" type="checkbox"/> | <input checked="" type="checkbox"/> | <input type="checkbox"/> |         |
| <input type="checkbox"/> | ExampleJob1 | <input type="checkbox"/> | <input type="checkbox"/>            | <input type="checkbox"/>            | <input type="checkbox"/> |         |

In this example, the **Platform.Dataflow** secured entity type is set to allow the View and Modify permissions but not the Delete permission. So by default, any user that has the RetentionDepartmentDesigner role would have these permissions for all dataflows. However, you want to prevent users with this role from modifying the ExampleJob1 dataflow only. So, you clear the checkbox in the Modify column for ExampleJob1. Now users with this role will not be able to modify this dataflow but will still be able to modify other dataflows.

### Related concepts

[Security Model](#) on page 28

[Security for the Location Intelligence Module](#) on page 48

[Users](#) on page 29

[Roles](#) on page 34

### Related tasks

[Disabling Role-Based Security](#) on page 37


[Adding a User](#) on page 30

[Creating a Role](#) on page 35

## Deleting Access Control Settings

When you delete access control settings for a user or role, the permission overrides defined by the access control settings are removed from the user or role. For users, this means that the permissions granted by the user's role will take effect without any overrides. For roles, this means that the permissions defined in the role itself will take effect without overrides.

1. Open Management Console.
2. Go to **System > Security**.

3. Click **Access Control**.
4. Check the box next to the user or role for whom you want to remove access control then click the Delete button .

## Security for the Location Intelligence Module

The Location Intelligence Module uses the role-based security that is used for the Spectrum™ Technology Platform. Because security is handled at the platform level, Management Console can be used to manage all Location Intelligence Module security activities. This includes setting permissions for named resources in addition to managing user accounts (that is, creating, modifying, and deleting user accounts).

### *Predefined Spatial Roles*

After you install the Location Intelligence Module, three predefined roles are available in Management Console:

- spatial-admin** The spatial-admin role provides full permissions (Create/View/Modify/Delete) for all named resources and datasets associated with named tables. These permissions are controlled using the Location Intelligence Module's secured entity types, Location Intelligence.Named Resources and Location Intelligence.Dataset.DML. Users of Location Intelligence Module services must have at least View permissions for the resources they use as well as for any dependent resources. See [Access Control for Datasets](#) on page 49 for more information on controlling dataset permissions.
- spatial-user** The spatial-user role provides View permissions to named resources only. These permissions are controlled using the Location Intelligence Module's secured entity type, Location Intelligence.Named Resources. Users of Location Intelligence Module services must have at least View permissions for the resources they use as well as for any dependent resources.
- spatial-dataset-editor** The spatial-dataset-editor role provides full permissions (Create/View/Modify/Delete) on datasets. These permissions are controlled using the Location Intelligence Module's secured entity type, Location Intelligence.Dataset.DML. See [Access Control for Datasets](#) on page 49 for more information on this role and controlling permissions on datasets.

Dataflow designers who require access to named resources need additional permissions beyond that of the "designer" role. For instructions on creating a spatial dataflow designer, see [Creating a Spatial Dataflow Designer](#) on page 52.

### *Custom Spatial Roles and Access Control Settings*

You can create custom roles based on the predefined spatial roles, assign them to user accounts, then fine-tune access to named resources for those roles and users by applying access control settings (overrides) to individual named resources, datasets, or to folders or directories. A typical



scenario and best practice for setting security for the Location Intelligence Module involves creating a role with no permissions, applying access control settings to that role (for example, allowing modify and delete permissions for named resources in a specific folder), then assigning that custom role as well as one of the predefined spatial roles to a user. Another common scenario involves establishing override permissions for a single user; for example, creating a user account which has view-only permissions to named resources, then applying access control settings to that user that allow modifying and deleting of named resources in a specific folder.

### *Folders*

Folder permissions are inherited by the resources and folders underneath as long as those resources and folders do not have any specific access control settings that override them. This is useful when you want to set permissions on a set of resources. You can make a folder accessible only to specified users or roles; other users will not see that folder or anything underneath it. For the Location Intelligence.Named Resources entity type, all listed resources that end with a forward slash (/) are folders or directories in the repository.

Permissions at the folder level, however, do not override permissions set at the lower, individual resource level. For example, if a folder has Create permissions for a specific role or user, but a single resource in the folder (such as a named table) has an access control setting to View permissions for that same role or user, the View (read-only) permissions for the single resource take precedence over the Create permissions for the folder.

#### **Related concepts**

[Users](#) on page 29

[Roles](#) on page 34

#### **Related tasks**

[Configuring Access Control](#) on page 45

## Access Control for Datasets

### *What is a Dataset?*

A dataset is a collection of data values in a tabular form that typically consists of rows (or records) and columns. In the Location Intelligence Module, a dataset can take the form of a .TAB file, a shapefile, a GeoPackage file, or a JDBC-based table such as an MS SQL Server table.

### *Benefits of Dataset Access Control*

Dataset access control allows administrators to disassociate the permissions of a named table from the editing permissions of the dataset that the named table points to. For example, as an administrator you can grant full editing (Create/Modify/Delete) permissions to a dataset while keeping read-only (View) permissions on the named table. When a user attempts to perform a data manipulation language (DML) operation (an insert, update, or delete operation using the Feature service or the Write Spatial Data stage), the user's permissions will be verified not only against the specified named table in the Location Intelligence.Named Resources entity type but also against the Location

Intelligence.Dataset.DML entity type. If View permissions are denied, the named table will not appear in the user's repository.

### *What is a Dataset Secured Entity?*

The LocationIntelligence.Dataset.DML secured entity is one of the two types of secured entities for the Location Intelligence Module. It controls DML permissions to datasets that are associated with named tables. When a named table is created or uploaded (using any tool, including Spatial Manager, the Administration Utility, the Named Resource Service, and WebDAV), a new LocationIntelligence.Dataset.DML secured entity is automatically created for the associated dataset of that named table. A user must have View permissions on a named table *and* Create/Modify/Delete permissions on the dataset in order to perform DML operations on writable (JDBC-based) tables. DML operations include insert, update, and delete operations performed using the Write Spatial Data stage or the Feature Service.

**Note:** Although you can set Create/Modify/Delete permissions on dataset secured entities for non-writable datasets such as .TAB files or shapefiles, you still cannot perform DML operations on these datasets.

**Tip:** The View permission on the secured entity for the dataset has no impact on its permissions. If you turn View permission off on a dataset secured entity you will still be able to view the data in the table. If you do not want a user to see a table, remove View permissions on the secured entity for the named resource instead.

When a named table is renamed, moved, or deleted, Spectrum Spatial will rename or delete the associated secured entity for the dataset.

### *Spatial Roles and Dataset Access*




Roles are used to grant or deny access to different parts of the system and help make permissions management easier. Three predefined roles for users of the Location Intelligence Module are available in Management Console:

- spatial-admin** The spatial-admin role provides full permissions (Create/View/Modify/Delete) for all named resources and datasets. A user with a spatial-admin role can view named resources as well as edit datasets.
- spatial-user** The spatial-user role provides View permissions to named resources only. A user with a spatial-user role can view resources but cannot edit datasets.
- spatial-dataset-editor** The spatial-dataset-editor role provides full permissions (Create/View/Modify/Delete) on datasets. For example, an administrator can easily grant full permissions to datasets by adding the spatial-dataset-editor role to a user who currently has the spatial-user role.

These predefined roles cannot be modified. You can, however, create custom roles based on the predefined spatial roles, assign them to user accounts, then fine-tune access on those roles and users by applying access control settings (overrides) to datasets, individual named resources, or folders containing named resources. See [Configuring Access Control](#) on page 45 for more information.

## Creating a Named Resources Administrator

To manage named resources in the repository using Management Console, a user must have an assigned role that allows full access to those resources in addition to the access that is provided by the predefined spatial roles. The predefined spatial roles cannot be modified and a predefined "Named Resources Administrator" role is not provided by the Spectrum™ Technology Platform; however, you can create such a role using a predefined spatial role as a base.

1. Open Management Console.
2. Go to **System > Security**.
3. Click **Roles**.
4. Check the box next to either the spatial-admin or spatial-user role to use as a starting point then click the Copy button . The spatial-admin role provides View, Modify, Create, and Delete permissions for the Location Intelligence Module.Named Resources and Location Intelligence Module.Dataset secured entity types; the spatial-user role provides View permissions.
5. In the **Role name** field, enter the name you want to give to this role (for example, "resource-admin").
6. Set additional permissions as follows for these secured entity types:  
**Database Resources:**
  - **Centrus Database Resources** to View/Modify/Create/Delete/Execute (if required)
  - **Enterprise Routing** to View/Modify/Create/Delete/Execute (if required)**Platform:**
  - **Resources - File Servers** to View
  - **Resources - JDBC Drivers** to View
  - **Services** to View/Modify/Execute
  - **System - Version Information** to View
7. Click **Save** to save the new resource-admin role.
8. Click **Users**.
9. Either select an existing user and click the Edit button  to modify it, or click the Add button  to create a new user.
10. Assign the new "resource-admin" role to the user account to allow it to manage named resources in Management Console.

The user now has the access required to manage named resources in Management Console.

## Creating a Spatial Dataflow Designer

To create dataflows for Location Intelligence Module stages and services, a user must have both the designer and spatial-user roles assigned. The spatial-user role provides View access to named resources under the Location Intelligence.Named Resources secured entity type. The designer role provides the necessary access to Platform secured entity types such as Dataflows.

1. In Management Console, go to **System > Security**.
2. Either select an existing user and click the Edit  button, or click the Add button  to create a new user.
3. In the Roles section, assign both the designer and spatial-user roles to the user account.

The user now has permission to view named resources and design dataflows using those resources for Location Intelligence Module stages and services.

## Limiting WebDAV Access to the Repository

WebDAV is used as a protocol to access resources within the Spectrum Spatial repository. By default, accessing the repository using WebDAV is not restricted to a particular server, rather open to all servers that can access the repository. You can restrict access to particular servers by modifying the spatial java property file. You can do this by adding the following property that includes a list of hostnames (IPs) that WebDAV is open to (comma separated). A Spectrum™ Technology Platform server restart is required after the change.

To limit repository access using WebDAV:

1. Open the `modules/spatial/java.properties` file in an editor.
2. Add the following property to the file.

```
repository.accesscontrol.allows=
```

3. Include a list of IP addresses that you want to allow WebDAV access. Multiple servers can be added using a comma separated list of IP addresses. Leaving the property empty disables all access using WebDAV for all servers except the machine where Spectrum™ Technology Platform is installed.

```
repository.accesscontrol.allows=192.168.2.1,192.168.2.2
```

- 
4. Restart the server.

Once finished, WebDAV access is limited for the repository.

## Using WebDAV with HTTPS

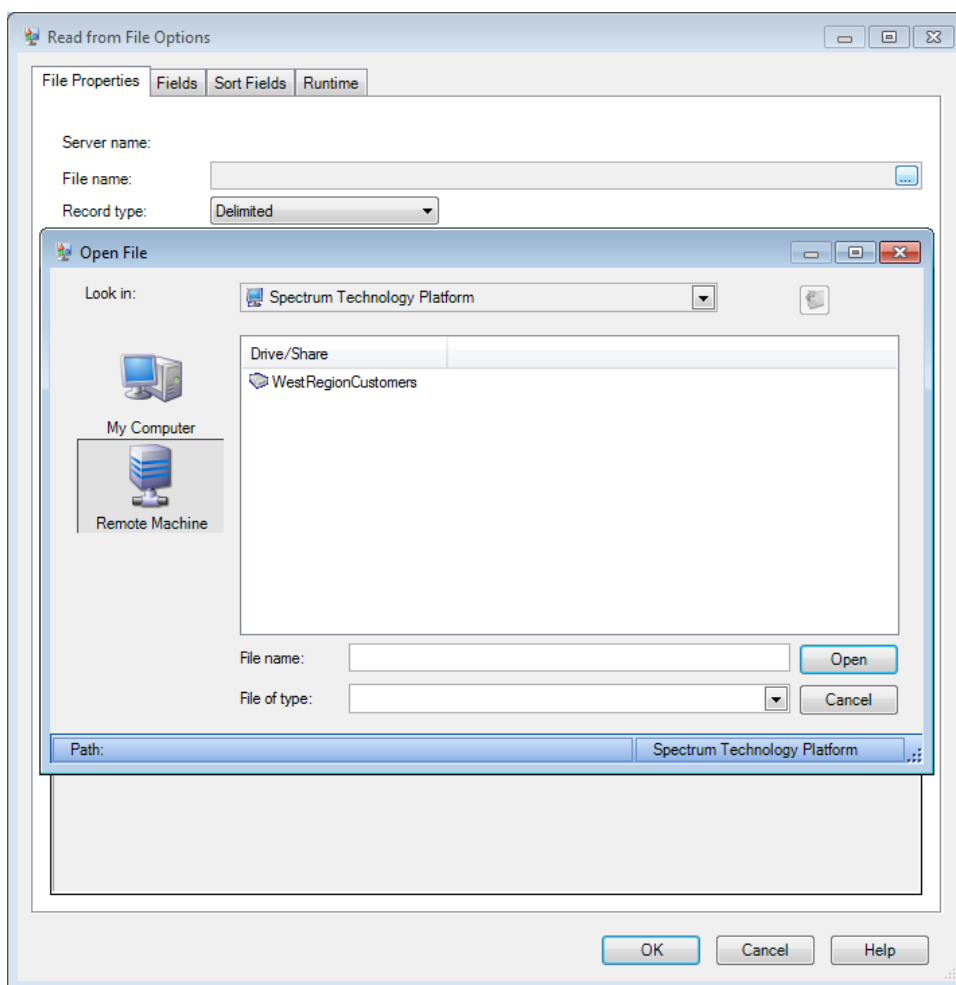
When communicating to the server over HTTPS to map a drive to the repository, a WebDAV client is required to use the TLS v1.2 protocol. For client machines running on Windows 7 SP1, Windows Server 2008 R2 SP1, and Windows Server 2012, you must apply a security patch and registry update to leverage this protocol.

1. On the client machine, apply the appropriate patch for the operating system from the Microsoft Knowledge Base: <https://support.microsoft.com/en-us/kb/3140245>
2. Follow the instructions in the KB article to update the registry to include support for TLS v1.2. The DefaultSecureProtocols value must be at least 0x00000800.
3. Restart the client machine after changing the registry entry.

## Limiting Server Directory Access

Users can browse the Spectrum™ Technology Platform server's folders when performing tasks that require them to select a file. For example, users can browse the server when selecting an input or output file in a source or sink stage in Enterprise Designer. As an administrator, you may want to restrict access so that sensitive portions of the server cannot be browsed or modified.

One way to prevent access to the server's file system by making sure that users do not have the Platform security permission **Security - Directory Paths**. This prevents access to all folders on the server. You can also prevent access to some folders on the server while allowing access to others. When you grant limited access, the folders you allow access to appear as the top-level folders in users' file browse windows. For example, if you allow users to only access a folder on the server named WestRegionCustomers, when users browse the server they would only see that folder, as shown here:




**Important:** There are two situations where users can view the server's entire file system even if you have granted only limited access:

- When browsing for a database file while creating a Spectrum database in Management Console
- When browsing for a JDBC driver file while creating a driver in Management Console

To prevent users from browsing the server's entire file system, use roles to limit the user's access to Spectrum databases and JDBC drivers.

To provide access to some folders on the server while restricting access to others, follow this procedure.

1. Open Management Console.
2. Go to **System > Security**.
3. Click **Directory Access**.
4. Click the Add button .
5. In the **Name** field, give a meaningful name for the folder to which you are granting access.

The name you provide here appears as the root name of the directory to users when browsing the server. In the example shown at the beginning of this topic, the name given to the accessible directory is WestRegionCustomers.

6. In the **Path** field, specify the folder to which you want to grant access. Users will be able to access all file and subfolders contained in the folder you specify.
7. Click **Save**.
8. If you want to grant access to additional folders, repeat the previous steps as needed.
9. Enforce the restrictions by setting the **Limit access to server directories** switch to **On**.

Users now have access only to the folders you have specified. Note that users must have the Platform security permission **Security - Directory Paths** in order to access server directories.

**Note:** If there are any dataflows that had previously accessed files that are no longer available because of file browsing restrictions, those dataflows will fail.

## Configuring HTTPS Communication

By default the Spectrum™ Technology Platform server uses HTTP for communication with Enterprise Designer, browser applications such as Management Console and Metadata Insights, as well as for handling web service requests and API calls, and for remote server communication. You can configure Spectrum™ Technology Platform to use HTTPS if you want to secure these network communications.

**Note:** Spectrum™ Technology Platform uses TLS 1.2 to encrypt communication. Applications that access Spectrum™ Technology Platform web services or the API must support TLS 1.2 in order to connect over HTTPS.

This procedure describes how to enable HTTPS communication on a single-server installation of Spectrum™ Technology Platform. If you want to use HTTPS and you are running Spectrum™ Technology Platform in a cluster, do not follow this procedure. Instead, configure the load balancer to use HTTPS for communication with clients. Communication between the load balancer and the Spectrum™ Technology Platform nodes, and between the nodes themselves, will be unencrypted because Spectrum™ Technology Platform clustering does not support HTTPS. The load balancer and the Spectrum™ Technology Platform servers in the cluster must be behind a firewall to provide a secure environment.

To configure HTTPS communication for a single-server installation of Spectrum™ Technology Platform:

1. Stop the Spectrum™ Technology Platform server.

- To stop the server on Windows, right-click the Spectrum™ Technology Platform icon in the Windows system tray and select **Stop Spectrum™**. Alternatively, you can use the Windows Services control panel and stop the Pitney Bowes Spectrum™ Technology Platform service.
  - To stop the server on Unix or Linux, source the `SpectrumLocation/server/bin/setup` script then execute the `SpectrumLocation/server/bin/server.stop` script.
2. Create a certificate signed by a trusted Certificate Authority (CA).
  3. Load the certificate into a JKS keystore. For more information, see [www.eclipse.org/jetty/documentation/current/configuring-ssl.html#loading-keys-and-certificates](http://www.eclipse.org/jetty/documentation/current/configuring-ssl.html#loading-keys-and-certificates).
  4. Create an XML file named `spectrum-override-container-ssl.xml` containing the following:

```
<beans xmlns="http://www.springframework.org/schema/beans"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:util="http://www.springframework.org/schema/util"
  xsi:schemaLocation="http://www.springframework.org/schema/beans
    http://www.springframework.org/schema/beans/spring-beans-3.0.xsd

    http://www.springframework.org/schema/util
    http://www.springframework.org/schema/util/spring-util-3.0.xsd">

  <bean id="defaultWebServerConnector"
    class="org.eclipse.jetty.server.ServerConnector">
    <constructor-arg ref="webServer"/>
    <constructor-arg>
      <bean class="org.eclipse.jetty.util.ssl.SslContextFactory">

        <property name="keyStorePath"
value="/SpectrumKeystore"/>
        <property name="keyManagerPassword" value="password"/>

        <property name="keyStorePassword" value="password"/>
      </bean>
    </constructor-arg>
    <property name="host" value="${spectrum.bind.address}"/>
    <property name="port" value="${spectrum.http.port}"/>
    <property name="idleTimeout" value="-1"/>
  </bean>
</beans>
```

5. Modify these lines as needed to reflect your environment:

|  |  |
|--|--|
| <code>&lt;property name="keyStorePath" value="/SpectrumKeystore"/&gt;</code> | Modify the value to be the full path to the Java keystore. |
| <code>&lt;property name="keyManagerpassword" value="password"/&gt;</code>    | Modify the value to be the password to the keystore.       |



```
<property name="keyStorePassword" value="password"/>
```

Modify the value to be the password to the key within the keystore.

6. Save the `spectrum-override-container-ssl.xml` file to `SpectrumLocation/server/app/conf/spring`.
7. Using a text editor, open the file `spectrum-container.properties` located in `SpectrumLocation/server/app/conf`. Uncomment and set these properties:

```
spectrum.http.port=port
spectrum.runtime.port=port
spectrum.runtime.hostname=dnsname
```

Where *port* is the network port to use for communication with the clients (for example 8443) and *dnsname* is the hostname of the Spectrum™ Technology Platform server. The port you specify must be the same for both `spectrum.http.port` and `spectrum.runtime.port`.

8. If you are configuring HTTPS communication for the Location Intelligence Module and Spectrum Spatial services, you must perform additional configuration steps prior to restarting the Spectrum™ Technology Platform server:
  - a) Modify the `java.properties` file (`SpectrumLocation\server\modules\spatial`) by changing all hostnames and ports to be exactly the same as the ones used for the Spectrum™ Technology Platform server. The hostname must match the DNS name of the server and the CN in the certificate. Set property `repository.useSecureConnection` to true. For example,

```
repository.host=www.example.com
repository.port=8443
repository.useSecureConnection=true
```

- b) In , change the URLs in these service configurations to use HTTPS:
    - Mapping (only necessary when accessing the Mapping Service via SOAP and when the `ReturnImage` parameter for a `RenderMap` request is `False`)
    - WFS
    - WMS
    - WMTS

For instructions, see the "Spatial Manager Guide" in the Utilities section of the *Spectrum Spatial Guide*.

9. Start the Spectrum™ Technology Platform server.
  - To start the server on Windows, right-click the Spectrum™ Technology Platform icon in the Windows system tray and select **Start Spectrum™**. Alternatively, you can use the Windows Services control panel to start the Pitney Bowes Spectrum™ Technology Platform service.
  - To start the server on Unix or Linux, execute the `SpectrumLocation/server/bin/server.start` script.

## Using WebFolders to Access Spectrum Spatial Repository Resources

To add or modify a named resource, you can copy it to or from the repository using a WebDAV tool. Using WebFolders is an easy way to access the Spectrum Spatial repository and the resources contained in it.

**Note:** To access the repository, you must be on the same machine where Spectrum™ Technology Platform and the repository are installed.

To configure a WebFolder on Windows 7:

1. Using Windows Explorer, select `Map Network Drive...`
2. In the pop-up window, click on the link 'Connect to a website...' to open the Add Network Location Wizard.
3. Click `Next` and select `Choose a custom network location`. Click `Next`.
4. In the `Internet or network address` field add the repository URL; for example, `http://<server>:<port>/RepositoryService/repository/default/`. Click `Next`.
5. Enter your credentials (username and password) if you are prompted for them.
6. Give this connection a name; for example, `Spectrum Spatial Repository`. Click `Next`.

Once finished, you will have a folder connection to the contents of the repository under your network places.

The WebFolder connection to the repository can be used like any other Windows Explorer folder.

**Note:** Be aware that if you use WebDAV to make changes to named resources or metadata resource records such that they are not located in the same folder or do not have the same base name, then Spatial Manager will no longer make matching changes to metadata records for move, rename or delete operations done on a resource.

## Web Service Authentication

Spectrum™ Technology Platform web services™ require authentication with valid user credentials. There are two methods for authenticating: Basic authentication and token authentication.

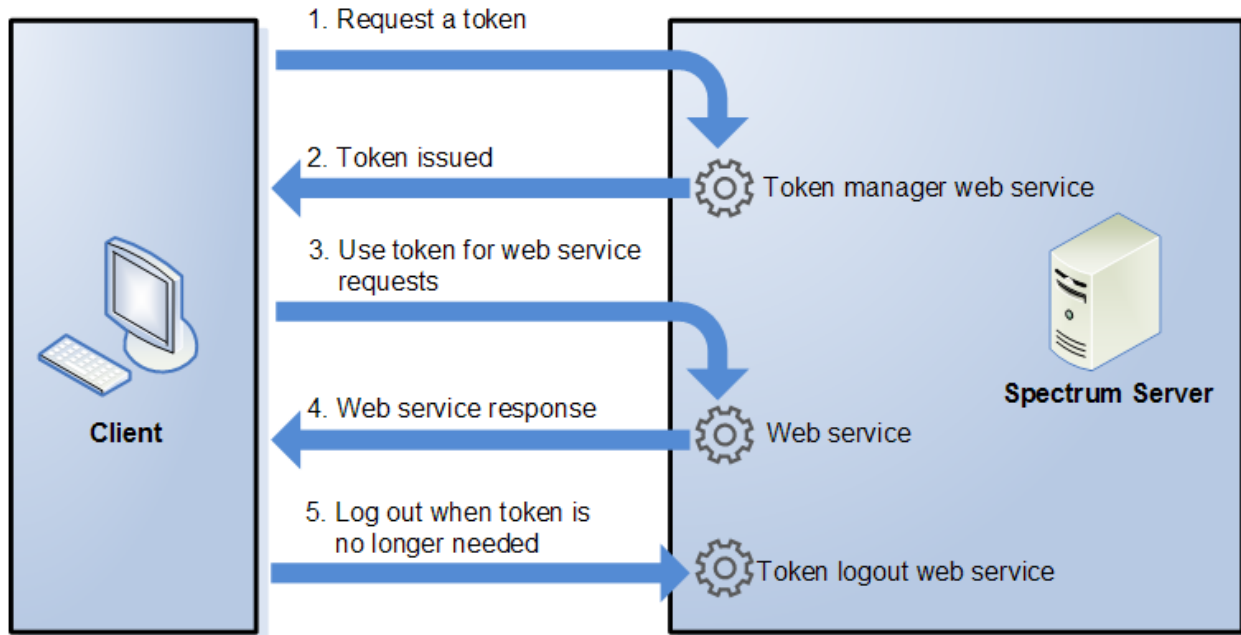
### *Basic Authentication*

With Basic authentication, the user ID and password are passed to Spectrum™ Technology Platform in the HTTP header of each request to the web service. Basic authentication is allowed by default, can be disabled. If Basic authentication is disabled you must use token authentication to access web services.

### Token Authentication

With token authentication, the requester obtains a token from the Spectrum™ Technology Platform server, then uses the token when sending a request to the web service. Instead of sending user credentials in each request, the token is sent to the server and the server determines if the token is valid.

This diagram illustrates the process:



1. Obtain a token from the Spectrum™ Technology Platform server by sending a request to the token manager service.
2. The token manager service issues a token. If you requested a session token it also issues a session ID.
3. Send a request to the desired web service with the token in the HTTP header. For session tokens, include the session ID in the HTTP header.
4. The web service issues a response. You can use the token to make additional web service requests to either the same web service or any other web service on the Spectrum™ Technology Platform server. There is no limit to the number of web service requests you can make with a token, but if the token has an expiration limit (also known as a time-to-live) it will become invalid after the time-to-live has elapsed. If the token is a session token, it will become invalid after 30 minutes of inactivity.
5. When the token is no longer needed you should log out by sending a request to the token logout web service. This will remove the token from the list of valid tokens on the Spectrum™ Technology Platform server.

## Disabling Basic Authentication for Web Services

Spectrum™ Technology Platform supports two types of authentication for web service requests: Basic authentication and token authentication. By default, both methods are enabled. If you want to require web service requests to use token authentication instead of Basic authentication, you can disable Basic authentication by following these steps.

**Note:** Disabling Basic authentication will cause existing clients to fail. For the Location Intelligence Module, WMS, WMTS, and WFS clients will either be expecting Basic authentication or no authentication. Leaving only token authentication will likely cause those clients to fail.

1. Stop the Spectrum™ Technology Platform server.
2. Open this file in a text editor:

```
SpectrumLocation/server/app/conf/spectrum-container.properties
```

3. Set this property to false:

```
spectrum.security.authentication.webservice.basicauth.enabled=false
```

4. Start the server.

## Disabling Authentication for Web Services

All services and access to resources used by Spectrum™ Technology Platform are configured, by default, with authentication turned on.

Service-level authentication can be disabled for all SOAP or REST web services (or both). This is useful if you have your own high-level authentication built into the solution that is using, for example, the Location Intelligence Module services.

To disable authentication for web services on the Spectrum™ Technology Platform :

1. Stop the Spectrum™ Technology Platform server.
2. Open the following file in a text editor:

```
SpectrumLocation\server\app\conf\spectrum-container.properties
```

3. Change the value of each property as needed. For example, to disable authentication for all SOAP services:

```
spectrum.security.authentication.webservice.enabled.REST=true  
spectrum.security.authentication.webservice.enabled.SOAP=false
```

**Note:** For the Location Intelligence Module, REST services also include OGC web services.

4. Save and close the properties file.
5. Start the Spectrum™ Technology Platform server.

Once finished, authentication is turned off for the type of web services that you specified.

### Related tasks

[Disabling Role-Based Security](#) on page 37

## Enabling CORS

Cross-Origin Resource Sharing (CORS) is a W3C standard that allows data sharing between domains. CORS enables web applications running in one domain to access data from another domain. By enabling CORS on your Spectrum™ Technology Platform server, you can allow web applications hosted in another domain to access Spectrum™ Technology Platform web services.

For example, say you have a web application hosted at **webapp.example.com**. This web application contains a JavaScript function that calls a Spectrum™ Technology Platform web service hosted at **spectrum.example.com**. Without CORS, you would need to use a proxy server to facilitate this request, which would add complexity to your implementation. With CORS, you do not need to use a proxy server. Instead, you can designate **webapp.example.com** as an "allowed origin", thus permitting Spectrum™ Technology Platform to respond to web service requests that originate from the domain **webapp.example.com**.

To enable CORS on your Spectrum™ Technology Platform server:

1. Stop the Spectrum™ Technology Platform server.
2. Open this file in a text editor:

```
SpectrumLocation/server/app/conf/spectrum-advanced.properties
```

3. Edit the following parameters.

#### **spectrum.jetty.cors.enabled**

Set this property to true to enable CORS. The default is false.

#### **spectrum.jetty.cors.allowedOrigins**

A comma separated list of origins that are allowed to access resources on the Spectrum™ Technology Platform server. The default value is `http://localhost:8080,http://localhost:443`, which allows access to resources using the default HTTP port 8080 and the default HTTPS port of 443.

If an allowed origin contains one or more asterisks ("\*"), for example `http://*.domain.com`, then asterisks are converted to `.*` and dots characters (".") are escaped to `\.` and the resulting allowed origin is interpreted as a regular expression.

Allowed origins can therefore be more complex expressions such as `https?://*.domain.[a-z]{3}` that matches http or https, multiple subdomains and any three-letter top-level domain (.com, .net, .org, etc.).

#### **spectrum.jetty.cors.allowedMethods**

A comma separated list of HTTP methods that are allowed to be used when accessing resources on the Spectrum™ Technology Platform server. The default value is POST,GET,OPTIONS,PUT,DELETE,HEAD.

#### **spectrum.jetty.cors.allowedHeaders**

A comma separated list of HTTP headers that are allowed when accessing resources on the Spectrum™ Technology Platform server. The default value is X-PINGOTHER, Origin, X-Requested-With, Content-Type, Accept. If the value is a single asterisk ("\*"), all headers will be accepted.

#### **spectrum.jetty.cors.preflightMaxAge**

The number of seconds that preflight requests can be cached by the client. The default value is 1800 seconds, or 30 minutes.

#### **spectrum.jetty.cors.allowCredentials**

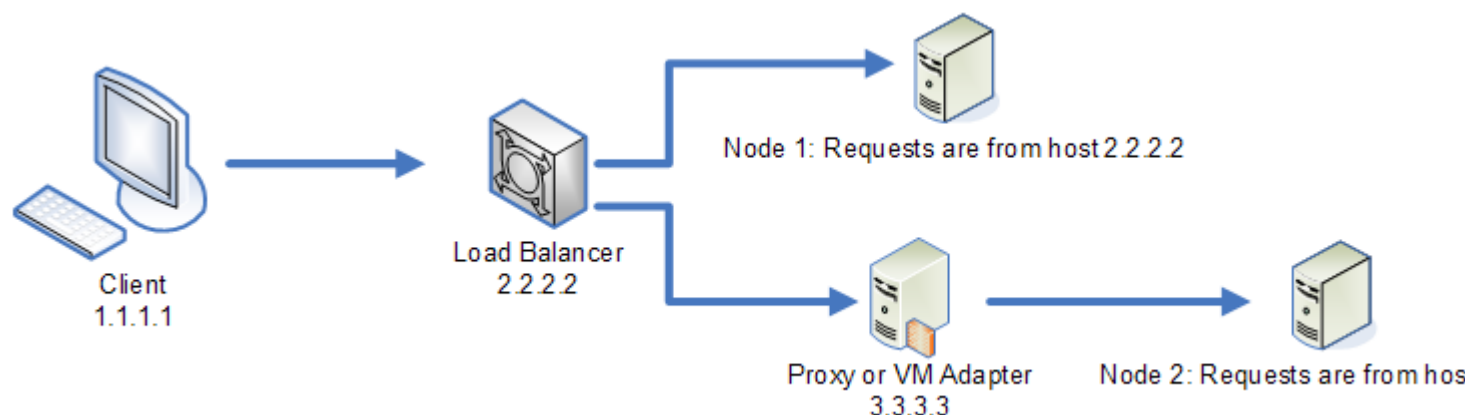
Indicates whether the resource allows requests with credentials. The default value is true.

4. Save and close the file.
5. Start the Spectrum™ Technology Platform server.

## Disabling Host Checks in Token Authentication

In token authentication, the Spectrum™ Technology Platform server examines the token presented by the client before responding to the request. The server checks the token to see if it has expired, if it is encrypted correctly, and if it is from the correct host. For session tokens, the server also checks the session ID. If any of these checks fail, the token is rejected and the server does not respond to the request.

In a clustered environment, it is possible that requests may be redirected in a way that makes the request appear to be coming from a different host than is specified in the token, resulting in "invalid token" errors. For example, say you have a cluster with two nodes as shown here:



Let's say that the client makes a request and the request is routed to Node 1. A token is created and tied to host 2.2.2.2 (the load balancer) since the node views the request as coming from the load balancer. If the next request from the client is routed to Node 2, the token will still be tied to host 2.2.2.2 but the request will appear to be coming from the proxy server, 3.3.3.3. In this case the node will reject the token because it appears that it is not associated with the host making the request.

In this situation you must configure the Spectrum™ Technology Platform server to ignore the host information included in the token. This should only be done if you have an environment where there are different network devices between the load balancer and the nodes. If all nodes are behind the same network device, there is no need to disable the host check.

**Note:** If you follow this procedure, client tokens will in effect become open tokens since the host check will be disabled. Session tokens will continue to be tied to a specific session ID but not a specific host.

1. Open the following properties file on the Spectrum™ Technology Platform server:

```
SpectrumLocation/server/app/conf/spectrum-container.properties
```

2. Set the following property to false.

```
spectrum.security.authentication.token.remoteClientCheck.enabled=false
```

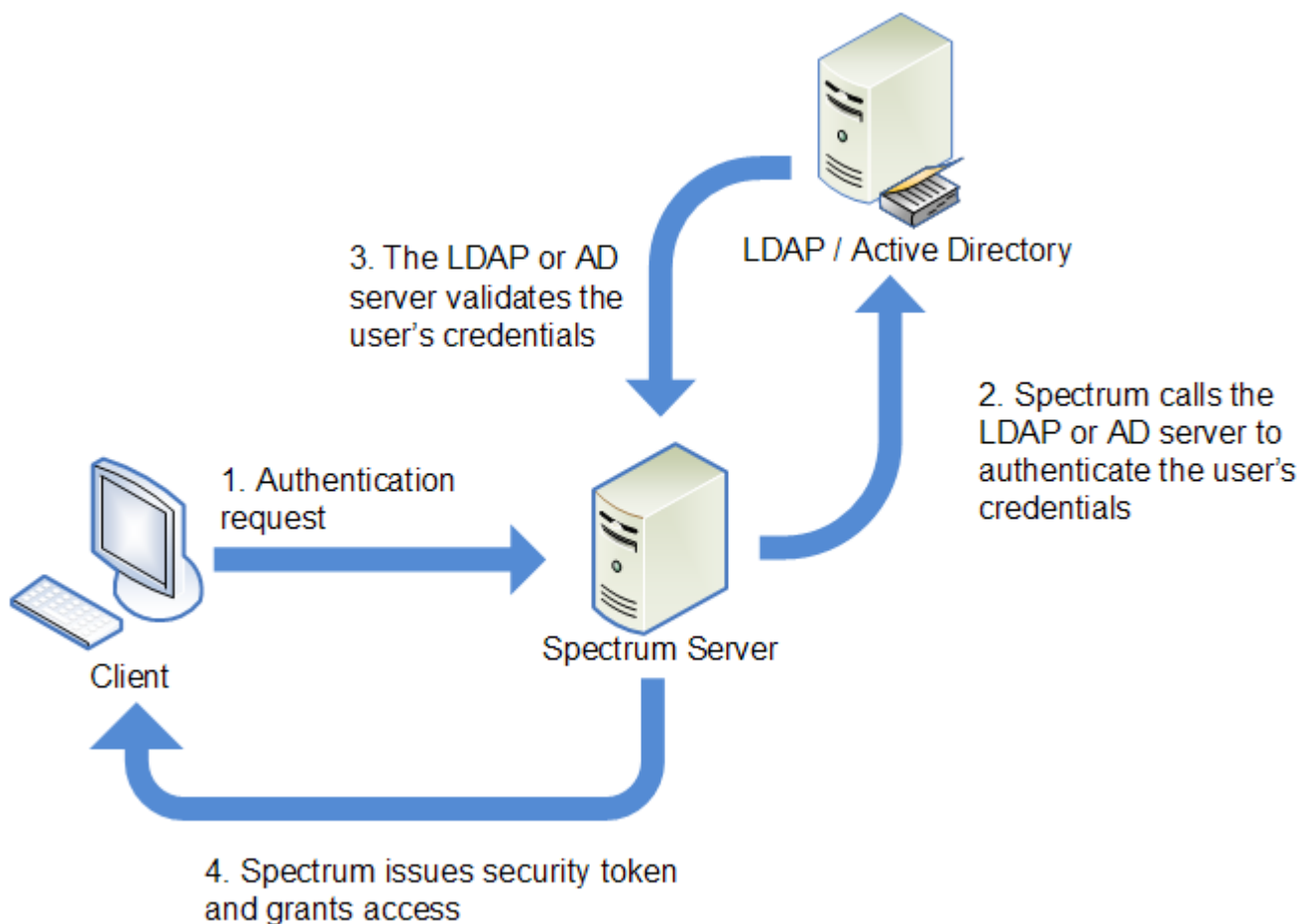
3. Save and close the properties file.
4. Repeat this process on all the nodes in the cluster.

## Using LDAP or Active Directory for Authentication

Spectrum™ Technology Platform can be configured to use an LDAP or Active Directory server for authentication. When a user logs in to Spectrum™ Technology Platform, the user's credentials are

verified using LDAP or AD. The system then checks to see if there is a Spectrum™ Technology Platform user with the same name. If there is, the user is logged in. If there is not, then a Spectrum™ Technology Platform user account is automatically created for the user and given the role `user`.

The following diagram illustrates this process:



Before configuring Spectrum™ Technology Platform to use a directory service for authentication, confirm that your directory service meets these requirements:

- For LDAP, the directory server must be LDAP Version 3 compliant.
- There are no specific requirements for the Active Directory server.
- Spectrum™ Technology Platform does not support SSL communication with the LDAP or Active directory server.

**Note:** We recommend that you contact Pitney Bowes Technical Support or Professional Services to guide you through this process.

1. If there are existing users configured in Management Console and you want to use them after you enable LDAP or Active Directory authentication, create those users in your LDAP or Active Directory system. Be sure to use the same user name as in Spectrum™ Technology Platform.



**Note:** You do not need to create the "admin" user in LDAP or Active Directory since this user will continue to use Spectrum™ Technology Platform for authentication after you enable LDAP or Active Directory authentication.

2. Stop the Spectrum™ Technology Platform server.
3. Turn on LDAP or Active Directory authentication:

- a) Open this configuration file in a text editor:

```
server\app\conf\spectrum-container.properties
```

- b) Set the property `spectrum.security.authentication.basic.authenticator` to LDAP:

```
spectrum.security.authentication.basic.authenticator=LDAP
```

The setting `LDAP` is used to enable Active Directory as well as LDAP.

- c) Save and close the file.

4. Configure the connection properties:

- a) Open this configuration file in a text editor:

```
server\app\conf\spring\security\spectrum-config-ldap.properties
```

- b) Modify these properties.

#### **spectrum.ldap.url**

The URL, including port, of the LDAP or Active Directory server. For example,

```
spectrum.ldap.url=ldap://ldapservers.example.com:389/
```

#### **spectrum.ldap.dn.format**

The format to use to search for the user account in LDAP or Active Directory. Use the variable `%s` for the user name. For example,

LDAP:

```
spectrum.ldap.dn.format=uid=%s,ou=users,dc=example,dc=com
```

Active Directory:

```
spectrum.ldap.dn.format=%s@example.com
```

#### **spectrum.ldap.dn.base**

The distinguished name (dn) to search for user accounts in LDAP or Active Directory. For example,

LDAP:

```
spectrum.ldap.dn.base=ou=users,dc=example,dc=com
```

Active Directory:

```
spectrum.ldap.dn.base=cn=Users,dc=example,dc=com
```

### **spectrum.ldap.search.filter**

A search filter to use when searching for attributes such as roles. The search filter can contain these variables:

- {user} is the user name logging into Spectrum™ Technology Platform
- {dn} is the distinguished name specified in `spectrum.ldap.dn.base`.

For example:

LDAP:

```
spectrum.ldap.search.filter=uid={user}
```

Active Directory:

```
spectrum.ldap.search.filter=userPrincipalName={dn}
```

### **spectrum.ldap.attribute.roles**

Optional. Specifies the LDAP or Active Directory attribute that contains the name of the Spectrum™ Technology Platform roles for the user. The role name you specify in the LDAP or Active Directory attribute must match the name of the role defined in Spectrum™ Technology Platform.

For example, to apply the roles defined in the attribute `spectrumroles` you would specify:

```
spectrum.ldap.attribute.roles=spectrumroles
```

If this attribute contains a role named `designer` then the `designer` role would be granted to the user.

You can only specify one attribute but the attribute may contain multiple roles. To specify multiple roles inside an attribute, separate each with a comma. You can also specify a multi-value attribute, with each instance of the attribute containing a different role. Only the roles specified in this one attribute are used in Spectrum™ Technology Platform. No other LDAP or Active Directory attributes will have any impact on Spectrum™ Technology Platform roles.

If the user has roles assigned to it in Spectrum™ Technology Platform, the user's permissions are the union of the roles from LDAP or Active Directory and the roles from Spectrum™ Technology Platform.

**Note:** When a user logs in for the first time, if the user does not have a Spectrum™ Technology Platform user account one is created automatically and given the role `user`. The effective permissions for the user are the union of the permissions in the `user` role and the roles specified in the attributes listed in the `spectrum.ldap.attribute.roles` property.

**Note:** When you view the user's roles in Management Console you will not see the roles assigned to the user by the `spectrum.ldap.attribute.roles` property.

#### **spectrum.ldap.pool.min**

The minimum size of the connection pool for connections to the LDAP or Active Directory server.

#### **spectrum.ldap.pool.max**

The maximum number of simultaneous connections to the LDAP or Active Directory server.

#### **spectrum.ldap.timeout.connect**

Specifies how long to wait to establish a connection to the LDAP or Active Directory server, in milliseconds. The default is 1000 milliseconds.

#### **spectrum.ldap.timeout.response**

Specifies how long to wait for a response from the LDAP or Active Directory server after the connection is established, in milliseconds. The default is 5000 milliseconds.

c) Save and close the properties file.

5. Start the Spectrum™ Technology Platform server.

If you are running Spectrum™ Technology Platform in a cluster, you must modify the `spectrum-container.properties` file and the `spectrum-config-ldap.properties` file on each of the servers in the cluster. Stop the server before modifying the file, then start the server after you are done modifying the file. If you mapped an LDAP attribute value to a role, this mapping will replicate to all nodes in the cluster, so you do not need to repeat the mapping procedure in the JMX console.

## Mapping LDAP Attribute Values to Roles

Before performing this procedure you must enable LDAP authentication. For more information, see [Using LDAP or Active Directory for Authentication](#) on page 63.

When you configure Spectrum™ Technology Platform to use LDAP or Active Directory for authentication, one of the configuration properties that you configure (the `spectrum.ldap.attribute.roles` property in the file `spectrum-config-ldap.properties`) specifies an LDAP attribute whose values determine the role to grant to a user. By default, the

attribute values must match the Spectrum™ Technology Platform role names exactly in order for the role to be granted. For example to grant the designer role, the attribute you specify must contain the value designer.

If the LDAP attribute value that you want to use does not match the role name in Spectrum™ Technology Platform, you can map the LDAP attribute value to a role name. You can also map an LDAP attribute value that has the same name as a Spectrum™ Technology Platform role to a different role. For example, one of the built-in roles is designer. If you have an LDAP attribute value named designer but you want it to map to another role, you could create a mapping.

1. Open a web browser and go to `http://server:port/jmx-console`

Where:

*server* is the IP address or hostname of your Spectrum™ Technology Platform server.

*port* is the HTTP port used by Spectrum™ Technology Platform. The default is 8080.

2. Click this property:

```
com.pb.spectrum.platform.common.security.ldap:mappings
```

**Note:** This property is only visible after you enable LDAP authentication and the server is fully started. If you have not enabled LDAP authentication, see [Using LDAP or Active Directory for Authentication](#) on page 63.

3. In the **addMapping** section, in the **ldapValue** field, enter the LDAP attribute value that you want to map to a Spectrum™ Technology Platform role.
4. In the **roleName** field, enter the Spectrum™ Technology Platform role that you want to map to the LDAP attribute value.
5. Click **Invoke**.

Users who have the LDAP attribute will now be granted the role you specified when they log in to Spectrum™ Technology Platform.

To remove a mapping, enter the LDAP attribute you want to un-map in the **ldapValue** field in the **removeMapping** section.

#### Example

For example, say you want to use a value in the `gecos` attribute to assign a role in Spectrum™ Technology Platform. If `gecos` contains the value `data-quality-user`, you want the user to be granted the designer role when logging in to Spectrum™ Technology Platform.

To accomplish this, you would specify the `gecos` attribute as the attribute to use assign roles by specifying this in the file `spectrum-config-ldap.properties`:

```
spectrum.ldap.attribute.roles=gecos
```

Then, you would map the data-quality-user value to the designer role in the JMX console:

JMX Console

Checking authority: superuser

**MBean: com.pb.spectrum.platform.common.security.ldap:mappings=LdapRoleMappings** All MBeans  
**Description: Handles mapping LDAP role attribute values to Spectrum roles**

| Attributes     |       |  |                  |  |
|----------------|-------|--|------------------|--|
| Name           | Value | Description                              | Type             |  |
| MappingsString |       | Defined role mappings (ldap -> spectrum) | java.lang.String |  |

| Operations           |             |                                       |                      |                                  |
|----------------------|-------------|---------------------------------------|----------------------|----------------------------------|
| Name                 | Return type | Description                           |                      |                                  |
| <b>addMapping</b>    | void        | Add an LDAP role mapping              |                      |                                  |
| Parameters           |             | <b>Name</b>                           | <b>Value</b>         | <b>Description</b>               |
|                      | ldapValue   |                                       | data-quality-user    | LDAP attribute value to map from |
|                      | roleName    |                                       | designer             | Spectrum role name to map to     |
|                      |             | <input type="button" value="Invoke"/> |                      |                                  |
| <b>removeMapping</b> | void        | Remove an LDAP role mapping           |                      |                                  |
| Parameters           |             | <b>Name</b>                           | <b>Value</b>         | <b>Description</b>               |
|                      | ldapValue   |                                       | <input type="text"/> | LDAP attribute value to map from |
|                      |             | <input type="button" value="Invoke"/> |                      |                                  |

As a result, any user that has the value data-quality-user in the `gecos` attribute will be granted the role designer.

# 3 - Data Sources

## In this section

---

|                              |     |
|------------------------------|-----|
| Introduction to Data Sources | 71  |
| Connection Types             | 75  |
| Deleting a Connection        | 124 |

## Introduction to Data Sources

A data source is a database, file server, cloud service, or other source of data that you want to process through Spectrum™ Technology Platform. Spectrum™ Technology Platform can connect to over 20 types of data sources.

To connect Spectrum™ Technology Platform to a data source, use Management Console to define the connection. For example, if you want to read data from an XML file into a dataflow, and the XML file is located on a remote file server, you would have to define a connection to the file server before you can define the input XML file in a dataflow. Similarly, if you want to write dataflow output to a database, you must first define the database as an external resource.

**Note:** If you want to read from or write to data located in a file on the Spectrum™ Technology Platform server itself there is no need to define a connection.

## Compression Support for Cloud File Servers

The file servers Amazon S3, Google cloud storage, and MS Azure Blobstore support the compressed formats `gzip (.gz)` and `zip (.zip)`.

The Spectrum™ Technology Platform handles the compression and decompression of the files written to and read from the file servers.

**Note:** You can use the same file server to handle both normal reads and writes of files as well as compression and decompression of files.

### *Reading a Compressed Format File*

While reading a file from the server, its compression format is derived from the metadata key property `Content-Encoding` received from the server.

### *Writing a Compressed Format File*

While writing a file to a server, mention the required compression format: `.gz` or `.zip`. The file is compressed based on the specified compression extension.

The metadata key property `Content-Encoding` is also set according to the selected compression format. This property value is passed to the cloud file server while writing the file to it.

## Supported Entities and Operations

Spectrum™ Technology Platform supports these entities and operations for each connection type:

| Connection Type | Table Properties   |
|-----------------|--|
| Marketo         | <p>The entities are of the following types:</p> <ol style="list-style-type: none"> <li>Entity</li> <li>Entity Update</li> </ol> <p><b>Note:</b> Entity Update is a virtual table used for Update on Lead entity. For example, <b>Merge_Leads</b> should be used for merging different Marketo Leads.</p>   |
| Netsuite        | <p>The entities are of the types:</p> <ul style="list-style-type: none"> <li>Standard Records</li> <li>Custom Records</li> <li>Joins</li> <li>Saved Searches</li> </ul> <p>On viewing the schema of the created physical model, for each entity type, the resultant schema is displayed.</p> <p>For example, for a Saved Search record, the schema of the search result is displayed. For a Join record, the schema of the join's result is displayed.</p> <p><b>Note:</b> In a NetSuite connection table, the primary key column is <code>internalId</code>.</p>  |
| SAP NetWeaver   | <p>The entity columns are of two types:</p> <ul style="list-style-type: none"> <li><b>Native:</b> Columns with native data types are displayed with their respective data types.</li> <li><b>Custom-defined:</b> Columns with custom-defined data types are displayed with a blank data type.</li> </ul> <p>To deploy a model store derived from a SAP connection, ensure its logical and physical models include only such entities whose columns are of native data types. If the physical and logical models have entities of custom-defined data types, the model store cannot be deployed.</p> <p>SAP OData Services supports a limited sets of operators listed here: <a href="http://www.odata.org/documentation/odata-version-2-0/uri-conventions/">http://www.odata.org/documentation/odata-version-2-0/uri-conventions/</a>. Example: <code>IsNull</code> or <code>Is not Null</code> operators are not supported.</p> |
| Siebel          | <p>The business components are displayed in the format <code>Business Object.Business Component</code>.</p>  |



| Connection Type | Table Properties   |
|-----------------|--|
| Splunk          | <p><b>Supported Operations</b> LIKE, ORDER BY, LIMIT, IN, BETWEEN, !=, &lt;=, &gt;=, &lt;, &gt;, multiple AND/OR operators</p> <p><b>Supported Functions</b></p> <p><b>String Functions</b> upper, lower, length, len, ltrim, rtrim, substring, max, min</p> <p><b>Mathematical Functions</b> abs, ceil, exp, floor, sqrt, round</p> <p><b>Note:</b> For all other query operations, use the Splunk <code>search</code> column as explained below.</p> <p>Metadata Insights provides a column <code>search</code> in the Splunk table using which you can look up the required data in the Splunk connection.</p> <p>While executing a <code>select</code> query on the <code>SplunkTable</code>, use the <code>search</code> column in the <code>where</code> clause in any of the below scenarios:</p> <ol style="list-style-type: none"> <li>1. To include such search criteria which cannot be specified using ANSI SQL syntax.</li> <li>2. To include such Splunk-specific search criteria which cannot be included as part of the main SQL query.</li> </ol> <p>For example, the below query looks for such a <code>_raw</code> value which contains the key <code>opp</code> with the value <code>ACC</code>.</p> <pre>select "_raw" from SplunkTable where "search"='search opp=ACC'</pre> |

| Connection Type | Table Properties   |
|-----------------|--|
| SuccessFactors  | <p>The entities are of two types:</p> <ol style="list-style-type: none"> <li>1. <b>Entity:</b> Represents a table representing a business entity.</li> <li>2. <b>Join:</b> Represents a mapping between any two <i>Entity</i> type tables: a parent table and any of its child tables.</li> </ol> <p><b>Note:</b> Links are not present between tables in the physical model schema derived from a SuccessFactors connection. This is because foreign keys are not present in SuccessFactors tables, and joins between tables are indicated by <i>Join</i> type tables in the Spectrum™ Technology Platform.</p> <p>The features of Join tables are:</p> <ol style="list-style-type: none"> <li>1. The name of a <i>Join</i> table indicates the two <i>Entity</i> tables which have been mapped together.</li> <li>2. Each record of a <i>Join</i> table contains the primary key from the parent entity and the columns of the respective child entity. Thus the particular parent entity's primary key is mapped to the child entity's details.<br/><br/>For example, <i>User#HR</i> is a <i>Join</i> table in which <i>User</i> is the parent entity and <i>Hr</i> is the child entity. This join represents all the users and their respective HR representatives. The join table <i>User#HR</i>, therefore, has the parent table <i>User</i>'s primary key <i>UserId</i>, which is mapped to the columns of the child table <i>HR</i>, like <i>hr_userId</i>, <i>hr_username</i>, <i>hr_email</i>, and so on.</li> <li>3. In case of <i>Join</i> tables, the <i>insert</i> and <i>update</i> functions work like the <i>upsert</i> function. This is because <i>Join</i> tables are not actual entities in SuccessFactors, but are mappings between entities and their navigation properties or child tables.<br/><br/>To <i>insert/update</i> any <i>Join</i> table, the parent entity is updated, while in child table a new record is inserted or the existing record is updated corresponding to the parent record.<br/><br/><b>Note:</b> While updating, the mapping between a parent and a child is modified. It is also possible to modify the individual attributes of the child as required.</li> </ol> |
| SugarCRM        | <p><b>Supported operations:</b></p> <p>LIKE (its operation is limited to picking up options starting with the specified value, such as in the statement WHERE name LIKE 's%', which picks up all the names starting with the alphabet S.), IS NULL, IS NOT NULL, IN, NOT IN, &gt;, &gt;=, &lt;, &lt;=, =, &lt;&gt;, AND, OR</p>  |


# Connection Types

## Connecting to Amazon

### Connecting to Amazon DynamoDB

In order for Spectrum™ Technology Platform to access data in Amazon DynamoDB you must define a connection to Amazon DynamoDB using Management Console. Once you do this, you can create flows in Enterprise Designer that can read data from, and write data to, Amazon DynamoDB.

**Note:** This connection is for use in the Metadata Insights module.

1. Open Management Console.
2. Go to **Resources > Data Sources**.
3. Click the Add button .
4. In the **Name** field, enter a name for the connection. The name can be anything you choose.

**Note:** Once you save a connection you cannot change the name.


5. In the **Type** field, choose **Amazon DynamoDB**.
6. In the **Access Key ID** field, enter the 20-character alpha-numeric sequence provided to you to access your Amazon AWS account.
7. In the **Secret Access Key** field, enter the 40-character key needed to authenticate the connection.
8. In the **Region** field, select the region of the Amazon AWS account.
9. To test the connection, click **Test**.
10. Click **Save**.

### Amazon DynamoDB Limitations

1. Hierarchical data types like lists, sets and maps are interpreted as String data types. This is because these data types are not supported.
2. Null values in a DynamoDB data source are interpreted as empty column values.

## Connecting to Amazon S3

In order for Spectrum™ Technology Platform to access data in Amazon S3 you must define a connection to Amazon S3 using Management Console. Once you do this, you can create flows in Enterprise Designer that can read data from, and write data to, Amazon S3.

1. Open Management Console.
2. Go to **Resources > Data Sources**.
3. Click the Add button .
4. In the **Name** field, enter a name for the connection. The name can be anything you choose.

**Note:** Once you save a connection you cannot change the name.

5. In the **Type** field, choose **Cloud**.
6. In the **Cloud service** field, choose **AmazonS3**.
7. In the **Bucket name** field, enter the bucket name as defined in your Amazon S3 cloud service. This is the bucket where Spectrum™ Technology Platform will read and write files.
8. Enter your access key and secret key assigned to you by Amazon..
9. In the **Storage Type**, field select the level of redundancy that you want to allow for data storage.

**Standard**                                      The default level of redundancy provided by Amazon S3.

**Reduced redundancy**                      Stores non-critical and easily-reproducible data at lower levels of redundancy. This provides fairly reliable storage at a lower cost.

10. In the **Encryption** section, select the encryption method for the data. You can select server side encryption, client side encryption, or both.

**Server side key**                      The data is encrypted and decrypted at the server side. Your data is transmitted in plain text to the Amazon cloud service where it is encrypted and stored. On retrieval, the data is decrypted by the Amazon cloud service then transmitted in plain text to your system.

You have two options for specifying the key:

- **AWS managed:** The key is automatically generated by the Amazon S3 cloud service.
- **Customer provided:** Enter the key to be used by the Amazon S3 cloud service to encrypt and decrypt the data on the server side.

**Client side key**                      The data is encrypted and decrypted at the client side. The data is encrypted locally on your client system then transmitted to the Amazon S3 cloud storage. On retrieval, the data is transmitted back in an encrypted format to your system and is decrypted on the client system.

**Client side key:** Enter the key to be used by your client system to encrypt and decrypt the data.

If you select both **Server side key** and **Client side key**, encryption and decryption is performed at both server and client sides. Data is first encrypted with your client side key and transmitted in an encrypted format to Amazon, where it is again encrypted with the server side key and stored. On retrieval, Amazon first decrypts the data with the server side key, transmitting the data in an encrypted format to your system, where it is finally decrypted with the client side key.

**Note:** To use the encryption feature of Amazon S3 cloud, you need to install the Amazon S3 Security JAR files. For more information, see [Using Amazon S3 Cloud Encryption](#) on page 77.

For more information about Amazon S3 encryption features, see:

[docs.aws.amazon.com/AmazonS3/latest/dev/UsingEncryption.html](https://docs.aws.amazon.com/AmazonS3/latest/dev/UsingEncryption.html)

11. If you want to set access permissions, in the **Permissions** section, click .

The three kinds of Grantees are:

|                           |   |
|---------------------------|---|
| <b>Everyone</b>           | Every one else other than Authenticated Users and Log Delivery group.                       |
| <b>AuthenticatedUsers</b> | For users who are logged into Amazon.   |
| <b>LogDelivery</b>        | For users who write activity logs in a user-specified Bucket, if Bucket Logging is enabled. |

For each Grantee, select the desired permissions:

|                      |   |
|----------------------|---|
| <b>Open/Download</b> | Allow the user to download the file.                          |
| <b>View</b>          | Allow the user to view the current permissions on the file.   |
| <b>Edit</b>          | Allow the user to modify and set the permissions on the file. |

12. To test the connection, click **Test**.  
13. Click **Save**.

### [Using Amazon S3 Cloud Encryption](#)

To use the encryption security feature of the Amazon S3 cloud service, you need to download security JAR files and place them on the Spectrum™ Technology Platform server. Using encryption is optional.

1. Go to the download site.

For Windows and Linux platforms using Java 7, the JAR files can be downloaded from:

<http://www.oracle.com/technetwork/java/javase/downloads/jce-7-download-432124.html>

For AIX platforms using Java 7, the JAR files can be downloaded from:

<https://www14.software.ibm.com/webapp/iwm/web/preLogin.do?source=jcesdk>

2. Download these two JAR files:

- local\_policy.jar
- US\_export\_policy.jar

3. Place the JAR files in the location:


```
SpectrumFolder\Pitney Bowes\Spectrum\java64\jre\lib\security
```

4. Restart the server.

### Connecting to Amazon SimpleDB

In order for Spectrum™ Technology Platform to access data in Amazon SimpleDB you must define a connection to Amazon SimpleDB using Management Console. Once you do this, you can create flows in Enterprise Designer that can read data from, and write data to, Amazon SimpleDB.

**Note:** This connection is for use in the Metadata Insights module.

1. Open Management Console.
2. Go to **Resources > Data Sources**.
3. Click the Add button .
4. In the **Name** field, enter a name for the connection. The name can be anything you choose.

**Note:** Once you save a connection you cannot change the name.

5. In the **Type** field, choose **Amazon SimpleDB**.
6. In the **Access Key ID** field, enter the 20-character alpha-numeric sequence provided to you to access your Amazon AWS account.
7. In the **Secret Access Key** field, enter the 40-character key needed to authenticate the connection.
8. To test the connection, click **Test**.
9. Click **Save**.

### Amazon SimpleDB Limitations


In the Write to DB stage, the write mode **Update** is not available when writing to an Amazon SimpleDB table. The **Insert** option handles both insert and update operations. It differentiates between an insert and an update using the unique value of the `ItemName` column which is present in all Amazon SimpleDB tables.

This limitation is due to the fact that an update query requires a Primary Key for each record of the table to be updated, which is not supported by Amazon SimpleDB databases.

## Connecting to Apache Cassandra

In order for Spectrum™ Technology Platform to access data in a Cassandra database, you must define a connection to the Cassandra database using Management Console. Once you do this, you can create flows in Enterprise Designer that can read data from, and write data to, the Cassandra database.

**Note:** This connection is for use in the Metadata Insights module.


1. Open Management Console.
2. Go to **Resources > Data Sources**.
3. Click the Add button .
4. In the **Name** field, enter a name for the connection. The name can be anything you choose.

**Note:** Once you save a connection you cannot change the name.

5. In the **Type** field, choose **Apache Cassandra**.
6. In the **Host** field, enter the machine name or the IP on which the Apache Cassandra database is installed.
7. In the **Keyspace** field, enter the name of the keyspace of the data center you wish to access.
8. In the **Port** field, enter the port on which the Apache Cassandra database is configured.
9. Enter the user name and password to use to authenticate to the Cassandra database.
10. In the **Consistency Level** field, select how consistent data rows must be across replica nodes for a successful data transaction. This can be at least one, or all, or a combination of available nodes.
11. In the **Fetch Size**, enter the number of resultset rows you wish to fetch on each read transaction.
12. To test the connection, click **Test**.
13. Click **Save**.

## Connecting to Azure Cloud

In order for Spectrum™ Technology Platform to access data in Microsoft Azure, you must define a connection to Microsoft Azure using Management Console. Once you do this, you can create flows in Enterprise Designer that can read data from, and write data to, Microsoft Azure.

1. Open Management Console.
2. Go to **Resources > Data Sources**.
3. Click the Add button .
4. In the **Name** field, enter a name for the connection. The name can be anything you choose.

**Note:** Once you save a connection you cannot change the name.


5. In the **Type** field, choose **Cloud**.
6. In the **Cloud service** field, choose **AzureBlobStorage**.
7. In the **Protocol** field select whether you want the connection between Azure and Spectrum™ Technology Platform to use HTTP or HTTPS.
8. In the **Account Name** field, enter the name of your Azure storage account.
9. In the **Access Key** field, enter the access key to your Azure account.
10. To test the cloud connection, click **Test**.
11. Click **Save**.

## Connecting to a Flat File

### Connecting to a Delimited Flat File

To add a new Delimited Flat File connection navigate to **Connections > Flat File**, and select the **Record Type** as **Delimited**. Enter the file's access details and content type details to allow the Data Federation Module to read it correctly.

**Note:** This connection is for use in the Metadata Insights module.

1. Go to **Connections > Flat File**.
2. By default, the screen opens in the create mode. Otherwise, click  to add a new Flat File connection.
3. Enter a **Connection Name** for the Flat File data connection.
4. Enter the **File Path** by clicking **Browse** and selecting the directory of the file.
5. Select the **Character Encoding** of the flat file from the drop-down.
6. Select the **Record Type** as **Delimited**.
7. In **Field Delimiter**, select the expected separator between any two fields of a file record.
8. Select the **Text Qualifier (optional)**, if any, that encloses the field values of a file record.
9. In **Line Separator**, the value `Default` is selected, indicating that the expected line separator depends on whether Spectrum™ Technology Platform is running on a Unix or Windows system.
10. To specify whether the first row of the file is a header row, shift the **First Row is Header Row** slider to either **Yes** or **No**.
11. To specify whether the data type of the various fields in any record of the file should be automatically detected, shift the **Detect data type from file** slider to either **Yes** or **No**.
12. To skip malformed records during file parsing, shift the **Skip Malformed Records** slider to **On**.
13. Click **Test**.  
A message confirms the successful test of the connection.



14. Click **Save**.

A message confirms the successful creation of the connection.

In order to view a sample record fetched using the created Delimited Flat File connection, click **Preview** in the header bar. File records will be fetched and the Fields sorted according to the details provided by you.

**Related reference**


[Date Time Formats in a File Connection](#) on page 82

While reading date and time values from files using a File Connection in the Module, the values need to adhere to certain specific date-time formats.

**Connecting to a Fixed Width Flat File**

To add a new Fixed Width Flat File connection navigate to **Connections > Flat File** and select the **Record Type** as **Fixed Width**. Enter the file's access details and content type details to allow the Data Federation Module to read it correctly.

**Note:** This connection is for use in the Metadata Insights module.

1. Go to **Connections > Flat File**.
2. By default, the screen opens in the create mode. Otherwise, click  to add a new Flat File connection.
3. Enter a **Connection Name** for the Flat File data connection.
4. Enter the **File Path** by clicking **Browse** and selecting the directory of the file.
5. Select the **Character Encoding** of the flat file from the drop-down.
6. Select the **Record Type** as **Fixed Width**.
7. In the **Record Length** field, enter the total number of characters in a file record.  
Repeat Step 8 to Step 13 to enter details of all fields expected in a file record.
8. Click **Add Field** to add a row for a field in a file record.
9. In the **Name** column, enter the name for the field value.
10. In the **Type** column, select the data type of the field value.
11. In the **Start Position** column, enter the position in the file record at which of the field value begins.  
For the first field in a file record, the **Start Position** counting begins from 1.
12. In the **Length** field, enter the total number of characters the field covers, including the character at the **Start Position**.

The sum of the **Start Position** and **Length** values for any field should be less than or equal to the **Record Length**

If the File Record is:

```
01234Rob Smith29PitneyBowes
```

**Record Length = 27**

For the field 'Name':

**Start Position** = 6

**Length** = 9

```
Name = Rob Smith
```

13. Check the **Trim** checkbox if you wish to trim any white spaces at the beginning and/or end of a field value.
14. Click **Test**.  
A message confirms the successful test of the connection.
15. Click **Save**.  
A message confirms the successful creation of the connection.

In order to view a sample record fetched using the created Fixed Width Flat File connection, click **Preview** in the header bar. File records will be fetched and the Fields sorted according to the details provided by you.

#### **Related reference**

[Date Time Formats in a File Connection](#) on page 82

While reading date and time values from files using a File Connection in the Module, the values need to adhere to certain specific date-time formats.

#### [Date Time Formats in a File Connection](#)

While reading date and time values from files using a File Connection in the Module, the values need to adhere to certain specific date-time formats.

#### *Accepted Date Time Formats*

- Date: "yyyy-mm-dd"
- Datetime: "yyyy-mm-dd HH:mm:ss"
- Time: "HH:mm:ss"

These are as per standard date-time notations.

#### *Delimited Files*

If the **Detect type** feature is turned on while configuring the Delimited File Connection, then the date and time values in the file records, which adhere to the above formats, are automatically detected as Date type.

If a date-time value does not adhere to one of the accepted formats, the value is read as a String type value instead of a Date type value.

#### *Fixed Width Files*

For Fixed Width files, date type values are configured while creating the Fixed Width File Connection. Hence these values are read as Date type values, irrespective of whether they adhere to the accepted formats or not.

If the date-time value in a Fixed Width file does not adhere to the accepted formats, it needs to be handled using **Transformations** at the logical model creation stage by applying the below *Conversion* category function to the value:

```
parsedate(String date, String format)
```

In this, the *date* is the value received from the file, while the *format* is the date-time format in which the value is received from the file. This helps to parse the date-time value correctly.

For example, if the *date* = 23-Feb-2008, then the *format* = dd-*MMM*-*yyyy*.

### Resulting Value Formats

While previewing data in a model store:

- If the value has been read as a date/time value, it is reflected in one of the accepted date/time formats in the preview.
- If the value has been read as a String value, it is reflected as it is in the preview.

### Related tasks

[Connecting to a Delimited Flat File](#) on page 80


[Connecting to a Fixed Width Flat File](#) on page 81

## Connecting to an FTP Server

In order for Spectrum™ Technology Platform to access files on an FTP server you must define a connection to the FTP server using Management Console. Once you do this, you can create dataflows in Enterprise Designer that can read data from, and write data to, files on the FTP server.

Before connecting to an FTP server, verify that the timeout settings on your FTP server are appropriate for the jobs that will use this connection. Depending on the design of a job, there may be periods of time when the connection is idle, which may cause the connection to time out. For example, you may have a dataflow with two Read from File stages connected to an Import To Hub stage. While the Import To Hub stage is reading records from one Read from File stage, the other will be idle, possibly causing its connection to the FTP server to time out. Consider setting the timeout value on your FTP server to 0 to prevent connections from timing out.

