

Spectrum™ Technology Platform

Version 12.0 SP2

Guía de Enterprise Data Integration



Contents

1 - Introducción

Arquitectura del proceso de gestión de datos empresariales	5
El diseño del almacén de datos con esquema de estrella	8

2 - Conexión a orígenes de datos y a almacenes de datos

Conexiones al origen de datos	14
Definición de conexiones	14
Soporte de compresión para servidores de archivos de nube	68
Eliminación de una conexión	69

3 - Completar el almacén de datos

Preparación de los datos	71
Cómo completar una tabla de dimensiones de tiempo	72
Cómo completar una tabla de dimensiones	73
Cómo completar una tabla de hechos	75
Agregar una marca de tiempo a los registros de almacenamiento de datos	80

4 - Actualizar el almacén de datos

Definición de un programa de actualización del almacén de datos	83
---	----

Cómo actualizar una tabla de hechos	84
Uso de una memoria caché global para las consultas	89
Uso de una caché local para consultas	90

5 - Referencia de etapas

Invocar procedimiento almacenado	94
DB Change Data Reader	96
DB Loader	100
Field Parser	110
Field Combiner	113
Field Selector	114
Generate Time Dimension	115
Query Cache	123
Query DB	136
Query NoSQL DB	140
Read from DB	143
Read From File	152
Read from Hadoop Sequence File	170
Read From Hive File	175
Archivo Read From HL7	178
Read from NoSQL DB	190
Read from SAP	194
Leer de la hoja de cálculo	200
Read From Variable Format File	202
Read from XML	217
SQL Command	225
Transposer	235
Unique ID Generator	239
Write to Cache	248
Write to DB	250
Write to File	255
Write to Hadoop Sequence File	274
Escribir en archivo Hive	278
Write to NoSQL DB	285
Write to Spreadsheet	288

Write to Variable Format File	292
Write to XML	302
Patrones de fecha y número	312

6 - Configuraciones

Configuraciones de Oracle LogMiner	319
------------------------------------	-----

7 - Cómo optimizar el rendimiento

Determinar un tamaño de recuperación óptimo	322
---	-----

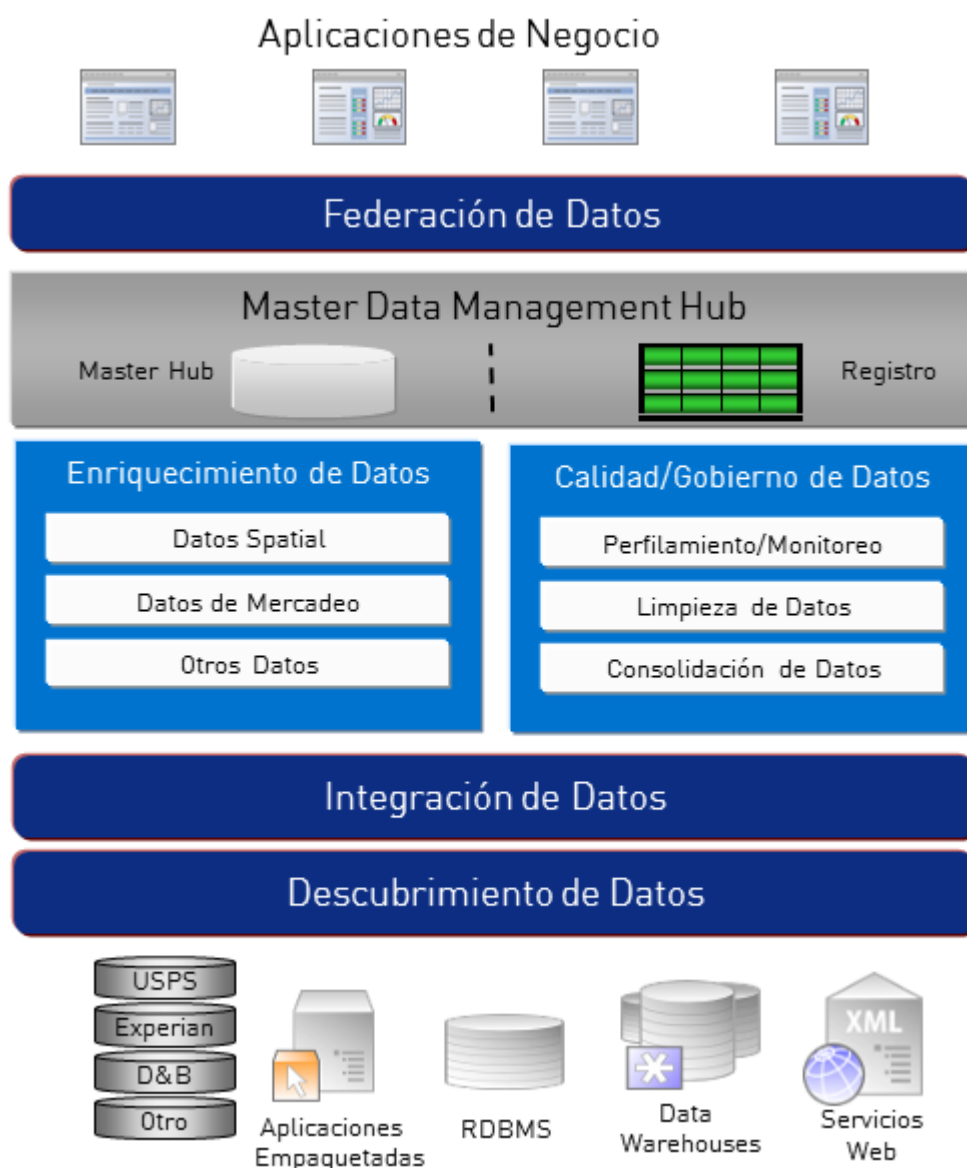
1 - Introducción

In this section

Arquitectura del proceso de gestión de datos empresariales	5
El diseño del almacén de datos con esquema de estrella	8

Arquitectura del proceso de gestión de datos empresariales

Con Spectrum™ Technology Platform, puede desarrollar un proceso integral para la gestión de los datos empresariales o bien usarlo como una solución más específica. El siguiente diagrama muestra una solución completa, que toma los datos desde su origen y los guía a través de los procesos de calidad y enriquecimiento para incorporarlos a un centro de gestión de datos maestros (Master Data Management, MDM), que ofrece una única vista centralizada de los datos disponibles para múltiples aplicaciones comerciales.



Descubrimiento de Datos

El descubrimiento de datos es el proceso de análisis de los recursos de datos para obtener un inventario completo de su entorno de datos. Spectrum™ Technology Platform puede analizar datos estructurados, datos no estructurados y datos semiestructurados valiéndose de una amplia gama de técnicas de perfiles de datos. Los resultados de la exploración se utilizan para generar automáticamente una biblioteca de documentación en la que se describan los datos de los activos de su empresa y para crear un repositorio de metadatos. Esta documentación y el repositorio de metadatos que la acompaña proporcionan la información que necesita antes de comenzar proyectos de integración de datos, calidad de datos, administración de datos, o de administración de datos maestros.

Para obtener más información sobre el módulo de Descubrimiento de Datos Spectrum™ Technology Platform, póngase en contacto con su ejecutivo de cuenta.

Data Integration

Una vez que haya un inventario de su entorno de datos, tiene que tener en cuenta la forma en que se tendrá acceso a los datos que necesita administrar. Spectrum™ Technology Platform puede conectarse a datos de varios orígenes ya sea directamente o por medio de la integración con sus actuales tecnologías de acceso a datos. Admite capacidades de integración de datos por lotes y en tiempo real para una variedad de requisitos empresariales, como almacenamiento de datos, calidad de datos, integración de sistemas y migración. Spectrum™ Technology Platform puede tener acceso a los datos de bases de datos RDBMS, almacenes de datos, archivos XML, archivos planos, etc. Spectrum™ Technology Platform admite consultas SQL con combinaciones y adiciones complejas y agregaciones y proporciona una herramienta de desarrollo de consulta visual. Además, Spectrum™ Technology Platform puede acceder a los datos a través de los servicios web REST y SOAP.

Spectrum™ Technology Platform puede activar el procesamiento por lotes basado en el aspecto de uno o más archivos de origen en una carpeta especificada. Esta activación por carpeta resulta útil para supervisar las cargas en FTP y procesarlas a medida que se producen.

Algunas de estas funciones de integración de datos requieren una licencia para el módulo Enterprise Data Integration. Para obtener más información, contacte a su ejecutivo de cuenta.

Por último, Spectrum™ Technology Platform puede integrarse en aplicaciones en paquete como SAP.

Calidad y administración de datos

Los procesos de calidad y administración de datos permiten verificar los datos para detectar registros duplicados e información contradictoria o inexacta.

El cruce de duplicados identifica posibles registros o relaciones duplicadas entre registros, ya sea que los datos sean de índole "nombre y dirección", o cualquier otro tipo de información del cliente. Spectrum™ Technology Platform permite especificar un conjunto coherente de reglas de cruce comerciales usando métodos de cruce booleanos, métodos de calificación, umbrales, algoritmos y ponderaciones para determinar si un grupo de registros contiene duplicados. Spectrum™

Technology Platform admite amplias opciones de personalización para que pueda adaptar las reglas a las necesidades únicas de su empresa.

Una vez identificados los registros duplicados, es posible que desee consolidarlos. Spectrum™ Technology Platform permite especificar la forma de vincular o combinar registros duplicados para que pueda crear el registro más preciso y completo a partir de cualquier colección de información de clientes. Por ejemplo, es posible elaborar un solo registro best-of-breed (mejor de la especie) a partir de todos los registros correspondientes a un hogar. El módulo Advanced Matching se usa para identificar duplicados y eliminarlos.

Los procesos de calidad de datos también estandarizan los datos. La estandarización es un proceso crítico porque los elementos de datos estandarizados son necesarios para lograr los mejores resultados posibles al comparar e identificar relaciones entre registros. Si bien varios módulos ejecutan funciones de estandarización de uno u otro tipo, el módulo Spectrum™ Technology Platform Data Normalization (Normalización de datos) ofrece el más completo conjunto de funciones de estandarización. Además, el módulo Universal Name ofrece funciones específicas de calidad de datos para el manejo de datos en relación con nombres personales y de empresas.

Los datos estandarizados no son necesariamente exactos. Spectrum™ Technology Platform puede comparar sus datos con datos de referencia conocidos y actualizados para su corrección. Las fuentes utilizadas para este proceso pueden incluir organismos regulatorios (como el Servicio Postal de los Estados Unidos, proveedores de datos de terceros como Experian o D&B o fuentes de referencia interna de la empresa, como los datos de contabilidad. Spectrum™ Technology Platform es particularmente potente en la validación de datos de dirección. Puede validar o estandarizar direcciones de 250 países y territorios de todo el mundo. Hay dos módulos que realizan validación de direcciones: los módulos Address Now y Universal Addressing.

Para determinar cuál es el módulo adecuado para su caso, analice sus necesidades con su ejecutivo de cuentas.

Si bien Spectrum™ Technology Platform puede manejar de forma automática una amplia variedad de problemas de calidad de datos, en algunas situaciones es más adecuado que un administrador de datos realice una revisión manual. Para permitir este procedimiento, el módulo Business Steward ofrece un método para especificar las reglas que activan una revisión manual y brinda una herramienta web para analizar los registros de excepciones. Esto incluye el acceso integrado a herramientas de terceros tales como mapas Bing y datos a fin de asistir a los administradores de datos en el proceso de revisión y resolución.

Enriquecimiento de datos

Los procesos de enriquecimiento de datos amplían los datos mediante información adicional. El enriquecimiento puede tomar como base datos espaciales, datos de marketing o datos de otras fuentes que se desean utilizar para agregar detalles adicionales a los datos. Por ejemplo, si usted tiene una base de datos de direcciones de clientes, puede geocodificar una dirección para determinar las coordenadas de latitud/ longitud de esa dirección y almacenar esas coordenadas como parte del registro. Los datos de los clientes pueden utilizarse entonces para llevar a cabo una variedad de cálculos espaciales, tales como encontrar la sucursal bancaria más cercana al cliente. Spectrum™ Technology Platform permite enriquecer sus datos con una gran variedad de información, que incluye geocodificación (con el módulo Enterprise Geocoding), asignación de jurisdicción tributaria

(con el módulo Enterprise Tax), cálculos geoespaciales (con el módulo Location Intelligence) y la direcciones a pie y de conducción entre puntos (con el módulo Enterprise Routing).

Centro de gestión de datos maestros

El centro de gestión de datos maestros (Master Data Management, MDM) permite elaborar rápidamente modelos de las entidades y sus complejas relaciones a través de funciones, procesos e interacciones. Su capacidad integrada de análisis de redes sociales ayuda a comprender a los usuarios influyentes, predecir la rotación, detectar relaciones que no son obvias y patrones fraudulentos, y ofrecer recomendaciones.

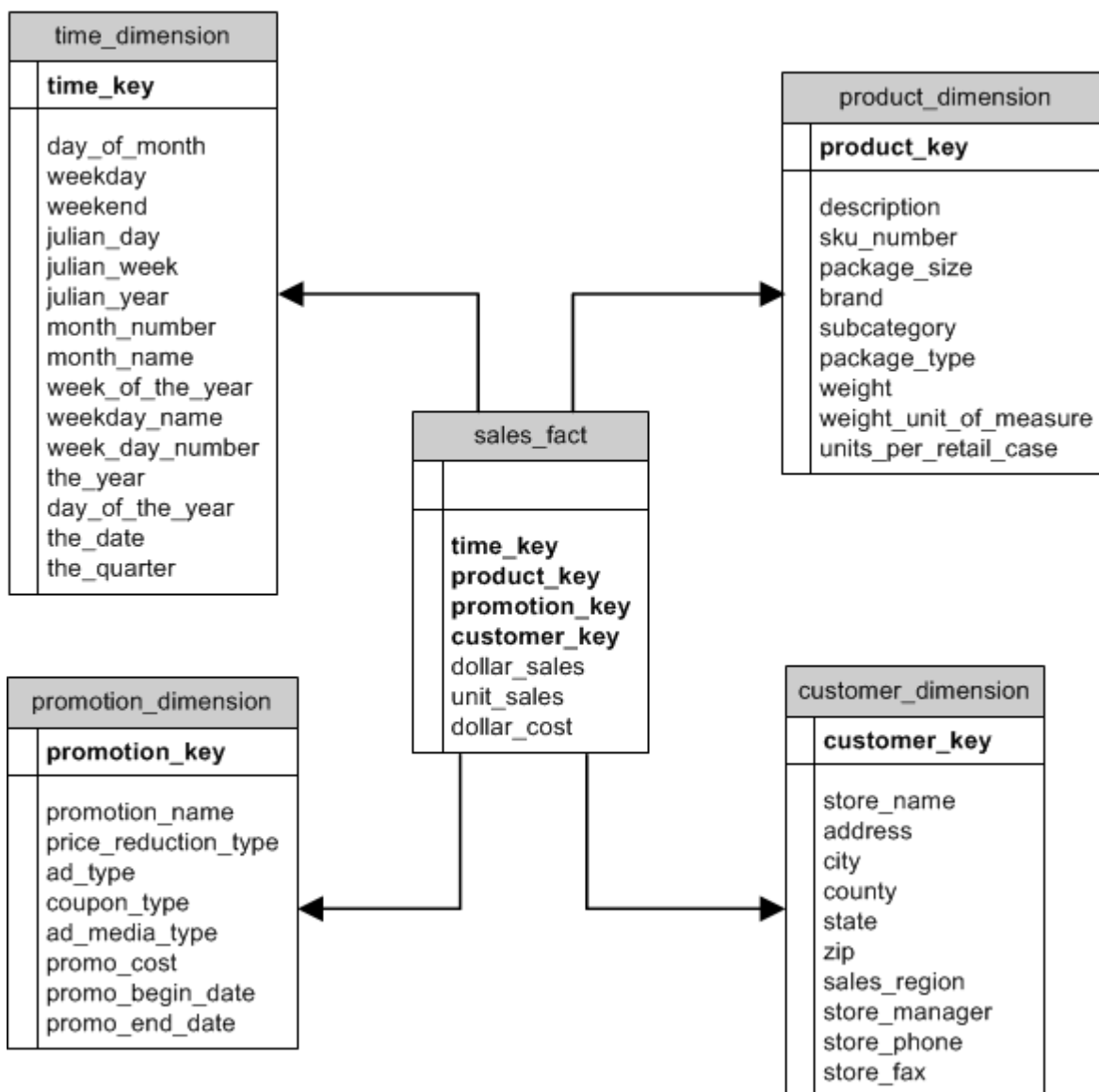
Spectrum™ Technology Platform admite dos enfoques respecto del centro MDM. El enfoque de centro maestro, los datos se mantienen en una única base de datos de MDM, y las aplicaciones acceden a los datos desde la base de datos de MDM. En el enfoque de registro, los datos se mantienen en cada aplicación empresarial y el registro del centro MDM contiene claves que se usan para encontrar registros relacionados. Por ejemplo, el registro de un cliente puede existir en una base de datos de entrada de pedidos y en una base de datos de asistencia al cliente. En ese caso, el registro MDM incluirá una sola clave que puede usarse para acceder a los datos del cliente en ambos lugares.

El módulo Data Hub brinda funciones MDM.

El diseño del almacén de datos con esquema de estrella

Spectrum™ Technology Platform admite la creación y el mantenimiento de almacenes de datos que utilizan un diseño de esquema de estrella. En un esquema en estrella, los datos se almacenan ya sea como hechos, es decir, las descripciones específicas de un evento, o como atributos dimensionales, es decir, las descripciones de los hechos en la tabla de hechos. Los hechos cambian en forma periódica y las dimensiones, o bien cambian lentamente, o no cambian nunca.

La siguiente ilustración muestra el diseño de un esquema de estrella:



Esta ilustración muestra las características principales de un esquema de estrella: una tabla de hechos, tablas de dimensión y las combinaciones.

Tabla de hechos

Las tablas de hechos son las tablas centrales del esquema de estrella de su almacén de datos. Las tablas de hechos suelen contener información numérica o cuantitativa (denominada medidas) que describe un evento específico. Por ejemplo, si usted tiene un almacén de datos que se utiliza para generar un informe sobre los ingresos de la compañía, tendrá como columnas **dollar_sales** y **dollar_cost** en su tabla de hechos, tal como se indicó en la anterior ilustración. Normalmente, los datos se evalúan en forma continua, y son aditivos. "Evaluado en forma continua" significa que el hecho es una medida numérica que arroja un valor cada vez que se mide. "Aditivo" significa que el hecho puede resumirse mediante la adición.

Las tablas de hechos contienen además un conjunto de columnas que forman una clave concatenada, o compuesta. Cada una de las columnas de la clave concatenada es una clave externa extraída de una clave principal de una tabla de dimensión. Por ejemplo, en la ilustración anterior, la tabla de hechos contiene una columna `product_key` que asocia el hecho a un producto específico en la tabla `product_dimension`.

El nivel de detalle de una tabla de hechos se llama grano. Cada fila de la tabla de hechos debe registrarse con el mismo nivel de detalle. En el diagrama anterior, las mediciones en la tabla de hechos son los totales diarios de ventas en dólares, ventas por unidades, y el costo en dólares de cada producto vendido. El grano es diario. Cada registro de la tabla de hechos representa el total de las ventas de un producto específico durante un único día en una tienda minorista. Cada nueva combinación de producto, tienda, o día genera un registro diferente en la tabla de hechos.

Las tablas de hechos se completan con los datos extraídos de un origen de datos. El origen de datos puede ser un sistema OLTP o un almacén de datos. Spectrum™ Technology Platform toma una instantánea de los datos de origen en un horario regular y mueve los datos hacia el almacén de datos, por lo general a la misma hora cada día, semana o mes.

Un esquema de estrella puede tener varias tablas de hechos. Utilice un esquema con varias tablas de hechos para separar conjuntos de mediciones que comparten un subconjunto común de tablas de dimensiones, o para realizar un seguimiento de mediciones con diferentes granos.

Tabla de dimensiones

Las tablas de dimensiones almacenan datos que describen la información contenida en la tabla de hechos. Por ejemplo, si el total de ventas, o `sales_total`, de un mes difiere del mes siguiente, se consultan las dimensiones para averiguar la razón. La misma tabla de dimensiones puede utilizarse con diferentes tablas de hechos.

Las tablas de dimensión tienen atributos y una clave principal de una sola parte que une la tabla de dimensión con la tabla de hechos. Los atributos son las columnas de la tabla de dimensiones. La clave principal de una sola parte permite que explore con rapidez una única tabla de dimensión. Explorar una tabla de dimensión puede ayudar a determinar la mejor manera de consultar la tabla de hechos.

Las tablas de dimensiones de tiempo son necesarias para realizar cálculos precisos basados en tiempo porque a veces no es fácil extraer los datos de fecha necesarios de los registros. Por ejemplo, los siguientes registros se encuentran en una base de datos de ventas. Tenga en cuenta que hay diferencias de tiempo entre registros. Por ejemplo, no hay ningún registro para el día 1/4/2012.

Fecha	Producto	Monto
03/01/2012	Red Shirt	\$10.00
05/01/2012	Red Shirt	\$5.00
07/01/2012	Red Shirt	\$15.00

Si consulta estos registros y calcula el promedio de ventas por día, la respuesta sería \$10.00 (\$30 / 3 registros). Sin embargo, esto no es correcto porque los tres registros en realidad abarcan un período de cinco días. Si tiene una tabla de dimensiones de tiempo con un registro para cada día, puede combinar esta tabla con la tabla anterior para obtener:

Fecha	Producto	Monto
03/01/2012	Red Shirt	\$10.00
04/01/2012		
05/01/2012	Red Shirt	\$5.00
06/01/2012		
07/01/2012	Red Shirt	\$15.00

Al calcular el promedio de ventas por día utilizando estos registros, obtendrá la respuesta correcta: \$6.00 (\$30 / 5 días).

Además, podría tener en cuenta los atributos de tiempo arbitrarios, tales como días festivos, fines de semana y trimestres en el cálculo. Por ejemplo, si 1/6/2012 fue un día festivo y a usted solo le interesan las ventas promedio por día laboral, entonces la respuesta sería \$7.50.

Combinaciones

Combinaciones define las relaciones entre una tabla de hechos y las tablas de dimensión en el esquema de estrella. La clave principal de la tabla de dimensiones es la clave externa de la tabla de hechos. La tabla de hechos debe contener un valor de clave primaria de cada tabla de dimensiones. La referencia de la clave externa con respecto de la clave principal es el mecanismo para verificar los valores entre las dos tablas. Las relaciones de combinación de este tipo garantizan la integridad referencial de un almacén de datos. Para asegurar la validez de los resultados debe mantenerse la integridad referencial.

Cada registro de una tabla de dimensiones puede describir muchos registros en la tabla de hechos, lo que hace que la lleva de uno a muchos la cardinalidad de la combinación de las tablas de dimensiones respecto de las tablas de hechos.

En la ilustración anterior, `product_key` es la clave principal de la tabla `product_dimension` y la clave externa en la tabla `sales_fact`. Esta combinación representa la relación entre los productos de la compañía y sus ventas.

Ventajas de un esquema en estrella

Un esquema bien diseñado le permite comprender, navegar y analizar rápidamente grandes conjuntos de datos multidimensionales. Las principales ventajas de los esquemas en estrella, en un entorno de ayuda para la toma de decisiones, son las siguientes:

Rendimiento de las consultas

Las consultas se ejecutan más rápido en una base de datos con esquema en estrella que en un sistema de Procesamiento de transacciones en línea (OLTP), ya que el esquema en estrella tiene menos tablas y rutas de unión claras. En un diseño de esquema en estrella, las dimensiones se vinculan a través de la tabla de hechos central. Las dimensiones se vinculan entre sí a través de una ruta de unión que se cruza con tabla de hechos. Esta característica de diseño logra que se generen resultados coherentes y precisos para las consultas.

Rendimiento de carga y administración

La estructura del esquema en estrella reduce el tiempo que se requiere para cargar grandes cantidades de datos en una base de datos. Al definir los hechos y las dimensiones, y dividirlos en diferentes tablas, se reduce el impacto de la operación de carga. Las tablas de dimensiones puede llenarse una vez, y actualizarse de vez en cuando. Pueden agregarse nuevos hechos de manera regular y selectiva, anexando registros a una tabla de hechos.