**Note:** The FTP server must be running in active connection mode. Passive connection mode is not supported.


1. Open Management Console.
2. Go to **Resources > Data Sources**.
3. Click the Add button .
4. In the **Name** field, enter a name for the connection. The name can be anything you choose.

**Note:** Once you save a connection you cannot change the name.

5. In the **Type** field, choose **FTP**.
6. In the **User name** and **Password** fields, enter the credentials to use to authenticate to the FTP server. This is required only if the FTP server requires it.
7. In the **Host** field, enter the hostname or IP address of the FTP server.
8. In the **Port** field, enter the network port number the server uses for FTP.
9. Click **Test** to verify that the Spectrum™ Technology Platform server can connect to the FTP server.
10. Click **Save**.

## Connecting to Google Cloud Storage

In order for Spectrum™ Technology Platform to access data in Google Cloud Storage, you must define a connection to Google Cloud Storage using Management Console. Once you do this, you can create flows in Enterprise Designer that can read data from, and write data to, Google Cloud Storage.

1. Open Management Console.
2. Go to **Resources > Data Sources**.
3. Click the Add button .
4. In the **Name** field, enter a name for the connection. The name can be anything you choose.

**Note:** Once you save a connection you cannot change the name.

5. In the **Type** field, choose **Cloud**.
6. In the **Cloud service** field, choose **GoogleCloudStorage**.
7. In the **Bucket name** field, enter the bucket name as defined in your Google cloud service. This is the bucket where Spectrum™ Technology Platform will read and write files.
8. Enter your the application name, service account, and private key file provided by Google.

**Note:** Ensure the private key file is present on the Spectrum™ Technology Platform server.

9. You can set access permissions in the **Permissions** section.

**Manage your data and permission**      Allows the user to manage the data and permissions.

**View your data**                              Allows the user to view data.

**Manage your data**                          Allows the user to manage data.

10. To test the connection, click **Test**.


11. Click **Save**.

For more information, see Google's [Service Account Authentication](#) documentation.

## Connecting to Hadoop

In order for Spectrum™ Technology Platform to access data in Hadoop, you must define a connection to Hadoop using Management Console. Once you do this, you can create flows in Enterprise Designer that can read data from, and write data to, Hadoop.

**Attention:** Spectrum™ Technology Platform does not support Hadoop 2.x for Kerberos on Windows.

1. Open Management Console.
2. Go to **Resources > Data Sources**.
3. Click the Add button .
4. In the **Name** field, enter a name for the connection. The name can be anything you choose.

**Note:** Once you save a connection you cannot change the name.

5. In the **Type** field, choose **HDFS**.
6. In the **Host** field, enter the hostname or IP address of the NameNode in the HDFS cluster.
7. In the **Port** field, enter the network port number.
8. In the **User** field, select the method for authenticating to HDFS:

**Server user**      Select this option if authentication is enabled in your HDFS cluster. This option will use the user credentials that the Spectrum™ Technology Platform server runs under to authenticate to HDFS.

**User name**      Select this option if authentication is disabled in your HDFS cluster.

9. Check **Kerberos** if you want to enable Kerberos authentication feature for this HDFS file server connection.
10. If you have opted to enable **Kerberos** authentication, then enter the path of the keytab file in the **Keytab file path** field.

**Note:** The keytab file must be on the Spectrum™ Technology Platform server.

11. In the **Protocol** field, select the method of communication with HDFS:

**WEBHDFS**      Select this option if the HDFS cluster is running HDFS 1.0 or later. This protocol supports both read and write operations.

**HFTP**          Select this option if the HDFS cluster is running a version older than HDFS 1.0, or if your organization does not allow the WEBHDFS protocol. This protocol only supports the read operation.

**HAR** Select this option to access Hadoop archive files. If you choose this option, specify the path to the archive file in the **Path** field. This protocol only supports the read operation.

12. If you selected the WEBHDFS protocol, expand **Advanced server options**. Review the settings and make any changes that are necessary.

**Replication factor** Specifies how many data nodes to replicate each block to. For example, the default setting of 3 replicates each block to three different nodes in the cluster. The maximum replication factor is 1024.

**Block size** Specifies the size of each block. HDFS breaks up a file into blocks of the size you specify here. For example, if you specify the default 64 MB, each file is broken up into 64 MB blocks. Each block is then replicated to the number of nodes in the cluster specified in the **Replication factor** field.

**File permissions** Specifies the level of access to files written to the HDFS cluster by Spectrum™ Technology Platform. You can specify read and write permissions for each of these options:


**Note:** The **Execute** permission is not applicable to Spectrum™ Technology Platform.

**User** This is the user specified above, either **Server user** or the user specified in the **User name** field.

**Group** This refers to any group of which the user is a member. For example, if the user is john123, then Group permissions apply to any group of which john123 is a member.

**Other** This refers to any other users as well as groups of which the specified user is not a member.

In the grid below the **File permissions** table, specify the server properties for Hadoop to ensure that the sorting and filtering features work as desired when the connection is used in a stage or activity.

To add a new property, click . Then, define the properties, as described in this table, depending on the stage or activity that will use the Hadoop connection, and whether Hadoop 1.x or Hadoop 2.x is being used.

## Stage or Activity using the Required Server Properties HDFS Connection

---

- Stage **Read from Sequence File**
- Activity **Run Hadoop Pig**

## Stage or Activity using the Required Server Properties HDFS Connection

---

### *Hadoop 1.x Parameters*

#### **fs.default.name**

Specifies the node and port on which Hadoop runs.

For example,

```
hdfs://152.144.226.224:9000
```

#### **mapred.job.tracker**

Specifies the hostname or IP address, and port on which the MapReduce job tracker runs. If the host name is entered as local, then jobs are run as a single map and reduce task.

For example,

```
152.144.226.224:9001
```

#### **dfs.namenode.name.dir**

Specifies where on the local files system a DFS name node should store the name table. If this is a comma-delimited list of directories, then the name table is replicated in all of the directories, for redundancy.

For example,

```
file:/home/hduser/Data/namenode
```

#### **dfs.datanode.data.dir**

Specifies where on the local file system a DFS data node should store its blocks. If this is a comma-delimited list of directories, then data will be stored in all the named directories that are usually on different devices. Directories that do not exist are ignored.

For example,

```
file:/home/hduser/Data/datanode
```

#### **hadoop.tmp.dir**

Specifies the base location for other temporary directories.



Stage or Activity using the Required Server Properties  
HDFS Connection

---

For example,

`/home/hduser/Data/tmp`

*Hadoop 2.x Parameters*

## Stage or Activity using the Required Server Properties HDFS Connection

---

### **fs.defaultFS**

Specifies the node and port on which Hadoop runs.

For example,

```
hdfs://152.144.226.224:9000.
```

**Note:** We recommend that the parameter name `fs.defaultFS` be used in Spectrum™ Technology Platform 11 SP1 and later.

### **yarn.resourcemanager.resource-tracker.address**

Specifies the hostname or IP-address of the Resource Manager.

For example,

```
152.144.226.224:8025
```

### **yarn.resourcemanager.scheduler.address**

Specifies the address of the Scheduler Interface.

For example,

```
152.144.226.224:8030
```

### **yarn.resourcemanager.address**

Specifies the address of the Applications Manager interface that is contained in the Resource Manager.

For example,

```
152.144.226.224:8041
```

### **mapreduce.jobhistory.address**

Specifies the host name or IP address, and port on which the MapReduce Job History Server is running.

For example,

```
152.144.226.224:10020
```

### **mapreduce.application.classpath**

Specifies the CLASSPATH for Map

## Stage or Activity using the Required Server Properties HDFS Connection

---

Reduce applications. This CLASSPATH denotes the location where classes related to Map Reduce applications are found. Separate entries with a comma.

For example

```
$HADOOP_CONF_DIR,  
$HADOOP_COMMON_HOME/share/hadoop/common/*,  
$HADOOP_COMMON_HOME/share/hadoop/common/lib/*,  
$HADOOP_HDFS_HOME/share/hadoop/hdfs/*,  
$HADOOP_HDFS_HOME/share/hadoop/hdfs/lib/*,  
$HADOOP_MAPRED_HOME/share/hadoop/mapreduce/*,  
$HADOOP_MAPRED_HOME/share/hadoop/mapreduce/lib/*,  
$HADOOP_YARN_HOME/share/hadoop/yarn/*,  
$HADOOP_YARN_HOME/share/hadoop/yarn/lib/*
```

### **mapreduce.app-submission.cross-platform**

Handles various platform issues that arise if your Spectrum server runs on a Windows machine and you install Cloudera on it. If your Spectrum server and Cloudera are running on different operating systems, then enter the value of this parameter as `true`. If not, mark it as `false`.

**Note:** Cloudera does not support Windows clients. Configuring this parameter is a workaround and not a solution platform-related issues.

If you have checked the **Kerberos** checkbox, then add these Kerberos configuration properties:

Stage or Activity using the Required Server Properties  
HDFS Connection

---

**hadoop.security.authentication**

Specifies the type of authentication security being used. Enter the value `kerberos`.

**yarn.resourcemanager.principal**

Specifies the Kerberos principal being used for the resource manager for your Hadoop YARN resource negotiator.

For example,

`yarn/_HOST@HADOOP.COM`

**dfs.namenode.kerberos.principal**

Specifies the Kerberos principal being used for the namenode of your Hadoop Distributed File System (HDFS).

For example,

`hdfs/_HOST@HADOOP.COM`

**dfs.datanode.kerberos.principal**

Specifies the Kerberos principal being used for the datanode of your Hadoop Distributed File System (HDFS).

For example,

`hdfs/_HOST@HADOOP.COM`

---

## Stage or Activity using the Required Server Properties HDFS Connection

---

- Stage **Read from File**
  - Stage **Write to File**
  - Stage **Read from Hive ORC File**
  - Stage **Write to Hive ORC File**
- Hadoop 1.x Parameters*
- fs.default.name**
- Specifies the node and port on which Hadoop runs.
- For example,  
hdfs://152.144.226.224:9000

*Hadoop 2.x Parameters*

**fs.defaultFS**

Specifies the node and port on which Hadoop runs.

For example,  
hdfs://152.144.226.224:9000.

**Note:** It is recommended that the parameter name `fs.defaultFS` be used Spectrum™ Technology Platform 11 SP1 onwards.

---

**Table 1: Properties for Read from File, Write to File, Read from Hive ORC File, and Write to Hive ORC File**

| Hadoop 1.x Properties                                    | Hadoop 2.x Properties   |
|--|---|
| <b>fs.default.name</b>                                   | <b>fs.defaultFS</b>   |
| Specifies the node and port on which Hadoop runs.        | Specifies the node and port on which Hadoop runs.   |
| For example,<br><code>hdfs://152.144.226.224:9000</code> | For example,<br><code>hdfs://152.144.226.224:9000</code>  |
|  | <b>Note:</b> We recommend that the parameter name <code>fs.defaultFS</code> be used for Spectrum™ Technology Platform 11 SP1 onwards. |

13. To test the connection, click **Test**.

14. Click **Save**.


After you have defined a connection to an HDFS cluster, it becomes available in source and sink stages in Enterprise Designer, such as Read from File and Write to File. You can select the HDFS cluster when you click **Remote Machine** when defining a file in a source or sink stage.

### Compression Support for Hadoop

Spectrum™ Technology Platform supports the compression formats `gzip (.gz)` and `bzip2 (.bz2)` on Hadoop. While using the **Read from File** and **Write to File** stages with an HDFS connection, include the extension corresponding to the required compression format (`.gz` or `.bz2`) in the **File name** field. The file is decompressed or compressed based on the specified compression extension. Spectrum™ Technology Platform handles the compression and decompression of the files.

## Connecting to JDBC

In order for Spectrum™ Technology Platform to access data in a database using a JDBC connection, you must define the JDBC connection using Management Console. Once you do this, you can create flows in Enterprise Designer that can read data from, and write data to a database using JDBC.

1. Open Management Console.
2. Go to **Resources > Data Sources**.
3. Click the Add button .
4. In the **Name** field, enter a name for the connection. The name can be anything you choose.

**Note:** Once you save a connection you cannot change the name.


5. In the **Type** field, select type of database you want to connect to.

Spectrum™ Technology Platform comes with JDBC drivers for SQL Server and Oracle databases. If you want to connect to a different database type, you must add the JDBC driver before defining a connection.

6. In the **URL** field, enter the JDBC connection URL. Your database administrator can provide this URL.

For example, to connect to a MySQL database named "SampleDatabase" hosted on a server named "MyServer" you would enter:

```
jdbc:mysql://MyServer/SampleDatabase
```

7. There may be additional fields you need to fill in depending on the JDBC driver. The fields represent properties in the connection string for the JDBC driver you selected in the **Type** field. See the JDBC driver provider's documentation or your database administrator for information about the specific connection properties and values required by the connection type.
8. Click **Save**.
9. Test the connection by checking the box next to the new connection and clicking the Test button .

### Related tasks

[Importing a JDBC Driver](#) on page 95

[Manually Adding a JDBC Driver](#) on page 96

### Importing a JDBC Driver

Spectrum™ Technology Platform can access data from any database using a JDBC driver. Drivers for SQL and Oracle are provided with Spectrum™ Technology Platform, and some modules include drivers for other types of databases. If Spectrum™ Technology Platform does not come with a driver for the type of database you need, you can add a JDBC driver.

In this procedure you will import a JDBC driver by copying the driver files to the Spectrum™ Technology Platform server. When you complete this procedure the driver will be available to use when defining a JDBC connection in Management Console.

**Note:** This procedure works for JDBC 4.x drivers. If the driver you want to add uses an older version of JDBC you must add the driver manually in Management Console. For more information, see [Manually Adding a JDBC Driver](#) on page 96

1. Put all the JDBC driver files for the database into a folder named:

*Name*.jdbc

Where *Name* is any name you want. The folder name must end with .jdbc.

2. Log in to the server running Spectrum™ Technology Platform.
3. Copy the folder containing the driver to this folder:

*Spectrum Location*\server\app\drivers

The driver is automatically imported.

4. To verify that the driver was successfully imported, log in to Management Console and go to **System > Drivers**. The driver should be listed.

If the driver is not listed, open the System Log in Management Console and look for errors related to deploying JDBC drivers.

#### Related tasks

[Connecting to JDBC](#) on page 95

[Manually Adding a JDBC Driver](#) on page 96

#### Manually Adding a JDBC Driver


Spectrum™ Technology Platform can access data from any database using a JDBC driver. Drivers for SQL and Oracle are provided with Spectrum™ Technology Platform, and some modules include drivers for other types of databases. If Spectrum™ Technology Platform does not come with a driver for the type of database you need, you can add a JDBC driver.

In this procedure you will add JDBC driver files to the server then manually define the connection string and connection properties. Before you begin, be sure that you understand the connection string format and properties required by the driver. You must define these accurately in order for the driver to function. You can typically find information about a driver's connection string and properties from the driver provider's website.

**Note:** We recommend that you use this procedure only when adding a JDBC driver that uses JDBC 1.x, 2.x, or 3.x. If the driver uses JDBC 4.x, we recommend that you use the import method to add the driver. For more information, see [Importing a JDBC Driver](#) on page 95.

1. Open Management Console.



2. Go to **System > Drivers**.
3. Click the Add button .
4. In the **Name** field, enter a name for the driver. The name can be anything you choose.
5. In the **JDBC driver class name** field, enter the Java class name of the driver. You can typically find the class name in your JDBC driver's documentation.

For example, to use the Microsoft JDBC driver, you might enter the following:

```
com.microsoft.sqlserver.jdbc.SQLServerDriver
```

6. In the **Connection string template** field, enter the JDBC connection URL to use to connect to the database, including any properties you want to set in the connection string. Different database vendors use different connection strings so check your database's documentation for more information about the connection string.

If the driver will be used by more than one database connection, consider using property tokens in the connection string instead of hard-coding property values that may be different for each connection. For example, if you want to have some connections use encryption and others not, you may want to define a property token for the encryption property.

To use a property token in the connection string, use this syntax:

```
#{PropertyToken}
```

Any property tokens you include in the connection string template will be required fields when defining a database connection.

For example, this connection string for SQL contains property tokens for host, port, instance, and encryption:

```
jdbc:sqlserver://#{host}:#{port};databaseName=#{instance};encrypt=#{encryption};  
TrustServerCertificate=true
```

These tokens are required fields when defining a database connection that uses this driver:

## Add Data Source

\*Name

### Connection

\*Type


\*Host

\*Port

\*Instance

\*encryption



7. If there are properties that you want to make optional for database connections, define them in the **Connection Properties** section.
  - a) In the **Connection properties** section, click the Add button .
  - b) In the **Label** field, enter a user-friendly description of the property. The label you enter here is used as the field label in the connections window when creating a connection using this driver.
  - c) In the **Property token** field, enter the token for the optional property. See the database driver's documentation for the properties supported by the driver.


For example, if you want to make encryption optional for database connections that use this driver, you could define the encryption property like this:

Home / System: Drivers / Edit Driver

## Edit Deployed Driver

\*Name

com.mysql.jdbc.Driver.5.1


\*JDBC driver class name 

com.mysql.jdbc.Driver

\*Connection string template 

jdbc:mysql://\${host}/\${instance}

### Properties & Drivers

Connection properties 



| Label                             | Property Token |
|-----------------------------------|----------------|
| <input type="checkbox"/> username | user           |
| <input type="checkbox"/> password | password       |
| <input type="checkbox"/> Use SSL  | useSSL         |

When a database connection uses this driver, the encryption property would be displayed as an optional property in the database connection:

[Home](#) / [Data Sources](#) / [Add Data Source](#)

## Add Data Source

\*Name

### Connection

\*Type


\*Host

\*Instance

User Name

Password

Use SSL

- Log in to the server running Spectrum™ Technology Platform and place the database driver file in a folder on the server. The location does not matter.
- In the **Driver files** section, click the Add button .
- In the **File path** field, enter the path to the database driver file on the server.
- Click **Save**.


### Related tasks

[Connecting to JDBC](#) on page 95

[Importing a JDBC Driver](#) on page 95

## Editing a JDBC Driver

You can edit any JDBC driver that you have added to Spectrum™ Technology Platform. You cannot edit or delete any of the drivers that come with The Spectrum™ Technology Platform Data Integration Module, including the drivers for SQL Server and Oracle databases.

1. Open Management Console.
2. Go to **System > Drivers**.
3. Select the driver you want to modify then click the Edit button .
4. In the **JDBC driver class name** field, enter the Java class name of the driver. You can typically find the class name in your JDBC driver's documentation.

For example, to use the Microsoft JDBC driver, you might enter this:

```
com.microsoft.sqlserver.jdbc.SQLServerDriver
```

**Note:** You cannot edit the class name if you imported the driver files by copying them to the *Spectrum Location\server\app\drivers* folder.

5. In the **Connection string template** field, enter the JDBC connection URL to use to connect to the database, including any properties you want to set in the connection string. Different database vendors use different connection strings so check your database's documentation for more information about the connection string.

If the driver will be used by more than one database connection, consider using property tokens in the connection string instead of hard-coding property values that may be different for each connection. For example, if you want to have some connections use encryption and others not, you may want to define a property token for the encryption property.

To use a property token in the connection string, use this syntax:

```
${PropertyToken}
```

Any property tokens you include in the connection string template will be required fields when defining a database connection.

For example, this connection string for SQL contains property tokens for host, port, instance, and encryption:

```
jdbc:sqlserver://${host}:${port};databaseName=${instance};encrypt=${encryption};TrustServerCertificate=true
```

These tokens are required fields when defining a database connection that uses this driver:

## Add Data Source

\*Name

### Connection

\*Type


\*Host

\*Port

\*Instance

\*encryption



6. If there are properties that you want to make optional for database connections, define them in the **Connection Properties** section.
  - a) In the **Connection properties** section, click the Add button .
  - b) In the **Label** field, enter a user-friendly description of the property. The label you enter here is used as the field label in the connections window when creating a connection using this driver.
  - c) In the **Property token** field, enter the token for the optional property. See the database driver's documentation for the properties supported by the driver.


For example, if you want to make encryption optional for database connections that use this driver, you could define the encryption property like this:

Home / System: Drivers / Edit Driver

## Edit Deployed Driver

\*Name

com.mysql.jdbc.Driver.5.1


\*JDBC driver class name 

com.mysql.jdbc.Driver

\*Connection string template 

jdbc:mysql://\${host}/\${instance}

### Properties & Drivers

Connection properties 



| Label                             | Property Token |
|-----------------------------------|----------------|
| <input type="checkbox"/> username | user           |
| <input type="checkbox"/> password | password       |
| <input type="checkbox"/> Use SSL  | useSSL         |

When a database connection uses this driver, the encryption property would be displayed as an optional property in the database connection:

[Home](#) / [Data Sources](#) / [Add Data Source](#)

## Add Data Source

\*Name

### Connection

\*Type

\*Host

\*Instance

User Name

Password

Use SSL

7. Click **Save**.

### Deleting an Imported JDBC Driver

JDBC drivers cannot be deleted using Management Console if the JDBC driver was imported to Spectrum™ Technology Platform rather than being added manually in Management Console. Instead, follow this procedure to delete the driver.

**Important:** Before deleting a driver, verify that there are no database connections using the driver.

1. Stop the Spectrum™ Technology Platform server.
2. Go to this folder:



*Spectrum Location\server\app\drivers*

3. In the `drivers` folder, delete folder containing the driver.
4. Start the Spectrum™ Technology Platform server.
5. To verify that the driver has been deleted, log in to Management Console, go to **System > Drivers**, and verify that the driver is no longer listed.

## Supported Database Data Types

Spectrum™ Technology Platform supports these data types commonly used in databases:

|                   |  |
|-------------------|--|
| <b>bigdecimal</b> | A numeric data type that supports 38 decimal points of precision. Use this data type for data that will be used in mathematical calculations requiring a high degree of precision, especially those involving financial data. The bigdecimal data type supports more precise calculations than the double data type. |
| <b>boolean</b>    | A logical type with two values: true and false.  |
| <b>date</b>       | A data type that contains a month, day, and year. For example, 2012-01-30 or January 30, 2012. You can specify a default date format in Management Console.  |
| <b>datetime</b>   | A data type that contains a month, day, year, and hours, minutes, and seconds. For example, 2012/01/30 6:15:00 PM.   |
| <b>double</b>     | A numeric data type that contains both negative and positive double precision numbers between $2^{-1074}$ and $(2-2^{-52}) \times 2^{1023}$ . In E notation, the range of values is -1.79769313486232E+308 to 1.79769313486232E+308.   |
| <b>float</b>      | A numeric data type that contains both negative and positive single precision numbers between $2^{-149}$ and $(2-2^{-23}) \times 2^{127}$ . In E notation, the range of values -3.402823E+38 to 3.402823E+38.  |
| <b>integer</b>    | A numeric data type that contains both negative and positive whole numbers between $-2^{31}$ (-2,147,483,648) and $2^{31}-1$ (2,147,483,647).  |
| <b>long</b>       | A numeric data type that contains both negative and positive whole numbers between $-2^{63}$ (-9,223,372,036,854,775,808) and $2^{63}-1$ (9,223,372,036,854,775,807).  |
| <b>string</b>     | A sequence of characters.  |
| <b>time</b>       | A data type that contains the time of day. For example, 21:15:59 or 9:15:59 PM.  |
| <b>Raw</b>        | An Oracle datatype for storing variable length binary data. Maximum size is 2000 bytes (the maximum length in Oracle 7 was 255 bytes).   |

Other database data types are automatically mapped to one of the supported data types as follows:

| Database Data Type     | Supported Data Type |
|------------------------|---------------------|
| <b>Date/Time Types</b> |                     |
| TIMESTAMP              | datetime            |

| Database Data Type   | Supported Data Type |
|----------------------|---------------------|
| <b>String Types</b>  |                     |
| CHAR                 | string              |
| CLOB                 | string              |
| LONGVARCHAR          | string              |
| NCHAR                | string              |
| NVARCHAR             | string              |
| VARCHAR              | string              |
| <b>Numeric Types</b> |                     |
| BIGINT               | long                |
| DECIMAL              | double              |
| FLOAT                | double              |
| NUMERIC              | bigdecimal          |
| REAL                 | float               |
| SMALLINT             | integer             |
| TINYINT              | integer             |
| <b>Boolean Types</b> |                     |
| BIT                  | boolean             |

### Supported Database Data Types for the Location Intelligence Module

These database data types are automatically mapped to one of the supported data types for the Location Intelligence Module.

| Database Data Type | Supported Data Type |
|--------------------|---------------------|
| <b>SQL Server</b>  |                     |
| tinyint            | SHORT_INTEGER       |
| smallint           | SHORT_INTEGER       |
| int                | INTEGER             |
| bigint             | LONG_INTEGER        |
| float              | DOUBLE              |
| real               | DOUBLE              |
| decimal(10, 5)     | DOUBLE              |
| numeric(10, 5)     | DOUBLE              |
| date               | DATE                |
| time               | TIME                |
| datetime           | DATE_TIME           |
| smalldatetime      | DATE_TIME           |
| char(10)           | STRING              |
| varchar(10)        | STRING              |
| nchar(10)          | STRING              |
| nvarchar(10)G      | STRING              |
| binary(10)         | BINARY              |
| varbinary(10)      | BINARY              |


| Database Data Type    | Supported Data Type |
|-----------------------|---------------------|
| <b>PostGIS</b>        |                     |
| smallint              | SHORT_INTEGER       |
| integer               | INTEGER             |
| bigint                | LONG_INTEGER        |
| numeric(10, 5)        | DOUBLE              |
| real                  | DOUBLE              |
| double precision      | DOUBLE              |
| serial                | INTEGER             |
| bigserial             | LONG_INTEGER        |
| bytea                 | BINARY              |
| date                  | DATE                |
| time                  | TIME                |
| timestamp             | DATE_TIME           |
| character(10)         | STRING              |
| character varying(10) | STRING              |
| nchar(10)             | STRING              |
| <b>Oracle</b>         |                     |
| NUMBER                | DOUBLE              |
| CHAR(10)              | STRING              |

| Database Data Type | Supported Data Type |
|--------------------|---------------------|
| VARCHAR(10)        | STRING              |
| VARCHAR2(10)       | STRING              |
| NCHAR(10)          | STRING              |
| NVARCHAR2(10)      | STRING              |
| DATE               | DATE_TIME           |
| TIMESTAMP          | DATE_TIME           |
| BLOB               | BINARY              |
| <b>SAP HANA</b>    |                     |
| tinyint            | SHORT_INTEGER       |
| smallint           | SHORT_INTEGER       |
| integer            | INTEGER             |
| bigint             | LONG_INTEGER        |
| smalldecimal       | DOUBLE              |
| decimal(10, 5)     | DOUBLE              |
| real               | DOUBLE              |
| double             | DOUBLE              |
| float(30)          | DOUBLE              |
| varchar(30)        | STRING              |
| nchar(10)          | STRING              |

| Database Data Type | Supported Data Type |
|--------------------|---------------------|
| nvarchar(30)       | STRING              |
| alphanum(30)       | STRING              |
| date               | DATE                |
| time               | TIME                |
| seconddate         | DATE_TIM            |
| timestamp          | DATE_TIM            |
| varbinary(30)      | BINARY              |

## Connecting to Knox

An Apache Knox Gateway allows you to access a Hadoop service through the Knox security layer. In order for Spectrum™ Technology Platform to access data in Hadoop via Knox, you must define a connection to Knox using Management Console. Once you do this, you can create flows in Enterprise Designer that can read data from, and write data to, Hadoop via Knox.

1. Open Management Console.
2. Go to **Resources > Data Sources**.
3. Click the Add button .
4. In the **Name** field, enter a name for the connection. The name can be anything you choose.
 

**Note:** Once you save a connection you cannot change the name.
5. In the **Type** field, choose *Gateway*.
6. In the **Gateway Type** field, choose **Knox**.
7. In the **Host** field, enter the hostname or IP address of the node in the HDFS cluster running the gateway.
8. In the **Port** field, enter the port number for the Knox gateway.
9. In the **User Name** field, enter the user name for the Knox gateway.
10. In the **Password** field, enter the password to authorize you access to the Knox gateway.
11. In the **Gateway Name** field, enter the name of the Knox gateway you wish to access.
12. In the **Cluster Name** field, enter the name of the Hadoop cluster to be accessed.


13. In the **Protocol** field, choose `webhdfs`.
14. In the **Service Name** field, enter the name of the Hadoop service to be accessed.
15. To test the connection, click **Test**.
16. Click **Save**.

After you have defined a Knox connection to an HDFS cluster, the connection can be used in Enterprise Designer, in the stages **Read from File** and **Write to File**. You can select the HDFS cluster when you click **Remote Machine** when defining a file in a source or sink stage.

## Connecting to Marketo

In order for Spectrum™ Technology Platform to access data in Marketo, you must define a connection to Marketo using Management Console. Once you do this, you can create flows in Enterprise Designer that can read data from, and write data to, Marketo.

**Note:** This connection is for use in the Metadata Insights module.

1. Open Management Console.
2. Go to **Resources > Data Sources**.
3. Click the Add button .
4. In the **Name** field, enter a name for the connection. The name can be anything you choose.

**Note:** Once you save a connection you cannot change the name.

5. In the **Type** field, choose **Marketo**.
6. In the **Endpoint URL** field, enter the endpoint URL of your Marketo account.

To find your endpoint URL, log in to your Marketo account and go to **Admin > Integration > Web Services**. The endpoint URL is under the heading **REST API** and is in this format:

```
https://AccountID.mktorest.com/rest
```

Copy the portion of the URL before `/rest`. For example,  
`https://AccountID.mktorest.com`.

7. Enter the client ID and secret key for your Marketo account.

To find your client ID and secret key, log in to your Marketo account and go to **Admin > Integration > LaunchPoint > API Rest > View Details**. The pop-up window displays the details.

8. To test the connection, click **Test**.
9. Click **Save**.

## Marketo Limitations

1. This query applies only to `List` and `Activity_type` entities. For others, provide the filter type.

```
Select * from Marketo_Table
```


2. Marketo does not support joins except between `Lead` and `Lead_List` entities. The join query between `Lead` and `Lead_List` for a `List_Id` is as follows:

```
Select Lead.* from Lead Inner Join Lead_List
On Lead.ID = Lead_List.Lead_ID
And Lead_List.List_ID = <List ID>
```

## Connecting to Microsoft Dynamics 365 Online

In order for Spectrum™ Technology Platform to access data in Microsoft Dynamics 365 Online, you must define a connection to Microsoft Dynamics 365 Online using Management Console. Once you do this, you can create flows in Enterprise Designer that can read data from, and write data to, Microsoft Dynamics 365 Online.

**Note:** This connection is for use in the Metadata Insights module.

1. Open Management Console.
2. Go to **Resources > Data Sources**.
3. Click the Add button .
4. In the **Name** field, enter a name for the connection. The name can be anything you choose.

**Note:** Once you save a connection you cannot change the name.

5. In the **Type** field, choose **Dynamics 365 Online**.
6. In the **Username** field, enter your Microsoft Dynamics user name.
7. In the **Password** field, enter your Microsoft Dynamics password.
8. In the **Organization Name** field, enter your organization unique name, which identifies your CRM instance.

To find your organization unique name, log in to Microsoft Dynamics and go to **Settings > Customization > Customizations > Developer Resources**. Your organization unique name is displayed.


9. In the **Region** field, select the geographical region of your Microsoft Dynamics account.



10. To test the connection, click **Test**.
11. Click **Save**.

## Connecting to a Model Store

Connect to a model store to use the data federated from various sources such as databases, file servers, and cloud services. Once a connection is defined, you can use the data in the logical and physical models of a model store in the **Read from DB** and **Write to DB** stages of Enterprise Designer. For more information about model stores, see [Model Store](#).

1. Open Management Console.
2. Go to **Resources > Data Sources**.
3. Click the Add button .
4. In the **Name** field, enter a name for the connection. The name can be anything you choose.

**Note:** Once you save a connection you cannot change the name.

5. In the **Type** field, choose **Model Store**.
6. In the **Model store** field, enter the name of the model store that you are establishing connection with.

To find the names of the available model stores, open Metadata Insights, to Modeling, and click the **Model Store** tab.

7. In the **Host** field, enter the hostname or IP address of the Spectrum™ Technology Platform server.

**Note:** You can leave the **Host** field blank if the server is running on localhost.

8. In the **User Name** field, enter the user name to access the server.
9. In the **Password** field, enter the password to authorize your access to the server.
10. To test the connection, click **Test**.
11. Click **Save**.

## Connecting to NetSuite


In order for Spectrum™ Technology Platform to access data in NetSuite, you must define a connection to NetSuite using Management Console. Once you do this, you can create flows in Enterprise Designer that can read data from, and write data to, NetSuite. While reading from and writing to a NetSuite connection, both interactive and batch modes are supported.

**Note:** This connection is for use in the Metadata Insights module.

Spectrum™ Technology Platform supports these NetSuite entity types:

- Standard records
- Custom records
- Saved searches
- Joins between Standard records

To connect to NetSuite:

1. Open Management Console.
2. Go to **Resources > Data Sources**.
3. Click the Add button .
4. In the **Name** field, enter a name for the connection. The name can be anything you choose.

**Note:** Once you save a connection you cannot change the name.

5. In the **Type** field, choose **NetSuite**.
6. In the **Email** field, enter the e-mail linked to the NetSuite account to be used for the connection.
7. In the **Password** field, enter the password of the NetSuite account.
8. In the **Account** field, enter the user name for the NetSuite account.
9. In the **Role** field, select the appropriate role for this connection from the roles mapped to the particular NetSuite user account.

The **Role** field is optional. If you leave the **Role** field blank, the default role is used to log in through the connection.

**Attention:** Only standard roles are supported. Custom roles are not supported.

10. To test the connection, click **Test**.
11. Click **Save**.

**Note:** To `INSERT` a record using a NetSuite connection, use an `UPSERT` query with the primary key (`internalId`) blank.

### NetSuite Limitations

1. When querying using joins, you must cite specific columns. For example, this query is not supported:

```
select * from CUSTOMER_M
```

2. Simultaneous connections to NetSuite are not supported because NetSuite allows only a single login per account.

3. You can only write Standard and Custom records.
4. For both `UPDATE` and `UPSERT` queries, an `UPSERT` operation is performed.
5. In the Write to DB stage, the maximum batch size permitted for an `insert` operation is 200 and the maximum batch size for an `update` operation is 100.
- 6.


## Connecting to NoSQL

In order for Spectrum™ Technology Platform to access data in a NoSQL database, you must define a connection to the NoSQL database using Management Console.

The supported NoSQL database types are:

1. Couchbase
2. MongoDB

Once the desired NoSQL connection is defined, you can create flows in Enterprise Designer that can read data from and write data to this database.

1. Open Management Console.
2. Go to **Resources > Data Sources**.
3. Click the Add button .
4. In the **Name** field, enter a name for the connection. The name can be anything you choose.


**Note:** Once you save a connection you cannot change the name.

5. In the **Type** field, select any one of:
  - Couchbase
  - MongoDB
6. Specify the **Host, Port, Database, Username** and **Password** of the specific NoSQL database you wish to access.
7. Click **Test** to check that the connection to the database is successful.
8. Click **OK**.

## Connecting to Salesforce

In order for Spectrum™ Technology Platform to access data in Salesforce, you must define a connection to Salesforce using Management Console. Once you do this, you can create flows in Enterprise Designer that can read data from, and write data to, Salesforce.

**Note:** This connection is for use in the Metadata Insights module.

1. Open Management Console.
2. Go to **Resources > Data Sources**.
3. Click the Add button .
4. In the **Name** field, enter a name for the connection. The name can be anything you choose.

**Note:** Once you save a connection you cannot change the name.

5. In the **Type** field, choose **Salesforce**.
6. In the **Username** field, enter the email ID registered on the Salesforce data store.
7. In the **Password** field, enter a combination of the Salesforce portal password and the security token generated through the Salesforce portal.

For example, if your password is Sales@Test, and the security token provided to you by Salesforce is 56709367, then the Password to authenticate this Salesforce connection would be Sales@Test56709367.

8. To test the connection, click **Test**.
9. Click **Save**.

**Note:** Audit fields are enabled on all tables by default. The Salesforce audit fields are:

- created date
- last modified date
- created by
- last modified by


**Attention:** Physical model created in Spectrum™ Technology Platform version 10 and earlier using Salesforce connections need to be opened and saved again in order to enable audit fields on their tables.

## Connecting to SAP NetWeaver

Creating a SAP NetWeaver connection in the Management Console using OData Services allows you to read, write and synchronize your CRM and ERP data. While reading from and writing to a SAP connection, both interactive and batch modes are supported.

**Note:** This connection is for use in the Metadata Insights module.

1. Open Management Console.
2. Go to **Resources > Data Sources**.

3. Click the Add button .
4. In the **Name** field, enter a name for the connection. The name can be anything you choose.
 

**Note:** Once you save a connection you cannot change the name.
5. In the **Type** field, choose **SAP**.
6. In the **Username** field, enter the username to access the SAP web service.
7. In the **Password** field, enter the password of the SAP web service.
8. In the OdataURL field, enter the address of the Odata web service to be used for this connection.
9. Click **Test**.  
A message confirms the successful test of the connection.
10. Click **Save**.  
A message confirms the successful creation of the connection.


**Note:** To perform fetch operations, an OData service must support the `$skip` and `$top` operations. If the service does not support these operations, the fetched records show inconsistencies in the model store preview.

### SAP NetWeaver Limitations

For both `UPDATE` and `UPSERT` operations, an `UPDATE` operation is performed.

## Connecting to SharePoint

In order for Spectrum™ Technology Platform to access data in SharePoint, you must define a connection to SharePoint using Management Console. Once you do this, you can create flows in Enterprise Designer that can read data from, and write data to, SharePoint.

1. Open Management Console.
2. Go to **Resources > Data Sources**.
3. Click the Add button .
4. In the **Name** field, enter a name for the connection. The name can be anything you choose.
 

**Note:** Once you save a connection you cannot change the name.
5. In the **Type** field, choose **Cloud**.
6. In the **Cloud service** field, choose **Sharepoint**.
7. In the **Version** field, select **v2010**. Spectrum™ Technology Platform currently supports SharePoint version 2010.
8. In the **Protocol** field, select the protocol required to connect SharePoint.

9. In the **Server address** field, enter the host name or IP address of the SharePoint server you want to connect to.
10. Enter the user name and password to use to authenticate to SharePoint.
11. In the **Project** field, enter the specific project whose SharePoint location you want to access.
12. To test the connection, click **Test**.
13. Click **Save**.

**Example**

For example, say you want to create a connection to this SharePoint URL:

```
https://sharepoint.example.com/sites/myportal
```


You would fill in the **Protocol**, **Server address**, and **Project** fields as follows:

- **Protocol:** https
- **Server address:** sharepoint.example.com
- **Project:** myportal

## Connecting to Siebel

In order for Spectrum™ Technology Platform to access data in Siebel, you must define a connection to Siebel using Management Console. Once you do this, you can create flows in Enterprise Designer that can read data from, and write data to Siebel.

**Note:** This connection is for use in the Metadata Insights module.

1. Open Management Console.
2. Go to **Resources > Data Sources**.
3. Click the Add button .
4. In the **Name** field, enter a name for the connection. The name can be anything you choose.

**Note:** Once you save a connection you cannot change the name.

5. In the **Type** field, choose **Siebel**.
6. In the **Host** field, enter the host name of the server where Siebel is installed.
7. In the **Port** field, enter the listening port number for the SCBroker component.

**Note:** Use 2321 for SCBPort.

8. In the **Server** field, enter the name of your Siebel Enterprise Server.
9. In the **Object Manager** field, enter the name of the Object Manager.

10. Enter the user name and password to use to authenticate to Siebel.
11. In the **Locale** field, select the type of Object Manager and Language Pack for the Siebel Business application.

**Note:** For locales other than **English**, you must install locale-specific JAR files. For more information, see [Required JAR Files](#).

12. To test the connection, click **Test**.
13. Click **Save**.

Within a physical model, Siebel business components are displayed in the format `<Business Object>.<Business Component>`.

**Note:** Business Components that are not a part of any Business Object are not displayed in physical model.

### Siebel Limitations

1. Siebel Business Components fields with the same names but differing cases are handled by leaving the first occurrence unchanged, while appending all subsequent occurrences with `_PB_` and an increasing numeral value.

For example, fields like `DeDup Token`, `Dedup Token`, and `DEdup Token` in a Siebel business component are renamed to `DeDup Token`, `Dedup Token_PB_1`, and `DEdup Token_PB_2` respectively.

2. If a field name contains a dot (.) in a Siebel schema, the dots are removed from the field name for use in Spectrum™ Technology Platform.
3. While creating a physical model from a Siebel connection, links between any two Siebel Business Components are displayed only if the link name matches the Business Component name in the Siebel schema.


For example, in a Siebel schema, if a link is defined between the Business Components `Account` and `Contact`, while another Business Component `Contact Custom` uses the links of `Custom`, then in the physical model created for this Siebel connection, no link is displayed between `Account` and `Contact Custom`. Links are displayed only between `Account` and `Contact`, as the original link is between these two in the Siebel schema.

4. Fields of a Business Component which are marked as inactive in the Siebel schema are not displayed in physical model, logical model, and model store.

## Connecting to Splunk

In order for Spectrum™ Technology Platform to access data in Splunk, you must define a connection to Splunk using Management Console. Once you do this, you can create flows in Enterprise Designer that can read data from, and write data to, Splunk.

**Note:** This connection is for use in the Metadata Insights module.

1. Open Management Console.
2. Go to **Resources > Data Sources**.
3. Click the Add button .
4. In the **Name** field, enter a name for the connection. The name can be anything you choose.

**Note:** Once you save a connection you cannot change the name.

5. In the **Type** field, choose **Splunk**.
6. In the **Username** field, enter the Splunk account user name to authenticate the Splunk instance.
7. In the **Password** field, enter the password of the Splunk account.
8. In the **Host** field, enter the address or host name of the server on which the Splunk data source is hosted.
9. In the **Port** field, enter the port number of the Splunk data source.
10. To test the connection, click **Test**.
11. Click **Save**.

### Splunk Limitations

This query is not supported:

```
select count(*) from SplunkTable
```


## Connecting to SuccessFactors

In order for Spectrum™ Technology Platform to access data in SuccessFactors, you must define a connection to SuccessFactors using Management Console. Once you do this, you can create flows in Enterprise Designer that can read data from, and write data to, SuccessFactors.

**Note:** This connection is for use in the Metadata Insights module.

1. Open Management Console.



2. Go to **Resources > Data Sources**.
3. Click the Add button .
4. In the **Name** field, enter a name for the connection. The name can be anything you choose.  
**Note:** Once you save a connection you cannot change the name.
5. In the **Type** field, choose **SuccessFactors**.
6. In the **Company ID** field, enter your company ID, which identifies your company's unique instance in a specific SuccessFactors data center.
7. In the **Service URL** field, enter the URL for the SuccessFactors server you want to connect to. This URL is specific to the global data center to which your company ID is mapped.
8. Enter your user name and password for your SuccessFactors client instance.
9. To test the connection, click **Test**.
10. Click **Save**.


### SuccessFactors Limitations

1. Batch operations can only be performed using `upsert` queries. So, `insert` and `update` queries are also performed as `upsert` queries in batch operations.
2. The table/column properties, as displayed in the physical model schema of a SuccessFactors connection, might not function as expected during use of the corresponding operation. For example, a column, which has been marked updatable, may throw a system exception when you attempt to update that column.

## Connecting to SugarCRM

For Spectrum™ Technology Platform to access data in SugarCRM, you must define a connection to SugarCRM using the Management Console. Once you do this, you can create flows in Enterprise Designer that can read data from, and write data to, SugarCRM. Both on-line and on-premise versions of SugarCRM are supported.

**Note:** This connection is for use in the Metadata Insights module.

1. Open Management Console.
2. Go to **Resources > Data Sources**.
3. Click the Add button .
4. In the **Name** field, enter a name for the connection. The name can be anything you choose.  
**Note:** Once you save a connection you cannot change the name.
5. In the **Type** field, choose **SugarCRM**.

6. Enter your SugarCRM user name and password.
7. In the **URL** field, enter the URL of the SugarCRM account to be used for this connection.
8. Enter the **Client Id** and **Client Secret** of your SugarCRM account.
9. To test the connection, click **Test**.
10. Click **Save**.

### SugarCRM Limitations

1. For both `UPDATE` and `UPSERT` queries, an `UPSERT` operation is performed.
2. When writing records in batches, if a job takes more time than specified in the installer, you may face rollback exceptions. In such cases, you need to either decrease the batch size or increase the timeout limit.

**Note:** The timeout limit needs to be set as

```
com.atomikos.icatch.default_jta_timeout = 300000 in the
federation_transaction.properties file here: <Spectrum Install
Directory>\server\modules\datafederation\conf.
```

3. The **Nullable** and **Updatable** columns in the table properties, as displayed in the **physical model Schema** of the connection, may not represent the correct operation. For example, a column not marked as **Updatable** may throw system exception when you try to update it and conversely, a column marked as **Nullable** might not throw an exception in updating.
4. When querying using joins you need to use an alias.

## Connecting to a Windows Mapped Drive

When Spectrum™ Technology Platform is running on a Windows server, it can access data on the server's mapped drives. Since the Spectrum™ Technology Platform server runs as a Windows service under a particular user account (often the Local System account) you need to define the mapped drive in the server's start-up process in order for it to be visible in Enterprise Designer and Management Console.

1. Stop the Spectrum™ Technology Platform server.
2. Under the folder where the Spectrum™ Technology Platform server is installed, go to `server\bin\wrapper`. For example, `C:\Program Files\Pitney Bowes\Spectrum\server\bin\wrapper`.
3. Open the file `wrapper.conf` in a text editor.

**Important:** In the following steps you will add new properties to this file. It is important that you follow these instructions precisely and only add and modify the properties described in the following steps. Do not modify any of the other properties in this file.

4. Add these lines:

```
wrapper.share.1.location  
wrapper.share.1.target  
wrapper.share.1.type  
wrapper.share.1.account  
wrapper.share.1.password
```

5. In the `wrapper.share.1.location` property, specify the location of the mapped drive in UNC format.

**Note:** Do not include a trailing backslash in the UNC.

For example,

```
wrapper.share.1.location=\\myserver\share
```

6. In the `wrapper.share.1.target` property, specify the drive letter to assign to this mapped drive.

For example,

```
wrapper.share.1.target=Y:
```

7. In the `type` property, specify `DISK`.

For example,

```
wrapper.share.1.type=DISK
```

8. If the share you are connecting to requires a user name and password, specify the user name in the `wrapper.share.1.account` property and specify the password in the `wrapper.share.1.password` property.

For example,

```
wrapper.share.1.account=domain\user123  
wrapper.share.1.password=mypassword1
```


**Note:** If the Spectrum™ Technology Platform server service is running under the Local System user, you cannot specify a user name and password. If the share requires a user name and password you must modify the service to run under a different account.

**Example**

This example shows two mapped drives being defined in the `wrapper.conf` file.

```
wrapper.share.1.location=\\myserver\data
wrapper.share.1.target=Y:
wrapper.share.1.type=DISK
wrapper.share.1.account=sample\user
wrapper.share.1.password=samplepass
wrapper.share.2.location=\\myserver\moredata
wrapper.share.2.target=Z:
wrapper.share.2.type=DISK
wrapper.share.2.account=sample\user
wrapper.share.2.password=samplepass
```

## Deleting a Connection

1. Open the Management Console.
2. Go to **Resources > Data Sources**.
3. Check the box next to the connection you want to delete, then click the Delete button .

# 4 - Spectrum Databases

## In this section


---

|  |     |
|--|-----|
| Introduction to Spectrum Databases       | 126 |
| Adding a Spectrum Database               | 126 |
| Database Pool Size and Runtime Instances | 126 |
| Deleting a Spectrum Database             | 129 |

## Introduction to Spectrum Databases

Spectrum databases contain reference data from trusted data providers that is used to enhance and validate your data. For example, to perform address validation, Spectrum™ Technology Platform uses official address data from the postal authority to compare your address to the official address. Other types of processing that use Spectrum databases include geocoding, location intelligence functions like point-in-polygon or travel directions, and tax jurisdiction assignment for a given address.

We update Spectrum databases periodically to provide you with the most up-to-date data from third-party data providers. Database updates occur independently from software updates, in some cases quarterly or even monthly. When a database update is available you will receive an email notification that includes a link to download the updated database. You should install it as soon as possible so that you are using the most accurate data available.

Only some modules use Spectrum databases. Modules that use Spectrum databases include the Enterprise Tax Module, the Enterprise Geocoding Module, the Global Sentry Module, and the Universal Addressing Module. To view the modules you have installed that use Spectrum databases, open Management Console and go to **Resources > Spectrum Databases** then click the Add button . If you have modules that use Spectrum databases, they are listed in the **Modules** field.

## Adding a Spectrum Database

A Spectrum database contains reference data such as address data used to validate an address, or spatial data used for geocoding. To add a Spectrum database resource, see the instructions for the specific module you are working with.

## Database Pool Size and Runtime Instances

In most Spectrum™ Technology Platform environments there are multiple flows running at the same time, whether they are batch jobs or services responding to web service or API requests. To optimize concurrent processing, you can use the database pool size setting, which limits the number of concurrent requests a Spectrum database handles, and runtime instances, which controls the number of instances of a flow stage that run concurrently. These two settings should be tuned together to achieve optimal performance.

### Database Pool Size

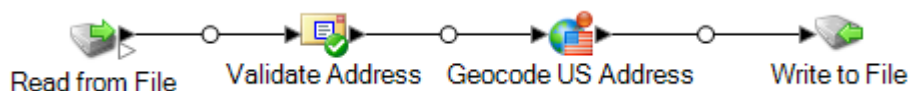
Spectrum databases contain reference data used by certain stages, such as postal data used to validate addresses, or geocoding data used to geocode addresses. These databases can be configured to accept multiple concurrent requests from the dataflow stages or services that use them, thereby improving the performance of the dataflows or service requests. The database pool size sets the maximum number of concurrent requests that a Spectrum database will process. By default, Spectrum databases have a pool size of 4, meaning the database can process four requests simultaneously.

The optimal pool size varies by module. You will generally see the best results by setting the pool size between one-half to twice the number of CPUs on the server, with the optimal pool size for most modules being the same as the number of CPUs. For example, if your server has four CPUs you may want to experiment with a pool size between 2 (one-half the number of CPUs) and 8 (twice the number of CPUs) with the optimal size possibly being 4 (the number of CPUs).

When modifying the pool size you must also consider the number of runtime instances specified in the dataflow for the stages accessing the database. Consider for example a dataflow that has a Geocode US Address stage that is configured to use one runtime instance. If you set the pool size for the US geocoding database to four, you will not see a performance improvement because there would be only one runtime instance and therefore there would only be one request at a time to the database. However, if you were to increase the number of runtime instances of Geocode US Address to four, you might then see an improvement in performance since there would be four instances of Geocode US Address accessing the database simultaneously, therefore using the full pool.

### Runtime Instances

Each stage in a dataflow operates asynchronously in its own thread and is independent of any other stage. This provides for parallel processing of stages in a dataflow, allowing you to utilize more than one runtime instance for a stage. This is useful in dataflows where some stages process data faster than others. This can lead to an unbalanced distribution of work among the threads. For example, consider a dataflow consisting of the following stages:



Depending on the configuration of the stages, it may be that the Validate Address stage processes records faster than the Geocode US Address stage. If this is the case, at some point during the execution of the dataflow all the records will have been processed by Validate Address, but Geocode US Address will still have records to process. In order to improve performance of this dataflow, it is necessary to improve the performance of the slowest stage - in this case Geocode US Address. One way to do that is to specify multiple runtime instances of the stage. Setting the number of runtime instances to two, for example, means that there will be two instances of that stage, each running in its own thread, available to process records.

As a general rule, the number of runtime instances should be at least equal to the number of instances of the remote component. See the *Administration Guide* for information about remote components.


While specifying multiple runtime instances can help improve performance, setting this value too high can strain your system resources, resulting in decreased performance.

**Note:** Using multiple runtime instances only improves performance when running jobs or when running service requests with more than one record.

### *Tuning Procedure*

Finding the right settings for database pool size and runtime instances is a matter of experimenting with different settings to find the ones maximize available server resources without overloading resources and causing reduced performance.

**Note:** You should optimize the dataflow pool size before tuning the database pool size. For information about optimizing the dataflow pool size, see [Dataflow Pool Size](#) on page 193.

1. Begin by finding sample data to use as you test different settings. The sample dataset should be large enough that execution time is measurable and can be validated for consistency. The sample data should also be representative of the actual data you want to process. For example, if you are doing performance testing for geocoding, be sure that your test data has an equal number of records for all the countries you intend to geocode.
2. If you are testing a service or dataflow that requires the use of a database resource, such as postal databases or geocoding databases, make sure that you have the latest version of the database installed.
3. With sample data ready and the latest database resources installed, create a simple dataflow that reads data from a file, processes it with the stage you want to optimize, and writes to a file. For example, if you want to test performance settings for Validate Address, create a dataflow consisting of Read from File, Validate Address, and Write to File.
4. Set the database resource pool size to 1:
  - a. Open Management Console.
  - b. Go to **Resources > Spectrum Databases**.
  - c. Select the database resource you want to optimize and click the Modify button .
  - d. In the **Pool size** field, specify 1.
  - e. Click **OK**.
5. Set the stage's runtime instances to 1:
  - a. Open the dataflow in Enterprise Designer.
  - b. Double-click the stage that you want to set to use multiple runtime instances.
  - c. Click **Runtime**.
 

**Note:** Not all stages are capable of using multiple runtime instances. If there is no **Runtime** button at the bottom of the stage's window, the stage is not capable of using multiple runtime instances.
  - d. Select **Local** and specify 1.
  - e. Click **OK** to close the **Runtime Performance** window, then click **OK** to close the stage.



6. Calculate baseline performance by running the dataflow several times and recording the average values for:

- Elapsed time
- CPU utilization
- Memory utilization

**Tip:** You can use the JMX console to monitor performance. For more information, see [Monitoring Performance with the JMX Console](#) on page 212.

7. Run multiple instances of the job concurrently, if this is a use case that must be supported. Record elapsed time, CPU utilization, and memory utilization for each scenario.


**Tip:** You can use a file monitor to run multiple instances of a job at once. For more information, see [Triggering a Flow with a Control File](#) on page 168.

8. Increment the database resource pool size and the stage runtime instances setting.
9. Restart the server.
10. Run the dataflow again, recording the elapsed time, CPU utilization, and memory utilization.
11. Continue to increment the database resource pool size and the stage runtime instances until you begin to see diminishing performance.
12. If you are testing geocoding performance, repeat this procedure using single country and multi-country input.

## Deleting a Spectrum Database

A Spectrum database contains reference data such as address data used to validate an address, or spatial data used for geocoding. You can delete Spectrum databases that are no longer used on your system. For example, you may want to delete a Spectrum database after installing an updated version of the database.

**Important:** Before deleting any resource, verify that there are no jobs or services using it. Deleting a resource that is referenced by jobs or services will cause those jobs or services to fail.

1. Open the Management Console.
2. Go to **Resources > Spectrum Databases**.
3. Put a check mark next to the Spectrum database you want to delete then click the Delete button .

Deleting a Spectrum database resource does not delete the database files themselves. After deleting the resource you must delete the database files if you want to free up space on your system.

# 5 - Services

## In this section

---

|                       |     |
|-----------------------|-----|
| Spectrum Services     | 131 |
| External Web Services | 139 |


# Spectrum Services

A service is a processing capability that you access through the REST or SOAP web service interface or through the Spectrum™ Technology Platform API. You pass one or more records to the service and optionally specify the options to use when processing the record. The service processes the data and returns the data.

Some services become available when you install a module. For example, when you install the Universal Addressing Module the service ValidateAddress becomes available on your system. In other cases, you must create a service in Enterprise Designer then expose that service on your system as a user-defined service. For example, the Location Intelligence Module's stages are not available as services unless you first create a service using the module's stages.

## Specifying Default Service Options

Default service options control the default behavior of each service on your system. You can specify a default value for each option in a service. The default option setting takes effect when an API call or web service request does not explicitly define a value for a given option. Default service options are also the settings used by default by services in a dataflow.

1. Open Management Console.
2. Click **Services**.
3. Check the box next to the service you want then click the Edit button .
4. Set the options for the service. For information about the service's options, see the guide for the service's module.
5. Click **Save**.

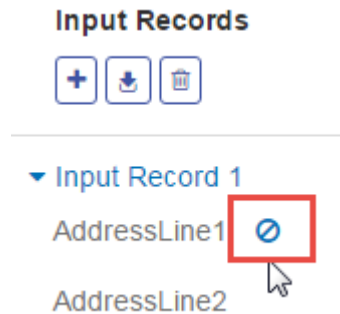
## Previewing a Service



You can preview the results of a service in Management Console using the service's Preview tab. Preview can be useful in helping you decide what options to specify because you can immediately see the effect that different options have on the data returned by the service.

1. Open Management Console.
2. Go to the **Services** menu and select the service you want to preview.
3. Click the **Preview** tab.
4. Enter the test data into each field.

Here are some tips for using preview:

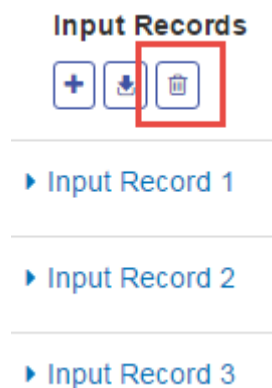
- You do not have to enter data in every field. Leaving a field empty results in an empty string being used for preview.
- If you want to preview the effect of passing a null value in a field, click the Disable icon next to the field:



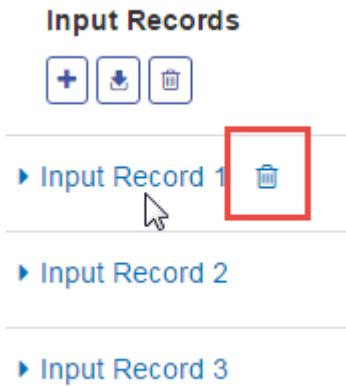
- You can preview multiple records at once. To add a record, click the Add button .
- You can import test data from a file. To import data, click the Import button . Note the following:
  - The first row in the file must be a header record. The field names in the header must match the field names required by the service.
  - If the file uses a space as the field separator, field values must be surrounded by quotes. Here is an example of a file that uses a space as the field separator:

```
AddressLine1 AddressLine2 City StateProvince PostalCode
"One Global View" "" "Troy" "NY" "12180"
"3001 Summer St" "" "Stamford" "CT" "06926"
"224 N Michigan Ave" "Suite 300" "Chicago" "IL" ""
```

- To delete all records, click the Delete button at the top of the preview area:



- To delete an individual record, hover over the input record name (for example, "Input Record 1") and click the Delete button next to the record name:



- If the service takes hierarchical input data:
  - To add child records, hover over the parent record and click the Add button.
  - To delete all children of a parent, hover over the parent record and click the Delete button.
  - To delete individual child records, hover over the child record and click the Delete button.

5. Click **Run Preview**.

The service processes the input records and displays the results:

**Input Records**

+ ↓ 🗑️

▼ Input Record 1

|               |           |
|---------------|-----------|
| AddressLine1  | 33 Monroe |
| AddressLine2  | Suite 20  |
| AddressLine3  |           |
| AddressLine4  |           |
| AddressLine5  |           |
| City          | Chicago   |
| StateProvince |           |
| PostalCode    |           |
| Country       |           |
| FirmName      |           |
| USUrbanName   |           |
| CanLanguage   |           |

**Run Preview**

**Output Records**

▼ Output Record 1

|                                |                          |
|--------------------------------|--------------------------|
| Confidence                     | 80                       |
| RecordType                     | HighRise                 |
| RecordType.Default             | Y                        |
| CountryLevel                   | A                        |
| ProcessedBy                    | USA                      |
| MatchScore                     | 0                        |
| AddressLine1                   | 33 W Monroe St Ste 20    |
| City                           | Chicago                  |
| StateProvince                  | IL                       |
| PostalCode                     | 60603-5300               |
| PostalCode.Base                | 60603                    |
| PostalCode.AddOn               | 5300                     |
| Country                        | United States Of America |
| AdditionalInputData.Base       |                          |
| AdditionalInputData.Unattached |                          |
| POBoxOnlyDeliveryZone          |                          |

6. Review your output data, making sure the results are what you intended to get from the service. If necessary you can make changes to the service's settings and click **Run Preview** again. (You do not need to input the data again.)

## Optimizing Services

There are a number of different ways to call a Spectrum™ Technology Platform service and some of them perform better than others. The different ways of calling Spectrum™ Technology Platform services, in approximate order from fastest to slowest are:

- Client API over SOCKET
- Client API over HTTP
- Client API over HTTPS
- XML over HTTP
- Web Services - SOAP and REST over HTTP

Invoking a service through the client API is generally faster than calling the web service. The network protocol can have a significant effect on the round-trip time for the service call. For example, using the persistent SOCKET connection instead of HTTP can improve response time by 30%-50%.

Performance in a real-time application calling Spectrum™ Technology Platform services also depends on whether the application is single-threaded or multi-threaded, and whether the server has the resources available to fulfill the service request. For a single-threaded client application, specifying additional instances of the remote component and/or additional runtime instances of a stage will have minimal impact on response time. A multi-threaded client application will generally benefit from multiple remote component instances and runtime instances, up to the number of concurrent threads.

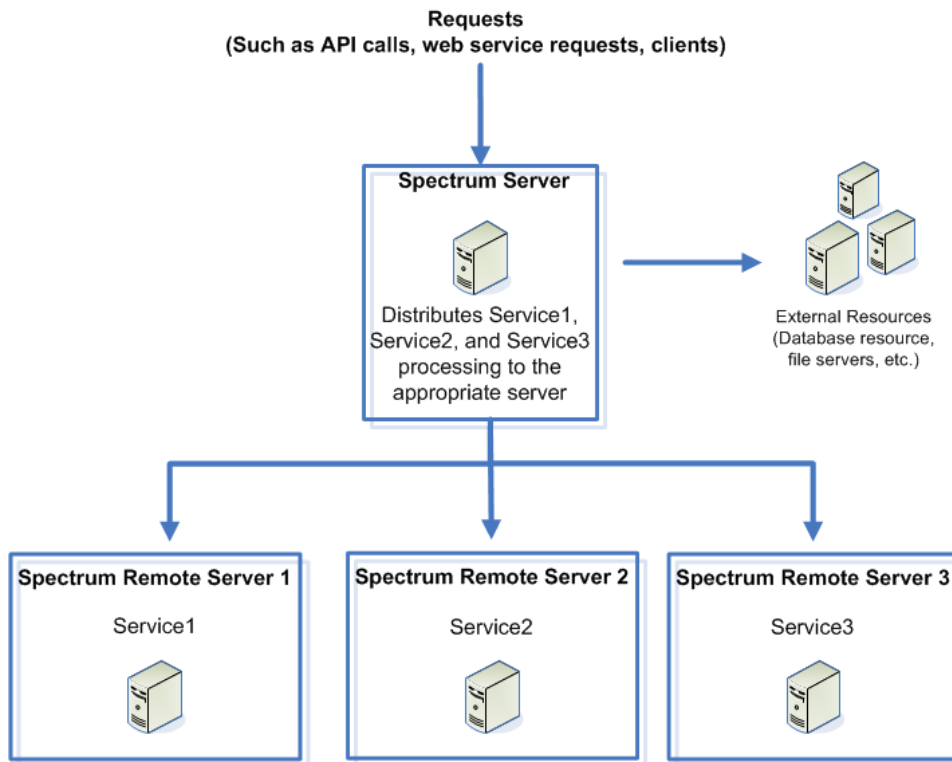
## Remote Servers

A remote server is a separate Spectrum™ Technology Platform server that handles processing for a specific service or services. When a processing request comes to your main Spectrum™ Technology Platform server, either through a service request or through the execution of a flow, the processing for the service is routed to the remote server.

Remote servers are useful if you have licensed many modules. To use remote servers, you install Spectrum™ Technology Platform on several different servers, installing different modules on each server. You then configure the individual services to execute on one of the remote servers. This approach has the following advantages:

- Improved performance, especially for web service and API calls. For jobs, there may be some performance improvement, but the time it takes to send data across the network to the remote server may offset any performance gains.

- Ability to perform database updates on individual modules without interrupting the availability of other modules. For example, if you need to update a postal database for the Universal Addressing Module, you could install the update by stopping just the remote server handling the Universal Addressing Module, allowing other modules to remain available on other remote servers.
- Startup time can be reduced. With all modules on one server, it can take a long time for the server to start up. With modules installed on separate servers, each server will start up more quickly.



To determine if a service is eligible for routing to a remote server, open the Management Console, click that service, and see if the **Routing** button at the bottom of the **Options** tab is enabled.


To view a list of remote servers, open Management Console and go to **Resources > Remote Servers**.

Tools deployed with a particular module, such as database resource tools, are available only on servers where that module is installed. For example, if the Enterprise Geocoding Module is installed only on a remote server, the tools, such as the database resource tools, will only be visible in Management Console only if you log in to the remote server. If you log in to Management Console on the local server, the Enterprise Geocoding tools would not be available.

### Adding a Remote Server

Follow the steps below to add a remote server.

1. Install Spectrum™ Technology Platform on another server. This will be the remote server. For installation instructions, see the *Spectrum™ Technology Platform Installation Guide*.
2. Open Management Console on the main server (not the new remote server you just installed).

3. Go to **Resources > Remote Servers**.
4. Click the Add button .
5. In the **Name** field, enter the name of the remote server.
6. In the **Host** field, enter the host name or IP address of the remote server.
7. In the **Port** field, enter the HTTP port used by the remote Spectrum™ Technology Platform server. The default port is 8080.
8. In the **User name** and **Password** fields, enter the credentials to use to run services on the remote server.
9. Check the **Use secure connection** box if the remote server you are connecting to is configured to use HTTPS.
10. If you wish to test the connection, click **Test**.

If the remote server test takes a long time (which implies that the test will fail), you can cancel the test by clicking **Stop**.

11. In the **Microbatch size** field, enter the number of records passed to the remote server in a single batch. The default is 50. Entering a higher number in this field will speed up input and output but slow down data transmission.

To determine the optimal microbatch size for your environment, create a test dataflow that uses the remote server in one of its stages. Run the dataflow with a microbatch size of 50 and record the time it takes to run the dataflow. Then run a test with the microbatch size at 60, and another one with a microbatch size of 70 and observe the results. You may also want to run tests with microbatch sizes of 40 and 30 to see if reducing the microbatch size improves execution time. Continue testing with different settings until you identify the best setting for your environment.

12. In the **Timeout** field, enter the number of seconds to allow the system to wait for a response from the remote server before the connection times out. The default is 2 seconds.
13. Click **Save**.

#### Related concepts

[Remote Servers](#) on page 134

#### Related tasks

[Routing a Service to a Remote Server](#) on page 136

#### Related reference

[Troubleshooting Remote Server Errors](#) on page 138

### Routing a Service to a Remote Server

When you route a service to a remote server, processing requests for the service are forwarded to another Spectrum™ Technology Platform server (the remote server).

**Note:** You can override a service's remote server setting in Enterprise Designer or the Spectrum™ Technology Platform API.

To route a service to a remote server:



1. Open the Management Console.
2. If you have not already done so, add the remote server in Management Console.
3. Select the service you want to route to a remote server.
4. In the **Routing** field, select the remote server to which you want route service requests for this service. If no **Routing** field is visible, there are no remote servers configured for your system. In order to route a service to a remote server, you must first configure a remote server in Management Console.
5. Click **Save**.

When you add a remote server that contains services that are not also installed on the local server, it brings with it the default options from that remote server.

When you make changes to services on a remote server, those changes will not be reflected on the local server. The local Management Console will reflect its own default options.

#### Related concepts

[Remote Servers](#) on page 134


#### Related tasks

[Adding a Remote Server](#) on page 135


#### Related reference

[Troubleshooting Remote Server Errors](#) on page 138

### Modifying a Remote Server


If you make changes on a remote server, such as adding or removing modules or creating new services, you must refresh the remote server information on the local server. To do this, open Management Console, go to **Resources > Remote Servers**, and click the Refresh button . The information for all your remote servers is updated with the latest configuration information from each remote server. The refresh process may take a few minutes.

If you want to modify the connection properties of a remote server, such as the user name and password used to connect to the remote server, follow these steps:

1. Open Management Console.
2. Go to **Resources > Remote Servers**.
3. Check the box next to the remote server you want to modify then click the Modify button .
4. Make the changes you want and click **Save**.

### Deleting a Remote Server

If you are no longer using a remote server, delete the remote server connection in Management Console so that users cannot route services and dataflow stages to it.

1. Open Management Console.
2. Go to **Resources > Remote Servers**.
3. Check the box next to the remote server you want to delete then click the Delete button .

If you delete a remote server that has services routed to it, those services will be automatically rerouted as follows:

- If the service is installed on the local server, services will be rerouted to the local server.
- If the service is not installed on the local server, services will be rerouted to a remote server that has the service installed.

If the service is not installed on either the local server or another remote server, you cannot delete the remote server until you either install the service locally or define another remote server where the service is installed.

## Troubleshooting Remote Server Errors

This section discusses possible errors you may experience when using remote servers.

### *Module Not Licensed*

The remote server must have the license for both the module and the mode of execution you are trying to run, either batch or real-time. The license on the remote server may be different from the license on the local server. Log in to the remote server using Management Console and verify that the correct license is installed. You must log in with an account that has administrative privileges in order to view license information.

### *Remote Server Not Available*

If the remote server is not running or is not reachable for any other reason, the remote services will become unavailable in Enterprise Designer and Management Console. You will see a yellow hazard icon in the status bar at the bottom of the screen:



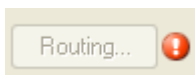
Click this icon to see an error message that describes which remote servers are not available.

In addition, in Enterprise Designer any stages that use a remote stage will be replaced with an icon showing you the stage is no longer available:



### *Routing Has Changed*

If you delete or undeploy a service that is installed both locally and remotely and has been routed through a remote server, and then click that service within Management Console, you will see a routing change indicator (a blinking exclamation point) next to the routing button on the Options tab for that service. This indicator means the routing has changed for that service.



**Related concepts**

[Remote Servers](#) on page 134

**Related tasks**

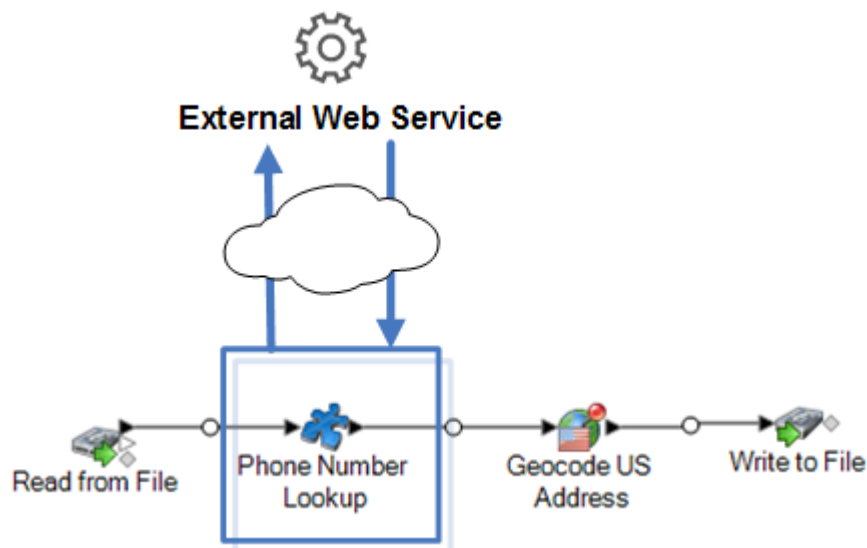
[Adding a Remote Server](#) on page 135

[Routing a Service to a Remote Server](#) on page 136

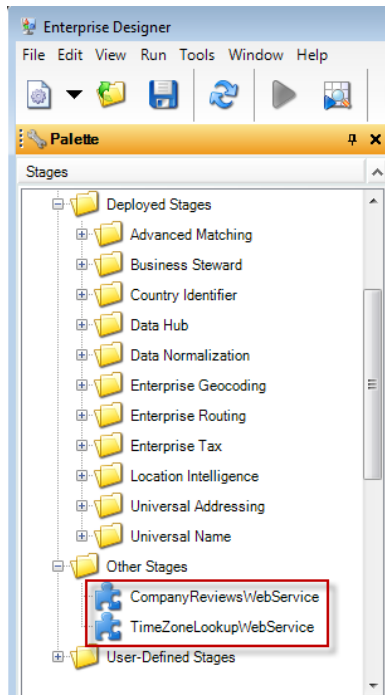
## External Web Services

External web services are data processing services provided over the Internet by a third party. You can define external web services in Management Console then use them as a stage in a dataflow. This allows you to incorporate almost any kind of processing into your Spectrum™ Technology Platform environment, since there are a wide variety of web services available on the Internet.

The following diagram illustrates the concept of external web services. Here, an external web service named Phone Number Lookup has been added to the dataflow. When the dataflow runs, Spectrum™ Technology Platform sends each record to the external web service. The external web service processes the record and returns the result to the stage. The updated record, with the phone number added, continues to the next stage in the dataflow, in this example Geocode US Address.



External web services show up in the palette in Enterprise Designer and you can work with them as you do other stages. The following shows an example of two external web services, `CompanyReviewsWebService` and `TimeZoneLookupWebService`.



### Requirements and Limitations


Spectrum™ Technology Platform supports external web services that use REST, SOAP 1.1, or SOAP 1.2 messaging, with these limitations:

- WADL requests and responses with more than one representation are not supported.
- Recursive schemas are not supported.

## Adding an External Web Service

External web services are data processing services provided over the Internet by a third party. You can use an external web service as a stage in a flow, allowing you to expand the capabilities of your Spectrum™ Technology Platform server.

This procedure defines a connection between your Spectrum™ Technology Platform server and a third-party web service. After completing this procedure, you will have a new stage in Enterprise Designer which represents the external web service. You can then use the external web service as you would any other stage in a flow.

1. Open the Management Console.
2. Go to **Resources > External Web Services**.
3. Click the Add button .
4. On the **Descriptor** step:
  - a) Specify the web service's WSDL or WADL by one of these methods:

|   |   |
|---|---|
| To load the web service's descriptor from a URL | In the <b>Load descriptor</b> field, choose <b>By URL</b> and in the <b>URL</b> field specify the URL of the WADL or WSDL. Enter your credentials for the web service if required by the web service. |
|---|---|

|  |  |
|--|--|
| To load the web service's descriptor from a file | In the <b>Load descriptor</b> field, choose <b>Upload</b> then select the file in the <b>Upload file</b> field. Some web service vendors prefer to provide the WSDL or WADL as a file rather than making it available via a URL. |
|--|--|

|   |   |
|---|---|
| For REST web services that have no WADL | In the <b>Load descriptor</b> field, choose <b>None</b> . |
|---|---|

- b) Click **Next**.
5. On the **Settings** step:
- a) In the **Name** field, enter the name you want to give the external web service when it is exposed on Spectrum™ Technology Platform. This will be the stage name shown in Enterprise Designer. The name can be anything you want but must not already be used by another web service on your system.
  - b) In the **Timeout** field, enter the number of seconds that are allowed to elapse before a request submitted to the web service times out.

**Note:** The timeout value you specify here applies to any request made to the web service. This includes not only transactions made against the exposed web service, but also requests made while configuring the web service. Such configuration requests are invoked by choosing a new item on the **Request** page and running preview. Timeouts can occur when performing any of these actions. If timeouts do occur, increasing the **Timeout** value can help to avoid them, provided the web service is actually up and running.

- c) If the external web service requires a user name and password, under **Security settings**, in the **Type** field, choose how to transmit the user name and password from the Spectrum™ Technology Platform server to the external web service.

|                             |   |
|-----------------------------|---|
| <b>None</b>                 | Choose this option if you do not need to provide a user name and password to use the external web service.  |
| <b>Basic Authentication</b> | Choose this option to pass user name and password information to the external web service in the HTTP header.   |
| <b>WS-Security</b>          | For SOAP services only, choose this option to pass user name and password information to the external web service through the header of the SOAP message. |

- d) Enter a user name and password if they are required to access the external web service.
- e) Click **Next**.

6. On the **Request** step, configure the parameters to include in requests to the external web service.

**For REST web services:**

**URL** If you did not supply a WADL on the **Descriptor** step, enter a sample request URL containing the path parameters (if any) and query parameters that you want to include in the request. For example:

```
http://example.com/rest/customers/{state}?age=31
```

Here, there is one path parameter, `{state}`, and one query parameter, `age`.

If you supplied a WADL on the **Descriptor** step, the **URL** field displays the endpoint based on the WADL you provided. You cannot edit the endpoint.

**Resource** This setting is only visible if you supplied a WADL on the **Descriptor** step. Select the web service resource that you want to expose on Spectrum™ Technology Platform.

**Note:** If you want to expose more than one resource, you must define a separate external web service for each resource.

**Method** Select the HTTP method to use to for the request to the external web service.

If you provided a WADL on the **Descriptor** step, only those HTTP methods supported by the external web service are listed.

**Path parameters** The **Path parameters** section lists the parameters that are part of the URL path, if the external web service uses path parameters. For example, this URL contains the path parameter `{state}`:

```
http://example.com/rest/customers/{state}?age=31
```


If you provided a WADL on the **Descriptor** step, the web service's path parameters are listed. If you did not supply a WADL on the **Descriptor** step, the list of path parameters is generated from the sample request URL you provided in the **URL** field. To add or remove a path parameter, add or remove it from the URL. Anything in the URL surrounded by curly braces is interpreted as a path parameter.

**Query parameters** The **Query parameters** section lists the parameters that appear after the "?" in the request URL. For example, this URL contains the query parameter `age`:

```
http://example.com/rest/customers/{state}?age=31
```

If you provided a WADL on the **Descriptor** step, the web service's query parameters are listed. If you did not provide a WADL on the **Descriptor** step, the list of query parameters is generated from the sample request URL you provided in the **URL** field. To add or remove a query parameter, add or remove it from the URL.

**Table 2: Settings for REST Path Parameters and Query Parameters**

| Setting              | Description  |
|----------------------|--|
| <b>Expose</b>        | Check the box in this column to make the parameter available in the Spectrum™ Technology Platform stage.   |
| <b>Request</b>       | This column lists the parameter name that is used in the request to the external web service.  |
| <b>Input</b>         | This column lists the input field names as they will appear in a flow. You can leave the names the same as those used by the third-party web service or you can change them by hovering over the name and clicking the edit button  .   |
| <b>Default Value</b> | <p>Check the box in this column if you want to specify a default value for the field. Enter the default value in to the field that appears after checking the box.</p> <p>If you want to provide a default value that cannot be overridden from a flow, check the box in the <b>Default Value</b> column and clear the corresponding box in the <b>Request</b> column. This hides the field from flows while providing the default value in the request to the external web service. This may be useful if you have an access key that you need to provide in each request. For example:</p> |


|   |               |
|---|---------------|
| <input type="checkbox"/> AccessKey                |               |
| <input checked="" type="checkbox"/> FirmName      | FirmName      |
| <input checked="" type="checkbox"/> AddressLine1  | AddressLine1  |
| <input checked="" type="checkbox"/> LastLine      | LastLine      |
| <input checked="" type="checkbox"/> City          | City          |
| <input checked="" type="checkbox"/> StateProvince | StateProvince |
| <input checked="" type="checkbox"/> PostalCode    | PostalCode    |
| <input checked="" type="checkbox"/> HouseNumber   | HouseNumber   |

**For SOAP web services:**

**URL** This field displays the endpoint based on the WSDL you provided on the **Descriptor** step. You cannot edit this field.

**Operation** Select the web service operation you want to perform.

**Note:** If you want to expose more than one operation, you must define a separate external web service for each operation.

- Request** In this column, select the fields and options you want to make available through Spectrum™ Technology Platform.
- Input** This column lists the input field names as they will appear in a flow. You can leave the names the same as those used by the third-party web service or you can change them by hovering over the name and clicking the edit button .
- Default Value** Check the box in this column if you want to specify a default value for the field. Enter the default value in to the field that appears after checking the box.

If you want to provide a default value that cannot be overridden from a flow, check the box in the **Default Value** column and clear the corresponding box in the **Request** column. This hides the field from flows while providing the default value in the request to the external web service. This may be useful if you have an access key that you need to provide in each request. For example:

|   |               |  |
|---|---------------|--|
| <input type="checkbox"/> AccessKey                |               | <input checked="" type="checkbox"/> 1234567890 |
| <input checked="" type="checkbox"/> FirmName      | FirmName      | <input type="checkbox"/>                       |
| <input checked="" type="checkbox"/> AddressLine1  | AddressLine1  | <input type="checkbox"/>                       |
| <input checked="" type="checkbox"/> LastLine      | LastLine      | <input type="checkbox"/>                       |
| <input checked="" type="checkbox"/> City          | City          | <input type="checkbox"/>                       |
| <input checked="" type="checkbox"/> StateProvince | StateProvince | <input type="checkbox"/>                       |
| <input checked="" type="checkbox"/> PostalCode    | PostalCode    | <input type="checkbox"/>                       |
| <input checked="" type="checkbox"/> HouseNumber   | HouseNumber   | <input type="checkbox"/>                       |


7. If you selected POST or PUT in the **Method** field, define the structure of the data you want to send to the web service in the POST or PUT operation. To do this, click the **Format** button and choose one of the following options:

- Upload schema** Choose this option if you have an XML schema that defines the structure of the data you to send to the web service in the POST or PUT operation. After selecting this option, browse to the schema file.
- Provide sample** Choose this option if you have a sample of the data you want to send to the web service in the POST or PUT operation. After selecting this option, you can enter the sample manually or paste the sample into the window.

After you provide a schema or sample, check the box next to each data element you want to make available in the flow stage.

- 8. Click **Next**.
- 9. On the **Headers** step:
  - a) Under **HTTP headers**, specify the value to pass to the external web service for each header. The values you specify here are used for all requests from Spectrum™ Technology Platform to the external web service. If no headers are listed, the external web service does not require any HTTP headers.
  - b) For SOAP web services, if the external web services supports SOAP headers you can select the headers you want to use under **SOAP headers**. You can specify a default value



for each SOAP header. The default value can be overridden in each request to Spectrum™ Technology Platform. The **Input** column shows the name of the header that will be used in requests to Spectrum™ Technology Platform. If you want to use a different name, hover over the name and click the edit button .

If a check box is checked and grayed out, it means that the header is required and you cannot disable it.

c) Click **Next**.


10. On the **Response** step:

- a) If you want the response from the external web service to be returned in a single field, check the box **Return payload as field**. All response elements will be placed in a single field instead of being returned in separate fields. For REST web services, the field name is RestReponse, and for SOAP web services the field name is SoapResponse.
- b) If you are configuring a REST web service, there is a **Format** button. Click this button to choose how you want to define the structure of the web service response that will be returned by Spectrum™ Technology Platform:

**Upload schema** Choose this option if you have an XML schema that defines the structure of the response you want Spectrum™ Technology Platform to return. After selecting this option, browse to the schema file.


**Provide sample** Choose this option if you have a sample response from the external web service and you want to use it to define the response you want from Spectrum™ Technology Platform. After selecting this option, you can enter the sample manually or paste the sample into the window.

- c) In the **Response** column, choose the fields you want to make available in Spectrum™ Technology Platform.

The icon  indicates that a field will be returned as a *list* field. A list is a Spectrum™ Technology Platform data type that contains hierarchical data that can be repeated. For example, a field PhoneNumbers may contain multiple Phone fields:

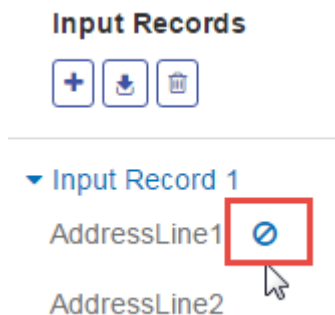
```
<PhoneNumbers>
  <Phone>
    <Type>Cell</Type>
    <Number>312-123-4567</Number>
  </Phone>
  <Phone>
    <Type>Home</Type>
    <Number>773-123-4567</Number>
  </Phone>
</PhoneNumbers>
```

In this case the PhoneNumbers field would be a list field that can contain a list of Phone elements.



- d) The **Output** column lists the output field names as they will appear in a flow. You can leave the names the same as those used by the third-party web service or you can change them by hovering over the name and clicking the edit button .
- e) Click **Next**.
11. On the **Preview** step, you can test the external web service by entering sample data then clicking **Run Preview**. This is optional.

Here are some tips for entering sample data:

- You do not have to enter data in every field. Leaving a field empty results in an empty string being used for preview.
- If you want to preview the effect of passing the default value or a null value in a field, click the Disable icon next to the field:

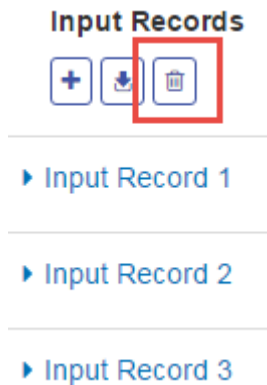


If you have defined a default value for the field on the **Request** tab, the default value will be used. If you have not defined a default value, a null value will be used.

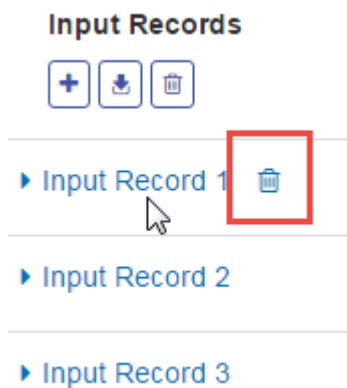
- You can preview multiple records at once. To add a record, click the Add button .
- You can import test data from a file. To import data, click the Import button . Note the following:
  - The first row in the file must be a header record. The field names in the header must match the field names required by the service.
  - The maximum number of records that can be imported is five.
  - If the file uses a space as the field separator, field values must be surrounded by quotes. Here is an example of a file that uses a space as the field separator:

```
AddressLine1 AddressLine2 City StateProvince PostalCode
"One Global View" "" "Troy" "NY" "12180"
"3001 Summer St" "" "Stamford" "CT" "06926"
"224 N Michigan Ave" "Suite 300" "Chicago" "IL" ""
```

- To delete all records, click the Delete button at the top of the preview area:



- To delete an individual record, hover over the input record name (for example, "Input Record 1") and click the Delete button next to the record name:



12. To make the external web service available to use in flows, switch the **Enabled** switch to **On**.
13. Click **Save**.
14. If you want to make the web service available to use in dataflows and to access as a Spectrum™ Technology Platform service:
  - a) Check the **Expose as service** box.
  - b) If you want the web service exposed as a SOAP web service, check the **SOAP** box. The web service will then be listed at `http://<ServerName>:<Port>/soap`.
  - c) If you want the web service exposed as a REST web service, check the **REST** box. The web service will then be listed at `http://<ServerName>:<Port>/rest`.

The external web service is now defined and available to use as a stage in a flow in Enterprise Designer.

### Related concepts

[External Web Services](#) on page 139

## Previewing an External Web Service

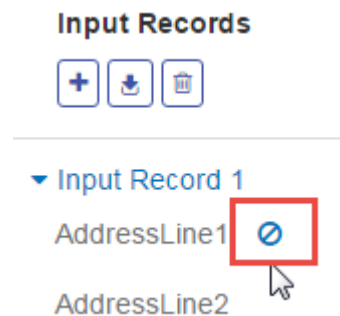
You can preview the response from an external web service by sending a test request. This can be done while adding a new external web service, in the last step of the "Add Web Service" wizard. You can also preview the response from external web services that have already been added to your Spectrum™ Technology Platform server. This topic describes how to preview an external web service that has already been added.

**Note:** In order to be able to preview an external web service, you must have View and Execute privileges for **Platform - Services** in addition to View and Modify privileges for **External Web Services - Connection**.

1. In Management Console, go to **Services > External Web Services**.
2. Click the external web service you want to preview.
3. Enter the sample data you want to use in the test request to the external web service.

Here are some tips for entering sample data:

- You do not have to enter data in every field. Leaving a field empty results in an empty string being used for preview.
- If you want to preview the effect of passing the default value or a null value in a field, click the Disable icon next to the field:



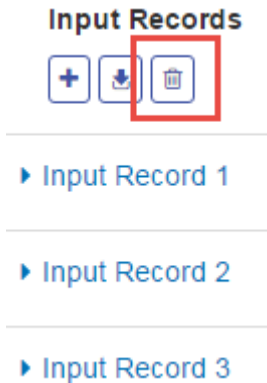
If you have defined a default value for the field on the **Request** tab, the default value will be used. If you have not defined a default value, a null value will be used.

- You can preview multiple records at once. To add a record, click the Add button
- You can import test data from a file. To import data, click the Import button . Note the following:
  - The first row in the file must be a header record. The field names in the header must match the field names required by the service.
  - The maximum number of records that can be imported is five.

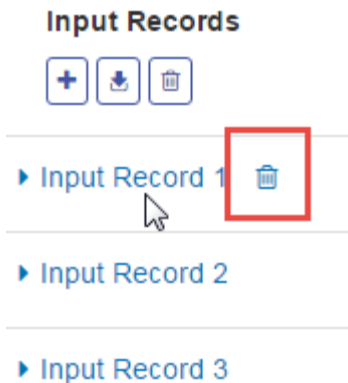
- If the file uses a space as the field separator, field values must be surrounded by quotes. Here is an example of a file that uses a space as the field separator:

```
AddressLine1 AddressLine2 City StateProvince PostalCode
"One Global View" "" "Troy" "NY" "12180"
"3001 Summer St" "" "Stamford" "CT" "06926"
"224 N Michigan Ave" "Suite 300" "Chicago" "IL" ""
```

- To delete all records, click the Delete button at the top of the preview area:



- To delete an individual record, hover over the input record name (for example, "Input Record 1") and click the Delete button next to the record name:




4. Click **Run Preview**.

The results from the external web service are displayed to the right of the input data.

## Exporting an External Web Service Definition

An external web service definition contains the connection properties that enable Spectrum™ Technology Platform access a third-party web service over the Internet. Connection properties include information like the external web service's URL and your credentials. You can save this

information to a file so that you can easily add the external web service to another Spectrum™ Technology Platform server.

1. Open Management Console.
2. Go to **Resources > External Web Services**.
3. Check the box next to the external web service definition that you want to export then click the Export button .
4. Choose a location to save the file.


The external web service definition is saved as a file with a `.ews` file extension.

#### Related tasks

[Importing an External Web Service Definition](#) on page 150

## Importing an External Web Service Definition

An external web service definition contains the connection properties that enable Spectrum™ Technology Platform access a third-party web service over the Internet. Connection properties include information like the external web service's URL and your credentials. You can import an external web services definition file, which enables you to take an external web services definition from one Spectrum™ Technology Platform server and implement it on another server.


1. Log in to Management Console on the server onto which you want to import the web service definition.
2. Go to **Resources > External Web Services**.
3. Click the Import button .
4. Select the external web services definition file you want to import. The external web services definition file has an `.ews` file extension.
5. Click **OK**

#### Related tasks

[Exporting an External Web Service Definition](#) on page 149

## Deleting an External Web Service

Deleting an external web service will break existing flows that reference the external web service. If no flows reference the external web service you can safely delete the external web service.

1. In Management Console, go to **Resources > External Web Services**.
2. Check the box next to the web service you want to delete then click the Delete button .
3. Click **OK** to confirm.

**Related concepts**

[External Web Services](#) on page 139

# 6 - Flows

## In this section

---

|                                       |     |
|---------------------------------------|-----|
| Configuring Flow Defaults             | 153 |
| Scheduling Flows                      | 163 |
| Viewing Flow Status and History       | 166 |
| Triggering a Flow with a Control File | 168 |
| Command Line Execution                | 172 |



# Configuring Flow Defaults

## Setting Data Type Conversion Defaults

Data type conversion occurs when a dataflow automatically converts a field to the data type needed by a stage. Data type conversion also occurs when within some stages. For example, in Read from DB you can choose to have a field use the string data type even though the source data is in a numeric data type. The data is converted into the string data type when it is read into the dataflow.

There are two settings that you can use to control data type conversions. First, there are settings that determine how to format numeric, date, and time data converted into a string. For example, you may want date data that is converted into a string to be represented in the format mm/dd/yyyy rather than dd/mm/yyyy. The other setting controls what should happen if the system is unable to convert a field from one data type to another.

You can set the default data type conversion settings for your system in Management Console. You can override the default formats for individual dataflows in Enterprise Designer.

To set the default data type conversion options for your system, follow this procedure.

1. Open Management Console.
2. Go to **Flows > Defaults**.
3. Click **Data Type Conversions**.
4. Specify the format you want to use when converting date and time data to a string.
  - a) In the **Locale** field, select the country whose format you want to use for dates converted to a string. Your selection will determine the default values in the **Date**, **Time**, and **DateTime** fields. Your selection will also determine the language used when a month is spelled out. For example, if you specify English the first month of the year would be "January" but if you specify French it would be "Janvier."
  - b) In the **Date** field, select the format to use for date data when it is converted to a string. A list of the most commonly used formats for the selected locale is provided.

For example, if you choose the format **M/D/YY** and a date field contains 2012-3-2, that date data would be converted to the string 3/2/12.
  - c) In the **Time** field, select the format to use for time data when it is converted to a string. A list of the most commonly used formats for the selected locale is provided.

For example, if you choose the format **h:mm a** and a time field contains 23:00, that time data would be converted to the string 11:00 PM.

- d) In the **DateTime** field, select the format to use for fields containing the DateTime data type when converted to a string. A list of the most commonly used formats for the selected locale is provided.

For example, if you choose the format **M/d/yy h:mm a** and a DateTime field contains 2012-3-2 23:00, that DateTime data would be converted to the string 3/2/12 11:00 PM.

- e) In the **Whole numbers** field, select the formatting you want to use for whole numbers (data types float and double).

For example, if you choose the format **#,###** then the number 4324 would be formatted as 4,324.

**Note:** If you leave this field blank, numbers will be formatted as follows:

- No thousands separator is used
- The dot (".") is used as the decimal separator
- Numbers less than  $10^{-3}$  or greater than or equal to  $10^7$  are formatted in scientific notation
- Negative numbers have a minus sign ("-") in front of them
- Numbers that use the bigdecimal data type are in the format **#,###.000**

- f) In the **Decimal numbers** field, select the formatting you want to use for numbers that contain a decimal value (data types integer and long).

For example, if you choose the format **#,##0.0#** then the number 4324.25 would be formatted as 4,324.25.

**Note:** If you leave this field blank, numbers will be formatted as follows:

- No thousands separator is used
- The dot (".") is used as the decimal separator
- Numbers less than  $10^{-3}$  or greater than or equal to  $10^7$  are formatted in scientific notation
- Negative numbers have a minus sign ("-") in front of them
- Numbers that use the bigdecimal data type are in the format **#,###.000**

You can also specify your own date, time, and number formats if the ones available for selection do not meet your needs. To specify your own date or time format, type the format into the field using the notation described in [Date and Time Patterns](#) on page 155. To specify your own number format, type the format into the file using the notation described in [Number Patterns](#) on page 158.

5. Under **Null handling**, choose whether to perform type conversion if a field contains a null value. If you select any of the following options, either the dataflow or the record containing the null value will fail based on your selection in the **Failure handling** field.

**Fail null string**                      Fail the dataflow or record if type conversion is needed on a string field that contains a null value.

|                          |  |
|--------------------------|--|
| <b>Fail null Boolean</b> | Fail the dataflow or record if type conversion is needed on a Boolean field that contains a null value.  |
| <b>Fail null numeric</b> | Fail the dataflow or record if type conversion is needed on a numeric field that contains a null value. Numeric fields include double, float, long, integer, and Big Decimal fields. |
| <b>Fail null date</b>    | Fail the dataflow or record if type conversion is needed on a date field that contains a null value. This includes date, time, and DateTime fields.                                  |

6. In the **Failure handling** field, specify what to do when a field's value cannot be automatically converted to the data type required by a stage.

|  |   |
|--|---|
| <b>Fail the dataflow</b>                         | If a field cannot be converted the dataflow will fail.  |
| <b>Fail the record</b>                           | If a field cannot be converted the record will fail but the dataflow will continue to run.  |
| <b>Initialize the field using default values</b> | If a field cannot be converted the field's value is replaced with the value you specify here. This option is useful if you know that some records contain bad data and you want to replace the bad data with a default value. Specify a value for each data type. |

### Date and Time Patterns

When defining data type options for date and time data, you can create your own custom date or time pattern if the predefined ones do not meet your needs. To create a date or time pattern, use the notation described in the following table. For example, this pattern:

dd MMMM yyyy

Would produce a date like this:

14 December 2012

| Letter | Description                | Example |
|--------|----------------------------|---------|
| G      | Era designator             | AD      |
| yy     | Two-digit year             | 96      |
| yyyy   | Four-digit year            | 1996    |
| M      | Numeric month of the year. | 7       |

| Letter | Description  | Example |
|--------|--|---------|
| MM     | Numeric month of the year. If the number is less than 10 a zero is added to make it a two-digit number.                    | 07      |
| MMM    | Short name of the month  | Jul     |
| MMMM   | Long name of the month   | July    |
| w      | Week of the year   | 27      |
| ww     | Two-digit week of the year. If the week is less than 10 an extra zero is added.  | 06      |
| W      | Week of the month  | 2       |
| D      | Day of the year  | 189     |
| DDD    | Three-digit day of the year. If the number contains less than three digits, zeros are added.                               | 006     |
| d      | Day of the month   | 10      |
| dd     | Two-digit day of the month. Numbers less than 10 have a zero added.  | 09      |
| F      | Day of the week in month   | 2       |
| E      | Short name of the day of the week  | Tue     |
| EEEE   | Long name of the day of the week   | Tuesday |
| a      | AM/PM marker   | PM      |
| H      | Hour of the day, with the first hour being 0 and the last hour being 23.   | 0       |
| HH     | Two-digit hour of the day, with the first hour being 0 and the last hour being 23. Numbers less than 10 have a zero added. | 08      |

| Letter | Description   | Example           |
|--------|---|-------------------|
| k      | Hour of the day, with the first hour being 1 and the last hour being 24.  | 24                |
| kk     | Two-digit hour of the day, with the first hour being 1 and the last hour being 24. Numbers less than 10 have a zero added.                            | 02                |
| K      | Hour hour of the morning (AM) or afternoon (PM), with 0 being the first hour and 11 being the last hour.  | 0                 |
| KK     | Two-digit hour of the day, with the first hour being 1 and the last hour being 24. Numbers less than 10 have a zero added.                            | 02                |
| h      | Hour of the morning (AM) or afternoon (PM), with 1 being the first hour and 12 being the last hour.   | 12                |
| hh     | Two-digit hour of the morning (AM) or afternoon (PM), with 1 being the first hour and 12 being the last hour. Numbers less than 10 have a zero added. | 09                |
| m      | Minute of the hour  | 30                |
| mm     | Two-digit minutes of the hour. Numbers less than 10 have a zero added.  | 05                |
| s      | Second of the minute  | 55                |
| ss     | Two-digit second of the minute. Numbers less than 10 have a zero added.   | 02                |
| S      | Millisecond of the second   | 978               |
| SSS    | Three-digit millisecond of the second. Numbers containing fewer than three digits will have one or two zeros added to make them three digits.         | 978<br>078<br>008 |
| z      | Time abbreviation of the time zone name. If the time zone does not have a name, the GMT offset.   | PST<br>GMT-08:00  |

| Letter | Description  | Example                            |
|--------|--|------------------------------------|
| zzzz   | The full time zone name. If the time zone does not have a name, the GMT offset.      | Pacific Standard Time<br>GMT-08:00 |
| Z      | The RFC 822 time zone.   | -0800                              |
| X      | The ISO 8601 time zone.  | -08Z                               |
| XX     | The ISO 8601 time zone with minutes.   | -0800Z                             |
| XXX    | The ISO 8601 time zone with minutes and a colon separator between hours and minutes. | -08:00Z                            |

### Related tasks

[Setting Data Type Conversion Defaults](#) on page 153

### Related reference

[Number Patterns](#) on page 158

## Number Patterns

When defining data type options for numeric data, you can create your own custom number pattern if the predefined ones do not meet your needs. A basic number pattern consists of the following elements:

- A prefix such as a currency symbol (optional)
- A pattern of numbers containing an optional grouping character (for example a comma as a thousands separator)
- A suffix (optional)

For example, this pattern:

```
$ ###,###.00
```

Would produce a number formatted like this (note the use of a thousands separator after the first three digits):

```
$232,998.60
```

### *Patterns for Negative Numbers*

By default, negative numbers are formatted the same as positive numbers but have the negative sign added as a prefix. The character used for the number sign is based on the locale. The negative sign is "-" in most locales. For example, if you specify this number pattern:

```
0.00
```

The number negative ten would be formatted like this in most locales:

```
-10.00
```

However, if you want to define a different prefix or suffix to use for negative numbers, specify a second pattern, separating it from the first pattern with a semicolon (";"). For example:

```
0.00; (0.00)
```

In this pattern, negative numbers would be contained in parentheses:

```
(10.00)
```

### Scientific Notation

If you want to format a number into scientific notation, use the character `E` followed by the minimum number of digits you want to include in the exponent. For example, given this pattern:

```
0.###E0
```

The number `1234` would be formatted like this:

```
1.234E3
```

In other words,  $1.234 \times 10^3$ .

Note the following:

- The number of digit characters after the exponent character gives the minimum exponent digit count. There is no maximum.
- Negative exponents are formatted using the localized minus sign, not the prefix and suffix from the pattern.
- Scientific notation patterns cannot contain grouping separators (for example, a thousands separator).

### Special Number Pattern Characters

The following characters are used to produce other characters, as opposed to being reproduced literally in the resulting number. If you want to use any of these special characters as literal characters in your number pattern's prefix or suffix, surround the special character with quotes.

| Symbol | Description  |
|--------|--|
| 0      | Represents a digit in the pattern including zeros where needed to fill in the pattern. For example, the number twenty-seven when applied to this pattern:<br>0000<br>Would be:<br>0027 |

| Symbol | Description   |
|--------|---|
| #      | <p>Represents a digit but zeros are omitted. For example, the number twenty-seven when applied to this pattern:</p> <pre>####</pre> <p>Would be:</p> <pre>27</pre>  |
| .      | <p>The decimal separator or monetary decimal separator used in the selected locale. For example, in the U.S. the dot (.) is used as the decimal separator but in France the comma (,) is used as the decimal separator.</p>   |
| -      | <p>The negative sign used in the selected locale. For most locals this is the minus sign (-).</p>   |
| ,      | <p>The grouping character used in the selected locale. The appropriate character for the selected locale will be used. For example, in the U.S., the comma (,) is used as a separator.</p> <p>The grouping separator is commonly used for thousands, but in some countries it separates ten-thousands. The grouping size is a constant number of digits between the grouping characters, such as 3 for 100,000,000 or 4 for 1,0000,0000. If you supply a pattern with multiple grouping characters, the interval between the last one and the end of the integer is the one that is used. For example, all the following patterns produce the same result:</p> <pre>#, ##, ###, ####</pre> <pre>#####, #####</pre> <pre>##, ####, #####</pre> |
| E      | <p>Separates mantissa and exponent in scientific notation. You do not need to surround the E with quotes in your pattern. See <a href="#">Scientific Notation</a> on page 159.</p>  |
| ;      | <p>Separates positive and negative subpatterns. See <a href="#">Patterns for Negative Numbers</a> on page 158.</p>  |



| Symbol | Description  |
|--------|--|
| %      | <p>Multiply the number by 100 and show the number as a percentage. For example, the number .35 when applied to this pattern:</p> <pre>##%</pre> <p>Would produce this result:</p> <pre>35%</pre>                           |
| ¤      | <p>The currency symbol for the selected locale. If doubled, the international currency symbol is used. If present in a pattern, the monetary decimal separator is used instead of the decimal separator.</p>               |
| '      | <p>Used to quote special characters in a prefix or suffix. For example:</p> <pre>"' #' #"</pre> <p>Formats 123 to:</p> <pre>"#123"</pre> <p>To create a single quote itself, use two in a row:</p> <pre>"# o''clock"</pre> |

#### Related tasks

[Setting Data Type Conversion Defaults](#) on page 153

#### Related reference

[Date and Time Patterns](#) on page 155

## Setting the Malformed Records Default

A malformed record is an input record that Spectrum™ Technology Platform cannot parse. By default, if the input data for a job contains one malformed record, the job will terminate. You can change this setting to allow more malformed input records, or even allow an unlimited number of malformed records. This procedure describes how to set a default malformed record threshold for jobs on your system.

**Note:** You can override the default malformed record threshold for a job by opening the job in Enterprise Designer and going to **Edit > Job Options**.

1. Open Management Console.
2. Go to **Flows > Defaults**.
3. Click **Malformed Records**.

- Select one of the following:

**Terminate jobs containing this many malformed records**

Select this option to terminate jobs if the input data contains one or more malformed records. Enter the number of malformed records that you want to trigger the termination of the job. The default is 1.

**Do not terminate flows with malformed records**

Select this option to allow an unlimited number of malformed records in the input data.

## Setting Report Defaults

Reports are generated by jobs that contain a report stage. Reports can include processing summaries such as the number of records processed by the job, or postal forms such as the USPS CASS 3553 form. Some modules come with predefined reports. You can also create custom reports. Setting report defaults establishes the default settings for saving reports. The default settings can be overridden for a job, or for a particular report within a job, by using Enterprise Designer.

This procedure describes how to set the default reporting options for your system.

- Open Management Console.
- Go to **Flows > Defaults**.
- Click **Reports**.
- Choose the format you want to use to save reports. Reports can be saved as HTML, PDF, or text.
- Choose where you want to save reports.

**Save reports to job history**

Saves reports on the server as part of the job history. This makes it convenient for Management Console and Enterprise Designer users to view reports since the reports are available in the execution history.

**Save reports to a file**

Saves reports to a file in the location you specify. This is useful if you want to share reports with people who are not Spectrum™ Technology Platform users. It is also useful if you want to create an archive of reports in a different location. To view reports saved in this manner you can use any tool that can open the report's format, such as a PDF viewer for PDF reports or a web browser for HTML reports.

- If you selected **Save reports to a file**, complete these fields.

**Report location**

The folder where you want to save reports.

**Append to report name**

Specifies variable information to include in the file name. You can choose one or more of the following:

**Job ID**

A unique ID assigned to a job execution. The first time you run a job on your system the job has an ID of 1. The second

time you run a job, either the same job or another job, it has a job ID of 2, and so on.

**Stage** The name of the stage that contributed data to the report, as specified in the report stage in Enterprise Designer.

**Date** The day, month, and year that the report was created.

**Overwrite existing reports** Replaces previous reports that have the same file name with the new report. If you do not select this option and there is an existing report that has the same name as the new report, the job will complete successfully but the new report will not be saved. A comment will appear in the execution history indicating that the report was not saved.

## Scheduling Flows

### Scheduling a Flow


Scheduling a flow enables a job or dataflow to run automatically at a specified time or times. In Management Console, you can schedule a flow to run:

- On a specific date and time
- On a recurring basis on the specified dates and times

**Note:** In order to create, edit, or view a schedule, you must have View permission for the secured entity type being scheduled, either **Dataflows** or **Process Flows**.

1. If you have not already done so, expose the flow.

You can expose a flow by opening it in Enterprise Designer and selecting **File > Expose/Unexpose and Save**.

2. Open Management Console.
3. Go to **Flows > Schedules**.
4. Click the Add button .
5. In the **Name** field, enter the name you want to give to this schedule. This is the name that will be displayed in the schedules listing.
6. In the **Flow** field, enter the job or process flow that you want to run. Only jobs and process flows that are saved and exposed are available here.

7. After you specify a flow, additional fields appear below the **Flow** field, one field for each of the flow's source stages (such as Read from File) and sink stages (such as Write to File). These fields show the files that will be used when the flow is executed by this schedule. By default, the flow will use the files specified in the flow's sources and sinks. You can specify different files to use when this schedule runs by replacing the file path with the path to another file. For example, if your flow has a Read from File stage that reads data from `C:\FlowInput\Customers.csv` but you want to use data from `C:\FlowInput\UpdatedCustomers.csv` when this schedule runs, you would specify `C:\FlowInput\UpdatedCustomers.csv` in the Read from File field.

**Note:** In order change the files used in the source and sink stages you must have Read permission for the **Resources - File Servers** secured entity type.

Note that when a flow is triggered by a schedule the files used by a flow must reside on the Spectrum™ Technology Platform server or on a file server defined as an external resource in Management Console. This applies both to jobs as well as job activities within a process flow. If a source or sink stage references a file on a client computer do one of the following:

| Option   | Description  |
|--|--|
| <b>Option 1: Modify the dataflow</b>   | <p>Move the file to the Spectrum™ Technology Platform server or file server then modify the dataflow:</p> <ol style="list-style-type: none"> <li>1. Open the dataflow in Enterprise Designer.</li> <li>2. Double-click the source or sink stage.</li> <li>3. In the <b>File name</b> field, click the browse button.</li> <li>4. Click <b>Remote Machine</b> then select the file you want.</li> </ol> <p><b>Note:</b> If you are running Enterprise Designer on the same machine as the Spectrum™ Technology Platform server, it will appear that clicking Remote Machine is no different than clicking My Computer. However, you must select the file using Remote Machine in order for the system to recognize the file as being on the Spectrum™ Technology Platform server.</p> |
| <b>Option 2: Override the dataflow file location when this schedule runs</b> | <p>You can override the file references contained in the flow when this schedule runs. To do this, replace the default file shown in each source and sink field with a path to a file on the Spectrum™ Technology Platform server or a file server resource defined in Management Console.</p>   |

8. In the **Trigger** field, select one of the following:

**Date/Time** Run the flow once at a specific date and time.

|                            |  |
|----------------------------|--|
| <b>Recurring Date/Time</b> | Run the flow on multiple dates and times using a pattern of recurrence.  |
| <b>Control File</b>        | Run the flow when a file appears in a specified directory. For information on using a control file, see <a href="#">Triggering a Flow with a Control File</a> on page 168. |

- Specify the date and time or the recurrence interval for running the flow.

**Note:** If you chose **Recurring Date/Time** in the **Trigger** field, be sure to select a start date that conforms to the recurrence pattern. For example, if you chose to run the flow on the first Monday of the month, be sure to select a date that is the first Monday of the month. If you select a date that does not meet the recurrence pattern, the flow may run at an unexpected time. Also, selecting a start date in the past may result in the flow running at an unexpected time.

- If the flow is configured to send email notifications you can specify additional recipients for the notifications that will be sent when this schedule runs. The recipients you specify here will receive notifications in addition to those recipients specified in the flow's notification settings. To configure a flow to send notifications, open the flow in Enterprise Designer and go to **Edit > Notifications**.
- Click **Save**.

#### Related tasks

[Configuring a Mail Server](#) on page 215

## Viewing Schedules

A flow schedule defines when a job or process flow will run. You can view the flow schedules on your system as well as the results of each run.


- Open Management Console.
- Go to **Flows > Schedules**.

A list of the flow schedules is displayed. To sort the list of schedules, click the column heading. You can filter the list by typing keywords into the **Filter** field. Note that the filter field only filters on the **Schedule Name**, **User**, **Next Run**, and **Last Run** columns.

## Deleting a Schedule

A scheduled flow runs automatically at a set time or on a repeating schedule. If you no longer want a flow to run on a schedule, you can delete the schedule. Deleting a schedule does not delete the flow itself.

To delete a schedule:

1. Open Management Console.
2. Go to **Flows > Schedules**.
3. Check the box next to the schedule you want to delete then click the Delete button .

## Viewing Flow Status and History


You can view a history of job, process flow, and service execution in Management Console and Enterprise Designer.

### *In Management Console*

To view flow status and history in Management Console, go to **Flows > History**. The **Flows** tab shows job and process flow history, and the **Transactions** tab shows services history.

**Note:** For flow history, the record counts shown when you hover over the **Results** column reflect the total number of records written as output by all the dataflow's sinks. This number may differ from the number of input records if the dataflow combines records, splits records, or creates new records.

By default, transaction history is disabled because enabling transaction history can have an adverse impact on performance. If you want to see transaction history you must turn on transaction history logging by clicking the **Transaction logging** switch. To view user activity, consider using the audit log which you can access under **System > Logs**.

The flow history list updates automatically every 30 seconds. If you want to update it sooner, click the Refresh button .

### *In Enterprise Designer*

To view flow status and history in Enterprise Designer, go to **View > Execution History**.

The flow history list updates automatically every 30 seconds. If you experience slowness when viewing execution history uncheck the **Auto refresh** box.

The **Jobs** tab is used to monitor job status and to pause, resume, or cancel jobs that are running as well as delete completed jobs.

**Note:** The record counts shown on the **Jobs** tab reflect the total number of records written as output by all the dataflow's sinks. This number may differ from the number of input records if the dataflow combines records, splits records, or creates new records.

- The **Succeeded** column shows the total number of records written as output by all the dataflow's sinks that have an empty value in the Status field.
- The **Failed** column shows the total number of records written as output by the dataflow's sinks that have a value of F in the Status field.


- The **Malformed** column shows the total number of records coming out of all source stage error ports.

The **Process Flows** tab is used to monitor process flow status and to cancel process flows that are running as well as delete completed process flows. If you click the plus sign next to any given process flow, you will view Activity Status information for the process flow. The following information is included in this area:

|                      |  |          |                          |          |   |           |                                |                      |  |
|----------------------|--|----------|--------------------------|----------|---|-----------|--------------------------------|----------------------|--|
| <b>ActivityName</b>  | Includes the names of all activities, including any success activities, that make up the process flow.   |          |                          |          |   |           |                                |                      |  |
| <b>State</b>         | The status of the activity (failed, succeeded, running, canceled).   |          |                          |          |   |           |                                |                      |  |
| <b>ReturnCode</b>    | A code that indicates the result of the process flow. One of the following: <table> <tr> <td><b>1</b></td> <td>The process flow failed.</td> </tr> <tr> <td><b>0</b></td> <td>The process flow finished successfully.</td> </tr> <tr> <td><b>-1</b></td> <td>The process flow was canceled.</td> </tr> <tr> <td><b>Other numbers</b></td> <td>If the process flow contains a Run Program activity, the external program may return codes of its own. Any values in the ReturnCode column other than 1, 0, and -1 are from the external program. See the external program's documentation for an explanation of its return codes.</td> </tr> </table> | <b>1</b> | The process flow failed. | <b>0</b> | The process flow finished successfully. | <b>-1</b> | The process flow was canceled. | <b>Other numbers</b> | If the process flow contains a Run Program activity, the external program may return codes of its own. Any values in the ReturnCode column other than 1, 0, and -1 are from the external program. See the external program's documentation for an explanation of its return codes. |
| <b>1</b>             | The process flow failed.   |          |                          |          |   |           |                                |                      |  |
| <b>0</b>             | The process flow finished successfully.  |          |                          |          |   |           |                                |                      |  |
| <b>-1</b>            | The process flow was canceled.   |          |                          |          |   |           |                                |                      |  |
| <b>Other numbers</b> | If the process flow contains a Run Program activity, the external program may return codes of its own. Any values in the ReturnCode column other than 1, 0, and -1 are from the external program. See the external program's documentation for an explanation of its return codes.   |          |                          |          |   |           |                                |                      |  |
| <b>Started</b>       | The date and time the activity started.  |          |                          |          |   |           |                                |                      |  |
| <b>Finished</b>      | The date and time the activity ended.  |          |                          |          |   |           |                                |                      |  |
| <b>Comment</b>       | Any comments associated with the activity.   |          |                          |          |   |           |                                |                      |  |

## Downloading Flow History

You can download the information shown in the History page in Management Console to a Microsoft Excel file.

1. Open Management Console.
2. Go to **Flows > History**.
3. To download history information for services, click **Transaction History**. To download history for jobs and process flows, leave the **Flows** tab selected.
4. Click the Download button .

**Tip:** If you want to download only certain entries in the history list, modify the filter settings to show only the history you want to download.


## Triggering a Flow with a Control File

A flow can run automatically when a control file is detected in a monitored directory. This feature is useful in situations where the flow needs another process to complete before running. For example, you may have a flow that needs an input file generated by another business process. You can set up the other process to place a control file into a folder, and configure Spectrum™ Technology Platform to run a flow when that control file appears.

**Note:** Be sure that the control file is placed in the monitored folder only after all files required by the flow are in place and ready for processing.

1. If you have not already done so, expose the flow.

You can expose a flow by opening it in Enterprise Designer and selecting **File > Expose/Unexpose and Save**.

2. Open Management Console.
3. Go to **Flows > Schedules**.
4. Click the Add button .
5. In the **Name** field, enter the name you want to give to this schedule. This is the name that will be displayed in the schedules listing.
6. In the **Flow** field, enter the job or process flow that you want to run. Only jobs and process flows that are saved and exposed are available here.
7. After you specify a flow, additional fields appear below the **Flow** field, one field for each of the flow's source stages (such as Read from File) and sink stages (such as Write to File). These fields show the files that will be used when the flow is executed by this schedule. By default, the flow will use the files specified in the flow's sources and sinks. You can specify different files to use when this schedule runs by replacing the file path with the path to another file. For example, if your flow has a Read from File stage that reads data from `C:\FlowInput\Customers.csv` but you want to use data from `C:\FlowInput\UpdatedCustomers.csv` when this schedule runs, you would specify `C:\FlowInput\UpdatedCustomers.csv` in the Read from File field.

**Note:** In order change the files used in the source and sink stages you must have Read permission for the **Resources - File Servers** secured entity type.

Note that when a flow is triggered by a schedule the files used by a flow must reside on the Spectrum™ Technology Platform server or on a file server defined as an external resource in Management Console. This applies both to jobs as well as job activities within a process flow. If a source or sink stage references a file on a client computer do one of the following:



| Option   | Description  |
|--|--|
| <b>Option 1: Modify the dataflow</b>   | <p>Move the file to the Spectrum™ Technology Platform server or file server then modify the dataflow:</p> <ol style="list-style-type: none"> <li>1. Open the dataflow in Enterprise Designer.</li> <li>2. Double-click the source or sink stage.</li> <li>3. In the <b>File name</b> field, click the browse button.</li> <li>4. Click <b>Remote Machine</b> then select the file you want.</li> </ol> <p><b>Note:</b> If you are running Enterprise Designer on the same machine as the Spectrum™ Technology Platform server, it will appear that clicking Remote Machine is no different than clicking My Computer. However, you must select the file using Remote Machine in order for the system to recognize the file as being on the Spectrum™ Technology Platform server.</p> |
| <b>Option 2: Override the dataflow file location when this schedule runs</b> | <p>You can override the file references contained in the flow when this schedule runs. To do this, replace the default file shown in each source and sink field with a path to a file on the Spectrum™ Technology Platform server or a file server resource defined in Management Console.</p>   |

8. In the **Trigger** field, choose **Control File**.
9. In the **Control file** field, specify the full path and name of the control file that will trigger the flow. You can specify an exact file name or you can use the asterisk (\*) as a wild card. For example, \*.trg would trigger the flow when any file with a .trg extension appears in the folder.

The presence of a control file indicates that all files required for the flow are in place and ready to be used in the flow.

The control file can be a blank file. For jobs, the control file can specify overrides to file paths configured in the Write to File or Read from File stages. To use a control file to override the file paths, specify the Read from File or Write from File stage names along with the input or output file as the last arguments like this:

```
stagename=filename
```

For example:

```
Read\ from\ File=file:C:/myfile_input.txt
Write\ to\ File=file:C:/myfile_output.txt
```

The stage name specified in the control file must match the stage label shown under the stage's icon in the dataflow. For example, if the input stage is labeled "Read From File" you would specify:

```
Read\ From\ File=file:C:/inputfile.txt
```

If the input stage is labeled "Illinois Customers" you would specify:

```
Illinois\ Customers=file:C:/inputfile.txt
```

When overriding a Read from File or Write to File location be sure to follow these guidelines:

- Start the path with the "file:" protocol. For example, on Windows specify "file:C:/myfile.txt" and on Unix or Linux specify "file:/testfiles/myfile.txt".
- The contents of the file must use an ASCII-based ISO-8559-1 (Latin-1) compatible character encoding.
- You must use forward slashes (/) in file paths, not backslashes.
- Spaces in stage names need to be escaped with a backslash.
- Stage names are case sensitive.

**Note:** If there are multiple schedules that use a control file trigger, it is important that they each monitor different control files. Otherwise, the same control file may trigger multiple jobs or process flows causing unexpected behavior. For organizational purposes we recommend putting all required files and the control file in a dedicated directory.

10. In the **Poll interval** field, specify how frequently to check for the presence of the control file. For example, if you specify 10, the monitor will look every 10 seconds to see if the control file is in the monitored folder.

The default is 60 seconds.

11. In the **Working folder** field, specify a folder where the control file will reside temporarily while the flow runs. Spectrum™ Technology Platform copies the file from the monitored folder to the working folder before running the flow. This clears out the monitored folder, which prevents the flow from being kicked off again by the same control file.
12. In the **Working folder options** field, specify what to do with the files in the working folder when the flow finishes running.

**Keep** Leaves the files in their current location with their current name. If you select this option, the files in the working folder will be overwritten each time this schedule runs.

**Move to** Moves the files from the working folder to a folder you specify. This allows you to preserve the files that were in the working folder by moving them to another location so that they are not overwritten the next time the file monitor runs. You can also use this option to move the files to another monitored folder to trigger a downstream process, like another dataflow or some other process.

**Rename with time stamp** Adds a time stamp to the file name in the working folder. This allows you to preserve a copy of the files in the working folder since the renamed file will have a unique name and so will not be overwritten the next time the monitor runs a dataflow.

**Delete** Deletes the files from the working folder after the flow finishes running.

13. If the flow is configured to send email notifications you can specify additional recipients for the notifications that will be sent when this schedule runs. The recipients you specify here will receive notifications in addition to those recipients specified in the flow's notification settings. To configure a flow to send notifications, open the flow in Enterprise Designer and go to **Edit > Notifications**.
14. Click **Save**.

#### Example: Monitored Folder and Working Folder

Let's say you have a car repair shop. Each day you want to mail the previous day's customers a coupon for a discount on future service. To accomplish this, you have a dataflow that takes the list of customers for the day, ensures the names have the correct casing, and validates the address. The list of customers for the day is generated by another system every evening. This other system generates a file containing the customer list, and you want to use this file as the input to the dataflow.

The system that generates the customer list puts it in a folder named `DailyCustomerReport`. It also places a blank trigger file in the folder when it is done. So you configure Spectrum™ Technology Platform to monitor this folder, specifying the following as the trigger file:

```
C:\DailyCustomerReport\*.trg
```

This tells Spectrum™ Technology Platform to run the dataflow whenever any file with a `.trg` extension appears in this folder. You could also specify a specific file name, but in this example we are using a wild card.

When a `.trg` file is detected in the `DailyCustomerReport` folder, Spectrum™ Technology Platform needs to move it to another folder before running the dataflow. The file must be moved because otherwise it would be detected again at the next polling interval, and this would result in the dataflow running again. So the file is moved to a "working folder" where it resides during the execution of the dataflow. You choose as the working folder `C:\SpectrumWorkingFolder`.

After the dataflow is finished processing the customer list, you want the trigger file to be moved to another location where it will trigger another process for billing. So, you select the **Move to** option and choose a folder named `C:\DailyBilling`.

So in this example, the trigger file starts off in `C:\DailyCustomerReport` and is then moved to the working folder `C:\SpectrumWorkingFolder`. After the dataflow is done, the trigger file is moved to `C:\DailyBilling` to initiate the billing process.

# Command Line Execution

## Running A Job from the Command Line

Before you can run a job from the command line, it must be exposed. To expose a job, open the job in Enterprise Designer and select **File > Expose/Unexpose and Save**.

To run a job from the command line, you must install the job executor utility on the system where you want to run the job. The Job Executor is available from the Spectrum™ Technology Platform Welcome page on the Spectrum™ Technology Platform server (for example, <http://myserver:8080>).

### Usage

```
java -jar jobexecutor.jar -u UserID -p Password -j Job [Optional Arguments]
```

| Required | Argument                | Description  |
|----------|-------------------------|--|
| No       | -?                      | Prints usage information.  |
| No       | -d <i>delimiter</i>     | Sets instance/status delimiter. This appears only when running in synchronous mode. For information about synchronous mode, see the description of the -w argument.                                  |
| No       | -e                      | Use a secure HTTPS connection for communication with the Spectrum™ Technology Platform server.   |
| No       | -f <i>property file</i> | Specifies a path to a job property file. A job property file contains job executor arguments. For more information on job property files, see <a href="#">Using a Job Property File</a> on page 179. |
| No       | -h <i>host name</i>     | Specifies the name or IP address of the Spectrum™ Technology Platform server.  |
| No       | -i <i>poll interval</i> | Specifies how often to check for completed jobs, in seconds. This applies only in synchronous mode.  |
| Yes      | -j <i>job name</i>      | A comma-separated list of jobs to run. Job names are case-sensitive. Jobs are started in the order listed.   |

| Required | Argument                      | Description   |
|----------|-------------------------------|---|
| No       | <code>-n email list</code>    | Specifies a comma-separated list of additional email addresses for configured job notifications.  |
| No       | <code>-o property file</code> | <p>Specifies a path to a dataflow options property file. Use a dataflow options property file to set options for stages in the dataflow. In order to set dataflow options using a property file, you must configure the dataflow to expose stage options at runtime. For more information, see <a href="#">Adding Dataflow Runtime Options</a>.</p> <p>For example, a dataflow options properties file for a dataflow that contains an Assign GeoTAX Info stage may look like this:</p> <pre data-bbox="868 835 1437 961">OutputCasing=U UseStreetLevelMatching=N TaxKey=T Database.GTX=gsl</pre>   |
| Yes      | <code>-p password</code>      | The password of the user.   |
| No       | <code>-r</code>               | <p>Specify this argument to return a detailed report about the job. This option only works if you also specify <code>-w</code>. The report contains this information:</p> <ul style="list-style-type: none"> <li>• <b>Position 1</b>—Name of job</li> <li>• <b>Position 2</b>—Job process ID</li> <li>• <b>Position 3</b>—Status</li> <li>• <b>Position 4</b>—Start Date/Time (MM/DD/YYYY HH:MM:SS)</li> <li>• <b>Position 5</b>—End Date/Time (MM/DD/YYYY HH:MM:SS)</li> <li>• <b>Position 6</b>—Number of successful records</li> <li>• <b>Position 7</b>—Number of failed records</li> <li>• <b>Position 8</b>—Number of malformed records</li> <li>• <b>Position 9</b>—Currently unused</li> </ul> <p>For example:</p> <pre data-bbox="852 1732 1437 1827">MySimpleJob 4 succeeded 04/09/2010 14:50:47 04/09/2010 14:50:47 100 0 0 </pre> |

| Required Argument   | Description   |
|---|---|
|   | The information is delimited using the delimiter specified in the <code>-d</code> argument.   |
| No <code>-s <i>port</i></code>                              | The socket (port) on which the Spectrum™ Technology Platform server is running. The default is 8080.  |
| No <code>-t <i>timeout</i></code>                           | Sets the timeout (in seconds) for synchronous mode. For information about synchronous mode, see the description of the <code>-w</code> argument. The default is 3600. The maximum is 2147483. This is a global, aggregate timeout and represents the maximum time to wait for all spawned jobs to complete.   |
| Yes <code>-u <i>user name</i></code>                        | The login name of the user.   |
| No <code>-v</code>  | Return verbose output.  |
| No <code>-w</code>  | Runs job executor in synchronous mode. This means that job executor remains running until the job completes. If you do not specify <code>-w</code> , job executor exits after starting the job. However, if the job reads from or writes to files on the server, job executor will run in synchronous mode even if you do not specify <code>-w</code> . |
| No <code><i>StageName=Protocol:FileName</i></code>          | Overrides the input or output file specified in Read from File or Write to File. For more information, see <a href="#">Overriding Job File Locations</a> on page 175.   |
| No <code><i>StageName:schema=Protocol:SchemaFile</i></code> | Overrides the file layout definition specified in Read from File or Write to File with one defined in a schema file. For more information, see <a href="#">Overriding the File Format at the Command Line</a> on page 177.  |

### Example Use of Job Executor

This example shows command line invocation and output:

```
D:\spectrum\job-executor>java -jar jobexecutor.jar -u user123
-p "mypassword" -j validateAddressJob1 -h spectrum.example.com
-s 8888 -w -d "%" -i 1 -t 9999

validateAddressJob1%105%succeeded
```

In this example, the output indicates that the job named 'validateAddressJob1' ran (with identifier 105) with no errors. Other possible results include "failed" or "running."

### Related reference

[Using a Job Property File](#) on page 179

### Overriding Job File Locations

When you run a job at the command line using job executor or the Administration Utility, you can override the input file specified in the dataflow's source stage (such as Read from File), as well as the output file specified in the dataflow's sink stage (such as Write to File).

To do this in job executor, specify the following at the end of the job executor command line command:

```
StageName=Protocol:FileName
```

In the Administration Utility, use the `--l` argument in the `job execute` command:

```
--l StageName=Protocol:FileName
```

Where:

#### **StageName**

The stage label shown under the stage's icon in the dataflow in Enterprise Designer. For example, if the stage is labeled "Read from File" you would specify `Read from File` for the stage name.

To specify a stage within an embedded dataflow or a subflow, preface the stage name with the name of the embedded dataflow or subflow, followed by a period then the stage name:

```
EmbeddedOrSubflowName.StageName
```

For example, to specify a stage named Write to File in a subflow named Subflow1, you would specify:

```
Subflow1.Write to File
```

To specify a stage in an embedded dataflow that is within another embedded dataflow, add the parent dataflow, separating each with a period. For example, if Embedded Dataflow 2 is inside Embedded Dataflow 1, and you want to specify the Write to File stage in Embedded Dataflow 2, you would specify this:

```
Embedded Dataflow 1.Embedded Dataflow 2.Write to File
```

#### **Protocol**

A communication protocol. One of the following:

**file** Use the file protocol if the file is on the same machine as the Spectrum™ Technology Platform server. For example, on Windows specify:

```
"file:C:/myfile.txt"
```

On Unix or Linux specify:

```
"file:/testfiles/myfile.txt"
```

**esclient** Use the esclient protocol if the file is on the computer where you are executing the job if it is a different computer from the one running the Spectrum™ Technology Platform server. Use the following format:

```
esclient:ComputerName/path to file
```

For example,

```
esclient:mycomputer/testfiles/myfile.txt
```

**Note:** If you are executing the job on the server itself, you can use either the file or esclient protocol, but are likely to have better performance using the file protocol.

If the host name of the Spectrum™ Technology Platform server cannot be resolved, you may get the error "Error occurred accessing file". To resolve this issue, open this file on the server:

```
SpectrumLocation/server/app/conf/spectrum-container.properties.
```

Set the `spectrum.runtime.hostname` property to the IP address of the server.

**esfile** Use the esfile protocol if the file is on a file server. The file server must be defined in Management Console as a resource. Use the following format:

```
esfile://file server/path to file
```

For example,

```
esfile://myserver/testfiles/myfile.txt
```

Where myserver is an FTP file server resource defined in Management Console.

**webhdfs** Use the webhdfs protocol if the file is on a Hadoop Distributed File Server. The HDFS server must be defined in Management Console as a resource. Use the following format:

```
webhdfs://file server/path to file
```

For example,

```
webhdfs://myserver/testfiles/myfile.txt
```

Where myserver is an HDFS file server resource defined in Management Console.

### **FileName**

The full path to the file you want to use as input or output.

**Note:** You must use forward slashes (/) in file paths, not backslashes.

To specify multiple overrides, separate each override with a comma.



**Example File Override**

The following job executor command would use the file `C:/myfile_input.txt` as the input file for the Read from File stage and would use the file `C:/myfile_output.txt` as the output file for the Write to File stage.

```
java -jar jobexecutor.jar -j Job1 -u Bob1234 -p "" "Read from
File"="file:C:/myfile_input.txt" "Write to
File"="file:C:/myfile_output.txt"
```

**Overriding the File Format at the Command Line**

When you run a job using job executor or the Administration Utility, you can override the file layout (or schema) of the file specified in the dataflow's Read from File stage and Write to File stage.

To do this in job executor, specify the following at the end of the job executor command line command:

```
StageName:schema=Protocol:SchemaFile
```

In the Administration Utility, use the `--l` argument in the `job execute` command:

```
--l StageName:schema=Protocol:SchemaFile
```

Where:

***StageName***

The stage label shown under the stage's icon in the dataflow in Enterprise Designer. For example, if the stage is labeled "Read from File" you would specify `Read from File` for the stage name.

To specify a stage within an embedded dataflow or a subflow, preface the stage name with the name of the embedded dataflow or subflow, followed by a period then the stage name:

```
EmbeddedOrSubflowName.StageName
```

For example, to specify a stage named Write to File in a subflow named Subflow1, you would specify:

```
Subflow1.Write to File
```

To specify a stage in an embedded dataflow that is within another embedded dataflow, add the parent dataflow, separating each with a period. For example, if Embedded Dataflow 2 is inside Embedded Dataflow 1, and you want to specify the Write to File stage in Embedded Dataflow 2, you would specify this:

```
Embedded Dataflow 1.Embedded Dataflow 2.Write to File
```

***Protocol***

A communication protocol. One of the following:

**file** Use the file protocol if the file is on the same machine as the Spectrum™ Technology Platform server. For example, on Windows specify:

```
"file:C:/myfile.txt"
```

On Unix or Linux specify:

```
"file:/testfiles/myfile.txt"
```

**esclient** Use the esclient protocol if the file is on the computer where you are executing the job if it is a different computer from the one running the Spectrum™ Technology Platform server. Use the following format:

```
esclient:ComputerName/path to file
```

For example,

```
esclient:mycomputer/testfiles/myfile.txt
```

**Note:** If you are executing the job on the server itself, you can use either the file or esclient protocol, but are likely to have better performance using the file protocol.

If the host name of the Spectrum™ Technology Platform server cannot be resolved, you may get the error "Error occurred accessing file". To resolve this issue, open this file on the server:

```
SpectrumLocation/server/app/conf/spectrum-container.properties.  
Set the spectrum.runtime.hostname property to the IP address of the server.
```

**esfile** Use the esfile protocol if the file is on a file server. The file server must be defined in Management Console as a resource. Use the following format:

```
esfile://file server/path to file
```

For example,

```
esfile://myserver/testfiles/myfile.txt
```

Where myserver is an FTP file server resource defined in Management Console.

**webhdfs** Use the webhdfs protocol if the file is on a Hadoop Distributed File Server. The HDFS server must be defined in Management Console as a resource. Use the following format:

```
webhdfs://file server/path to file
```

For example,

```
webhdfs://myserver/testfiles/myfile.txt
```

Where myserver is an HDFS file server resource defined in Management Console.

### SchemaFile

The full path to the file that defines the layout you want to use.

**Note:** You must use forward slashes (/) in file paths, not backslashes.

To create a schema file, define the layout you want in Read from File or Write to File, then click the **Export** button to create an XML file that defines the layout.

**Note:** You cannot override a field's data type in a schema file when using job executor. The value in the <Type> element, which is a child of the <FieldSchema> element, must match the field's type specified in the dataflow's Read from File or Write to File stage.

#### Example File Format Override

The following job executor command would use the file `C:/myschema.xml` as the layout definition for the file read in by the Read from File stage.

```
java -jar jobexecutor.jar -j Job1 -u Bob1234 -p "" "Read from File":schema="file:C:/myschema.xml"
```

### Using a Job Property File

A job property file contains arguments that control the execution of jobs when you use the job executor or the Administration Utility to run a job. Use a job property file if you want to reuse arguments by specifying a single argument at the command line (`-f`) rather than specifying each argument individually at the command line.

To create a property file, create a text file with one argument on each line. For example:

```
d %
h spectrum.mydomain.com
i 30
j validateAddressJob1
u user
p password
s 8888
t 9999
w true
```

The job property file can contain these arguments:

| Required | Argument           | Description  |
|----------|--------------------|--|
| No       | ?                  | Prints usage information.  |
| No       | d <i>delimiter</i> | Sets instance/status delimiter. This appears in synchronous output only.                       |
| No       | e                  | Use a secure HTTPS connection for communication with the Spectrum™ Technology Platform server. |

| Required | Argument              | Description   |
|----------|-----------------------|---|
| No       | <i>h hostname</i>     | Specifies the name or IP address of the Spectrum™ Technology Platform server.   |
| No       | <i>i pollinterval</i> | Specifies how often to check for completed jobs, in seconds. This applies only in synchronous mode.   |
| Yes      | <i>j jobname</i>      | A comma-separated list of jobs to run. Job names are case-sensitive. Jobs are started in the order listed.  |
| No       | <i>n emailist</i>     | Specifies a comma-separated list of additional email addresses for configured job notifications.  |
| Yes      | <i>p password</i>     | The password of the user.   |
| No       | <i>r</i>              | <p>Returns a delimited list with the following information about the job written to standard output:</p> <ul style="list-style-type: none"> <li>• <b>Position 1</b>—Name of job</li> <li>• <b>Position 2</b>—Job process ID</li> <li>• <b>Position 3</b>—Status</li> <li>• <b>Position 4</b>—Start Date/Time (MM/DD/YYYY HH:MM:SS)</li> <li>• <b>Position 5</b>—End Date/Time (MM/DD/YYYY HH:MM:SS)</li> <li>• <b>Position 6</b>—Number of successful records</li> <li>• <b>Position 7</b>—Number of failed records</li> <li>• <b>Position 8</b>—Number of malformed records</li> <li>• <b>Position 9</b>—Currently unused</li> </ul> <p>The information is delimited using the delimiter specified in the <i>-d</i> argument. For example:</p> <pre>MySimpleJob 4 succeeded 04/09/2010 14:50:47 04/09/2010 14:50:47 100 0 0 </pre> |
| No       | <i>s port</i>         | The socket (port) on which the Spectrum™ Technology Platform server is running. The default is 8080.  |
| No       | <i>t timeout</i>      | Sets the timeout (in seconds) for synchronous mode. The default is 3600. The maximum is 2147483. This is a global, aggregate timeout and represents the maximum time to wait for all spawned jobs to complete.  |
| Yes      | <i>u username</i>     | The login name of the user.   |
| No       | <i>v</i>              | Return verbose output.  |
| No       | <i>w</i>              | Specifies to wait for jobs to complete in a synchronous mode.   |

### Using Both Command Line Arguments and a Property File

A combination of both command-line entry and property file entry is also valid. For example:

```
java -jar jobexecutor.jar -f /dcb/job.properties -j job1
```

In this case command line arguments take precedence over arguments specified in the properties file. In the above example, the job `job1` would take precedence over a job specified in the properties file.