Integridad referencial incorporada

Un esquema en estrella está diseñado para forzar la integridad referencial de los datos cargados. La integridad referencial se posibilita mediante el uso de claves principales y externas. Las claves principales de las tablas de dimensiones se vuelven claves externas en las tablas de hechos, para vincular cada uno de los registros en las tablas de hechos y en las de dimensiones.

Navegación eficaz por los datos

Se navega eficientemente por los datos, ya que las dimensiones se unen mediante tablas de hechos. Estas uniones son importantes porque representan las relaciones fundamentales de los procesos reales de la empresa. Puede explorar una tabla de una sola dimensión para seleccionar los valores de los atributos, para realizar una consulta eficaz.

2 - Conexión a orígenes de datos y a almacenes de datos

In this section

Conexiones al origen de datos	14
Definición de conexiones	14
Soporte de compresión para servidores de archivos de nube	68
Eliminación de una conexión	69

Conexiones al origen de datos

Una fuente de datos es una base de datos, un servidor de archivos, un servicio de nube u otra fuente de datos que desee procesar a través de Spectrum™ Technology Platform. Spectrum™ Technology Platform se puede conectar a más de 20 tipos de fuentes de datos.

Si desea conectar Spectrum™ Technology Platform con una fuente de datos, primero debe definir la conexión. Por ejemplo, si desea leer datos de un archivo XML en un flujo de datos, y el archivo XML se encuentra ubicado en un servidor de archivos remoto, tendría que definir una conexión con el servidor de archivos antes de poder definir el archivo XML de entrada en un flujo de datos. De manera similar, si desea escribir datos de salida del flujo de datos en una base de datos, primero debe definir la base de datos como un recurso externo.

Definición de conexiones

Para definir una nueva conexión en Spectrum™ Technology Platform, utilice uno de los siguientes módulos:

- Management Console
- Pestaña de fuentes de datos de Metadata Insights

Nota: Si desea leer o escribir datos situados en un archivo en el servidor de Spectrum™ Technology Platform mismo, no hay necesidad de definir una conexión.

Conexión con Amazon

Conexión con Amazon DynamoDB

Puede usar esta conexión para los siguientes fines:

- Para escribir y leer datos en Enterprise Designer
- Para crear Physical Models en Metadata Insights

1. Acceda a la página **Fuentes de datos** mediante uno de estos módulos:

Management Console: Acceda a Management Console a través de la dirección URL: `http://server.port/managementconsole`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.


Vaya a **Recursos > Fuentes de datos**.

Metadata Insights:

Acceda a Metadata Insights a través de la dirección URL: `http://server.port/metadata-insights`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Fuentes de datos**.

- Haga clic en el botón **Agregar** .
- En el campo **Nombre**, ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.

Nota: Una vez que guardó una conexión, no puede cambiarle el nombre.

- En el campo **Tipo**, seleccione **Amazon DynamoDB**.
- En el campo **ID de clave de acceso**, ingrese la secuencia de 20 caracteres alfanuméricos que se le proporcionó para acceder a la cuenta de AWS de Amazon.
- En el campo **Clave de acceso secreta**, ingrese la clave de 40 caracteres necesaria para autenticar la conexión.
- En el campo **Región**, seleccione la región de la cuenta AWS de Amazon.
- Para probar la conexión, haga clic en **Probar**.
- Haga clic en **Guardar**.

Limitaciones de Amazon DynamoDB

- Los tipos de datos jerárquicos, como las listas, los conjuntos y los mapas, se interpretan como tipos de datos de Cadena. Esto se debe a que no se admiten estos tipos de datos.
- Los valores nulos en un origen de datos DynamoDB se interpretan como valores de columna vacíos.
- No se admite la función agregado de `recuento` para la consulta en Model Store.

Conexión con Amazon S3

Puede usar esta conexión en los trabajos de calidad de datos Hadoop MapReduce y Spark para grandes conjuntos de datos.

- Acceda a la página **Fuentes de datos** mediante uno de estos módulos:

Management Console:

Acceda a Management Console a través de la dirección URL: `http://server.port/managementconsole`, donde *server* es el nombre del

servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Recursos > Fuentes de datos**.


Metadata Insights:

Acceda a Metadata Insights a través de la dirección URL:

`http://server.port/metadata-insights`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Fuentes de datos**.

2. Haga clic en el botón **Agregar** .
3. En el campo **Nombre**, ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.

Nota: Una vez que guardó una conexión, no puede cambiarle el nombre.

4. En el campo **Tipo**, seleccione **Nube**.
5. En el campo **Servicio de nube**, seleccione **AmazonS3**.
6. En el campo **Nombre de depósito**, ingrese el nombre de depósito según la definición en el servicio de nube Amazon S3. Este es el depósito donde Spectrum™ Technology Platform leerá y escribirá archivos.
7. Ingrese la clave de acceso y la clave secreta que le asignó Amazon.
8. En el **Tipo de almacenamiento**, seleccione el nivel de redundancia que desea permitir para el almacenamiento de datos.

Estándar El nivel predeterminado de redundancia que proporciona Amazon S3.

Redundancia reducida Almacena datos no fundamentales y se reproducen datos con facilidad en niveles bajos de redundancia. Esto proporciona un almacenamiento bastante confiable a un menor costo.

9. En la sección **Cifrado**, seleccione el método de cifrado para los datos. Puede seleccionar un enfoque de cifrado del lado del servidor o de cifrado del lado del cliente o ambos.

Clave del lado del servidor El cifrado y descifrado de los datos se produce en el lado del servidor. Sus datos se transmiten en texto sin formato al servicio de nube Amazon, donde se cifran y almacenan. Al realizar la recuperación, el servicio de nube Amazon descifra los datos y luego los transmite en texto sin formato a su sistema.

Tiene dos opciones para especificar la clave:

- **Administrado por AWS:** el servicio de nube Amazon S3 genera la clave automáticamente.
- **Proporcionada por el cliente:** ingrese la clave que va a usar el servicio de nube Amazon S3 para cifrar y descifrar los datos en el lado del servidor.

Clave del lado del cliente El cifrado y descifrado de los datos se produce en el lado del cliente. Los datos se cifran localmente en el sistema del cliente y luego se transmiten al almacenamiento de la nube Amazon S3. Al realizar la recuperación, los datos se vuelven a transmitir en un formato cifrado a su sistema y se descifran en el sistema del cliente.


Clave del lado del cliente: escriba la clave que va a usar el servicio del cliente para cifrar y descifrar los datos.

Si selecciona una **Clave del lado del servidor** y una **Clave del lado del cliente**, el proceso de cifrado y descifrado se realiza tanto en el lado del cliente como en el del servidor. En primer lugar, los datos se cifran con la clave del lado del cliente y se transmiten en un formato cifrado a Amazon, donde se vuelven a cifrar con la clave del lado del servidor y se almacenan. Al realizar la recuperación, primero Amazon descifra los datos con la clave del lado del servidor, transmite los datos en un formato cifrado a su sistema, donde finalmente se descifran con la clave del lado del cliente.

Nota: Para usar la función de cifrado de la nube Amazon S3, debe instalar los archivos JAR de seguridad de Amazon S3. Para obtener más información, consulte [Uso del cifrado Amazon S3 Cloud](#) en la página 18.

Para obtener más información sobre las funciones de cifrado de Amazon S3, consulte:

docs.aws.amazon.com/AmazonS3/latest/dev/UsingEncryption.html

10. Si desea establecer permisos de acceso, en la sección **Permisos**, haga clic en .

Los tres tipos de Beneficiarios son los siguientes:

Todos	Todos además del grupo Usuarios autenticados y Entrega de registros.
Usuarios autenticados	Para usuarios que han iniciado sesión en Amazon.
Entrega de registros	Para los usuarios que escriben registros de actividad en un Depósito especificado por el usuario, si está activada la opción Inicio de sesión en depósito.

Para cada Beneficiario, seleccione los permisos deseados:

Abrir/Descargar	Permite que el usuario descargue el archivo.
Ver	Permite que el usuario visualice los permisos actuales en el archivo.
Editar	Permite que el usuario modifique y establezca los permisos en el archivo.

11. Para probar la conexión, haga clic en **Probar**.
12. Haga clic en **Guardar**.

Uso del cifrado Amazon S3 Cloud

Para utilizar la función de seguridad de cifrado del servicio de nube Amazon S3, debe descargar los archivos JAR de seguridad y colocarlos en el servidor de Spectrum™ Technology Platform. El uso de cifrado es opcional.

1. Vaya al sitio de descarga.

Para plataformas Windows y Linux que usan Java 7, los archivos JAR se pueden descargar de:

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

Para plataformas AIX que usan Java 7, los archivos JAR se pueden descargar de:

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

2. Descargue estos dos archivos JAR:

- local_policy.jar
- US_export_policy.jar

3. Coloque los archivos JAR en la ubicación:

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

4. Reinicie el servidor.

Conexión con Amazon SimpleDB

Puede usar esta conexión para los siguientes fines:

- Para escribir y leer datos en Enterprise Designer
- Para crear Physical Models en Metadata Insights

1. Acceda a la página **Fuentes de datos** mediante uno de estos módulos:

Management Console: Acceda a Management Console a través de la dirección URL: `http://server:port/managementconsole`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.


Vaya a **Recursos > Fuentes de datos**.

Metadata Insights: Acceda a Metadata Insights a través de la dirección URL: `http://server:port/metadata-insights`, donde *server* es el nombre del servidor

o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Fuentes de datos**.

2. Haga clic en el botón **Agregar** .
3. En el campo **Nombre**, ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.

Nota: Una vez que guardó una conexión, no puede cambiarle el nombre.

4. En el campo **Tipo**, seleccione **Amazon SimpleDB**.
5. En el campo **ID de clave de acceso**, ingrese la secuencia de 20 caracteres alfanuméricos que se le proporcionó para acceder a la cuenta de AWS de Amazon.
6. En el campo **Clave de acceso secreta**, ingrese la clave de 40 caracteres necesaria para autenticar la conexión.
7. Para probar la conexión, haga clic en **Probar**.
8. Haga clic en **Guardar**.

Limitaciones de Amazon SimpleDB

Limitación de escritura

En la etapa Write to DB, el modo de escritura **Actualizar** no está disponible cuando escribe en una tabla Amazon SimpleDB. La opción **Insertar** gestiona las operaciones de inserción y actualización. Esta opción diferencia entre una inserción y una actualización usando el valor único de la columna `ItemName`, que está presente en todas las tablas Amazon SimpleDB.

Motivo: una consulta de actualización requiere una clave principal para cada registro de la tabla que se va a actualizar, lo que las bases de datos de Amazon SimpleDB no admiten.

Limitación de lectura

No se admiten las funciones agregadas `SUM` ni `AVG` mientras se ejecutan consultas en Model Store.

Conexión con Apache Cassandra

Puede usar esta conexión para los siguientes fines:

- Para escribir y leer datos en Enterprise Designer
- Para crear Physical Models en Metadata Insights

1. Acceda a la página **Fuentes de datos** mediante uno de estos módulos:

Management Console: Acceda a Management Console a través de la dirección URL: `http://server.port/managementconsole`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.


Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Recursos > Fuentes de datos**.

Metadata Insights: Acceda a Metadata Insights a través de la dirección URL: `http://server.port/metadata-insights`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Fuentes de datos**.

2. Haga clic en el botón **Agregar** .
3. En el campo **Nombre**, ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.

Nota: Una vez que guardó una conexión, no puede cambiarle el nombre.

4. En el campo **Tipo**, seleccione **Apache Cassandra**.
5. En el campo **Host**, ingrese el nombre del equipo o IP en el que está instalada la base de datos de Apache Cassandra.
6. En el campo **Keyspace**, ingrese el nombre del keyspace del centro de datos al que desea acceder..
7. En el campo **Puerto**, ingrese el nombre del puerto en el que está configurada la base de datos de Apache Cassandra.
8. Ingrese el nombre de usuario y la contraseña que usará para autenticarse en la base de datos de Cassandra.
9. En el campo **Nivel de uniformidad**, seleccione cómo deben ser las filas de datos uniformes en nodos de réplica para una transacción de datos correcta. Esto puede ser al menos un nodo, todos los nodos o una combinación de los nodos disponibles.
10. En **Tamaño de recuperación**, ingrese la cantidad de filas de conjuntos de resultados que desea recuperar en cada transacción de lectura.
11. Para probar la conexión, haga clic en **Probar**.
12. Haga clic en **Guardar**.

Limitación de Apache Cassandra

No se admite la función agregado de `recuento` para la consulta en Model Store.

Conexión con Azure Cloud

1. Acceda a la página **Fuentes de datos** mediante uno de estos módulos:

Management Console: Acceda a Management Console a través de la dirección URL: `http://server.port/managementconsole`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.


Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Recursos > Fuentes de datos**.

Metadata Insights: Acceda a Metadata Insights a través de la dirección URL: `http://server.port/metadata-insights`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Fuentes de datos**.

2. Haga clic en el botón **Agregar** .
3. En el campo **Nombre**, ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.

Nota: Una vez que guardó una conexión, no puede cambiarle el nombre.


4. En el campo **Tipo**, seleccione **Nube**.
5. En el campo **Servicio de nube**, seleccione **AzureBlobStorage**.
6. En el campo **Protocolo** seleccione si desea que en la conexión entre Azure y Spectrum™ Technology Platform se use HTTP o HTTPS.
7. En el campo **Nombre de cuenta**, ingrese el nombre de su cuenta de almacenamiento Azure.
8. En el campo **Clave de acceso**, ingrese la clave de acceso a su cuenta Azure.
9. Para probar la conexión de nube, haga clic en **Probar**.
10. Haga clic en **Guardar**.

Conexión con un archivo sin formato

Conexión con archivo sin formato delimitado

Puede usar esta conexión para los siguientes fines:

- Para escribir y leer datos en Enterprise Designer
- Para crear Physical Models en Metadata Insights


1. Vaya a **Conexiones > Archivo sin formato**.
2. De forma predeterminada, la pantalla se abre en modo crear. Si esto no sucede, haga clic en  para agregar una nueva conexión de archivo sin formato.
3. Ingrese un **Nombre de conexión** para la conexión de archivo sin formato.
4. Ingrese la **Ruta del archivo**. Para ello, haga clic en **Examinar** y seleccione el directorio del archivo.
5. Del menú desplegable, seleccione la **Codificación de caracteres** del archivo sin formato.
6. Seleccione el **Tipo de registro** como **Delimitado**.
7. En **Delimitador de campo**, seleccione el separador esperado entre dos campos cualesquiera del registro del archivo.
8. Si existe, seleccione el **Calificador de texto (opcional)** que contiene los valores de campo del registro de un archivo.
9. En **Separador de línea**, el valor `Predeterminado` está seleccionado, lo que indica que el separador de línea esperado depende de si el sistema en el que se ejecuta Spectrum™ Technology Platform es Unix o Windows.
10. Para especificar si la primera fila de un archivo es una fila de encabezado, mueva el control deslizante **La primera fila es fila de encabezado** a **Sí** o **No**.
11. Para especificar si el tipo de datos de diferentes campos en cualquier registro del archivo debe detectarse en forma automática, mueva el control deslizante **Detectar el tipo de datos desde el archivo** a **Sí** o **No**.
12. Para omitir registros malformados durante el análisis del archivo, cambie el control deslizante **Omitir registros malformados** a **Activado**.
13. Haga clic en **Probar**.
Un mensaje confirmará la prueba exitosa de la conexión.
14. Haga clic en **Guardar**.
Un mensaje confirmará la creación exitosa de la conexión.

Para ver un registro de muestra recopilado a través de la conexión de archivo sin formato delimitado creado, haga clic en **Vista previa** en la barra de encabezado. Esto recopilará los registros de archivos y los campos ordenados según los detalles que usted especificó.

Conexión con archivo sin formato con ancho fijo

Puede usar esta conexión para los siguientes fines:

- Para escribir y leer datos en Enterprise Designer
- Para crear Physical Models en Metadata Insights

1. Vaya a **Conexiones > Archivo sin formato**.
2. De forma predeterminada, la pantalla se abre en modo crear. Si esto no sucede, haga clic en  para agregar una nueva conexión de archivo sin formato.
3. Ingrese un **Nombre de conexión** para la conexión de archivo sin formato.
4. Ingrese la **Ruta del archivo**. Para ello, haga clic en **Examinar** y seleccione el directorio del archivo.
5. Del menú desplegable, seleccione la **Codificación de caracteres** del archivo sin formato.
6. Seleccione el **Tipo de registro** como **Ancho fijo**.
7. En el campo **Longitud de registro**, ingrese la cantidad total de caracteres que hay en un registro de archivo.

Repita las instrucciones del Paso 8 al Paso 13 para ingresar los detalles para todos los campos que se esperan en un registro de archivo.

8. Haga clic en **Agregar campo** para agregar una fila para un campo en un registro de archivo.
9. En la columna **Nombre**, ingrese el nombre para el valor del campo.
10. En la columna **Tipo**, seleccione el tipo de datos para el valor del campo.
11. En la columna **Posición de inicio**, ingrese la posición en el registro de archivo en la que empieza el valor del campo.

Para el primer campo en un registro de archivo, el conteo de **Posición de inicio** comienza en 1.

12. En el campo **Longitud**, ingrese la cantidad total de caracteres que cubre el campo, incluyendo el carácter en la **Posición de inicio**.

La suma de los valores de la **Posición de inicio** y de la **Longitud** para cualquier campo debe ser menor o igual a la **Longitud de registro**

Si el registro de archivo es:

```
01234Rob Smith29PitneyBowes
```

Longitud de registro = 27

Para el campo "Nombre":

Posición de inicio = 6

Longitud = 9

```
Name = Rob Smith
```

13. Marque la casilla de verificación **Recortar** si desea recortar cualquier espacio en blanco al principio y/o al final de cada valor de campo.
14. Haga clic en **Probar**.
Un mensaje confirmará la prueba exitosa de la conexión.
15. Haga clic en **Guardar**.
Un mensaje confirmará la creación exitosa de la conexión.

Para ver un registro de muestra recopilado a través de la conexión de archivo sin formato con ancho fijo creado, haga clic en **Vista previa** en la barra de encabezado. Esto recopilará los registros de archivos y los campos ordenados según los detalles que usted especificó.

Formatos de hora y fecha en una conexión de archivo

Al leer los valores de hora y fecha de archivos a través de una conexión de archivo en Spectrum™ Technology Platform, los valores deben adherirse a ciertos formatos específicos de hora y fecha.

Formatos de hora y fecha aceptados

- Fecha: "yyyy-mm-dd"
- Fecha y hora: "aaaa-mm-dd HH:mm:ss"
- Hora: "HH:mm:ss"

Archivos delimitados

Si la función **Detectar tipo** está activada al configurar la conexión de archivo delimitado, los valores de hora y fecha en los registros de archivo, que se adhieren a los formatos mencionados arriba, se detectan automáticamente como tipo de fecha.

Si un valor fecha-hora no se adhiere a uno de los formatos aceptados, el valor se leerá como un valor de tipo cadena en lugar de un valor de tipo fecha.

Archivos de ancho fijo

Para los archivos de ancho fijo, los valores de tipo fecha se configuran al crear la conexión de archivo de ancho fijo. Por lo tanto, estos valores se leen como valores de tipo fecha, independientemente de si se adhieren a los formatos aceptados o no.

Si el valor fecha-hora de un archivo de ancho fijo no se adhiere a los formatos aceptados, se deberá administrar a través de **Transformaciones** en la etapa de creación de Logical Models mediante la aplicación de esta función de categoría de *Conversión* al valor:

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

En este caso, *fecha* es el valor recibido del archivo, mientras que *formato* es el formato hora-fecha en la que el valor se recibe del archivo. Esto ayuda a analizar el valor fecha-hora correctamente.

Por ejemplo, si la fecha es 23-Feb-2008, el formato será dd-*MMM*-aaaa.

Formatos de valores resultantes

Al previsualizar los datos en una tienda de modelos:

- Si el valor se leyó como un valor de fecha-hora, se reflejará en uno de los formatos de fecha-hora aceptados en la vista previa.
- Si el valor se leyó como un valor de cadena, se reflejará como tal en la vista previa.

Conexión con un servidor FTP

Para que Spectrum™ Technology Platform acceda a los archivos de un servidor FTP, debe definir una conexión con el servidor FTP mediante Management Console. Una vez que lo hace, puede crear flujos de datos en Enterprise Designer que puede leer datos de los archivos en el servidor FTP, o bien escribir datos en ellos.

Antes de conectarse con un servidor FTP, verifique que la configuración de tiempo de espera en su servidor FTP sea correcta para los trabajos que usarán esta conexión. Según el diseño de los trabajos, es posible que haya periodos cuando la conexión esté inactiva, lo que puede provocar que se agote su tiempo de espera. Por ejemplo, puede tener un flujo de datos con dos etapas Read from File conectadas a una etapa Import To Hub. Mientras la etapa Import To Hub está leyendo los registros desde una etapa Read from File, la otra estará inactiva, lo que posiblemente cause que se agote el tiempo de espera de su conexión al servidor FTP. Considere configurar el valor de tiempo de espera en su servidor FTP en 0, para evitar que las conexiones agoten su tiempo de espera.

Nota: El servidor FTP debe estar ejecutándose en modo de conexión activa. El modo de conexión pasiva no es compatible.

1. Acceda a la página **Fuentes de datos** mediante uno de estos módulos:

Management Console: Acceda a Management Console a través de la dirección URL: `http://server.port/managementconsole`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.


Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Recursos > Fuentes de datos**.

Metadata Insights: Acceda a Metadata Insights a través de la dirección URL: `http://server.port/metadata-insights`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Fuentes de datos**.

2. Haga clic en el botón **Agregar** .
3. En el campo **Nombre**, ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.

Nota: Una vez que guardó una conexión, no puede cambiarle el nombre.

4. En el campo **Tipo**, seleccione **FTP**.
5. En los campos **Nombre de usuario** y **Contraseña**, ingrese las credenciales para autenticarse en el servidor FTP. Esto se requiere solo si el servidor FTP lo exige.
6. En el campo **Host**, ingrese el nombre del host o la dirección IP del servidor FTP.
7. En el campo **Puerto**, ingrese el número de puerto de red que utiliza el servidor para FTP.
8. Haga clic en **Prueba** para verificar que el servidor Spectrum™ Technology Platform se pueda conectar con el servidor FTP.
9. Haga clic en **Guardar**

Conexión a Google Cloud Storage

1. Acceda a la página **Fuentes de datos** mediante uno de estos módulos:

Management Console: Acceda a Management Console a través de la dirección URL: `http://server.port/managementconsole`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.


Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Recursos > Fuentes de datos**.

Metadata Insights: Acceda a Metadata Insights a través de la dirección URL: `http://server.port/metadata-insights`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Fuentes de datos**.

2. Haga clic en el botón **Agregar** .
3. En el campo **Nombre**, ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.

Nota: Una vez que guardó una conexión, no puede cambiarle el nombre.

4. En el campo **Tipo**, seleccione **Nube**.
5. En el campo **Servicio de nube**, seleccione **GoogleCloudStorage**.
6. En el campo **Nombre de depósito**, ingrese el nombre de depósito según la definición en el servicio de nube Google. Este es el depósito donde Spectrum™ Technology Platform leerá y escribirá archivos.
7. Ingrese el nombre de su aplicación, la cuenta de servicio y el archivo de clave privada proporcionado por Google.

Nota: Asegúrese de que el archivo de clave privada esté en el servidor de Spectrum™ Technology Platform.

8. Puede establecer permisos de acceso en la sección **Permisos**.

Administrar sus datos y permisos	Permite al usuario administrar los datos y los permisos.
Ver sus datos	Permite al usuario visualizar datos.
Administrar sus datos	Permite al usuario visualizar datos.

9. Para probar la conexión, haga clic en **Probar**.
10. Haga clic en **Guardar**.

Para obtener más información, consulte la documentación [Autenticación de cuenta de servicio](#) de Google.

Conexión con Hadoop

Puede crear flujos en Enterprise Designer mediante las siguientes etapas del módulo Enterprise Big Data, a fin de leer datos de la base de datos y escribirlos en ella.

- **Read from Hadoop Sequence File**
- **Write to Hadoop Sequence File**

Atención: Spectrum™ Technology Platform no admite *Hadoop 2.x* para Kerberos en plataformas Windows.

1. Acceda a la página **Fuentes de datos** mediante uno de estos módulos:

Management Console:	Acceda a Management Console a través de la dirección URL: <code>http://server.port/managementconsole</code> , donde <i>server</i> es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y <i>port</i> es el puerto HTTP que usa Spectrum™ Technology Platform.
----------------------------	---


Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Recursos > Fuentes de datos**.