### Related reference

[Running A Job from the Command Line](#) on page 172

### Overriding Input and Output Files Using a Job Property File

You can override the input file specified in the dataflow's source stage (such as Read from File), as well as the output file specified in the dataflow's sink stage (such as Write to File) in a job executor property file. To do this, specify the following in the property file:

```
StageName\ :file=Protocol:FileName
```

Where:

#### **StageName**

The stage label shown under the stage's icon in the dataflow in Enterprise Designer. Use a backslash before any spaces, colons, or equal signs in the stage name. For example, if the stage is labeled "Read from File" you would specify `Read\ from\ File` for the stage name.

```
Embedded\ Dataflow\ 1.Embedded\ Dataflow\ 2.Write\ to\ File
```

To specify a stage within an embedded dataflow or a subflow, preface the stage name with the name of the embedded dataflow or subflow, followed by a period then the stage name:

```
EmbeddedOrSubflowName.StageName
```

For example, to specify a stage named Write to File in a subflow named Subflow1, you would specify:

```
Subflow1.Write\ to\ File
```

To specify a stage in an embedded dataflow that is within another embedded dataflow, add the parent dataflow, separating each with a period. For example, if Embedded Dataflow 2 is inside Embedded Dataflow 1, and you want to specify the Write to File stage in Embedded Dataflow 2, you would specify this:

```
Embedded\ Dataflow\ 1.Embedded\ Dataflow\ 2.Write\ to\ File
```

**Note:** You must include `:file` after the stage name. For example, `Read\ from\ File:file`. This is different from the syntax used to override files at the command line where `:file` is not specified after the stage name.

#### **Protocol**

A communication protocol. One of the following:

**file** Use the file protocol if the file is on the same machine as the Spectrum™ Technology Platform server. For example, on Windows specify:

```
"file:C:/myfile.txt"
```

On Unix or Linux specify:

```
"file:/testfiles/myfile.txt"
```

**esclient** Use the esclient protocol if the file is on the computer where you are executing the job if it is a different computer from the one running the Spectrum™ Technology Platform server. Use the following format:

```
esclient:ComputerName/path to file
```

For example,

```
esclient:mycomputer/testfiles/myfile.txt
```

**Note:** If you are executing the job on the server itself, you can use either the file or esclient protocol, but are likely to have better performance using the file protocol.

If the host name of the Spectrum™ Technology Platform server cannot be resolved, you may get the error "Error occurred accessing file". To resolve this issue, open this file on the server:

```
SpectrumLocation/server/app/conf/spectrum-container.properties.
```

Set the `spectrum.runtime.hostname` property to the IP address of the server.

**esfile** Use the esfile protocol if the file is on a file server. The file server must be defined in Management Console as a resource. Use the following format:

```
esfile://file server/path to file
```

For example,

```
esfile://myserver/testfiles/myfile.txt
```

Where myserver is an FTP file server resource defined in Management Console.

**webhdfs** Use the webhdfs protocol if the file is on a Hadoop Distributed File Server. The HDFS server must be defined in Management Console as a resource. Use the following format:

```
webhdfs://file server/path to file
```

For example,

```
webhdfs://myserver/testfiles/myfile.txt
```

Where myserver is an HDFS file server resource defined in Management Console.

**Example**

The last two lines of the following property file specify the files for the Read from File stage and the Write to File stage.

```
j=testJob
h=myspectrumserver.example.com
s=8080
u=david1234
p=mypassword1234
Read\ from\ File\:file=file:C:/myfile_input.txt
Write\ to\ File\:file=file:C:/myfile_output.txt
```

**Overriding File Format Using a Job Property File**

You can use a property file to override the file layout (or schema) of the file specified in the dataflow's Read from File stage and Write to File stage. To do this, specify the following in the property file:

```
StageName\ :schema=Protocol:SchemaFile
```

Where:

**StageName**

The stage label shown under the stage's icon in the dataflow in Enterprise Designer. Use a backslash before any spaces, colons, or equal signs in the stage name. For example, if the stage is labeled "Read from File" you would specify `Read\ from\ File` for the stage name.

```
Embedded\ Dataflow\ 1.Embedded\ Dataflow\ 2.Write\ to\ File
```

To specify a stage within an embedded dataflow or a subflow, preface the stage name with the name of the embedded dataflow or subflow, followed by a period then the stage name:

```
EmbeddedOrSubflowName.StageName
```

For example, to specify a stage named Write to File in a subflow named Subflow1, you would specify:

```
Subflow1.Write\ to\ File
```

To specify a stage in an embedded dataflow that is within another embedded dataflow, add the parent dataflow, separating each with a period. For example, if Embedded Dataflow 2 is inside Embedded Dataflow 1, and you want to specify the Write to File stage in Embedded Dataflow 2, you would specify this:

```
Embedded\ Dataflow\ 1.Embedded\ Dataflow\ 2.Write\ to\ File
```

**Note:** You must include `:file` after the stage name. For example, `Read\ from\ File:file`. This is different from the syntax used to override

files at the command line where `:file` is not specified after the stage name.

### Protocol

A communication protocol. One of the following:

**file** Use the file protocol if the file is on the same machine as the Spectrum™ Technology Platform server. For example, on Windows specify:

```
"file:C:/myfile.txt"
```

On Unix or Linux specify:

```
"file:/testfiles/myfile.txt"
```

**esclient** Use the esclient protocol if the file is on the computer where you are executing the job if it is a different computer from the one running the Spectrum™ Technology Platform server. Use the following format:

```
esclient:ComputerName/path to file
```

For example,

```
esclient:mycomputer/testfiles/myfile.txt
```

**Note:** If you are executing the job on the server itself, you can use either the file or esclient protocol, but are likely to have better performance using the file protocol.

If the host name of the Spectrum™ Technology Platform server cannot be resolved, you may get the error "Error occurred accessing file". To resolve this issue, open this file on the server:

```
SpectrumLocation/server/app/conf/spectrum-container.properties.  
Set the spectrum.runtime.hostname property to the IP address of the server.
```

**esfile** Use the esfile protocol if the file is on a file server. The file server must be defined in Management Console as a resource. Use the following format:

```
esfile://file server/path to file
```

For example,

```
esfile://myserver/testfiles/myfile.txt
```

Where myserver is an FTP file server resource defined in Management Console.

**webhdfs** Use the webhdfs protocol if the file is on a Hadoop Distributed File Server. The HDFS server must be defined in Management Console as a resource. Use the following format:

```
webhdfs://file server/path to file
```

For example,

```
webhdfs://myserver/testfiles/myfile.txt
```



Where myserver is an HDFS file server resource defined in Management Console.

### SchemaFile

The full path to the file that defines the layout you want to use.

**Note:** You must use forward slashes (/) in file paths, not backslashes.

To create a schema file, define the layout you want in Read from File or Write to File, then click the **Export** button to create an XML file that defines the layout.

**Note:** You cannot override a field's data type in a schema file when using job executor. The value in the <Type> element, which is a child of the <FieldSchema> element, must match the field's type specified in the dataflow's Read from File or Write to File stage.

#### Example

In the following example properties file, the last line overrides the file layout defined in the Read from File stage with the layout defined in the file `inputSchema.xml`. A backslash is used before the spaces in the stage's name.

```
j=testJob
h=myspectrumserver.example.com
s=8080
u=david1234
p=mypassword1234
Read\ from\ File\ :file=esclient:c:/MyData/testInput.txt
Read\ from\ File\ :schema=esclient:c:/MyData/inputSchema.xml
```

## Running a Process Flow from the Command Line

To run a process flow from the command line, use the Process Flow Executor. You can install the Process Flow Executor from the Spectrum™ Technology Platform Welcome page (for example, <http://myserver:8080>).

**Note:** You can also use the Administration Utility to execute process flows from the command line.

### Usage

```
java -jar pflowexecutor.jar -r ProcessFlowName -u UserID -p Password [Optional Arguments]
```

| Required | Argument                        | Description  |
|----------|---------------------------------|--|
| No       | -?                              | Prints usage information.  |
| No       | -d<br><i>DelimiterCharacter</i> | Sets a delimiter to use to separate the status information displayed in the command line when you execute the command. The default is " ". For example, using the default character, the following would be displayed at the command line when you execute a processflow named "MyProcessflow":<br><br>MyProcessflow 1 Succeeded |
| No       | -e                              | Use an HTTPS connection for communication with the Spectrum™ Technology Platform server.<br><br><b>Note:</b> If you specify any file overrides this argument must not be the last argument specified.  |
| No       | -f <i>PropertyFile</i>          | Specifies a path to a property file. For more information on property files, see <a href="#">Using a Process Flow Property File</a> on page 187.   |
| No       | -h <i>HostName</i>              | Specifies the name or IP address of the Spectrum™ Technology Platform server.  |
| No       | -i <i>PollInterval</i>          | Specifies how often to check for completed jobs, in seconds. The default is "5".   |
| Yes      | -p <i>Password</i>              | The password of the user. Required.  |
| Yes      | -r<br><i>ProcessFlowNames</i>   | A comma-separated list of process flows to run. Required.<br><br><b>Note:</b> If you specify any file overrides this argument must not be the last argument specified.   |
| No       | -s <i>Port</i>                  | The socket (port) on which the Spectrum™ Technology Platform server is running. The default is 8080.   |
| No       | -t <i>Timeout</i>               | This option is deprecated and will be ignored.   |
| Yes      | -u <i>UserName</i>              | The login name of the user. Required.  |
| No       | -v <i>Verbose</i>               | Return verbose output where <i>Verbose</i> is one of the following:<br><br><b>true</b> Return verbose output.<br><b>false</b> Do not return verbose output.<br><br><b>Note:</b> If you specify any file overrides this argument must not be the last argument specified.   |
| No       | -w<br><i>WaitToComplete</i>     | This option is deprecated and will be ignored.   |

| Required Argument                  | Description  |
|------------------------------------|--|
| No <code>StageName=FileName</code> | Overrides the input or output file specified in the job. For more information, see <a href="#">Overriding Process Flow File Locations</a> on page 188. |

### Examples

This is a basic command-line entry, with a process flow name and user ID, and password:

```
java -jar pflowexecutor.jar -r MyFlow1 -u Bob1234 -p
"mypassword1"
```

This example shows the same information as above but with additional arguments:

```
java -jar pflowexecutor.jar -r Flow1 -u Bob1234 -p
"mypassword1" -h spectrum.example.com -s 8888 -w -d "%" -i
1 -t 9999
```

The following example shows command line invocation and output.

```
D:\spectrum\pflow-executor>java -jar pflowexecutor.jar -u
Bob1234 -p "mypassword1" -r
validateAddressFlow1 -h spectrum.example.com -s 8888 -w -d
"%" -i
1 -t 9999
validateAddressJob1%111%succeeded
```

In this example, the process flow named `validateAddressFlow1` ran (with identifier 111). No errors occurred. Other possible results include "failed" or "running."

### Using a Process Flow Property File

A property file contains arguments that you can reuse by specifying the path to the property file with the `-f` argument in the process flow executor. The property file must contain, at minimum, the process flow (`r`), user ID (`u`), and password (`p`).

1. Open a text editor.
2. Specify one argument on each line as shown in the following example. See [Running a Process Flow from the Command Line](#) on page 185 for a list of arguments.

**Note:** You cannot use a property file to override input or output files. Overriding input and output files can only be done using command line arguments.

```
d=%
h=myserver.mydomain.com
i=30
```

```
u=user
p=password
r=MyFlow1
s=8888
```

3. Save the file with a file extension of .properties (for example, "example.properties").
4. When you run the process flow executor, specify the path to the property file using the `-f` argument. A combination of both command-line entry and property file entry is also valid. Command line arguments take precedence over arguments specified in the properties file.

```
java -jar pflowexecutor.jar -f /dcbg/flow.properties -r MyFlow2
```

In the above example, the process flow MyFlow2 would take precedence over a process flow specified in the properties file.

### Overriding Process Flow File Locations

When you execute a process flow using the Process Flow Executor command line tool, you can specify that the process flow should use different input and output files than those specified in the job referenced by the process flow. To do this, specify the Read from File or Write from File stage names along with the input or output file as the last arguments like this:

```
"<jobname>|<stagename>"="<filename>"
```

Where:

#### **JobName**

The name of a job referenced in the process flow.

#### **StageName**

The name of a Read from File or Write to File stage in the job as shown in the stage label under the stage's icon in the dataflow. For example, if the input stage is labeled "Read From File" you would specify:

```
"Job1|Read From File"="file:C:/inputfile.txt"
```

If the input stage is labeled "Illinois Customers" you would specify:

```
"Job1|Illinois Customers"="file:C:/inputfile.txt"
```

#### **File**

The protocol and full path to the file. You must use forward slashes (/) in file paths, not backslashes. The protocol must be one of the following:

##### **file:**

If the file is on the same machine as the Spectrum™ Technology Platform server, start the path with the "file:" protocol. For example, on Windows specify

```
file:C:/myfile.txt
```

```
file:/testfiles/myfile.txt.
```

**Note:** If the client and server are running on the same machine, you can use either the "file:" or "esclient:" protocol, but are likely to have get better performance using the "file:" protocol.

**esclient:**

If the file is on the same machine as Process Flow Executor, start the path with the "esclient:" protocol. For example, on Windows specify `esclient:C:/myfile.txt` and on Unix or Linux specify `esclient:/testfiles/myfile.txt`.

**Note:** If the machine running process flow executor cannot resolve the host name of the Spectrum™ Technology Platform server, you may get an error "Error occurred accessing file". To resolve this issue, open this file on the server:

*SpectrumLocation/server/app/conf/spectrum-container.properties.*

Set the `spectrum.runtime.hostname` property to the IP address of the server.

**ftp:**

To use a file server defined in the Management Console, use the following format: `ftp:NameOfFileServer/PathToFile`. For example, `ftp://FS/testfiles/myfile.txt` where FS is a file server resource defined in Management Console.

For example:

```
java -jar pflowexecutor.jar -r Flow1 -u Bob1234 -p "mypassword1" -h
spectrum.example.com -s 8888 -w -d "%" -i 1 -t 9999 "Job1|Read from
File"="file:C:/myfile_input.txt" "Job1|Write to
File"="file:C:/myfile_output.txt"
```

# 7 - Performance

## In this section

---

|                              |     |
|------------------------------|-----|
| Performance Tuning Checklist | 191 |
| Monitoring Performance       | 209 |

## Performance Tuning Checklist

This checklist describes the approach we recommend for getting optimal performance from your Spectrum™ Technology Platform environment. The techniques are listed in order from those having the most significant impact on performance to those having the least.

| Performance Setting                            | Description  | Additional Information  |
|--|--|---|
| Dataflow design                                | The most significant impact on performance is the design of the dataflow. There are several best practices you can follow to ensure that your dataflows have good performance.   | <a href="#">Design Guidelines for Optimal Performance</a> on page 192 |
| Dataflow pool size                             | Dataflow pool size controls how many instances of a service dataflow can be running at a time. The default pool size for a dataflow is eight.  | <a href="#">Dataflow Pool Size</a> on page 193                        |
| Database pool size and stage runtime instances | Database pool size and stage runtime instances control the ability of the system to handle multiple requests concurrently. They need to be adjusted in tandem.   | <a href="#">Database Pool Size and Runtime Instances</a> on page 126  |
| Sort performance                               | Sorting large data sets can be one of the most time-consuming operations performed during batch processing. Sort performance options control memory and disk utilization, allowing you to take full advantage of the available memory and disk capacity. | <a href="#">Setting Default Sort Performance Options</a> on page 197  |
| Remote component options                       | These settings control memory usage for certain stages such as address validation and geocoding stages.  | <a href="#">Configuring Remote Component Options</a> on page 198      |
| Individual stages                              | There are some actions you can take to optimize specific types of processing. Review the best practices for performance to make sure that you have configured your dataflow to achieve optimal performance.  | <a href="#">Optimizing Individual Stages</a> on page 199              |
| JVM settings                                   | There are JVM options that will yield performance gains with certain hardware.   | <a href="#">JVM Performance Tuning</a> on page 204                    |

| Performance Setting                        | Description   | Additional Information  |
|--|---|---|
| Micro-batch processing                     | Micro-batch processing is a technique where you include more than one record in a single service request.   | <a href="#">Micro-Batch Processing</a> on page 204                    |
| Heap size configuration for elastic search | Increase the Elasticsearch heap size if you are performing memory-intensive operations with the profiling tool in Metadata Insights or with any of the matching stages in a flow. | <a href="#">Heap Size Configuration for Elasticsearch</a> on page 199 |

### Related tasks

[Monitoring Performance with the Administration Utility](#) on page 209

### Related reference

[performancemonitor enabled get](#) on page 297

[performancemonitor enabled set](#) on page 298

## Design Guidelines for Optimal Performance

Carefully designing your dataflows to optimize performance is the most important thing you can do to achieve good performance on Spectrum™ Technology Platform. These guidelines describe techniques you can use to optimize dataflow performance.

### *Minimize the Number of Stages*

Spectrum™ Technology Platform achieves high performance through parallel processing. Each stage in a flow runs asynchronously in its own thread. However, it is possible to overthread the processors when executing certain types of dataflows. When this happens, the system spends as much or more time managing threads as doing "real work". We have seen dataflows with as many as 130 individual stages that perform very poorly on smaller servers with one or two processors.

So the first consideration in designing dataflows that perform well is to use as many stages as needed, but no more. Some examples of using more stages than needed are:

- Using multiple conditional routers where one would suffice
- Defining multiple transformer stages instead of combining the transforms in a single stage

Fortunately it is usually possible to redesign these dataflows to remove redundant or unneeded stages and improve performance.

For complex flows, consider using embedded flows to reduce clutter on the canvas and make it easier to view and navigate the flow. Using embedded flows does not have a performance benefit at runtime, but it does make it easier to work with flows in Enterprise Designer.



### *Reduce Record Length*

Since data is being passed between concurrently executing stages, another consideration is the length of the input records. Generally input with a longer record length will take longer to process than input with a shorter record length, simply because there is more data to read, write, and sort. Dataflows with multiple sort operations will particularly benefit from a reduced record length. In the case of very large record lengths it can be faster to remove the unnecessary fields from the input prior to running the Spectrum™ Technology Platform job then append them back to the resulting output file.

### *Use Sorting Appropriately*

Another consideration is to minimize sort operations. Sorting is often more time consuming than other operations, and can become problematic as the number and size of input records increases. However, many Spectrum™ Technology Platform stages either require or prefer sorted input data. The Universal Addressing Module and Enterprise Geocoding Module, for example, perform optimally when the input is sorted by country and postal code. Stages such as Intraflow Match and Interflow Match require that the input be sorted by the "group by" field. In some cases you can use an external sort application to presort the input data and this can be faster than sorting within the Spectrum™ Technology Platform dataflow.

## Dataflow Pool Size

Dataflow pool size controls how many instances of each service dataflow can be running at a time. You can increase the pool size to improve performance to a point, but increased pool size may result in reduced performance if the server does not have the processor or memory resources available to handle several instances of each service dataflow running concurrently. If processor and memory resources are being used to their limit, you may find that reducing dataflow pool size, which limits the number of concurrent instances of each service dataflow, may provide more acceptable performance overall.

When finding the right pool size for your system, keep in mind that the dataflow pool size limits the number of instances of each service dataflow, not the total number of concurrent service dataflows. For example, with the default setting of 8, each service dataflow is allowed to have eight instances running at a time. If there are two service dataflows each utilizing the maximum of 8 concurrent instances, there would be 16 total instances of service dataflows running concurrently on your system.

**Note:** Dataflow pool size affects the performance of services only, not jobs.

To configure the dataflow pool size:

1. Enable the JMX Console performance monitor. For more information, see [Monitoring Performance with the JMX Console](#) on page 212.
2. Send multiple concurrent requests to the service through web services or the API.

3. In the JMX Console, note the time shown in **ServiceRuntimeManager.borrow.DataflowName**. This shows the amount of time that a service request is waiting for space in the pool before running. The time is shown in milliseconds.
4. Open this file:  

```
SpectrumLocation\server\app\conf\dataflowpool-pool-sizes.properties
```
5. Increase or decrease the pool size based on what you think will result in improved performance. The default pool size is eight.
  - If the wait time shown in **ServiceRuntimeManager.borrow.DataflowName** was high, consider increasing the dataflow pool size to reduce the amount of time that service requests are waiting run. Or, you may want a high wait time in order to throttle requests and not overload server resources.
  - If the wait time shown in **ServiceRuntimeManager.borrow.DataflowName** was not high and you are experiencing slow performance, it may indicate that your server does not have the processor or memory resources to handle many concurrent service dataflows. Consider reducing the dataflow pool size to throttle requests.
6. Save and close the file.
7. Send multiple concurrent requests again, and observe the wait time shown in **ServiceRuntimeManager.borrow.DataflowName** in the JMX Console.
8. Experiment with additional pool sizes until you find the optimal setting.

## Database Pool Size and Runtime Instances

In most Spectrum™ Technology Platform environments there are multiple flows running at the same time, whether they are batch jobs or services responding to web service or API requests. To optimize concurrent processing, you can use the database pool size setting, which limits the number of concurrent requests a Spectrum database handles, and runtime instances, which controls the number of instances of a flow stage that run concurrently. These two settings should be tuned together to achieve optimal performance.

### Database Pool Size

Spectrum databases contain reference data used by certain stages, such as postal data used to validate addresses, or geocoding data used to geocode addresses. These databases can be configured to accept multiple concurrent requests from the dataflow stages or services that use them, thereby improving the performance of the dataflows or service requests. The database pool size sets the maximum number of concurrent requests that a Spectrum database will process. By default, Spectrum databases have a pool size of 4, meaning the database can process four requests simultaneously.

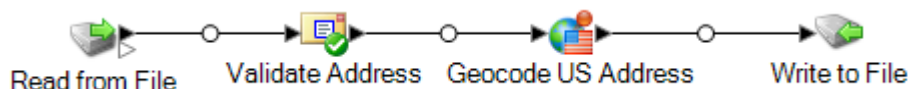
The optimal pool size varies by module. You will generally see the best results by setting the pool size between one-half to twice the number of CPUs on the server, with the optimal pool size for

most modules being the same as the number of CPUs. For example, if your server has four CPUs you may want to experiment with a pool size between 2 (one-half the number of CPUs) and 8 (twice the number of CPUs) with the optimal size possibly being 4 (the number of CPUs).

When modifying the pool size you must also consider the number of runtime instances specified in the dataflow for the stages accessing the database. Consider for example a dataflow that has a Geocode US Address stage that is configured to use one runtime instance. If you set the pool size for the US geocoding database to four, you will not see a performance improvement because there would be only one runtime instance and therefore there would only be one request at a time to the database. However, if you were to increase the number of runtime instances of Geocode US Address to four, you might then see an improvement in performance since there would be four instances of Geocode US Address accessing the database simultaneously, therefore using the full pool.

### Runtime Instances

Each stage in a dataflow operates asynchronously in its own thread and is independent of any other stage. This provides for parallel processing of stages in a dataflow, allowing you to utilize more than one runtime instance for a stage. This is useful in dataflows where some stages process data faster than others. This can lead to an unbalanced distribution of work among the threads. For example, consider a dataflow consisting of the following stages:



Depending on the configuration of the stages, it may be that the Validate Address stage processes records faster than the Geocode US Address stage. If this is the case, at some point during the execution of the dataflow all the records will have been processed by Validate Address, but Geocode US Address will still have records to process. In order to improve performance of this dataflow, it is necessary to improve the performance of the slowest stage - in this case Geocode US Address. One way to do that is to specify multiple runtime instances of the stage. Setting the number of runtime instances to two, for example, means that there will be two instances of that stage, each running in its own thread, available to process records.


As a general rule, the number of runtime instances should be at least equal to the number of instances of the remote component. See the *Administration Guide* for information about remote components. While specifying multiple runtime instances can help improve performance, setting this value too high can strain your system resources, resulting in decreased performance.

**Note:** Using multiple runtime instances only improves performance when running jobs or when running service requests with more than one record.

### Tuning Procedure

Finding the right settings for database pool size and runtime instances is a matter of experimenting with different settings to find the ones maximize available server resources without overloading resources and causing reduced performance.

**Note:** You should optimize the dataflow pool size before tuning the database pool size. For information about optimizing the dataflow pool size, see [Dataflow Pool Size](#) on page 193.

1. Begin by finding sample data to use as you test different settings. The sample dataset should be large enough that execution time is measurable and can be validated for consistency. The sample data should also be representative of the actual data you want to process. For example, if you are doing performance testing for geocoding, be sure that your test data has an equal number of records for all the countries you intend to geocode.
2. If you are testing a service or dataflow that requires the use of a database resource, such as postal databases or geocoding databases, make sure that you have the latest version of the database installed.
3. With sample data ready and the latest database resources installed, create a simple dataflow that reads data from a file, processes it with the stage you want to optimize, and writes to a file. For example, if you want to test performance settings for Validate Address, create a dataflow consisting of Read from File, Validate Address, and Write to File.
4. Set the database resource pool size to 1:
  - a. Open Management Console.
  - b. Go to **Resources > Spectrum Databases**.
  - c. Select the database resource you want to optimize and click the Modify button .
  - d. In the **Pool size** field, specify 1.
  - e. Click **OK**.
5. Set the stage's runtime instances to 1:
  - a. Open the dataflow in Enterprise Designer.
  - b. Double-click the stage that you want to set to use multiple runtime instances.
  - c. Click **Runtime**.
 

**Note:** Not all stages are capable of using multiple runtime instances. If there is no **Runtime** button at the bottom of the stage's window, the stage is not capable of using multiple runtime instances.
  - d. Select **Local** and specify 1.
  - e. Click **OK** to close the **Runtime Performance** window, then click **OK** to close the stage.
6. Calculate baseline performance by running the dataflow several times and recording the average values for:
  - Elapsed time
  - CPU utilization
  - Memory utilization

**Tip:** You can use the JMX console to monitor performance. For more information, see [Monitoring Performance with the JMX Console](#) on page 212.
7. Run multiple instances of the job concurrently, if this is a use case that must be supported. Record elapsed time, CPU utilization, and memory utilization for each scenario.

**Tip:** You can use a file monitor to run multiple instances of a job at once. For more information, see [Triggering a Flow with a Control File](#) on page 168.

8. Increment the database resource pool size and the stage runtime instances setting.
9. Restart the server.
10. Run the dataflow again, recording the elapsed time, CPU utilization, and memory utilization.
11. Continue to increment the database resource pool size and the stage runtime instances until you begin to see diminishing performance.
12. If you are testing geocoding performance, repeat this procedure using single country and multi-country input.

## Setting Default Sort Performance Options

Sorting large data sets can be one of the most time-consuming operations performed during batch processing, so setting appropriate sort performance options can have a significant impact on the performance of your jobs. Sort performance options control memory and disk utilization, allowing you to take full advantage of the available memory and disk capacity.

There are two places where you can configure sort performance settings. The first is in Management Console. This is where you specify default sort performance options for your system. The second place is in dataflow stages that perform a sort. The Sorter stage, Read from File, Write to File, and all other stages that include a sort operation, contain sort performance options. When you specify a sort performance option in a stage, you override the default sort performance options, choosing different settings to apply to individual stages in a dataflow.

This procedure describes how to set the default sort performance options.

1. Open Management Console.
2. Go to **Flows > Defaults**.
3. Use these settings to control sort performance:

|                               |  |
|-------------------------------|--|
| <b>In memory record limit</b> | Specifies the maximum number of data rows a sorter will hold in memory before it starts paging to disk. By default, a sort of 10,000 records or less will be done in memory and a sort of more than 10,000 records will be performed as a disk sort. The maximum limit is 100,000 records. Typically an in-memory sort is much faster than a disk sort, so this value should be set high enough so that most of the sorts will be in-memory sorts and only large sets will be written to disk. |
|-------------------------------|--|

**Note:** Be careful in environments where there are jobs running concurrently because increasing the **In memory record limit** setting increases the likelihood of running out of memory.

|  |   |
|--|---|
| <b>Maximum number of temporary files</b> | Specifies the maximum number of temporary files that may be used by a sort process. Using a larger number of temporary files can result in better performance. However, the optimal number is highly dependent on the |
|--|---|

configuration of the server running Spectrum™ Technology Platform. You should experiment with different settings, observing the effect on performance of using more or fewer temporary files. To calculate the approximate number of temporary files that may be needed, use this equation:

$$\frac{(NumberOfRecords \times 2)}{InMemoryRecordLimit} = NumberOfTempFiles$$

Note that the maximum number of temporary files cannot be more than 1,000.

### **Enable compression**

Specifies that temporary files are compressed when they are written to disk.

**Note:** The optimal sort performance settings depends on your server's hardware configuration. Nevertheless, the following equation generally produces good sort performance:

$$\frac{(InMemoryRecordLimit \times MaxNumberOfTempFiles \div 2)}{TotalNumberOfRecords} \geq$$

## Configuring Remote Component Options

A remote components is an underlying engine that performs a specific processing function, such as address validation, geocoding, or tax jurisdiction assignment. Some remote components can be configured to maximize performance. For example, a remote component might have options controlling how much reference data is cached in memory or how incoming data is matched to the reference data.

Each remote component is deployed into its own JVM. This means that JVM configuration can be done per remote component and independent of the server itself, allowing for flexibility of memory allocation and tuning of performance based on the characteristics of the remote component.

Remote component options affect all instances of the component as well as any stages that use that component. This is different from stage options, which can be modified at design time and at runtime.

### *Universal Addressing Module Component Configuration*

For U.S. address processing, there are several options controlling which reference data is cached in memory. These options are set in this configuration file:

`server/modules/clp/java.properties.`

- `DpvMemoryModel`: Controls which DPV files are in memory
- `LacsLinkMemoryModel`: Controls which LACS<sup>Link</sup> files are in memory

- SuiteLinkMemoryModel: Controls which Suite<sup>Link</sup> files are in memory

Refer to the `java.properties` configuration file for specifics on the different values for these options.

### *Enterprise Geocoding Module Component Configuration*

The Enterprise Geocoding Module has several options that can affect the performance of U.S. geocoding. The options are in this configuration file:

`server/modules/geostan/java.properties`. Of particular interest are:

- `egm.us.multimatch.max.records`: Specifies the maximum number of possible matches to return. A smaller number results in better performance, but at the expense of matches.
- `egm.us.multimatch.max.processing`: Specifies the number of searches to perform. A smaller number results in better performance, but at the expense of matches.
- `FileMemoryLimit`: Controls how much of the reference data is initially loaded into memory.

## Heap Size Configuration for Elasticsearch

Elasticsearch is an underlying search technology used when performing data profiling in Metadata Insights and when performing matching using the matching stages in Enterprise Designer. Consider increasing the Elasticsearch heap size in these situations:

- In Metadata Insights, you run multiple profiles concurrently or you run a profile having multiple tables
- You have flows that run multiple search index queries in parallel, each of which returns 1,000 or more candidates

To increase the Elasticsearch heap size, open this file in a text editor:

`SpectrumLocation\index\spectrum.vargs`

Increase the value in the `-Xmx` property. The default heap size is `-Xmx2048m`.

## Optimizing Individual Stages

### Optimizing Matching

Matching is typically one of the most time-consuming operations in any data quality implementation, making it important to ensure that matching is operating as efficiently as possible. There is always a balance between matching results and performance. If every record in a file is compared to every other record, you can be quite confident that all matches will be identified. However, this approach is unsustainable as the volume of data grows. For example, given an input file of 1 million records,

matching each record to every other record would require nearly 1 trillion comparisons to evaluate each match rule.

Given that most records in a file do not match, the general approach to solving this problem is to define a match key and only compare those records that have the same match key. Proper match key definition is the most critical variable affecting performance of the matching engine. To define a proper match key, you must understand how the matching engine processes records and the options that are available.

The default matching method performs an exhaustive comparison of the record in a match queue to identify the maximum number of matches. Because of this, it is often the most time consuming way to do matching. Under the default matching method, the first record in the match queue becomes the suspect record. The next record is compared, and if it matches it is written out as a duplicate. If it does not match, it is added as a suspect, and the next record is compared to the two active suspects. Consider the following match queue:

| Unique ID | Match Key |
|-----------|-----------|
| 1         | 123A      |
| 2         | 123A      |
| 3         | 123A      |
| 4         | 123A      |
| 5         | 123A      |
| 6         | 123A      |
| 7         | 123A      |
| 8         | 123A      |
| 9         | 123A      |
| 10        | 123A      |

First, record 2 would be compared to record 1. Assuming it does not match, record 2 would be added as a suspect. Then record 3 would be compared to records 1 and 2, and so on. If there are no matching records, the total number of comparisons would be 45. If some records match, the number of comparisons will be less. For a match queue of a given size  $N$ , the maximum number of comparisons will be  $N \times (N-1) \div 2$ . When the queue size is small this is not noticeable, but as the queue



size grows the impact is significant. For example, a queue size of 100 could result in 4,450 comparisons, and a queue size of 500 could result in 124,750 comparisons.

### *Defining an Appropriate Match Key*

To define an appropriate match key, consider the following:

- The most important thing to remember is most records do not match. Therefore you want to compare only records that are likely to match.
- Only records with the same match key will be compared.
- Performance is a key consideration:
  - The match key determines the size of the match queue.
  - For a given number of records, as the match queue size doubles, execution time doubles.
  - A "tight" match key results in better performance. A "tight" match key is one that is specific, containing more characters from possibly more fields.
  - A "loose" match key may result in more matches. A "loose" match key is one that is less specific, containing fewer characters from possibly fewer fields.

### *Finding a Balance Between Performance and Match Results*

To find a good balance between performance and results, consider the match rule and the density of the data.

- Consider the match rules:
  - Fields requiring an exact match could be included in the match key.
  - Build an appropriate key for the match rule. For example, for a phonetic match rule, a phonetic match key is probably appropriate.
  - A match key will often consist of parts of all the fields being matched.
  - Be aware of the effects of missing data.
- Consider the density of the data:
  - For example, in address matching, the match key would likely be tighter if all the records are in a single town instead of a national dataset.
  - Consider the largest match queue, not just the average. Review the Match Summary report to find the largest match queue.
- When using transactional match, the same considerations apply to the SELECT statement in Candidate Finder.

### *Express Match Key*

In a typical file, most of the duplicate records match either exactly or nearly exactly. Defining an express match key allows the matching engine to perform an initial comparison of the express match keys to determine that two records are duplicates. This can significantly improve performance by avoiding the need to evaluate all the field level match rules.

### Intraflow Match Methods

The default Intraflow Match match method compares all records having the same match key. For a match queue size of  $N$ , the default method performs anywhere from  $N-1$  to  $N \times (N-1)$  comparisons. If all records match, the number of comparisons is  $N-1$ . If no records match the number of comparisons is  $N \times (N-1)$ . Usually the number of comparisons is somewhere in the upper part of this range.

If performance is a priority, consider using the sliding window match method instead of the default method. The sliding window match method compares each record to the next  $W$  records (where  $W$  is the window size). For a given file size  $N$ , the sliding window method performs no more than  $N \times W$  comparisons. This can lead to better performance, but some matches may be missed.

### Optimizing Candidate Finder

Candidate Finder selects candidate records from a database for comparison by Transactional Match. Since transactional match compares the suspect record to all of the candidate records returned by Candidate Finder, the performance of Transactional Match is proportional to the number of comparisons.

However, there are things you can do to improve the performance of Candidate Finder. To maximize the performance of Candidate Finder, a database administrator, or developer with extensive knowledge of the database schema and indexes, should tune the SQL SELECT statement in Candidate Finder. One of the most common performance problems is a query that contains a JOIN that requires a full table scan. In this case, consider adding an index or using a UNION instead of a JOIN. As a general rule, SQL queries should be examined and optimized by qualified individuals.

### Optimizing Transforms

The Transformer stage provides a set of predefined operations that can be performed on the input data. Generally, these predefined transforms execute faster than custom transforms, since they are already compiled. However, when defining a large number of transforms, a custom transform will frequently execute faster. For example, to trim a number of fields, the following custom transform will typically execute faster than nine separate trim transforms.

```
data['AddressLine1'] = (data['AddressLine1'] != null) ?
data['AddressLine1'].trim() : null;
data['AddressLine2'] = (data['AddressLine2'] != null) ?
data['AddressLine2'].trim() : null;
data['AddressLine3'] = (data['AddressLine3'] != null) ?
data['AddressLine3'].trim() : null;
data['AddressLine4'] = (data['AddressLine4'] != null) ?
data['AddressLine4'].trim() : null;
data['City'] = (data['City'] != null) ? data['City'].trim() : null;
data['StateProvince'] = (data['StateProvince'] != null) ?
data['StateProvince'].trim() : null;
data['PostalCode'] = (data['PostalCode'] != null) ?
data['PostalCode'].trim() : null;
data['LastName'] = (data['LastName'] != null) ? data['LastName'].trim()
```

```

: null;
data['FirstName'] = (data['FirstName'] != null) ?
data['FirstName'].trim() : null;

```

## Optimizing Write to DB

By default the Write to DB stage commits after each row is inserted into the table. However, to improve performance enable the **Batch commit** option. When this option is enabled, a commit will be done after the specified number of records. Depending on the database this can significantly improve write performance.

When selecting a batch size, consider the following:

- **Data arrival rate to Write To DB stage:** If data is arriving at slower rate than the database can process then modifying batch size will not improve overall dataflow performance. For example, dataflows with address validation or geocoding may not benefit from an increased batch size.
- **Network traffic:** For slow networks, increasing batch size to a medium batch size (1,000 to 10,000) will result in better performance.
- **Database load and/or processing speed:** For databases with high processing power, increasing batch size will improve performance.
- **Multiple runtime instances:** If you use multiple runtime instances of the Write to DB stage, a large batch size will consume a lot of memory, so use a small or medium batch size (100 to 10,000).
- **Database roll backs:** Whenever a statement fails, the complete batch is rolled back. The larger the batch size, the longer it will take to perform the rollback.

## Optimizing Address Validation

Validate Address provides the best performance when the input records are sorted by postal code. This is because of the way the reference data is loaded in memory. Sorted input will sometimes perform several times faster than unsorted input. Since there will be some records that do not contain data in the postal code field, the following sort order is recommended:

1. Country (Only needed when processing records for multiple countries)
2. PostalCode
3. StateProvince
4. City

## Optimizing Geocoding

Geocoding stages provide the best performance when the input records are sorted by postal code. This is because of the way the reference data is loaded in memory. Sorted input will sometimes perform several times faster than unsorted input. Since there will be some records that do not contain data in the postal code field, the following sort order is recommended:

1. PostalCode
2. StateProvince

### 3. City

You can also optimize geocoding stages by experimenting with different match modes. The match mode controls how the geocoding stage determines if a geocoding result is a close match. Consider setting the match mode to the **Relaxed** setting and seeing if the results meet your requirements. The **Relaxed** mode will generally perform better than other match modes.

#### Optimizing Geocode US Address

The Geocode US Address stage has several options that affect performance. These options are in this file:

```
SpectrumLocation\server\modules\geostan\java.properties
```

|   |   |
|---|---|
| <b>egm.us.multimatch.max.records</b>    | Specifies the maximum number of matches to return. A smaller number results in better performance, but at the expense of matches. |
| <b>egm.us.multimatch.max.processing</b> | Specifies the number of searches to perform. A smaller number results in better performance, but at the expense of matches.       |
| <b>FileMemoryLimit</b>                  | Controls how much of the reference data is initially loaded into memory.  |

## JVM Performance Tuning

There are JVM options that will yield performance gains with certain hardware. These are advanced options and can cause unexpected behavior when used incorrectly (possibly even JVM crashes). We recommend that you contact Pitney Bowes Technical Support if you want to explore increasing performance by tuning the JVM.

- On servers with multiple CPUs, you can add the option `-XX:+UseParallelGC` to improve GC processing.
- These options may improve performance, although on some hardware they may cause JVM crashes.
  - `-Xmn512m`
  - `-XX:+AggressiveOpts`

## Micro-Batch Processing

Micro-batch processing is a technique where you include more than one record in a single service request. By including multiple records in a request instead of issuing separate requests for each record, you can significantly improve performance when processing a large collection of records

through a service. Spectrum™ Technology Platform supports micro-batch processing for REST and SOAP web services as well for the Client SDK.

### Micro-Batch Size

There is no limit to the number of records you can include in a request, but in general you will see the best performance when sending between 50 and 100 records in a micro-batch. We recommend that you test micro-batches of various sizes to determine the optimal micro-batch size for your environment. Keep in mind that in some cases you may get multiple records in the response for each input record. For example, if you are performing address validation and include 10 addresses in the micro-batch, and each address matches to two possible validated addresses, you would get 20 records in the response, not just 10.

Use caution when using both micro-batches and multiple threads for requests to Spectrum™ Technology Platform. Multiple threads can overwhelm the system if each thread's micro-batch size is too large.

The maximum number of records allowed in a request depends on the service's category.

| Service Category | Maximum Input Entities |
|------------------|------------------------|
| Addressing       | 25                     |
| Analytic         | 10                     |
| Data Quality     | 10                     |
| Geocoding        | 25                     |
| Identity Insight | 10                     |
| PSAP             | 25                     |
| Routing          | 1                      |
| Spatial          | 1                      |
| Tax Jurisdiction | 25                     |
| Telco            | 1                      |

Each micro-batch call must be synchronous. Send the request and wait for the response before sending another request. Asynchronous micro-batch calls can flood Spectrum™ Technology Platform

which may result in decreased performance. If you use asynchronous calls or micro-batches that are too large, your account may be disabled.

### *Using a Record ID*

You may find it helpful to assign an ID to each record in a micro-batch so that you can correlate the records in the request with the records returned in the response. Use user fields to do this. For information about user fields, see [The REST Interface](#). For information about user fields, see [The SOAP Interface](#).

### *Micro-Batch Processing in REST*

To perform micro-batch processing with a REST web service, include two or more records as XML or JSON in the body of the request and send the request using the POST method. For more information about sending a POST request to a Spectrum™ Technology Platform web service, see [JSON POST Request](#) and [XML POST Request](#).

For example, this request includes two records as XML in the body of the request:

```
POST
http://spectrum.example.com:8080/rest/ValidateAddressPOST/results.xml
HTTP/1.1
Accept-Encoding: gzip,deflate
Content-Type: application/xml
Authorization: Basic YWRtaW46YWRtaW4=
Content-Length: 533
Host: config813vm0:8080
Connection: Keep-Alive
User-Agent: Apache-HttpClient/4.1.1 (java 1.5)

<ValidateAddressPOSTRequest
xmlns:svc="http://www.pb.com/spectrum/services/ValidateAddressPOST">
  <svc:Input>
    <svc:Row>
      <svc:AddressLine1>3001 Summer</svc:AddressLine1>
      <svc:City>Stamford</svc:City>
      <svc:StateProvince>CT</svc:StateProvince>
    </svc:Row>
    <svc:Row>
      <svc:AddressLine1>33 west monroe</svc:AddressLine1>
      <svc:City>Chicago</svc:City>
      <svc:StateProvince>IL</svc:StateProvince>
    </svc:Row>
  </svc:Input>
</ValidateAddressPOSTRequest>
```

**Note:** Services do not have POST support enabled by default. To perform micro-batch processing with these services you must enable POST support. For more information, see [Adding POST Support to a REST Web Service](#).

### Micro-Batch Processing in SOAP

To perform micro-batch processing in a SOAP web service, include two or more records in the SOAP request. For example, this request contains two records:

```
POST http://spectrum.example.com:8080/soap/ValidateAddress HTTP/1.1
Accept-Encoding: gzip,deflate
Content-Type: text/xml;charset=UTF-8
SOAPAction: ""
Authorization: Basic YWRtaW46YWRtaW4=
Content-Length: 782
Host: config813vm0:8080
Connection: Keep-Alive
User-Agent: Apache-HttpClient/4.1.1 (java 1.5)

<soapenv:Envelope
xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/"
xmlns:val="http://www.pb.com/spectrum/services/ValidateAddress">
  <soapenv:Header/>
  <soapenv:Body>
    <val:ValidateAddressRequest>
      <val:input_port>
        <val:Address>
          <val:AddressLine1>1 N. State St.</val:AddressLine1>
          <val:City>Chicago</val:City>
          <val:StateProvince>IL</val:StateProvince>
        </val:Address>
        <val:Address>
          <val:AddressLine1>3001 summer</val:AddressLine1>
          <val:City>stamford</val:City>
          <val:StateProvince>ct</val:StateProvince>
        </val:Address>
      </val:input_port>
    </val:ValidateAddressRequest>
  </soapenv:Body>
</soapenv:Envelope>
```

### Micro-Batch Processing in the Client API

To perform micro-batch processing in an API request to a service, send multiple data rows in the request. For example, this .NET class sends two rows in the request:

```
using System;
using System.Collections.Generic;
using System.Diagnostics;
using System.Linq;
using System.Net;
using System.Text;
using ConsoleApplication1.ValidateAddress_Reference;

namespace Test
{
```

```

class Program
{
    static void Main(string[] args)
    {
        var validateClient = new ValidateAddress {Credentials = new
NetworkCredential("admin", "admin")};

        var address1 = new input_portAddress
        {
            AddressLine1 = "1825B Kramer Lane",
            AddressLine2 = "Suite 100",
            PostalCode = "78758",
            City = "Austin",
            StateProvince = "Texas"
        };

        var address2 = new input_portAddress
        {
            AddressLine1 = "100 Congress",
            PostalCode = "78701",
            City = "Austin",
            StateProvince = "Texas"
        };

        var addresses = new input_portAddress[2];
        addresses[0] = address1;
        addresses[1] = address2;

        var options = new options {OutputCasing = OutputCasing.M};
        output_portAddress[] results =
validateClient.CallValidateAddress(options, addresses);

        for (int i = 0; i < results.Length; i++)
        {
            System.Console.WriteLine("Record " + (i+1) + ":");
            System.Console.WriteLine("AddressLine1=" +
results[i].AddressLine1);
            System.Console.WriteLine("City=" + results[i].City);
            System.Console.WriteLine("StateProvince=" +
results[i].StateProvince);
            System.Console.WriteLine("PostalCode=" +
results[i].PostalCode + "\n");
        }

        System.Console.Write("Press any key to continue...");
        System.Console.ReadKey();
    }
}

```

**Tip:** Do not disconnect between requests. Disconnecting and connecting can reduce performance.



## Monitoring Performance

Monitoring the performance of your dataflows enables you to tune performance by identifying performance bottlenecks. There are two ways you can monitor performance: the Administration Utility and the JMX Console.

The Administration Utility is a command line tool that provides access to many administrative functions, including a performance monitor. When enabled, the performance monitor writes performance data to a log file each time a dataflow is executed. The log file contains performance data for each stage in the dataflow.

The JMX console is browser-based tool. It has a performance monitoring tool that records performance statistics for each stage in a dataflow.

### Monitoring Performance with the Administration Utility

The Administration Utility is a command line tool that provides access to many administrative functions, including a performance monitor. When enabled, the performance monitor writes performance data to a log file each time a dataflow is executed. The log file contains performance data for each stage in the dataflow.

1. Open the Administration Utility.
2. Type the following command:

```
performancemonitor enabled set --e True --d DataflowName
```

Where *DataflowName* is the name of the job or service you want to monitor.

Performance monitoring is now enabled for the dataflow you specified. When the dataflow runs, performance information will be written to the performance log.

#### Related concepts

[Performance Tuning Checklist](#) on page 191

#### Related reference

[performancemonitor enabled get](#) on page 297

[performancemonitor enabled set](#) on page 298

#### The Performance Log

The performance log contains details about how long it takes for a job or service to run. It includes overall performance information for the job or service as well as performance information for each stage in the job or service dataflow. You can use this information to identify bottlenecks in your dataflow by looking at the execution time and processing time for each stage. A large difference

between execution time and processing time means that the stage is spending time waiting for data from upstream stages. This may indicate that an upstream stage is a bottleneck in the dataflow. Note that for sinks, a large difference between execution time and processing time does not necessarily indicate a performance problem because sinks typically have to wait for the first records from the rest of the dataflow.

To enable performance monitoring for a job or service, use the `performancemonitor enabled set` command in the Administration Utility.

The performance log is located on your Spectrum™ Technology Platform server in the following location:

```
SpectrumLocation\server\app\Repository\logs\performance.log
```

The performance log contains one row for each run of a monitored job or service. It is a rolling log that consists of a maximum of five files. Each file is limited to 10 MB in size. Once this limit is reached, the oldest performance data is deleted when new performance data is logged.

Each entry in the performance log contains the following information.

**Note:** For ease of reading, line breaks and indentation are shown below. In the actual log, the entry is on one line.

```
Date Time [performance]
{
  "username": "UserName",
  "dataflowId": "DataflowName",
  "runMode": "BatchOrRealTime",
  "remote": TrueOrFalse,
  "elapsedTime": Nanoseconds,
  "stageInfo": [
    {
      "stageName": "Name",
      "stageLabel": "Label",
      "options": {
        OptionsList
      },
      "recordsRead": Count,
      "recordsWritten": Count,
      "executionTime": Nanoseconds,
      "processingTime": Nanoseconds
    }
  ]
}
```

Where:

**username**

The user who executed the job or service.

**dataflowID**

The name of the service or job as defined in Enterprise Designer.

**runMode**

Indicates whether the log entry is for a job or a service. One of the following:

- |                 |                                 |
|-----------------|---------------------------------|
| <b>Batch</b>    | The log entry is for a job.     |
| <b>RealTime</b> | The log entry is for a service. |

**remote**

Indicates whether the job or service was executed in the local server or on a remote server. For jobs that contain one or more stages that run on a remote server, the performance log on the server running the job will indicate that the job is not remote (a value of "false") while the performance log on the remote server that executes one or more of the job's stages will indicate that the job is remote (a value of "true"). One of the following:

- |              |  |
|--------------|--|
| <b>true</b>  | The job or service was executed on a remote server.  |
| <b>false</b> | The job or service was executed on the local server. |

**elapsedTime**

The time in nanoseconds that it took to run the job or service request.

**stageInfo**

Lists execution information for each stage in the dataflow. The following information is listed for each stage:

**stageName**

The permanent name of the stage.

**stageLabel**

The user-defined name of the stage. The stage label is shown on the canvas in Enterprise Designer.

**options**

If any options were specified at runtime, those options and their settings are listed here.

**recordsRead**

The total number of records that passed into the stage through all the stage's input ports.

**recordsWritten**

The total number of records that the stage wrote to all its output ports.

**executiontime**

The elapsed time from when the stage processed its first record to when it processed its last record. This includes the time the stage was idle while waiting for data from other stages in the dataflow.

**processingtime**

The amount of time the stage spent actively processing records, not including the time it was idle while waiting for other stages in the dataflow.

## Monitoring Performance with the JMX Console

The JMX console is browser-based tool. It has a performance monitoring tool that records performance statistics for each stage in a dataflow.

1. Open a web browser and go to `http://server:port/jmx-console`

Where:

*server* is the IP address or hostname of your Spectrum™ Technology Platform server.

*port* is the HTTP port used by Spectrum™ Technology Platform. The default is 8080.

2. Log in using the admin account.
3. Under **Domain: com.pb.spectrum.platform.performance**, click **com.pb.spectrum.platform.performance:service=PerformanceMonitorManager**.
4. Click the **Invoke** button next to **enable**.
5. Click **Return to MBean View** to go back to the PerformanceMonitorManager screen.

Performance monitoring is now enabled. When a dataflow runs, the performance statistics will display at the top of the PerformanceMonitorManager screen. Note the following:

- You must refresh the screen to see updates.
- To reset the counters, click the **Invoke** button next to **reset**.
- If you stop the Spectrum™ Technology Platform server, performance monitoring will be turned off. You will have to turn it back on when you start the server again.

### JMX Performance Monitor Statistics

The JMX Console's Performance Monitor Manager displays statistics about the performance of different parts of a dataflow execution, including the overall execution time, throughput, and execution time of individual stages. The statistics are reported in a semicolon-delimited format:

JMX Console

Checking authority: superuser

MBean: `com.pb.spectrum.platform.performance:service=PerformanceMonitorManager` All MBeans  
 Description: Performance Monitor Services

| Attributes |   |                                       |                  |
|------------|---|---------------------------------------|------------------|
| Name       | Value   | Description                           | Type             |
| Enabled    | true  | Is Performance Monitoring enabled     | boolean          |
| Report     | Monitor:Hits;Avg Delta;Min;Max;Total<br>JobManagerService:1,443,9629;443,9629;443,9629;443,9629<br>JobRuntimeManager:1,440,5076;440,5076;440,5076;440,5076<br>JobProcess;1,1,166,7749;1,166,7749;1,166,7749;1,166,7749<br>DataflowProcess;1,1,113,7199;1,113,7199;1,113,7199;1,113,7199<br>Dataflow initialize:1,6,4529;6,4529;6,4529;6,4529<br>Stage[Write to Null]:1,26,5067;26,5067;26,5067;26,5067<br>Stage[Read from File]:1,944,1421;944,1421;944,1421;944,1421<br>Dataflow:1,1,042,9381;1,042,9381;1,042,9381;1,042,9381 | Get the Performance Monitoring Report | java.lang.String |

| Operations |             |                                   |
|------------|-------------|-----------------------------------|
| Name       | Return type | Description                       |
| isEnabled  | boolean     | Is Performance Monitoring enabled |
| reset      | void        | Reset Performance Monitor         |
| enable     | void        | Enable Performance Monitoring     |
| disable    | void        | Disable Performance Monitoring    |

**Tip:** Put the data into a spreadsheet for easier viewing.

The first row is a header record consisting of the following columns:

|                |   |
|----------------|---|
| <b>Monitor</b> | The item whose performance is being measured.   |
| <b>Hits</b>    | The number of times the item was executed.  |
| <b>Avg</b>     | The average amount of time that the item spent processing a request, in milliseconds. |
| <b>Delta</b>   | This statistic is not used.   |
| <b>Min</b>     | The shortest amount time that the item spent processing a request, in milliseconds.   |
| <b>Max</b>     | The longest amount of time that the item spent processing a request, in milliseconds. |
| <b>Total</b>   | The total time spent processing, in milliseconds.                                     |

The most important items to review are the following.

|   |   |
|---|---|
| <b>Dataflow</b>                                   | The overall data flow response times and throughput.  |
| <b>ServiceRuntimeManager.borrow.DataflowName</b>  | Response time for handling concurrent requests, in milliseconds. This response time may be improved by modifying the dataflow's pool size.  |
| <b>RemoteComponent.borrow.RemoteComponentName</b> | Database resource performance directly impacted by the database resource pool size. This response time may go up as you increase your dataflow pool size. If this number increases to something significant such as ½ a millisecond increase the database resource pool size.   |
| <b>Stage[StageName]</b>                           | The response times for each stage. This can help you identify bottleneck stages that are taking significantly longer to process than other stages. Bottleneck stages can potentially be addressed by changing how the functionality in the existing stage is implemented and/or by increasing the runtime instances for that stage. |

# 8 - Monitoring

## In this section

---

|   |     |
|---|-----|
| Email Notification                            | 215 |
| Audit Log                                     | 218 |
| System Log                                    | 220 |
| Logging the Record that Caused a Flow to Fail | 222 |
| Transaction Limit Warnings                    | 222 |
| Viewing Version Information                   | 223 |
| Viewing and Exporting License Information     | 223 |

# Email Notification

## Configuring a Mail Server

Spectrum™ Technology Platform can send email alerts to notify you of system events. Email notifications can be sent as a result of conditions within dataflows and process flows, and when licenses, databases, and other items are about to expire.

Spectrum™ Technology Platform does not have a built-in mail server, so in order to enable email notification you must configure it to use an external SMTP server.

1. Open Management Console.
2. Go to **System > Mail Server**.
3. In the **Host** field, enter the host name or IP address of the SMTP server you want to use to send email notifications.
4. In the **Port** field, enter a port number or range to use for network communication between the Spectrum™ Technology Platform server and the SMTP server.

The default port is 25.

5. In the **User name** and **Password** fields, enter the credentials that the Spectrum™ Technology Platform server should use to authenticate with the SMTP server.
6. In the **From address** field, enter the email address from which notification email will be sent.
7. To confirm that you have correctly configured the mail server, you can send a test email. Enter the email address you want to send the test to in the **Test address** field then click **Test**.
8. Click **Save**.

The Spectrum™ Technology Platform server is now connected to an SMTP server and can use that server to send notification email.

### Example: Configuring a Mail Server

You have an SMTP server named mail.example.com. You want to use this mail server to handle email notifications sent from the Spectrum™ Technology Platform server. You have created an account on the SMTP server called Spectrum123 with a password of Example123, and the email address for this account is spectrum.notification@example.com.

To configure notification with this information, you would complete the fields as follows:

|                     |                                   |
|---------------------|-----------------------------------|
| <b>Host</b>         | mail.example.com                  |
| <b>From address</b> | spectrum.notification@example.com |

|                  |             |
|------------------|-------------|
| <b>User name</b> | Spectrum123 |
| <b>Password</b>  | Example123  |

**Related tasks**

[Scheduling a Flow](#) on page 163

[Selecting Items for Expiration Notification](#) on page 217

[Configuring Expiration Notification](#) on page 216

## Configuring Expiration Notification

Spectrum™ Technology Platform can send an email notification when a license, database, or software component is about to expire. This allows you to take the necessary action to ensure that your business processes are not disrupted by an expiration. Some of the components that have expiration dates include:

- Licenses

**Note:** Email notifications are not available for transaction-based licenses. If you are approaching the maximum number of transactions for a license, a message appears in the system log in Management Console.

- Databases, such as U.S. postal databases used for CASS processing
- Certain software components, such as the engine used to validate U.S. addresses in the Universal Addressing Module

**Tip:** To view the items that have expiration dates, open Management Console and go to **System > Licensing and Expiration**.

This procedure describes how to specify when to send expiration notifications, and the recipients of the notification emails.

1. Open the Management Console.
2. Go to **System > Licensing and Expiration**.
3. Click **Configure Notification**.
4. Check the **Send notification** box.
5. In the **Days before expiration** field, specify the number of days in advance that you want to be notified of a pending license, software, or data expiration. This is the default value. You can specify a different notification period for each license item on the **System > Licensing and Expiration** page.

For example, if you want to be notified 30 days before items expire, specify 30.

6. Under **Recipients**, click the Add button  and enter the email address you want to receive the expiration notification email. You can multiple email addresses if needed.



7. Click **Save**.

You have now specified recipients for the notifications and how far in advance of expiration to send the notification email. If you have not already done so, you must configure a mail server to use to send the emails. Notifications will not be sent until a mail server has been configured.

**Note:** By default the system will send expiration notifications for all items that expire (licenses, databases, software components, and so on). You can disable expiration notifications for specific items by going to **System > Licensing and Expiration**.

#### Related tasks

[Selecting Items for Expiration Notification](#) on page 217

[Configuring a Mail Server](#) on page 215

## Selecting Items for Expiration Notification

Spectrum™ Technology Platform can send an email notification when a license, database, or software component is about to expire. This allows you to take the necessary action to ensure that your business processes are not disrupted by an expiration. Some of the components that have expiration dates include:

- Licenses

**Note:** Email notifications are not available for transaction-based licenses. If you are approaching the maximum number of transactions for a license, a message appears in the system log in Management Console.

- Databases, such as U.S. postal databases used for CASS processing
- Certain software components, such as the engine used to validate U.S. addresses in the Universal Addressing Module

**Tip:** To view the items that have expiration dates, open Management Console and go to **System > Licensing and Expiration**.

You can choose which items you want to be notified about so that you only receive notifications for those items that concern you.

1. Open the Management Console.
2. Go to **System > Licensing and Expiration**.
3. To receive an expiration notification email for an item, check the box in the **Send Notification** column. If you want to be notified earlier or later than the default, specify the number of days in advance of the expiration that you want to be notified.

#### Related tasks

[Configuring Expiration Notification](#) on page 216

[Configuring a Mail Server](#) on page 215

# Audit Log

## Viewing the Audit Log

The audit log records the activities of users. It records events that occur when users create and modify objects on your system, as well as events that occur when users execute jobs or access services through the API or web services. Some examples of events recorded in the audit log include creating a dataflow, modifying a database connection, or running a job.

1. Open Management Console.
2. Go to **System > Logs**.
3. Click **Audit Log**.

The audit log lists the following information.

| Column    | Description   |
|-----------|---|
| Severity  | <p><b>Error</b> Errors indicate an isolated problem that causes part of the system to become unusable. For example, a problem that causes a single service to not work would generate an error.</p> <p><b>Warning</b> Warnings indicate problems that do not stop the system from working. For example, when loading a service where a parameter has an invalid value, a warning is issued and the default parameter is used. During the use of a service, if results are returned but there is a problem, a warning will be logged.</p> <p><b>Info</b> Info events are typically seen during startup and initialization, providing information such as version information and which services were loaded.</p> |
| Date/Time | The date and time of the event in the time zone of the Spectrum™ Technology Platform server.  |
| User      | The user account that performed the action.   |
| Source    | The software component that generated the event. This could be the name of a module or "Platform".  |

| Column      | Description   |
|-------------|---|
| Event       | <p>The action that occurred. The platform events are listed below. In addition to these events, other events may appear in your audit log depending on the modules you have installed.</p> <p><b>Create</b>                    The object was created and saved to the server.</p> <p><b>Create Version</b>        A new version of the dataflow was created in Enterprise Designer.</p> <p><b>Delete</b>                    The object was removed from the server.</p> <p><b>Delete Version</b>        The dataflow version was removed. Other versions may still exist.</p> <p><b>Expose</b>                    The dataflow was exposed, making it available for execution.</p> <p><b>Item added</b>                The object was added to a folder on the server.</p> <p><b>Item moved</b>              The object was moved to a different folder on the server.</p> <p><b>Rename</b>                    The object's name was modified and the object was saved.</p> <p><b>Unexpose</b>                The dataflow was made unavailable for execution. It can still be edited in Enterprise Designer.</p> <p><b>Update</b>                    The object was modified and saved.</p> |
| Type        | <p>The part of the system that was modified by the event. Examples of types include the dataflow type (job, service, subflow, process flow) file servers, and access control settings.</p> <p>In some situations an object of the same name may appear multiple times with different values in the <b>Type</b> column. This is because one user action may generate multiple events in the system. For example, when you create a job in Enterprise Designer, you will see a "Create" event for the job in the audit log, plus an "Item added" event for a FolderItem type that has the same name as the job. This indicates that the job was saved and placed in a folder on the system. The saving of the job and placement of the job into a folder are treated as two separate system events.</p>   |
| Object Name | <p>The name of the item that generated the log entry. For example, the name of a dataflow. Object names may be user-defined, such as the name of a dataflow, or they may be defined by the system.</p>  |

## Audit Log Archiving

Audit log events are archived on a monthly basis to help prevent the audit log from growing too large. Every month, events that are six months old and older are moved to a compressed file in this folder:

```
Spectrum Location\server\app\repository\store\archive
```

You can move the compressed file to another location for permanent archiving if needed.


## System Log

### Viewing System Events

The system log displays messages from the Spectrum™ Technology Platform server's wrapper log. These messages include information about server operations as well as requests made to services from the API and through web services. View the system log when you experience trouble and are looking for information about possible causes.

If you are running Spectrum™ Technology Platform in a cluster, the system log that you will get will be the one from the node you happen to be connected to. You can view the system log for a specific node by using a text editor to open this file on the node you want:

```
ServerLocation\server\app\repository\logs\wrapper.log.
```

1. Open the Management Console.
2. Go to **System > Logs**.
3. Click the Download icon  to download the system log file.
4. Open the downloaded file in a text editor.

### Setting Logging Levels for Services

You can specify the default logging level as well as logging levels for each service on your system. When you change logging levels the change will not be reflected in the log entries made before the change.

**Note:** The logging levels you specify for services do not affect the audit log. They only control the level of logging for the event log which you can view in Management Console. At this time you cannot view the event log in the web version of Management Console.

1. Open the Management Console.
2. Go to **System > Logs**.
3. In the **System default logging level** field, select a default event logging level for services on your system.

|                 |   |
|-----------------|---|
| <b>Disabled</b> | No event logging enabled.   |
| <b>Fatal</b>    | Minimal logging. Only fatal errors are logged. Fatal errors are those that make the system unusable.  |
| <b>Error</b>    | Errors and fatal errors are logged. Errors indicate an isolated problem that causes part of the system to become unusable. For example, a problem that causes a single service to not work would generate an error.   |
| <b>Warn</b>     | Event warnings, errors, and fatal errors are logged. Warnings indicate problems that do not stop the system from working. For example, when loading a service where a parameter has an invalid value, a warning is issued and the default parameter is used. During the use of a service, if results are returned but there is a problem, a warning will be logged. |
| <b>Info</b>     | High-level system information is logged. This is the most detailed logging level suitable for production. Info events are typically seen during startup and initialization, providing information such as version information and which services were loaded.   |
| <b>Debug</b>    | A highly detailed level of logging, suitable for debugging problems with the system.  |
| <b>Trace</b>    | The most detailed level of logging, tracing program execution (method entry and exit). It provides detailed program flow information for debugging.   |

Each logging level includes the ones above it on the list. In other words, if Warning is selected as the logging level, errors and fatal errors will also be logged. If Info is selected, informational messages, warnings, errors, and fatal errors will be logged.

**Note:** Selecting the most intensive logging level can affect system performance. Therefore, you should select the least intensive setting that meets your particular logging requirements.

4. If you want to specify different logging levels for each service choose the logging level you want.

## Logging the Record that Caused a Flow to Fail

When troubleshooting the cause of a flow failure it can be useful to examine the record that caused the failure. Flow failure records are written to a log file on the Spectrum™ Technology Platform server. The log file contains records that cause a stage within the flow to fail. The log does not capture records when the flow failure is due to other causes, such as malformed input records or expired licenses.

To enable the logging of records that cause flow failures:

1. Open a web browser and go to `http://server:port/jmx-console`

Where:

*server* is the IP address or hostname of your Spectrum™ Technology Platform server.

*port* is the HTTP port used by Spectrum™ Technology Platform. The default is 8080.

2. Log in using the admin account.
3. Scroll down to this entry and click it:

**com.pb.spectrum.platform.config.manager=LoggingConfigurationManager**

4. Set the attribute **LogLastRecordReadOnError** to **true** then click **set**.

Records that cause a flow to fail will now be logged in a new log file on the server:

`SpectrumLocation/server/app/repository/logs/error_records.log`

**Note:** Since this log may contain sensitive data, consider deleting the log file when you are done troubleshooting.

## Transaction Limit Warnings

Transaction-based licenses place a limit on the number of transactions you can perform before you need to renew the license. When you have approximately 10% of your transaction limit remaining, warning messages begin appearing in the event log in Management Console. For example, if you have a license that allows 1,000,000 transactions for the Universal Addressing Module's Validate Address service, and you have performed 900,000 transactions, you will begin to see messages like this in the event log:

```
WARN [ValidateAddress] license for feature(s): UNC/USA/RealTime has
100,000 transactions remaining
```

When you reach the limit, the feature is disabled and you will see messages like this in the event log:

```
ERROR [ValidateAddress] Usage limit exceeded for feature(s):  
UNC/USA/RealTime
```

**Note:** The system calculates the number of remaining transactions every few minutes and logs the warning message if necessary. If a job or a large number of transactions occurs and uses up the final 10% of remaining transactions at once, the remaining transactions may be used up before the system can display the warning message. In this situation, the warning message will not appear before the feature is disabled.

To view the number of transactions remaining on your license, open Management Console, expand **System**, click **Licensing and Expiration**, then click the **License Info** tab.

To renew your license, contact your Pitney Bowes Account Executive.

## Viewing Version Information

1. In a web browser go to this URL:

`http://server.port/managementconsole`

Where *server* is the server name or IP address of your Spectrum™ Technology Platform server and *port* is the HTTP port used by Spectrum™ Technology Platform. By default, the HTTP port is 8080.

2. Click **System > Version**.

## Viewing and Exporting License Information

You can export information about your license to an XML file. This may be necessary when resolving license issues with technical support.

1. In a web browser go to this URL:

`http://server.port/managementconsole`

Where *server* is the server name or IP address of your Spectrum™ Technology Platform server and *port* is the HTTP port used by Spectrum™ Technology Platform. By default, the HTTP port is 8080.

2. Click **System > Licensing and Expiration**.

3. Click the export icon.

Your license information is saved to an XML file with a `.lic` extension.



# 9 - Backup and Restore

## In this section

---

|                            |     |
|----------------------------|-----|
| Scheduling Backups         | 226 |
| Creating a Backup Manually | 230 |
| Restoring a Server         | 231 |

## Scheduling Backups

To back up your Spectrum™ Technology Platform server, you need to create a backup copy of the server's configuration database. The configuration database contains your security settings, dataflows, service options, data resource definitions, and other configuration settings. If you were to lose your server due to a severe system failure or other disaster, you could use the backup of the configuration database to restore your configuration to another Spectrum™ Technology Platform server.

This procedure describes how to configure Spectrum™ Technology Platform to create a backup on a regular schedule.

**Important:** Schedule backups to occur when there is little or no activity on the Spectrum™ Technology Platform server. While the backup is in progress, calls to services may time out and jobs may fail to run successfully.

1. Stop the Spectrum™ Technology Platform server.

**Note:** If you are running Spectrum™ Technology Platform in a clustered environment, choose a node on which to configure scheduled backups and stop the server on that node. You will configure this node for scheduled backups then apply the configuration to the other nodes.

2. Open this file in a text editor:

```
SpectrumLocation\server\app\conf\spectrum-container.properties
```

3. Specify these parameters:

```
spectrum.backup.enabled=true
spectrum.backup.cron=Interval
spectrum.backup.directory=Destination
```

Where:

### ***Interval***

A cron expression that specifies how often to create the backup database. A cron expression consists of six space-separated values, with an optional seventh value:

| Field   | Valid Values | Valid Special Characters |
|---------|--------------|--------------------------|
| Seconds | 0-59         | , - * /                  |
| Minutes | 0-59         | , - * /                  |

| Field            | Valid Values    | Valid Special Characters |
|------------------|-----------------|--------------------------|
| Hours            | 0-23            | , - * /                  |
| Day of the month | 1-31            | , - * ? / L W            |
| Month            | 1-12 or JAN-DEC | , - * /                  |
| Day of the week  | 1-7 or SUN-SAT  | , - * ? / L #            |
| Year (Optional)  | 1970 - 2099     | , - * /                  |

For example, this expression would back up the configuration database every day at 10:00 AM:

```
spectrum.backup.cron=0 0 10 * * ?
```

This expression would back up the configuration database on the first day of the month at 2 AM:

```
spectrum.backup.cron=0 0 2 1 * ?
```

The special characters are:

\*

Specifies all values. For example, if you use \* in the day-of-the-month field, it means every day of the month.

?

Specifies no specific value. This is used in combination with other fields. For example, if you want to run a backup on the first day of the month and don't care which day of the week the first is on, you would specify ? in the day-of-the-week field and 1 in the month field.

-

Specifies a range of values. For example, SAT-SUN means the Saturday through Sunday.

,

Separates multiple values. For example, 15, 30 in the day-of-the-month field means the 15th day of the month and the 30th day of the month.

/

Specifies increments. For example, 0/3 in the hour field means the backup will occur at midnight then every three hours.

**L**

Specifies "last", which has different meaning depending on the field in which it is used. When used in the day-of-the-month field, it means the last day of the month. When used alone in the day-of-the-week field, it means Saturday. However, when used in the day-of-the-week field in combination with a day, it means the last day-of-the-week in the month. For example, `6L` means the last Friday of the month.

**W**

Use this value in the day-of-the-month field to specify the weekday nearest to a given day. For example, `15W` means the nearest weekday to the 15th day of the month.

For more information about cron expressions, including examples, see [quartz-scheduler.org](http://quartz-scheduler.org).

**Destination**

The directory where you want to save the backup database. For example,

```
spectrum.orientdb.backup.directory\\\\"exampleserver1\\Shared\\Backup
```

You must use the escape character `\` when specifying a backslash in the path.

**Note:** If you are using Spectrum™ Technology Platform in a clustered environment, you should specify a centralized location as the backup destination. This is because in a clustered environment, scheduled backups occur on a random node in the cluster. By specifying a centralized location it will be easier to retrieve the latest backup from the cluster.

- If you are running the configuration database on a separate server from the Spectrum™ Technology Platform server, configure these properties. If not, leave them with their default values.

```
spectrum.backup.databaseURL=Host
spectrum.backup.http.port=Port
```

Where:

**Host**

The host name or IP address of one of the configuration databases. Typically the configuration database is run in a cluster. You can use any node in the cluster as the source for the backup.

**Port**

The HTTP database backup port configured on your database.

- Save and close the properties file.
- Start the Spectrum™ Technology Platform server.
- Optional: If you are using Spectrum™ Technology Platform in a clustered environment, repeat this procedure for each node in the cluster.

**Note:** You must specify identical values for `spectrum.backup.cron` and `spectrum.backup.directory` on all nodes in the cluster.

- Some modules store additional data that is not backed up as part of the Spectrum™ Technology Platform scheduled backup process. You must back up this data manually, or create a separate process to back up this data.

| Module   | Items to Back Up   |
|--|--|
| Advanced Matching Module, Data Normalization Module, and Universal Name Module | <p>Back up the contents of these subfolders located in <code>SpectrumLocation/server/modules</code>:</p> <ul style="list-style-type: none"> <li>• <code>cdqdb</code></li> <li>• <code>lucene</code></li> <li>• <code>matcher</code></li> <li>• <code>parser</code></li> <li>• <code>searchindex</code></li> <li>• <code>tables</code></li> <li>•</li> </ul>                        |
| Data Hub Module  | <p>Open the Relationship Analysis Client and click <b>Manage</b>. Select the model you want to back up then click <b>Backup</b>.</p> <p>In addition to backing up your models, back up these two property files:</p> <ul style="list-style-type: none"> <li>• <code>server\modules\hub\hub.properties</code></li> <li>• <code>server\modules\db\neo4j.properties</code></li> </ul> |
| Location Intelligence Module   | <p>Back up your named resources, data, and configuration files.</p>  |

### Related concepts

[Creating a Backup Manually](#) on page 230

### Related tasks

[Restoring a Server](#) on page 231

### Related reference

[server backup](#) on page 366

## Creating a Backup Manually

To back up your Spectrum™ Technology Platform server, you need to create a backup copy of the server's configuration database. The configuration database contains your security settings, dataflows, service options, data resource definitions, and other configuration settings. If you were to lose your server due to a severe system failure or other disaster, you could use the backup of the configuration database to restore your configuration to another Spectrum™ Technology Platform server.

To manually create a backup of the Spectrum™ Technology Platform configuration database, use the Administration Utility's `server backup` command. For more information, see [server backup](#) on page 366.

In addition, some modules have data that is not included in the Administration Utility backup process. You must back up this data separately:

| Module   | Items to Back Up   |
|--|--|
| Advanced Matching Module, Data Normalization Module, and Universal Name Module | <p>Back up the contents of these subfolders located in <i>SpectrumLocation/server/modules</i>:</p> <ul style="list-style-type: none"> <li>• cdqdb</li> <li>• lucene</li> <li>• matcher</li> <li>• parser</li> <li>• searchindex</li> <li>• tables</li> <li>.</li> </ul>  |
| Data Hub Module  | <p>Open the Relationship Analysis Client and click <b>Manage</b>. Select the model you want to back up then click <b>Backup</b>.</p> <p>In addition to backing up your models, back up these two property files:</p> <ul style="list-style-type: none"> <li>• server\modules\hub\hub.properties</li> <li>• server\modules\db\neo4j.properties</li> </ul> |
| Location Intelligence Module   | <p>Back up your named resources, data, and configuration files.</p>  |

### Related tasks

[Scheduling Backups](#) on page 226

[Restoring a Server](#) on page 231

**Related reference**

[server backup](#) on page 366

## Restoring a Server

If you lose your server due to a severe system failure or other disaster, you can restore your server using a backup of the configuration database. In order to have a backup you must have either created a backup manually or have configured Spectrum™ Technology Platform to create backups on a regular schedule. By default, Spectrum™ Technology Platform does not create backups of the configuration database.

**Note:** This procedure is intended to be used in situations where you have a single Spectrum™ Technology Platform server. If you have a cluster of Spectrum™ Technology Platform servers and you need to restore a single node, install a new server and add it to the node. The configuration of the cluster will automatically be applied to the new node, in effect restoring the node. The only scenario where you would need to restore from a backup in a clustered environment would be in the event of a complete loss of all nodes in the cluster.

1. Install a new Spectrum™ Technology Platform server. For more information, see the *Installation Guide*.
2. If the server is running, stop the server.
3. Obtain the backup zip file and unzip it to this location, overwriting the existing files:

```
SpectrumLocation\repository\data\databases
```

This will replace the existing `graph.db` folder.

4. Restore the module-specific data for any modules you have installed.

| Module   | Items to Back Up   |
|--|--|
| Advanced Matching Module, Data Normalization Module, and Universal Name Module | Restore the contents of these subfolders located in <i>SpectrumLocation/server/modules</i> : <ul style="list-style-type: none"> <li>• cdqdb</li> <li>• lucene</li> <li>• matcher</li> <li>• parser</li> <li>• searchindex</li> <li>• tables</li> </ul> |

| Module                       | Items to Back Up   |
|------------------------------|--|
| Data Hub Module              | <p>Restore your models.</p> <p>In addition to restoring your models, restore these two property files:</p> <ul style="list-style-type: none"> <li>• <code>server\modules\hub\hub.properties</code></li> <li>• <code>server\modules\db\neo4j.properties</code></li> </ul> |
| Location Intelligence Module | <p>Restore your named resources, data, and configuration files.</p>  |

5. Start the server.
6. Wait for the server to fully start.
7. Stop the server.
8. Apply all the updates for the platform and any modules you have installed. For a listing of updates, see the [Update Summary](#) on the Pitney Bowes support website.

#### Related concepts

[Creating a Backup Manually](#) on page 230

#### Related tasks

[Scheduling Backups](#) on page 226

#### Related reference

[server backup](#) on page 366



# 10 - Administration Utility

## In this section

---

|   |     |
|---|-----|
| Getting Started with the Administration Utility | 234 |
| Using a Script with the Administration Utility  | 235 |
| Data Sources                                    | 236 |
| Dataflows                                       | 245 |
| Enterprise Routing Module                       | 251 |
| Folders   | 270 |
| Jobs  | 273 |
| Location Intelligence Module                    | 282 |
| Match Rules                                     | 288 |
| Notification                                    | 291 |
| Open Parser Cultures                            | 295 |
| Open Parser Domains                             | 296 |
| Performance Monitor                             | 297 |
| Permissions                                     | 301 |
| Process Flows                                   | 301 |
| Remote Servers                                  | 310 |
| Roles   | 314 |
| Search Indexes                                  | 318 |
| Services  | 322 |
| Spectrum Databases                              | 326 |
| System  | 362 |
| Tables  | 370 |
| Tokens  | 373 |
| User Accounts                                   | 375 |

## Getting Started with the Administration Utility

The Administration Utility provides command line access to administrative functions. You can use it in a script, allowing you to automate certain administrative tasks. You can also use it interactively. Not all administrative functions are available in the Administration Utility. Use Management Console to access the functions that are not available in the Administration Utility.

**Note:** The Administration Utility requires Java 8 or later. If you are installing the Administration Utility on the same computer where the Spectrum™ Technology Platform server is installed, Java 8 is already installed since it is included with the server. If you want to install the Administration Utility on another computer, make sure that the computer has Java 8 or later installed.

1. Open a web browser and go to the Spectrum™ Technology Platform Welcome Page at:  
`http://<servername>:<port>`  
For example, if you installed Spectrum™ Technology Platform on a computer named "myspectrumplatform" and it is using the default HTTP port 8080, you would go to:  
`http://myspectrumplatform:8080`
2. Click **Platform Client Tools**.
3. Click **Command Line Clients**.
4. Under **Administration Utility**, click **Download** and download the zip file to the computer where you want to use the Administration Utility.
5. Extract the contents of the zip file.
6. To launch the command line interface, do one of the following:
  - If you are running the server on a Unix or Linux system, execute `cli.sh`.
  - If you are running the server on a Windows system, execute `cli.cmd`.

**Note:** If necessary, modify the `.sh` or `.cmd` file to use the path to your Java installation.
7. Connect to the Spectrum™ Technology Platform server by typing this command:  
`connect --h <servername>:<port> --u <username> --p <password> --s <SSLTrueOrFalse>`  
For example,  
`connect --h myserver:8080 --u admin --p myPassword1 --s true`
8. Once you are connected you can run commands. Some tips:
  - For a list of available commands, type `help` or press the tab key.

- To auto-complete a command, type the first few characters then press the tab key. For example, typing `us` then pressing the tab key automatically completes the command `user`. Pressing the tab key again will display a list of all the `user` commands.
  - If you specify an option value that contains a space, enclose the value in double quotes.
9. When you are done, type `exit` to exit the Administration Utility.

## Using a Script with the Administration Utility

The Administration Utility can execute a series of commands from a script file. This is useful if you want to automate or standardize administrative actions through the use of a script instead of manually executing commands through the Administration Utility or by using the Management Console.

1. Using a text editor, create a script file. A script file contains the commands that you want to execute.

To add a command to a script file, type the command and the necessary parameters as you would if you were entering the command at the command prompt. Enter one command per line.

To insert comments into a script file, use the following notation:

|                 |   |
|-----------------|---|
| <code>/*</code> | Indicates the start of a block comment.                       |
| <code>*/</code> | Indicates the end of a block comment.                         |
| <code>//</code> | Indicates an inline comment. Use at the start of a line only. |
| <code>;</code>  | Indicates an inline comment. Use at the start of a line only. |

2. Save the script either on the computer where you run the Administration Utility or in a location that is accessible from the computer where you run the Administration Utility. You can use any file name and extension you choose. We recommend that you use the file extension `.cli`.
3. To run the script, do one of the following:

| Option   | Description  |
|--|--|
| <b>To run the script at the command line</b>             | Run this command at the command line or in a batch or shell script:<br><br><code>cli.cmd --cmdfile <i>ScriptFile</i></code>  |
| <b>To run the script from the Administration Utility</b> | Open the Administration Utility and connect to the Spectrum™ Technology Platform server using the <code>connect</code> command. Then, use the <code>script</code> command to execute the script. For more information on this command, see <a href="#">script</a> on page 367. |

**Example: Moving Dataflows from Staging to Production**

You have three dataflows: Deduplication, AddressValidation, and DrivingDirections. You have a staging server where you make changes to these dataflows and test them, and a production environment where the dataflows are made available for execution. You want to have a consistent and automated way to move these dataflows from your staging server to your production server so you decide to use an Administration Utility script to accomplish this. The script might look like this:

```
// Connect to the staging server
connect --h stagingserver:8080 --u allan12 --p something123

// Export from staging
dataflow export --d "Deduplication" --e true --o exported
dataflow export --d "AddressValidation" --e true --o exported
dataflow export --d "DrivingDirections" --e true --o exported

// Close connection to the staging server
close

// Connect to the production server
connect --h productionserver:8080 --u allan12 --p something123

// Import to production
dataflow import --f exported\Deduplication.df
dataflow import --f exported\AddressValidation.df
dataflow import --f exported\DrivingDirections.df

// Close the connection to the production server
close
```

**Related reference**  
[script](#) on page 367

## Data Sources

### FTP

#### **data source ftp add**

The `data source ftp add` command creates a connection between Spectrum™ Technology Platform and an FTP server.

### Usage

```
data source ftp add --n ConnectionName --h Host --o Port --u Username --p Password
```

| Required | Argument                  | Description   |
|----------|---------------------------|---|
| Yes      | --n <i>ConnectionName</i> | Specifies the name for the connection. The name can be anything you choose. |
| Yes      | --h <i>Host</i>           | Specifies the host name or IP address of the FTP server.                    |
| No       | --o <i>Port</i>           | Specifies the network port to use for communication with the FTP server.    |
| No       | --u <i>Username</i>       | The username to use to connect to the FTP server, if required.              |
| No       | --p <i>Password</i>       | The password to use to connect to the FTP server, if required.              |

#### Example

This example creates a connection to the FTP server named MyFTPServer.

```
data source ftp add --n NorthernRegionCustomers --h MyFTPServer
--u ExampleUsername --p Example123
```

### data source ftp delete

The `data source ftp delete` command deletes a connection between Spectrum™ Technology Platform and an FTP server.

### Usage

```
data source ftp delete --n ConnectionName
```

| Required | Argument                  | Description  |
|----------|---------------------------|--|
| Yes      | --n <i>ConnectionName</i> | Specifies the name of the connection you want to delete. To view a list of connections, use the <code>data source ftp list</code> command. |

#### Example

This example deletes a connection named NorthernRegionCustomers.

```
data source ftp delete --n NorthernRegionCustomers
```

### data source ftp list

The `data source ftp list` command returns a list of the FTP connections defined on the Spectrum™ Technology Platform server.

### Usage

```
data source ftp list
```

### data source ftp test

The `data source ftp test` command tests a connection between Spectrum™ Technology Platform and an FTP server.

### Usage

```
data source ftp test --n ConnectionName
```

| Required | Argument                  | Description  |
|----------|---------------------------|--|
| Yes      | --n <i>ConnectionName</i> | Specifies the name of the connection you want to test. To view a list of connections, use the <code>data source ftp list</code> command. |

#### Example

This example test the connection NorthernRegionCustomers.

```
data source ftp test --n NorthernRegionCustomers
```

### data source ftp update

The `data source ftp update` command modifies a connection between Spectrum™ Technology Platform and an FTP server.

### Usage

```
data source ftp update --n ConnectionName --h Host --o Port --u Username --p Password
```

| Required | Argument                  | Description   |
|----------|---------------------------|---|
| Yes      | --n <i>ConnectionName</i> | Specifies the name for the connection you want to modify. To view a list of connections, use the <code>data source ftp list</code> command. |
| Yes      | --h <i>Host</i>           | Specifies the host name or IP address of the FTP server.  |
| No       | --o <i>Port</i>           | Specifies the network port to use for communication with the FTP server.  |
| No       | --u <i>Username</i>       | The username to use to connect to the FTP server, if required.  |
| No       | --p <i>Password</i>       | The password to use to connect to the FTP server, if required.  |

**Example**

This example modifies an FTP connection named NorthernRegionCustomers. It changes the host to MyFTPServer2.

```
data source ftp update --n NorthernRegionCustomers --h
MyFTPServer2
```

## JDBC Database


### Connections

#### *dbconnection add*

The `dbconnection add` command creates a connection between Spectrum™ Technology Platform and a database.

#### *Usage*

```
dbconnection add --n ConnectionName --d Driver --h Host --o Port --i Instance --u
Username --p Password --l "property:value"
```

| Required | Argument                      | Description  |
|----------|-------------------------------|--|
| Yes      | --n <i>ConnectionName</i>     | Specifies the name for the connection. The name can be anything you choose.  |
| Yes      | --d <i>Driver</i>             | Specifies the driver for the type of database you want to connect to. To view a list of database drivers available on your server, use the <code>dbdriver list</code> command.   |
| Yes      | --h <i>Host</i>               | Specifies the host name or IP address of the database server.  |
| No       | --o <i>Port</i>               | Specifies the network port to use for communication with the database server.  |
| No       | --i <i>Instance</i>           | Specifies the database instance to connect to.   |
| No       | --u <i>Username</i>           | The username to use to connect to the database, if required.   |
| No       | --p <i>Password</i>           | The password to use to connect to the database, if required.   |
| No       | --l " <i>property:value</i> " | Specifies a comma-separated list of connection property and value pairs for the driver. To view the list of valid properties for a driver, open Management Console, go to <b>Resources &gt; Data Sources</b> , then click the <b>Drivers</b> tab. Select the driver you want then click the Edit button  to view its connection properties. |

**Example**

This example creates a connection to a database located on the host MyServer. The name of the connection will be NorthernRegionCustomers. It will use the driver ExampleSQLDriver which takes two connection properties: ExampleProp1, which is given a value of 123, and ExampleProp2, which is given a value of 456.

```
dbconnection add --n NorthernRegionCustomers --d
ExampleSQLDriver --h MyServer --l
"ExampleProp1:123,ExampleProp2:456"
```

**dbconnection delete**

The `dbconnection delete` command deletes a connection between Spectrum™ Technology Platform and a database.

**Usage**

```
dbconnection delete --n ConnectionName
```

| Required | Argument                  | Description   |
|----------|---------------------------|---|
| Yes      | --n <i>ConnectionName</i> | Specifies the name of the connection you want to delete. To view a list of connections, use the <code>dbconnection list</code> command. |

**Example**

This example deletes a connection named NorthernRegionCustomers.

```
dbconnection delete --n NorthernRegionCustomers
```

**dbconnection export**

The `dbconnection export` command exports a database connection definition to a JSON file.

**Usage**

```
dbconnection export --n ConnectionName --o OutputDirectory
```

| Required | Argument                   | Description   |
|----------|----------------------------|---|
| Yes      | --n <i>ConnectionName</i>  | Specifies the name of the database connection you want to export. If the connection name contains spaces, enclose the name in quotes.<br><br><b>Tip:</b> If you are unsure of the connection name you can use the <code>dbconnection list</code> command to get a list of the connection names. |
| No       | --o <i>OutputDirectory</i> | Specifies the directory to which you want to export the database connection. The path you specify here is relative to the directory where you are running the Administration Utility. If you omit this  |



| Required Argument | Description   |
|-------------------|---|
|                   | argument, the database connection is exported to the directory containing the Administration Utility. |

**Example**

This example exports the database connection named "My Connection" to a folder named `exported` which is a subfolder in the location where you have installed the Administration Utility.

```
dbconnection export --n "My Connection" --o exported
```

**dbconnection import**

The `dbconnection import` command imports a database connection definition file into the server. Database connection definition files are created by exporting a database connection from the server using the `dbconnection export` command. You can only import database connections that were exported from the same version of Spectrum™ Technology Platform.

**Usage**

```
dbconnection import --f DatabaseConnectionFile --u TrueOrFalse
```

| Required Argument                                  | Description   |
|--|---|
| Yes <code>--f <i>DatabaseConnectionFile</i></code> | Specifies the database connection file you want to import. Connection files have a <code>.json</code> extension. Directory paths you specify here are relative to the location where you are running the Administration Utility.  |
| No <code>--u <i>TrueOrFalse</i></code>             | Specifies whether to overwrite the existing database connection if a connection with the same name is already on the server, where <i>TrueOrFalse</i> is one of the following: <p><b>true</b></p> <p>If there is a database connection on the server with the same name as the one you are importing, the connection on the server will be overwritten. This is the default setting.</p> <p><b>false</b></p> <p>If there is a database connection on the server with the same name as the one you are importing, the connection will not be imported.</p> |

**Example**

This example imports the database connection definition named `MyDatabaseConnection.json` which is located in a subfolder named `exported` in the location where you are running the Administration Utility.

```
dbconnection import --f exported\MyDatabaseConnection.json
```

**dbconnection list**

The `dbconnection list` command returns a list of the database connections defined on the Spectrum™ Technology Platform server.

**Usage**

```
dbconnection list
```

**dbconnection test**

The `dbconnection test` command tests a connection between Spectrum™ Technology Platform and a database.

**Usage**

```
dbconnection test --n ConnectionName
```

| Required | Argument                  | Description   |
|----------|---------------------------|---|
| Yes      | --n <i>ConnectionName</i> | Specifies the name of the connection you want to test. To view a list of connections, use the <code>dbconnection list</code> command. |

**Example**

This example test the connection `NorthernRegionCustomers`.

```
dbconnection test --n NorthernRegionCustomers
```


**dbconnection update**

The `dbconnection update` command modifies a connection between Spectrum™ Technology Platform and a database.

**Usage**

```
dbconnection update --n ConnectionName --d Driver --h Host --o Port --i Instance --u Username --p Password --l "property:value"
```

| Required | Argument                  | Description  |
|----------|---------------------------|--|
| Yes      | --n <i>ConnectionName</i> | Specifies the name for the connection you want to modify. To view a list of connections, use the <code>dbconnection list</code> command. |

| Required | Argument                      | Description  |
|----------|-------------------------------|--|
| Yes      | --d <i>Driver</i>             | Specifies the driver for the type of database you want to connect to. To view a list of database drivers available on your server, use the <code>dbdriver list</code> command.   |
| Yes      | --h <i>Host</i>               | Specifies the host name or IP address of the database server.  |
| No       | --o <i>Port</i>               | Specifies the network port to use for communication with the database server.  |
| No       | --i <i>Instance</i>           | Specifies the database instance to connect to.   |
| No       | --u <i>Username</i>           | The username to use to connect to the database, if required.   |
| No       | --p <i>Password</i>           | The password to use to connect to the database, if required.   |
| No       | --l " <i>property:value</i> " | Specifies a comma-separated list of connection property and value pairs for the driver. To view the list of valid properties for a driver, open Management Console, go to <b>Resources &gt; Data Sources</b> , then click the <b>Drivers</b> tab. Select the driver you want then click the Edit button  to view its connection properties. |

#### Example

This example modifies a database connection named NorthernRegionCustomers. It changes the driver to `MSSQLServer2`, changes the host to `MyServer2`, and changes the instance to `MyInstance2`.

```
dbconnection update --n NorthernRegionCustomers --d MSSQLServer2
--h MyServer2 --i MyInstance2
```

## Drivers

### *dbdriver delete*

The `dbdriver delete` command deletes a JDBC driver definition. It does not delete driver files, only the definition created in Management Console.

### Usage

```
dbdriver delete --n DriverName
```

| Required | Argument              | Description  |
|----------|-----------------------|--|
| Yes      | --n <i>DriverName</i> | Specifies the name of the JDBC driver you want to delete. To view a list of drivers, use the <code>dbdriver list</code> command. |

#### Example

This example deletes a JDBC driver named MyDriver.

```
dbconnection delete --n MyDriver
```

### dbdriver export

The `dbdriver export` command exports a JDBC driver definition to a JSON file. It does not export driver files, only the driver definition created in Management Console.

#### Usage

```
dbdriver export --n DriverName --o OutputDirectory
```

| Required | Argument                                | Description  |
|----------|---|--|
| Yes      | <code>--n <i>DriverName</i></code>      | Specifies the name of the database driver you want to export. If the driver name contains spaces, enclose the name in quotes.<br><br><b>Tip:</b> If you are unsure of the driver name you can use the <code>dbdriver list</code> command to get a list of the driver names.                          |
| No       | <code>--o <i>OutputDirectory</i></code> | Specifies the directory to which you want to export the database driver. The path you specify here is relative to the directory where you are running the Administration Utility. If you omit this argument, the database driver is exported to the directory containing the Administration Utility. |

#### Example

This example exports the database driver named "My Driver" to a folder named `exported` which is a subfolder in the location where you have installed the Administration Utility.

```
dbdriver export --n "My Driver" --o exported
```

### dbdriver import

The `dbdriver import` command imports a JDBC database driver definition file into the server. Database driver definition files are created by exporting a database driver definition from the server using the `dbdriver export` command. You can only import database driver definitions that were exported from the same version of Spectrum™ Technology Platform.

#### Usage

```
dbdriver import --f DriverDefinitionFile --u TrueOrFalse
```

| Required | Argument                                     | Description   |
|----------|--|---|
| Yes      | <code>--f <i>DriverDefinitionFile</i></code> | Specifies the database driver JSON file you want to import. Directory paths you specify here are relative to the location where you are running the Administration Utility. |

| Required Argument                      | Description  |
|--|--|
| No <code>--u <i>TrueOrFalse</i></code> | <p>Specifies whether to overwrite the existing database driver definition if a database driver with the same name is already on the server, where <i>TrueOrFalse</i> is one of the following:</p> <p><b>true</b></p> <p>If there is a database driver on the server with the same name as the one you are importing, the driver on the server will be overwritten. This is the default setting.</p> <p><b>false</b></p> <p>If there is a database driver on the server with the same name as the one you are importing, the driver will not be imported.</p> |

**Example**

This example imports the database driver definition named `MyDatabaseDriver.json` which is located in a subfolder named `exported` in the location where you are running the Administration Utility.

```
dbdriver import --f exported\MyDatabaseDriver.json
```

***dbdriver list***

The `dbdriver list` command returns a list of the database drivers defined on the Spectrum™ Technology Platform server.

***Usage***

```
dbdriver list
```

## Dataflows

### dataflow delete

The `dataflow delete` command removes a dataflow from your system.

***Usage***

```
dataflow delete --d DataflowName
```

| Required | Argument                             | Description  |
|----------|--------------------------------------|--|
| Yes      | <code>--d <i>DataflowName</i></code> | Specifies the dataflow to delete. If the dataflow name contains spaces, enclose the dataflow name in quotes. |

**Example**

This example deletes the dataflow named My Dataflow.

```
dataflow delete --d "My Dataflow"
```

## dataflow export

The `dataflow export` command exports a dataflow from the server to a `.df` file. The dataflow can then be imported to another server.

**Note:** Dataflows can only be exchanged between identical versions of Spectrum™ Technology Platform.

### Usage

```
dataflow export --d DataflowName --e TrueOrFalse --o OutputDirectory
```

| Required | Argument                                | Description  |
|----------|---|--|
| Yes      | <code>--d <i>DataflowName</i></code>    | Specifies the name of the dataflow you want to export. If the dataflow name contains spaces, enclose the name in quotes.<br><br><b>Tip:</b> If you are unsure of the exact dataflow name you can use the <code>dataflow list</code> command to get a list of the dataflow names.       |
| Yes      | <code>--e <i>TrueOrFalse</i></code>     | Specifies whether to export the exposed version of the dataflow, where <i>TrueOrFalse</i> is one of the following:<br><br><b>true</b><br><br>Export the exposed version of the dataflow.<br><br><b>false</b><br><br>Export the most recently saved version of the dataflow.            |
| No       | <code>--o <i>OutputDirectory</i></code> | Specifies the directory to which you want to export the dataflow. The path you specify here is relative to the directory where you are running the Administration Utility. If you omit this argument, the dataflow is exported to the directory containing the Administration Utility. |

**Example**

This example exports the exposed version of a dataflow named "My Dataflow" to a folder named `exported` which is a subfolder in the location where you have installed the Administration Utility.

```
dataflow export --d "My Dataflow" --e true --o exported
```

## dataflow expose

The `dataflow expose` command makes the dataflow available for execution. For service dataflows, exposing the dataflow makes the service available to web service requests and API calls, and makes it available for setting logging levels. For subflows, exposing the dataflow makes the subflow available for use in a dataflow. For job dataflows, exposing the dataflow makes it possible to run the job through the Job Executor command line tool. To expose a process flow use the `processflow expose` command.

**Note:** If you use dataflow visioning in Enterprise Designer, the `dataflow expose` command exposes the most recently saved version of the dataflow.

### Usage

```
dataflow expose --d DataflowName
```

| Required | Argument                             | Description  |
|----------|--------------------------------------|--|
| Yes      | <code>--d <i>DataflowName</i></code> | Specifies the name of the dataflow you want to expose. If the dataflow name contains spaces, enclose the name in quotes.<br><br><b>Tip:</b> If you are unsure of the exact dataflow name you can use the <code>dataflow list</code> command to get a list of the dataflow names. |

**Example**

This example exposes the dataflow named "My Dataflow".

```
dataflow expose --d "My Dataflow"
```

## dataflow import

The `dataflow import` command imports a dataflow file (a `.df` file) into the server. Dataflow files are created by exporting a dataflow from the server using the `dataflow export` command

### Usage

```
dataflow import --f DataflowFile --u TrueOrFalse --p Path --c TrueOrFalse
```

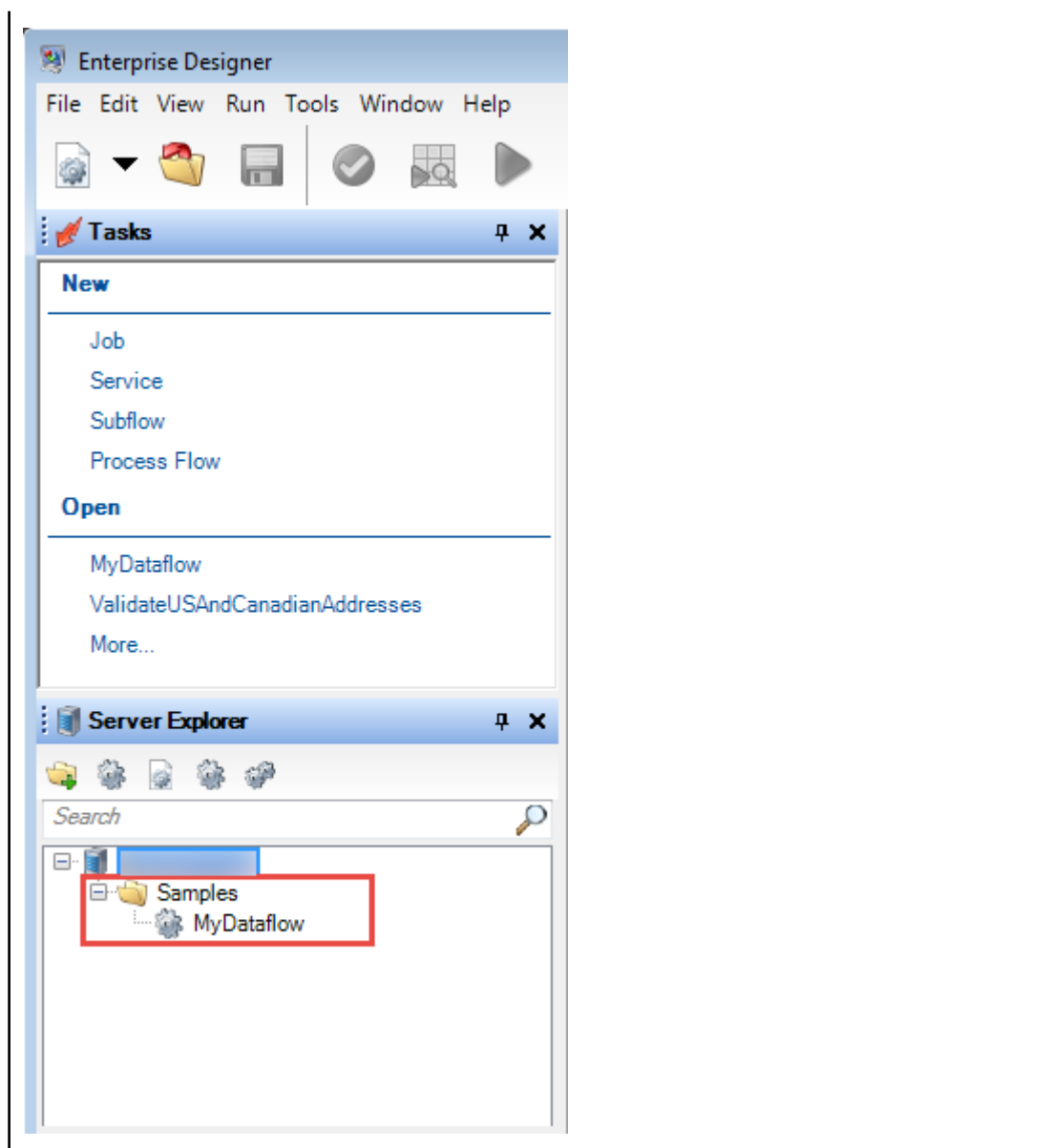
| Required | Argument                | Description  |
|----------|-------------------------|--|
| Yes      | --f <i>DataflowFile</i> | Specifies the dataflow file (the <code>.df</code> file) you want to import. Relative directory paths are relative to the location where you are running the Administration Utility. You can also specify an absolute path.   |
| No       | --u <i>TrueOrFalse</i>  | Specifies whether to overwrite the existing dataflow if a dataflow with the same name is already on the server, where <i>TrueOrFalse</i> is one of the following:<br><b>true</b><br>If there is a dataflow on the server with the same name as the dataflow you are importing, the dataflow on the server will be overwritten. This is the default setting.<br><b>false</b><br>If there is a dataflow on the server with the same name as the dataflow you are importing, the dataflow will not be imported. |
| No       | --p <i>Path</i>         | Specifies the Enterprise Designer Server Explorer folder to import the flow into.  |
| No       | --c <i>TrueOrFalse</i>  | Specifies whether to create the folder specified in --p if it does not exist.<br><b>true</b><br>Create the folder specified in --p if it does not exist. Default.<br><b>false</b><br>Do not create the folder specified in --p if it does not exist. The flow will not be imported unless the folder specified in --p exists.  |

#### Example

This example imports the dataflow named `MyDataflow.df` which is located in a subfolder named `exported` in the location where you are running the Administration Utility. The dataflow will be imported into the `Samples` folder in Enterprise Designer.

```
dataflow import --f exported\MyDataflow.df --p Samples
```





## dataflow sourcesink list

The `dataflow sourcesink list` command lists the stages in a dataflow that specify the input for the dataflow and the stages that specify the output from the dataflow.

### Usage

```
dataflow sourcesink list --d DataflowName --e TrueOrFalse --o TrueOrFalse
```

| Required | Argument                             | Description   |
|----------|--------------------------------------|---|
| Yes      | <code>--d <i>DataflowName</i></code> | <p>Specifies the name of the dataflow whose sources and sinks you want to list. If the dataflow name contains spaces, enclose the name in quotes.</p> <p><b>Tip:</b> If you are unsure of the exact dataflow name you can use the <code>dataflow list</code> command to get a list of the dataflow names.</p>   |
| Yes      | <code>--e <i>TrueOrFalse</i></code>  | <p>Specifies whether to list the sources and sinks for the exposed version of the dataflow or the latest saved version.</p> <p><b>true</b></p> <p>List the sources and sinks in the exposed version of the dataflow.</p> <p><b>false</b></p> <p>List the sources and sinks in the most recently saved version of the dataflow.</p>  |
| No       | <code>--o <i>TrueOrFalse</i></code>  | <p>Specifies whether to list only those sources and sinks that allow file overrides at runtime. A file override is when you specify at runtime a different file for the stage to read from or write to, overriding the file specified in the stage itself. Stages that support file overrides include Read from File and Write to File. Stages that do not support file overrides include Write to Null and Terminator.</p> <p><b>true</b></p> <p>List only those stages that support file overrides.</p> <p><b>false</b></p> <p>List all sources and sinks. This is the default setting.</p> |

**Example**

This example lists the sources and sinks in the exposed version of a dataflow named "My Dataflow". All sources and sinks are listed, even those that do not allow file overrides.

```
dataflow sourcesink list --d "My Dataflow" --e true
```

## dataflow unexpose

The `dataflow unexpose` command makes a dataflow unavailable for execution as either a service or as a job.

### Usage

```
dataflow unexpose --d DataflowName
```

| Required | Argument                | Description  |
|----------|-------------------------|--|
| Yes      | --d <i>DataflowName</i> | Specifies the name of the dataflow you want to unexpose. If the dataflow name contains spaces, enclose the name in quotes.<br><br><b>Tip:</b> If you are unsure of the exact dataflow name you can use the <code>dataflow list</code> command to get a list of the dataflow names. |

#### Example

This example unexposes the dataflow named "My Dataflow".

```
dataflow unexpose --d "My Dataflow"
```

## Enterprise Routing Module

### ermdb list

**Note:** For instructions on installing and running the Administration Utility, see [Getting Started with the Administration Utility](#) on page 234.

The `ermdb list` command retrieves a list of all the existing routing database resource on the server. You must have the Enterprise Routing Module installed to use this command.

### Usage

```
ermdb list
```

#### Example

This example returns all the database resources on the server.

```
ermdb list
```

## ermdb get

**Note:** For instructions on installing and running the Administration Utility, see [Getting Started with the Administration Utility](#) on page 234.

The `ermdb get` command allows you to return information on the routing databases configured on the server. Information returned is the name of the database, location of the database on the file system (path), and the pool size configured for the database. You must have the Enterprise Routing Module installed to use this command.

### Usage

```
ermdb get --name database_name
```

**Note:** To see a list of parameters, type `help ermdb get`.

| Required | Argument   | Description  |
|----------|--|--|
| Yes      | <code>--name</code> or <code>--n</code> <i>database_name</i> | Specifies the name of the database resource to return information. The name must be a unique name on the server. For a list of existing routing database resources, use the <code>ermdb list</code> command. |

### Example

This example returns the information for the database resources US from the server.

```
ermdb get --name US
```

## ermdb add

**Note:** For instructions on installing and running the Administration Utility, see [Getting Started with the Administration Utility](#) on page 234.

The `ermdb add` command creates a new routing database resource on the server. You must have the Enterprise Routing Module installed to use this command.

**Note:** The `ermdb add` command requires a unique name be used for each of the databases being added.

### Usage

```
ermdb add --name database_name --poolsize pool_size --path database_path
```

**Note:** To see a list of parameters, type `help ermdb add`.

| Required | Argument   | Description  |
|----------|--|--|
| Yes      | <code>--name</code> or <code>--n <i>database_name</i></code> | Specifies the name of the database resource to be added. The name must be a unique name on the server. For a list of existing routing database resources, use the <code>ermdb list</code> command. |
| No       | <code>--poolsize</code> or <code>--s <i>pool_size</i></code> | Indicates the maximum number of concurrent requests the database should handle. The default if not specified is 4. The accepted range for concurrent requests is any integer between 1 and 128.    |
| YES      | <code>--path <i>database_path</i></code>                     | Specifies the location of the routing database on the file server.   |

#### Example

This example adds the database resources US from `E:/ERM-US/2014.09/driving/south` into the server.

```
ermdb add --name US --poolsize 10 --path E:/ERM-US/2014.09/driving/south
```

## ermdb delete

**Note:** For instructions on installing and running the Administration Utility, see [Getting Started with the Administration Utility](#) on page 234.

The `ermdb delete` command removes an existing routing database resource from the server. You must have the Enterprise Routing Module installed to use this command.

#### Usage

```
ermdb delete --name database_name
```

**Note:** To see a list of parameters, type `help ermdb delete`.

| Required | Argument   | Description  |
|----------|--|--|
| Yes      | <code>--name</code> or <code>--n <i>database_name</i></code> | Specifies the name of the database resource to be deleted. For a list of existing routing database resources, use the <code>ermdb list</code> command. |

**Example**

This example removes the database resources US from the server.

```
ermdb delete --name US
```

## ermdb modify

**Note:** For instructions on installing and running the Administration Utility, see [Getting Started with the Administration Utility](#) on page 234.

The `ermdb modify` command changes an existing routing database resource on the server. You must have the Enterprise Routing Module installed to use this command.

### Usage

```
ermdb modify --name database_name --poolsize pool_size --path database_path
```

**Note:** To see a list of parameters, type `help ermdb modify`.

| Required | Argument   | Description  |
|----------|--|--|
| Yes      | <code>--name</code> or <code>--n</code> <i>database_name</i> | Specifies the name of the database resource to be modified. For a list of existing routing database resources, use the <code>ermdb list</code> command.  |
| No       | <code>--poolsize</code> or <code>--s</code> <i>pool_size</i> | Indicates the maximum number of concurrent requests the database should handle. The accepted range for concurrent requests is any integer between 1 and 128. You must specify either a new pool size or a new database path. |
| No       | <code>--path</code> <i>database_path</i>                     | Specifies the new location of the routing database on the file server. You must specify either a new pool size or a new database path.   |

**Example**

This example modifies both the pool size and the database path for a new vintage.

```
ermdb modify --name US --poolsize 20 --path E:  
/ERM-US/2015.03/driving/south
```

## ermdb import

**Note:** For instructions on installing and running the Administration Utility, see [Getting Started with the Administration Utility](#) on page 234.

The `ermdb import` command allows you to import a file consisting of routing database configurations and creates the database resources on the server. You can either create the import file, or use the file created by the `ermdb export` command. You must have the Enterprise Routing Module installed to use this command.

The import file format is as follows:

```
[ { "product": "Spatial", "module": "routing", "name": "US", "maxActive": 4, "properties": {
  "DatasetPaths": "E:/ERM-US/2014.09/driving/northeast" } } ]
```

Where `product` and `module` must be `Spatial` and `routing`, `name` is the name of the database, `maxActive` is the maximum number of concurrent requests you want this database to handle (or the pool size), and `DatasetPaths` is the path to the data sets for the database resource.

You can add multiple databases in an import file (duplicate the example above), and add multiple datasets for each database resource separating them using semi colons.

**Note:** If you want to specify UTF-8 characters in import file, you must add the JVM parameter `file.encoding` to the value `UTF-8` in the startup of the CLI command prompt. E.g.,  
`-Dfile.encoding=UTF-8`

### Usage

```
ermdb import --file file_name
```

**Note:** To see a list of parameters, type `help ermdb import`.

| Required | Argument   | Description  |
|----------|--|--|
| YES      | <code>--file</code> or <code>--f</code> <i>file_name</i> | Specifies the directory and name of the import file. |

### Example

This example imports two databases US1 and US2 each consisting of multiple datasets.

```
ermdb import --file E:/ERM-US/export/ermDbResource.txt
```

The input file is defined as the following:

```
[ { "product": "Spatial", "module": "routing", "name": "US1", "maxActive": 4, "properties": {
  "DatasetPaths":
  "E:/ERM-US/2014.09/driving/northeast;E:/ERM-US/2014.09/driving/south" } }, {
"product": "Spatial", "module": "routing", "name": "US2", "maxActive": 4, "properties":
```

```
{ "DatasetPaths":
  "E:/ERM-US/2014.09/driving/northwest;E:/ERM-US/2014.09/driving/central" } }
```

## ermdb export

**Note:** For instructions on installing and running the Administration Utility, see [Getting Started with the Administration Utility](#) on page 234.

The `ermdb export` command allows you to export the routing databases configured on the server to a file. This file can then be used to import into another instance using the `ermdb import` command, either as a backup, or for migration from one instance to another. You must have the Enterprise Routing Module installed to use this command.

**Note:** The `ermdb export` command will always create an export filename name `ermDbResource.txt`

### Usage

```
ermdb export --directory directory_name
```

**Note:** To see a list of parameters, type `help ermdb export`.

| Required | Argument  | Description   |
|----------|---|---|
| No       | <code>--directory</code> or <code>--o</code><br><i>directory_name</i> | Specifies the name of the directory on the file system where to export the database file. The export command will always create an export filename name <code>ermDbResource.txt</code> . If this parameter is not specified, the export file will be created in the directory where the export command is being executed. |

### Example

This example creates an export database file in the `E:/ERM-US/export` directory.

```
ermdb export --directory E:/ERM-US/export
```

## erm getpointdata

**Note:** For instructions on installing and running the Administration Utility, see [Getting Started with the Administration Utility](#) on page 234.



The `erm getpointdata` command returns segments information for a point. The closest segment(s) is returned to the specified point. Types of information returned are; segment ID, road type, length, speed, direction, time, road name, etc. You must have the Enterprise Routing Module installed to use this command.

### Usage

```
erm getpointdata --datasource db_resource --point "x,y,coordsys"
```

**Note:** To see a list of parameters, type `help erm getpointdata`.

| Required | Argument                                     | Description  |
|----------|--|--|
| Yes      | <code>--datasource <i>db_resource</i></code> | Specifies the name of the database resource to return data. For a list of existing routing database resources, use the <code>ermdb list</code> command.                          |
| Yes      | <code>--point "x,y,coordsys"</code>          | Indicates the point to return the closest segment information. The point is specified in the format "x,y,coordsys", where <i>coordsys</i> is the coordinate system of the point. |

### Example

This example returns the closest segment data to the specified point from the US\_NE database resources configured on the server.

```
erm getpointdata --datasource US_NE --point "-72,40,epsg:4326"
```

## erm getsegmentdata

**Note:** For instructions on installing and running the Administration Utility, see [Getting Started with the Administration Utility](#) on page 234.

The `erm getsegmentdata` command returns segments information for a given segment ID. Types of information returned are; segment ID, road type, length, speed, direction, time, road name, etc. You must have the Enterprise Routing Module installed to use this command.

### Usage

```
erm getsegmentdata --datasource db_resource --segmentid "segment_id"
```

**Note:** To see a list of parameters, type `help erm getsegmentdata`.

| Required | Argument                                     | Description   |
|----------|--|---|
| Yes      | <code>--datasource <i>db_resource</i></code> | Specifies the name of the database resource to return data. For a list of existing routing database resources, use the <code>ermdb list</code> command.       |
| Yes      | <code>--segmentid "<i>segment_id</i>"</code> | Indicates the segment to return the information. The segment is specified in the format specified in the data. For example, " <code>7e3396fc:6e5251</code> ". |

**Example**

This example returns data for the specified segment from the US\_NE database resources configured on the server.

```
erm getpointdata --datasource US_NE --segmentid
"7e3396fc:6e5251"
```

## erm createpointupdate

**Note:** For instructions on installing and running the Administration Utility, see [Getting Started with the Administration Utility](#) on page 234.

The `erm createpointupdate` command overrides the routing data of the closest segment for a given point. This command allows you to set or change the speed, or exclude a section of the route. You must have the Enterprise Routing Module installed to use this command.

**Note:** The type of persistent update is valid only for the specified data resource and may not be valid after a data update.

**Usage**

```
erm createpointupdate --datasource db_resource --point "x,y,coordsys" --exclude
--velocity velocity_value --velocityunit velocity_unit --velocityadjustment
velocity_adjustment_value --velocitypercentage velocity_percentage_value
```

**Note:** To see a list of parameters, type `help erm createpointupdate`.

| Required | Argument                                     | Description   |
|----------|--|---|
| Yes      | <code>--datasource <i>db_resource</i></code> | Specifies the name of the database resource to override the data. For a list of existing routing database resources, use the <code>ermdb list</code> command. |

| Required | Argument   | Description  |
|----------|--|--|
| Yes      | <code>--point "x,y,coordsys"</code>                                | Indicates the point to override the closest segment information. The point is specified in the format "x,y,coordsys", where <i>coordsys</i> is the coordinate system of the point.   |
| No       | <code>--exclude</code>   | Excludes the specified point from all route calculations. Having this parameter in the command specifies whether to exclude the point.   |
| No       | <code>--velocity <i>velocity_value</i></code>                      | Defines a speed update where you specify the new speed of the point by specifying the new velocity. The default unit is mph(miles per hour) unless you specify the <code>velocityunit</code> parameter.  |
| No       | <code>--velocityunit <i>velocity_unit</i></code>                   | Defines a unit of speed for the <code>velocity</code> or <code>velocityadjustment</code> overrides. The default value is mph(miles per hour). For speed updates, the velocity unit can have one of the following values: kph (kilometers per hour), mps(meters per second), or mph (miles per hour).               |
| No       | <code>--velocityadjustment <i>velocity_adjustment_value</i></code> | Defines a speed update where you define a change in the speed of the point by specifying the change in velocity (unit and value). Speed values can be increased (positive value) or decreased(negative value). The default unit is mph(miles per hour) unless you specify the <code>velocityunit</code> parameter. |
| No       | <code>--velocitypercentage <i>velocity_percentage_value</i></code> | Defines a speed update where you define an increase in the speed of the point by specifying a percentage to increase(positive value) or decrease(negative value) the speed.  |

### Examples

This example overrides the speed of the point to 15 mph, from the US\_NE database resources configured on the server.

```
erm createpointupdate --datasource US_NE --point
"-72,40,epsg:4326" --velocity 15 --velocityunit mph
```

This example excludes the specified point from the US\_NE database resources configured on the server.

```
erm createpointupdate --datasource US_NE --point
"-72,40,epsg:4326" --exclude
```

This example overrides the speed of the point by increasing the speed by 45 kph, from the US\_NE database resources configured on the server.

```
erm createpointupdate --datasource US_NE --point
"-72,40,epsg:4326" --velocityadjustment 45 velocityunit kph
```

This example overrides the speed of the point by decreasing the speed by 60 percent, from the US\_NE database resources configured on the server.

```
erm createpointupdate --datasource US_NE --point
"-72,40,epsg:4326" --velocitypercentage -60
```

## erm resetpointupdate

**Note:** For instructions on installing and running the Administration Utility, see [Getting Started with the Administration Utility](#) on page 234.

The `erm resetpointupdate` command returns any overrides to the original state of the data. You must have the Enterprise Routing Module installed to use this command.

### Usage

```
erm resetpointupdate --datasource db_resource --point "x,y,coordsys" --resettype reset_type
```

**Note:** To see a list of parameters, type `help erm resetpointupdate`.

| Required       | Argument                                     | Description  |              |                         |                |                            |
|----------------|--|--|--------------|-------------------------|----------------|----------------------------|
| Yes            | <code>--datasource <i>db_resource</i></code> | Specifies the name of the database resource that has the overrides. For a list of existing routing database resources, use the <code>ermdb list</code> command.  |              |                         |                |                            |
| Yes            | <code>--point "<i>x,y,coordsys</i>"</code>   | Indicates the point where the existing overrides are located. The point is specified in the format " <i>x,y,coordsys</i> ", where <i>coordsys</i> is the coordinate system of the point.   |              |                         |                |                            |
| Yes            | <code>--resettype <i>reset_type</i></code>   | The type of override to remove (undo). <table border="0" style="margin-left: 20px;"> <tr> <td><b>speed</b></td> <td>Removes a speed update.</td> </tr> <tr> <td><b>exclude</b></td> <td>Removes an exclude update.</td> </tr> </table> | <b>speed</b> | Removes a speed update. | <b>exclude</b> | Removes an exclude update. |
| <b>speed</b>   | Removes a speed update.                      |  |              |                         |                |                            |
| <b>exclude</b> | Removes an exclude update.                   |  |              |                         |                |                            |

**Example**

This example resets an existing exclude override for the given point, from the US\_NE database resources configured on the server.

```
erm resetpointupdate --datasource US_NE --point
"-72,40,epsg:4326" --resettype exclude
```

## erm createsegmentupdate

**Note:** For instructions on installing and running the Administration Utility, see [Getting Started with the Administration Utility](#) on page 234.

The `erm createsegmentupdate` command overrides the routing data of the specified segment. This command allows you to set or change the speed, exclude a section of the route, or change the road type. You must have the Enterprise Routing Module installed to use this command.

**Note:** The type of persistent update is valid only for the specified data resource and may not be valid after a data update.

### Usage

```
erm createsegmentupdate --datasource db_resource --segmentid "segment_id"
--exclude --velocity velocity_value --velocityunit velocity_unit --velocityadjustment
velocity_adjustment_value --velocitypercentage velocity_percentage_value --roadtype
road_type
```

**Note:** To see a list of parameters, type `help erm createsegmentupdate`.

| Required | Argument                                     | Description   |
|----------|--|---|
| Yes      | <code>--datasource <i>db_resource</i></code> | Specifies the name of the database resource to override the data. For a list of existing routing database resources, use the <code>ermdb list</code> command. |
| Yes      | <code>--segmentid "<i>segment_id</i>"</code> | Indicates the segment to override. The segment is specified in the format specified in the data. For example, " <code>7e3396fc:6e5251</code> ".               |
| No       | <code>--exclude</code>                       | Excludes the specified segment from all route calculations. Having this parameter in the command specifies whether to exclude the segment.                    |

| Required | Argument   | Description  |
|----------|--|--|
| No       | <code>--velocity <i>velocity_value</i></code>                      | Defines a speed update where you specify the new speed of the segment by specifying the new velocity. The default unit is mph(miles per hour) unless you specify the <code>velocityunit</code> parameter.  |
| No       | <code>--velocityunit <i>velocity_unit</i></code>                   | Defines a unit of speed for the <code>velocity</code> or <code>velocityadjustment</code> overrides. The default value is mph(miles per hour). For speed updates, the velocity unit can have one of the following values: kph (kilometers per hour), mps(meters per second), or mph (miles per hour).                 |
| No       | <code>--velocityadjustment <i>velocity_adjustment_value</i></code> | Defines a speed update where you define a change in the speed of the segment by specifying the change in velocity (unit and value). Speed values can be increased (positive value) or decreased(negative value). The default unit is mph(miles per hour) unless you specify the <code>velocityunit</code> parameter. |
| No       | <code>--velocitypercentage <i>velocity_percentage_value</i></code> | Defines a speed update where you define an increase in the speed of the segment by specifying a percentage to increase(positive value) or decrease(negative value) the speed.  |
| No       | <code>--roadtype <i>road_type</i></code>                           | Defines the new road type for the segment.   |

### Examples

This example overrides the speed of the segment to 15 mph, from the US\_NE database resources configured on the server.

```
erm createsegmentupdate --datasource US_NE --segmentid
"7e3396fc:6e5251" --velocity 15 --velocityunit mph
```

This example excludes the specified segment from the US\_NE database resources configured on the server.

```
erm createsegmentupdate --datasource US_NE --segmentid
"7e3396fc:6e5251" --exclude
```

This example overrides the speed of the segment by increasing the speed by 45 kph, from the US\_NE database resources configured on the server.

```
erm createsegmentupdate --datasource US_NE --segmentid
"7e3396fc:6e5251" --velocityadjustment 45 velocityunit kph
```

This example overrides the speed of the segment by decreasing the speed by 60 percent, from the US\_NE database resources configured on the server.

```
erm createsegmentupdate --datasource US_NE --segmentid
"7e3396fc:6e5251" --velocitypercentage -60
```

This example overrides the road type of the segment to ferry, from the US\_NE database resources configured on the server.

```
erm createsegmentupdate --datasource US_NE --segmentid
"7e3396fc:6e5251" --roadtype ferry
```

## erm resetsegmentupdate

**Note:** For instructions on installing and running the Administration Utility, see [Getting Started with the Administration Utility](#) on page 234.

The `erm resetsegmentupdate` command returns any overrides to the original state of the data. You must have the Enterprise Routing Module installed to use this command.

### Usage

```
erm resetsegmentupdate --datasource db_resource --segmentid "segment_id"
--resettype reset_type
```

**Note:** To see a list of parameters, type `help erm resetsegmentupdate`.

| Required | Argument                                     | Description   |
|----------|--|---|
| Yes      | <code>--datasource <i>db_resource</i></code> | Specifies the name of the database resource that has the overrides. For a list of existing routing database resources, use the <code>ermdb list</code> command.   |
| Yes      | <code>--segment "<i>segment_id</i>"</code>   | Indicates the segment where the existing overrides are located. The segment is specified in the format specified in the data. For example, <code>"7e3396fc:6e5251"</code> .   |
| Yes      | <code>--resettype <i>reset_type</i></code>   | The type of override to remove (undo). <ul style="list-style-type: none"> <li><b>speed</b>                Removes a speed update.</li> <li><b>exclude</b>            Removes an exclude update.</li> <li><b>roadtype</b>           Removes a road type update.</li> </ul> |

**Example**

This example resets an existing road type override for the given segment, from the US\_NE database resources configured on the server.

```
erm resetpointupdate --datasource US_NE --point
"7e3396fc:6e5251" --resettype roadtype
```

## erm getsegmentupdates

**Note:** For instructions on installing and running the Administration Utility, see [Getting Started with the Administration Utility](#) on page 234.

The `erm getsegmentupdates` command returns a list of overrides in the routing data for the specified segment(s). You must have the Enterprise Routing Module installed to use this command.

**Note:** `segmentids` is an optional parameter. If no segment ids are specified, then overrides for all available segments are returned.

### Usage

```
erm getsegmentupdates --datasource db_resource --segmentids "segment_ids"
```

**Note:** To see a list of parameters, type `help erm getsegmentupdates`.

| Required | Argument                                       | Description   |
|----------|--|---|
| Yes      | <code>--datasource <i>db_resource</i></code>   | Specifies the name of the database resource that has overrides. For a list of existing routing database resources, use the <code>ermdb list</code> command.                     |
| No       | <code>--segmentids "<i>segment_ids</i>"</code> | A comma separated list of segment ids to return override information. Segments are specified in the format specified in the data. For example, <code>"7e3396fc:6e5251"</code> . |

**Example**

This example returns the overrides for a segment, from the US\_NE database resources configured on the server.

```
erm getsegmentupdates --datasource US_NE --segmentids
"7e3396fc:6e5251"
```



## erm createroadtypeupdate

**Note:** For instructions on installing and running the Administration Utility, see [Getting Started with the Administration Utility](#) on page 234.

The `erm createroadtypeupdate` command overrides the routing data of the specified road type. This command allows you to set or change the speed of the route for the particular road type. You must have the Enterprise Routing Module installed to use this command.

**Note:** The type of persistent update is valid only for the specified data resource and may not be valid after a data update.

### Usage

```
erm createroadtypeupdate --datasource db_resource --roadtype "road_type"
--velocity velocity_value --velocityunit velocity_unit --velocityadjustment
velocity_adjustment_value --velocitypercentage velocity_percentage_value --roadtype
road_type
```

**Note:** To see a list of parameters, type `help erm createroadtypeupdate`.

| Required | Argument                                     | Description   |
|----------|--|---|
| Yes      | <code>--datasource <i>db_resource</i></code> | Specifies the name of the database resource to override the data. For a list of existing routing database resources, use the <code>ermdb list</code> command.   |
| Yes      | <code>--roadtype "<i>road_type</i>"</code>   | Indicates the road type to override. The road type can be one of the following: <ul style="list-style-type: none"> <li>• access way</li> <li>• back road</li> <li>• connector</li> <li>• ferry</li> <li>• footpath</li> <li>• limited access dense urban</li> <li>• limited access rural</li> <li>• limited access suburban</li> <li>• limited access urban</li> <li>• local road dense urban</li> <li>• local road rural</li> <li>• local road suburban</li> <li>• local road urban</li> </ul> |

| Required Argument                                   | Description  |
|---|--|
|   | <ul style="list-style-type: none"> <li>• major local road dense urban</li> <li>• major local road rural</li> <li>• major local road suburban</li> <li>• major local road urban</li> <li>• major road dense urban</li> <li>• major road rural</li> <li>• major road suburban</li> <li>• major road urban</li> <li>• minor local road dense Urban</li> <li>• minor local road rural</li> <li>• minor local road suburban</li> <li>• minor local road urban</li> <li>• normal road dense urban</li> <li>• normal road rural</li> <li>• normal road rural</li> <li>• normal road urban</li> <li>• primary highway dense urban</li> <li>• primary highway rural</li> <li>• primary highway suburban</li> <li>• primary highway urban</li> <li>• ramp dense urban</li> <li>• ramp limited access</li> <li>• ramp major road</li> <li>• ramp primary highway</li> <li>• ramp rural</li> <li>• ramp secondary highway</li> <li>• ramp urban</li> <li>• ramp suburban</li> <li>• secondary highway dense urban</li> <li>• secondary highway rural</li> <li>• secondary highway suburban</li> <li>• secondary highway urban</li> </ul> |
| No <code>--velocity <i>velocity_value</i></code>    | Defines a speed update where you specify the new speed of the road type by specifying the new velocity. The default unit is mph(miles per hour) unless you specify the <code>velocityunit</code> parameter.  |
| No <code>--velocityunit <i>velocity_unit</i></code> | Defines a unit of speed for the <code>velocity</code> or <code>velocityadjustment</code> overrides. The default value is mph(miles per hour). For speed updates,   |

| Required Argument   | Description   |
|---|---|
| No<br><code>--velocityadjustment</code><br><code>velocity_adjustment_value</code> | the velocity unit can have one of the following values: kph (kilometers per hour), mps(meters per second), or mph (miles per hour).<br><br>Defines a speed update where you define a change in the speed of the road type by specifying the change in velocity (unit and value). Speed values can be increased (positive value) or decreased(negative value). The default unit is mph(miles per hour) unless you specify the <code>velocityunit</code> parameter. |
| No<br><code>--velocitypercentage</code><br><code>velocity_percentage_value</code> | Defines a speed update where you define an increase in the speed of the road type by specifying a percentage to increase(positive value) or decrease(negative value) the speed.   |

### Examples

This example overrides the speed of a road type to 25 kph, from the US\_NE database resources configured on the server.

```
erm createsegmentupdate --datasource US_NE --roadtype "normal
road suburban" --velocity 25 --velocityunit kph
```

This example increases the speed of the specified road type by 50 kph, from the US\_NE database resources configured on the server.

```
erm createsegmentupdate --datasource US_NE --roadtype "normal
road suburban" --velocityadjustment 50 --velocityunit mph
```

This example overrides the speed of the road type by decreasing the speed by 65 percent, from the US\_NE database resources configured on the server.

```
erm createsegmentupdate --datasource US_NE --roadtype "normal
road suburban" --velocitypercentage -65
```

## erm resetroadtypeupdate

**Note:** For instructions on installing and running the Administration Utility, see [Getting Started with the Administration Utility](#) on page 234.

The `erm resetroadtypeupdate` command returns any overrides to the original state of the data. You must have the Enterprise Routing Module installed to use this command.

### Usage

```
erm resetroadtypeupdate --datasource db_resource --roadtype "road_type"
```

**Note:** To see a list of parameters, type `help erm resetroadtypeupdate`.

| Required | Argument                                     | Description   |
|----------|--|---|
| Yes      | <code>--datasource <i>db_resource</i></code> | Specifies the name of the database resource that has the overrides. For a list of existing routing database resources, use the <code>ermdb list</code> command. |
| Yes      | <code>--roadtype "<i>road_type</i>"</code>   | Indicates the road type that has the existing overrides. For a list of road types, see <a href="#">erm createroadtypeupdate</a> on page 265.                    |

#### Example

This example resets the "normal road suburban" road type override, from the US\_NE database resources configured on the server.

```
erm resetpointupdate --datasource US_NE --roadtype "normal road suburban"
```

## erm getroadtypeupdates

**Note:** For instructions on installing and running the Administration Utility, see [Getting Started with the Administration Utility](#) on page 234.

The `erm getroadtypeupdates` command returns a list of overrides in the routing data for the specified road type(s). You must have the Enterprise Routing Module installed to use this command.

**Note:** `roadtypes` is an optional parameter. If no road types are specified, then overrides for all available road types are returned.

### Usage

```
erm getsegmentupdates --datasource db_resource --roadtypes "road_types"
```

**Note:** To see a list of parameters, type `help erm getsegmentupdates`.

| Required | Argument                                     | Description   |
|----------|--|---|
| Yes      | <code>--datasource <i>db_resource</i></code> | Specifies the name of the database resource that has overrides. For a list of existing routing database resources, use the <code>ermdb list</code> command. |

| Required | Argument                              | Description  |
|----------|---------------------------------------|--|
| No       | <code>--roadtypes "road_types"</code> | A comma separated list of road types to return override information. For a list of road types, see <a href="#">erm createroadtypeupdate</a> on page 265. |

**Example**

This example returns the overrides for the "normal road urban" road type, from the US\_NE database resources configured on the server.

```
erm getsegmentupdates --datasource US_NE --roadtypes "normal
road urban"
```

## erm getallupdates

**Note:** For instructions on installing and running the Administration Utility, see [Getting Started with the Administration Utility](#) on page 234.

The `erm getallupdates` command returns a list of overrides for a specified routing database resource. You must have the Enterprise Routing Module installed to use this command.

*Usage*

```
erm getsegmentupdates --datasource db_resource --segmentids "segment_ids"
```

**Note:** To see a list of parameters, type `help erm getallupdates`.

| Required | Argument                                     | Description   |
|----------|--|---|
| Yes      | <code>--datasource <i>db_resource</i></code> | Specifies the name of the database resource that has the overrides. For a list of existing routing database resources, use the <code>ermdb list</code> command. |

**Example**

This example returns all the overrides from the US\_NE database resources configured on the server.

```
erm getallupdates --datasource US_NE
```

## erm resetallupdates

**Note:** For instructions on installing and running the Administration Utility, see [Getting Started with the Administration Utility](#) on page 234.

The `erm resetallupdates` command returns all overrides to the original state of the data. You must have the Enterprise Routing Module installed to use this command.

### Usage

```
erm resetallupdates --datasource db_resource
```

**Note:** To see a list of parameters, type `help erm resetallupdates`.

| Required | Argument                                     | Description   |
|----------|--|---|
| Yes      | <code>--datasource <i>db_resource</i></code> | Specifies the name of the database resource that has the overrides. For a list of existing routing database resources, use the <code>ermdb list</code> command. |

### Example

This example resets all overrides from the US\_NE database resources configured on the server.

```
erm resetallupdates --datasource US_NE
```

## Folders

### folder browse

The `folder browse` command lists the contents of a Server Explorer folder.

### Usage

```
folder browse --p Path
```

| Required | Argument              | Description  |
|----------|-----------------------|--|
| No       | <code>--p Path</code> | Specifies the folder whose contents you want to list. If you omit this parameter the contents of the root folder are listed. |

## folder create

The `folder create` command creates a folder in Server Explorer.

### Usage

```
folder create --p Path
```

| Required | Argument              | Description  |
|----------|-----------------------|--|
| Yes      | <code>--p Path</code> | Specifies the path of the folder you want to create. |

### Example

This example creates a folder named `Example123` inside the folder `ExampleABC`.

```
folder create --p ExampleABC/Example123
```

## folder delete

The `folder delete` command deletes a folder from Server Explorer.

### Usage

```
folder delete --p Path --r TrueFalse
```

| Required | Argument                   | Description  |
|----------|----------------------------|--|
| Yes      | <code>--p Path</code>      | Specifies the path of the folder you want to delete.   |
| No       | <code>--r TrueFalse</code> | Specifies whether to delete the folder if it contains a flow or subfolders.<br><b>true</b><br>Delete the folder if it contains flows or subfolders.<br><b>false</b><br>Do not delete the folder if it contains flows or subfolders. Default. |

**Example**

This example deletes a folder named `Example123`. The folder will be deleted even if it contains flows or subfolders.

```
folder delete --p ExampleABC/Example123 --r true
```

## folder move

The `folder move` command moves a folder in Server Explorer to another location.

*Usage*

```
folder move --p Path --t Target
```

| Required | Argument          | Description  |
|----------|-------------------|--|
| Yes      | --p <i>Path</i>   | Specifies the path of the folder you want to move.       |
| Yes      | --t <i>Target</i> | Specifies the path to which you want to move the folder. |

**Example**

This example moves a folder named `ExampleABC` into the folder `Example123`.

```
folder move --p ExampleABC --t Example123
```

## folder rename

The `folder rename` command changes the name of a folder in Server Explorer.