Metadata Insights: Acceda a Metadata Insights a través de la dirección URL: `http://server.port/metadata-insights`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Fuentes de datos**.

2. Haga clic en el botón **Agregar** .
3. En el campo **Nombre**, ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.

Nota: Una vez que guardó una conexión, no puede cambiarle el nombre.

4. En el campo **Tipo**, seleccione **HDFS**.
5. En el campo **Host**, ingrese el nombre de host o la dirección IP de NameNode en el clúster HDFS.
6. En el campo **Puerto** ingrese el número de puerto de red.
7. En **Usuario**, seleccione una de las siguientes opciones:

Usuario del servidor	Escoja esta opción si está activada la autenticación en su clúster HDFS. Esta opción utilizará las credenciales de usuario que ejecuta el servidor de Spectrum™ Technology Platform para autenticar en HDFS.
Nombre de usuario	Escoja esta opción si está desactivada la autenticación en su clúster HDFS.
8. Marque la casilla **Kerberos** si desea activar la función de autenticación de Kerberos para esta conexión de servidor de archivos HDFS.
9. Si ha optado por activar la autenticación de **Kerberos**, entonces ingrese la ruta en el archivo keytab en el campo **Ruta del archivo keytab**.

Nota: Asegúrese de que el archivo keytab esté en el servidor de Spectrum™ Technology Platform.

10. En el campo **Protocolo**, seleccione una de las siguientes opciones:

WEBHDFS	Seleccione esta opción si el clúster HDFS está ejecutando HDFS 1.0 o una versión posterior. Este protocolo admite operaciones de lectura y escritura.
HFTP	Seleccione esta opción si el clúster HDFS está ejecutando una versión anterior a HDFS 1.0 o si su organización no permite el protocolo WEBHDFS. Este protocolo solo admite la operación de lectura.

HAR Seleccione esta opción para acceder a archivos de Hadoop. Si escoge esta opción, especifique la ruta de acceso hacia el archivo en el campo **Ruta de acceso**. Este protocolo solo admite la operación de lectura.

11. Expanda las **opciones avanzadas**.

12. Si seleccionó el protocolo WEBHDFS, puede especificar las siguientes opciones avanzadas según se requiera:

Factor de replicación Especifica cuántos datos replicará cada nodo en cada bloque. Por ejemplo, la configuración predeterminada de 3 replica cada bloque en tres nodos diferentes en el clúster. El factor de replicación máximo es 1024.

Tamaño del bloque Especifica el tamaño de cada bloque. HDFS divide un archivo en bloques del tamaño que especifique aquí. Por ejemplo, si especifica el valor predeterminado de 64 MB, cada archivo se divide en bloques de 64 MB. Luego, cada bloque se replica en la cantidad de nodos en el clúster que se especifica en el campo **Factor de replicación**.

Permisos de archivo Especifica el nivel de acceso a los archivos que Spectrum™ Technology Platform escribe en el clúster HDFS. Puede especificar los permisos de lectura y escritura para cada una de las siguientes opciones:



Nota: El permiso de *ejecución* no es válido para Spectrum™ Technology Platform.

Usuario Este es el usuario que se especificó arriba, ya sea el **Usuario del servidor** o el usuario especificado en el campo **Nombre del usuario**.

Grupo Esto se refiere a cualquier grupo al que pertenece el usuario. Por ejemplo, si el usuario es john123, entonces los permisos de grupo se aplican a cualquier grupo al que pertenece john123.

Otros Esto se refiere a cualquier otro usuario o también grupos de usuarios a los cuales no pertenece el usuario especificado.

13. Use las descripciones de **Permisos de archivo** a continuación para definir las propiedades de servidor para que Hadoop garantice que las funciones de clasificación y filtrado funcionen como se desea cuando se usa la conexión en una actividad o etapa. Para agregar propiedades, complete uno de los siguientes pasos:

- Haga clic en  y agregue las propiedades y sus valores respectivos en los campos **Propiedad** y **Valor**.
- Haga clic en  y cargue su archivo XML de configuración. El archivo XML debe ser similar a hdfs-site.xml, yarn-site.xml, o core-site.xml.

Nota: Coloque el archivo de configuración en el servidor.

Permisos y parámetros de archivo: Hadoop 1.x

Esta sección es válida para esta etapa y actividad:

- Etapa: **Read from Sequence File**
- Actividad: **Run Hadoop Pig**

fs.default.name

Especifica el nodo y el puerto donde se ejecuta Hadoop. Por ejemplo,
`hdfs://152.144.226.224:9000`

mapred.job.tracker

Especifica el nombre de host o la dirección IP y el puerto en el que se ejecuta el rastreador de trabajo MapReduce. Si el nombre de host se ingresa como local, los trabajos se ejecutan como un mapa simple y reducen la tarea. Por ejemplo,
`152.144.226.224:9001`

dfs.namenode.name.dir

Especifica en qué lugar del sistema de archivos locales, un nodo de nombre DFS debe guardar la tabla con nombre. Si es una lista de directorios delimitada por comas, la tabla con nombre se replica en todos los directorios, para propósitos de redundancia. Por ejemplo, `file:/home/hduser/Data/namenode`

hadoop.tmp.dir

Especifica la ubicación base de otros directorios temporales. Por ejemplo,
`/home/hduser/Data/tmp`

Permisos y parámetros de archivo: Hadoop 2.x

Esta sección es válida para esta etapa y actividad:

- Etapa: **Read from Sequence File**
- Actividad: **Run Hadoop Pig**

fs.defaultFS

Especifica el nodo y el puerto donde se ejecuta Hadoop. Por ejemplo,
`hdfs://152.144.226.224:9000.`

NOTA: En el caso de las versiones de Spectrum 11.0 y anteriores, se debe utilizar el nombre del parámetro `fs.defaultfs`. Tenga en cuenta la diferencia entre mayúsculas y minúsculas. En el caso de las versiones 11 SP2 y posteriores, tanto `fs.defaultfs` como `fs.defaultFS` son válidos. Recomendamos usar el nombre del parámetro `fs.defaultFS` con las versiones 11.0 SP2 y posteriores.

yarn.resourcemanager.resource-tracker.address

Especifica el nombre de host o la dirección IP del administrador de recursos. Por ejemplo, `152.144.226.224:8025`

yarn.resourcemanager.scheduler.address

Especifica la dirección de la interfaz del programador. Por ejemplo,
`152.144.226.224:8030`

yarn.resourcemanager.address

Especifica la dirección de la interfaz del administrador de aplicaciones, que se incluye en el administrador de recursos. Por ejemplo, `152.144.226.224:8041`

mapreduce.jobhistory.address

Especifica el nombre de host o la dirección IP y el puerto en el que se ejecuta el servidor del historial de trabajos de MapReduce. Por ejemplo, `152.144.226.224:10020`

mapreduce.application.classpath

Especifica el elemento CLASSPATH para las aplicaciones de Map Reduce. El elemento CLASSPATH denota la ubicación donde se encuentran las clases relacionadas con las aplicaciones de Map Reduce. Las entradas deben estar separadas por comas.

Por ejemplo:

```
$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

Maneja diversos problemas de plataforma que surgen si el servidor Spectrum se ejecuta en un equipo con Windows y se le instala Cloudera. Si el servidor Spectrum y Cloudera se están ejecutando en distintos sistemas operativos, entonces ingrese el valor de este parámetro como `verdadero`. De lo contrario, márkelo como `falso`.

Nota: Cloudera no es compatible con clientes de Windows. La configuración de este parámetro es una solución temporal, y no una solución a todos los problemas de la plataforma.

Permisos y parámetros de archivo: Kerberos

Esta sección es válida para esta etapa y actividad:

- Etapa: **Read from Sequence File**
- Actividad: **Run Hadoop Pig**

Si marcó la casilla de verificación **Kerberos**, agregue las siguientes propiedades de configuración de Kerberos:

hadoop.security.authentication

El tipo de seguridad con autenticación que se usará. Ingrese el valor de `kerberos`.

yarn.resourcemanager.principal

El Kerberos principal que se utiliza para el administrador de recursos para su negociador de recursos Hadoop YARN. Por ejemplo: `yarn/_HOST@HADOOP.COM`

dfs.namenode.kerberos.principal

El Kerberos principal que se utiliza para el namenode de su Hadoop Distributed File System (HDFS). Por ejemplo, `hdfs/_HOST@HADOOP.COM`

dfs.datanode.kerberos.principal

El Kerberos principal que se utiliza para el datanode de su Hadoop Distributed File System (HDFS). Por ejemplo, `hdfs/_HOST@HADOOP.COM`

Permisos y parámetros de archivo: Hadoop 1.x

Esta sección es válida para todas las etapas:

- Etapa **Read from File**
- Etapa **Write to File**
- Nueva etapa **Read from Hive ORC File**
- Etapa **Write to Hive ORC File**

fs.default.name

Especifica el nodo y el puerto donde se ejecuta Hadoop. Por ejemplo, `hdfs://152.144.226.224:9000`

Permisos y parámetros de archivo: Hadoop 2.x

Esta sección es válida para todas las etapas:

- Etapa **Read or write from File**
- Etapa **Read or write from Hive ORC File**

fs.defaultFS

Especifica el nodo y el puerto donde se ejecuta Hadoop. Por ejemplo, `hdfs://152.144.226.224:9000`

NOTA: En el caso de las versiones de Spectrum 11.0 y anteriores, se debe utilizar el nombre del parámetro `fs.defaultfs`. Tenga en cuenta la diferencia entre mayúsculas y minúsculas. En el caso de las versiones 11 SP2 y posteriores, tanto `fs.defaultfs` como `fs.defaultFS` son válidos. Recomendamos usar el nombre del parámetro `fs.defaultFS` con las versiones 11.0 SP2 y posteriores.

14. Para probar la conexión, haga clic en **Probar**.

15. Haga clic en **Guardar**.

Después de definir una conexión a un clúster HDFS, queda disponible en las etapas de origen y recepción en Enterprise Designer, como Read from File y Write to File. Puede seleccionar el clúster HDFS cuando hace clic en **Equipo remoto** para definir un archivo en una etapa de origen o recepción.

Soporte de compresión para Hadoop

Spectrum™ Technology Platform admite los formatos de compresión `gzip (.gz)` y `bzip2 (.bz2)` en Hadoop. Cuando esté usando las etapas **Read from File** y **Write to File** con una conexión HDFS, incluya la extensión correspondiente al formato de compresión requerido (`.gz` o `.bz2`) en el campo **Nombre de archivo**. El archivo se comprime o descomprime según la extensión de compresión especificada. Spectrum™ Technology Platform maneja la compresión y la descompresión de los archivos.

Conexión con una base de datos JDBC

Defina una conexión mediante la página **Fuentes de datos**. Puede acceder a esta página a través del módulo **Management Console** o el módulo **Metadata Insights**.

1. Acceda a la página **Fuentes de datos** mediante uno de estos módulos:

Management Console: Acceda a Management Console a través de la dirección URL: `http://server.port/managementconsole`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.


Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Recursos > Fuentes de datos**.

Metadata Insights: Acceda a Metadata Insights a través de la dirección URL: `http://server.port/metadata-insights`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Fuentes de datos**.

2. Haga clic en el botón **Agregar** .
3. En el campo **Nombre**, ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.

Nota: Una vez que guardó una conexión, no puede cambiarle el nombre.


4. En el campo **Tipo**, seleccione el tipo de base de datos con la que desea conectarse.

En el módulo Data Integration de Spectrum™ Technology Platform se incluyen controladores JDBC para las bases de datos SQL Server, Oracle y PostgreSQL. Si desea conectarse con un tipo de base de datos diferente, debe agregar el controlador JDBC antes de definir una conexión.

5. En el campo **URL**, ingrese la dirección URL de la conexión JDBC: Su administrador de bases de datos puede proporcionarle esta dirección URL.

Por ejemplo, para conectarse con una base de datos MySQL denominada "SampleDatabase", hospedada en un servidor denominado "MyServer", debe ingresar:

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

6. Es posible que deba completar campos adicionales, lo que dependerá del controlador JDBC. Los campos representan las propiedades de la cadena de conexión del controlador JDBC que seleccionó en el campo **Tipo**. Para obtener información sobre los valores y propiedades específicos de conexión que requiere el tipo de conexión, consulte la documentación del proveedor del controlador JDBC que posee su administrador de bases de datos.
7. Haga clic en **Guardar**.
8. Pruebe la conexión al marcar la casilla junto a la nueva conexión y al hacer clic en el botón Probar .

Importación de un controlador JDBC

Spectrum™ Technology Platform puede acceder a los datos de cualquier base de datos con un controlador JDBC. Los controladores para SQL, Oracle y PostgreSQL se proporcionan con el módulo Data Integration de Spectrum™ Technology Platform, que también incluye controladores para otros tipos de bases de datos. Si Spectrum™ Technology Platform no posee un controlador para el tipo de base de datos que necesita, puede agregar un controlador JDBC.

Este procedimiento consiste en importar un controlador JDBC copiando los archivos del controlador en el servidor Spectrum™ Technology Platform. Después de realizar este procedimiento, el controlador estará disponible cuando defina una conexión de base de datos JDBC en Management Console.

Nota: Este procedimiento funciona para los controladores *JDBC 4.x*. Si el controlador que desea agregar usa una versión de JDBC anterior, debe agregarlo manualmente en Management Console. Para obtener más información, consulte [Adición manual de un controlador JDBC](#) en la página 35

1. Coloque todos los archivos del controlador JDBC para la base de datos en una carpeta con el nombre:

```
Name.jdbc
```

Donde *Name* es el nombre que desea usar. El nombre de la carpeta debe finalizar con `.jdbc`.

2. Inicie sesión el servidor que está ejecutando Spectrum™ Technology Platform.
3. Copie la carpeta que contiene el controlador a esta carpeta:

```
Spectrum Location\server\app\drivers
```

El controlador se importa automáticamente.

- Para verificar que el controlador se haya importado correctamente, inicie sesión en Management Console y vaya a **Sistema > Controladores**. Debería aparecer el controlador.


Si el controlador no aparece, abra el registro del sistema en Management Console y busque errores relacionados con la implementación de controladores JDBC.

Adición manual de un controlador JDBC

Spectrum™ Technology Platform puede acceder a los datos de cualquier base de datos con un controlador JDBC. Los controladores para SQL, Oracle y PostgreSQL se proporcionan con el módulo Data Integration de Spectrum™ Technology Platform, que también incluye controladores para otros tipos de bases de datos. Si Spectrum™ Technology Platform no posee un controlador para el tipo de base de datos que necesita, puede agregar un controlador JDBC.

En este procedimiento, agregará los archivos del controlador JDBC al servidor y, luego, definirá manualmente la cadena y las propiedades de conexión. Antes de comenzar, asegúrese de comprender el formato y las propiedades de la cadena de conexión que requiere el controlador. Debe definir estos parámetros con precisión para que funcione el controlador. Por lo general, puede encontrar información sobre la cadena y las propiedades de conexión del controlador en el sitio web del proveedor.

Nota: Recomendamos que use este procedimiento solo cuando agregue un controlador JDBC que use *JDBC 1.x, 2.x o 3.x*. Si el controlador usa *JDBC 4.x*, recomendamos que use el método de importación para agregarlo. Para obtener más información, consulte [Importación de un controlador JDBC](#) en la página 34.

- Abra Management Console.
- Vaya a **Sistema > Controladores**.
- Haga clic en el botón Agregar .
- En el campo **Nombre**, ingrese un nombre para el controlador. El nombre puede ser cualquiera que elija.
- En el campo **Nombre de clase de controlador JDBC**, ingrese el nombre de clase Java del controlador. Generalmente, puede encontrar el nombre de clase en la documentación de su controlador JDBC.

Por ejemplo, para utilizar el controlador JDBC de Microsoft, puede ingresar lo siguiente:

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

- En el campo **Plantilla de cadena de conexión**, escriba la dirección URL de la conexión JDBC que se usará para conectarse con la base de datos, incluidas todas las propiedades que desee establecer en la cadena de conexión. Los distintos proveedores de bases de datos utilizan diferentes cadenas de conexión, de modo que debe revisar la documentación de su base de datos para obtener más información sobre la cadena de conexión.

Si más de una conexión de base de datos va a usar el controlador, considere utilizar muestras de propiedad en la conexión, en lugar de valores de propiedad con codificado de forma rígida,

que pueden ser distintos para cada conexión. Por ejemplo, si desea disponer que algunas conexiones usen cifrado y otras no, es posible que desee definir una muestra de propiedad para la propiedad de cifrado.

Para usar una muestra de propiedad en la cadena de caracteres de conexión, use esta sintaxis:

```
${PropertyToken}
```

Todas las muestras de propiedad que incluya en la plantilla de cadena de conexión, serán campos obligatorios cuando se defina una conexión de base de datos.

Nota: Use el nombre de muestra de propiedad `${password}` para la propiedad que contendrá la contraseña de la base de datos. Si utiliza este nombre de muestra, la contraseña se ocultará en el campo de Management Console y se cifrará en la base de datos.

Por ejemplo, la siguiente cadena de conexión para SQL contiene muestras de propiedad para host, puerto, instancia y cifrado:

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

Estas muestras son campos obligatorios al definir una conexión de base de datos que usa este controlador:

Inicio > Recursos: fuentes de datos > Agregar origen de datos

Agregar origen de datos

*Nombre

ExampleConnection

Conexión

*Tipo

ExampleDriver


*Host

*Port

*Instance

*encryption

Prueba

7. Si existen propiedades que desee dejar como opcionales para las conexiones de base de datos, definalas en la sección **Propiedades de conexión**.
 - a) En la sección **Propiedades de conexión**, haga clic en el botón Agregar .
 - b) En el campo **Rótulo**, escriba una descripción sencilla de la propiedad. El rótulo que ingrese aquí se usa como el rótulo de campo en la ventana de conexiones, al momento de crear una conexión por medio de este controlador.
 - c) En el campo **Muestra de propiedad** ingrese la muestra de la propiedad opcional. Consulte la documentación del controlador de la base de datos, para conocer las propiedades que admite el controlador.

Nota: Use el nombre de muestra de propiedad `password` para la propiedad que contendrá la contraseña de la base de datos. Si utiliza este nombre de muestra, la contraseña se ocultará en el campo de Management Console y se cifrará en la base de datos.

Por ejemplo, si desea dejar el cifrado como opcional para las conexiones de bases de datos que usen este controlador, debe definir la propiedad de cifrado de esta manera:

Inicio > Sistema: controladores > Editar controlador

Editar controlador implementado

*Nombre

com.mysql.jdbc.Driver.5.1

*Nombre de clase del controlador JDBC ?

com.mysql.jdbc.Driver

*Planilla de cadena de caracteres de la conexión ?

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

Propiedades y controladores

Propiedades de conexión ?



Rótulo	Muestra de propiedad
<input type="checkbox"/> username	user
<input type="checkbox"/> password	password
<input type="checkbox"/> Use SSL	useSSL

Cuando una conexión de base de datos use este controlador, se mostrará la propiedad de cifrado como una propiedad opcional en la conexión de base de datos:

Inicio > Recursos: fuentes de datos > Agregar origen de datos

Agregar origen de datos

*Nombre

MyConnection

Conexión

*Tipo

com.mysql.jdbc.Driver.5.1


*Host

*Instance

User Name

Password

Use SSL

8. Inicie sesión en el servidor ejecutando Spectrum™ Technology Platform, y coloque el archivo del controlador de la base de datos en una carpeta del servidor. La ubicación no es importante.
9. En la sección **Archivos de controlador**, haga clic en el botón Agregar .
10. En el campo **Ruta de archivo**, ingrese la ruta al archivo del controlador de la base de datos en el servidor.
11. Haga clic en **Guardar**.

Eliminación de un controlador JDBC importado

Los controladores JDBC no se pueden eliminar con Management Console si fueron importados a Spectrum™ Technology Platform en vez de agregados manualmente en Management Console. En cambio, siga este procedimiento para eliminar el controlador.

Importante: Antes de eliminar un controlador, verifique que no haya conexiones de bases de datos usándolo.

1. Detenga el servidor de Spectrum™ Technology Platform.
2. Vaya a esta carpeta:

```
Spectrum Location\server\app\drivers
```

3. En la carpeta `drivers`, elimine la carpeta que contiene el controlador.
4. Inicie el servidor de Spectrum™ Technology Platform.
5. Para verificar que el controlador se haya eliminado, inicie sesión en Management Console, vaya a **Sistema > Controladores** y compruebe que ya no aparezca.

Tipos de datos de la base de datos admitidos

Spectrum™ Technology Platform admite los siguientes tipos de datos, comúnmente utilizados en las bases de datos:

bigdecimal	Un tipo de dato numérico que admite 38 puntos decimales de precisión. Utilice este tipo de datos para datos que se usarán en cálculos matemáticos que requieren un alto grado de precisión, especialmente aquellos que implican datos financieros. El tipo de datos bigdecimal es compatible con cálculos más precisos que el tipo de datos double.
boolean	Tipo de dato lógico con dos valores: verdadero y falso
fecha	Un tipo de dato que contiene un mes, día y año. Por ejemplo, 2012-01-30 o 30 de enero, 2012. Puede especificar un formato de fecha predeterminado en Management Console.
fecha y hora	Un tipo de dato que contiene un mes, día, año y horas, minutos y segundos. Por ejemplo, 2012/01/30 6:15 PM.
double	Un tipo de dato numérico que contiene números de doble precisión tanto negativos como positivos entre 2^{-1074} y $(2-2^{-52}) \times 2^{1023}$. En notación científica, el rango de valores es - 1.79769313486232E+308 a 1.79769313486232E+308.
float	Un tipo de dato numérico que contiene números de precisión simple tanto negativos como positivos entre 2^{-149} y $(2-2^{-23}) \times 2^{127}$. En notación científica, el rango de valores es - 3.402823E+38 a 3.402823E+38.
integer	Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{31} (-2,147,483,648) y $2^{31}-1$ (2,147,483,647).
long	Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{63} (-9,223,372,036,854,775,808) y $2^{63}-1$ (9,223,372,036,854,775,807).
string	Secuencia de caracteres.

hora Un tipo de dato que contiene la hora del día. Por ejemplo, 21:15:59 o 9:15:59 PM.

En bruto Un tipo de datos de Oracle para guardar datos binarios de longitud variable. El tamaño máximo es 2.000 bytes (la longitud máxima en Oracle 7 era 255 bytes).

Otros tipos de datos de bases de datos se asignan automáticamente a uno de los tipos de datos admitidos de la siguiente manera:

Tipos de datos de bases de datos	Tipo de datos admitidos
Tipos de fecha/hora	
TIMESTAMP	fecha y hora
Tipos de cadenas	
char	string
CLOB	string
LONGVARCHAR	string
NCHAR	string
NVARCHAR	string
VARCHAR	string
Tipos numéricos	
bigint	long
Decimal	double
FLOAT	double
Numérico	bigdecimal
real	float
SMALLINT	integer

Tipos de datos de bases de datos	Tipo de datos admitidos
tinyint	integer

Tipos de datos booleanos

BIT	boolean
-----	---------

Tipos de datos de la base de datos admitidos para el módulo Location Intelligence

Los siguientes tipos de datos de bases de datos se asignan automáticamente a uno de los tipos de datos compatibles con el módulo Location Intelligence.

Tipos de datos de bases de datos	Tipo de datos admitidos
SQL Server	
tinyint	SHORT_INTEGER
smallint	SHORT_INTEGER
int	INTEGER
bigint	LONG_INTEGER
float	DOUBLE
real	DOUBLE
decimal(10, 5)	DOUBLE
numérico(10, 5)	DOUBLE
fecha	DATE
hora	HOUR
fecha y hora	DATE_TIME
smalldatetime	DATE_TIME

Tipos de datos de bases de datos	Tipo de datos admitidos
char(10)	STRING
varchar(10)	STRING
nchar(10)	STRING
nvarchar(10)G	STRING
binary(10)	BINARY
varbinary(10)	BINARY
PostGIS	
smallint	SHORT_INTEGER
integer	INTEGER
bigint	LONG_INTEGER
numérico(10, 5)	DOUBLE
real	DOUBLE
doble precisión	DOUBLE
serial	INTEGER
bigserial	LONG_INTEGER
bytea	BINARY
fecha	DATE
hora	HOUR
timestamp	DATE_TIME

Tipos de datos de bases de datos	Tipo de datos admitidos
carácter(10)	STRING
variación de carácter(10)	STRING
nchar(10)	STRING
Oracle	
NUMBER	DOUBLE
CHAR(10)	STRING
VARCHAR(10)	STRING
VARCHAR2(10)	STRING
NCHAR(10)	STRING
NVARCHAR2 (10)	STRING
DATE	DATE_TIME
TIMESTAMP	DATE_TIME
BLOB	BINARY
SAP HANA	
tinyint	SHORT_INTEGER
smallint	SHORT_INTEGER
integer	INTEGER
bigint	LONG_INTEGER

Tipos de datos de bases de datos	Tipo de datos admitidos
smalldecimal	DOUBLE
decimal(10, 5)	DOUBLE
real	DOUBLE
double	DOUBLE
float(30)	DOUBLE
varchar(30)	STRING
nchar(10)	STRING
nvarchar(30)	STRING
alphanum(30)	STRING
fecha	DATE
hora	HOUR
seconddate	DATE_TIM
timestamp	DATE_TIM
varbinary(30)	BINARY

Limitaciones

- No se admiten los conectores MongoDB/Cassandra a través de PrestoDB en Metadata-Insights. Existen conectores independientes para MongoDB y Cassandra.
- %Presto DB no recomienda; la escritura en cualquier base de datos a través de Presto; por lo tanto, no es compatible con el conector Presto JDBC.

Conexión a Knox

Una puerta de enlace Apache Knox le permite acceder a un servicio Hadoop a través de una capa de seguridad Knox.

Con esta conexión, puede crear flujos en Enterprise Designer mediante etapas del módulo Enterprise Big Data, a fin de leer datos de Hadoop a través de Knox, o escribir datos en este.

1. Acceda a la página **Fuentes de datos** mediante uno de estos módulos:

Management Console: Acceda a Management Console a través de la dirección URL: `http://server:port/managementconsole`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.


Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Recursos > Fuentes de datos**.

Metadata Insights: Acceda a Metadata Insights a través de la dirección URL: `http://server:port/metadata-insights`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Fuentes de datos**.

2. Haga clic en el botón **Agregar** .
3. En el campo **Nombre**, ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.

Nota: Una vez que guardó una conexión, no puede cambiarle el nombre.

4. En el campo **Tipo**, seleccione *Puerta de enlace*.
5. En el campo **Tipo de puerta de enlace**, seleccione **Knox**.
6. En el campo **Host**, ingrese el nombre de host o la dirección IP del nodo del clúster HDFS que ejecuta la puerta de enlace.
7. En el campo **Puerto**, ingrese el número de puerto para la puerta de enlace Knox.
8. En el campo **Nombre de usuario**, ingrese el nombre de usuario para la puerta de enlace Knox.
9. En el campo **Contraseña**, ingrese la contraseña para autorizar su acceso a la puerta de enlace Knox.
10. En el campo **Nombre de puerta de enlace**, ingrese el nombre de la puerta de enlace Knox a la que desea acceder.

11. En el campo **Nombre de clúster**, ingrese el nombre del clúster Hadoop al que se accederá.
12. En el campo **Protocolo**, seleccione `WEBHDFS`.
13. En el campo **Nombre de servicio**, ingrese el nombre del servicio Hadoop al que se accederá.
14. Para probar la conexión, haga clic en **Probar**.
15. Haga clic en **Guardar**.

Después de definir una conexión de Knox en un clúster HDFS, puede usar la conexión en Enterprise Designer, en las etapas **Read from File** y **Write to File**. Puede seleccionar el clúster HDFS cuando hace clic en **Equipo remoto** para definir un archivo en una etapa de origen o recepción.

Conexión con Marketo

1. Acceda a la página **Fuentes de datos** mediante uno de estos módulos:

Management Console: Acceda a Management Console a través de la dirección URL: `http://server.port/managementconsole`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.


Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Recursos > Fuentes de datos**.

Metadata Insights: Acceda a Metadata Insights a través de la dirección URL: `http://server.port/metadata-insights`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Fuentes de datos**.

2. Haga clic en el botón **Agregar** .
3. En el campo **Nombre**, ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.

Nota: Una vez que guardó una conexión, no puede cambiarle el nombre.

4. En el campo **Tipo**, seleccione **Marketo**.
5. En el campo **URL del endpoint**, ingrese la URL del endpoint de su cuenta Marketo.

Para encontrar la URL del endpoint, inicie sesión en su cuenta Marketo y vaya a **Administrador > Integración > Servicios web**. La URL del endpoint se encuentra en el encabezado **REST API** y se presenta en el siguiente formato:

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

Copie la parte de la URL antes de `/rest`. Por ejemplo, `Spatial:name=Administration,type=WMS Service.<parmname>Spatial:name=Administration,type=WMS Service</parmname>https://AccountID.mktorest.com`

6. Ingrese el Id. del cliente y la clave secreta de su cuenta Marketo.

Para encontrar su ID de cliente y su clave secreta, inicie sesión en su cuenta Marketo y vaya a **Administrador > Integración > LaunchPoint > API Rest > Ver detalles**. En la ventana emergente aparecen los detalles.

7. Para probar la conexión, haga clic en **Probar**.
8. Haga clic en **Guardar**.

Limitaciones de Marketo

1. Esta consulta solo se aplica a las entidades `List` y `Activity_type`. Para otras, proporcione el tipo de filtro.

```
Select * from Marketo_Table
```

2. No se admiten combinaciones, excepto entre entidades `Lead` y `Lead_List`. La consulta de combinación entre `Lead` y `Lead_List` para `List_Id` es la siguiente:

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

Entidades y operaciones admitidas

Las entidades pueden ser de los siguientes tipos:

1. Entidad
2. Actualizar entidad: esta es una tabla virtual usada para actualizar la entidad principal. Por ejemplo, **Merge_Leads** debe utilizarse para fusionar diferentes Principales de Marketo.

Conexión con Microsoft Dynamics 365

Conexión con Microsoft Dynamics 365 Online

Spectrum™ Technology Platform admite la conexión a Microsoft Dynamics 365 (versión 9) solamente.

1. Acceda a la página **Fuentes de datos** mediante uno de estos módulos:

Management Console: Acceda a Management Console a través de la dirección URL: `http://server.port/managementconsole`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.


Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Recursos > Fuentes de datos**.

Metadata Insights: Acceda a Metadata Insights a través de la dirección URL: `http://server.port/metadata-insights`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Fuentes de datos**.

2. Haga clic en el botón **Agregar** .
3. En el campo **Nombre**, ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.

Nota: Una vez que guardó una conexión, no puede cambiarle el nombre.

4. En el campo **Tipo**, seleccione **Microsoft Dynamics 365**.
5. En el campo **Tipo de implementación**, seleccione **En línea**.
6. En el campo **Nombre de usuario**, ingrese su nombre de usuario de Microsoft Dynamics.
7. En el campo **Contraseña**, ingrese su contraseña de Microsoft Dynamics.
8. En el campo **Nombre de la organización**, ingrese su nombre único de la organización asignado, el cual identifica su instancia CRM.

Para encontrar su nombre único de la organización, inicie sesión en Microsoft Dynamics y vaya a **Configuración > Personalización > Personalizaciones > Recursos de desarrollador**. Aparece el nombre único de la organización.

9. En el campo **Región**, seleccione la región geográfica de su cuenta de Microsoft Dynamics.
10. Para probar la conexión, haga clic en **Probar**.
11. Haga clic en **Guardar**

Conexión con Microsoft Dynamics 365 On Premises

Este conector de Spectrum™ Technology Platform es compatible con la autenticación basada en reclamaciones para Microsoft Dynamics 365 On Premises.

Requisitos previos:

Importación del certificado en el archivo de almacén de claves: para importar los certificados del servidor Dynamics CRM al almacén de claves de distribución Java Spectrum, siga los pasos que se describen a continuación:

1. Copie los certificados del servidor a una carpeta local
2. Siga esta ruta para acceder a la distribución de Spectrum JAVA:
`<SPECTRUM_HOME>\java\jre\lib\security`
3. Ejecute el siguiente comando para importar los certificados: `keytool -importcert -alias <certificate`

```
alias name> -file " <certificate path>\<certificate name>" -keystore
keystore.jks en Windows y keytool -import -alias <certificate alias name>
-file "<certificate path>/<certificate name>" -keystore keystore.jks
en Unix.
```

Definición de la conexión con Microsoft Dynamics 365 On Premises

Siga los pasos que se describen a continuación para definir una conexión **Microsoft Dynamics 365 On Premises**:

1. Acceda a la página **Fuentes de datos** mediante uno de estos módulos:

Management Console: Acceda a Management Console a través de la dirección URL:
`http://server:port/managementconsole`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.


Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Recursos > Fuentes de datos**.

Metadata Insights: Acceda a Metadata Insights a través de la dirección URL:
`http://server:port/metadata-insights`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Fuentes de datos**.

2. Haga clic en el botón **Agregar** .
3. En el campo **Nombre**, ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.

Nota: Una vez que guardó una conexión, no puede cambiarle el nombre.

4. Haga clic en **Microsoft Dynamics 365** en el **Tipo**.

5. Haga clic en la opción **En las instalaciones** en el **Tipo de implementación**.
6. Ingrese su **Nombre de usuario** de Microsoft Dynamics en el campo respectivo.
7. Ingrese su **Contraseña** de Microsoft Dynamics en el campo respectivo.
8. Ingrese el **Nombre del host** en el campo respectivo.
9. Ingrese el **Nombre del puerto** en el campo respectivo.
10. Ingrese la **URL de STS** en el campo respectivo.
11. Haga clic en **Probar** para comprobar la conexión.
12. Haga clic en **Guardar**.

Limitaciones

Crear/Actualizar: puede haber errores en la operación Crear/Actualizar si se asigna una columna de una Entidad a varias entidades de referencia. Por ejemplo, 'ParentCustomerId' en Cliente puede estar asociado a Cuenta, Prospecto y mucho más. Para solucionarlo, los datos de esta columna deben tener el siguiente formato: "ReferenceEntityName:GUID" en lugar de "GUID".

Entidades y operaciones admitidas

Las entidades pueden ser de los siguientes tipos:

1. Propiedad del usuario
2. Propiedad de la organización
3. Propiedad de la empresa
4. Ninguno

Conexión a una tienda de modelos

Conecte con una tienda de modelos para usar los datos federados de diversos orígenes, como bases de datos, servidores de archivos y servicios de nube. Después de definir una conexión, puede usar los datos en los Physical Models y Logical Models de una Model Store (creada e implementada en Metadata Insights) en las etapas **Read from DB** y **Write to DB** de Enterprise Designer.

1. Acceda a la página **Fuentes de datos** mediante uno de estos módulos:

Management Console: Acceda a Management Console a través de la dirección URL: `http://server.port/managementconsole`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.


Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Recursos > Fuentes de datos**.

Metadata Insights: Acceda a Metadata Insights a través de la dirección URL: `http://server.port/metadata-insights`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Fuentes de datos**.

- Haga clic en el botón **Agregar** .
- En el campo **Nombre**, ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.

Nota: Una vez que guardó una conexión, no puede cambiarle el nombre.

- En el campo **Tipo**, elija **Tienda de modelos**.
- En el campo **Tienda de modelos**, ingrese el nombre de la tienda de modelos con la que está estableciendo una conexión.

Para buscar nombres de las tiendas de modelos disponibles, abra Metadata Insights, diríjase a Modelado y haga clic en la pestaña **Tienda de modelos**.

- Para probar la conexión, haga clic en **Probar**.
- Haga clic en **Guardar**.

Conexión con NetSuite

Durante la lectura y escritura en una conexión NetSuite, se admiten los modos interactivo y de lote. Spectrum™ Technology Platform admite estos tipos de entidad de NetSuite:

- Registros estándar
- Registros personalizados
- Búsquedas guardadas
- Combinaciones entre registros estándar

Para conectarse a NetSuite:

- Acceda a la página **Fuentes de datos** mediante uno de estos módulos:

Management Console: Acceda a Management Console a través de la dirección URL: `http://server.port/managementconsole`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.


Vaya a **Recursos > Fuentes de datos**.

Metadata Insights:

Acceda a Metadata Insights a través de la dirección URL: `http://server:port/metadata-insights`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Fuentes de datos**.

- Haga clic en el botón **Agregar** .
- En el campo **Nombre**, ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.

Nota: Una vez que guardó una conexión, no puede cambiarle el nombre.

- En el campo **Tipo**, seleccione **NetSuite**.
- En el campo **Correo electrónico**, ingrese el correo electrónico vinculado a la cuenta de NetSuite que se utilizará para la conexión.
- En el campo **Contraseña**, ingrese la contraseña de la cuenta de NetSuite.
- En el campo **Cuenta**, ingrese el nombre de usuario para la cuenta de NetSuite.
- En el campo **Función**, seleccione la función que corresponda a esta conexión entre las funciones asignadas a la cuenta de usuario específica de NetSuite.

El campo **Función** es opcional. Si deja el campo **Función** en blanco, se utilizará la función predeterminada para iniciar sesión mediante la conexión.

Atención: Solo se admiten funciones estándar. No se admiten funciones personalizadas.

- Para probar la conexión, haga clic en **Probar**.
- Haga clic en **Guardar**.

Nota: Para realizar una operación `INSERT` en un registro mediante una conexión NetSuite, utilice una consulta `UPSERT` con la clave principal (`internalId`) en blanco.

Limitaciones de NetSuite

- Cuando consulte usando combinaciones, debe mencionar columnas específicas. Por ejemplo, no se admite esta consulta:

```
select * from CUSTOMER_M
```

- Las conexiones simultáneas a NetSuite no se admiten debido a que NetSuite admite un único inicio de sesión para una cuenta.
- Solo puede escribir registros estándar y personalizados.

4. Tanto para las consultas `UPDATE` como para las `UPSERT`, se realiza una operación `UPSERT`.
5. En la etapa Write to DB, el tamaño máximo de lote permitido para una operación `insert` es 200 y el tamaño máximo de lote para una operación `update` es 100.

Entidades y operaciones admitidas

Las entidades pueden ser de los siguientes tipos:

1. Registros estándar
2. Registros personalizados
3. Combinaciones
4. Búsquedas guardadas

Nota: En una tabla de conexión NetSuite, la columna de clave principal es `internalId`.

Conexión con NoSQL

Los tipos de base de datos NoSQL admitidos son:

- Couchbase
- MongoDB

Puede crear flujos en Enterprise Designer mediante las siguientes etapas del módulo Enterprise Big Data, a fin de leer datos de la base de datos y escribirlos en ella.

- Query NoSQL DB
- Read from NoSQL DB
- Write to NoSQL DB

1. Acceda a la página **Fuentes de datos** mediante uno de estos módulos:

Management Console: Acceda a Management Console a través de la dirección URL: `http://server:port/managementconsole`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.


Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Recursos > Fuentes de datos**.

Metadata Insights: Acceda a Metadata Insights a través de la dirección URL: `http://server:port/metadata-insights`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Fuentes de datos**.

2. Haga clic en el botón **Agregar** .
3. En el campo **Nombre**, ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.

Nota: Una vez que guardó una conexión, no puede cambiarle el nombre.

4. En el campo **Tipo**, seleccione cualquiera de las siguientes opciones:
 - Couchbase
 - MongoDB
5. Especifique el **Host**, el **Puerto**, la **Base de datos**, el **Nombre de usuario** y la **Contraseña** de la base de datos NoSQL específica a la cual desea acceder.
6. Haga clic en **Probar** para comprobar que la conexión a la base de datos sea correcta.
7. Haga clic en **Aceptar**.

Conexión con Salesforce

Puede usar esta conexión para los siguientes fines:

- Para escribir y leer datos en Enterprise Designer
- Para crear Physical Models en Metadata Insights

1. Acceda a la página **Fuentes de datos** mediante uno de estos módulos:

Management Console: Acceda a Management Console a través de la dirección URL: `http://server.port/managementconsole`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Recursos > Fuentes de datos**.

Metadata Insights: Acceda a Metadata Insights a través de la dirección URL: `http://server.port/metadata-insights`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Fuentes de datos**.

2. Haga clic en el botón **Agregar** .

3. En el campo **Nombre**, ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.

Nota: Una vez que guardó una conexión, no puede cambiarle el nombre.

4. En el campo **Tipo**, seleccione **Salesforce**.
5. En el campo **Nombre de usuario**, ingrese el ID del correo electrónico registrado en el almacenamiento de datos de Salesforce.
6. En el campo **Contraseña**, ingrese una combinación de la contraseña del portal de Salesforce y de la muestra de seguridad generada a través del portal de Salesforce.

Por ejemplo, si su contraseña es `ventas@prueba` y la muestra de seguridad que le proporcionó Salesforce es `56709367`, la contraseña para autenticar esta conexión de Salesforce sería `ventas@prueba56709367`.

7. Cambie el botón de alternancia de **Usar el punto final predeterminado** a **No** si desea utilizar una URL de punto final específico para acceder a los datos de Salesforce. Ingrese la URL requerida en el campo **URL de Salesforce** que aparece debajo del botón de alternancia **Usar el punto final predeterminado**.

Nota: La **URL de Salesforce** es un campo obligatorio.

8. Cambie el botón de alternancia **Utilice lectura masiva** a **Sí** si desea recuperar datos masivos desde Salesforce. El valor predeterminado es **No**.

Nota: Las consultas de obtención masiva de datos no se admiten para campos combinados de dirección y ubicación geográfica. Para obtener más información acerca de las consideraciones y limitaciones de la consulta masiva, lea los documentos [Consideraciones y limitaciones de los campos combinados](#) y [Límites de operaciones masivas en la API](#).

9. Para probar la conexión, haga clic en **Probar**.
10. Haga clic en **Guardar**.

Nota: Los campos de auditoría están habilitados en todas las tablas en forma predeterminada. Los campos de auditoría de Salesforce son:

- fecha de creación
- fecha de última modificación
- creado por
- modificado por última vez por

Atención: El modelo físico creado en Spectrum™ Technology Platform, versiones 10 y anteriores, mediante las conexiones de Salesforce se debe abrir y guardar de nuevo para habilitar los campos de auditoría en sus tablas.

Limitación de Salesforce

No se admiten las funciones agregadas mientras se ejecutan consultas en Model Store.

Conexión con SAP NetWeaver

Crear una conexión SAP NetWeaver en Management Console mediante OData Services le permite leer, escribir y sincronizar sus datos de CRM y ERP. Durante la lectura y escritura en una conexión SAP, se admiten los modos interactivo y de lote.

Para definir una conexión SAP NetWeaver, haga lo siguiente:

1. Acceda a la página **Fuentes de datos** mediante uno de estos módulos:

Management Console: Acceda a Management Console a través de la dirección URL: `http://server.port/managementconsole`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.


Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Recursos > Fuentes de datos**.

Metadata Insights: Acceda a Metadata Insights a través de la dirección URL: `http://server.port/metadata-insights`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Fuentes de datos**.

2. Haga clic en el botón **Agregar** .
3. En el campo **Nombre**, ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.

Nota: Una vez que guardó una conexión, no puede cambiarle el nombre.

4. En el campo **Tipo** elija **SAP**.
5. En el campo **Nombre de usuario**, ingrese el nombre de usuario para acceder al servicio web de SAP.
6. En el campo **Contraseña**, ingrese la contraseña del servicio web de SAP.
7. En el campo **OdataURL**, ingrese la dirección del servicio web de Odata que se utilizará para la conexión.
8. Haga clic en **Probar**.

Un mensaje confirmará la prueba exitosa de la conexión.

9. Haga clic en **Guardar**.

Un mensaje confirmará la creación exitosa de la conexión.

Nota: Para realizar operaciones de recopilación, un servicio OData debe ser compatible con las operaciones `$skip` y `$top`. Si el servicio no es compatible con estas operaciones, los registros recopilados presentarán incoherencias en la vista previa de la tienda de modelos.

Limitaciones de SAP NetWeaver

Para las consultas `UPDATE` y `UPSERT`, se ejecuta una operación `UPDATE`.

Entidades y operaciones admitidas

Las columnas de entidades son de dos tipos:

- Nativa: las columnas con tipos de datos nativos se muestran con sus tipos de datos correspondientes
- Definidas por el cliente: las columnas con tipos de datos definidos por el cliente se muestran con un tipo de datos en blanco

Para implementar una Model Store derivada de una conexión SAP, asegúrese de que los Physical Models y Logical Models incluyan solo entidades cuyas columnas sean de tipos de datos nativos. Si los modelos tienen entidades de tipos de datos definidos por el cliente, no es posible implementar la Model Store.

Conexión con SharePoint

1. Acceda a la página **Fuentes de datos** mediante uno de estos módulos:

Management Console: Acceda a Management Console a través de la dirección URL: `http://server.port/managementconsole`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.


Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Recursos > Fuentes de datos**.

Metadata Insights: Acceda a Metadata Insights a través de la dirección URL: `http://server.port/metadata-insights`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Fuentes de datos**.

- Haga clic en el botón **Agregar** .
- En el campo **Nombre**, ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.

Nota: Una vez que guardó una conexión, no puede cambiarle el nombre.

- En el campo **Tipo**, seleccione **Nube**.
- En el campo **Servicio de nube**, seleccione **Sharepoint**.
- En el campo **Versión** seleccione **v2010**. Spectrum™ Technology Platform actualmente admite la versión 2010 de SharePoint.
- En el campo **Protocolo**, seleccione el protocolo necesario para conectarse a SharePoint.
- En el campo **Dirección del servidor**, ingrese el nombre del host o la dirección IP del servidor SharePoint al cual desea conectarse.
- Ingrese el nombre de usuario y contraseña que se van a usar para autenticarse en SharePoint.
- En el campo **Proyecto**, ingrese el proyecto específico a cuya ubicación SharePoint desea acceder.
- Para probar la conexión, haga clic en **Probar**.
- Haga clic en **Guardar**.

Ejemplo

Por ejemplo, imagine que desea crear una conexión a la siguiente URL de SharePoint:

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

Debería completar los campos **Protocolo**, **Dirección del servidor** y **Proyecto** de la siguiente manera:

- **Protocolo:** https
- **Dirección del servidor:** sharepoint.example.com
- **Proyecto:** myportal

Conexión con Splunk

- Acceda a la página **Fuentes de datos** mediante uno de estos módulos:

Management Console: Acceda a Management Console a través de la dirección URL: `http://server.port/managementconsole`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.


Vaya a **Recursos > Fuentes de datos**.

Metadata Insights:

Acceda a Metadata Insights a través de la dirección URL: `http://server.port/metadata-insights`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Fuentes de datos**.

- Haga clic en el botón **Agregar** .
- En el campo **Nombre**, ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.

Nota: Una vez que guardó una conexión, no puede cambiarle el nombre.

- En el campo **Tipo**, seleccione **Splunk**.
- En el campo **Nombre de usuario**, ingrese el nombre de usuario de la cuenta de Splunk para autenticar la instancia de Splunk.
- En el campo **Contraseña**, ingrese la contraseña de la cuenta de Splunk.
- En el campo **Host**, ingrese la dirección o el nombre del host del servidor donde está alojado el origen de datos de Splunk.
- En el campo **Puerto**, ingrese el número de puerto del origen de datos de Splunk.
- Para probar la conexión, haga clic en **Probar**.
- Haga clic en **Guardar**.

Limitaciones de Splunk

No se admite esta consulta:

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

Entidades y operaciones admitidas

Operaciones admitidas

LIKE, ORDER BY, LIMIT, IN, BETWEEN, !=, <=, >=, <, >, varios operadores AND/OR.

Funciones admitidas

- **Funciones de cadenas de caracteres:** upper, lower, length, len, ltrim, rtrim, substring, max, min

- Funciones matemáticas: `abs`, `ceil`, `exp`, `floor`, `sqrt`, `round`

Nota: Para todas las demás consultas, utilice la columna `search` de Splunk tal como se explica a continuación.

Spectrum™ Technology Platform proporciona una columna `search` en la tabla de Splunk con la que puede buscar los datos necesarios en la conexión Splunk.

Mientras ejecuta una consulta `select` en `SplunkTable`, utilice la columna `search` en la cláusula `where` en cualquiera de las siguientes situaciones:

1. Para incluir los criterios de búsqueda que no se pueden especificar a través de sintaxis ANSI SQL.
2. Para incluir los criterios de búsqueda específicos de Splunk que no se pueden incorporar como parte de la consulta SQL principal.

Por ejemplo, esta consulta busca un valor `_raw` que contenga la clave `opp` con el valor `ACC`.