*Usage*

```
folder rename --p Path --n NewName
```

| Required | Argument           | Description  |
|----------|--------------------|--|
| Yes      | --p <i>Path</i>    | Specifies the path of the folder you want to rename. |
| Yes      | --n <i>NewName</i> | Specifies the new name of the folder.                |



# Jobs

## job history list

The `job history list` command shows the execution history for a job.

### Usage

```
job status list --j JobName --f FromDateTime --t ToDateTime
```

| Required | Argument                | Description  |
|----------|-------------------------|--|
| Yes      | --j <i>JobName</i>      | Specifies the name of the job whose history you want to get. If the job name contains spaces, enclose the name in quotes.<br><br><b>Tip:</b> If you are unsure of the exact dataflow name you can use the <code>dataflow list</code> command to get a list of the dataflow names.  |
| No       | --f <i>FromDateTime</i> | If you want to see the history for a specific date and time range, specify the starting date and time for the range, in the format MM-dd-yyyy HH:mm:ss. For example, December 31, 2014 1:00 PM would be specified as <code>12-31-2014 13:00:00</code> .<br><br>When you specify a date and time range, the history list will include jobs that started execution on or after the date you specified in the --f argument and before the date you specify in the --t argument.<br><br>If you omit this argument the history will include jobs that started execution on the current date.                                  |
| No       | --t <i>ToDateTime</i>   | If you want to see the history for a specific date and time range, specify the ending date and time for the range, in the format MM-dd-yyyy HH:mm:ss. For example, December 31, 2014 1:00 PM would be specified as <code>12-31-2014 13:00:00</code> .<br><br>When you specify a date and time range, the history list will include jobs that started execution on or after the date you specified in the --f argument and before the date you specify in the --t argument.<br><br>If you omit this argument the history will include all jobs that started execution on or after the date specified in the --f argument. |

**Example**

This example gets the status of the job named "My Job".

```
job history list --j "My Job"
```

## job execute

The `job execute` command runs one or more jobs. After the job runs, the job name and job ID are returned in the format:

```
<JobName=JobID>
```

### Usage

```
job execute --j JobNames --f JobPropertyFile --i PollInterval --d ReportDelimiter --n NotificationEmails --o OptionPropertyFile --r ReportTrueOrFalse --t Timeout --w WaitTrueOrFalse --l FileOverrides --v VerboseTrueOrFalse
```

| Required | Argument                                   | Description  |
|----------|--|--|
| Yes      | <code>--j <i>JobNames</i></code>           | Specifies the name of one or more jobs to run. If you specify more than one job, separate each job with a comma. Jobs run in the order you list them. If the job name contains spaces, enclose the name in quotes.<br><br><b>Tip:</b> If you are unsure of the exact dataflow name you can use the <code>dataflow list</code> command to get a list of the dataflow names. |
| No       | <code>--f <i>JobPropertyFile</i></code>    | Specifies the path to a job property file. A job property file contains arguments that control the execution of jobs. For more information, see <a href="#">Using a Job Property File</a> on page 179.   |
| No       | <code>--i <i>PollInterval</i></code>       | When <code>--w</code> is set to <code>true</code> , use this option to specify how often to check for completed jobs, in seconds. The default is 5.  |
| No       | <code>--d <i>ReportDelimiter</i></code>    | Sets the delimiter character to use in the report output when you also specify <code>--w true</code> or <code>--r true</code> . The default is the pipe character ( <code> </code> ).  |
| No       | <code>--n <i>NotificationEmails</i></code> | Specifies one or more email addresses to receive notifications about the status of jobs, as configured in Management Console. Separate each email address with a comma.  |
| No       | <code>--o <i>OptionPropertyFile</i></code> | Specifies a path to a dataflow options property file. Use a dataflow options property file to set options for stages in the dataflow. In order to set dataflow options using a property file, you must configure the dataflow to expose stage options at   |

| Required Argument   | Description  |
|---|--|
|   | <p>runtime. For more information, see <a href="#">Adding Dataflow Runtime Options</a>.</p> <p>For example, a dataflow options properties file for a dataflow that contains an Assign GeoTAX Info stage may look like this:</p>   |
|   | <pre data-bbox="636 472 1421 619">OutputCasing=U UseStreetLevelMatching=N TaxKey=T Database.GTX=gsl</pre>  |
| <p>No      <code>--r <i>ReportTrueOrFalse</i></code></p>  | <p>Specify <code>true</code> to return a detailed report about the job. This option only works if you also specify <code>--w true</code>. The report contains the following information:</p> <ul style="list-style-type: none"> <li>• <b>Position 1</b>—Name of job</li> <li>• <b>Position 2</b>—Job process ID</li> <li>• <b>Position 3</b>—Status</li> <li>• <b>Position 4</b>—Start Date/Time (MM/DD/YYYY HH:MM:SS)</li> <li>• <b>Position 5</b>—End Date/Time (MM/DD/YYYY HH:MM:SS)</li> <li>• <b>Position 6</b>—Number of successful records</li> <li>• <b>Position 7</b>—Number of failed records</li> <li>• <b>Position 8</b>—Number of malformed records</li> <li>• <b>Position 9</b>—Currently unused</li> </ul> <p>For example,</p> <pre data-bbox="636 1207 1421 1281">MySimpleJob 4 succeeded 04/09/2010 14:50:47 04/09/2010 14:50:47 100 0 0 </pre> <p>The information is delimited using the delimiter specified in the <code>--d</code> argument.</p> |
| <p>No      <code>--t <i>Timeout</i></code></p>            | <p>Sets the timeout value for synchronous mode, in seconds. The default is 3600.</p>   |
| <p>No      <code>--w <i>WaitTrueOrFalse</i></code></p>    | <p>Specify <code>true</code> to run jobs one at a time in synchronous mode. Specify <code>false</code> to run all the jobs at the same time. The default is false.</p>   |
| <p>No      <code>--l <i>FileOverrides</i></code></p>      | <p>Overrides the input and output file and file format. For more information see <a href="#">Overriding Job Files</a> on page 278 and <a href="#">Overriding File Format</a> on page 280.</p>  |
| <p>No      <code>--v <i>VerboseTrueOrFalse</i></code></p> | <p>Specify <code>true</code> to return information about the arguments used to run the job and other details about the job execution.</p>  |

**Example**

This example runs a job named Example1. It returns a comma-delimited report. Note that `--w true` is specified because this is required to return a report even if only one job is running. The input file specified in the Read from File stage is changed from what is specified in the stage to a different file named `CandidateHomes2.csv`. Verbose output is also returned.

```
job execute --j Example1 --w true --d "," --r true --l "Read
from
File=file:///e:/SampleDataflows/DataFiles/DataFiles/CandidateHomes2.csv"
--v true
```

**Using a Job Property File**

A job property file contains arguments that control the execution of jobs when you use the job executor or the Administration Utility to run a job. Use a job property file if you want to reuse arguments by specifying a single argument at the command line (`-f`) rather than specifying each argument individually at the command line.

To create a property file, create a text file with one argument on each line. For example:

```
d %
h spectrum.mydomain.com
i 30
j validateAddressJob1
u user
p password
s 8888
t 9999
w true
```

The job property file can contain these arguments:

| Required | Argument              | Description  |
|----------|-----------------------|--|
| No       | ?                     | Prints usage information.  |
| No       | d <i>delimiter</i>    | Sets instance/status delimiter. This appears in synchronous output only.                                   |
| No       | e                     | Use a secure HTTPS connection for communication with the Spectrum™ Technology Platform server.             |
| No       | h <i>hostname</i>     | Specifies the name or IP address of the Spectrum™ Technology Platform server.                              |
| No       | i <i>pollinterval</i> | Specifies how often to check for completed jobs, in seconds. This applies only in synchronous mode.        |
| Yes      | j <i>jobname</i>      | A comma-separated list of jobs to run. Job names are case-sensitive. Jobs are started in the order listed. |

| Required | Argument           | Description  |
|----------|--------------------|--|
| No       | n <i>emaillist</i> | Specifies a comma-separated list of additional email addresses for configured job notifications.   |
| Yes      | p <i>password</i>  | The password of the user.  |
| No       | r                  | <p>Returns a delimited list with the following information about the job written to standard output:</p> <ul style="list-style-type: none"> <li>• <b>Position 1</b>—Name of job</li> <li>• <b>Position 2</b>—Job process ID</li> <li>• <b>Position 3</b>—Status</li> <li>• <b>Position 4</b>—Start Date/Time (MM/DD/YYYY HH:MM:SS)</li> <li>• <b>Position 5</b>—End Date/Time (MM/DD/YYYY HH:MM:SS)</li> <li>• <b>Position 6</b>—Number of successful records</li> <li>• <b>Position 7</b>—Number of failed records</li> <li>• <b>Position 8</b>—Number of malformed records</li> <li>• <b>Position 9</b>—Currently unused</li> </ul> <p>The information is delimited using the delimiter specified in the -d argument. For example:</p> <pre>MySimpleJob 4 succeeded 04/09/2010 14:50:47 04/09/2010 14:50:47 100 0 0 </pre> |
| No       | s <i>port</i>      | The socket (port) on which the Spectrum™ Technology Platform server is running. The default is 8080.   |
| No       | t <i>timeout</i>   | Sets the timeout (in seconds) for synchronous mode. The default is 3600. The maximum is 2147483. This is a global, aggregate timeout and represents the maximum time to wait for all spawned jobs to complete.   |
| Yes      | u <i>username</i>  | The login name of the user.  |
| No       | v                  | Return verbose output.   |
| No       | w                  | Specifies to wait for jobs to complete in a synchronous mode.  |

### Using Both Command Line Arguments and a Property File

A combination of both command-line entry and property file entry is also valid. For example:

```
java -jar jobexecutor.jar -f /dcg/job.properties -j job1
```

In this case command line arguments take precedence over arguments specified in the properties file. In the above example, the job job1 would take precedence over a job specified in the properties file.

## Overriding Job Files

When you run a job at the command line using job executor or the Administration Utility, you can override the input file specified in the dataflow's source stage (such as Read from File), as well as the output file specified in the dataflow's sink stage (such as Write to File).

To do this in job executor, specify the following at the end of the job executor command line command:

```
StageName=Protocol:FileName
```

In the Administration Utility, use the `--l` argument in the `job execute` command:

```
--l StageName=Protocol:FileName
```

Where:

### **StageName**

The stage label shown under the stage's icon in the dataflow in Enterprise Designer. For example, if the stage is labeled "Read from File" you would specify `Read from File` for the stage name.

To specify a stage within an embedded dataflow or a subflow, preface the stage name with the name of the embedded dataflow or subflow, followed by a period then the stage name:

```
EmbeddedOrSubflowName.StageName
```

For example, to specify a stage named Write to File in a subflow named Subflow1, you would specify:

```
Subflow1.Write to File
```

To specify a stage in an embedded dataflow that is within another embedded dataflow, add the parent dataflow, separating each with a period. For example, if Embedded Dataflow 2 is inside Embedded Dataflow 1, and you want to specify the Write to File stage in Embedded Dataflow 2, you would specify this:

```
Embedded Dataflow 1.Embedded Dataflow 2.Write to File
```

### **Protocol**

A communication protocol. One of the following:

**file** Use the file protocol if the file is on the same machine as the Spectrum™ Technology Platform server. For example, on Windows specify:

```
"file:C:/myfile.txt"
```

On Unix or Linux specify:

```
"file:/testfiles/myfile.txt"
```

**esclient** Use the esclient protocol if the file is on the computer where you are executing the job if it is a different computer from the one running the Spectrum™ Technology Platform server. Use the following format:

```
esclient:ComputerName/path to file
```

For example,

```
esclient:mycomputer/testfiles/myfile.txt
```

**Note:** If you are executing the job on the server itself, you can use either the file or esclient protocol, but are likely to have better performance using the file protocol.

If the host name of the Spectrum™ Technology Platform server cannot be resolved, you may get the error "Error occurred accessing file". To resolve this issue, open this file on the server:

```
SpectrumLocation/server/app/conf/spectrum-container.properties.
```

Set the `spectrum.runtime.hostname` property to the IP address of the server.

**esfile** Use the esfile protocol if the file is on a file server. The file server must be defined in Management Console as a resource. Use the following format:

```
esfile://file server/path to file
```

For example,

```
esfile://myserver/testfiles/myfile.txt
```

Where myserver is an FTP file server resource defined in Management Console.

**webhdfs** Use the webhdfs protocol if the file is on a Hadoop Distributed File Server. The HDFS server must be defined in Management Console as a resource. Use the following format:

```
webhdfs://file server/path to file
```

For example,

```
webhdfs://myserver/testfiles/myfile.txt
```

Where myserver is an HDFS file server resource defined in Management Console.

### **FileName**

The full path to the file you want to use as input or output.

**Note:** You must use forward slashes (/) in file paths, not backslashes.

To specify multiple overrides, separate each override with a comma.

**Example File Override**

This example executes a job named TestJob. Instead of writing the output to the file specified in the Write to File stage, it will write the output to `outputoverride.txt`.

```
job execute --j TestJob --l "Write to
File=file:/Users/me/outputoverride.txt"
```

**Overriding File Format**

When you run a job using job executor or the Administration Utility, you can override the file layout (or schema) of the file specified in the dataflow's Read from File stage and Write to File stage.

To do this in job executor, specify the following at the end of the job executor command line command:

```
StageName:schema=Protocol:SchemaFile
```

In the Administration Utility, use the `--l` argument in the `job execute` command:

```
--l StageName:schema=Protocol:SchemaFile
```

Where:

**StageName**

The stage label shown under the stage's icon in the dataflow in Enterprise Designer. For example, if the stage is labeled "Read from File" you would specify `Read from File` for the stage name.

To specify a stage within an embedded dataflow or a subflow, preface the stage name with the name of the embedded dataflow or subflow, followed by a period then the stage name:

```
EmbeddedOrSubflowName.StageName
```

For example, to specify a stage named Write to File in a subflow named Subflow1, you would specify:

```
Subflow1.Write to File
```

To specify a stage in an embedded dataflow that is within another embedded dataflow, add the parent dataflow, separating each with a period. For example, if Embedded Dataflow 2 is inside Embedded Dataflow 1, and you want to specify the Write to File stage in Embedded Dataflow 2, you would specify this:

```
Embedded Dataflow 1.Embedded Dataflow 2.Write to File
```

**Protocol**

A communication protocol. One of the following:

**file** Use the file protocol if the file is on the same machine as the Spectrum™ Technology Platform server. For example, on Windows specify:

```
"file:C:/myfile.txt"
```



On Unix or Linux specify:

```
"file:/testfiles/myfile.txt"
```

**esclient** Use the esclient protocol if the file is on the computer where you are executing the job if it is a different computer from the one running the Spectrum™ Technology Platform server. Use the following format:

```
esclient:ComputerName/path to file
```

For example,

```
esclient:mycomputer/testfiles/myfile.txt
```

**Note:** If you are executing the job on the server itself, you can use either the file or esclient protocol, but are likely to have better performance using the file protocol.

If the host name of the Spectrum™ Technology Platform server cannot be resolved, you may get the error "Error occurred accessing file". To resolve this issue, open this file on the server:

```
SpectrumLocation/server/app/conf/spectrum-container.properties.
```

Set the `spectrum.runtime.hostname` property to the IP address of the server.

**esfile** Use the esfile protocol if the file is on a file server. The file server must be defined in Management Console as a resource. Use the following format:

```
esfile://file server/path to file
```

For example,

```
esfile://myserver/testfiles/myfile.txt
```

Where myserver is an FTP file server resource defined in Management Console.

**webhdfs** Use the webhdfs protocol if the file is on a Hadoop Distributed File Server. The HDFS server must be defined in Management Console as a resource. Use the following format:

```
webhdfs://file server/path to file
```

For example,

```
webhdfs://myserver/testfiles/myfile.txt
```

Where myserver is an HDFS file server resource defined in Management Console.

### SchemaFile

The full path to the file that defines the layout you want to use.

**Note:** You must use forward slashes (/) in file paths, not backslashes.

To create a schema file, define the layout you want in Read from File or Write to File, then click the **Export** button to create an XML file that defines the layout.

**Note:** You cannot override a field's data type in a schema file when using job executor. The value in the <Type> element, which is a child of the <FieldSchema> element, must match the field's type specified in the dataflow's Read from File or Write to File stage.

#### Example File Format Override

This example executes a job named TestJob. Instead of writing the output to the file specified in the Write to File stage, it will write the output to `outputoverride.txt`. Instead of using the file schema specified in the Write to File stage in the flow, the job will use the schema specified in `output-data.xml`.

```
job execute --j TestJob --l "Write to
File=file:/Users/me/outputoverride.txt,Write to
File:schema=file:/Users/me/output-data.xml"
```

## Location Intelligence Module

### limrepo export

**Note:** For instructions on installing and running the Administration Utility, see [Getting Started with the Administration Utility](#) on page 234.

The `limrepo export` command exports named resources (such as named tables) from the Spectrum Spatial repository to a local file system. You must have the Location Intelligence Module installed to use this command.

Resources are exported with their full repository paths in the target folder. For example, if you run `limrepo export --s /Samples/NamedTables --o C:\export`, the tool creates `C:\export\Samples\NamedTables\WorldTable`, and so on for each named table under the `NamedTables` folder or directory.

**Note:** The `limrepo export` command will always recursively export all folders, including empty ones.

#### Usage

```
limrepo export --s SourceRepositoryPath --o OutputFilePath
```

**Note:** To see a list of parameters, type `help limrepo export`.

| Required | Argument   | Description  |
|----------|--|--|
| Yes      | <code>--s SourceRepositoryPath</code>              | Specifies the path to the resource or a folder to be exported.   |
| Yes      | <code>--o OutputFilePath</code>                    | Specifies the path to a folder on the local file system where you want to export. This can be a new folder or an existing folder; however, an existing folder must be empty otherwise the export will fail.  |
| No       | <code>--q</code> or <code>--quiet</code>           | Disables the display of the resources copied during the export; that is, operates in quiet mode.<br><br>If the flag is specified, the default value is true. If the flag is not specified, the default value is false.   |
| No       | <code>--f</code> or <code>--fullpaths</code>       | Prints the full source and output paths.<br><br>If the flag is specified, the default value is true. If the flag is not specified, the default value is false.   |
| No       | <code>--r</code> or <code>--recursive</code>       | Recursively exports subfolders (children of the specified source).<br><br>If the flag is specified, the default value is true. If the flag is not specified, the default value is true.  |
| No       | <code>--c</code> or <code>--continueonerror</code> | Continues with the export if an error occurs.<br><br>If the flag is specified, the default value is true. If the flag is not specified, the default value is false.  |
| No       | <code>--a</code> or <code>--acl</code>             | Preserves existing permissions for the exported resources in the export folder on the local file system. An access control list (ACL) indicates the operations each user or role can perform on a named resource, such as create, view, edit, or delete.<br><br>If the flag is specified, the default value is true. If the flag is not specified, the default value is false. |

**Example**

This example exports the named resources in the repository's \Samples folder to C:\myrepository\samples on your local file system.

```
limrepo export --s /Samples --o C:\myrepository\samples
```

## limrepo import

**Note:** For instructions on installing and running the Administration Utility, see [Getting Started with the Administration Utility](#) on page 234.

The `limrepo import` command imports named resources (such as named tables) from a local file system into the Spectrum Spatial repository. You must have the Location Intelligence Module installed to use this command.

When importing, you must point to the same folder or directory you exported to previously. For example, if you run `limrepo export --s /Samples/NamedTables --o C:\export`, the tool creates `C:\export\Samples\NamedTables\WorldTable`, and so on for each named table under the `NamedTables` folder or directory. Resources are exported with their full repository paths in the target folder. Running `limrepo import --s C:\export` then imports `WorldTable` back to `/Samples/NamedTables/WorldTable`.

**Note:** The `limrepo import` command will always recursively import all folders, including empty ones.

After performing an import, in many cases, you will need to adjust the named connections to point to their new path using Spatial Manager. For example, if your Native TAB files were installed on “C:\myfiles” in your test instance and the same files are installed on “E:\ApplicationData\Spectrum\Spatial\Spring2016” then that connection would have to be corrected in Spatial Manager after import. See the Utilities section of the *Spectrum Spatial Guide* for instructions on using Spatial Manager to edit a named connection.

**Note:** If you are using `limrepo import` to restore service configuration files that you exported from a pre-12.0 version of Spectrum™ Technology Platform, the files will automatically be modified to be compliant with version 12.0 and later (for example, the repository URLs will be removed).

### Usage

```
limrepo import --s SourceFilePath
```

**Note:** To see a list of parameters, type `help limrepo import`.

| Required | Argument                               | Description   |
|----------|--|---|
| Yes      | <code>--s <i>SourceFilePath</i></code> | Specifies the path to the resource or a folder on the local file system that is to be imported. This must be the root folder of a previous export on the local file system. |

| Required | Argument   | Description   |
|----------|--|---|
| No       | <code>--q</code> or <code>--quiet</code>           | <p>Disables the display of the resources copied during the import; that is, operates in quiet mode.</p> <p>If the flag is specified, the default value is true. If the flag is not specified, the default value is false.</p>   |
| No       | <code>--u</code> or <code>--update</code>          | <p>Specifies whether to overwrite existing resources if resources with the same name are already on the server.</p> <p><b>true</b> If there is a resource on the server with the same name as a resource you are importing, the resource on the server will be overwritten. This is the default setting if the flag is not specified or if the flag is specified without a value.</p> <p><b>false</b> If there is a resource on the server with the same name as a resource you are importing, the resource will not be imported.</p>   |
| No       | <code>--f</code> or <code>--fullpaths</code>       | <p>Prints the full source and output paths.</p> <p>If the flag is specified, the default value is true. If the flag is not specified, the default value is false.</p>   |
| No       | <code>--c</code> or <code>--continueonerror</code> | <p>Continues with the import if an error occurs.</p> <p>If the flag is specified, the default value is true. If the flag is not specified, the default value is false.</p>  |
| No       | <code>--a</code> or <code>--acl</code>             | <p>Preserves any previously exported permissions and merges them with existing permissions when importing resources. An access control list (ACL) indicates the operations each user or role can perform on a named resource, such as create, view, modify, or delete.</p> <p>For example, a user has read and write permissions on a resource when exporting. If the user only has read permissions on the resource when importing, write permission will be granted again after the import finishes successfully.</p> <p>Conflicting permissions cannot be merged and will be ignored. ACL entries for users and roles that do not exist in the target repository are also ignored.</p> |

| Required Argument | Description   |
|-------------------|---|
|                   | <p>If the flag is specified, the default value is true. If the flag is not specified, the default value is false.</p> <p><b>Tip:</b> When using this flag, the user on the server you exported from should also exist on the server to which you are importing. For example, you have "testuser" with access control settings and export the resources with ACL from one server, then import those named resources to another server that does not have "testuser". In this case, named resources will be uploaded but not the ACL.</p> |

**Example**

This example imports the named resources from C:\myrepository\samples on your local file system.

```
limrepo import --s C:\myrepository\samples
```

## limrepo mwsimport

**Note:** For instructions on installing and running the Administration Utility, see [Getting Started with the Administration Utility](#) on page 234.

The `limrepo mwsimport` command in the Spectrum™ Technology Platform Administration Utility allows you to provision a map from a MapInfo Workspace (MWS) file that has been created either by MapInfo Pro or the MapXtreme Workspace Manager into the Spectrum Spatial repository. The import will create the named map and all its dependent resources (layers, tables and connections). The connection is named by appending 'Connection' to the map name. The named tables and named layers are created in subfolders (NamedTables and NamedLayers, respectively).

You must have the Location Intelligence Module installed to use this command.

**Usage**

```
limrepo mwsimport --s MWSFilePath --o Output --p ServerPath
```

**Note:** To see a list of parameters, type `help limrepo mwsimport`.

| Required Argument                       | Description  |
|---|--|
| Yes <code>--s <i>MWSFilePath</i></code> | Specifies the path to an MWS file on the local file system that is to be imported. |

| Required | Argument              | Description  |
|----------|-----------------------|--|
| Yes      | --o <i>Output</i>     | Specifies the path to the named map on the repository. All resources will be created within the same folder as the named map.  |
| Yes      | --p <i>ServerPath</i> | Specifies the file path to the location of the data on the server. This path is used to create a named connection which is then referenced by all the named tables that are created. These tables will use file paths relative to that named connection.   |
| No       | --l <i>LocalPath</i>  | Specifies the file path to the location of the data on the local file system, if the MWS contains file paths that do not exist on the server file system. Any occurrences of the specified value in the MWS file will be substituted with the specified server path. If you have partial paths in the MWS file, this is not required; this is usually the case with anything created from MapXtreme. |

### Example

This example imports an MWS file on the D: drive (where the data on the server exists at C:\mydata) and provisions the named resources into /Europe/Countries in the repository.

```
limrepo mwsimport --s D:\europe.mws --o /Europe/Countries --p C:\mydata
```

### Result

The following named resources are created:

```
/Europe/Countries/Europe (named map)
/Europe/Countries/EuropeConnection (named connection)
/Europe/Countries/NamedTables/austria (named table)
/Europe/Countries/NamedTables/belgium (named table)
.
/Europe/Countries/NamedLayers/austria (named layer)
/Europe/Countries/NamedLayers/belgium (named layer)
..
```

## Match Rules

### matchrule delete

The `matchrule delete` command removes a match rule from your system. For more information, see the "Matching" section of the *Data Quality Guide*.

#### Usage

```
matchrule delete matchRuleName
```

| Required | Argument                                 | Description                         |
|----------|--|-------------------------------------|
| Yes      | <code>--d</code><br><i>matchRuleName</i> | Specifies the match rule to delete. |

#### Example

This example deletes the match rule named My Match Rule.

```
matchrule delete My Match Rule
```

### matchrule export

The `matchrule export` command exports a match rule that was created using one of the matching stages (Interflow Match, Intraflow Match, Transactional Match) in the Enterprise Designer. The match rule can then be imported to another server. You can export the match rule as `.mr` or `.json` files.

For more information, see the "Matching" section of the *Data Quality Guide*.

#### Usage

```
matchrule export matchRuleName --o OutputDirectory
```



| Required | Argument                                   | Description  |
|----------|--|--|
| Yes      | <code>--m</code><br><i>matchRuleName</i>   | Specifies the name of the match rule you want to export in the MapReduce file format.<br><br><b>Tip:</b> If you are unsure of the exact match rule name you can use the <code>matchrule list</code> command to get a list of the match rule names.   |
| No       | <code>--j</code><br><i>matchRuleName</i>   | Specifies the name of the match rule you want to export in JSON file format.<br><br><b>Tip:</b> If you are unsure of the exact match rule name you can use the <code>matchrule list</code> command to get a list of the match rule names.  |
| No       | <code>--o</code><br><i>OutputDirectory</i> | Specifies the directory to which you want to export the match rule. The path you specify here is relative to the directory where you are running the Administration Utility. If you omit this argument, the match rule is exported to the directory containing the Administration Utility. |

**Example: Match rule export in MapReduce format**

This example exports a match rule named "My Match Rule" in MapReduce format to a subfolder named `export` in the location where you are running the Administration Utility.

```
matchrule export My Match Rule --o export
```

**Example: Match rule export in JSON format**

This example exports a match rule named "My Match Rule" in JSON format to the directory containing the Administration Utility.

```
matchrule export My Match Rule --j
```

## matchrule import

The `matchrule import` command imports a match rule file into the server. Match rules are created using one of the matching stages (Interflow Match, Intraflow Match, Transactional Match) in Enterprise Designer. For more information, see the "Matching" section of the *Data Quality Guide*.

### Usage

```
matchrule import MatchRule --u TrueOrFalse
```

| Required | Argument                       | Description  |
|----------|--------------------------------|--|
| Yes      | <code>--f MatchRuleFile</code> | Specifies the match rule file you want to import. Directory paths you specify here are relative to the location where you are running the Administration Utility.  |
| No       | <code>--u TrueOrFalse</code>   | Specifies whether to overwrite the existing match rule file if a match rule file with the same name is already on the server, where <i>TrueOrFalse</i> is one of the following:<br><b>true</b><br>If there is a match rule file on the server with the same name as the match rule file you are importing, the match rule file on the server will be overwritten. This is the default setting.<br><b>false</b><br>If there is a match rule file on the server with the same name as the match rule file you are importing, the match rule file will not be imported. |

**Example**

This example imports the match rule named `MyMatchRule.mr`, which is located in a subfolder named `exported` in the location where you are running the Administration Utility. Because no `--u` command is specified, any existing match rule file of the same name on the server will be overwritten.

```
matchrule import exported\MyMatchRule.mr
```

## matchrule list

The `matchrule list` command lists all the match rules on the server. For each match rule, the following information is displayed: the match rule name and whether the match rule is exposed. For more information, see the "Matching" section of the *Data Quality Guide*.

### Usage

```
matchrule list
```

## Notification

### notification daystoexpire set

The `notification daystoexpire set` command specifies the number of days in advance that you want to be notified of a pending license or data expiration.

#### Usage

```
notification daystoexpire set --d Days
```

| Required | Argument        | Description   |
|----------|-----------------|---|
| Yes      | --d <i>Days</i> | Specifies the number of days in advance that you want to be notified of a pending license or data expiration. |

#### Example

This example sets notifications to be sent 30 days before a license or data expires.

```
notification daystoexpire set --d 30
```

### notification enabled set

The `notification enabled set` command enables or disables email notifications of pending license or data expiration.

#### Usage

```
notification enabled set --e TrueOrFalse
```

| Required | Argument               | Description  |
|----------|------------------------|--|
| Yes      | --e <i>TrueOrFalse</i> | Enables or disables license expiration notification where <i>TrueOrFalse</i> is one of the following: <ul style="list-style-type: none"> <li><b>true</b> Enables license expiration notification.</li> <li><b>false</b> Disables license expiration notification.</li> </ul> |

**Example**

This example enables notifications:

```
notification enabled set --e true
```

## notification expirationsettings list

The `notification expirationsettings list` command displays the expiration notification settings in effect on your system. The command displays the number of days before expiration that notifications will be sent, the users who are subscribed to notifications, and whether notifications are enabled or disabled.

### Usage

```
notification expirationsettings list
```

## notification smtp get

The `notification smtp get` command displays the email settings used to send license expiration notifications.

### Usage

```
notification smtp get
```

## notification smtp set

The `notification smtp set` command specifies the email settings to use to send license expiration notification emails.

### Usage

```
notification smtp set --h Host --o Port --u UserName --p Password --e FromEmail
```

| Required | Argument            | Description   |
|----------|---------------------|---|
| No       | --h <i>Host</i>     | Specifies the host name or IP address of the SMTP server to use to send the notification email.         |
| No       | --o <i>Port</i>     | Specifies the port number or range used by the SMTP server. The default is 25.                          |
| No       | --u <i>UserName</i> | Specifies the user name to use to send the email. This must be a valid user account on the SMTP server. |

| Required | Argument             | Description   |
|----------|----------------------|---|
| No       | --p <i>Password</i>  | Specifies the password for the user account specified in the --u parameter. |
| No       | --e <i>FromEmail</i> | Specifies the email address from which the notification email will be sent. |

**Example**

This example sets SMTP settings for email notifications.

```
notification smtp set --h mail.example.com --o 25 --u me123
--p MyPassword --e spectrum@example.com
```

## notification smtp test

The `notification smtp test` command sends a test email to an email address you specify. Use this command to verify the SMTP settings used for license expiration notification.

*Usage*

```
notification smtp test --e DestinationEmail
```

| Required | Argument                    | Description   |
|----------|-----------------------------|---|
| Yes      | --e <i>DestinationEmail</i> | Specifies the email address to which you want to send a test email message. |

**Example**

This example sends a test email address to `me@example.com`.

```
notification smtp test --e me@example.com
```

## notification subscriber add

The `notification subscriber add` command adds an email address to receive license expiration notifications.

*Usage*

```
notification subscriber add --e Email
```

| Required | Argument                      | Description   |
|----------|-------------------------------|---|
| Yes      | <code>--e <i>Email</i></code> | Specifies an email address to receive license expiration notifications. |

**Example**

This example adds the email address `me@example.com` to receive license expiration notifications.

```
notification subscriber add --e me@example.com
```

## notification subscriber delete

The `notification subscriber delete` command removes an email address from the list of email addresses that receive license expiration notifications.

*Usage*

```
notification subscriber delete --e Email
```

| Required | Argument                      | Description  |
|----------|-------------------------------|--|
| Yes      | <code>--e <i>Email</i></code> | Specifies an email address to remove from the list of email addresses that receive license expiration notifications. |

**Example**

This example removes the email address `me@example.com` from the list of email addresses that receive license expiration notifications.

```
notification subscriber delete --e me@example.com
```

## Open Parser Cultures

### openparser culture export

The `openparser culture export` command exports all Open Parser cultures. The cultures can then be imported to another server. For more information, see the "Parsing" section of the *Data Quality Guide*.

#### Usage

```
openparser culture export OutputFile
```

| Required | Argument                           | Description  |
|----------|------------------------------------|--|
| No       | <code>--f <i>OutputFile</i></code> | Specifies the file containing the cultures you want to export. If you omit this argument, the Administration Utility automatically exports the <code>cultures.xml</code> file. |

#### Example

This example exports the `culturesFR.xml` file.

```
openparser culture export culturesFR.xml
```

### openparser culture import

The `openparser culture import` command imports an Open Parser culture file into the server. For more information, see the "Parsing" section of the *Data Quality Guide*.

#### Usage

```
openparser culture import CultureFileName
```

| Required | Argument                                | Description   |
|----------|---|---|
| Yes      | <code>--f <i>CultureFileName</i></code> | Specifies the file containing the culture you want to import. |

#### Example

This example imports the file named `MyCulture.xml`.

```
openparser culture import MyCulture.xml
```

## Open Parser Domains

### openparser domain export

The `openparser domain export` command exports an Open Parser domain. The domain can then be imported to another server. For more information, see the "Parsing" section of the *Data Quality Guide*.

#### Usage

```
openparser domain export DomainName --o OutputDirectory
```

| Required Argument | Description   |
|-------------------|---|
| Yes               | <p><code>--d <i>DomainName</i></code> Specifies the name of the domain you want to export.</p> <p><b>Tip:</b> If you are unsure of the exact domain name you can use the <code>openparser domain list</code> command to get a list of the domain names.</p>   |
| No                | <p><code>--o <i>OutputDirectory</i></code> Specifies the directory to which you want to export the domain. The path you specify here is relative to the directory where you are running the Administration Utility. If you omit this argument, the domain is exported to the directory containing the Administration Utility.</p> |

#### Example

This example exports a domain named "MyDomain" to a folder named `exported`, which is a subfolder in the location where you have installed the Administration Utility.

```
openparser domain export MyDomain --o exported
```

### openparser domain import

The `openparser domain import` command imports an Open Parser domain into the server. For more information, see the "Parsing" section of the *Data Quality Guide*.



### Usage

```
openparser domain import DomainFileName
```

| Required | Argument                                  | Description  |
|----------|---|--|
| Yes      | <code>--f</code><br><i>DomainFileName</i> | Specifies the file containing the domain you want to import. Directory paths you specify here are relative to the location where you are running the Administration Utility. |

#### Example

This example imports the file named `MyDomain.xml`, which is located in a subfolder named `exported` in the location where you are running the Administration Utility.

```
openparser domain import exported\MyDomain.xml
```

## openparser domain list

The `openparser domain list` command lists all the Open Parser domains on the server. For each domain, the following information is displayed: the domain name and whether the domain is exposed.

### Usage

```
openparser domain list
```

## Performance Monitor

### performancemonitor enabled get

The `performancemonitor enabled get` command displays the jobs and services that have performance monitoring enabled.

### Usage

```
performancemonitor enabled get
```

#### Related concepts

[Performance Tuning Checklist](#) on page 191

#### Related tasks

[Monitoring Performance with the Administration Utility](#) on page 209

## performancemonitor enabled set

The `performancemonitor enabled set` command enables and disables performance monitoring for a job or service. When performance monitoring is enabled, performance information for the job or service is written to the performance log. The performance log includes overall performance information for the job or service as well as performance information for each stage in the job or service dataflow.

### Usage

```
performancemonitor enabled set --e TrueOrFalse --d DataflowName
```

| Required | Argument                             | Description  |
|----------|--------------------------------------|--|
| Yes      | <code>--e <i>TrueOrFalse</i></code>  | Enables or disables performance monitoring where <i>TrueOrFalse</i> is one of the following: <ul style="list-style-type: none"> <li><b>true</b>            Enables performance monitoring.</li> <li><b>false</b>          Disables disables performance monitoring.</li> </ul> |
| Yes      | <code>--d <i>DataflowName</i></code> | Specifies the name of the dataflow for which you want to enable or disable performance monitoring. <p><b>Tip:</b> If you are unsure of the exact dataflow name you can use the <code>dataflow list</code> command to get a list of the dataflow names.</p>                     |

#### Example

This example turns on performance monitoring for the dataflow `MyNameParsingDataflow`:

```
performancemonitor enabled set --e true --d
"MyNameParsingDataflow"
```

#### Related concepts

[Performance Tuning Checklist](#) on page 191

#### Related tasks

[Monitoring Performance with the Administration Utility](#) on page 209

### The Performance Log

The performance log contains details about how long it takes for a job or service to run. It includes overall performance information for the job or service as well as performance information for each stage in the job or service dataflow. You can use this information to identify bottlenecks in your dataflow by looking at the execution time and processing time for each stage. A large difference between execution time and processing time means that the stage is spending time waiting for data.

from upstream stages. This may indicate that an upstream stage is a bottleneck in the dataflow. Note that for sinks, a large difference between execution time and processing time does not necessarily indicate a performance problem because sinks typically have to wait for the first records from the rest of the dataflow.

To enable performance monitoring for a job or service, use the `performancemonitor enabled set` command in the Administration Utility.

The performance log is located on your Spectrum™ Technology Platform server in the following location:

```
SpectrumLocation\server\app\Repository\logs\performance.log
```

The performance log contains one row for each run of a monitored job or service. It is a rolling log that consists of a maximum of five files. Each file is limited to 10 MB in size. Once this limit is reached, the oldest performance data is deleted when new performance data is logged.

Each entry in the performance log contains the following information.

**Note:** For ease of reading, line breaks and indentation are shown below. In the actual log, the entry is on one line.

```
Date Time [performance]
{
  "username": "UserName",
  "dataflowId": "DataflowName",
  "runMode": "BatchOrRealTime",
  "remote": TrueOrFalse,
  "elapsedTime": Nanoseconds,
  "stageInfo": [
    {
      "stageName": "Name",
      "stageLabel": "Label",
      "options": {
        OptionsList
      },
      "recordsRead": Count,
      "recordsWritten": Count,
      "executionTime": Nanoseconds,
      "processingTime": Nanoseconds
    }
  ]
}
```

Where:

**username**

The user who executed the job or service.

**dataflowID**

The name of the service or job as defined in Enterprise Designer.

**runMode**

Indicates whether the log entry is for a job or a service. One of the following:

- Batch**                                    The log entry is for a job.  
**RealTime**                                The log entry is for a service.

### **remote**

Indicates whether the job or service was executed in the local server or on a remote server. For jobs that contain one or more stages that run on a remote server, the performance log on the server running the job will indicate that the job is not remote (a value of "false") while the performance log on the remote server that executes one or more of the job's stages will indicate that the job is remote (a value of "true"). One of the following:

- true**                                    The job or service was executed on a remote server.  
**false**                                    The job or service was executed on the local server.

### **elapsedTime**

The time in nanoseconds that it took to run the job or service request.

### **stageInfo**

Lists execution information for each stage in the dataflow. The following information is listed for each stage:

#### **stageName**

The permanent name of the stage.

#### **stageLabel**

The user-defined name of the stage. The stage label is shown on the canvas in Enterprise Designer.

#### **options**

If any options were specified at runtime, those options and their settings are listed here.

#### **recordsRead**

The total number of records that passed into the stage through all the stage's input ports.

#### **recordsWritten**

The total number of records that the stage wrote to all its output ports.

#### **executiontime**

The elapsed time from when the stage processed its first record to when it processed its last record. This includes the time the stage was idle while waiting for data from other stages in the dataflow.

#### **processingtime**

The amount of time the stage spent actively processing records, not including the time it was idle while waiting for other stages in the dataflow.

## Permissions

### permission list

The `permission list` command lists the names of the entities to which you can assign permissions.

#### Usage

```
permission list
```

## Process Flows

### processflow delete

The `processflow delete` command removes a process flow from your system.

#### Usage

```
processflow delete --n ProcessFlowName
```

| Required | Argument                                | Description  |
|----------|---|--|
| Yes      | <code>--n <i>ProcessFlowName</i></code> | Specifies the process flow to delete. If the process flow name contains spaces, enclose the process flow name in quotes. |

#### Example

This example deletes the process flow named My Process Flow.

```
processflow delete --n "My Process Flow"
```

## processflow execute

The `processflow execute` command runs one or more process flows. This command is one of two ways you can execute process flows from the command line. The other way is to use the Process Flow Executor, which is a command line utility available from the Spectrum™ Technology Platform welcome page on your server. The advantage of using the `processflow execute` command in the Administration Utility is that it allows you to also include other commands in a single script or batch file. For example, you could expose the process flow using the `processflow expose` command then execute it using the `processflow execute` command. The `processflow execute` command provides the same features as the Process Flow Executor.

### Usage

```
processflow execute --r ProcessFlowNames --f propertyFile --i PollInterval --d DelimiterCharacter --t Timeout --w WaitToComplete --o StageName=File
```

| Required | Argument                      | Description  |
|----------|-------------------------------|--|
| No       | --?                           | Prints usage information.  |
| No       | --d <i>DelimiterCharacter</i> | Sets a delimiter to use to separate the status information displayed in the command line when you execute the command. The default is " ". For example, using the default character, the following would be displayed at the command line when you execute a processflow named "MyProcessflow":<br><br>MyProcessflow 1 Succeeded |
| No       | --f <i>PropertyFile</i>       | Specifies a path to a property file. For more information on property files, see <a href="#">Using a Process Flow Property File</a> on page 187.   |
| No       | --i <i>PollInterval</i>       | Specifies how often to check for completed jobs, in seconds. The default is "5".   |
| Yes      | --r <i>ProcessFlowNames</i>   | A comma-separated list of process flows to run. Required.<br><br><b>Note:</b> If you specify any file overrides this argument must not be the last argument specified.   |
| No       | --t <i>Timeout</i>            | This option is deprecated and will be ignored.   |
| No       | --v <i>Verbose</i>            | Return verbose output where <i>Verbose</i> is one of the following:<br><b>true</b> Return verbose output.<br><b>false</b> Do not return verbose output.<br><br><b>Note:</b> If you specify any file overrides this argument must not be the last argument specified.   |

| Required | Argument                               | Description  |
|----------|--|--|
| No       | <code>--w <i>WaitToComplete</i></code> | This option is deprecated and will be ignored.   |
| No       | <code>--o <i>StageName=File</i></code> | Overrides the input or output file specified in the job. For more information, see <a href="#">Overriding Process Flow File Locations</a> on page 303. |

**Example**

This example executes the process flow named "My Process Flow".

```
processflow execute --r "My Process Flow"
```

**Overriding Process Flow File Locations**

When you execute a process flow using the `process flow execute` command in the Administration Utility, you can specify that the process flow should use different input and output files than those specified in the job. To do this use the `--o` argument:

```
--o "JobName|StageName=File"
```

Where:

***JobName***

The name of a job referenced in the process flow.

***StageName***

The name of a Read from File or Write to File stage in the job as shown in the stage label under the stage's icon in the dataflow. For example, if the input stage is labeled "Read From File" you would specify:

```
"Job1|Read From File=file:C:/inputfile.txt"
```

If the input stage is labeled "Illinois Customers" you would specify:

```
"Job1|Illinois Customers=file:C:/inputfile.txt"
```

***File***

The protocol and full path to the file. You must use forward slashes (/) in file paths, not backslashes. The protocol must be one of the following:

**file:**

If the file is on the same machine as the Spectrum™ Technology Platform server, start the path with the "file:" protocol. For example, on Windows specify

```
file:C:/myfile.txt
```

```
and on Unix or Linux specify
```

```
file:/testfiles/myfile.txt.
```

**Note:** If the client and server are running on the same machine, you can use either the "file:" or "esclient:" protocol, but are likely to have get better performance using the "file:" protocol.

**esclient:**

If the file is on the same machine as Process Flow Executor, start the path with the "esclient:" protocol. For example, on Windows specify `esclient:C:/myfile.txt` and on Unix or Linux specify `esclient:/testfiles/myfile.txt`.

**Note:** If the machine running process flow executor cannot resolve the host name of the Spectrum™ Technology Platform server, you may get an error "Error occurred accessing file". To resolve this issue, open this file on the server: `SpectrumLocation/server/app/conf/spectrum-container.properties`. Set the `spectrum.runtime.hostname` property to the IP address of the server.

**ftp:**

To use a file server defined in the Management Console, use the following format: `ftp:NameOfFileServer/PathToFile`. For example, `ftp://FS/testfiles/myfile.txt` where FS is a file server resource defined in Management Console.

The following example illustrates how to override file locations using the `--o` argument:

```
--o "Job1|Read from File=file:C:/myfile_input.txt" "Job1|Write to
File=file:C:/myfile_output.txt"
```

## processflow export

The `processflow export` command exports a process flow from the server to a `.pf` file. The process flow can then be imported to another server.

**Note:** Process flows can only be exchanged between identical versions of Spectrum™ Technology Platform.

### Usage

```
processflow export --n ProcessFlowName --o OutputFile
```

| Required Argument                           | Description   |
|---|---|
| Yes <code>--n <i>ProcessFlowName</i></code> | Specifies the name of the process flow you want to export. If the process flow name contains spaces, enclose the name in quotes.<br><br><b>Tip:</b> If you are unsure of the exact process flow name you can use the <code>processflow list</code> command to get a list of the process flow names. |
| No <code>--o <i>OutputFile</i></code>       | Specifies the directory to which you want to export the process flow. The path you specify here is relative to the directory where you are running the Administration Utility. If you omit this   |



| Required Argument | Description  |
|-------------------|--|
|                   | argument, the process flow is exported to the directory containing the Administration Utility. |

**Example**

This example exports the process flow named "My Process Flow" to a folder named `exported` which is a subfolder in the location where you have installed the Administration Utility.

```
processflow export --n "My Process Flow" --o exported
```

## processflow expose

The `processflow expose` command makes the process flow available for execution.

**Note:** If you use dataflow versioning in Enterprise Designer, the `processflow expose` command exposes the most recent version of the dataflow.

### Usage

```
processflow expose --n ProcessFlowName
```

| Required | Argument                                | Description  |
|----------|---|--|
| Yes      | <code>--n <i>ProcessFlowName</i></code> | Specifies the name of the processflow you want to expose. If the process flow name contains spaces, enclose the name in quotes.<br><br><b>Tip:</b> If you are unsure of the exact process flow name you can use the <code>processflow list</code> command to get a list of the process flow names. |

**Example**

This example exposes the process flow named "My Process Flow".

```
processflow expose --n "My Process Flow"
```

## processflow import

The `processflow import` command imports a process flow file (a `.pf` file) into the server. Process flow files are created by exporting a process flow from the server using the `processflow export` command

### Usage

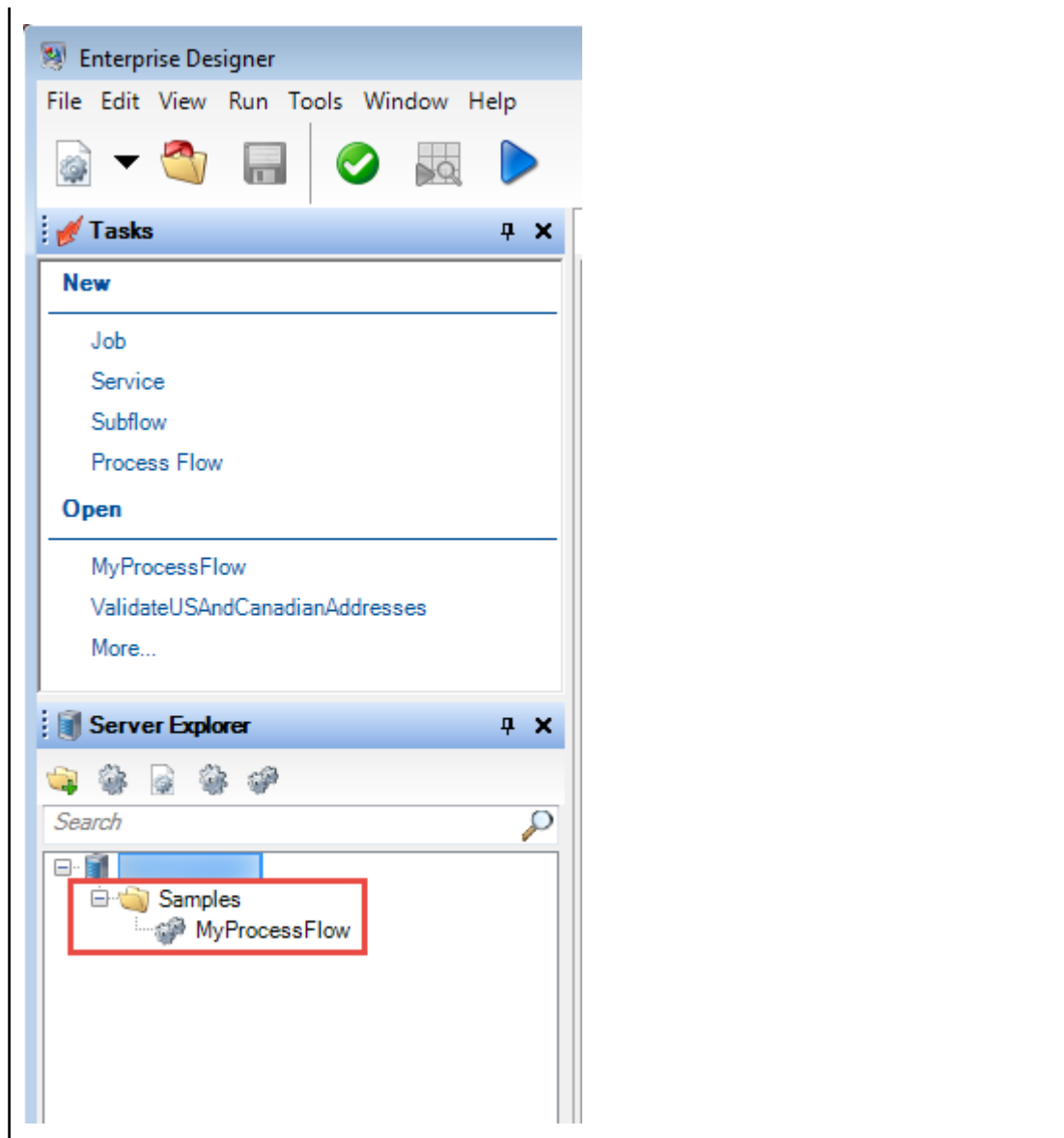
```
processflow import --f ProcessFlowFile --u TrueOrFalse --p Path --c TrueOrFalse
```

| Required | Argument                   | Description  |
|----------|----------------------------|--|
| Yes      | --f <i>ProcessFlowFile</i> | Specifies the process flow file (the .pf file) you want to import. Relative directory paths are relative to the location where you are running the Administration Utility. You can also specify an absolute path.  |
| No       | --u <i>TrueOrFalse</i>     | Specifies whether to overwrite the existing process flow if a process flow with the same name is already on the server, where <i>TrueOrFalse</i> is one of the following:<br><b>true</b><br>If there is a process flow on the server with the same name as the process flow you are importing, the process flow on the server will be overwritten. This is the default setting.<br><b>false</b><br>If there is a process flow on the server with the same name as the process flow you are importing, the process flow will not be imported. |
| No       | --p <i>Path</i>            | Specifies the Enterprise Designer Server Explorer folder to import the flow into.  |
| No       | --c <i>TrueOrFalse</i>     | Specifies whether to create the folder specified in --p if it does not exist.<br><b>true</b><br>Create the folder specified in --p if it does not exist. Default.<br><b>false</b><br>Do not create the folder specified in --p if it does not exist. The flow will not be imported unless the folder specified in --p exists.  |

#### Example

This example imports the process flow named `MyProcessFlow.pf` which is located in a subfolder named `exported` in the location where you are running the Administration Utility. The process flow will be imported into the `Samples` folder in Enterprise Designer.

```
processflow import --f exported\MyProcessFlow.pf --p Samples
```



## processflow list

The `processflow list` command lists all the process flows on the server. For each process flow, the process flow name is displayed as well as whether the process flow is exposed.

### Usage

```
processflow list
```

## processflow history list

The `processflow history list` command shows the execution history for a process flow.

### Usage

```
processflow history list --n ProcessFlowName --f FromDateTime --t ToDateTime
```

| Required | Argument                   | Description   |
|----------|----------------------------|---|
| Yes      | --n <i>ProcessFlowName</i> | <p>Specifies the name of the process flow whose status you want to get. If the process flow name contains spaces, enclose the name in quotes.</p> <p><b>Tip:</b> If you are unsure of the exact process flow name you can use the <code>processflow list</code> command to get a list of the process flow names.</p>  |
| No       | --f <i>FromDateTime</i>    | <p>If you want to see the history for a specific date and time range, specify the starting date and time for the range, in the format MM-dd-yyyy HH:mm:ss. For example, December 31, 2014 1:00 PM would be specified as <code>12-31-2014 13:00:00</code>.</p> <p>When you specify a date and time range, the history list will include process flows that started execution on or after the date you specified in the <code>--f</code> argument and before the date you specify in the <code>--t</code> argument.</p> <p>If you omit this argument the history will include process flows that started execution on the current date.</p>   |
| No       | --t <i>ToDateTime</i>      | <p>If you want to see the history for a specific date and time range, specify the ending date and time for the range, in the format MM-dd-yyyy HH:mm:ss. For example, December 31, 2014 1:00 PM would be specified as <code>12-31-2014 13:00:00</code>.</p> <p>When you specify a date and time range, the history list will include process flows that started execution on or after the date you specified in the <code>--f</code> argument and before the date you specify in the <code>--t</code> argument.</p> <p>If you omit this argument the history will include all process flows that started execution on or after the date specified in the <code>--f</code> argument.</p> |

**Example**

This example gets the history of the process flow named "My Process Flow".

```
processflow history list --n "My Process Flow"
```

## processflow unexpose

The `processflow unexpose` command makes a process flow unavailable for execution.

### Usage

```
processflow unexpose --n ProcessFlowName
```

| Required | Argument                   | Description   |
|----------|----------------------------|---|
| Yes      | --n <i>ProcessFlowName</i> | Specifies the name of the process flow you want to unexpose. If the process flow name contains spaces, enclose the name in quotes.<br><br><b>Tip:</b> If you are unsure of the exact process flow name you can use the <code>processflow list</code> command to get a list of the process flow names. |

**Example**

This example unexposes the process flow named "My Process Flow".

```
processflow unexpose --n "My Process Flow"
```

## processflow version list

The `processflow version list` command displays version information for all the saved versions of the process flow, including the version number, creator, creation date, comments, and which version is the exposed version.

### Usage

```
processflow version list --n ProcessFlowName
```

| Required | Argument                                | Description  |
|----------|---|--|
| Yes      | <code>--n <i>ProcessFlowName</i></code> | Specifies the name of the processflow whose version information you want to view. If the process flow name contains spaces, enclose the name in quotes.<br><br><b>Tip:</b> If you are unsure of the exact process flow name you can use the <code>processflow list</code> command to get a list of the process flow names. |

**Example**

This example displays the version information for the process flow named "My Process Flow".

```
processflow version list --n "My Process Flow"
```

## Remote Servers

### remoteserver list

The `remoteserver list` command lists the remote servers defined on the local server.

*Usage*

```
remoteserver list
```

### remoteserver add

*Usage*

```
remoteserver add --n RemoteServerNameName --h Host --o Port --u Username --p Password --s UseHTTPS --t timeout --m MicrobatchSize
```

| Required | Argument                               | Description  |
|----------|--|--|
| Yes      | <code>--n <i>ConnectionName</i></code> | Specifies a name for the remote server. The name can be anything you choose. |
| Yes      | <code>--h <i>Host</i></code>           | Specifies the host name or IP address of the remote server.                  |

| Required | Argument                               | Description   |
|----------|--|---|
| No       | <code>--o <i>Port</i></code>           | Specifies the network port to use for communication with the remote server.   |
| Yes      | <code>--u <i>Username</i></code>       | The username to use to connect to the remote server.  |
| Yes      | <code>--p <i>Password</i></code>       | The password to use to connect to the remote server.  |
| No       | <code>--s <i>UseHTTPS</i></code>       | Specifies whether to use HTTPS communication between the local server and the remote server. HTTPS must be enabled on the remote server in order for HTTPS to work.   |
| No       | <code>--t <i>timeout</i></code>        | The number of seconds to allow the system to wait for a response from the remote server before the connection times out. The default is 2 seconds.  |
| No       | <code>--m <i>MicrobatchSize</i></code> | <p>The number of records passed to the remote server in a single batch. The default is 50. Entering a higher number in this field will speed up input and output but slow down data transmission.</p> <p>To determine the optimal microbatch size for your environment, create a test dataflow that uses the remote server in one of its stages. Run the dataflow with a microbatch size of 50 and record the time it takes to run the dataflow. Then run a test with the microbatch size at 60, and another one with a microbatch size of 70 and observe the results. You may also want to run tests with microbatch sizes of 40 and 30 to see if reducing the microbatch size improves execution time. Continue testing with different settings until you identify the best setting for your environment.</p> |

### Example

This example creates a connection to a remote server named SpectrumServer2.

```
remoteserver add --n SpectrumServer2 --h MyServer --u john123
--p examplepassword1
```

## remoteserver delete

The `remoteserver delete` command deletes a remote server from your system.

### Usage

```
remoteserver delete --n RemoteServerName
```

| Required | Argument                          | Description   |
|----------|-----------------------------------|---|
| Yes      | <code>--n RemoteServerName</code> | Specifies the remote server to delete. If the remote server name contains spaces, enclose the name in quotes. |

**Example**

This example deletes the remote server named MyRemoteServer.

```
remoteserver delete --n MyRemoteServer
```

## remoteserver test

The `remoteserver test` command tests a connection between the local server and a remote server.

*Usage*

```
remoteserver test --n RemoteServerName
```

| Required | Argument                          | Description   |
|----------|-----------------------------------|---|
| Yes      | <code>--n RemoteServerName</code> | Specifies the name of the remote server you want to test. To view a list of remote servers, use the <code>remoteserver list</code> command. |

**Example**

This example test the connection RemoteServer1.

```
remoteserver test --n RemoteServer1
```

## remoteserver update

The `remoteserver update` command modifies a remote server definition. For example, you can update the password used to connect to the remote server if the password on the remote server has changed..

*Usage*

```
remoteserver update --n RemoteServerName --h Host --o Port --u Username --p Password
```



| Required | Argument                  | Description   |
|----------|---------------------------|---|
| Yes      | --n <i>ConnectionName</i> | Specifies the name for the remote server you want to modify. To view a list of remote servers, use the <code>remoteserver list</code> command.                      |
| Yes      | --h <i>Host</i>           | Specifies the host name or IP address of the remote server.   |
| No       | --o <i>Port</i>           | Specifies the network port to use for communication with the remote server.   |
| Yes      | --u <i>Username</i>       | The username to use to connect to the remote server.  |
| Yes      | --p <i>Password</i>       | The password to use to connect to the remote server.  |
| No       | --s <i>UseHTTPS</i>       | Specifies whether to use HTTPS communication between the local server and the remote server. HTTPS must be enabled on the remote server in order for HTTPS to work. |
| No       | --t <i>timeout</i>        | The number of seconds to allow the system to wait for a response from the remote server before the connection times out. The default is 2 seconds.                  |

**Example**

This example modifies a database connection named RemoteServer1. It changes the password to `NewPassword4`.

```
remoteserver update --p NewPassword4
```

## remoteserver refresh

The `remoteserver refresh` command lists the remote servers defined on the local server. If you make changes on a remote server, such as adding or removing modules or creating new services, you must refresh the remote server information on the local server.

### Usage

```
remoteserver refresh
```

# Roles

## role create

The `role create` command creates a new role with the permissions defined in a JSON file.

### Usage

```
role create --r RoleName --f JSONFile
```

| Required | Argument            | Description   |
|----------|---------------------|---|
| Yes      | --r <i>RoleName</i> | The name of the new role.   |
| Yes      | --f <i>JSONFile</i> | The path to a JSON file containing the definition for the new role. |

### Role File Format

The easiest way to create a role definition file is to use the `role permission export` command to generate a file of an existing role, then modify it. In this file, permissions are grouped as they are in the Management Console list of permissions. For each secured entity, you can specify the permission for EXECUTE, DELETE, CREATE, MODIFY, and VIEW. The valid values are:

#### true

Grants the permission.

#### false

Does not grant the permission.

#### null

A null value indicates a permission that does not apply to the secured entity.

The following example creates a new role named `MyNewRole`. This role has permissions for the permission group `Matching`. The permissions are `Open Parser Cultures`, `Open Parser Domains`, and `OpenParser Tables`.

```
{
  "name" : "MyNewRole",
  "userNames" : [ ],
  "groups" : [ {
    "name" : "Matching",
    "permissions" : [ {
      "name" : "Open Parser Cultures",
      "permissions" : {
```

```

        "EXECUTE" : "",
        "DELETE" : "",
        "CREATE" : "",
        "MODIFY" : "true",
        "VIEW" : "true"
    }
}, {
    "name" : "Open Parser Domains",
    "permissions" : {
        "EXECUTE" : "",
        "DELETE" : "",
        "CREATE" : "",
        "MODIFY" : "false",
        "VIEW" : "false"
    }
}, {
    "name" : "OpenParser Tables",
    "permissions" : {
        "EXECUTE" : "",
        "DELETE" : "false",
        "MODIFY" : "false",
        "CREATE" : "false",
        "VIEW" : "false"
    }
} ]
} ],
}

```

**Example**

This example creates a new role named SalesAnalyst and uses a role definition in the file c:\roles\SalesAnalyst.json.

```
role create --r SalesAnalyst --f C:\roles\SalesAnalyst.json
```

## role delete

The `role delete` command deletes a role.

**Usage**

```
role delete --r RoleName
```

| Required | Argument            | Description                               |
|----------|---------------------|---|
| Yes      | --r <i>RoleName</i> | Specifies the name of the role to delete. |

**Example**

This example deletes the role named SalesAnalyst.

```
role delete --r SalesAnalyst
```

## role export

The `role export` command exports all role definitions to a JSON file named `roles.json`.

### Usage

```
role export --o Folder
```

| Required | Argument          | Description   |
|----------|-------------------|---|
| No       | --o <i>Folder</i> | Specifies the name of the folder to which you want to export the roles. If you specify a relative path the path is relative to the location of the Administration Utility. If you do not specify a path the roles are exported to the folder where the Administration Utility is located. |

**Example**

This example exports the roles to the RoleExports folder.

```
role export --o RoleExports
```

## role list

The `role list` command lists the names of the all the roles on the system.

### Usage

```
role list
```

## role permission export

The `role permission export` command exports a role definition to a JSON file.

### Usage

```
role permission export --r RoleName --o OutputFolder
```

| Required | Argument                             | Description  |
|----------|--------------------------------------|--|
| Yes      | <code>--r <i>RoleName</i></code>     | Specifies the name of the role you want to export.   |
| No       | <code>--o <i>OutputFolder</i></code> | Specifies the name of the folder to which you want to export the role. If you specify a relative path the path is relative to the location of the Administration Utility. If you do not specify a path the role is exported to the folder where the Administration Utility is located. |

**Example**

This example exports the role named SalesAnalyst to the folder c:\roles.

```
role permission export --r SalesAnalyst --o C:\roles
```

## role permission import

The `role permission import` command modifies an existing role by importing permission settings from a JSON file.

**Usage**

```
role permission import --r RoleName --f PermissionsFile
```

| Required | Argument                                | Description   |
|----------|---|---|
| Yes      | <code>--r <i>RoleName</i></code>        | Specifies the name of the role you want to modify.  |
| Yes      | <code>--f <i>PermissionsFile</i></code> | Specifies the name of the JSON file that contains the permissions you want to add to the role. If you specify a relative path the path is relative to the location of the Administration Utility. |

**Example**

This example modifies the role named SalesAnalyst to have the permissions defined in c:\roles\permissions.json.

```
role permission import --r SalesAnalyst --f
C:\roles\permissions.json
```

## Search Indexes

### index list

The `index list` command returns a list of all Advanced Matching Module search indexes in tabular format. The details include name of the index and its type. The indexes can be of two types: **Legacy** and **Clustered**.

**Note:** The **Clustered** indexes can be backed up and restored.

#### Usage

```
index list
```

#### Example

This example lists all search indexes.

```
index list
```

### index delete

The `index delete` command deletes an Advanced Matching Module search index.

#### Usage

```
index delete --d Name
```

| Required | Argument        | Description  |
|----------|-----------------|--|
| Yes      | --d <i>Name</i> | Specifies the name of the search index you want to delete. |

#### Example

This example deletes a search index named "CustomerIndex".

```
index delete --d CustomerIndex
```

## index snapshot create

The `index snapshot create` command creates a snapshot of a search index. For a snapshot to be created successfully, the entire data set in the index needs to be valid. Any missing primary shard will result in failure of snapshot creation. Once a snapshot is created, its subsequent back up takes much less time since it requires incremental addition and deletion of data.

Snapshot creation is not affected by any operation being performed on the search index in parallel. However, only the records present in the index at that particular time gets recorded in the snapshot.

To view the status of the snapshot you created, use the command `index snapshot list`. For more information see [index snapshot list](#) on page 319.

**Note:** You need to create an index snapshot repository before you can create snapshot of a search index. For more information, see [index snapshot repository](#) on page 320.

### Usage

```
index snapshot create
```

| Required | Argument                              | Description   |
|----------|---------------------------------------|---|
| Yes      | <code>--i <i>Index name</i></code>    | Name of the search index in the snapshot to be created. |
| Yes      | <code>--s <i>Snapshot name</i></code> | Name of the snapshot to be created.                     |

**Note:** The name needs to be unique.

**Note:** It will be in lowercase irrespective of the casing you use in the command.

#### Example

This example creates a new snapshot "my\_snapshot" of search index "customer\_index".

```
index snapshot create --i customer_index --s my_snapshot
```

### index snapshot list

The `index snapshot list` command returns a list of all search index snapshots. The details displayed are:

- Name of the snapshot
- Name of the search index
- If the snapshot was created successfully
- Reason of failure, if the snapshot was not created successfully

- The start time, end time and total time taken in creating the snapshot
- Total shards in the snapshot and successful and failed shards, if any

### Usage

```
index snapshot list
```

#### Example

This example lists all the search index snapshots.

```
index snapshot list
```

## index snapshot delete

The `index snapshot delete` command deletes a search index snapshot.

### Usage

```
index snapshot delete
```

| Required | Argument                              | Description   |
|----------|---------------------------------------|---|
| Yes      | <code>--i <i>Index name</i></code>    | Name of the search index, the snapshot of which is to be deleted. |
| Yes      | <code>--s <i>Snapshot name</i></code> | Name of the snapshot to be deleted.                               |

#### Example

This example deletes the snapshot "my\_snapshot" of search index "customer\_index".

```
index snapshot delete --i customer_index --s my_snapshot
```

## index snapshot repository

The `index snapshot repository` command sets the search index snapshot repository shared file system path. You can set any number of repositories. But, the Search Index Engine always uses the currently set file path for backing up and restoring data.

To use the index snapshot repository for search index (in CLI), you need to first specify the repository path, `path.repo` at the location given below and restart the server. In case of cluster set-up this modification needs to be done at all the nodes. This path is used while creating the repository.

```
<Spectrum Installation Folder>\index\elasticsearch.template.
```

For example, the path can be:

- `path.repo: ["/mount/backups"]`



- `path.repo: ["C:/SIbackups"]`

Examples of repositories created using the above path are:

- `index snapshot repository --p /mount/backups/index_customer`
- `index snapshot repository --p C:/SIbackups/index_customer`

### Usage

`index snapshot repository`

| Required | Argument                   | Description  |
|----------|----------------------------|--|
| Yes      | <code>--p Name</code>      | Specifies the path where the index snapshot repository is to be set.<br><b>Note:</b> It is mandatory to use a shared file system path. |
| No       | <code>--o Overwrite</code> | Overwrites the already set path with the new one.  |

#### Example

This example sets a new index snapshot repository at the path "c:\\MyIndexRepository".

```
index snapshot repository --p c:\\MyIndexRepository
```

## index snapshot restore

The `index snapshot restore` command restores a search index snapshot.

**Note:** No operation is allowed on the search index while you are restoring its snapshot.

To view the status of snapshot restore use the command `index restore list`. For more information, see [index restore list](#) on page 322

### Usage

`index snapshot restore`

| Required | Argument                       | Description  |
|----------|--------------------------------|--|
| Yes      | <code>--i Index name</code>    | Name of the search index, the snapshot of which is being restored. |
| Yes      | <code>--s Snapshot name</code> | Name of the snapshot to be restored.                               |

#### Example

This example restores the snapshot "my\_snapshot" of search index "customer\_index".

```
index snapshot restore --i customer_index --s my_snapshot
```

### index restore list

The `index restore list` command returns a list of all the restored index snapshots of Advanced Matching Module. The details in the tabular format includes the name of the index and its restore status. It also gives the shard-wise restoration details, such as the total time taken in restoring each shard, the status, and restoration description.

#### Usage

```
index restore list
```

#### Example

This example lists all the restored search index snapshots.

```
index restore list
```

## Services

### service list

The `service list` command lists the services that are exposed on your server. Services that are not exposed will not be shown in the list.

**Note:** You can expose services using the `dataflow expose` command.

#### Usage

```
service list
```

### service loglevel list

The `service loglevel list` command lists the level of detail included in the log for each service. The log levels are:

- Default**      The service uses the system default logging level. You can set the system log level by using the `systemloglevel set` command.
- Disabled**     No event logging enabled.

|              |   |
|--------------|---|
| <b>Fatal</b> | Minimal logging. Only fatal errors are logged. Fatal errors are those that make the system unusable.  |
| <b>Error</b> | Errors and fatal errors are logged. Errors indicate an isolated problem that causes part of the system to become unusable. For example, a problem that causes a single service to not work would generate an error.   |
| <b>Warn</b>  | Event warnings, errors, and fatal errors are logged. Warnings indicate problems that do not stop the system from working. For example, when loading a service where a parameter has an invalid value, a warning is issued and the default parameter is used. During the use of a service, if results are returned but there is a problem, a warning will be logged. |
| <b>Info</b>  | High-level system information is logged. This is the most detailed logging level suitable for production. Info events are typically seen during startup and initialization, providing information such as version information and which services were loaded.   |
| <b>Debug</b> | A highly detailed level of logging, suitable for debugging problems with the system.  |
| <b>Trace</b> | The most detailed level of logging, tracing program execution (method entry and exit). It provides detailed program flow information for debugging.   |

### Usage

```
service loglevel list
```

## service loglevel set

The `service loglevel set` command specifies the level of detail included in the service log.

You can specify the default logging level as well as logging levels for each service on your system. When you change logging levels the change will not be reflected in the log entries made before the change.

### Usage

```
service loglevel set --s ServiceName --l LogLevel
```

| Required | Argument                            | Description  |
|----------|-------------------------------------|--|
| Yes      | <code>--s <i>ServiceName</i></code> | Specifies the name of the service whose logging level you want to set.   |
| Yes      | <code>--l <i>LogLevel</i></code>    | Specifies the logging level for the service, where <i>LogLevel</i> is one of the following: <ul style="list-style-type: none"> <li><b>Default</b> The service uses the system default logging level. You can set the system log level by using the <code>systemloglevel set</code> command.</li> <li><b>Disabled</b> No event logging enabled.</li> <li><b>Fatal</b> Minimal logging. Only fatal errors are logged. Fatal errors are those that make the system unusable.</li> </ul> |

| Required Argument | Description   |
|-------------------|---|
| <b>Error</b>      | Errors and fatal errors are logged. Errors indicate an isolated problem that causes part of the system to become unusable. For example, a problem that causes a single service to not work would generate an error.   |
| <b>Warn</b>       | Event warnings, errors, and fatal errors are logged. Warnings indicate problems that do not stop the system from working. For example, when loading a service where a parameter has an invalid value, a warning is issued and the default parameter is used. During the use of a service, if results are returned but there is a problem, a warning will be logged. |
| <b>Info</b>       | High-level system information is logged. This is the most detailed logging level suitable for production. Info events are typically seen during startup and initialization, providing information such as version information and which services were loaded.   |
| <b>Debug</b>      | A highly detailed level of logging, suitable for debugging problems with the system.  |
| <b>Trace</b>      | The most detailed level of logging, tracing program execution (method entry and exit). It provides detailed program flow information for debugging.   |

Each logging level includes the ones above it on the list. In other words, if Warning is selected as the logging level, errors and fatal errors will also be logged. If Info is selected, informational messages, warnings, errors, and fatal errors will be logged.

**Note:** Selecting the most intensive logging level can affect system performance. Therefore, you should select the least intensive setting that meets your particular logging requirements.

#### Example

This example sets the logging level for ValidateAddress to Warn:

```
service loglevel set --s ValidateAddress --l Warn
```

## service option list

The `service option list` command lists the options in effect for a service. For a description of each service's options and their values, see one of the following: *API Guide*, *REST Web Services Guide*, or *SOAP Web Services Guide*.

### Usage

```
service option list --s ServiceName
```

| Required | Argument               | Description   |
|----------|------------------------|---|
| Yes      | --s <i>ServiceName</i> | Specifies the name of the service whose options you want to view. Service names are case sensitive. |

### Example

This example lists the options in effect for the `ValidateAddress` service:

```
service option list --s ValidateAddress
```

## service option set

The `service option set` command specifies a default setting for a service option.

Default service options control the default behavior of each service on your system. You can specify a default value for each option in a service. The default option setting takes effect when an API call or web service request does not explicitly define a value for a given option. Default service options are also the settings used by default by services in a dataflow.

### Usage

```
service option set --s ServiceName --o OptionName --v Value
```

| Required | Argument               | Description   |
|----------|------------------------|---|
| Yes      | --s <i>ServiceName</i> | Specifies the name of the service for which you want to set an option. Service names are case sensitive.  |
| Yes      | --o <i>OptionName</i>  | The name of the option you want to set. For a description of each service's options and their values, see one of the following: <i>API Guide</i> , <i>REST Web Services Guide</i> , or <i>SOAP Web Services Guide</i> . |
| Yes      | --v <i>Value</i>       | The value you want to set for the option. For a description of each service's options and their values, see one of the following:   |

| Required | Argument | Description  |
|----------|----------|--|
|          |          | <i>API Guide, REST Web Services Guide, or SOAP Web Services Guide.</i> |

**Example**

This example sets the MaximumResults option for the ValidateAddress service to 15:

```
service option set --s ValidateAddress --o MaximumResults --v 15
```

## Spectrum Databases

### Enterprise Geocoding Module for Global Databases

#### egmglobaldb create sample file

The `egmglobaldb create_sample_file` command creates sample json file of single and double database resource. These generated files can be used as reference for providing configurations for creation of database resource. It creates *egmGlobalSingleDictDbResource.txt* and *egmGlobalDoubleDictDbResource.txt* json files in the current directory or at specified folder location.

#### Usage

```
egmglobaldb create_sample_file outputpath
```

| Required | Argument          | Description   |
|----------|-------------------|---|
| No       | <i>outputpath</i> | All the sample database resources json files will be created at provided output path else all will be exported to current folder. |

**Example**

This example creates the sample global database resources json files to current folder. The second example will export all the database resources to C:\OutputFolder.

```
egmglobaldb create_sample_file
```

```
egmglobaldb create_sample_file C:\OutputFolder
```

**Sample JSON for database resource file**

```
[{"product":"InternationalGeocoder GLOBAL",
"module":"igeocode-global", "name":"$$DATABASE_NAME$$",
"maxActive":4,
"properties":{"COUNTRY_CODE1":"$$COUNTRY_CODE1$$",
"$$COUNTRY_CODE1$$_DICTIONARY_PATH1":"$$DICTIONARY_PATH1$$",
"COUNTRY_COUNT":"1",
"$$COUNTRY_CODE1$$_DICTIONARY_PATH_NAME1":"$$DICTIONARY_PATH_NAME1$$"}}]
```

**egmglobaldb delete**

The `egmglobaldb delete` command deletes a configured Enterprise Geocoding Module Global database.

*Usage*

```
egmglobaldb delete --n Name
```

| Required Argument                | Description                         |
|----------------------------------|-------------------------------------|
| Yes <code>--n <i>Name</i></code> | Specifies the name of the database. |

**Example**

This example deletes the global database from the global module.

```
egmgbrdb delete --n Global
```

**egmglobaldb import**

The `egmglobaldb import` command imports an Enterprise Geocoding Module Global database property file. This configures the Global database resources on the current system.

*Usage*

```
egmglobaldb import --f File
```

| Required Argument                | Description  |
|----------------------------------|--|
| Yes <code>--f <i>File</i></code> | Specifies the JSON-formatted database property file. This file is mandatory. |

**Example**

This example creates a Global database resource as defined by the configuration provided in the `egmGlobalSingleDictDbResource.txt` JSON-formatted file.

```
egmglobaldb import --f egmGlobalSingleDictDbResource.txt
```

### egmglobaldb export

The `egmglobaldb export` command exports all of the Global database resource information to a database properties file, `EgmGlobalDbResource.txt`, at the specified location. If the location for the output file is not provided, the `EgmGlobalDbResource.txt` file is written to the current folder. The database properties file can subsequently be used with the `egmglobaldb import` command to configure the databases on another system.

#### Usage

```
egmglobaldb export --o outputpath
```

| Required | Argument                           | Description  |
|----------|------------------------------------|--|
| No       | <code>--o <i>outputpath</i></code> | All the database resources will be exported to the provided output path. If path is not specified, all resources will be exported to current folder. The exported <code>EgmGlobalDbResource.txt</code> file JSON-formatted output file contains the database properties information. |

#### Example

This example exports the database resource information to the designated location.

```
egmglobaldb export --o C:\DBs\
```

### egmglobaldb get

The `egmglobaldb get` command returns information about a Global Enterprise Geocoding Module database.

#### Usage

```
egmglobaldb get --n Name
```

| Required | Argument                     | Description   |
|----------|------------------------------|---|
| Yes      | <code>--n <i>Name</i></code> | Specifies the name of the database. This file is mandatory. |

#### Example

This example displays all the information for the configured Global database resource.



```
egmglobaldb get --n Global
```

### egmglobaldb list

The `egmglobaldb list` command displays all the configured Global Enterprise Geocoding Module databases and their pool sizes.

#### Usage

`egmglobaldb list` This command does not have any properties.

#### Example

This example lists the Enterprise Geocoding Module Global database and the pool size.

```
egmglobaldb list
```

### egmglobaldb poolsize set

The `egmglobaldb poolsize set` command sets the pool size for a configured global database resource. The pool size is the maximum number of concurrent requests allowed to a database.

#### Usage

`egmglobaldb poolsize set --n Name --s Poolsize`

| Required | Argument                         | Description   |
|----------|----------------------------------|---|
| Yes      | <code>--n <i>Name</i></code>     | Specifies the name of the database.   |
| No       | <code>--s <i>Poolsize</i></code> | Sets the pool size (specified as an integer value) of the database. The default is 4. |

#### Example

This example sets the poolsize of an already configured Global database resource to 10.

```
egmglobaldb poolsize set --n global --s 10
```

## Enterprise Geocoding Module for US Databases

### egmusadb add

The `egmusadb add` command creates a new US Enterprise Geocoding Module database resource on the server. You must have the US Enterprise Geocoding Module installed to use this command.

#### Usage

```
egmusadb add --f File
```

| Required | Argument                     | Description  |
|----------|------------------------------|--|
| Yes      | <code>--f <i>File</i></code> | Specifies the directory and name of the database resource to be added. |

### egmusadb delete

The `egmusadb delete` command deletes a configured Enterprise Geocoding Module US database.

#### Usage

```
egmusadb delete --n Name
```

| Required | Argument                     | Description                         |
|----------|------------------------------|-------------------------------------|
| Yes      | <code>--n <i>Name</i></code> | Specifies the name of the database. |

#### Example

This example deletes a database named "EGM\_CENTRUS\_POINTS".

```
egmusadb delete --n EGM_CENTRUS_POINTS
```

### egmusadb import

The `egmusadb import` command imports a Enterprise Geocoding Module US database property file created by the `egmusadb export` command. The `egmusadb import` command configures the US database resource on the current system.

#### Usage

```
egmusadb import --f File
```

| Required | Argument                     | Description  |
|----------|------------------------------|--|
| Yes      | <code>--f <i>File</i></code> | Specifies the JSON-formatted database property file. |

**Example**

This example will create a United States database resource as defined by the configuration provided in the EgmDbResource.txt JSON-formatted file.

```
egmusadb import --f EgmDbResource.txt
```

**egmusadb export**

The `egmusadb export` command exports all of the United States database resource information to a database property file, `EgmDbResource.txt`, at the specified location. If the location for the output file is not provided, the `EgmDbResource.txt` file is written to the current folder. The database property file can subsequently be used with the `egmusadb import` command to configure the databases on another system.

*Usage*

```
egmusadb export --o Directory
```

| Required | Argument             | Description  |
|----------|----------------------|--|
| No       | --o <i>Directory</i> | Specifies the output directory to store the JSON-formatted output file, <code>EgmDbResource.txt</code> , which contains the database properties information. |

**Example**

This example exports the database information to the designated location.

```
egmusadb export --o C:\DBs\
```

The `EgmDbResource.txt` output file contains database property information similar to the following:

```
[{"product": "GeoStan",
  "module": "geostan",
  "name": "TomTomStreets",
  "maxActive": 4,
  "properties": {"BASE_DB_PATHS": "C:/Dataset/DVDGDT",
    "DataSetName": "TomTomStreets"}},
 {"product": "GeoStan",
  "module": "geostan",
  "name": "CentrusPoints",
  "maxActive": 4,
  "properties": {"BASE_DB_PATHS": "C:/Dataset/DVDCPoints;C:/Dataset/DVDGDT",
    "DataSetName": "CentrusPoints"}}]
```

## egmusadb get

The `egmusadb get` command returns information about a US Enterprise Geocoding Module database.

### Usage

```
egmusadb get --n Name
```

| Required Argument                | Description                         |
|----------------------------------|-------------------------------------|
| Yes <code>--n <i>Name</i></code> | Specifies the name of the database. |

### Example

This example retrieves information for a database named "EGM\_CENTRUS\_PTS".

```
egmusadb get --n EGM_CENTRUS_PTS
```

The returned information may be similar to the following:

```

DATABASE NAME = EGM_CENTRUS_PTS
POOL SIZE = 4
BASE_DB_PATH = C:\DBs\EGM\CENTRUS_JUL14
DataSetName = EGM_CENTRUS_PTS

```

## egmusadb list

The `egmusadb list` command displays all the configured US Enterprise Geocoding Module databases and their pool size.

### Usage

```
egmusadb list
```

This command does not have any properties.

### Example

This example lists all the Enterprise Geocoding Module US databases.

```
egmusadb list
```

The returned information may be similar to the following:

```

+-----+-----+
| DATABASE NAME | POOL SIZE |
+-----+-----+
| TomTomStreets |         4 |
| TomTomPoints  |         4 |
| NAVTEQStreets |         4 |

```

```
| CentrusPoints |      4 |
+-----+-----+
```

### egmusadb poolsize set

The `egmusadb poolsize set` command sets the pool size for a configured US Enterprise Geocoding Module database. The pool size is the maximum number of concurrent requests allowed to a database.

#### Usage

```
egmusadb poolsize set --n Name --s Poolsize
```

| Required | Argument                         | Description   |
|----------|----------------------------------|---|
| Yes      | <code>--n <i>Name</i></code>     | Specifies the name of the database.   |
| No       | <code>--s <i>Poolsize</i></code> | Sets the pool size (specified as an integer value) of the database. The default is 4. |

#### Example

This example sets the poolsize to '3' for the "EGM\_CENTRUS\_PTS" database.

```
egmusadb poolsize set --n EGM_CENTRUS_PTS --s 3
```

## Enterprise Geocoding Module for Great Britain Databases

### egmgbrdb create sample file

The `egmgbrdb create_sample_file` command creates sample json file of single and double database resource. These generated files can be used as reference for providing configurations for creation of database resource. This command creates `egmGbrSingleDictDbResource.txt` json file in the current directory or at specified folder location.

#### Usage

```
egmgbrdb create_sample_file outputpath
```

| Required | Argument                       | Description  |
|----------|--------------------------------|--|
| No       | <code><i>outputpath</i></code> | All the sample database resources json file will be created at provided output path else all will be exported to current folder. |

**Example**

The first example creates the sample global database resources json file in the current folder. The second example creates the database resource file in C:\OutputFolder.

```
egmgbrdb create_sample_file
```

```
egmgbrdb create_sample_file C:\outputFolder
```

**egmgbrdb delete**

The `egmgbrdb delete` command deletes a configured Enterprise Geocoding Module GBR database.

*Usage*

```
egmgbrdb delete --n Name
```

| Required | Argument                     | Description                         |
|----------|------------------------------|-------------------------------------|
| Yes      | <code>--n <i>Name</i></code> | Specifies the name of the database. |

**Example**

This example deletes the configured GBR database from the GBR module.

```
egmgbrdb delete --n GBR
```

**egmgbrdb import**

The `egmgbrdb import` command imports an Enterprise Geocoding Module GBR database property file. This configures the GBR database resource on the current system.

*Usage*

```
egmgbrdb import --f File
```

| Required | Argument                     | Description  |
|----------|------------------------------|--|
| Yes      | <code>--f <i>File</i></code> | Specifies the JSON-formatted database property file. This file is mandatory. |

**Example**

This example creates a GBR database resource as defined by the configuration provided in the `EgmGBRDbResource.txt` JSON-formatted file.

```
egmgbrdb import --f EgmGBRDbResource.txt
```

### egmgbrdb export

The `egmgbrdb export` command exports all of the GBR database resource information to a database properties file, `EgmGbrDbResource.txt`, at the specified location. If the location for the output file is not provided, all the GBR database resources are exported to current folder. The database properties file can subsequently be used with the `egmgbrdb import` command to configure the databases on another system.

#### Usage

```
egmgbrdb export --o outputpath
```

| Required Argument                     | Description   |
|---------------------------------------|---|
| No <code>--o <i>outputpath</i></code> | All the database resources will be exported to the provided output path. If path is not specified, all resources will be exported to current folder. The exported <code>EgmGbrDbResource.txt</code> file JSON-formatted output file contains the database properties information. |

#### Example

This example exports the database resource information to the designated location.

```
egmgbrdb export --o C:\DBs\
```

### egmgbrdb get

The `egmgbrdb get` command returns information about a GBR Enterprise Geocoding Module database.

#### Usage

```
egmgbrdb get --n Name
```

| Required Argument                | Description                         |
|----------------------------------|-------------------------------------|
| Yes <code>--n <i>Name</i></code> | Specifies the name of the database. |

#### Example

This example displays all the information of the configured GBR database resource.

```
egmgbrdb get --n GBR
```

### egmgbrdb list

The `egmgbrdb list` command displays all the configured GBR Enterprise Geocoding Module databases and their pool size.

#### Usage

`egmgbrdb list` This command does not have any properties.

#### Example

This example lists all the Enterprise Geocoding Module GBR databases.

```
egmgbrdb list
```

### egmgbrdb poolsize set

The `egmgbrdb poolsize set` command sets the pool size for a configured GBR Enterprise Geocoding Module database. The pool size is the maximum number of concurrent requests allowed to a database.

#### Usage

`egmgbrdb poolsize set --n Name --s Poolsize`

| Required | Argument                         | Description   |
|----------|----------------------------------|---|
| Yes      | <code>--n <i>Name</i></code>     | Specifies the name of the database.   |
| No       | <code>--s <i>Poolsize</i></code> | Sets the pool size (specified as an integer value) of the database. The default is 4. |

#### Example

This example sets the pool size of the already configured GBR database to 10.

```
egmgbrdb poolsize set --n GBR --s 10
```



## Enterprise Geocoding Module for Australia Databases

### egmausdb create sample file

The `egmausdb create_sample_file` command creates sample json file of single and double database resource. These generated files can be used as reference for providing configurations for creation of database resource. This command creates `egmAusSingleDictDbResource.txt` and `egmAusDoubleDictDbResource.txt` json files in the current directory or at specified folder location.

#### Usage

```
egmausdb create_sample_file outputpath
```

| Required | Argument          | Description  |
|----------|-------------------|--|
| No       | <i>outputpath</i> | All the sample database resources json file will be created at provided output path else all will be exported to current folder. |

#### Example

The first example creates the sample global database resources json files in the current folder. The second example creates the database resource files in C:\OutputFolder.

```
egmausdb create_sample_file
```

```
egmausdb create_sample_file C:\outputFolder
```

### egmausdb delete

The `egmausdb delete` command deletes a configured Enterprise Geocoding Module AUS database.

#### Usage

```
egmausdb delete--n Name
```

| Required | Argument                     | Description                         |
|----------|------------------------------|-------------------------------------|
| Yes      | <code>--n</code> <i>Name</i> | Specifies the name of the database. |

**Example**

This example displays all the information of the configured AUS database resource.

```
egmausdb delete --n AUS
```

**egmausdb import**

The `egmausdb import` command imports a Enterprise Geocoding Module AUS database property file. This configures the Australia database resource on the current system.

*Usage*

```
egmausdb import --f File
```

| Required | Argument                     | Description  |
|----------|------------------------------|--|
| Yes      | <code>--f <i>File</i></code> | Specifies the JSON-formatted database property file. This file is mandatory. |

**Example**

This example creates an AUS database resource as defined by the configuration provided in the `EgmAUSDbResource.txt` JSON-formatted file.

```
egmausdb import --f EgmAUSDbResource.txt
```

**egmausdb export**

The `egmausdb export` command exports all of the AUS database resource information to a database properties file, `EgmAusDbResource.txt`, at the specified location. If the location for the output file is not provided, the `EgmAusDbResource.txt` file is exported to the current folder. The database properties file can subsequently be used with the `egmausdb import` command to configure the databases on another system.

*Usage*

```
egmausdb export --o outputpath
```

| Required | Argument                           | Description   |
|----------|------------------------------------|---|
| No       | <code>--o <i>outputpath</i></code> | All the database resources will be exported to the provided output path. If path is not specified, all resources will be exported to current folder. The exported <code>EgmAusDbResource.txt</code> file JSON-formatted output file contains the database properties information. |

**Example**

This example exports the database resource information to the designated location.

```
egmausdb export --o C:\DBs\
```

### egmausdb get

The `egmausdb get` command returns information about an AUS Enterprise Geocoding Module database.

#### Usage

```
egmausdb get --n Name
```

| Required Argument                | Description                         |
|----------------------------------|-------------------------------------|
| Yes <code>--n <i>Name</i></code> | Specifies the name of the database. |

#### Example

This example displays all the information of the configured AUS database resource.

```
egmausdb get --n AUS
```

### egmausdb list

The `egmausdb list` command displays all the configured AUS Enterprise Geocoding Module databases and their pool size.

#### Usage

`egmausdb list` This command does not have any properties.

#### Example

This example lists all the Enterprise Geocoding Module AUS databases.

```
egmausdb list
```

### egmausdb poolsize set

The `egmausdb poolsize set` command sets the pool size for a configured US Enterprise Geocoding Module database. The pool size is the maximum number of concurrent requests allowed to a database.

#### Usage

```
egmausdb poolsize set --n Name --s Poolsize
```

| Required | Argument                         | Description   |
|----------|----------------------------------|---|
| Yes      | <code>--n <i>Name</i></code>     | Specifies the name of the database.   |
| No       | <code>--s <i>Poolsize</i></code> | Sets the pool size (specified as an integer value) of the database. The default is 4. |

**Example**

This example sets the pool size of the already configured AUS database to 10.

```
egmausdb poolsize set --n AUS --s 10
```

## Enterprise Geocoding Module for World Databases

### egmworlddb create sample file

The `egmworlddb create_sample_file` command creates sample json files of single and double database resource. These generated files can be used as reference for providing configurations for creation of database resource. It creates `egmWorldSingleDictDbResource.txt` and `egmWorldDoubleDictDbResource.txt` json files in the current directory or at specified folder location.

#### Usage

```
egmworlddb create_sample_file outputpath
```

| Required | Argument          | Description   |
|----------|-------------------|---|
| No       | <i>outputpath</i> | All the sample database resources json files will be created at provided output path else all will be exported to current folder. |

**Example**

This example creates the sample world database resources json files to current folder. The second example will export all the database resources to C:\OutputFolder.

```
egmworlddb create_sample_file
```

```
egmworlddb create_sample_file C:\OutputFolder
```

**Sample JSON for database resource file**

```
[{"product":"InternationalGeocoder WORLD",
"module":"igeocode-world", "name":"$$DATABASE_NAME$$",
"maxActive":4,
"properties":{"COUNTRY_CODE1":"$$COUNTRY_CODE1$$",
```

```

"$$COUNTRY_CODE1$$_DICTIONARY_PATH1":"$$DICTIONARY_PATH1$$",
"COUNTRY_COUNT":"1",
"$$COUNTRY_CODE1$$_DICTIONARY_PATH_NAME1":"$$DICTIONARY_PATH_NAME1$$"}]]

```

### egmworlddb delete

The `egmworlddb delete` command deletes a configured Enterprise Geocoding Module World database.

#### Usage

```
egmworlddb delete --n Name
```

| Required Argument                | Description                         |
|----------------------------------|-------------------------------------|
| Yes <code>--n <i>Name</i></code> | Specifies the name of the database. |

#### Example

This example deletes the World database from the World module.

```
egmgbrdb delete --n world
```

### egmworlddb import

The `egmworlddb import` command imports an Enterprise Geocoding Module World database property file. This configures the World database resource on the current system.

#### Usage

```
egmworlddb import --f File
```

| Required Argument                | Description  |
|----------------------------------|--|
| Yes <code>--f <i>File</i></code> | Specifies the JSON-formatted database property file. This file is mandatory. |

#### Example

This example creates a World database resource as defined by the configuration provided in the `egmGlobalSingleDictDbResource.txt` JSON-formatted file.

```
egmglobaldb import --f egmWorldSingleDictDbResource.txt
```

## egmworlddb export

The `egmworlddb export` command exports all of the World database resource information to a database properties file, `EgmWorldDbResource.txt`, at the specified location. If the location for the output file is not provided, the `EgmWorldDbResource.txt` file is written to the current folder. The database properties file can subsequently be used with the `egmworlddb import` command to configure the databases on another system.

### Usage

```
egmworlddb export --o outputpath
```

| Required | Argument                           | Description   |
|----------|------------------------------------|---|
| No       | <code>--o <i>outputpath</i></code> | All the database resources will be exported to the provided output path. If path is not specified, all resources will be exported to current folder. The exported <code>EgmWorldDbResource.txt</code> file JSON-formatted output file contains the database properties information. |

#### Example

This example exports the database resource information to the designated location.

```
egmworlddb export --o C:\DBs\
```

## egmworlddb get

The `egmworlddb get` command returns information about a Global Enterprise Geocoding Module database.

### Usage

```
egmworlddb get --n Name
```

| Required | Argument                     | Description   |
|----------|------------------------------|---|
| Yes      | <code>--n <i>Name</i></code> | Specifies the name of the database. This file is mandatory. |

#### Example

This example displays all the information for the configured World database resource.

```
egmworlddb get --n World
```

## egmworldb list

The `egmworlddb list` command displays all the configured World Enterprise Geocoding Module databases and their pool sizes.

### Usage

`egmworlddb list` This command does not have any properties.

**Example**

This example lists the Enterprise Geocoding Module World database and the pool size.

```
egmworlddb list
```

### egmworlddb poolsize set

The `egmworlddb poolsize set` command sets the pool size for a configured World database resource. The pool size is the maximum number of concurrent requests allowed to a database.

### Usage

`egmworlddb poolsize set --n Name --s Poolsize`

| Required | Argument                         | Description   |
|----------|----------------------------------|---|
| Yes      | <code>--n <i>Name</i></code>     | Specifies the name of the database.   |
| No       | <code>--s <i>Poolsize</i></code> | Sets the pool size (specified as an integer value) of the database. The default is 4. |

**Example**

This example sets the poolsize of an already configured Global database resource to 10.

```
egmworlddb poolsize set --n world --s 10
```

## Global Addressing Module Databases

### gamdb create

The `gamdb create` command creates and configures Global Addressing Module databases.

### Usage

`gamdb create --n Name --d Dataset Name --v Dataset Vintage --c Country --t Type --p Poolsize`

| Required | Argument                             | Description  |
|----------|--------------------------------------|--|
| Yes      | <code>--n <i>Name</i></code>         | Specifies the name of the database resource to create. |
| Yes      | <code>--d <i>Dataset Name</i></code> | Specifies the name of the SPD dataset.                 |

| Required | Argument                                | Description   |
|----------|---|---|
| Yes      | <code>--v <i>Dataset Vintage</i></code> | Specifies the vintage of the dataset.   |
| No       | <code>--c <i>Country</i></code>         | Specifies the three-digit ISO code for each country to include in the databases specified by the "t" option (type of SPD) where Countries is a list of three-digit ISO codes separated by semicolons. For more information about ISO codes, see the Spectrum™ <i>Addressing Guide</i> . |
| Yes      | <code>--t <i>Type</i></code>            | Specifies the type of dataset.<br><br><b>GAV</b> Global Address Validation dataset.<br><b>GTA</b> Global Type Ahead dataset.  |
| No       | <code>--p <i>Poolsize</i></code>        | Specifies the maximum number of concurrent requests you want this database to handle. The default is 4.   |

**Example**

This example creates a Global Addressing Validation database for Germany named "GAV\_DEU" using the database resource "GAV\_EMEA" with a December 2016 vintage. This example configures the GAV\_DEU database with a pool size of 5.

```
gamdb create --n GAV_DEU --d GAV_EMEA --v DEC2016 --c DEU --t
GAV --p 5
```

**Example**

This example creates a Global Type Ahead database for Austria named "GTA\_AUT" using the database resource "GTA\_EMEA" with a December 2016 vintage. This example configures the GTA\_AUT database with a pool size of 6.

```
gamdb create --n GTA_AUT --d GTA_EMEA --v DEC2016 --c AUT --t
GTA --p 6
```

**gamdb delete**

The `gamdb delete` command deletes a Global Addressing Module database.

*Usage*

```
gamdb delete --n Name
```

| Required | Argument                     | Description                         |
|----------|------------------------------|-------------------------------------|
| Yes      | <code>--n <i>Name</i></code> | Specifies the name of the database. |



**Example**

This example deletes a Global Address Validation database for Germany named "GAV\_DEU".

```
gamdb delete --n GAV_DEU
```

**Example**

This example deletes a Global Type Ahead database for Austria named "GTA\_AUT".

```
gamdb delete --n GTA_AUT
```

**gamdb export**

The `gamdb export` command exports all of the Global Addressing database resource information to a database properties file, `GlobalAddressingDbResource.txt`, either at a specified location, or if the location for the output file is not provided, `GlobalAddressingDbResource.txt` is written to the current folder. The database properties file can subsequently be used with the `gamdb import` command to configure the databases on another system.

*Usage*

```
gamdb export --o outputpath
```

| Required Argument                     | Description  |
|---------------------------------------|--|
| No <code>--o <i>outputpath</i></code> | The information on the Global Addressing database resources will be exported to <code>GlobalAddressingDbResource.txt</code> in the specified output directory. If the path is not provided, <code>GlobalAddressingDbResource.txt</code> will be written to the current folder. |

**Example**

This example exports the Global Addressing database resource information to the designated directory.

```
gamdb export --o C:\DBs\
```

**gamdb get info**

The `gamdb get info` command returns detailed information about a Global Addressing database.

*Usage*

```
gamdb get info --n Name
```

| Required Argument                | Description                         |
|----------------------------------|-------------------------------------|
| Yes <code>--n <i>Name</i></code> | Specifies the name of the database. |

**Example**

This example displays all the information for the configured Global Addressing Validation database for Germany.

```
gamdb get info --n GAV_DEU
```

The returned information may be similar to the following:

```
DATABASE NAME = GAV_DEU
POOL SIZE = 5
BASE_DB_PATH = C:\DBs\DEU\
```

**Example**

This example returns information in a table for Global Addressing Validation.

```
gamdb get info --n GAV
```

The returned information may be similar to the following:

```
+-----+-----+-----+
|   SPDNAME   | SPDTYPE | COUNTRY |
+-----+-----+-----+
|   GAV_APAC  |   GAV   |   ALL   |
|   GAV_EMEA  |   GAV   |   ALL   |
| GAV_AMERICAS |   GAV   |   ALL   |
+-----+-----+-----+
```

**Example**

This example displays all the information for the configured Global Type Ahead database for Austria.

```
gamdb get info --n GTA_AUT
```

The returned information may be similar to the following:

```
DATABASE NAME = GAV_AUT
POOL SIZE = 6
BASE_DB_PATH = C:\DBs\AUT\
```

**Example**

This example returns information in a table for Global Type Ahead.

```
gamdb get info --n GTA
```

The returned information may be similar to the following:

```
+-----+-----+-----+
|   SPDNAME   | SPDTYPE | COUNTRY |
+-----+-----+-----+
|   GTA_APAC  |   GTA  |   ALL  |
|   GTA_EMEA  |   GTA  |   ALL  |
| GTA_AMERICAS |   GTA  |   ALL  |
+-----+-----+-----+
```

## gamdb import

The `gamdb import` command imports a Global Addressing database property file that configures the database resources on the current system.

### Usage

```
gamdb import --f File
```

| Required | Argument                     | Description  |
|----------|------------------------------|--|
| Yes      | <code>--f <i>File</i></code> | Specifies the JSON-formatted database property file. This file is mandatory. |

### Example

This example creates a Global Addressing database resource as defined by the configuration provided in the `GlobalAddressingDbResource.txt` JSON-formatted file.

```
gamdb import --f GlobalAddressingDbResource.txt
```

## gamdb listdatasets

The `gamdb listdatasets` command displays the Global Addressing Module databases registered on the platform.

### Usage

```
gamdb listdatasets This command does not have any properties.
```

### Example

This example lists the Global Addressing Module databases registered on the platform.

```
gamdb listdatasets
```

## gamdb listdbresources

The `gamdb listdbresources` command displays all the configured Global Addressing Module databases and the pool size for each database.

### Usage

`gamdb listdbresources` This command does not have any properties.

#### Example

This example lists the Global Addressing Module databases and the pool size for each database.

```
gamdb listdbresources
```

## gamdb modify

The `gamdb modify` command modifies and updates Global Addressing Module databases.

### Usage

`gamdb modify --n Name --d Dataset Name --v Dataset Vintage --c Country --t Type --p Poolsize`

| Required | Argument                         | Description   |
|----------|----------------------------------|---|
| Yes      | <code>--n Name</code>            | Specifies the name of the database resource to modify.  |
| Yes      | <code>--d Dataset Name</code>    | Specifies the name of the SPD dataset.  |
| Yes      | <code>--v Dataset Vintage</code> | Specifies the vintage of the dataset.   |
| No       | <code>--c Country</code>         | Specifies the three-digit ISO code for each country to include in the databases specified by the "t" option (type of SPD) where Countries is a list of three-digit ISO codes separated by semicolons. For more information about ISO codes, see the <i>Addressing Guide</i> . |
| Yes      | <code>--t Type</code>            | Specifies the type of dataset.<br><b>GAV</b> Global Address Validation database.<br><b>GTA</b> Global Type Ahead database.  |
| No       | <code>--p Poolsize</code>        | Specifies the maximum number of concurrent requests you want this database to handle. The default is 4.   |

#### Example

This example modifies the poolsize of the Global Addressing Validation database for Germany named "GAV\_DEU".

```
gamdb modify --n GAV_DEU --d GAV_EMEA --v DEC2016 --c DEU --t
GAV --p 6
```

**Example**

This example modifies the poolsize of the Global Type Ahead database for Austria named "GTA\_AUT".

```
gamdb modify --n GTA_AUT --d GTA_EMEA --v DEC2016 --c AUT --t
GTA --p 3
```

**gamdb poolsize set**

The `gamdb poolsize set` command sets the pool size for a configured Global Addressing Module database resource. The pool size is the maximum number of concurrent requests allowed to a database.

*Usage*

```
gamdb poolsize set --n Name --s Poolsize
```

| Required | Argument            | Description   |
|----------|---------------------|---|
| Yes      | --n <i>Name</i>     | Specifies the name of the database.   |
| No       | --s <i>Poolsize</i> | Sets the pool size (specified as an integer value) of the database. The default is 4. |

**Example**

This example sets the poolsize of an already configured Global Addressing Validation database for Germany to 5.

```
gamdb poolsize set --n GAV_DEU --s 5
```

**Example**

This example sets the poolsize of an already configured Global Type Ahead database for Austria to 7.

```
gamdb poolsize set --n GTA_AUT --s 7
```

## Global Geocoding Databases

### globalgeocodedb create sample file

The `globalgeocodedb create_sample_file` command creates sample json files of the Global Geocoding database resources. These generated files can be used as a reference for providing configurations when creating a database resource. `GeocodeGlobalSingleDictDbResource.txt` and `GlobalGeocodeDoubleDictDbResource.txt` json files are created in the current directory or in a specified folder location.

#### Usage

```
globalgeocodedb create_sample_file --o outputpath
```

| Required Argument                     | Description  |
|---------------------------------------|--|
| No <code>--o <i>outputpath</i></code> | The sample database resources json files will be created at the designated output directory. If the output path is not specified, the sample json files will be written to the current folder. |

#### Example

This example creates the sample database resources json files in the current folder.

```
globalgeocodedb create_sample_file
```

The following example creates the database resources json files in `C:\OutputFolder\`.

```
globalgeocodedb create_sample_file --o C:\OutputFolder\
```

#### Sample JSON database resources file

```
[{"product": "GlobalGeocode",
  "module": "GlobalGeocode",
  "name": "$DATABASE_NAME$",
  "maxActive": 4,
  "properties":
  {"COUNTRY_CODE1": "$COUNTRY_CODE1$",
   "$COUNTRY_CODE1$ _DICTIONARY_PATH1": "$DICTIONARY_PATH1$",

   "COUNTRY_COUNT": "1",

   "$COUNTRY_CODE1$ _DICTIONARY_PATH_NAME1": "$DICTIONARY_PATH_NAME1$"}
}]
```

### globalgeocodedb delete

The `globalgeocodedb delete` command deletes a configured Global Geocoding database.

#### Usage

```
globalgeocodedb delete --n Name
```

| Required Argument                | Description                         |
|----------------------------------|-------------------------------------|
| Yes <code>--n <i>Name</i></code> | Specifies the name of the database. |

#### Example

This example deletes the TomTomUSA database.

```
globalgeocodedb delete --n TomTomUSA
```

### globalgeocodedb import

The `globalgeocodedb import` command imports a Global Geocoding database property file. This configures the database resources on the current system.

#### Usage

```
globalgeocodedb import --f File
```

| Required Argument                | Description  |
|----------------------------------|--|
| Yes <code>--f <i>File</i></code> | Specifies the JSON-formatted database property file. This file is mandatory. |

#### Example

This example creates a Global Geocoding database resource as defined by the configuration provided in the `GlobalGeocodeDbResource.txt` JSON-formatted file.

```
globalgeocodedb import --f GlobalGeocodeDbResource.txt
```

There are several cases that may occur in response to the `globalgeocodedb import` command.

- **Case 1:** The directories in the specified root folder are all invalid. In this case, no database is added.

```
spectrum> globalgeocodedb import --f ./GlobalGeocodeDbResource.txt
/managers/GlobalGeocode/verify?rootFolder=D:/SGI_Data/
```

The response is as follows:

```
Invalid Folder locations found.
["D:\\SGI_Data\\IGEO-AT1"]
```

```
"D:\\SGI_Data\\IGEO-CZ1"]
unable to add the database resource due to invalid paths
```

- **Case 2:** The provided root folder has at least one valid directory. In this case, the database is added.

```
spectrum> globalgeocodedb import --f ./GlobalGeocodeDbResource.txt
/managers/GlobalGeocode/verify?rootFolder=D:/SGI_Data/GEO-DB
```

The response is as follows:

```
Invalid Folder locations found.
["D:\\SGI_Data\\IGEO-CZ1"]
Database resource imported [./GlobalGeocodeDbResource.txt]
```

- **Case 3:** The provided root folder is invalid or doesn't exist. In this case, the database is added.

```
spectrum> globalgeocodedb import --f ./GlobalGeocodeDbResource.txt
```

The response is as follows:

```
unable to add the database resource due to invalid paths
```

### globalgeocodedb export

The `globalgeocodedb export` command exports all of the Global Geocoding database resource information to a database properties file, `GlobalGeocodeDbResource.txt`, either at a specified location, or if the location for the output file is not provided, `GlobalGeocodeDbResource.txt` is written to the current folder. The database properties file can subsequently be used with the `globalgeocodedb import` command to configure the databases on another system.

#### Usage

```
globalgeocodedb export --o outputpath
```

| Required Argument                     | Description   |
|---------------------------------------|---|
| No <code>--o <i>outputpath</i></code> | The information on the Global Geocoding database resources will be exported to <code>GlobalGeocodeDbResource.txt</code> in the specified output directory. If the path is not provided, <code>GlobalGeocodeDbResource.txt</code> will be written to the current folder. |

#### Example

This example exports the Global Geocoding database resource information to the designated directory.

```
globalgeocodedb export --o C:\DBs\
```



The `GlobalGeocodeDbResource.txt` output file contains database resource information similar to the following:

```
[{"product": "GlobalGeocode",
  "module": "GlobalGeocode",
  "name": "TomTomStreets",
  "maxActive": 4,
  "properties":
  {"BASE_DB_PATHS": "C:/Dataset/DVDGDT",
   "DataSetName": "TomTomStreets"}},
 {"product": "GlobalGeocode",
  "module": "GlobalGeocode",
  "name": "CentrusPoints",
  "maxActive": 4,
  "properties":
  {"BASE_DB_PATHS": "C:/Dataset/DVDCPoints;C:/Dataset/DVDGDT",
   "DataSetName": "CentrusPoints"}}]
```

### globalgeocodedb get

The `globalgeocodedb get` command returns information about a Global Geocoding database.

#### Usage

```
globalgeocodedb get --n Name
```

| Required Argument                | Description   |
|----------------------------------|---|
| Yes <code>--n <i>Name</i></code> | Specifies the name of the database. This file is mandatory. |

#### Example

This example displays all the information for the configured Global Geocoding database resource.

```
globalgeocodedb get --n CENTRUS_PTS
```

The returned information may be similar to the following:

```
DATABASE NAME = CENTRUS_PTS
POOL SIZE = 4
BASE_DB_PATH = C:\DBs\USA\
DataSetName = USA_POINTS
```

### globalgeocodedb list

The `globalgeocodedb list` command displays all the configured Global Geocoding databases and their pool sizes.

### Usage

`globalgeocodedb list` This command does not have any properties.

**Example**

This example lists the Global Geocoding databases and the pool size.

```
globalgeocodedb list
```

### globalgeocodedb poolsize set

The `globalgeocodedb poolsize set` command sets the pool size for a configured Global Geocoding database resource. The pool size is the maximum number of concurrent requests allowed to a database.

### Usage

`globalgeocodedb poolsize set --n Name --s Poolsize`

| Required | Argument                         | Description   |
|----------|----------------------------------|---|
| Yes      | <code>--n <i>Name</i></code>     | Specifies the name of the database.   |
| No       | <code>--s <i>Poolsize</i></code> | Sets the pool size (specified as an integer value) of the database. The default is 4. |

**Example**

This example sets the poolsize of an already configured Global Geocoding database resource to 10.

```
globalgeocodedb poolsize set --n DEU_DB -s 10
```

## Universal Addressing Module Databases

### uamdb create

The `uamdb create` command creates a new Universal Addressing Module database.

### Usage

`uamdb create --t Type --n Name --c CacheSize --i Country --pl PreloadingType --dt DatabaseType --b BasePath --d DPVPath --l LACSPPath --s SuiteLinkPath --r RDIPPath --e EWSPPath --p Poolsize`

| Required                 | Argument   | Description  |                          |  |                |  |              |  |               |                 |               |                                  |             |                     |
|--------------------------|--|--|--------------------------|--|----------------|--|--------------|--|---------------|-----------------|---------------|----------------------------------|-------------|---------------------|
| Yes                      | <code>--t Type</code>  | Specifies the type of database, where <i>Type</i> is one of the following: <table border="0"> <tr> <td><b>USA</b></td> <td>United States database</td> </tr> <tr> <td><b>CAN</b></td> <td>Canadian Database</td> </tr> <tr> <td><b>INTL</b></td> <td>International Database</td> </tr> <tr> <td><b>Loqate</b></td> <td>Loqate Database</td> </tr> <tr> <td><b>Global</b></td> <td>Validate Address Global Database</td> </tr> <tr> <td><b>Amas</b></td> <td>Australian Database</td> </tr> </table>  | <b>USA</b>               | United States database   | <b>CAN</b>     | Canadian Database  | <b>INTL</b>  | International Database   | <b>Loqate</b> | Loqate Database | <b>Global</b> | Validate Address Global Database | <b>Amas</b> | Australian Database |
| <b>USA</b>               | United States database   |  |                          |  |                |  |              |  |               |                 |               |                                  |             |                     |
| <b>CAN</b>               | Canadian Database  |  |                          |  |                |  |              |  |               |                 |               |                                  |             |                     |
| <b>INTL</b>              | International Database   |  |                          |  |                |  |              |  |               |                 |               |                                  |             |                     |
| <b>Loqate</b>            | Loqate Database  |  |                          |  |                |  |              |  |               |                 |               |                                  |             |                     |
| <b>Global</b>            | Validate Address Global Database   |  |                          |  |                |  |              |  |               |                 |               |                                  |             |                     |
| <b>Amas</b>              | Australian Database  |  |                          |  |                |  |              |  |               |                 |               |                                  |             |                     |
| Yes                      | <code>--n Name</code>  | Specifies the name of the database.  |                          |  |                |  |              |  |               |                 |               |                                  |             |                     |
| No                       | <code>--c CacheSize</code>   | Specifies the cache size of a Validate Address Global database, where <i>CacheSize</i> is one of the following: <table border="0"> <tr> <td><b>None</b></td> <td>No cache</td> </tr> <tr> <td><b>Small</b></td> <td>Small cache</td> </tr> <tr> <td><b>Large</b></td> <td>Large cache (default)</td> </tr> </table>  | <b>None</b>              | No cache   | <b>Small</b>   | Small cache  | <b>Large</b> | Large cache (default)  |               |                 |               |                                  |             |                     |
| <b>None</b>              | No cache   |  |                          |  |                |  |              |  |               |                 |               |                                  |             |                     |
| <b>Small</b>             | Small cache  |  |                          |  |                |  |              |  |               |                 |               |                                  |             |                     |
| <b>Large</b>             | Large cache (default)  |  |                          |  |                |  |              |  |               |                 |               |                                  |             |                     |
| No                       | <code>--i Country</code>   | Specifies the three-digit ISO code(s) for each country in a Validate Address Global database that you want to use, where <i>Country</i> is either "All" (default) or a list of codes separated by comma.   |                          |  |                |  |              |  |               |                 |               |                                  |             |                     |
| No                       | <code>--pl PreloadingType</code>   | Specifies the amount of a Validate Address Global database that is preloaded, where <i>PreloadingType</i> is one of the following: <table border="0"> <tr> <td><b>None</b></td> <td>No data is preloaded (default).</td> </tr> <tr> <td><b>Partial</b></td> <td>Loads the metadata and indexing structures into memory. The reference data itself will remain on the hard drive. Offers some performance enhancements and is an alternative when not enough memory is available to fully load the desired databases.</td> </tr> <tr> <td><b>Full</b></td> <td>Moves the entire reference database into memory. This may need a significant amount of memory for countries with large databases such as the USA or the United Kingdom, but it will significantly increase the processing speed.</td> </tr> </table> | <b>None</b>              | No data is preloaded (default).  | <b>Partial</b> | Loads the metadata and indexing structures into memory. The reference data itself will remain on the hard drive. Offers some performance enhancements and is an alternative when not enough memory is available to fully load the desired databases. | <b>Full</b>  | Moves the entire reference database into memory. This may need a significant amount of memory for countries with large databases such as the USA or the United Kingdom, but it will significantly increase the processing speed. |               |                 |               |                                  |             |                     |
| <b>None</b>              | No data is preloaded (default).  |  |                          |  |                |  |              |  |               |                 |               |                                  |             |                     |
| <b>Partial</b>           | Loads the metadata and indexing structures into memory. The reference data itself will remain on the hard drive. Offers some performance enhancements and is an alternative when not enough memory is available to fully load the desired databases. |  |                          |  |                |  |              |  |               |                 |               |                                  |             |                     |
| <b>Full</b>              | Moves the entire reference database into memory. This may need a significant amount of memory for countries with large databases such as the USA or the United Kingdom, but it will significantly increase the processing speed.                     |  |                          |  |                |  |              |  |               |                 |               |                                  |             |                     |
| No                       | <code>--dt DatabaseType</code>   | Specifies the processing mode for a Validate Address Global database, where <i>DatabaseType</i> is one of the following: <table border="0"> <tr> <td><b>Batch_Interactive</b></td> <td>Used in batch processing or interactive environments. It is optimized for speed and will terminate attempts to correct an address when</td> </tr> </table>  | <b>Batch_Interactive</b> | Used in batch processing or interactive environments. It is optimized for speed and will terminate attempts to correct an address when |                |  |              |  |               |                 |               |                                  |             |                     |
| <b>Batch_Interactive</b> | Used in batch processing or interactive environments. It is optimized for speed and will terminate attempts to correct an address when   |  |                          |  |                |  |              |  |               |                 |               |                                  |             |                     |

| Required | Argument                 | Description   |
|----------|--------------------------|---|
|          |                          | ambiguous data is encountered that cannot be corrected automatically (default).   |
|          | <b>Certified</b>         | Used in batch processing environments for Australian mail to standardize and validate mail against the Postal Address File. |
|          | <b>FastCompletion</b>    | Used to enter truncated data in address fields and have Validate Address Global generate suggestions.                       |
| Yes      | --b <i>BasePath</i>      | Specifies the base subscription database path.  |
| No       | --d <i>DPVPath</i>       | Specifies the DPV database path.  |
| No       | --l <i>LACSPath</i>      | Specifies the LACS database path.   |
| No       | --s <i>SuiteLinkPath</i> | Specifies the SuiteLink database path.  |
| No       | --r <i>RDIPath</i>       | Specifies the RDI database path.  |
| No       | --e <i>EWSPath</i>       | Specifies the EWS database path.  |
| No       | --p <i>Poolsize</i>      | Specifies the maximum number of concurrent requests you want this database to handle. The default is 4.                     |

**Example**

This example creates a United States database named "UAM\_US" found in C:\DBs\UAM\US\_JUL14 with a DPV database found in the same location.

```
uamdb create --t USA --n UAM_US --b C:\DBs\UAM\US_JUL14 --d
C:\DBs\UAM\US_JUL14
```

**uamdb modify**

The `uamdb modify` command updates an existing Universal Addressing Module database.

**Usage**

```
uamdb modify --t Type --n Name --b BasePath --d DPVPath --l LACSPath --s
SuiteLinkPath --r RDIPath --e EWSPath --p Poolsize
```

| Required | Argument        | Description  |
|----------|-----------------|--|
| Yes      | --t <i>Type</i> | Specifies the type of database, where <i>Type</i> is one of the following: |
|          | <b>USA</b>      | United States database   |
|          | <b>CAN</b>      | Canadian Database  |
|          | <b>INTL</b>     | International Database   |
|          | <b>Loqate</b>   | Loqate Database  |

| Required | Argument                 | Description   |
|----------|--------------------------|---|
|          |                          | <b>Global</b> Validate Address Global Database  |
|          |                          | <b>Amas</b> Australian Database   |
| Yes      | --n <i>Name</i>          | Specifies the name of the database.   |
| Yes      | --b <i>BasePath</i>      | Specifies the base subscription database path.  |
| No       | --d <i>DPVPath</i>       | Specifies the DPV database path.  |
| No       | --l <i>LACSPath</i>      | Specifies the LACS database path.   |
| No       | --s <i>SuiteLinkPath</i> | Specifies the SuiteLink database path.  |
| No       | --r <i>RDIPath</i>       | Specifies the RDI database path.  |
| No       | --e <i>EWSPath</i>       | Specifies the EWS database path.  |
| No       | --p <i>Poolsize</i>      | Specifies the maximum number of concurrent requests you want this database to handle. The default is 4. |

**Example**

This example updates a United States database named "UAM\_US" found in C:\DBs\UAM\US\_JUL14 with a LACS database found in the same location.

```
uamdb modify --n UAM_US --t USA --b C:\DBs\UAM\US_JUL14 --l
C:\DBs\UAM\US_JUL14
```

**uamdb delete**

The `uamdb delete` command deletes a Universal Addressing Module database.

**Usage**

```
uamdb delete --t Type --n Name
```

| Required | Argument        | Description  |
|----------|-----------------|--|
| Yes      | --t <i>Type</i> | Specifies the type of database, where <i>Type</i> is one of the following: <ul style="list-style-type: none"> <li><b>USA</b> United States database</li> <li><b>CAN</b> Canadian Database</li> <li><b>INTL</b> International Database</li> <li><b>Loqate</b> Loqate Database</li> <li><b>Global</b> Validate Address Global Database</li> <li><b>Amas</b> Australian Database</li> </ul> |
| Yes      | --n <i>Name</i> | Specifies the name of the database.  |

**Example**

This example deletes a Canadian database named "UAM\_CAN".

```
uamdb delete --t CAN --n UAM_CAN
```

**uamdb import**

The `uamdb import` command exports a Universal Addressing Module database.

*Usage*

```
uamdb import --t Type
```

| Required | Argument        | Description  |
|----------|-----------------|--|
| Yes      | --t <i>Type</i> | Specifies the type of database, where <i>Type</i> is one of the following: |
|          | <b>USA</b>      | United States database   |
|          | <b>CAN</b>      | Canadian Database  |
|          | <b>INTL</b>     | International Database   |
|          | <b>Loqate</b>   | Loqate Database  |
|          | <b>Global</b>   | Validate Address Global Database   |
|          | <b>Amas</b>     | Australian Database  |

**Example**

This example imports a United States database.

```
uamdb import --t USA
```

**uamdb export**

The `uamdb export` command exports a Universal Addressing Module database.

*Usage*

```
uamdb export --t Type
```

| Required | Argument        | Description  |
|----------|-----------------|--|
| Yes      | --t <i>Type</i> | Specifies the type of database, where <i>Type</i> is one of the following: |
|          | <b>USA</b>      | United States database   |
|          | <b>CAN</b>      | Canadian Database  |
|          | <b>INTL</b>     | International Database   |
|          | <b>Loqate</b>   | Loqate Database  |

| Required Argument | Description                      |
|-------------------|----------------------------------|
| <b>Global</b>     | Validate Address Global Database |
| <b>Amas</b>       | Australian Database              |

**Example**

This example exports an international database.

```
uamdb export --t INTL
```

**uamdb get resource info**

The `uamdb get resource info` command returns information about a database.

*Usage*

```
uamdb get resource info --t Type --n Name
```

| Required Argument       | Description  |
|-------------------------|--|
| Yes     --t <i>Type</i> | Specifies the type of database, where <i>Type</i> is one of the following: <ul style="list-style-type: none"> <li><b>USA</b>                   United States database</li> <li><b>CAN</b>                   Canadian Database</li> <li><b>INTL</b>                  International Database</li> <li><b>Loqate</b>                Loqate Database</li> <li><b>Global</b>                Validate Address Global Database</li> <li><b>Amas</b>                  Australian Database</li> </ul> |
| Yes     --n <i>Name</i> | Specifies the name of the database.  |

**Example**

This example retrieves information for a United States database named "UAM\_US".

```
uamdb get resource info --t USA --n UAM_US
```

It may return information similar to the following:

```
DATABASE NAME = UAM_US
POOL SIZE = 4
LACS_DB_PATH = Z:\UAM\US_AUG12
SUITELINK_DB_PATH = Z:\UAM\US_AUG12
BASE_DB_PATH = Z:\UAM\US_AUG12
DPV_DB_PATH = E:\UAM_US_MAY_14_DB
RDI_DB_PATH = E:\UAM_US_MAY_14_DB
EWS_DB_PATH = Z:\UAM\US_AUG12
```

## uamdb list

The `uamdb list` command returns all Universal Addressing Module databases of that type in tabular format.

### Usage

```
uamdb list --t Type
```

| Required      | Argument                         | Description   |            |                        |            |                   |             |                        |               |                 |               |                                  |             |                     |
|---------------|----------------------------------|---|------------|------------------------|------------|-------------------|-------------|------------------------|---------------|-----------------|---------------|----------------------------------|-------------|---------------------|
| Yes           | <code>--t <i>Type</i></code>     | Specifies the type of database, where <i>Type</i> is one of the following: <table border="0"> <tr> <td><b>USA</b></td> <td>United States database</td> </tr> <tr> <td><b>CAN</b></td> <td>Canadian Database</td> </tr> <tr> <td><b>INTL</b></td> <td>International Database</td> </tr> <tr> <td><b>Loqate</b></td> <td>Loqate Database</td> </tr> <tr> <td><b>Global</b></td> <td>Validate Address Global Database</td> </tr> <tr> <td><b>Amas</b></td> <td>Australian Database</td> </tr> </table> | <b>USA</b> | United States database | <b>CAN</b> | Canadian Database | <b>INTL</b> | International Database | <b>Loqate</b> | Loqate Database | <b>Global</b> | Validate Address Global Database | <b>Amas</b> | Australian Database |
| <b>USA</b>    | United States database           |   |            |                        |            |                   |             |                        |               |                 |               |                                  |             |                     |
| <b>CAN</b>    | Canadian Database                |   |            |                        |            |                   |             |                        |               |                 |               |                                  |             |                     |
| <b>INTL</b>   | International Database           |   |            |                        |            |                   |             |                        |               |                 |               |                                  |             |                     |
| <b>Loqate</b> | Loqate Database                  |   |            |                        |            |                   |             |                        |               |                 |               |                                  |             |                     |
| <b>Global</b> | Validate Address Global Database |   |            |                        |            |                   |             |                        |               |                 |               |                                  |             |                     |
| <b>Amas</b>   | Australian Database              |   |            |                        |            |                   |             |                        |               |                 |               |                                  |             |                     |

### Example

This example lists all Canadian databases.

```
uamdb list --t CAN
```

## uamdbglobalmultipath create\_sample\_file

The `uamdbglobalmultipath create_sample_file` command configures a database resource with multiple path elements and creates a sample JSON file (`UamDbGlobalMultiPath.txt`) that can be modified with place holders and data paths. This command should be followed by a `uamdb_import` command for additional database configuration.

**Note:** You must replace the token values in the text file with the absolute values and data paths.

### Usage

```
uamdbglobalmultipath create_sample_file --o OutputDirectory --n  
NumberOfPathElements
```

| Required | Argument  | Description  |
|----------|---|--|
| No       | <code>--o</code><br><i>OutputDirectory</i>      | Specifies the directory where the file should go. The current directory is the default location. |
| Yes      | <code>--n</code><br><i>NumberOfPathElements</i> | Specifies the number of elements in the path.  |



**Example**

This example creates a sample JSON file named "UamDbGlobalMultiPath.txt" with its properties in JSON key-value format. This database resource has three path elements.

```
uamdbglobalmultipath create_sample_file --n 3
```

**uamdb poolsize set**

The `uamdb poolsize set` command sets the default poolsize for a database.

*Usage*

```
uamdb poolsize set --t Type --n Name --s Size
```

| Required Argument       | Description  |            |                        |            |                   |             |                        |
|-------------------------|--|------------|------------------------|------------|-------------------|-------------|------------------------|
| Yes     --t <i>Type</i> | Specifies the type of database, where <i>Type</i> is one of the following: <table border="0" style="margin-left: 20px;"> <tr> <td><b>USA</b></td> <td>United States database</td> </tr> <tr> <td><b>CAN</b></td> <td>Canadian Database</td> </tr> <tr> <td><b>INTL</b></td> <td>International Database</td> </tr> </table> | <b>USA</b> | United States database | <b>CAN</b> | Canadian Database | <b>INTL</b> | International Database |
| <b>USA</b>              | United States database   |            |                        |            |                   |             |                        |
| <b>CAN</b>              | Canadian Database  |            |                        |            |                   |             |                        |
| <b>INTL</b>             | International Database   |            |                        |            |                   |             |                        |
| Yes     --n <i>Name</i> | Specifies the name of the database.  |            |                        |            |                   |             |                        |
| Yes     --s <i>Size</i> | Specifies the default poolsize.  |            |                        |            |                   |             |                        |

**Example**

This example sets the poolsize of a Canadian database named "UAM\_CAN" to 4.

```
uamdb poolsize set --t CAN --n UAM_CAN --s 4
```

**uamdbglobalmultipath create\_sample\_file**

The `uamdbglobalmultipath create_sample_file` command configures a database resource with multiple path elements and creates a sample JSON file (UamDbGlobalMultiPath.txt) that can be modified with place holders and data paths. This command should be followed by a `uamdb_import` command for additional database configuration.

**Note:** You must replace the token values in the text file with the absolute values and data paths.

*Usage*

```
uamdbglobalmultipath create_sample_file --o OutputDirectory --n NumberOfPathElements
```

| Required | Argument                           | Description  |
|----------|------------------------------------|--|
| No       | --o<br><i>OutputDirectory</i>      | Specifies the directory where the file should go. The current directory is the default location. |
| Yes      | --n<br><i>NumberOfPathElements</i> | Specifies the number of elements in the path.  |

**Example**

This example creates a sample JSON file named "UamDbGlobalMultiPath.txt" with its properties in JSON key-value format. This database resource has three path elements.

```
uamdbglobalmultipath create_sample_file --n 3
```

## System

### close

The `close` command closes the session with the Spectrum™ Technology Platform server. Use this command if you want to close the connection to the server without exiting the Administration Utility. You can close and exit by using the `exit` command.

*Usage*

```
close
```

### connect

The `connect` command opens a session with the Spectrum™ Technology Platform server you specify. You must issue the `connect` command before you can issue other commands.

*Usage*

```
connect --h HostName --u UserName --p Password --s TrueOrFalse
```

| Required | Argument               | Description  |
|----------|------------------------|--|
| Yes      | --h <i>HostName</i>    | The host name and port to connect to, separated by a colon. For example, to connect to MyServer on port 8080 you would specify --h MyServer:8080.  |
| Yes      | --u <i>UserName</i>    | The user name to use to authenticate to the server.  |
| Yes      | --p <i>Password</i>    | The password for the user.   |
| No       | --s <i>TrueOrFalse</i> | Specifies whether to create a secure connection using HTTPS. You can only connect to the server over a secure connection if the server has been configured to support HTTPS communication. <i>TrueOrFalse</i> must be one of the following:<br><b>true</b><br>Use HTTPS.<br><b>false</b><br>Do not use HTTPS. This is the default setting. |

**Example**

This example opens a connection to the server MyServer on port 8080 with the user name admin and the password myPassword1.

```
connect --h MyServer:8080 --u admin --p myPassword1
```

**Related tasks**

[Configuring HTTPS Communication](#) on page 55

## date

The `date` command displays the current date and time on the computer where you are running the Administration Utility.

**Usage**

```
date
```

## exit

The `exit` command closes the session and exits the Administration Utility. If you want to close your session without exiting the Administration Utility, use the `close` command.

### Usage

```
exit
```

## help

The `help` command displays a list of the commands you can use in the Administration Utility. You can also use the `help` command to get information about the parameters used in each command.

### Usage

```
help Command
```

| Required | Argument       | Description  |
|----------|----------------|--|
| No       | <i>Command</i> | If you specify a command name as a parameter to the <code>help</code> command, detailed information about the command you specify is displayed. If you do not specify a command name, a list of all commands is shown. |

#### Example

This command lists all the commands available in the Administration Utility:

```
help
```

This command displays detailed information about the `set serviceoption` command:

```
help set serviceoption
```

## license expirationinfo export

The `license expirationinfo export` command exports a list of licenses that are about to expire. The licenses that are included are those that are going to expire within the time specified in Management Console on the **Notification** tab.

### Usage

```
license expirationinfo export --o Directory
```

| Required | Argument             | Description   |
|----------|----------------------|---|
| No       | --o <i>Directory</i> | Specifies the directory to which you want to export the license expiration information. The path you specify here is relative to the directory where you are running the Administration Utility. If you |

| Required | Argument | Description  |
|----------|----------|--|
|          |          | omit this option the expiration information is placed in a file in the folder where you are running the Administration Tool. |

**Example**

This example exports license expiration information to the LicenseExpiration subfolder under the folder where the Administration Tool is located.

```
license expirationinfo export --o LicenseExpiration
```

## license expirationinfo list

The `license expirationinfo list` command returns a list of licenses that are about to expire. The licenses that are displayed are those that are going to expire within the time specified in Management Console on the **Notification** tab.

*Usage*

```
license expirationinfo list
```

## licenseinfo export

The `licenseinfo export` command exports license information to a file. A license file may be needed when resolving license issues with technical support.

*Usage*

```
licenseinfo export --o Directory
```

| Required | Argument             | Description  |
|----------|----------------------|--|
| No       | --o <i>Directory</i> | Specifies the directory to which you want to export the license file. The path you specify here is relative to the directory where you are running the Administration Utility. If you omit this option the license file is placed in the folder where you are running the Administration Tool. |

**Example**

This example exports license information to the License subfolder under the folder where the Administration Tool is located.

```
licenseinfo export --o License
```

## licenseinfo list

The `licenseinfo list` command displays license information such as which licenses are installed, the number of days remaining on the license, and the number of transactions remaining.

### Usage

```
licenseinfo list
```

## server backup

Use the `server backup` command to back up your Spectrum™ Technology Platform server.

To back up your Spectrum™ Technology Platform server, you need to create a backup copy of the server's configuration database. The configuration database contains your security settings, dataflows, service options, data resource definitions, and other configuration settings. If you were to lose your server due to a severe system failure or other disaster, you could use the backup of the configuration database to restore your configuration to another Spectrum™ Technology Platform server.

**Important:** Do not run a backup when there is activity on the Spectrum™ Technology Platform server. While the backup is in progress, calls to services may time out and jobs may fail to execute successfully.

### Usage

```
server backup --o Directory
```

| Required | Argument             | Description  |
|----------|----------------------|--|
| No       | --o <i>Directory</i> | Specifies the directory to which you want to save the backup copy of the server's database. The path you specify here is relative to the directory where you are running the Administration Utility. If you omit this option the backup is placed in the folder where you are running the Administration Tool. |

### Example

This example saves the backup to the LatestServerBackup subfolder under the folder where the Administration Tool is located.

```
server backup --o LatestServerBackup
```

### Related concepts

[Creating a Backup Manually](#) on page 230

### Related tasks

[Scheduling Backups](#) on page 226

[Restoring a Server](#) on page 231

## script

The `script` command directs the Administration Tool to execute a script containing a series of commands. You can use a script to automate administrative tasks.

### Usage

```
script --file ScriptFile --linenumbers TrueOrFalse
```

| Required | Argument                                      | Description   |
|----------|---|---|
| Yes      | <code>--file <i>ScriptFile</i></code>         | Specifies the path to the script file.  |
| No       | <code>--linenumbers <i>TrueOrFalse</i></code> | Specifies whether to display line numbers while executing the script, where <i>TrueOrFalse</i> is one of the following:<br><b>true</b><br>Displays line numbers while executing the script.<br><b>false</b><br>Does not display line numbers while executing the script. This is the default value. |

### Example

This example executes a script named `myscript.cli` located in the folder `scripts` which is a subfolder under the folder that contains the Administration Utility.

```
script --file scripts/myscript.cli
```

### Related tasks

[Using a Script with the Administration Utility](#) on page 235

## system loglevel get

The `system loglevel get` command returns the default logging level for services. The logging levels are:

- Off** No event logging enabled.
- Fatal** Minimal logging. Only fatal errors are logged. Fatal errors are those that make the system unusable.

|              |   |
|--------------|---|
| <b>Error</b> | Errors and fatal errors are logged. Errors indicate an isolated problem that causes part of the system to become unusable. For example, a problem that causes a single service to not work would generate an error.   |
| <b>Warn</b>  | Event warnings, errors, and fatal errors are logged. Warnings indicate problems that do not stop the system from working. For example, when loading a service where a parameter has an invalid value, a warning is issued and the default parameter is used. During the use of a service, if results are returned but there is a problem, a warning will be logged. |
| <b>Info</b>  | High-level system information is logged. This is the most detailed logging level suitable for production. Info events are typically seen during startup and initialization, providing information such as version information and which services were loaded.   |
| <b>Debug</b> | A highly detailed level of logging, suitable for debugging problems with the system.  |
| <b>Trace</b> | The most detailed level of logging, tracing program execution (method entry and exit). It provides detailed program flow information for debugging.   |

### Usage

```
system loglevel get
```

## system loglevel set

The `system loglevel set` command sets the default logging level for services on your system.

### Usage

```
system loglevel set --l Level
```

#### Required Argument Description

|     |                               |   |
|-----|-------------------------------|---|
| Yes | <code>--l <i>Level</i></code> | Specifies the default logging level for services on your system, where <i>Level</i> is one of the following:  |
|     | <b>Off</b>                    | No event logging enabled.   |
|     | <b>Fatal</b>                  | Minimal logging. Only fatal errors are logged. Fatal errors are those that make the system unusable.  |
|     | <b>Error</b>                  | Errors and fatal errors are logged. Errors indicate an isolated problem that causes part of the system to become unusable. For example, a problem that causes a single service to not work would generate an error.   |
|     | <b>Warn</b>                   | Event warnings, errors, and fatal errors are logged. Warnings indicate problems that do not stop the system from working. For example, when loading a service where a parameter has an invalid value, a warning is issued and the default parameter is used. During the use of a service, if results are returned but there is a problem, a warning will be logged. |



| Required Argument | Description   |
|-------------------|---|
| <b>Info</b>       | High-level system information is logged. This is the most detailed logging level suitable for production. Info events are typically seen during startup and initialization, providing information such as version information and which services were loaded. |
| <b>Debug</b>      | A highly detailed level of logging, suitable for debugging problems with the system.  |
| <b>Trace</b>      | The most detailed level of logging, tracing program execution (method entry and exit). It provides detailed program flow information for debugging.   |

---

**Note:** Selecting the most intensive logging level can affect system performance. Therefore, you should select the least intensive setting that meets your particular logging requirements.