```
select "_raw" from SplunkTable where "search"='search opp=ACC'
```

Conexión con SuccessFactors

1. Acceda a la página **Fuentes de datos** mediante uno de estos módulos:

Management Console: Acceda a Management Console a través de la dirección URL: `http://server:port/managementconsole`, donde `server` es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y `port` es el puerto HTTP que usa Spectrum™ Technology Platform.


Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Recursos > Fuentes de datos**.

Metadata Insights: Acceda a Metadata Insights a través de la dirección URL: `http://server:port/metadata-insights`, donde `server` es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y `port` es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Fuentes de datos**.

2. Haga clic en el botón **Agregar** .
3. En el campo **Nombre**, ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.

Nota: Una vez que guardó una conexión, no puede cambiarle el nombre.

4. En el campo **Tipo**, seleccione **SuccessFactors**.
5. En el campo **ID de la empresa**, ingrese su ID de la empresa, el cual permite identificar la instancia exclusiva de la empresa en un centro de datos SuccessFactors específico.
6. En el campo **URL del servicio**, ingrese la URL del servidor de SuccessFactors al que desea acceder. Esta URL es específica para el centro de datos global al cual se asigna la ID de la empresa.
7. Ingrese su nombre de usuario y contraseña para su instancia de cliente de SuccessFactors.
8. Para probar la conexión, haga clic en **Probar**.
9. Haga clic en **Guardar**

Limitaciones de SuccessFactors

1. Las operaciones por lote solo pueden ejecutarse si se utilizan `upsert` consultas. Por lo tanto, las consultas `insert` y `update` también se realizan como consultas `upsert` en las operaciones por lote.
2. Es posible que las propiedades de las tablas/columnas, según se muestran en el esquema del modelo físico de una conexión de SuccessFactors, no funcionen según lo esperado durante el uso de la operación correspondiente. Por ejemplo, es posible que una columna, que se había marcado como columna que admite actualizaciones, arroje una excepción del sistema cuando intente actualizar la columna.

Entidades y operaciones admitidas

Las entidades son de dos tipos:

- **Entidad:** esta tabla representa una entidad de negocios
- **Unión:** un mapping entre dos tablas de tipo *Entidad*: una tabla principal y cualquiera de sus tablas secundarias.

Nota: Los enlaces no están presentes entre tablas en el esquema de modelos físicos derivado de una conexión SuccessFactors. Esto se debe a que las claves externas no están presentes en las tablas SuccessFactors, y las uniones entre tablas se indican con tablas del tipo *Unión* en Spectrum™ Technology Platform.

Las características de las tablas de Unión son:

- El nombre de una tabla de *Unión* indica las dos tablas de *Entidad* que se asignaron juntas.
- Cada registro de una tabla de *Unión* contiene la clave principal de la entidad principal y las columnas de la entidad secundaria correspondiente. Por lo tanto, la clave principal particular de la entidad principal se asigna a los detalles de la entidad secundaria. Por ejemplo, `User#HR` es una tabla de *Unión* en la que `User` es la entidad principal y `Hr` es la entidad secundaria. Esta unión representa todos los usuarios y sus respectivos representantes de HR. Por lo tanto, la tabla

de unión `User#HR` tiene la clave principal de la tabla principal `User UserId`, que se asigna a las columnas de la tabla secundaria `HR`, tales como `hr_userId`, `hr_username` y `hr_email`.

- En el caso de las tablas de *unión*, las funciones `insert` y `update` funcionan como la función `upsert`. Esto se debe a que las tablas de *Unión* no son realmente entidades en SuccessFactors, sino asignaciones entre entidades y sus propiedades de navegación o tablas secundarias.

Para operaciones `insert` o `update` para cualquier tabla de *Unión*, se actualiza la entidad principal, mientras que en la tabla secundaria se inserta un nuevo registro o se actualiza el registro existente que corresponde con el registro principal.

Nota: Mientras se actualiza, se modifica la asignación entre un elemento principal y uno secundario. También es posible modificar los atributos individuales del elemento secundario según sea necesario.

Conexión con SugarCRM

1. Acceda a la página **Fuentes de datos** mediante uno de estos módulos:

Management Console: Acceda a Management Console a través de la dirección URL: `http://server:port/managementconsole`, donde `server` es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y `port` es el puerto HTTP que usa Spectrum™ Technology Platform.


Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Recursos > Fuentes de datos**.

Metadata Insights: Acceda a Metadata Insights a través de la dirección URL: `http://server:port/metadata-insights`, donde `server` es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y `port` es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Fuentes de datos**.

2. Haga clic en el botón **Agregar** .
3. En el campo **Nombre**, ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.

Nota: Una vez que guardó una conexión, no puede cambiarle el nombre.

4. En el campo **Tipo**, seleccione **SugarCRM**.
5. Ingrese su nombre de usuario y contraseña de SugarCRM.

6. En el campo **URL**, ingrese la URL de la cuenta de SugarCRM que se utilizará para la conexión.
7. Ingrese su **Id. del cliente** y **Secreto de cliente** de su cuenta SugarCRM.
8. Para probar la conexión, haga clic en **Probar**.
9. Haga clic en **Guardar**

Limitaciones de SugarCRM

1. Tanto para las consultas `UPDATE` como para las `UPSERT`, se realiza una operación `UPSERT`.
2. Las columnas **Anulable** y **Actualizable** en las propiedades de tabla, como se muestran en **Esquema de modelo físico** de la conexión, pueden no representar la operación correcta. Por ejemplo, una columna marcada como Actualizable puede arrojar una excepción del sistema cuando intenta actualizarla y, por el contrario, una columna marcada como Anulable puede no arrojar una excepción en la actualización.
3. Cuando realiza consultas usando combinaciones, debe usar un alias.

Entidades y operaciones admitidas

Operaciones admitidas

`LIKE` (su operación se limita a la elección de opciones que empiecen con el valor especificado, como en la afirmación `WHERE name LIKE 's%'`, en la cual se seleccionan todos los nombres que empiecen con la letra S), `ISNULL`, `IS NOT NULL`, `IN`, `NOT IN`, `>`, `>=`, `<`, `<=`, `=`, `<>`, `AND`, `OR`

Conexión con Oracle Eloqua

1. Acceda a la página **Fuentes de datos** mediante uno de estos módulos:

Management Console: Acceda a Management Console a través de la dirección URL: `http://server:port/managementconsole`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.


Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Recursos > Fuentes de datos**.

Metadata Insights: Acceda a Metadata Insights a través de la dirección URL: `http://server:port/metadata-insights`, donde *server* es el nombre del servidor o la dirección IP del Spectrum™ Technology Platform servidor y *port* es el puerto HTTP que usa Spectrum™ Technology Platform.

Nota: De manera predeterminada, el puerto HTTP es 8080.

Vaya a **Fuentes de datos**.

2. Haga clic en el botón **Agregar** .
3. En el campo **Nombre**, ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.

Nota: Una vez que guardó una conexión, no puede cambiarle el nombre.

4. En el campo **Tipo**, haga clic en **Oracle Eloqua**.
5. Ingrese el **Nombre de sitio** que es igual al Nombre de empresa.
6. Ingrese el **Nombre de usuario** en el campo correspondiente.
7. Ingrese la **Contraseña** en el campo correspondiente.
8. Haga clic en **Probar** para comprobar la conexión.
9. Haga clic en **Guardar**.

Operaciones especiales

1. Utilice esta consulta de combinación para recuperar los contactos en una lista de contactos:

```
select * from Contacts inner join ContactListMembers on
Contacts.Eloqua_Contact_ID = ContactListMembers.Contact_Id where
ContactListMembers.ContactList_Id = '<id>'
```

Utilice esta consulta de combinación para recuperar los contactos en un segmento de contactos:

```
select * from Contacts inner join ContactSegmentMembers on
Contacts.Eloqua_Contact_ID = ContactSegmentMembers.Contact_Id where
ContactSegmentMembers.Contactlist_Id = '<id>'
```

2. Utilice esta afirmación para insertar contactos en una lista de contactos:

```
insert into ContactListMembers (ContactList_ID,Contact_ID) values
('<contactlist_id>','<contact_id>')
```

3. Utilice esta afirmación para eliminar contactos de una lista de contactos:

```
delete from ContactListMembers where ContactList_ID =
'<contactlist_id>' and Contact_ID = '<contact_id>'
```

Limitaciones

A continuación, vea las limitaciones:

1. **Crear/Actualizar:**
 - a. Puede haber errores en la operación Insertar/Upsert (insertar y actualizar) si las columnas No es nulo están en blanco o no existen

- b. Puede haber errores en la operación `Insertar/Upsert` (insertar y actualizar) si los valores de las columnas `Único` no son únicos para un lote en específico
 - c. Para evitar una excepción de reversión, mantenga el valor de **Recuento de lote para realizar confirmación** en 1
2. **Leer:** en el caso de entidades personalizadas, la operación `Seleccionar` solo se puede aplicar en conjunto con la entidad `Contactos`
3. **Filtro:**
- a. Los filtros compatibles son los siguientes: `=, !=, >, <, >=, <=`
 - b. No se admiten los operadores condicionales `IN` y `NOT IN` cuando se proporciona más de un valor
 - c. No se admiten las `combinaciones` entre entidades
 - d. Solo se admiten los operadores condicionales de operación `AND` y `OR` para las entidades `Cuentas` y `Contactos`
 - e. Solo se admite el operador condicional `AND`. No funciona para las demás entidades de la columna
 - f. El filtro `=` no siempre funciona en los campos que tienen el tipo de datos `Timestamp`

Entidades y operaciones admitidas

En la siguiente tabla se describen las operaciones y las entidades compatibles con Oracle Eloqua.

Se proporciona compatibilidad con las siguientes entidades:

- **Entidad:** representa una tabla en la que se expresa una entidad empresarial.
- **Actividad:** representa una tabla en la que se expresa una entidad empresarial, cuyos datos se generan según una actividad.
- **Entidad personalizada:** representa entidades que se utilizan como parte de operaciones especiales que se proporcionan con el conector.

En esta tabla, se enumeran las entidades, junto con las operaciones admitidas para estas.

Nombre de entidad	Crear	Leer	Actualizar	Eliminar	Compatibilidad con lotes	Tamaño de lote máximo
Cuentas	X	X	X	X	Insertar/Actualizar*	1000
Grupos de cuenta		X				
Campaña		X				
Contactos	X	X	X	X	Insertar/Actualizar*	1000
Lista de contactos	X	X	X	X		

Nombre de entidad Crear Leer Actualizar Eliminar Compatibilidad con lotes Tamaño de lote máximo

Nombre de entidad	Crear	Leer	Actualizar	Eliminar	Compatibilidad con lotes	Tamaño de lote máximo
Segmento de contacto	X	X	X	X		
Correos electrónicos		X				
Carpetas de correo electrónico		X				
Grupos de correo electrónico		X				
Micrositios		X				
Usuarios		X				
Visitantes		X				
Actividad						
Apertura de correo electrónico		X				
Tasa de cantidad de clics de correo electrónico		X				
Envío de correo electrónico		X				
Suscripción		X				
Anulación de suscripción		X				
Rebote		X				
WebVisit		X				
PageView		X				
FormSubmit		X				

Nombre de entidad Crear Leer Actualizar Eliminar Compatibilidad con lotes Tamaño de lote máximo

Nombre de entidad	Crear	Leer	Actualizar	Eliminar	Compatibilidad con lotes	Tamaño de lote máximo
Entidades personalizadas						
ContactListMembers	X	X		X	Insertar/Eliminar	1000
ContactSegmentMembers		X				

* La operación de actualización funciona como “upsert” (insertar o actualizar).

Soporte de compresión para servidores de archivos de nube

Los servidores de archivos Amazon S3, almacenamiento en la nube de Google y MS Azure Blobstore admiten los formatos comprimidos `gzip (.gz)` y `zip (.zip)`.

Spectrum™ Technology Platform maneja la compresión y descompresión de los archivos que se escriben y se leen en los servidores de archivos.

Nota: Puede usar el mismo servidor de archivos para manejar las lecturas y escrituras normales de archivos y la compresión y descompresión de archivos.

Lectura de un archivo en formato comprimido

Durante la lectura de un archivo del servidor, su formato de compresión se deriva de la propiedad clave de metadatos `Content-Encoding` recibida del servidor.

Escritura de un archivo en formato comprimido


Durante la escritura de un archivo en un servidor, mencione el formato de compresión que se necesita: `.gz` o `.zip`. El archivo se comprime según la extensión de compresión especificada.

La propiedad clave de metadatos `Content-Encoding` también se establece según el formato de compresión seleccionado. Este valor de propiedad se traspasa al servidor de archivos en nube mientras el archivo se escribe en él.

Eliminación de una conexión

Puede eliminar la conexión mediante cualquiera de estos módulos:

- Management Console
- Metadata Insights

1. Acceda a la página de **Fuentes de datos** del módulo requerido.
 - En Management Console, haga clic en **Recursos > Fuentes de datos**.
 - En Metadata Insights, haga clic en **Fuentes de datos**.
2. Marque la casilla junto a la conexión que desea eliminar y, luego, haga clic en el botón **Eliminar** .

3 - Completar el almacén de datos

In this section

Preparación de los datos	71
Cómo completar una tabla de dimensiones de tiempo	72
Cómo completar una tabla de dimensiones	73
Cómo completar una tabla de hechos	75
Agregar una marca de tiempo a los registros de almacenamiento de datos	80

Preparación de los datos

Antes de llenar el depósito de datos, debe asegurarse de que los datos sean de buena calidad y de que tengan todos los atributos necesarios para ofrecer una perspectiva interesante para los usuarios de la empresa. Un enfoque común es utilizar un almacén de datos operacionales (ODS) para este fin. Un ODS es una base de datos en la que se pueden realizar operaciones de calidad de datos, con el fin de que se preparen los datos para que se carguen en el depósito de datos. Spectrum™ Technology Platform tiene una variedad de características que se pueden implementar en un ODS para mejorar la calidad de sus datos y también para ampliarlos con datos espaciales y demográficos, entre otros. Si actualmente no tiene con licencia las siguientes características, póngase en contacto con Pitney Bowes para obtener más información.

Análisis y estandarización y validación de nombres

Para llevar a cabo la normalización más precisa, puede ser necesario dividir las cadenas de datos en varios campos. Spectrum™ Technology Platform proporciona características de análisis avanzado que le permiten analizar nombres personales, nombres de empresas y muchos otros términos y abreviaturas. Además, puede crear su propia lista de términos personalizados para usar como base para las operaciones de exploración/ extracción. El módulo Universal Name proporciona esta funcionalidad.

Deduplicación y consolidación

Identificar entidades únicas le permite consolidar registros, eliminar duplicados y desarrollar registros "best-of-breed". Un registro "best-of-breed" es un registro compuesto que se construye con los datos de otros registros. El módulo Advanced Matching y el módulo Data Normalization proporcionan esta funcionalidad.

Validación de direcciones

La validación de direcciones aplica reglas de las autoridades postales correspondientes a fin de incluir cada dirección en un formato estándar e incluso validar que sea una dirección que permita la entrega. La validación de direcciones puede ayudar a reunir los requisitos necesarios para acceder a descuentos postales y mejorar la capacidad de entrega de su correo. El módulo Universal Addressing y el módulo Address Now proporcionan esta funcionalidad.

Geocodificación

La geocodificación es un proceso que consiste en tomar una dirección y determinar sus coordenadas geográficas (latitud y longitud). La geocodificación puede utilizarse para generar mapas pero eso es solo una de sus aplicaciones, ya que los datos de ubicación subyacentes también ayudan a tomar decisiones empresariales. En el proceso inverso, es posible ingresar un código geográfico (un punto representado por una coordenada de latitud y longitud) y recibir información de dirección acerca de ese código geográfico. El módulo Enterprise Geocoding proporciona esta funcionalidad.

Location Intelligence

Location Intelligence (Inteligencia de localización) crea nueva información acerca de sus datos mediante la evaluación, la examinación, el análisis y el modelado de relaciones geográficas. Por medio del proceso de inteligencia de localización, puede verificar ubicaciones y transformar la información en inteligencia empresarial de alto valor. El módulo Location Intelligence proporciona esta funcionalidad.

Asignación de jurisdicciones fiscales

La asignación de jurisdicciones fiscales consiste en tomar una dirección y determinar las jurisdicciones fiscales que le corresponden. Asignar jurisdicciones fiscales de la forma más precisa posible permite reducir los riesgos financieros y las responsabilidades reglamentarias.

El software Spectrum™ Technology Platform de Pitney Bowes™ integra los límites jurisdiccionales con las direcciones físicas exactas de los registros de los clientes, lo que le permite adjuntar a los registros los datos correctos de estado, condado, localidad, municipio y distritos fiscales especiales. Algunos ejemplos del uso de la asignación de jurisdicciones fiscales son:

- Impuestos de uso y ventas
- Impuestos a la propiedad personal
- Impuesto sobre las primas de seguros

El módulo Enterprise Tax proporciona esta funcionalidad.

Cómo completar una tabla de dimensiones de tiempo

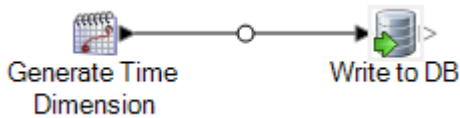
Una tabla de dimensiones de tiempo es una tabla en una base de datos que posibilita el análisis de los datos históricos, sin el uso de cálculos SQL complejos. Por ejemplo, puede analizar sus datos por jornadas laborales en comparación con días festivos, días de la semana en comparación con fines de semana, por períodos fiscales o por eventos especiales.

El siguiente procedimiento describe cómo usar Spectrum™ Technology Platform para completar una tabla de dimensiones de tiempo en su almacén de datos.

Nota: Antes de comenzar este procedimiento, debe tener definidas las conexiones al almacén de datos en el cual desee crear una tabla de dimensiones de tiempo. Si no ha definido las conexiones necesarias, consulte [Conexiones al origen de datos](#) en la página 14.

1. En Enterprise Designer, seleccione **Archivo > Nuevo > Flujo de datos > Trabajo**.
2. Arrastre la etapa Generate Time Dimension sobre el lienzo.
3. Arrastre una etapa Write To DB sobre el lienzo y conecte la etapa Generate Time Dimension a esta última.

El flujo de datos ahora debe tener la siguiente apariencia:



4. Haga doble clic en la etapa Generate Time Dimension y configúrela para producir las dimensiones de tiempo que desee. Para obtener más información, consulte [Generate Time Dimension](#) en la página 115.

Nota: El día juliano normalmente se usa como un valor clave para una tabla de dimensiones de tiempo, si el grano es un día o más. Si el grano es inferior a un día, puede generar una clave separada al añadir una etapa Unique ID Generator al flujo de datos. Si usa el día juliano como la clave, configure Generate Time Dimension para producir una columna de enteros para los valores de día juliano, y una columna con el tipo de dato o la fecha y hora para los valores de fecha.

5. Haga doble clic sobre la etapa Write to DB en el lienzo y configúrela para apuntar hacia la base de datos y la tabla donde desee crear la tabla de dimensiones de tiempo. Para obtener información acerca de la configuración de Write to DB, consulte [Write to DB](#) en la página 250.
6. Para obtener una vista previa de los valores de dimensión de tiempo antes de escribirlos en la tabla de dimensiones de tiempo:
 - a) Haga clic con el botón derecho del mouse en el canal que conecta la etapa Generate Time Dimension y la etapa Write to DB, y seleccione **Agregar punto de inspección..**
 - b) Seleccione **Ejecutar > Inspeccionar flujo actual.**
Aparece el panel de inspección en la parte inferior de la ventana de Enterprise Designer y muestra los datos que se escribirán en la tabla de dimensiones de tiempo. Si es necesario, puede realizar ajustes en la etapa Generate Time Dimension y luego volver a ejecutar el proceso de inspección, para ver el efecto de sus cambios.
7. Cuando esté satisfecho con el flujo de datos, seleccione **Ejecutar > Ejecutar flujo actual** para ejecutar el flujo de datos y completar la tabla de dimensiones de tiempo.

Cómo completar una tabla de dimensiones

Las tablas de dimensiones forman parte de un esquema de estrella e incluyen información detallada para las columnas en la tabla de hechos. Las tablas de dimensión tienen atributos y una clave principal de una sola parte que une la tabla de dimensión con la tabla de hechos. La clave principal de una sola parte permite que explore con rapidez una única tabla de dimensión. Explorar una tabla de dimensión puede ayudar a determinar la mejor manera de consultar la tabla de hechos.

El siguiente procedimiento describe cómo usar Spectrum™ Technology Platform para completar inicialmente una tabla de dimensiones en su almacén de datos.

Nota: Antes de comenzar este procedimiento, debe definir las conexiones a los recursos externos que desea usar como origen de la tabla de dimensiones si está usando una base de datos, un servidor de archivos o un servicio web como el origen de los datos. También debe definir una conexión al almacén de datos en el cual desee crear una tabla de dimensiones. Si no ha definido las conexiones necesarias, consulte [Conexiones al origen de datos](#) en la página 14.

1. En su almacén de datos, cree la tabla que desee usar como tabla de dimensiones.
2. En Management Console, cree las conexiones al origen de datos y al almacén de datos.
 - Para conectar a una base de datos, consulte [Conexión con una base de datos JDBC](#) en la página 33.
 - Para conectar a un servidor de archivos, consulte [Conexión con un servidor FTP](#) en la página 25.
3. En Enterprise Designer, seleccione **Archivo > Nuevo > Flujo de datos > Trabajo**.
4. Arrastre la etapa de origen sobre el lienzo.
 - Para usar los datos desde una base de datos para completar la tabla, arrastre la etapa **Read from DB** sobre el lienzo.
 - Para usar los datos desde un archivo plano para completar la tabla, arrastre la etapa **Read from File** sobre el lienzo.
 - Para usar los datos desde un archivo de formato variable para completar la tabla, arrastre la etapa **Read from Variable Format File** sobre el lienzo.
 - Para usar los datos desde un archivo XML para completar la tabla, arrastre la etapa **Read from XML** sobre el lienzo.
5. Haga doble clic en la etapa de origen que acaba de ubicar en el lienzo y configúrela para que apunte hacia el origen de los datos con que desea rellenar la tabla de dimensión.
 - Para obtener más información sobre la configuración de Read from DB, consulte [Read from DB](#) en la página 143
 - Para obtener más información sobre la configuración de Read From File, consulte [Read From File](#) en la página 152
 - Para obtener más información sobre la configuración de Read From Variable Format File, consulte [Read From Variable Format File](#) en la página 202
 - Para obtener más información sobre la configuración de Read From XML, consulte [Read from XML](#) en la página 217
6. Arrastre una etapa Unique ID Generator sobre el lienzo y conecte la etapa de origen a esta última. Por ejemplo, si está usando Read from DB como la etapa de origen, debe conectar Read from DB a Unique ID Generator.
7. Haga doble clic en la etapa Unique ID Generator en el lienzo y configúrela para crear una clave suplente.

Nota: Normalmente, la clave del sistema operativo no se usa como clave principal para una dimensión en el almacén. Esto ayuda a mantener la uniformidad histórica, ya que el valor clave puede cambiar en el sistema operativo.

8. Arrastre una etapa Write To DB sobre el lienzo y conecte la etapa Unique ID Generator a esta última.
9. Haga doble clic en la etapa Write to DB en el lienzo y configúrela para que apunte hacia la base de datos y la tabla de dimensión que desea rellenar. Para obtener información acerca de la configuración de Write to DB, consulte [Write to DB](#) en la página 250.
10. Seleccione **Archivo > Guardar** y guarde el flujo de datos.
11. Para ejecutar ahora el flujo de datos y completar la tabla de dimensiones, seleccione **Ejecutar > Ejecutar flujo actual**.

Cómo completar una tabla de hechos

Después de completar las tablas de dimensiones en su almacén de datos, ya está listo para completar la tabla de hechos. La tabla de hechos se completa con mediciones numéricas de las tablas en la base de datos OLTP.

Importante: Debe completar las tablas de dimensiones antes de completar la tabla de hechos.

El siguiente procedimiento describe cómo usar Spectrum™ Technology Platform para completar una tabla de hechos en su almacén de datos. En este procedimiento, se creará un flujo de datos que lee en los datos de origen desde una tabla en su base de datos de origen, reemplaza las claves naturales de las tablas de origen por las claves suplentes de las tablas de dimensiones, y luego carga el registro actualizado que incluye las claves suplentes y los datos de hechos desde las tablas de origen en la tabla de hechos.

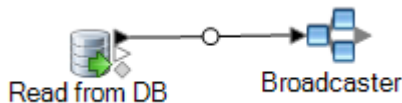
1. En Management Console, cree las conexiones al origen de datos y al almacén de datos.
 - Para conectar a una base de datos, consulte [Conexión con una base de datos JDBC](#) en la página 33.
 - Para conectar a un servidor de archivos, consulte [Conexión con un servidor FTP](#) en la página 25.
2. En Enterprise Designer, seleccione **Archivo > Nuevo > Flujo de datos > Trabajo**.
3. En función de los datos que desea escribir en la tabla de hechos, arrastre hacia el lienzo la etapa que corresponda.
 - Para usar los datos desde una base de datos para completar la tabla, arrastre la etapa **Read from DB** sobre el lienzo.
 - Para usar los datos desde un archivo plano para completar la tabla, arrastre la etapa **Read from File** sobre el lienzo.

- Para usar los datos desde un archivo de formato variable para completar la tabla, arrastre la etapa **Read from Variable Format File** sobre el lienzo.
 - Para usar los datos desde un archivo XML para completar la tabla, arrastre la etapa **Read from XML** sobre el lienzo.
4. Haga doble clic en la etapa de origen que acaba de ubicar en el lienzo y configúrela para que apunte hacia el origen de los datos con que desea rellenar la tabla de hechos.
- Para obtener más información sobre la configuración de Read from DB, consulte [Read from DB](#) en la página 143
 - Para obtener más información sobre la configuración de Read From File, consulte [Read From File](#) en la página 152
 - Para obtener más información sobre la configuración de Read From Variable Format File, consulte [Read From Variable Format File](#) en la página 202
 - Para obtener más información sobre la configuración de Read From XML, consulte [Read from XML](#) en la página 217

Nota: Normalmente, un flujo de datos de datos que completa una tabla de hechos, lee desde una base de datos en lugar de un archivo. Debido a que este es la situación más común, los ejemplos en el resto de este procedimiento usa Read From DB.

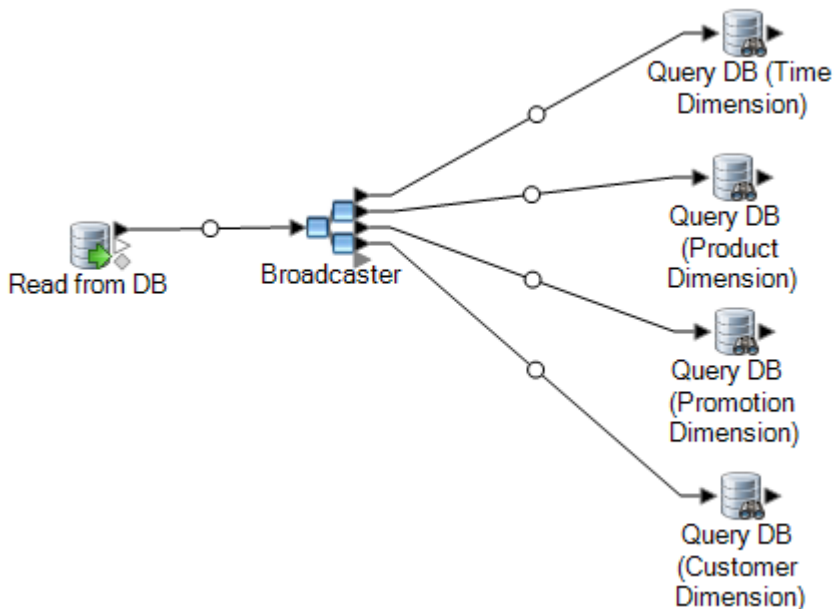
5. Arrastre una etapa Broadcaster sobre el lienzo y conecte la etapa de origen a esta última.

Su flujo de datos ahora tiene la siguiente apariencia:



6. Arrastre una etapa Query DB sobre el lienzo para cada tabla de dimensiones en su almacén de datos y conéctelas a la etapa Broadcaster.

Por ejemplo, si tiene cuatro tablas de dimensiones en su almacén de datos, arrastre cuatro etapas Query DB sobre el lienzo. Su flujo de datos debería tener la siguiente apariencia:



Las etapas Query DB se usarán para buscar la clave suplente para cada dimensión, mediante el uso de la clave natural desde el origen de datos. La clave suplente posteriormente reemplazará a la clave natural en cada registro que esté cargado en la tabla de hechos.

Consejo: Puede modificar el nombre de la etapa para facilitar ver qué tabla consulta cada etapa.

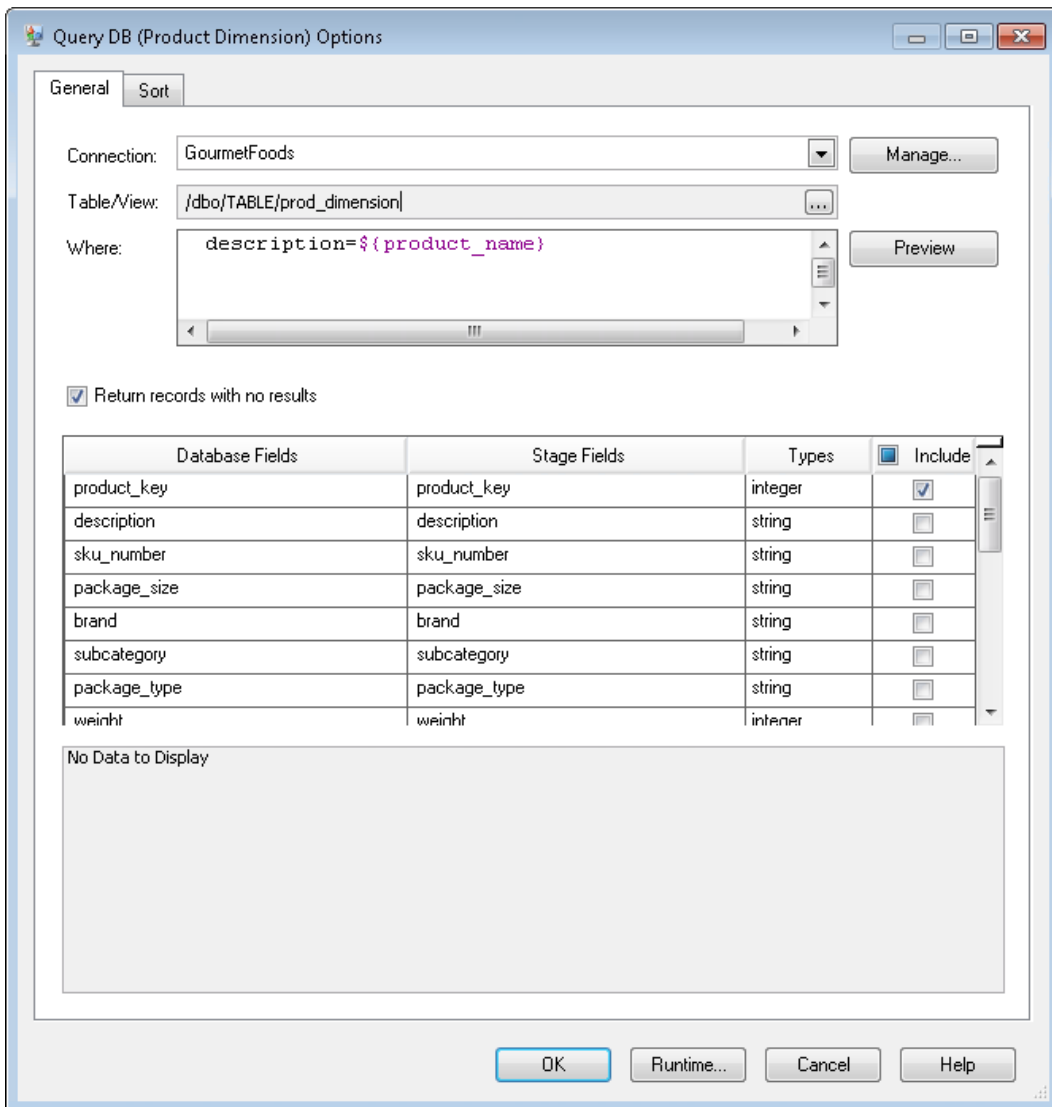
7. Configure cada etapa Query DB de modo que busque la clave suplente para cada clave natural desde el origen de datos. Para esto, haga lo siguiente:
 - a) En el campo **Conexión**, especifique la conexión al depósito de datos.
 - b) En el campo **Tabla/Vista** seleccione la tabla de dimensiones que desee que esta etapa consulte.
 - c) En el campo **Donde**, escriba una instrucción `WHERE` que busque la clave suplente según el valor en el campo de flujo de datos correspondiente.

Por ejemplo, esto buscaría la clave suplente de un producto, al encontrar el registro en la tabla de dimensiones cuyo valor en la columna `description` coincide con el valor en el campo `product_name` del origen de datos.