**Example**

This example sets the default logging level to Warn:

```
system loglevel set --l warn
```

## system properties

The `system properties` command displays information about the shell running the Administration Utility, such as Java properties and OS version. It does not display information about the Spectrum™ Technology Platform server.

### Usage

```
system properties
```

## versioninfo export

The `versioninfo export` command exports system, component, and service version information to a file.

### Usage

```
versioninfo export --o Directory
```

| Required | Argument             | Description   |
|----------|----------------------|---|
| No       | --o <i>Directory</i> | Specifies the directory to which you want to export the version information text file. The path you specify here is relative to the directory where you are running the Administration Utility. If you omit this option the version information file is placed in the folder where you are running the Administration Tool. |

**Example**

This example exports version information to the VersionInformation subfolder under the folder where the Administration Tool is located.

```
versioninfo export --o VersionInformation
```

## versioninfo list

The `versioninfo list` command displays information about the version of Spectrum™ Technology Platform installed on your system, its underlying components, as well some system information.

*Usage*

```
versioninfo list
```

## Tables

### table delete

The `table delete` command removes a table from your system. For more information, see the "Lookup Tables" section of the *Data Quality Guide*.

*Usage*

```
table delete TableName --t TableType
```

| Required | Argument             | Description   |
|----------|----------------------|---|
| Yes      | --n <i>TableName</i> | Specifies the table to delete.  |
| Yes      | --t <i>TableType</i> | Specifies the type of table to delete: AdvancedTransformer, OpenParser, or TableLookup. |

**Example**

This example deletes the Table Lookup table named My Table.

```
table delete My Table --t TableLookup
```

## table export

The `table export` command exports a custom table that was created using the Table Management feature in Enterprise Designer. The table can then be imported to another server. For more information, see the "Lookup Tables" section of the *Data Quality Guide*.

### Usage

```
table export TableName --t TableType --o OutputDirectory
```

| Required | Argument                   | Description  |
|----------|----------------------------|--|
| Yes      | --n <i>TableName</i>       | Specifies the name of the table you want to export.<br><br><b>Tip:</b> If you are unsure of the exact table name you can use the <code>table list</code> command to get a list of the table names.   |
| Yes      | --t <i>TableType</i>       | Specifies the type of table to export: AdvancedTransformer, OpenParser, or TableLookup.  |
| No       | --o <i>OutputDirectory</i> | Specifies the directory to which you want to export the table. The path you specify here is relative to the directory where you are running the Administration Utility. If you omit this argument, the table is exported to the directory containing the Administration Utility. |

**Example**

This example exports an Open Parser table named "My Table" to the location where you have installed the Administration Utility.

```
table export My Table --t OpenParser
```

## table import

The `table import` command imports a custom table into the server. Custom tables are created using the Table Management feature in Enterprise Designer. For more information, see the "Lookup Tables" section of the *Data Quality Guide*.

### Usage

```
table import CustomTable --u TrueOrFalse
```

| Required | Argument               | Description  |
|----------|------------------------|--|
| Yes      | --f <i>CustomTable</i> | Specifies the custom table you want to import. Directory paths you specify here are relative to the location where you are running the Administration Utility.   |
| No       | --u <i>TrueOrFalse</i> | Specifies whether to overwrite the existing table if a table with the same name is already on the server, where <i>TrueOrFalse</i> is one of the following:<br><b>true</b><br>If there is a table on the server with the same name as the table you are importing, the table on the server will be overwritten. This is the default setting.<br><b>false</b><br>If there is a table on the server with the same name as the table you are importing, the table will not be imported. |

#### Example

This example imports the table named `MyTable.db` which is located in a subfolder named `exported` in the location where you are running the Administration Utility. Because no `--u` command is specified, any existing table of the same name on the server will be overwritten.

```
table import exported\MyTable.db
```

## table list

The `table list` command lists all the tables on the server. For each table, the following information is displayed: the table name and whether the dataflow is exposed.

### Usage

```
table list --t TableType
```

| Required | Argument             | Description  |
|----------|----------------------|--|
| Yes      | --t <i>TableType</i> | Specifies the type of tables to list: AdvancedTransformer, OpenParser, or TableLookup. |

**Example**

This example lists all Advanced Transformer tables.

```
table list --t AdvancedTransformer
```

## Tokens

### token list

The `token list` command returns a list of the active tokens on the Spectrum™ Technology Platform server. For each token, the following information is provided:

- User name
- Date and time the token was created
- Date and time when the token was last used
- The IP address of the computer being used to access the Spectrum™ Technology Platform server
- The session ID
- The token

Each of these fields is separated by a pipe character ( | ).

#### Usage

```
token list --u UserName
```

| Required | Argument            | Description  |
|----------|---------------------|--|
| No       | --u <i>UserName</i> | Specifies the user whose tokens you want to view. If you do not specify this argument, all users' tokens are listed. |

**Example**

This example lists the tokens for the user amy123.

```
token list --u amy123
```

This example lists all tokens.

```
token list
```

## token refreshsecret

The `token refreshsecret` command refreshes the secret key. This has the effect of rendering all active tokens invalid. Users with active tokens will need to log in again to obtain a new token.

### Usage

```
token refreshsecret
```

## token revoke

The `token revoke` command make a token invalid. The user will need to log in again to obtain a new token.

### Usage

```
token revoke --t Token
```

| Required | Argument         | Description  |
|----------|------------------|--|
| Yes      | --t <i>Token</i> | Specifies the token you want to revoke. To see a list of active tokens, use the <code>token list</code> command. |

### Example

This example revokes the token specified in the `--t` argument.

```
token revoke --t
```

## token userrevoke

The `token userrevoke` command makes all of a user's tokens invalid. The user will need to log in again to obtain a new token.

### Usage

```
token userrevoke --u UserName
```

| Required | Argument            | Description   |
|----------|---------------------|---|
| Yes      | --u <i>UserName</i> | Specifies the user whose tokens you want to revoke. To see a list of active users, use the <code>token list</code> command. |

**Example**

This example revokes the tokens for the user amy123.

```
token userrevoke --u amy123
```

**Related tasks**

[Disabling a User Account](#) on page 32

[Deleting a User](#) on page 32

## User Accounts

### user create

The `user create` command creates a new user and assigns roles to it.

*Usage*

```
user create --u UserName --p Password --d Description --e EmailAddress --r Roles
```

| Required | Argument                | Description   |
|----------|-------------------------|---|
| Yes      | --u <i>UserName</i>     | Specifies the name of the new user.   |
| Yes      | --p <i>Password</i>     | Specifies the password for the new user.  |
| No       | --d <i>Description</i>  | Specifies a description for the user. If the description contains spaces, enclose the description in quotes.  |
| No       | --e <i>EmailAddress</i> | Specifies the email address of the new user.  |
| No       | --r <i>Roles</i>        | Specifies the roles for the user. Separate multiple roles with a comma. Do not use spaces. If the role name contains a space, enclose the role in quotes. |

**Example**

This example creates a new user named allan12, assigns a password of myPassword1, a description of Allan P. Smith, an email of allan@example.com, and assigns two roles, USBanking and California Users.

```
user create --u allan12 --p myPassword1 --d "Allan P. Smith"
--e allan@example.com --r USBanking,"California Users"
```

## user delete

The `user delete` command removes a user account from the system.

**Tip:** User accounts can also be disabled, which prevents the account from being used to access the system without deleting the account.

### Usage

```
user delete --u UserName
```

| Required | Argument            | Description  |
|----------|---------------------|--|
| Yes      | --u <i>UserName</i> | Specifies the name of the user you want to delete. |

### Example

This example deletes a user named allan12.

```
user delete --u allan12
```

## user description set

The `user description set` command changes the account description.

### Usage

```
user description set --u UserName --d Description
```

| Required | Argument               | Description  |
|----------|------------------------|--|
| Yes      | --u <i>UserName</i>    | Specifies the user account whose description you want to change.   |
| Yes      | --d <i>Description</i> | Specifies a description for the user. If the description contains spaces, enclose the description in quotes. |

### Example

This example changes the description of the user allan12 to "Allan P. Smith."

```
user description set --u allan12 --d "Allan P. Smith"
```



## user email set

The `user email set` command changes the email address associated with the user.

### Usage

```
user email set --u UserName --e EmailAddress
```

| Required | Argument                | Description  |
|----------|-------------------------|--|
| Yes      | --u <i>UserName</i>     | Specifies the user account whose email address you want to change. |
| Yes      | --e <i>EmailAddress</i> | Specifies the email address to associate with the user.            |

### Example

This example sets the email address for the user account `allan12` to `allan@example.com`.

```
user email set --u allan12 --e allan@example.com
```

## user enabled set

The `user enabled set` command enables or disables a user account.

You can disable a user account so that it cannot be used to gain access to Spectrum™ Technology Platform. Any jobs that run on a schedule using a disabled user account will not run.

**Note:** The user account "admin" cannot be disabled.

### Usage

```
user enabled set --u UserName --e TrueOrFalse
```

| Required | Argument               | Description   |
|----------|------------------------|---|
| Yes      | --u <i>UserName</i>    | Specifies the name of the user you want to disable.   |
| Yes      | --e <i>TrueOrFalse</i> | Specifies whether to enable or disable the user account where <i>TrueOrFalse</i> is one of the following:<br><b>true</b><br>Enables the user account.<br><b>false</b><br>Disables the user account. |

**Example**

This example disables the user account allan12.

```
user enabled set --u allan12 --e false
```

## user list

The `user list` command returns a list of users. For each user, the command lists the user's roles, email address, description, and whether the user is enabled or disabled.

**Usage**

```
user list
```

## user password set

The `user password set` command changes the password for a user account.

**Usage**

```
user password set --u UserName --p Password
```

| Required | Argument            | Description   |
|----------|---------------------|---|
| Yes      | --u <i>UserName</i> | Specifies the user account whose password you want to change. |
| Yes      | --p <i>Password</i> | Specifies the password to assign to the user account.         |

**Example**

This example sets the password for the user account allan12 to mypassword1.

```
user password set --u allan12 --p mypassword1
```

## user role grant

The `user role grant` command assigns one or more roles to a user.

**Usage**

```
user role grant --u UserName --r Roles
```

| Required | Argument            | Description   |
|----------|---------------------|---|
| Yes      | --u <i>UserName</i> | Specifies the name of the user to whom you want to assign a role.   |
| Yes      | --r <i>Roles</i>    | Specifies the roles for the user. Separate multiple roles with a comma. Do not use spaces. If the role name contains a space, enclose the role in quotes. |

**Example**

This example assigns two roles, USBanking and CaliforniaUsers, to the user allan12.

```
user role grant --u allan12 --r USBanking,CaliforniaUsers
```

## user role list

The `user role list` command returns a list of roles on your system.

**Usage**

```
user role list
```

## user role revoke

The `user role revoke` command removes a role from a user so that the user no longer has the privileges granted by the role.

**Usage**

```
user role revoke --u UserName --r Roles
```

| Required | Argument            | Description  |
|----------|---------------------|--|
| Yes      | --u <i>UserName</i> | Specifies the name of the user whose role you want to revoke.  |
| Yes      | --r <i>Roles</i>    | Specifies the role you want to revoke. Separate multiple roles with a comma. Do not use spaces. If the role name contains a space, enclose the role in quotes. |

**Example**

This example revokes the role USBanking from the user allan12.

```
user role revoke --u allan12 --r USBanking
```

# 11 - Clustering

## In this section

---

|   |     |
|---|-----|
| Clustered Architecture                                  | 381 |
| Using Enterprise Designer with a Cluster                | 383 |
| Starting a Cluster                                      | 384 |
| Stopping a Cluster                                      | 385 |
| Removing a Node from a Cluster                          | 385 |
| Managing a Cluster for the Location Intelligence Module | 386 |

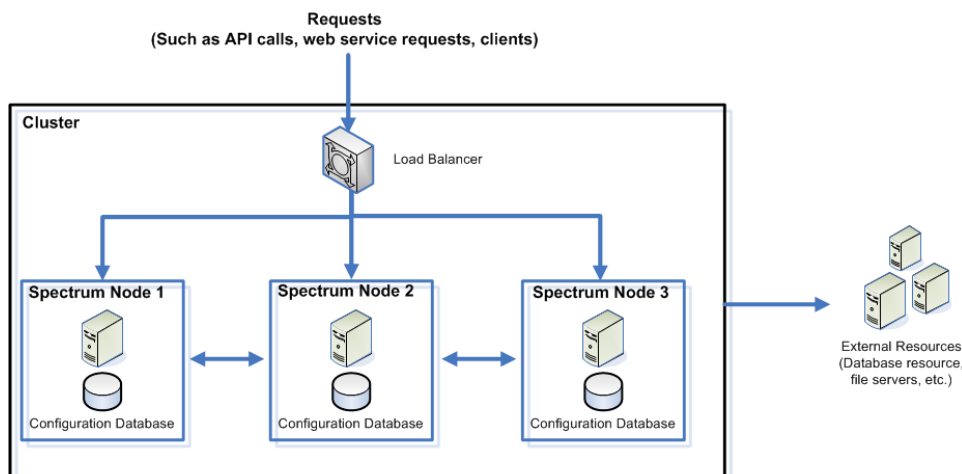
## Clustered Architecture

In a clustered environment, processing is shared among two or more instances of the server. All communication with Spectrum™ Technology Platform goes through a load balancer. Consider clustering if you require failover redundancy and high-volume, high-performance processing.

There are two ways to run Spectrum™ Technology Platform in a cluster. You can install the Spectrum™ Technology Platform server and configuration database on each node in the cluster, or you can separate the server from the configuration database.

### Cluster with Server and Configuration Database on Each Node

In this configuration, a complete Spectrum™ Technology Platform server is installed on each node so that each node has both the server software and the configuration database. The configuration database stores things like your user names and passwords, connections, flows, history, and audit log.



This configuration has these advantages:

- Processing is automatically distributed among the nodes in the cluster for optimal performance.
- Configuration settings are automatically synchronized across nodes.
- Uptime can be improved since if a node goes down the other nodes remain available to handle processing requests.
- Adding a new node is easy.

### Cluster with Separated Configuration Database

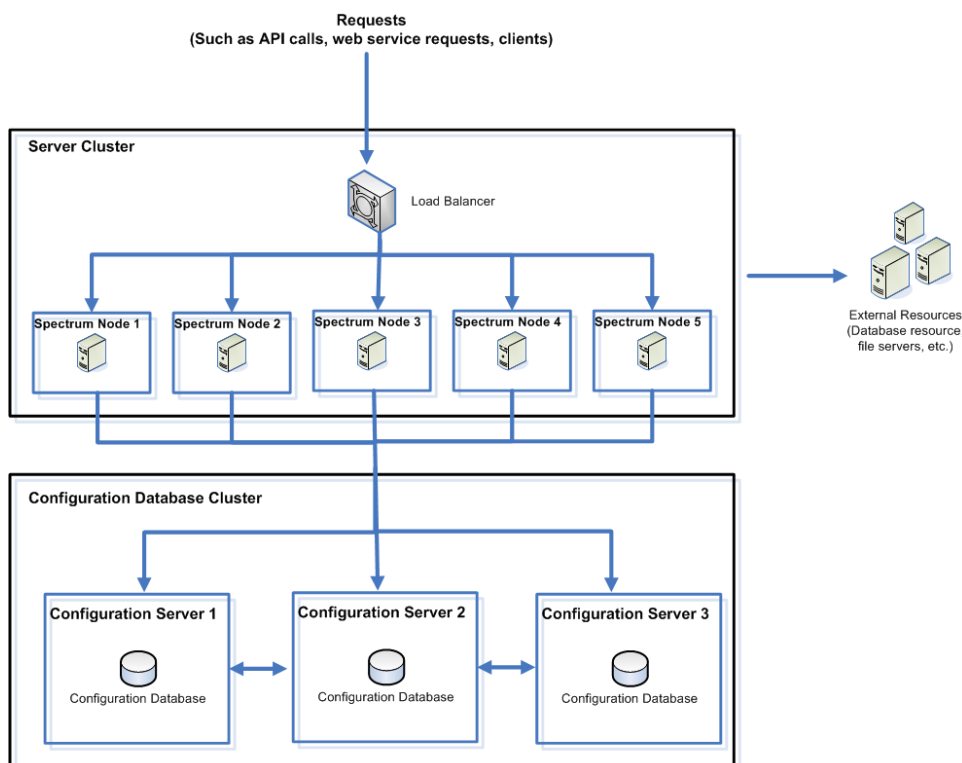
When you install the Spectrum™ Technology Platform server you are installing two things: server software and a configuration database. The server handles the execution of jobs and service requests, reporting, and logging. The configuration database stores data such as users and roles, data connections, dataflows, and the audit log.

For clusters consisting of only a few nodes, installing the server and configuration database together on each node provides acceptable performance. However, a cluster that consists of a large number of nodes may have its performance negatively affected by having a copy of the configuration database on each node. This is because an increasing number of nodes results in increased data replication between nodes during synchronization. Performance may also suffer if users are accessing the configuration database frequently through activities such as editing dataflows or job schedules. These activities require processing power, resulting in less CPU capacity available for running jobs and responding to service requests.

If you have an implementation with four or more nodes, you should consider installing the configuration database to separate servers. This is the most scalable installation scenario because you can add nodes without necessarily adding another instance of the configuration database. For example, say you have five nodes and three configuration database servers. You could add another five nodes, bringing the total number of nodes to 10, while still using just the three configuration databases to support all 10 nodes.

Installing the configuration database to a separate server has the following advantages:

- More CPU is available to the server for running jobs, process flows, and service requests because reading and writing to the configuration database is handled by a separate server.
- Reduced network traffic due to reduced configuration data replicated between nodes. For example, you could have five Spectrum™ Technology Platform servers for handling execution while having only three repositories. With just three repositories instead of five, data only needs to replicate between three servers rather than five.
- Most scalable installation scenario.



## Cluster Components

A cluster is comprised of these components:

- Load balancer** As requests come into the cluster, the load balancer identifies the best available Spectrum™ Technology Platform node to handle the request. The request is then passed to a Spectrum™ Technology Platform node.
- From the user's perspective, the distributed architecture is handled automatically behind the scenes. The user sends a request to the load balancer's URL and port for Spectrum™ Technology Platform (typically port 80 for a distributed environment) as if it were a single Spectrum™ Technology Platform server.
- Nodes** A node is a Spectrum™ Technology Platform server. A node can contain a copy of the server and configuration database, or if you installed the configuration database to a separate cluster, each node contains either the server or the configuration database. Nodes are continuously synchronized. This enables each node to share the same settings, such as license information, dataflows, and database resources.
- To configure the cluster, simply point Management Console or Enterprise Designer to the load balancer's URL and port for Spectrum™ Technology Platform (typically port 80 for a distributed environment).
- External resources** The definitions for external resources such as database resources (postal databases and geocoding databases for example), JDBC connections, and file servers, reside in the configuration database. The resources themselves (databases, files, web services) can reside anywhere you choose. Database resources can be installed either on each node in the cluster or on a shared network location.
- Because the database resources themselves reside outside the cluster, multiple clusters can share the same database resources. You must create the resource definitions in each cluster using Management Console. For example if you want multiple clusters to share the same geocoding database, you can install the geocoding database on a server accessible from each cluster, then in Management Console point each cluster to the geocoding database.

## Installing a Cluster

For instructions on installing a cluster, see the *Spectrum™ Technology Platform Installation Guide*.

# Using Enterprise Designer with a Cluster

1. Launch Enterprise Designer.
2. In the **Server name** field, enter the server name of the load balancer.
3. In the **Port** field, enter the port that you have configured the load balancer to listen on.

**Note:** Input files, output files and database resources must be on a shared drive, or file server, or some commonly-accessible location. Otherwise, all files must be loaded on each server that hosts a Spectrum™ Technology Platform server and must be located in the same path.

Once you have logged in you can use Enterprise Designer as normal. The actions you take will apply to all Spectrum™ Technology Platform instances in the cluster where you are logged in.

## Starting a Cluster

If all the nodes in a cluster are stopped, you must follow this procedure to start the cluster safely and avoid data loss.

1. On the last node that was stopped last, remove the seed nodes and start the server.

**Warning:** The first node that you start must be the last node that was stopped, and that node must be a seed node. Starting another node first may result in loss of data such as job history and configuration settings. If you do not know which node was stopped last, look in each node's wrapper log for the time stamp of the shutdown message. You can find the wrapper log in:

*Spectrum Location\server\app\repository\logs\wrapper.log.*

- a) Open this file in a text editor:

```
server/app/conf/spectrum-container.properties
```

- b) In the `spectrum.cluster.seeds` property, remove all host names and IP addresses except for the one for this server. Save the host names and IP addresses so that you can re-add them later.
- c) Save the file.
- d) Start the server.
- e) Wait for the Spectrum™ Technology Platform server to *completely* start.

You can tell when the Spectrum™ Technology Platform server has completely started by looking in the wrapper log: *Spectrum*

*Location\server\app\repository\logs\wrapper.log.* This message is displayed when the server is completely started:

```
Pitney Bowes Spectrum(TM) Technology Platform (Version Version Number) Started.
```

- f) In the properties file `spectrum-container.properties`, in the `spectrum.cluster.seeds` property, add the host names or IP addresses that you had removed, separating each with a comma.
- g) Save and close the file. You do not need to restart the server.



2. Start the other nodes in the cluster.

**Warning:** Be sure to wait for the first node to start *completely* before starting additional nodes. Starting additional nodes before the first one is started may result in loss of data.

## Stopping a Cluster

To stop an entire cluster:

1. Identify which nodes are seed nodes. To do this, open the file `SpectrumFolder/server/app/conf/spectrum-container.properties` and look at the nodes listed in the `spectrum.cluser.seeds` property.
2. Stop each Spectrum™ Technology Platform server in the cluster, making sure that the last node you stop is a seed node.
  - On Unix or Linux, change the working directory to the Spectrum™ Technology Platform server's `bin` directory, source the setup file, then type the following command: `./server.stop`.
  - On Windows, right-click the Spectrum™ Technology Platform icon in the Windows system tray and select **Stop Spectrum™**.

**Warning:** To prevent loss of data the last node you stop must be a seed node.

3. Make a note of which node was the last node you stopped. You will need this information when starting up the cluster.

**Warning:** To prevent loss of data when starting the cluster, the first node you start must be the last node that was stopped, and that node must be a seed node.

## Removing a Node from a Cluster

To remove a node from a cluster, stop the Spectrum™ Technology Platform server.

1. Stop the node you want to remove:
  - On Unix or Linux, change the working directory to the Spectrum™ Technology Platform server's `bin` directory, source the setup file, then type the following command: `./server.stop`.
  - On Windows, right-click the Spectrum™ Technology Platform icon in the system tray and select **Stop Spectrum™**.
2. Open the file `server/app/conf/spectrum-container.properties` in a text editor and set `spectrum.cluster.enabled` to `false`.

3. On each of the other nodes in the cluster, open the `spectrum-container.properties` file and remove the node from the `spectrum.cluster.seeds` property.

**For Location Intelligence Module users:** If you want to keep the node standalone and able to run outside the cluster, copy back the original `repository.xml` file and remove the following folders from the `/server/modules/spatial/jackrabbit` directory for each instance of Spectrum™ Technology Platform: `repository`, `version`, `workspaces`. Restart the server and import the repository content.

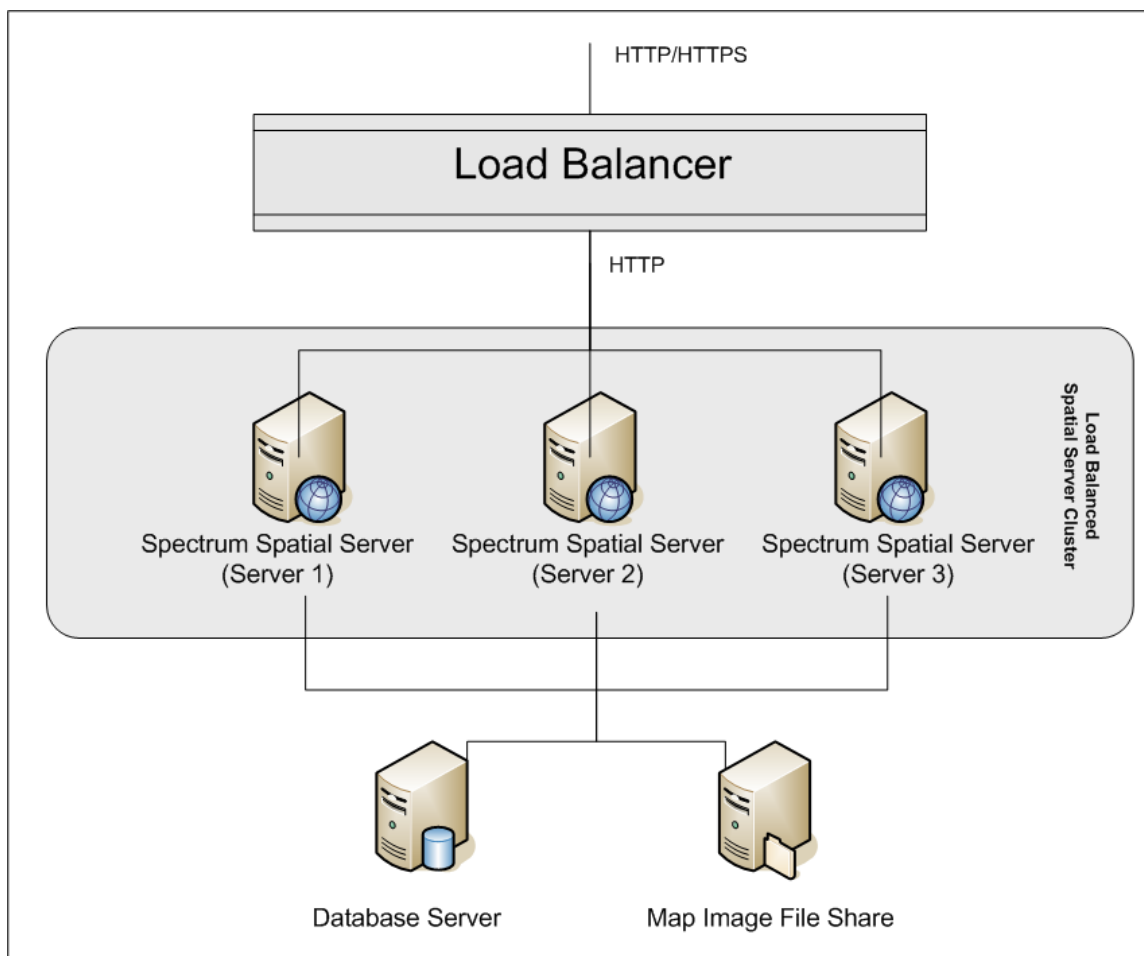
## Managing a Cluster for the Location Intelligence Module

### Clustered Architecture for the Location Intelligence Module

In a clustered environment, processing is shared among two or more instances of the server. The diagram below illustrates the deployment architecture of such a configuration. Load balancing can be used to support high availability and scaling. The deployment architecture includes a load balancer, a Spectrum Spatial cluster, a database, and a file share. With this approach it is possible to scale both horizontally and vertically. You can cluster the Location Intelligence Module with or without platform clustering.

**Note:** Setting up both a Spectrum™ Technology Platform cluster and a Location Intelligence Module cluster is recommended and has several benefits:

- Security (ACL) synchronization happens automatically for named resources .
- Dataflows, users, and roles created on one node will automatically synchronize to all nodes.
- All Location Intelligence Module demo pages and utilities (such as Spatial Manager) can and should point to the load balancer.



### *Load Balancer*

The load balancer spreads requests between the Spectrum Spatial instances. Any load balancer that supports load balancing HTTP/HTTPS requests can be used.

### *Spectrum Spatial Cluster*

The cluster is a collection of Spectrum instances with LIM sharing administration, named resources, geographical metadata content and configuration settings. Additional nodes can be added to the cluster for resilience or to deliver support for greater loads. Each node can be scaled vertically through additional hardware resources and/or additional instances should this be required for hardware with massive resources. Spectrum can be configured to use restricted numbers of CPUs.

### *Database*

Spectrum stores named resources (maps, layers, tables and styles), geographic metadata and configuration in a repository. In the default single server installation an embedded database is used to store these resources on the local server. To create a resilient scalable solution this embedded database should be replaced with a resilient independent database. Oracle, PostgreSQL/PostGIS and Microsoft SQL Server are the supported repository databases.

In the load balanced configuration, Spectrum nodes cache these resources in a local cache and search index in each node in the cluster. When a Spectrum node receives a request it uses the local cache and index to find resources. Named resources can be added through any node in the cluster. Each node keeps its cache current by checking for differences between its local cache and the central database. This check occurs every 2 seconds by default. Time frequency can be configured. This architecture ensures the server delivers high performance transactions and the load on the repository database is kept to a minimum. If a new Spectrum node is added to the cluster the cache and index are created automatically. Such a scenario can occur to remedy a node failure or grow the capability of the deployment.

### File Share

The file share provides a folder to hold map images generated by Spectrum. When maps are rendered using the web services the server supports the map images being returned through URLs or returned as a base 64 encoded image. When a URL is returned the map image is stored as a file and served on request of the URL. To ensure any Spectrum node can return the map image a file share is used to store the images.

## Setting Up a Common Repository Database

You must configure the Location Intelligence Module to use a common repository database for the cluster. This ensures that named resources, geographic metadata and configuration settings are managed across the cluster.

The repository is installed with a set of named resources, geographic metadata and configuration files. To migrate these resources to the common database repository the resources need to be exported from the default internal repository database and reimported into the new shared repository database.

For bulk export and import of repository content, use the `limrepo import` and `limrepo export` commands in the Administration Utility. These commands give you the option of preserving permissions (see the Administration section of the *Spectrum Spatial Guide* for instructions.)

These steps describe how to set up your repository on a common database, either PostgreSQL, Oracle, or Microsoft SQL Server:

1. Export all repository resources to a local folder using the `limrepo export` command in the Administration Utility (see the Administration section of the *Spectrum Spatial Guide* for instructions).

The contents of the installed repository must be exported. This step only needs to be performed once, as the contents of the repository should be the same at this point for all instances of Spectrum™ Technology Platform.

2. Stop the Spectrum™ Technology Platform server on all nodes (for instructions, see [Stopping a Cluster](#) on page 385.)

3. On all nodes of Spectrum™ Technology Platform modify the configuration to specify the common database.
  - a) Copy the contents of `repository.<databaseType>.xml` to `repository.xml` located under the `server/modules/spatial/jackrabbit` folder where `<databaseType>` is the appropriate type for your database (postgres, oracle, or mssql).
  - b) In `repository.xml`:
    - Modify the DataSource section with the server host name, port, database, user, and password.
    - Modify the Cluster section to assign a distinct cluster ID, like Node1. Ensure unique IDs are assigned to every subsequent node in the cluster (for example, Node2, Node3).
    - Save the changes to `repository.xml`.
  - c) Remove these folders from the `/server/modules/spatial/jackrabbit` folder: `repository`, `version`, `workspaces`.
4. If your database has previously contained any repository content, you must remove tables from it so that a clean repository can be created. These tables need to be removed:

|                                  |                              |
|----------------------------------|------------------------------|
| <code>default_binval</code>      | <code>security_binval</code> |
| <code>default_bundle</code>      | <code>security_bundle</code> |
| <code>default_names</code>       | <code>security_names</code>  |
| <code>default_refs</code>        | <code>security_refs</code>   |
| <code>rep_fsentry</code>         | <code>version_binval</code>  |
| <code>rep_global_revision</code> | <code>version_bundle</code>  |
| <code>rep_journal</code>         | <code>version_names</code>   |
| <code>rep_local_revisions</code> | <code>version_refs</code>    |

If using Oracle, then also delete `version_seq_names_id`, `security_seq_names_id`, and `default_seq_names_id`.

5. On the seed node only, import the backed up repository content.
  - a) Start the Spectrum™ Technology Platform server (for instructions, see [Starting a Cluster](#) on page 384).
  - b) Import the contents using the `limrepo import` command, pointing to the seed node.
6. Start the remaining nodes in the cluster (for instructions, see [Starting a Cluster](#) on page 384).

## Configuring Your System

Once the Spectrum™ Technology Platform is installed and you have configured a common repository, you need to configure your instance before you can replicate it to another virtual machine. If you are not using a virtual machine environment, you will need to perform these steps on each of your Spectrum™ Technology Platform installations.

### Configure the Map File Share

To configure the map file share (a shared image folder) to Spectrum™ Technology Platform you first need a shared map image directory. To create a map file share, see [Creating a Map Image File Share on Unix/Linux](#) on page 391 or [Creating a Map Image File Share on Windows](#) on page 391.

Once a map image directory has been created, configure the map file share:

1. Modify the Mapping Service configuration by pointing to a shared image folder and load balance server. In the ImageCache change the Directory parameter to a common image directory, and change the AccessBaseURL parameter to the load balancer machine image URL.

If you are using a virtual machine environment, remember this IP address, as you must set the load balancer VM to this IP address.

For Unix/Linux installations:

```
<ImageCache>
<Directory></spatial server
root>/server/modules/spatial/images</Directory>
<AccessBaseURL>http://<loadbalance_IP_address>/rest/Spatial/MappingService/internal/imageCache</AccessBaseURL>

<FileExpire>30</FileExpire>
<ScanInterval>30</ScanInterval>
</ImageCache>
```

For Windows installations:

```
<ImageCache>
<Directory>\\server\Share\images</Directory>
<AccessBaseURL>http://<loadbalance_IP_address>/rest/Spatial/MappingService/internal/imageCache</AccessBaseURL>

<FileExpire>30</FileExpire>
<ScanInterval>30</ScanInterval>
</ImageCache>
```

2. For Unix/Linux installations, you must set up a symbolic link to enable map images to go to the shared file system.

Create an `images` subfolder in the mounted share folder, e.g., `/mnt/<linux mount>/images`

```
cd /<spatial server root>/server/modules/spatial
rm -Rf images
ln -s /mnt/<linux mount>/images ./images
```

### Creating a Map Image File Share on Unix/Linux

The file share provides a folder to hold map images generated by Spectrum Spatial. Create a shared folder accessible to all Spectrum nodes. The file share is not required if maps are returned from the web services as Base64-encoded images.

To create a map image file share on Unix/Linux:

1. Mount a shared folder on each operating system hosting Spectrum. The commands below mount a drive on a Microsoft Windows Server or network drive supporting CIFS.

```
mkdir /mnt/<linux mount>
mount -t cifs //<windows host>/<windows share> /mnt/<linux mount>-o
username=shareuser,password=sharepassword,domain=pb
```

2. Set the image share to load at startup in `/etc/fstab`.

```
//<windows ip address for share>/share /path_to/mount cifs
username=server_user,password=secret,_netdev 0 0
```

### Creating a Map Image File Share on Windows

The file share provides a folder to hold map images generated by Spectrum Spatial. Create a shared folder accessible to all Spectrum nodes. The file share is not required if maps are returned from the web services as Base64-encoded images.

To create a map image file share on Windows:

1. In Windows Explorer, select the image folder you want to share.
2. Right-click, and then click **Share** or **Share with**.
3. Select the users who will be writing to the image folder. These users must have read/write privileges.

### Modifying OGC Service Configurations for Clustering

To ensure clustering works when you have both a Spectrum™ Technology Platform cluster and a Location Intelligence Module cluster, changes are required to the OGC services configuration files using Spatial Manager: From the WFS, WMS, and WMTS settings pages, change the online resource (service) URL to the IP address and port of the load balancer. See the *Spatial Manager Guide* in the Utilities section of the *Spectrum Spatial Guide* for more information.

## Modifying the Java Properties Files in All Nodes

You must change the java property file in all nodes of the cluster. To modify the java properties for Spectrum™ Technology Platform:

1. Modify the java.properties file, located in `<spectrum>/server/modules/spatial/java.properties`, to point repository.host to localhost.
2. Change the images.webapp.url and all of the service host and port numbers to point to the load balance server.

## Configuring Ports for Multiple Spectrum Instances

If you have multiple Spectrum™ Technology Platform instances on a single machine, you must change the port numbers.

To change the port numbers for each Spectrum™ Technology Platform instance:

1. Change all ports in `<Spectrum root>/server/app/conf/spectrum-container.properties` to new port values that are not in use. The http port reflects the port number entered in the installer.
2. Update the rmi port in bootstrap.properties in the `/<spectrum root>/server/modules/spatial` folder (for example, 11099). The default is 1099.

## Shared Spectrum Local Data

If you are using TAB file data on the file system, this data needs to be in a shared location accessible by all instances of Spectrum in the load balanced environment. It is also important to note that all named resources in the repository accessing data on the file system should point to this shared location.

Each VM or machine hosting Spectrum needs to have access to the mounted shared drive.

**Note:** Using named resources that point to database tables do not require a shared drive, as the named resources in the repository do not access the data using a file path; rather they use a named connection to the data in the database.



# 12 - About Spectrum™ Technology Platform

## In this section

---

|  |     |
|--|-----|
| What Is Spectrum™ Technology Platform?     | 394 |
| Enterprise Data Management Architecture    | 395 |
| Spectrum™ Technology Platform Architecture | 399 |
| Modules and Components                     | 404 |

## What Is Spectrum™ Technology Platform?

Spectrum™ Technology Platform is a system that improves the completeness, validity, consistency, timeliness, and accuracy of your data through data standardization, verification and enhancement. Ensuring that your data is accurate, complete, and up to date enables your firm to better understand and connect with your customers.

Spectrum™ Technology Platform aids in the design and implementation of business rules for data quality by performing the following functions.

### *Parsing, Name Standardization, and Name Validation*

To perform the most accurate standardization you may need to break up strings of data into multiple fields. Spectrum™ Technology Platform provides advanced parsing features that enable you to parse personal names, company names, and many other terms and abbreviations. In addition, you can create your own list of custom terms to use as the basis of scan/extract operations. The Universal Name Module provides this functionality.

### *Deduplication and Consolidation*

Identifying unique entities enables you to consolidate records, eliminate duplicates and develop "best-of-breed" records. A "best-of-breed" record is a composite record that is built using data from other records. The Advanced Matching Module and Data Normalization Module provide this functionality.

### *Address Validation*

Address validation applies rules from the appropriate postal authority to put an address into a standard form and even validate that the address is a deliverable address. Address validation can help you qualify for postal discounts and can improve the deliverability of your mail. The Universal Addressing Module and the Address Now Module provide this functionality.

### *Geocoding*

Geocoding is the process of taking an address and determining its geographic coordinates (latitude and longitude). Geocoding can be used for map generation, but that is only one application. The underlying location data can help drive business decisions. Reversing the process, you can enter a geocode (a point represented by a latitude and longitude coordinate) and receive address information about the geocode. The Enterprise Geocoding Module provides this functionality.

### *Location Intelligence*

Location intelligence creates new information about your data by assessing, evaluating, analyzing and modeling geographic relationships. Using location intelligence processing you can verify locations

and transform information into valuable business intelligence. The Location Intelligence Module provides this functionality.

### *Master Data Management*

Master data management enables you to create relationship-centric master data views of your critical data assets. The Data Hub Module helps you identify influencers and non-obvious relationships, detect fraud, and improve the quality, integration, and accessibility of your information.

### *Tax Jurisdiction Assignment*

Tax jurisdiction assignment takes an address and determines the tax jurisdictions that apply to the address's location. Assigning the most accurate tax jurisdictions can reduce financial risk and regulatory liability.

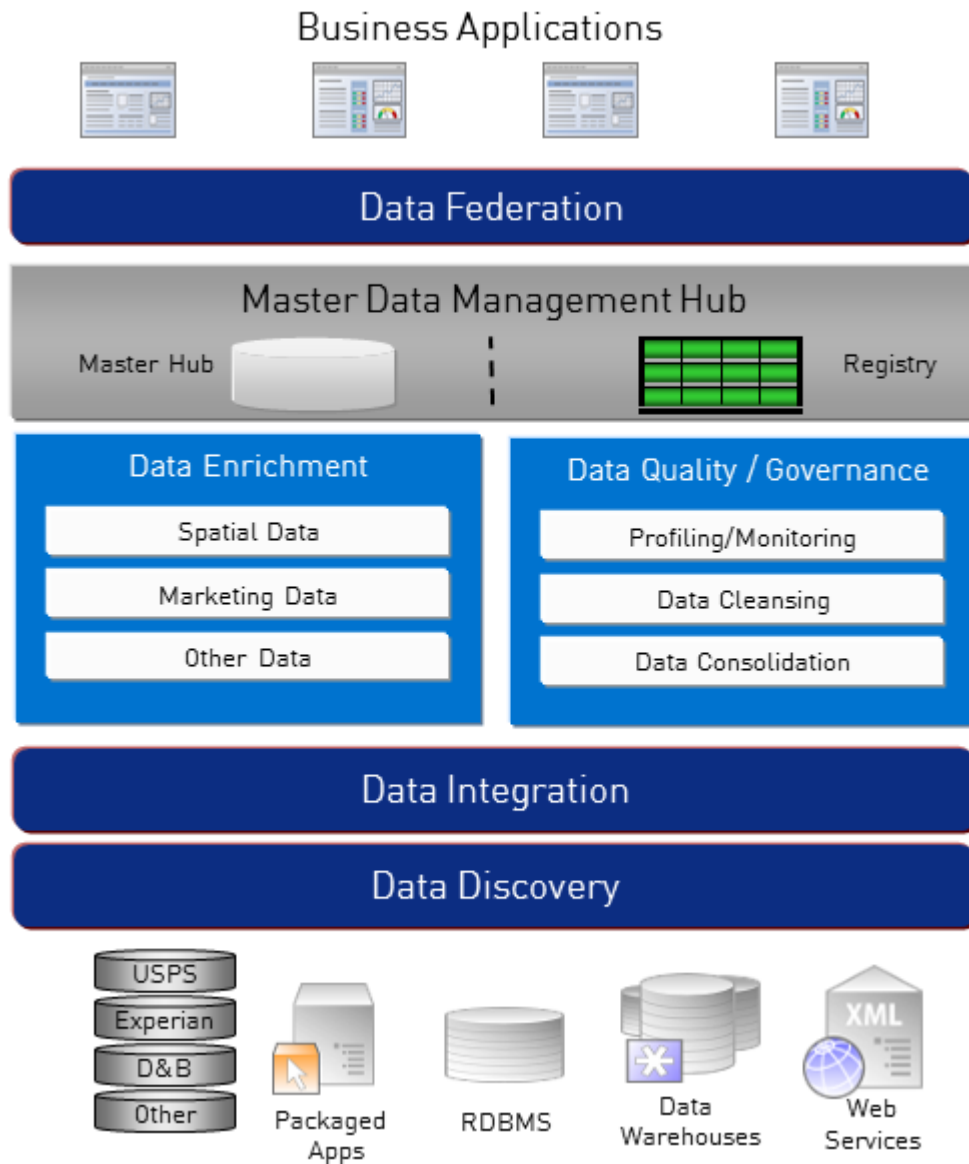
Spectrum™ Technology Platform software from Pitney Bowes integrates up-to-date jurisdictional boundaries with the exact street addresses of your customer records, enabling you to append the correct state, county, township, municipal, and special tax district information to your records. Some example uses of tax jurisdiction assignment are:

- Sales and use tax
- Personal property tax
- Insurance premium tax

The Enterprise Tax Module provides this functionality.

## Enterprise Data Management Architecture

With Spectrum™ Technology Platform, you can build a comprehensive enterprise data management process, or you can use it as a more targeted solution. The following diagram illustrates a complete solution that takes data from its source, through data enrichment and data quality processes, feeding a master data management hub which makes a single view of the data available to multiple business applications.



### Data Discovery

Data discovery is the process of scanning your data resources to get a complete inventory of your data landscape. Spectrum™ Technology Platform can scan structured data, unstructured data, and semi-structured data using a wide array of data profiling techniques. The results of the scan are used to automatically generate a library of documentation describing your company's data assets and to create a metadata repository. This documentation and accompanying metadata repository provide the insight you need before beginning data integration, data quality, data governance, or master data management projects.

For more information on the Spectrum™ Technology Platform Data Discovery Module, contact your account executive.

### *Data Integration*

Once you have an inventory of your data landscape, you need to consider how you will access the data you need to manage. Spectrum™ Technology Platform can connect to data in multiple sources either directly or through integration with your existing data access technologies. It supports batch and real time data integration capabilities for a variety of business needs including data warehousing, data quality, systems integration, and migration. Spectrum™ Technology Platform can access data in RDBMS databases, data warehouses, XML files, flat files, and more. Spectrum™ Technology Platform supports SQL queries with complex joins and aggregations and provides a visual query development tool. In addition, Spectrum™ Technology Platform can access data over REST and SOAP web services.

Spectrum™ Technology Platform can trigger batch processing based on the appearance of one or more source files in a specified folder. This "hot folder" trigger is useful for monitoring FTP uploads and processing them as they occur.

Some of these data integration capabilities require a license for the Enterprise Data Integration Module. For more information, contact your account executive.

Finally, Spectrum™ Technology Platform can integrate with packaged applications such as SAP and Siebel.

### *Data Quality/Governance*

Data quality and data governance processes check your data for duplicate records, inconsistent information, and inaccurate information.

Duplicate matching identifies potential duplicate records or relationships between records, whether the data is name and address in nature or any other type of customer information. Spectrum™ Technology Platform allows you to specify a consistent set of business match rules using boolean matching methods, scoring methods, thresholds, algorithms and weights to determine if a group of records contains duplicates. Spectrum™ Technology Platform supports extensive customization so you can tailor the rules to the unique needs of your business.

Once duplicate records have been identified, you may wish to consolidate records. Spectrum™ Technology Platform allows you to specify how to link or merge duplicate records so you can create the most accurate and complete record from any collection of customer information. For example, a single best-of-breed record can be built from all of the records in a household. The Advanced Matching Module is used to identify duplicates and eliminate them.

Data quality processes also standardize your data. Standardization is a critical process because standardized data elements are necessary to achieve the highest possible results for matching and identifying relationships between records. While several modules perform standardization of one type or another, the Spectrum™ Technology Platform Data Normalization module provides the most comprehensive set of standardization features. In addition, the Universal Name module provides specific data quality features for handling personal name and business name data.

Standardized data is not necessarily accurate data. Spectrum™ Technology Platform can compare your data to known, up-to-date reference data for correctness. The sources used for this process may include regulatory bodies such as the U.S. Postal Service, third-party data providers such as

Experian or D&B, or your company's internal reference sources, such as accounting data. Spectrum™ Technology Platform is particularly strong in address data validation. It can validate or standardize addresses in 250 countries and territories around the world. There are two modules that perform address validation: the Address Now Module and the Universal Addressing Module.

To determine which one is right for you, discuss your needs with your account executive.

While Spectrum™ Technology Platform can automatically handle a wide range of data quality issues, there are some situations where a manual review by a data steward is appropriate. To support this, the Business Steward Module provides a way to specify the rules that will trigger a manual review, and it provides a web-based tool for reviewing exception records. It includes integrated access to third-party tools such as Bing maps and Experian data to aid data stewards in the review and resolution process.

### *Data Enrichment*

Data enrichment processes augment your data with additional information. Enrichment can be based on spatial data, marketing data, or data from other sources that you wish to use to add additional detail to your data. For example, if you have a database of customer addresses, you could geocode the address to determine the latitude/longitude coordinates of the address and store those coordinates as part of the record. Your customer data could then be used to perform a variety of spatial calculations, such as finding the bank branch nearest the customer. Spectrum™ Technology Platform allows you to enrich your data with a variety of information, including geocoding (with the Enterprise Geocoding Module), tax jurisdiction assignment (with the Enterprise Tax Module), geospatial calculations (with the Location Intelligence Module), and driving and walking directions between points (with the Enterprise Routing Module).

### *Master Data Management Hub*

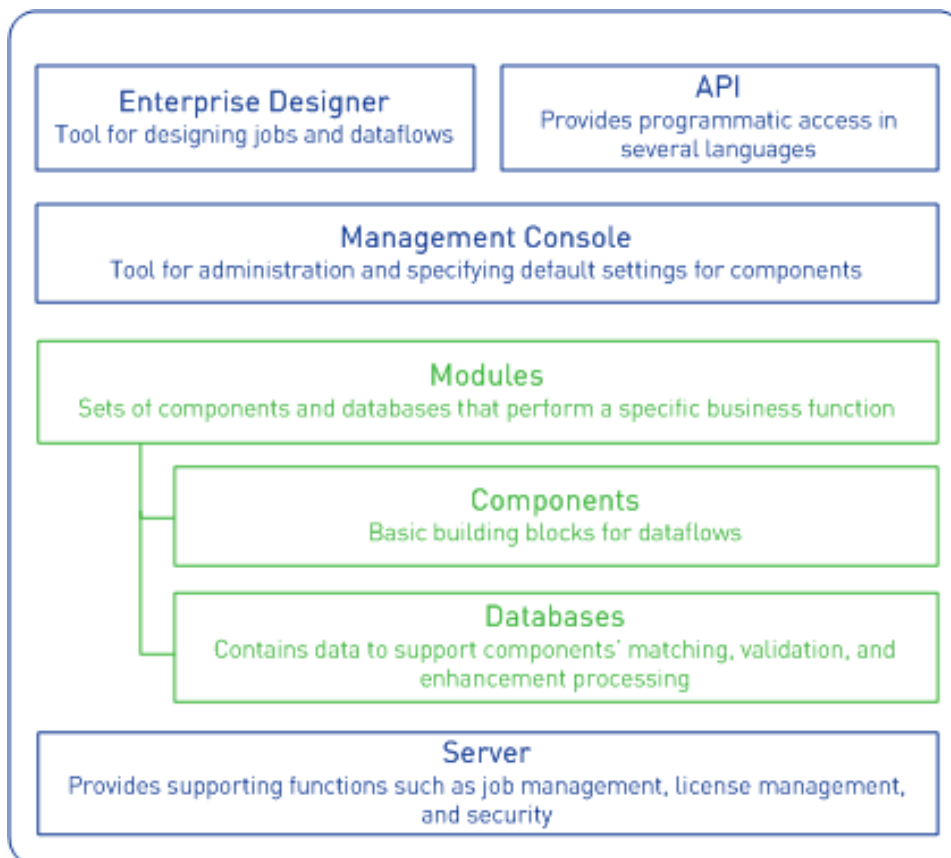
The Master Data Management (MDM) hub allows for rapid modeling of entities and their complex relationships across roles, processes and interactions. It provides built-in social network analysis capabilities to help you understand influencers, predict churn, detect non-obvious relationships and fraudulent patterns, and provide recommendations.

Spectrum™ Technology Platform supports two approaches to the MDM hub. In the master hub approach, the data is maintained in a single MDM database and applications access the data from the MDM database. In the registry approach, the data is maintained in each business application and the MDM hub registry contains keys which are used to find related records. For example, a customer's record may exist in an order entry database and a customer support database. The MDM registry would contain a single key which could be used to access the customer data in both places.

The Data Hub Module provides MDM capabilities.

# Spectrum™ Technology Platform Architecture

Spectrum™ Technology Platform from Pitney Bowes consists of a server that runs a number of modules. These modules provide different functions, such as address validation, geocoding, and advanced parsing, among others. The following diagram illustrates the Spectrum™ Technology Platform architecture.



## Server

The foundation of the Spectrum™ Technology Platform is the server. The server handles data processing, synchronizes repository data, and manages communication. It provides job management and security features.

## Modules

Modules are sets of features that perform a specific function. For example, the Universal Addressing Module standardizes addresses to conform to postal standards. The Enterprise Tax Module determines the tax jurisdictions that apply to a given address. Modules are grouped together to solve common business problems and licensed together as bundles.

## Components

Modules are comprised of components which perform a specific function in a flow or as a service. For example, the Enterprise Geocoding module's Geocode US Address component takes an address and returns the latitude and longitude coordinates for that address; the Universal Addressing module's Get City State Province takes a postal code and returns the city and state or province where that postal code is located.

The components that you have available on your system depend on which Spectrum™ Technology Platform bundle you have licensed.

## Databases

Some modules depend on databases containing reference data. For example, the Universal Addressing module needs to have access to U.S. Postal Service data in order to verify and standardize addresses in the U.S. Databases are installed separately and some are updated on a regular basis to provide you with the latest data.

Modules have both required and optional databases. Optional databases provide data needed for certain features that can enhance your Spectrum™ Technology Platform process.

## Management Console

Management Console is a tool for administering Spectrum™ Technology Platform. You can use Management Console to:

- Define the connections between Spectrum™ Technology Platform and your data
- Specify the default settings for services and flows
- Manage user accounts, including permissions and passwords
- View logs
- View licenses including license expiration information



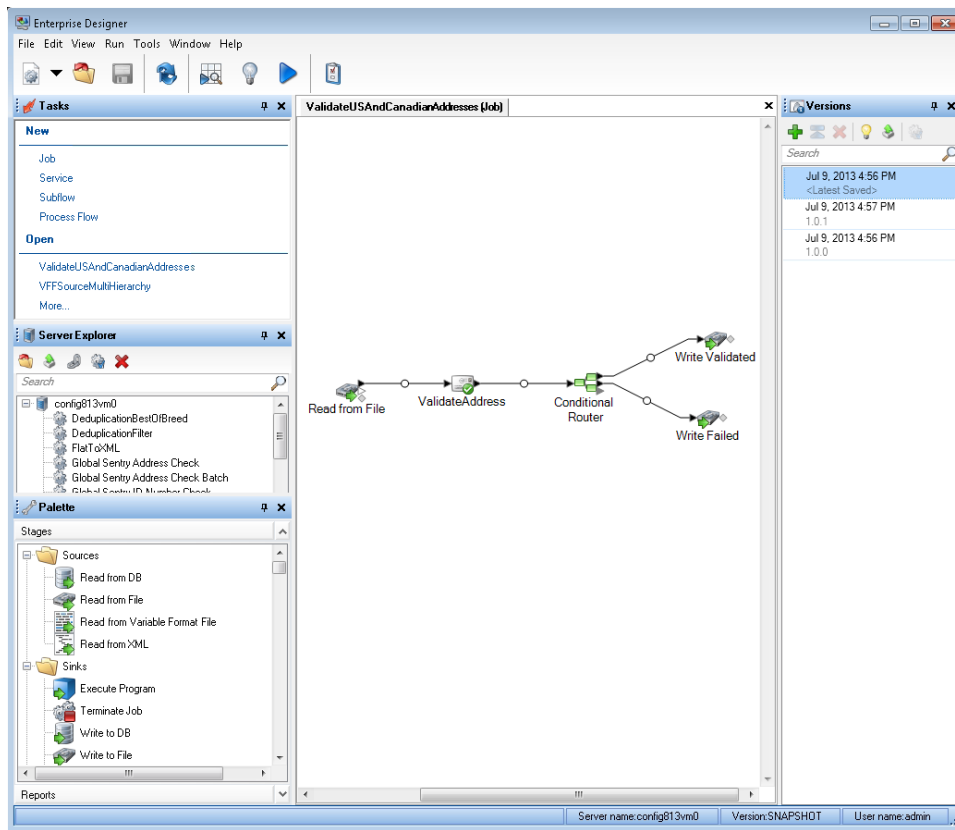
The screenshot shows the Management Console interface. The top navigation bar includes 'Management Console', 'Flows', 'Services', 'Resources', and 'System'. The user is logged in as 'admin'. The breadcrumb trail is 'Home > Resources: Data Sources'. The main heading is 'Data Sources'. Below the heading are several action icons (add, edit, refresh, delete, share) and a search filter box. A table lists five data sources with columns for 'Name' and 'Type'. The table content is as follows:

| Name          | Type  |
|---------------|-------|
| test1         | FTP   |
| test2         | FTP   |
| test4         | Cloud |
| test5HDFS     | HDFS  |
| mdg1teamcity1 | FTP   |

Below the table, it indicates 'Showing 5 of 5 records' and 'Rows per page' set to 10. The footer contains the Pitney Bowes logo and copyright information: '© 2017 Pitney Bowes Inc.'

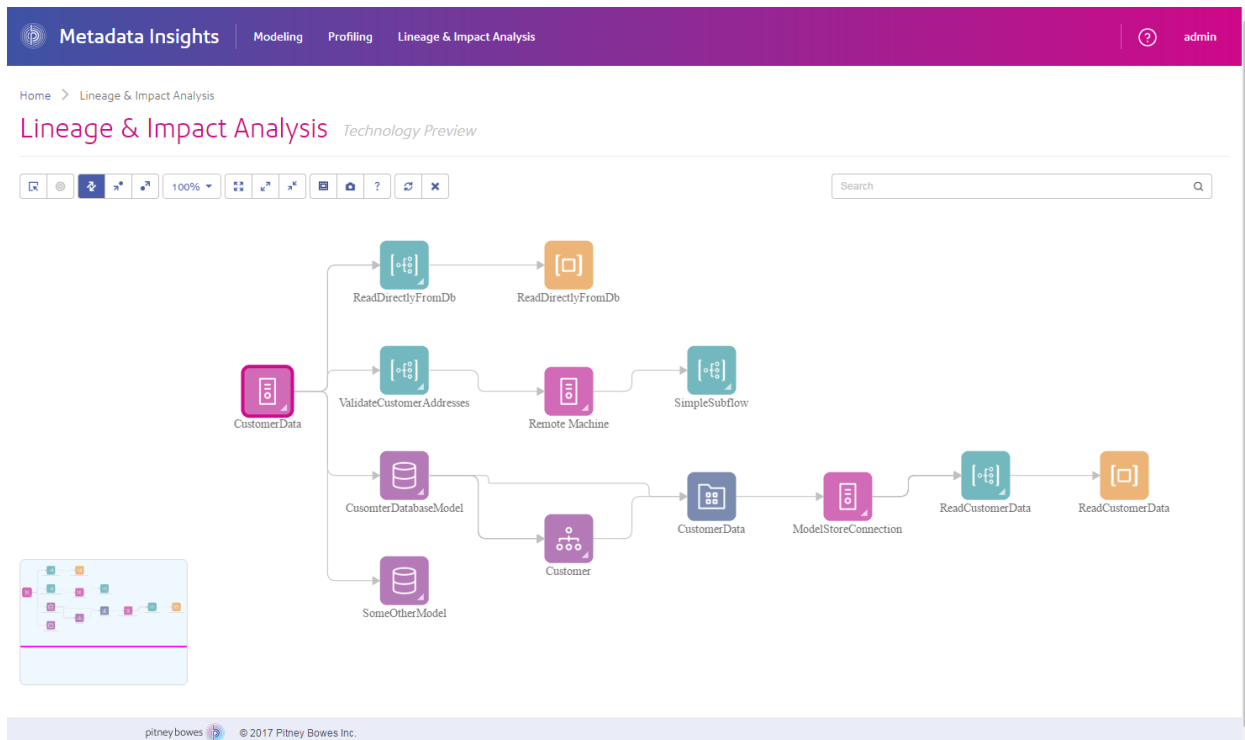
### *Enterprise Designer*

Enterprise Designer is a tool for creating Spectrum™ Technology Platform jobs, services, subflows, and process flows. It utilizes an easy drag-and-drop interface to allow you to graphically create complex dataflows.



## Metadata Insights

Metadata Insights gives you the control you need to deliver accurate and timely data-driven insights to your business. Use Metadata Insights to develop data models, view the flow of data from source to business application, and assess the quality of your data through profiling. With this insight, you can identify the data resources to use to answer particular business questions, adapt and optimize processes to improve the usefulness and consistency of data across your business, and troubleshoot data issues.



### Web Services and API

You can integrate Spectrum™ Technology Platform capabilities into your applications using web services and programming APIs. These interfaces provide simple integration, streamline record processing, and support backward compatibility of future versions.

The Spectrum™ Technology Platform API is available for these languages:

- C
- C++
- COM
- Java
- .NET

Web services are available via SOAP and REST.

## Modules and Components

| Module                   | Description  | Components  |
|--------------------------|--|---|
| Address Now Module       | Provides enhanced validation and standardization for addresses outside the U.S., and other address processing. | Build Global Address<br>Get Global Candidate Addresses<br>Validate Global Address   |
| Advanced Matching Module | Matches records within and between input files.  | Best Of Breed<br>Candidate Finder<br>Duplicate Synchronization<br>Filter<br>Interflow Match<br>Intraflow Match<br>Match Key Generator<br>Transactional Match<br>Write to Search Index |
| Analytics Scoring Module | Enriches data by adding predicted outputs derived through model scoring.                                       | PMML Model Scoring<br>Read from Miner Dataset<br>Write to Miner Dataset<br>Analytics Scoring Repository   |
| Business Steward Module  | Identifies exception records and provides a browser-based tool for manually reviewing exception records.       | Exception Monitor<br>Read Exceptions<br>Write Exceptions<br>Business Steward Portal   |

| Module                      | Description  | Components   |
|-----------------------------|--|--|
| Country Identifier          | Takes a country name or a combination of postal code and state/province and returns the two-character ISO country code, the three-character Universal Postal Union (UPU) code, and the English country name. | Country Identifier   |
| Data Hub Module             | Links and analyzes data, identifying relationships and trends.   | Delete from Hub<br>Import to Hub<br>Merge Entities<br>Query Hub<br>Read From Hub<br>Split Entity<br>Write to Hub<br>Data Hub Visualization<br>Data Hub Browser<br>Relationship Analysis Client |
| Data Normalization Module   | Removes inconsistencies in data.   | Advanced Transformer<br>Open Parser<br>Table Lookup<br>Transliterator  |
| Enterprise Data Integration | Connects to data in multiple sources for a variety of business needs including data warehousing, data quality, systems integration, and migration.   | Call Stored Procedure<br>Field Selector<br>Generate Time Dimension<br>Query Cache<br>Write to Cache  |

| Module                      | Description  | Components   |
|-----------------------------|--|--|
| Enterprise Geocoding Module | Determines the geographic coordinates for an address. Also determines the address of a given latitude and longitude.   | Geocode Address AUS<br>Geocode Address GBR<br>Geocode Address Global<br>Geocode Address World<br>Geocode US Address<br>GNAF PID Location Search<br>Reverse APN Lookup<br>Reverse Geocode Address Global<br>Reverse Geocode US Location |
| Enterprise Routing Module   | Obtains driving or walking directions, calculates drive time and drive distance, and identifies locations within a certain time or distance from a starting point.   | Get Route Data<br>Get Travel Boundary<br>Get Travel Cost Matrix<br>Get Travel Directions<br>Persistent Update  |
| Enterprise Tax Module       | Determines the tax jurisdictions that apply to a given location.   | Assign GeoTAX Info<br>Calculate Distance   |
| GeoComplete Module          | Automatically suggests addresses or points of interest based on your input. You can search on address, POI name or POI Categories. This "as you type" recognition allows for searching based on your location. | International GeoComplete  |
| GeoConfidence Module        | Determines the probability that an address or street intersection is within a given area.  | Geo Confidence Surface<br>CreatePointsConvexHull   |
| Global Addressing Module    | Provides enhanced address standardization and validation for addresses outside the U.S. Also, automatically suggests addresses as you type and immediately returns candidates based on your input.             | Global Address Validation<br>Global Type-Ahead   |

| Module                        | Description  | Components  |
|-------------------------------|--|---|
| Global Geocoding Module       | Determines the geographic coordinates for an address. Also determines the address of a given latitude and longitude. | Global Geocode<br>Global Reverse Geocode  |
| Global Sentry                 | Attempts to match transactions against government-provided watch lists that contain data from different countries.   | Global Sentry<br>Global Sentry Address Check<br>Global Sentry ID Number Check<br>Global Sentry Name Check<br>Global Sentry Other Data Check |
| Information Extraction Module | Provides advanced text processing capabilities and information extraction from natural language input text.          | Read from Documents<br>Entity Extractor<br>Text Categorizer<br>Relationship Extractor   |
| Location Intelligence Module  | Performs point in polygon and radial analysis against a variety of geospatial databases.                             | Closest Site<br>Find Nearest<br>Point In Polygon<br>Query Spatial Data<br>Read Spatial Data<br>Spatial Calculator<br>Spatial Union          |
| Machine Learning Module       | Provides the ability to fit supervised and unsupervised machine learning models.                                     | Binning<br>K Means Clustering<br>Logistic Regression<br>Java Model Scoring<br>Machine Learning Model Management                             |

| Module        | Description   | Components   |
|---------------|---|--|
| SAP Module    | Enables Spectrum™ Technology Platform to interface with SAP Customer Relationship Management Module applications. | <ul style="list-style-type: none"> <li>SAP Generate Match Key</li> <li>SAP Generate Match Score</li> <li>SAP Generate Search Key</li> <li>SAP Generate Search Key Constant</li> <li>SAP Generate Search Key Metaphone</li> <li>SAP Generate Search Key Substring</li> <li>SAP Validate Address With Candidates</li> </ul>  |
| Siebel Module | Enables Spectrum™ Technology Platform to interface Siebel applications.   | <ul style="list-style-type: none"> <li>Siebel Generate Match Key</li> <li>Siebel Generate Match Score</li> <li>Siebel Generate Search Key</li> <li>Siebel Business Name Standardization</li> <li>Siebel Standardize Name</li> <li>Siebel Geocode US Address With Candidates</li> <li>Siebel Geocode US Address With No Candidates</li> <li>Siebel Get Global Candidate Addresses</li> <li>Siebel Validate Address With Candidates</li> <li>Siebel Validate Address With No Candidates</li> </ul> |



| Module                      | Description  | Components  |
|-----------------------------|--|---|
| Universal Addressing Module | Standardizes and validates addresses according to the postal authority's standards.      | Auto Complete Loqate<br>Get Candidate Addresses<br>Get Candidate Addresses Loqate<br>Get City State Province<br>Get City State Province Loqate<br>Get Postal Codes<br>Get Postal Codes Loqate<br>Validate Address<br>Validate Address AUS<br>Validate Address Global<br>Validate Address Loqate |
| Universal Name Module       | Parses personal names, company names, addresses, and many other terms and abbreviations. | Name Parser (Deprecated)<br>Name Variant Finder<br>Open Name Parser   |

# Notices

© 2017 Pitney Bowes Software Inc. All rights reserved. MapInfo and Group 1 Software are trademarks of Pitney Bowes Software Inc. All other marks and trademarks are property of their respective holders.

### *USPS® Notices*

Pitney Bowes Inc. holds a non-exclusive license to publish and sell ZIP + 4® databases on optical and magnetic media. The following trademarks are owned by the United States Postal Service: CASS, CASS Certified, DPV, eLOT, FASTforward, First-Class Mail, Intelligent Mail, LACS<sup>Link</sup>, NCOA<sup>Link</sup>, PAVE, PLANET Code, Postal Service, POSTNET, Post Office, RDI, Suite<sup>Link</sup>, United States Postal Service, Standard Mail, United States Post Office, USPS, ZIP Code, and ZIP + 4. This list is not exhaustive of the trademarks belonging to the Postal Service.

Pitney Bowes Inc. is a non-exclusive licensee of USPS® for NCOA<sup>Link</sup>® processing.

Prices for Pitney Bowes Software's products, options, and services are not established, controlled, or approved by USPS® or United States Government. When utilizing RDI™ data to determine parcel-shipping costs, the business decision on which parcel delivery company to use is not made by the USPS® or United States Government.

### *Data Provider and Related Notices*

Data Products contained on this media and used within Pitney Bowes Software applications are protected by various trademarks and by one or more of the following copyrights:

© Copyright United States Postal Service. All rights reserved.  
 © 2014 TomTom. All rights reserved. TomTom and the TomTom logo are registered trademarks of TomTom N.V.

© 2016 HERE

Fuente: INEGI (Instituto Nacional de Estadística y Geografía)

Based upon electronic data © National Land Survey Sweden.

© Copyright United States Census Bureau

© Copyright Nova Marketing Group, Inc.

Portions of this program are © Copyright 1993-2007 by Nova Marketing Group Inc. All Rights Reserved

© Copyright Second Decimal, LLC

© Copyright Canada Post Corporation

This CD-ROM contains data from a compilation in which Canada Post Corporation is the copyright owner.

© 2007 Claritas, Inc.

The Geocode Address World data set contains data licensed from the GeoNames Project ([www.geonames.org](http://www.geonames.org)) provided under the Creative Commons Attribution License ("Attribution

License") located at <http://creativecommons.org/licenses/by/3.0/legalcode>. Your use of the GeoNames data (described in the Spectrum™ Technology Platform User Manual) is governed by the terms of the Attribution License, and any conflict between your agreement with Pitney Bowes Software, Inc. and the Attribution License will be resolved in favor of the Attribution License solely as it relates to your use of the GeoNames data.



3001 Summer Street  
Stamford CT 06926-0700  
USA

[www.pitneybowes.com](http://www.pitneybowes.com)