```
description=${product_name}
```

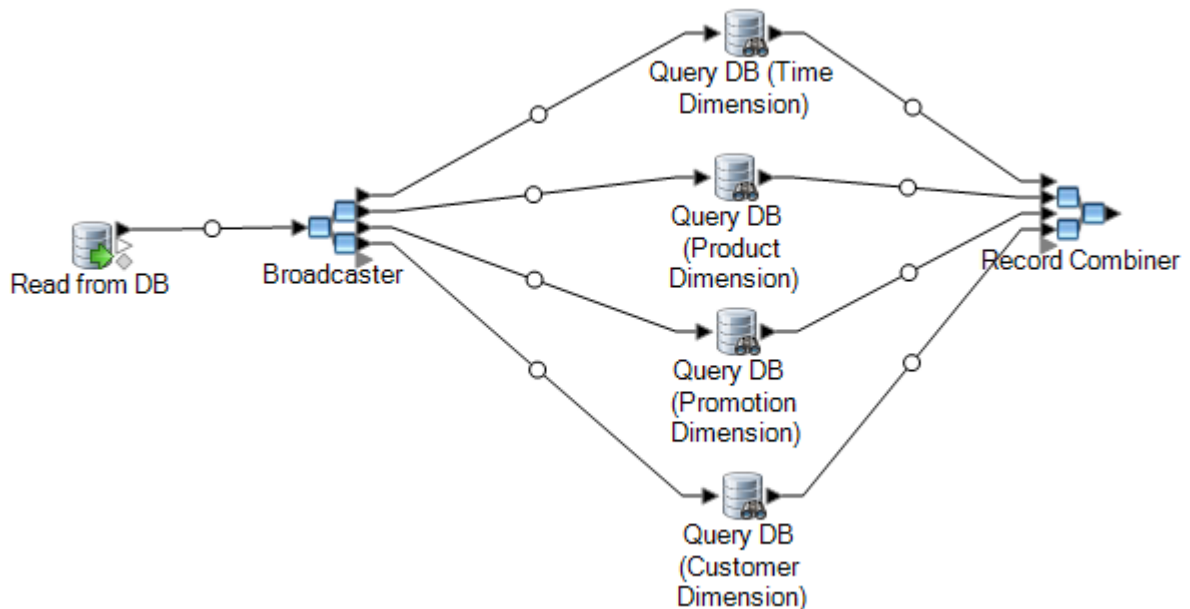
- d) En la columna **Incluir** seleccione la columna de la base de datos que contiene la clave suplente.

Por ejemplo, una etapa Query DB que busca la clave suplente de un nombre de producto tendría la siguiente apariencia:

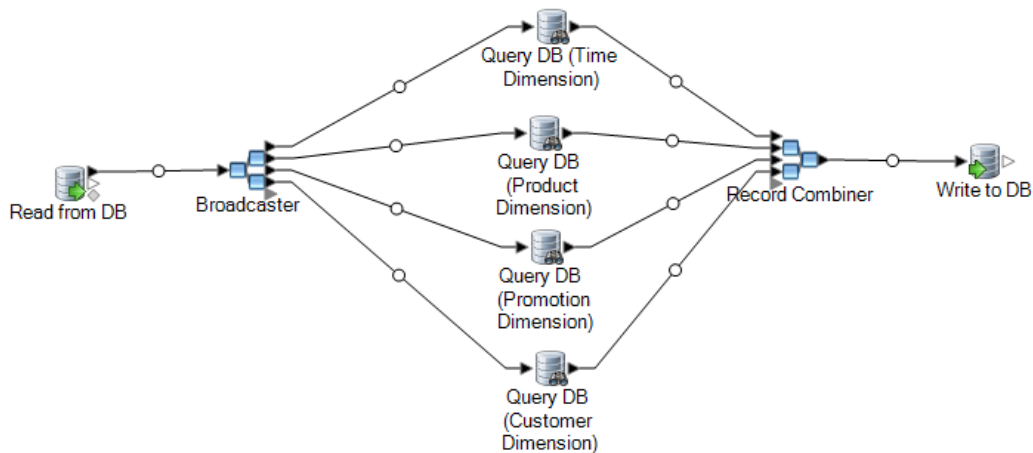


En este ejemplo, la consulta busca la clave de producto al encontrar el registro en la tabla `prod_dimension`, donde el valor en la columna `description` coincide con el valor en el campo del flujo de datos `product_name`. La etapa devuelve el campo `product_key` desde la tabla y lo agrega al flujo de datos, según se indique mediante la casilla marcada en la columna **Incluir**.

8. Arrastre una etapa Record Combiner sobre el lienzo y conecte la etapa Query DB a esta última. Su flujo de datos debería tener ahora el siguiente aspecto:



9. Arrastre una etapa Write To DB sobre el lienzo y conéctela a la etapa Record Combiner. Su flujo de datos debería tener ahora el siguiente aspecto:



10. Configure la etapa Write to DB para escribir los registros en la tabla de hechos. Para esto, haga lo siguiente:
- En el campo **Conexión**, especifique la conexión al depósito de datos.
 - En el campo **Tabla/Vista** seleccione la tabla de hechos que desee que esta etapa consulte. Si la tabla de hechos no existe en el almacén de datos, haga clic en **Crear tabla** para crear la tabla de hechos en el almacén de datos.
 - Marque la casilla en la columna **Incluir** por cada campo que desee escribir en la tabla de hechos.
 - En la ficha **Ejecución**, observe que la opción **Insertar** está seleccionada predeterminadamente para el modo de escritura. Normalmente la completación de la tabla

de hechos se ejecuta en el modo de inserción, de modo que puede dejar seleccionada esta opción.

11. Guarde y ejecute su flujo de datos.

Ejemplo de reemplazo de datos de origen por claves en la tabla de dimensiones

Considere este registro:

```
March 28 2013,Parsley Garlic Pasta,Mile High Gourmet
Market,78.35
```

En este ejemplo, existe un campo de fecha, seguido por un nombre de producto (Parsley Garlic Pasta), un cliente (Mile High Gourmet Market) y una cantidad (78.25). El almacén de datos tiene tablas de dimensiones para la fecha, el nombre de producto y el cliente, de modo que las claves naturales en el registro se deben reemplazar por claves suplentes desde las tablas de dimensiones. Para lograr esto, el flujo de datos debe tener tres etapas Query DB, una que busque la clave suplente para la fecha, una que busque la clave suplente para el nombre de producto y otra para que busque la clave suplente para el cliente.

Cada Query DB debe tener una instrucción `WHERE` que busque la clave suplente.

Como resultado de estas búsquedas, el registro puede tener la siguiente apariencia cuando se escriba en la tabla de hechos:

```
711,1,15,78.35
```

Observe que las claves naturales de la fecha, el nombre de producto y el cliente se reemplazaron por claves suplentes.

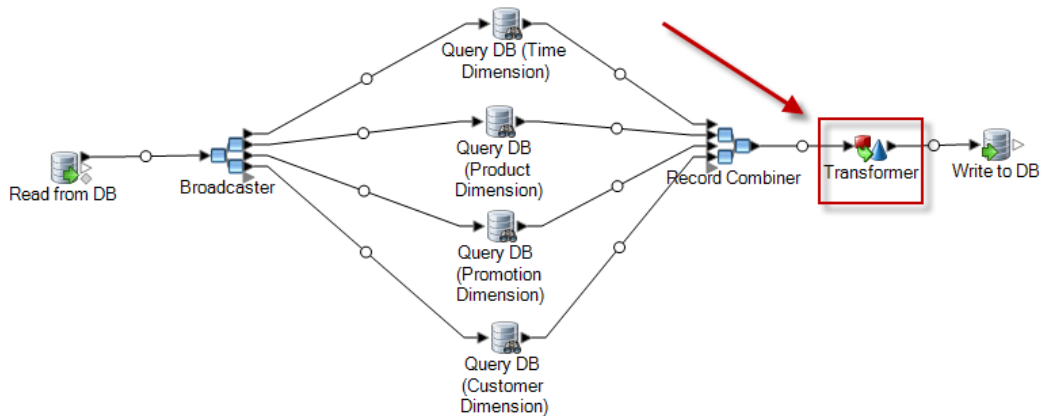
Agregar una marca de tiempo a los registros de almacenamiento de datos

Una forma conveniente de asegurar la calidad de los datos es marcar los registros almacenados de acuerdo a la fecha en que se cargaron. Si el proceso de carga no se completa o detecta problemas después de que los datos ya están cargados, una columna de marca de tiempo facilita la identificación de los registros afectados. Después, puede eliminar todos los registros procesados durante una determinada fase, volver al estado de precarga y manejar los problemas antes de intentar cargar los datos de nuevo. Puede poner una marca de tiempo en la operación de carga al agregar una columna adicional, como `load_date`, a su tabla de hechos con la etapa SQL Command.

Para agregar una marca en el flujo de datos al rellenar o actualizar el almacenamiento de datos:

1. En Enterprise Designer, abra el flujo de datos que rellena o actualiza el almacenamiento de datos:
2. Arrastre la etapa Transformer al lienzo y conéctela al flujo de datos antes de la etapa Write To DB.

Por ejemplo:



3. Haga doble clic en la etapa Transformer.
4. Haga clic en **Agregar**.
5. En **General**, seleccione **Personalizada**.
6. En el campo **Nombre de la transformación personalizada**, ingrese el nombre para esta transformación. El nombre puede ser cualquiera que elija. Por ejemplo, Agregar marca de tiempo.
7. En el campo **Secuencia de comandos personalizada**, ingrese lo siguiente:

```
data['<timestamp field>']=currentDateTime()
```

Donde <timestamp field> es el nombre del campo de flujo de datos donde desea incluir la marca de tiempo.

Por ejemplo, si desea poner la marca de tiempo en un campo de flujo de datos llamado Timestamp, entonces el comando personalizado sería:

```
data['Timestamp']=currentDateTime()
```

8. Haga clic en el botón **Agregar** que está en la parte inferior de la ventana.
9. Haga clic en **Cerrar**.
10. Haga clic en **Aceptar** para cerrar la ventana de **Opciones de Transformer**.

Ahora el flujo de datos agrega la hora actual en un campo de cada registro, lo que proporciona una marca de tiempo en el almacenamiento de datos, que muestra el momento en que se cargó cada registro.

4 - Actualizar el almacén de datos

In this section

Definición de un programa de actualización del almacén de datos	83
Cómo actualizar una tabla de hechos	84
Uso de una memoria caché global para las consultas	89
Uso de una caché local para consultas	90

Definición de un programa de actualización del almacén de datos

Puede programar Spectrum™ Technology Platform flujos de datos para extraer y transformar los datos de la estructura normalizada de un origen de datos en la estructura del esquema de estrella del almacenamiento de datos. Programar flujos de datos es útil, ya que la mayoría de las operaciones de carga requieren recursos del sistema que no están disponibles durante el día laboral.

A la hora de elegir el programa de actualización, se debe considerar lo siguiente:

- Frecuencia
- Secuencia
- Dependencias

Frecuencia

Debe programar flujos de datos que se ejecuten en base a la granularidad de la tabla de hechos más detallada. Por ejemplo:

- Si la granularidad de la tabla de hechos es diaria, programe el flujo de datos de llenado de la tabla de hechos para que se ejecute cada día.
- Si la granularidad es mensual, programe el flujo de datos de llenado de la tabla de datos para que se ejecute mensualmente y no antes, ya que los usuarios pueden trabajar solo con los datos de meses anteriores.

La mayor parte los flujos de datos de llenado procesan grandes cantidades de datos, por lo que es recomendable programar los flujos de datos e llenados para que se ejecuten cuando el uso del servidor Spectrum™ Technology Platform, el origen, las bases de datos del almacén de datos y la red es mínimo.

Rellene todas las dimensiones y tablas de hechos durante la carga inicial. Después de la carga inicial, actualice las tablas en base a lo que se ha agregado o cambiado. En general, las tablas de hecho se actualizan con más frecuencia que las tablas de dimensiones debido a que:

- Las tablas de dimensiones generalmente son estáticas, a menos que uno de los atributos de la fuente se haya modificado o agregado.
- Los datos de una tabla de hechos en una base de datos de apoyo, comúnmente contiene información histórica y requiere que se agreguen datos periódicamente y se mantenga actualizada. La carga inicial y la mayoría de las cargas incrementales afectan a las tablas de hechos.

Secuencia

Existen relaciones de dependencia entre los datos en las bases de datos del almacén de datos, por lo que debe determinar la secuencia en la que se debe ejecutar el llenado de los flujos de datos antes de fijar el cronograma de ejecución.

Rellene las tablas de dimensión antes de las tablas de hechos, ya que cada registro clave y de dimensión debe existir antes de que se llenen las tablas de hechos. Esta restricción es una función de la relación clave principalmente extranjera entre las tablas de dimensión y de hechos en un esquema de estrella.

Actualice las tablas de nivel básico antes de llenar las tablas agregadas en su base de datos de apoyo de decisiones. Esta secuencia garantiza que las tablas agregadas y de nivel básico permanecerán sincronizadas.

El orden correcto de ejecución de llenado de un flujo de datos es:

1. Planes de una tabla de dimensión de nivel básico
2. Planes de una tabla de hechos de nivel básico
3. Planes de una tabla de dimensión agregada
4. Planes de una tabla de hechos agregada

Dependencias

Puede crear dependencias de flujos de datos si se requiere llenar varios flujos de datos para ejecutar una orden en específico o si no se puede predecir la cantidad de tiempo para ejecutar un flujo de datos. Solo se debe ejecutar un flujo de datos si se cumplen ciertos requisitos, como que el flujo de datos anterior se haya completado o el flujo de datos anterior haya fallado.

Para crear dependencias de flujos de datos, cree un flujo de proceso en Enterprise Designer. Para obtener más información sobre los flujos de proceso, consulte la *Guía del diseñador de flujo de datos*.

Cómo actualizar una tabla de hechos

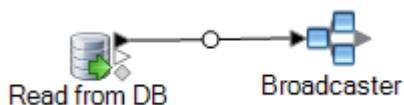
Este procedimiento describe cómo crear un flujo de datos que lee los datos de una base de datos o un archivo de origen y que los utiliza para actualizar una tabla de datos en el almacén de datos.

1. En Enterprise Designer, seleccione **Archivo > Nuevo > Flujo de datos > Trabajo**.
2. En función de los datos que desea escribir en la tabla de hechos, arrastre hacia el lienzo la etapa que corresponda.
 - Para usar los datos desde una base de datos para completar la tabla, arrastre la etapa **Read from DB** sobre el lienzo.
 - Para usar los datos desde un archivo plano para completar la tabla, arrastre la etapa **Read from File** sobre el lienzo.
 - Para usar los datos desde un archivo de formato variable para completar la tabla, arrastre la etapa **Read from Variable Format File** sobre el lienzo.
 - Para usar los datos desde un archivo XML para completar la tabla, arrastre la etapa **Read from XML** sobre el lienzo.

Nota: Si va a leer los datos de un archivo y no una base de datos, asegúrese de que el archivo solo contenga los nuevos registros que desea agregar a la tabla de hechos y no los registros que ya existen en la tabla de hechos. Si va a leer los datos de una base de datos, debe definir una consulta para filtrar los registros más adelante en este procedimiento.

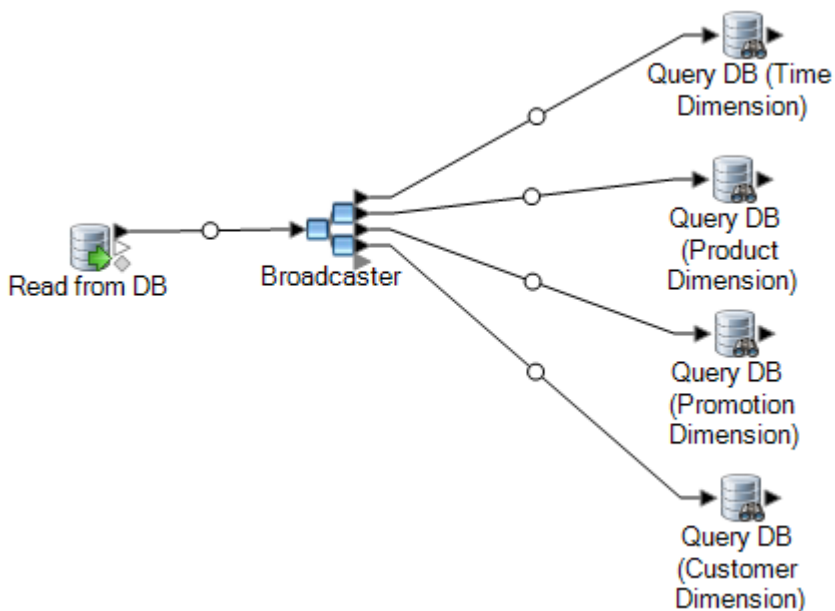
3. Haga doble clic en la etapa de origen que acaba de ubicar en el lienzo y configúrela para que apunte hacia el origen de los datos con que desea rellenar la tabla de hechos.
 - Para obtener más información sobre la configuración de Read from DB, consulte [Read from DB](#) en la página 143
 - Para obtener más información sobre la configuración de Read From File, consulte [Read From File](#) en la página 152
 - Para obtener más información sobre la configuración de Read From Variable Format File, consulte [Read From Variable Format File](#) en la página 202
 - Para obtener más información sobre la configuración de Read From XML, consulte [Read from XML](#) en la página 217
4. Si lee los datos de una base de datos, filtre los registros para que solo los registros nuevos se agreguen a la tabla de hechos. Para hacerlo, puede definir que la instrucción SQL SELECT solo lea los registros que se han modificado desde la última vez que la tabla de hechos se actualizó.
5. Arrastre una etapa Broadcaster sobre el lienzo y conecte la etapa de origen a esta última.

Su flujo de datos ahora tiene la siguiente apariencia:



6. Arrastre una etapa Query DB sobre el lienzo para cada tabla de dimensiones en su almacén de datos y conéctelas a la etapa Broadcaster.

Por ejemplo, si tiene cuatro tablas de dimensiones en su almacén de datos, arrastre cuatro etapas Query DB sobre el lienzo. Su flujo de datos debiera tener la siguiente apariencia:



Las etapas Query DB se usarán para buscar la clave suplente para cada dimensión, mediante el uso de la clave natural desde el origen de datos. La clave suplente posteriormente reemplazará a la clave natural en cada registro que esté cargado en la tabla de hechos.

Consejo: Puede modificar el nombre de la etapa para facilitar ver qué tabla consulta cada etapa.

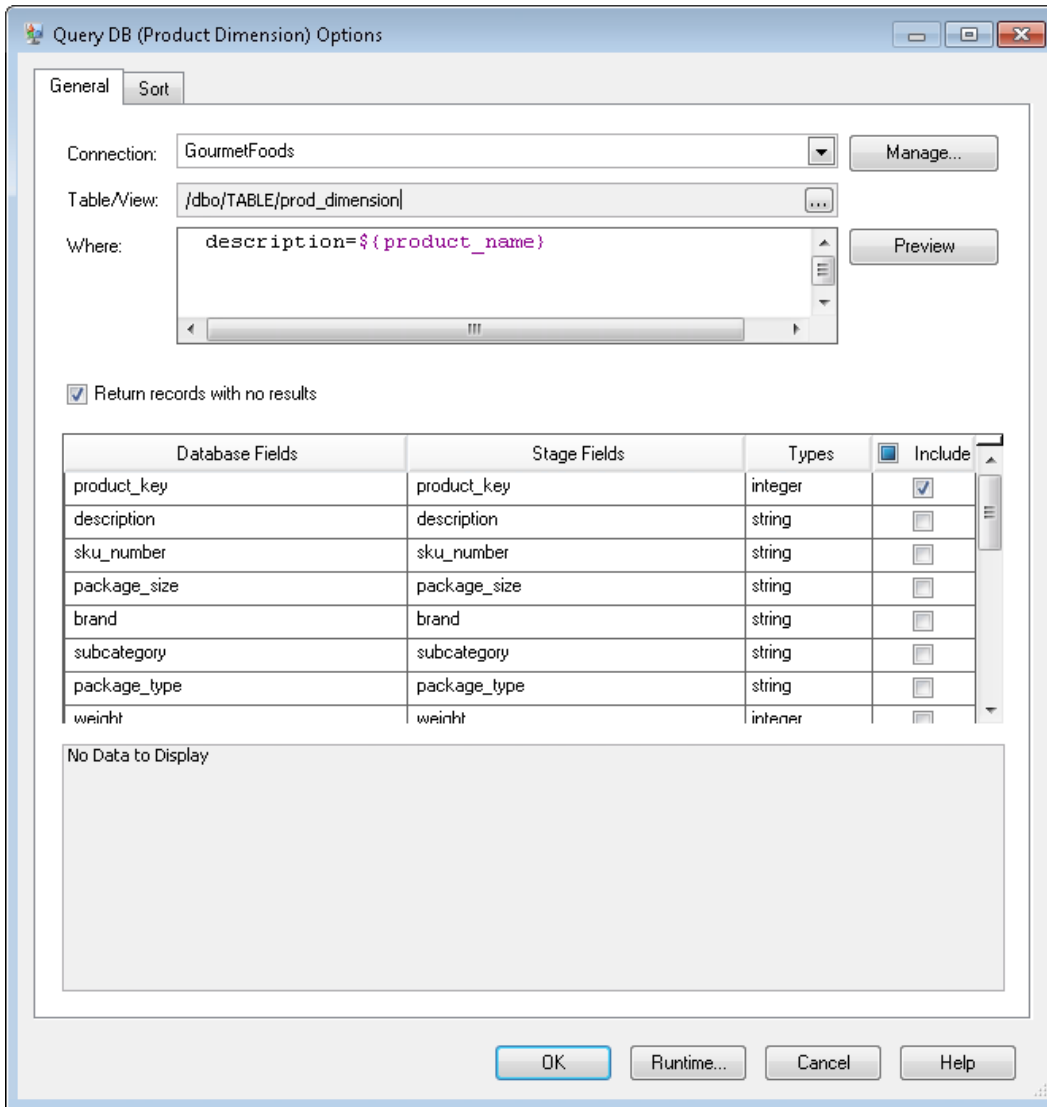
7. Configure cada etapa Query DB de modo que busque la clave suplente para cada clave natural desde el origen de datos. Para esto, haga lo siguiente:
 - a) En el campo **Conexión**, especifique la conexión al depósito de datos.
 - b) En el campo **Tabla/Vista** seleccione la tabla de dimensiones que desee que esta etapa consulte.
 - c) En el campo **Donde**, escriba una instrucción `WHERE` que busque la clave suplente según el valor en el campo de flujo de datos correspondiente.

Por ejemplo, esto buscaría la clave suplente de un producto, al encontrar el registro en la tabla de dimensiones cuyo valor en la columna `description` coincide con el valor en el campo `product_name` del origen de datos.

```
description=${product_name}
```

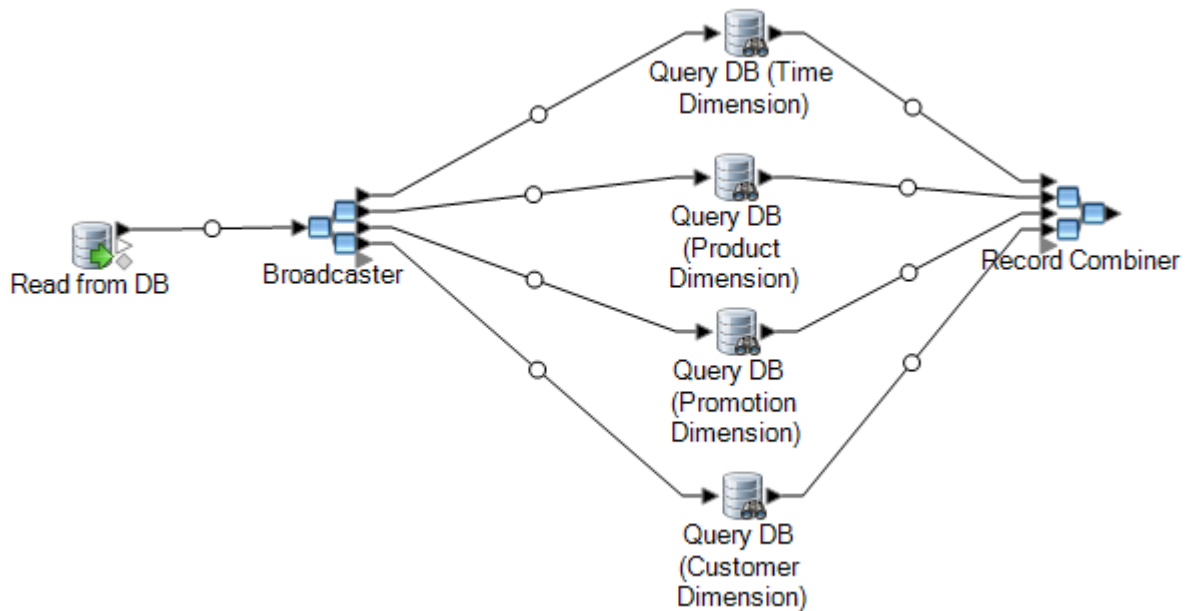
- d) En la columna **Incluir** seleccione la columna de la base de datos que contiene la clave suplente.

Por ejemplo, una etapa Query DB que busca la clave suplente de un nombre de producto tendría la siguiente apariencia:

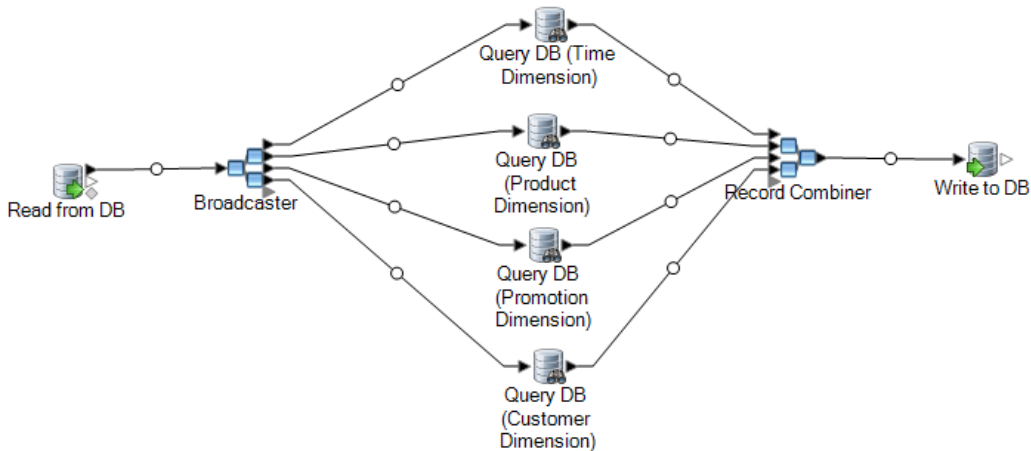


En este ejemplo, la consulta busca la clave de producto al encontrar el registro en la tabla `prod_dimension`, donde el valor en la columna `description` coincide con el valor en el campo del flujo de datos `product_name`. La etapa devuelve el campo `product_key` desde la tabla y lo agrega al flujo de datos, según se indique mediante la casilla marcada en la columna **Incluir**.

- Arrastre una etapa Record Combiner sobre el lienzo y conecte la etapa Query DB a esta última. Su flujo de datos debería tener ahora el siguiente aspecto:



9. Arrastre una etapa Write To DB sobre el lienzo y conéctela a la etapa Record Combiner. Su flujo de datos debería tener ahora el siguiente aspecto:



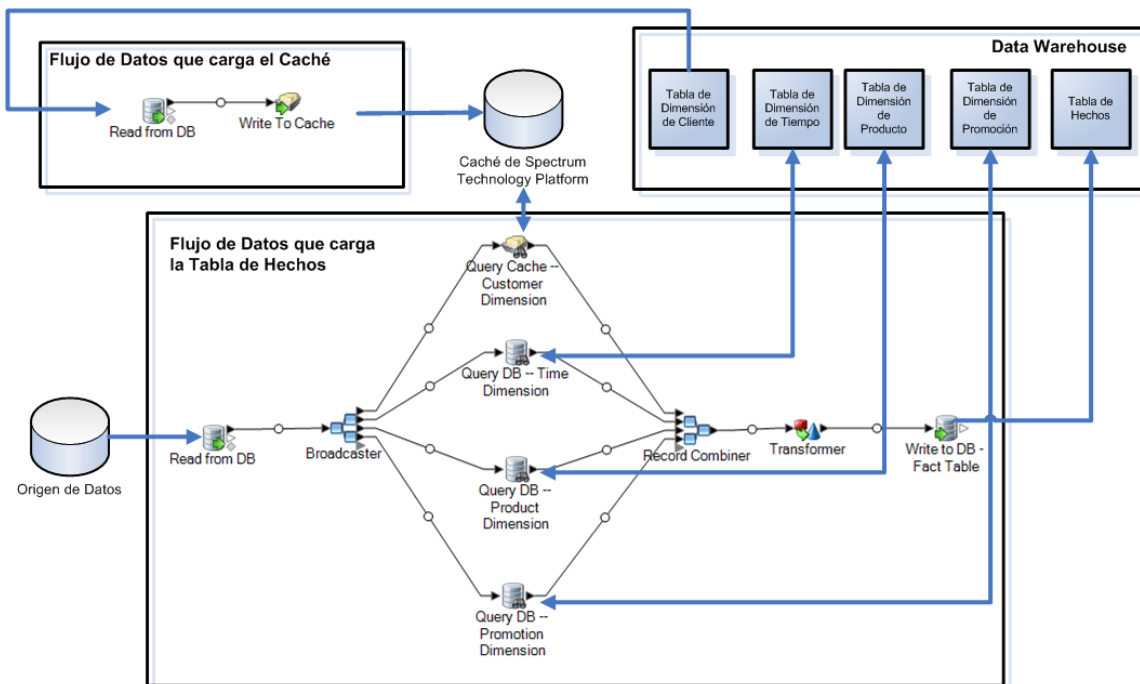
10. Configure la etapa Write to DB para escribir los registros en la tabla de hechos. Para esto, haga lo siguiente:
- En el campo **Conexión**, especifique la conexión al depósito de datos.
 - En el campo **Tabla/Vista** seleccione la tabla de hechos que desee que esta etapa consulte. Si la tabla de hechos no existe en el almacén de datos, haga clic en **Crear tabla** para crear la tabla de hechos en el almacén de datos.
 - Marque la casilla en la columna **Incluir** por cada campo que desee escribir en la tabla de hechos.
 - En la ficha **Ejecución**, observe que la opción **Insertar** está seleccionada predeterminadamente para el modo de escritura. Normalmente la completación de la tabla

de hechos se ejecuta en el modo de inserción, de modo que puede dejar seleccionada esta opción.

Uso de una memoria caché global para las consultas

Si tiene una tabla de grandes dimensiones, puede cargar los datos de la tabla de dimensiones en una memoria caché, y utilizarla para buscar claves de reemplazo. Usar una memoria caché mejora el rendimiento, en comparación con la búsqueda directa en la tabla de dimensiones mediante la base de datos de consulta.

Para utilizar una memoria caché debe crear dos flujos de datos: uno para llenar la memoria caché con los datos de la tabla de dimensiones y otro que utilice la memoria caché durante la actualización de la tabla de hechos. El siguiente diagrama ilustra cómo los dos flujos trabajar juntos:



1. Creación de un flujo de datos que llene la memoria caché con datos de la tabla de dimensiones de la tabla de grandes dimensiones.

Este flujo de datos debe tener dos etapas:

- Una etapa Read from DB que lea los datos de la tabla de dimensiones que desea cargar en la memoria caché.
- Una etapa Write to Cache que llene la memoria caché con los datos de la tabla dimensiones.

2. Ejecute este flujo de datos para llenar la memoria caché.

3. En el flujo de datos que llena la tabla de hechos, agregue una memoria caché de consulta.
4. En la etapa Query Cache, configure la etapa para llenar con consultas la caché creada en la etapa Write to Cache.
5. Ejecute este flujo de datos para llenar la tabla de hechos.


Si desea asegurarse de que la memoria caché se llena con los datos más recientes de la tabla de dimensiones cada vez que actualiza la tabla de hechos, puede crear un flujo de proceso que primero ejecute el trabajo para llenar la tabla de dimensiones y, a continuación, ejecute el trabajo para actualizar la tabla de hechos. Esto le permite activar el flujo de proceso con el fin de ejecutar los flujos de datos en sucesión. Para obtener más información sobre los flujos de proceso, consulte la *Guía del diseñador de flujo de datos*.

Eliminación de una memoria caché

En la búsqueda de datos, la etapa Query Cache, de caché de consulta, utiliza una memoria caché global. Si ya no se necesita la memoria caché, puede eliminarla mediante el siguiente procedimiento. La eliminación de una memoria caché libera espacio en la memoria.

Nota: Eliminación de una memoria caché se usa solamente para eliminar una memoria caché global de Management Console.

Aviso: Antes de eliminar una memoria caché, asegúrese de que ningún flujo de datos que contenga una etapa Query Cache la esté utilizando. Si elimina una memoria caché que está siendo utilizada por un flujo de datos, se producirá una falla en el flujo de datos.

1. Abra Management Console.
2. Expanda **Recursos > Administración de la memoria caché**. En la cuadrícula, se muestran las memorias caché existentes.
3. Seleccione la memoria caché que desee eliminar.
4. Haga clic en .

Uso de una caché local para consultas

Si tiene una tabla de grandes dimensiones, puede cargar los datos de la tabla de dimensiones en una memoria caché, y utilizarla para buscar claves de reemplazo. Usar una memoria caché mejora el rendimiento, en comparación con la búsqueda directa en la tabla de dimensiones mediante la base de datos de consulta.

Una caché local es una caché provisoria que se usa únicamente durante la ejecución de una etapa Query Cache. Luego, busca datos en la caché basándose en los campos clave y las condiciones de búsqueda y devuelve datos a partir de registros coincidentes en la caché, añadiendo los datos

del registro de la caché al registro en el flujo de datos. Utilice una caché local en lugar de una caché global si solo se va a utilizar en un flujo de datos o si la tabla de búsqueda cambia con frecuencia.

Para usar una caché local para consultas:

1. En Enterprise Designer, abra el flujo de datos en el que desea realizar una consulta con una caché.
2. Arrastre hacia el lienzo una etapa Query Cache y conéctela al flujo de datos.
3. Haga doble clic en la caché de consultas.
4. Seleccione **Caché local**.
5. Seleccione la conexión de base de datos que quiere utilizar. Las opciones pueden variar según las conexiones definidas en el Administrador de conexión de base de datos de Management Console. Si necesita hacer una conexión de base de datos nueva o modificar o eliminar una conexión de base de datos existente, haga clic en **Administrar**.

Si va a agregar o modificar una conexión de base de datos, complete los siguientes campos:

Nombre de conexión Ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.

Controlador de base de datos Seleccione el tipo de base de datos correspondiente.

Opciones de conexión Especifique el host, el puerto, la instancia, el nombre de usuario y la contraseña que se utilizarán para conectarse con la base de datos.

6. Especifique la tabla o vista en la base de datos que se desea consultar.
7. Seleccione una clave en la opción **Campo clave**.
8. Seleccione un campo de entrada en la opción **Campo de entrada**. El **Campo de entrada** contiene los campos procedentes de la etapa anterior. Si el valor de este campo coincide con la clave en **Campo clave** en la base de datos, la consulta devuelve los datos del registro en la base de datos.
9. Haga clic en **Aceptar**.
10. Ejecute el flujo de datos.

5 - Referencia de etapas

In this section

Invocar procedimiento almacenado	94
DB Change Data Reader	96
DB Loader	100
Field Parser	110
Field Combiner	113
Field Selector	114
Generate Time Dimension	115
Query Cache	123
Query DB	136
Query NoSQL DB	140
Read from DB	143
Read From File	152
Read from Hadoop Sequence File	170
Read From Hive File	175
Archivo Read From HL7	178
Read from NoSQL DB	190
Read from SAP	194
Leer de la hoja de cálculo	200
Read From Variable Format File	202
Read from XML	217
SQL Command	225
Transposer	235
Unique ID Generator	239
Write to Cache	248
Write to DB	250
Write to File	255
Write to Hadoop Sequence File	274
Escribir en archivo Hive	278
Write to NoSQL DB	285
Write to Spreadsheet	288

Write to Variable Format File	292
Write to XML	302
Patrones de fecha y número	312

Invocar procedimiento almacenado

Call Stored Procedure es una etapa de origen que ejecuta un procedimiento almacenado en una base de datos y devuelve los resultados de la invocación del procedimiento almacenado como entrada al flujo de datos. Use Call Stored Procedure cuando desee obtener datos desde una base de datos utilizando el procedimiento almacenado de la base de datos en lugar de una consulta a una tabla o vista.

Nota: Si desea leer los datos en un flujo de datos directamente desde una tabla o vista, utilice la etapa Read From DB.

Es posible que desee usar Invocar procedimiento almacenado para leer los datos en un flujo de datos, si tiene una lógica de negocio incorporada en el procedimiento almacenado y desee usar dicha lógica en su entorno Spectrum™ Technology Platform. Por ejemplo, muchos sistemas operativos no utilizan revisiones de integridad referenciales en la base de datos de grandes tablas con actualización constante, debido a la reducción del rendimiento que dichas revisiones provocarían. De modo que, para mantener la integridad referencial, se pueden crear procedimientos almacenados y usarlos para todas las actualizaciones en el sistema.

Los procedimientos almacenados también se pueden utilizar para simplificar la gestión del entorno Spectrum™ Technology Platform. Por ejemplo, si posee cientos de procesos ETL donde todos leen los mismos datos, es posible que desee poner la consulta en un procedimiento almacenado, para que esté en un solo lugar. Esto facilita el mantenimiento, dado que solo debe modificar un único procedimiento almacenado, en lugar de cientos de procedimientos diferentes.

Nombre de la opción	Descripción						
Conexión	<p>Seleccione la conexión de base de datos que quiere utilizar. Las opciones pueden variar según las conexiones definidas en el Administrador de conexión de base de datos de Management Console. Si necesita hacer una conexión de base de datos nueva o modificar o eliminar una conexión de base de datos existente, haga clic en Administrar.</p> <p>Si va a agregar o modificar una conexión de base de datos, complete los siguientes campos:</p> <table border="0"> <tr> <td>Nombre de conexión</td> <td>Ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.</td> </tr> <tr> <td>Controlador de base de datos</td> <td>Seleccione el tipo de base de datos correspondiente.</td> </tr> <tr> <td>Opciones de conexión</td> <td>Especifique el host, el puerto, la instancia, el nombre de usuario y la contraseña que se utilizarán para conectarse con la base de datos.</td> </tr> </table>	Nombre de conexión	Ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.	Controlador de base de datos	Seleccione el tipo de base de datos correspondiente.	Opciones de conexión	Especifique el host, el puerto, la instancia, el nombre de usuario y la contraseña que se utilizarán para conectarse con la base de datos.
Nombre de conexión	Ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.						
Controlador de base de datos	Seleccione el tipo de base de datos correspondiente.						
Opciones de conexión	Especifique el host, el puerto, la instancia, el nombre de usuario y la contraseña que se utilizarán para conectarse con la base de datos.						

Nombre de la opción	Descripción
Esquema	Especifica el esquema que contiene el procedimiento almacenado que desea invocar.
Procedimiento	Especifica el procedimiento almacenado que desea invocar.
Parámetros de procedimiento almacenado	<p>Esta tabla especifica los valores de los parámetros del procedimiento almacenado.</p> <p>Parámetros Esta columna muestra los parámetros definidos en el procedimiento almacenado.</p> <p>Campos de etapa Para los parámetros OUT, INOUT, y RETURN, esta columna muestra el nombre del campo del flujo de datos que incluirá los datos devueltos por el parámetro. Inicialmente, el nombre del campo es el mismo nombre el parámetro. Puede modificar el nombre del campo de la etapa, al hacer clic en el nombre del campo y escribir un nombre nuevo para los parámetros. Esta columna no se usa para los parámetros IN.</p> <p>Dirección Una de las siguientes:</p> <p>IN Es un parámetro de entrada. El valor que especifique para este parámetro se pasa al procedimiento almacenado como entrada.</p> <p>OUT Es un parámetro de salida. El procedimiento almacenado devuelve los datos a la etapa en este parámetro.</p> <p>INOUT El parámetro se puede usar como parámetro de entrada para pasar un valor al procedimiento almacenado, y como parámetro de salida para recibir los datos devueltos por el procedimiento almacenado.</p> <p>RETURN El parámetro contiene un código de devolución del procedimiento almacenado.</p> <p>Tipos Esta columna muestra el tipo de dato del valor del parámetro. Si el tipo de dato no es compatible con Spectrum™ Technology Platform, el tipo será "No compatible" y el procedimiento almacenado no se ejecutará satisfactoriamente.</p> <p>Valor En esta columna, ingrese el valor que desea establecer para el parámetro. Esta columna está deshabilitada para los parámetros OUT.</p>

Nombre de la opción	Descripción
Campos del conjunto de resultados	<p>Esta tabla especifica los campos de flujo de datos que se van a usar para los datos devueltos por el procedimiento almacenado.</p> <p>Tablas de base de datos Esta columna muestra las tablas desde las cuales el procedimiento almacenado devuelve los datos.</p> <p>Campos de base de datos Esta columna muestra los campos desde los cuales el procedimiento almacenado devuelve los datos.</p> <p>Campos de etapa Esta columna muestra el nombre del campo del flujo de datos que incluirá los datos desde el campo de base de datos.</p> <p>Tipos Esta columna muestra el tipo de dato del campo. Si el tipo de dato no es compatible con Spectrum™ Technology Platform, el tipo será "No compatible".</p> <p>Incluir Marque el cuadro en esta columna, para incluir el campo en el flujo de datos. Si el cuadro no se marca, el campo no se usará en el flujo de datos.</p>
Obtener campos	Haga clic en este botón para completar la tabla Campos del conjunto de resultados con el esquema del conjunto de resultados devueltos por el procedimiento almacenado. Esto ejecutará el procedimiento almacenado y obtendrá el esquema del conjunto de resultados.
Agregar	Haga clic en este botón para agregar un campo de conjunto de resultados manualmente.
Quitar	Haga clic en este botón para quitar un campo de conjunto de resultados de la lista de campos disponibles.

DB Change Data Reader

La etapa **DB Change Data Reader** le permite seleccionar las columnas que se incluirán en el flujo de trabajo actual, donde las columnas tienen la función Change Data Capture habilitada.

En la etapa puede crear un recurso de Change Data Capture (CDC), que es la tabla de origen de datos necesaria. Las columnas del recurso CDC en las que está habilitada la función Change Data Capture se reflejan mediante casillas de verificación.

Captura de datos modificados

La función Change Data Capture permite capturar todos los cambios realizados en una columna. Para cada columna seleccionada, se capturan todas las inserciones, actualizaciones y eliminaciones.

Bases de datos compatibles

Actualmente, Spectrum™ Technology Platform admite la función Change Data Capture (CDC) solo para bases de datos MS SQL y Oracle.

MS SQL Para orígenes de datos MS SQL, la función Change Data Capture se puede habilitar o deshabilitar para columnas de tablas desde el servidor back-end. Consulte [aquí](#) para ver los pasos necesarios.

Nota: La función Change Data Capture de Spectrum™ no es compatible con la edición Express de SQL Server.

Oracle Para los orígenes de datos Oracle, Spectrum™ registra los cambios en los datos en las columnas de la tabla por medio de la utilidad LogMiner de Oracle. CDC no se puede habilitar o deshabilitar en columnas de tablas de un origen de datos Oracle a través de Spectrum™ Technology Platform.

Los cambios en los datos derivados de consultas `insert`, `update` y `delete` se registran y capturan a partir de una fecha y hora de inicio ingresadas hasta la fecha y hora actuales. Esta hora de inicio ingresada se aplica durante la ejecución de la primera consulta con CDC activado.

En las siguientes ejecuciones en la misma conexión Oracle, los cambios en los datos se capturan a partir de la hora de la última ejecución hasta la hora actual.

Nota: Este es un proceso incremental. En la primera captura, se capturan los cambios desde la fecha y hora de inicio ingresadas hasta la fecha y hora actual. En capturas posteriores, los cambios se capturan después de la fecha y hora de finalización de la captura previa a la fecha y hora actual.

Adición de un recurso de CDC

Nota: Para utilizar la función Change Data Capture de Spectrum™, asegúrese de que SQL Server Agent se esté ejecutando en el servidor MS SQL.

1. Abra la ventana emergente **Administración de Captura de datos modificados** a través de una de las siguientes dos vías:
 - Navegue hasta **Herramientas > Administración de Captura de datos modificados**.
 - Agregue la etapa **DB Change Data Reader** a una tarea, abra la configuración de la etapa y haga clic en **Administrar**.
2. Haga clic en **Agregar**.
3. Ingrese un **Nombre** para el recurso de CDC.

4. En el campo **Conexión** seleccione la conexión de base de datos SQL que desea utilizar. Para realizar una nueva conexión de base de datos, haga clic en **Manage** (Administrar). Para obtener más información sobre la creación de conexiones de bases de datos, consulte [Administrador de conexión de base de datos](#).
5. En el campo **Tabla/Vista** especifique la tabla cuyas columnas se van a incluir en el flujo de trabajo. Haga clic en el botón de examinar ([...]) para navegar hasta la tabla o la vista que desea utilizar.
La siguiente cuadrícula muestra todas las columnas de la tabla seleccionada, junto con los tipos de datos de cada columna.
6. Si selecciona una conexión Oracle en el campo **Conexión**, el campo **Fecha de inicio** queda disponible.

El campo se completa con el valor predeterminado de la fecha actual y la hora 00:00 h. Puede ingresar la fecha de inicio y hora que prefiera.

La fecha y hora de finalización se consideran la fecha y hora actuales.

Atención: Debe tener permisos de *ejecución* para la utilidad Oracle LogMiner a fin de poder usar la función CDC en una conexión Oracle. Para obtener más información, consulte [Configuraciones de Oracle LogMiner](#) en la página 319.

Nota: El campo **Fecha de inicio** no está disponible si se selecciona una conexión MS SQL en el campo **Conexión**.

7. Haga clic en **Aceptar**.

La tabla Recurso CDC creada ahora está lista para su uso en la etapa **DB Change Data Reader**, cuyas columnas se pueden incluir o excluir del flujo de trabajo.

La etapa muestra las columnas de la tabla en las que se activó la función CDC.

Editar un recurso de CDC

1. Abra la ventana emergente **Administración de captura de datos modificados** a través de una de las siguientes vías:
 - Navegue hasta **Herramientas > Administración de Captura de datos modificados**.
 - Agregue la etapa **DB Change Data Reader** a una tarea, abra la configuración de la etapa y haga clic en **Administrar**.
2. Seleccione el recurso de CDC que necesita modificar.
3. Haga clic en **Editar**.
4. Modifique los detalles del recurso de CDC agregado según sea necesario.
5. Haga clic en **Aceptar**.

Eliminar un recurso de CDC

1. Abra la ventana emergente **Administración de captura de datos modificados** a través de una de las siguientes vías:
 - Navegue hasta **Herramientas > Administración de Captura de datos modificados**.
 - Agregue la etapa **DB Change Data Reader** a una tarea, abra la configuración de la etapa y haga clic en **Administrar**.
2. Seleccione el recurso de CDC que quiere eliminar.
3. Haga clic en **Eliminar**.

Seleccionar las opciones de Change Data Reader

Las **Opciones de DB Change Data Reader** reflejan las columnas de la tabla del recurso CDC seleccionado, para el cual se habilita la función CDC.

Puede seleccionar qué columnas incluir o excluir en el flujo de trabajo actual.

1. En un trabajo, agregue la etapa **DB Change Data Reader**.
2. Abra las **Opciones de DB Change Data Reader** haciendo doble clic en el ícono de etapa.
3. Seleccione el recurso de CDC del menú desplegable **Seleccionar un recurso..**

Puede modificar o agregar un recurso de CDC haciendo clic en **Administrar**.

La siguiente cuadrícula muestra todas las columnas de tabla con sus tipos de datos. También muestra si una columna en particular se incluye en el flujo de trabajo y si la columna está seleccionada para la función Captura de datos modificados.

4. A través de las casillas de verificación debajo de la columna **Incluir** de la cuadrícula, selecciona las columnas de tabla que quiere incluir en el flujo de trabajo.
5. Las casillas de verificación debajo de la columna **CDC habilitado** de la cuadrícula reflejan las columnas de la tabla para las que la función de CDC está habilitada.

Las casillas de verificación de solo lectura de **CDC habilitado** están marcadas para las columnas que tienen la función de CDC habilitada.

Nota: En relación con los orígenes de datos MS SQL, consulte [aquí](#) para obtener instrucciones sobre cómo habilitar o deshabilitar la función Change Data Capture en columnas de tabla en particular del servidor back-end.

6. Haga clic en **Aceptar**.

Los datos de las columnas de tablas, seleccionadas para Captura de datos modificados, se recopilan y se guardan.

DB Loader

La etapa **DB Loader** le permite acceder a los datos y cargarlos desde y en las bases de datos configuradas en el módulo Spectrum™ Data Integration. Esta etapa proporciona una interfaz para una utilidad de carga de datos de alta velocidad. Actualmente la plataforma Spectrum™ Data Integration admite **Oracle Loader**, **DB2 Loader**, **PostgreSQL Loader** y **Teradata Loader**.

Oracle Loader

Oracle Loader permite cargar datos en cualquier base de datos Oracle configurada en la plataforma Spectrum™ Data Integration.

Nota: El cliente Oracle debe estar instalado con la configuración del administrador antes de utilizar Oracle Loader.

Nombre de la opción	Descripción
Conexión	<p>Seleccione la conexión de base de datos que quiere utilizar. Las opciones pueden variar según las conexiones definidas en el Administrador de conexión de base de datos de Management Console. Si necesita hacer una conexión de base de datos nueva o modificar o eliminar una conexión de base de datos existente, haga clic en Administrar.</p> <p>Si va a agregar o modificar una conexión de base de datos, complete los siguientes campos:</p> <p>Nombre de conexión Ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.</p> <p>Controlador de base de datos Seleccione el tipo de base de datos correspondiente.</p> <p>Opciones de conexión Especifique el host, el puerto, la instancia, el nombre de usuario y la contraseña que se utilizarán para conectarse con la base de datos.</p>
Tabla/Vista	<p>Después de seleccionar una conexión, especifique la tabla o la vista donde se escribirá. Haga clic en el botón de examinar ([...]) para navegar hasta la tabla o la vista que quiere utilizar o haga clic en Create Table (Crear tabla) para crear una nueva tabla en la base de datos.</p> <p>Nota: Si escribe en una base de datos SQL, no puede escribir en vistas que hagan referencia a más de una tabla. Ello se debe a una limitación en el servidor SQL.</p>

Nombre de la opción	Descripción
Receptor	Especifique un nombre de variable que contiene la información de dirección y conexión necesaria para establecer una conexión con la base de datos de Oracle. Por ejemplo, "XE". Esta variable está presente en <code>tnsnames.ora</code> , un archivo que contiene parámetros de configuración de red del lado del cliente.
Campos de etapa	Esta columna lista los nombres de campo que se usan en el flujo de datos. No puede modificar estos nombres de campos.
Tipos	Esta columna muestra el tipo de dato de cada campo.

Ficha Tiempo de ejecución

Nombre de la opción	Descripción
Método de carga	<p>Especifica la manera en que desea escribir datos en la base de datos de Oracle.</p> <p>Adjuntar Incorpora los datos en la tabla de destino sin borrar los datos existentes en la tabla.</p> <p>Insertar Carga los datos en la base de datos. La tabla debe estar vacía antes de la carga. No funciona en instancias de ejecución múltiples.</p> <p>Truncar e insertar Elimina las filas existentes, si las hay, y luego carga los datos de entrada en la tabla. Esto no funciona en múltiples instancias de ejecución.</p>
Usar cargador de ruta directa	Seleccione esta opción para cargar datos directamente en la base de datos de Oracle pasando por alto gran parte del procesamiento de datos que normalmente se realiza.
No recuperable	Esta casilla de verificación se activa cuando selecciona Usar cargador de ruta directa . Seleccione esta opción si no desea escribir registros Rehacer en la base de datos. Para obtener más información sobre registros Rehacer, consulte http://docs.oracle.com/cd/B28359_01/server.111/b28310/onlineredo001.htm#ADMIN11302
Carpeta de archivo de registro	Especifica la ruta a la carpeta. Haga clic en el botón de los puntos suspensivos (...) para buscar la carpeta que desea. El archivo de registro contiene un registro de las actividades de esta etapa durante una sesión de carga.
Carpeta de archivo malo	Especifica la ruta a la carpeta. Haga clic en el botón de los puntos suspensivos (...) para buscar la carpeta que desea. El archivo malo contiene una lista de registros que la etapa no pudo cargar en la base de datos.

Nombre de la opción	Descripción
Máximo de errores permitidos	Especifique el número máximo de errores que se permiten antes de detener la operación de carga. Para detener la operación de carga en el primer error, configure el valor en 0. Se permite un máximo de 32767 errores.

Nota: Puede lograr mejoras de rendimiento importantes mediante el uso de varias instancias de tiempo de ejecución de esta operación. Para llevar a cabo esto, haga clic en el botón **Tiempo de ejecución** e ingrese el valor obligatorio en el campo **Instancias de tiempo de ejecución**.

DB2 Loader

DB2 Loader permite cargar datos en cualquier base de datos DB2 configurada en la plataforma Spectrum™ Data Integration.

Nota: El cliente en tiempo de ejecución DB2 debe estar instalado con la configuración del administrador antes de utilizar DB2 Loader.

Nombre de la opción	Descripción
Conexión	<p>Seleccione la conexión de base de datos que quiere utilizar. Las opciones pueden variar según las conexiones definidas en el Administrador de conexión de base de datos de Management Console. Si necesita hacer una conexión de base de datos nueva o modificar o eliminar una conexión de base de datos existente, haga clic en Administrar.</p> <p>Si va a agregar o modificar una conexión de base de datos, complete los siguientes campos:</p> <p>Nombre de conexión Ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.</p> <p>Controlador de base de datos Seleccione el tipo de base de datos correspondiente.</p> <p>Opciones de conexión Especifique el host, el puerto, la instancia, el nombre de usuario y la contraseña que se utilizarán para conectarse con la base de datos.</p>
Tabla/Vista	Después de seleccionar una conexión, especifique la tabla o la vista donde se escribirá. Haga clic en el botón de examinar ([...]) para ir hasta la tabla o la vista que quiere utilizar, o haga clic en Crear tabla para crear una nueva tabla en la base de datos.

Nombre de la opción	Descripción
Base de datos/Alias	<p>Esta es una variable que cataloga la base de datos y el servidor DB2.</p> <p>Para catalogar el servidor DB2 Utilice el procesador de la línea de comandos DB2 o el equipo servidor de Spectrum e ingrese el comando:</p> <pre>CATALOG TCPIP NODE <nodename> REMOTE <hostname> SERVER <port></pre> <p>Donde:</p> <p>nodename: nombre de la conexión</p> <p>hostname: nombre de TCP/IP del equipo servidor DB2</p> <p>port: puerto del servidor</p> <p>Para catalogar la base de datos Use el comando:</p> <pre>CATALOG DATABASE <databasename> AS <local_database_alias> AT NODE <nodename></pre> <p>Donde:</p> <p>databasename: nombre de la base de datos en el servidor DB2</p> <p>local_database_alias: nombre local dado a la base de datos al conectarse desde el equipo servidor</p> <p>nodename: nombre utilizado en el comando CATALOG TCP/IP anterior</p>
Campos de etapa	Esta columna lista los nombres de campo que se usan en el flujo de datos. No puede modificar estos nombres de campos.
Tipos	Esta columna muestra el tipo de dato de cada campo.

Ficha Tiempo de ejecución

Nombre de la opción	Descripción
Método de carga	<p>Indica el modo de escritura de datos en una tabla DB2.</p> <p>Insertar Inserta los datos cargados en la tabla, mientras los datos de tabla existentes permanecen inalterados.</p> <p>Reemplazar Inserta los datos cargados en la tabla luego de eliminar los datos existentes de esta.</p> <p>El esquema de tabla y las definiciones del índice permanecen inalterados.</p> <p>Reiniciar Reinicia la carga de datos, en caso de que el intento de carga previo se vea interrumpido.</p>
No recuperable	<p>Indica si esta transacción de carga no es recuperable.</p> <p>Si selecciona esta opción, la transacción de carga se marca como no recuperable. Los espacios de tabla no se ponen en estado Backup Pending luego de la carga, ni se realiza una copia de los datos cargados durante la carga. Por lo tanto, una transacción no recuperable no se puede recuperar en caso de errores en la carga de datos, incluso si se intenta usar la utilidad <code>rollforward</code> posteriormente.</p> <p>Si selecciona esta opción, no podrá realizar una recuperación desde la transacción incluso si usa la utilidad DB2 <code>rollforward</code>, ya que dicha utilidad ignora esos tipos de transacciones no recuperables, lo que marcará la tabla como "no válida". Además, las transacciones posteriores ejecutadas contra la tabla también son ignoradas por <code>rollforward</code>.</p> <p>Para restaurar una tabla con transacciones no recuperables, debe usar ya sea una copia de seguridad de nivel de espacio de tabla o una copia de seguridad completa realizada en un punto de confirmación posterior a la carga no recuperable.</p> <p>Nota: No seleccione esta opción si los datos contienen columnas Datalink con el atributo File Link Control.</p>
CPU	La cantidad de procesos paralelos que la utilidad de carga puede generar y sostener para cargar, analizar y formatear los registros, mientras se generan objetos de tabla en cada partición de la base de datos.
Disco	La cantidad de procesos paralelos que la utilidad de carga puede generar y sostener para escribir datos en los contenedores de espacio de tabla.

Nombre de la opción	Descripción
Modo de indexación	<p>Indica el modo de manejo de índices de la utilidad de carga.</p> <p>Autoselect La utilidad de carga decide si se debe aplicar o no un modo Reconstruir o Incremental, según la cantidad de datos y la profundidad del árbol del índice.</p> <p>Reconstruir Se reconstruyen todos los índices.</p> <p>Incremental Se agregan nuevos datos a los índices existentes.</p> <p>Este modo se puede aplicar solo si el objeto de índice es válido y accesible durante el inicio de la operación de carga.</p> <p>Nota: La indexación incremental no es compatible cuando TODAS las siguientes condiciones son verdaderas:</p> <ol style="list-style-type: none"> 1. La opción Copia de carga está especificada (<code>logretain</code> o <code>userexit</code> está activado). 2. La tabla reside en un espacio de tabla DMS. 3. El objeto de índice reside en un espacio de tabla que se comparte con otros objetos de tabla, que pertenecen a la tabla que se está cargando. <p>Para superar esta limitación, coloque los índices en espacios de tabla separados.</p> <p>Diferido La utilidad de carga no intenta crear un índice. Los índices existentes tienen activada la opción de actualización.</p> <p>Nota: La construcción de índices requiere más tiempo en modo Diferido que en modo Reconstruir. Por lo tanto, mientras realiza múltiples operaciones de carga, permita que la última operación de carga realice una reconstrucción de los índices en lugar de reconstruirlos en el primer acceso de una operación no destinada a carga.</p> <p>Nota: Este modo solo se admite para tablas con índices no exclusivos.</p>
Análisis rápido	<p>Indica si se debe omitir la validación sintáctica de los valores de columna, para mejorar el rendimiento.</p> <p>Si se selecciona, se ignorarán los errores sintácticos en los datos en pro de un rendimiento optimizado.</p> <p>Por ejemplo, de encontrarse un valor de cadena de caracteres <code>12wxvg56</code> en un campo asignado a una columna de valores enteros en un archivo ASCII, la utilidad de carga normalmente generaría un error de sintaxis. Pero si se selecciona la opción Análisis rápido, el error de sintaxis se ignora y se carga un número arbitrario en el campo de valores enteros.</p> <p>Nota: Asegúrese de usar esta opción solo con datos correctos y limpios.</p>

Nombre de la opción	Descripción
Nombre de esquema	El esquema donde se almacenan las tablas de excepción.
Nombre de tabla	La tabla de excepción en la cual se copian las filas en las cuales se detectaron errores durante la carga.
Carpeta de archivo de registro	<p>La ruta del directorio donde se almacenan los archivos de registro.</p> <p>Un archivo de registro contiene una lista de transacciones de carga de base de datos ejecutadas por una etapa DB Loader en una sesión de carga.</p> <p>Haga clic en el botón de elipsis (...) para especificar el directorio preferido para los archivos de registro.</p>
Carpeta de archivo malo	<p>La ruta del directorio en el servidor DB2 donde se almacenan los archivos de elementos defectuosos.</p> <p>El archivo de elementos defectuosos contiene una lista de registros que la etapa DB Loader no pudo cargar en la base de datos.</p> <p>Haga clic en el botón de elipsis (...) para especificar el directorio preferido para los archivos de elementos defectuosos.</p>
Máximo de errores permitidos	<p>La cantidad máxima de errores permitidos antes de abortar una operación de carga.</p> <p>Para abortar una operación de carga en el momento de encontrar el primer error, establezca el valor de este campo en 0.</p> <p>Nota: Se permite un máximo de 32767 errores.</p>
Paralelismo	<p>Una base de datos DB2 se puede dividir en múltiples particiones mediante la clonación del entorno en diferentes nodos físicos.</p> <p>Las solicitudes de base de datos independientes para captura y actualización de datos se dividen automáticamente entre las diferentes particiones y se ejecutan en paralelo para optimizar el rendimiento.</p>
Manejo de excepciones	<p>Una base de datos DB2 le permite registrar los errores y las excepciones detectados durante la ejecución de consultas y procedimientos, así como manejarlos correctamente.</p> <p>Para esto, una base de datos DB2 proporciona tablas y esquemas de excepción específicos donde se almacenan el origen y los seguimientos de registro de cada excepción de base de datos.</p>

PostgreSQL Loader

PostgreSQL Loader permite cargar datos en cualquier base de datos PostgreSQL configurada en la plataforma Spectrum™ Data Integration.

Nombre de la opción	Descripción
Conexión	<p>Seleccione la conexión de base de datos que quiere utilizar. Las opciones pueden variar según las conexiones definidas en el Administrador de conexión de base de datos de Management Console. Si necesita hacer una conexión de base de datos nueva o modificar o eliminar una conexión de base de datos existente, haga clic en Administrar.</p> <p>Si va a agregar o modificar una conexión de base de datos, complete los siguientes campos:</p> <p>Nombre de conexión Ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.</p> <p>Controlador de base de datos Seleccione el tipo de base de datos correspondiente.</p> <p>Opciones de conexión Especifique el host, el puerto, la instancia, el nombre de usuario y la contraseña que se utilizarán para conectarse con la base de datos.</p>
Tabla/Vista	Haga clic en el botón de examinar ([...]) para navegar hasta la tabla o la vista que desea utilizar.
Campos de base de datos	Esta columna enumera el nombre de campo en la base de datos. No puede modificar estos nombres de campos.
Campos de etapa	Esta columna lista los nombres de campo que se usan en el flujo de datos. No puede modificar estos nombres de campos.
Tipos	Esta columna muestra el tipo de dato de cada campo.

Ficha Tiempo de ejecución

Nombre de la opción	Descripción
Método de carga	<p>Especifica el método de escritura de datos en las tablas de la base de datos PostgreSQL.</p> <ul style="list-style-type: none"> • Seleccione Insertar para escribir los datos en una tabla vacía o unirla a una tabla de datos existente. • Seleccione Truncar e insertar para truncar los datos antes de cargarlos en la tabla de la base de datos.

Nota: Puede lograr mejoras de rendimiento importantes mediante el uso de varias instancias de tiempo de ejecución de esta operación. Para llevar a cabo esto, haga clic en el botón **Tiempo de ejecución** e ingrese el valor obligatorio en el campo **Instancias de tiempo de ejecución**.

Teradata Loader

Teradata Loader permite cargar datos en cualquier base de datos Teradata configurada en la plataforma Spectrum™ Data Integration.

Nota: El cargador es compatible únicamente en sistemas Windows.

Nombre de la opción	Descripción						
Conexión	<p>Seleccione la conexión de base de datos que quiere utilizar. Las opciones pueden variar según las conexiones definidas en el Administrador de conexión de base de datos de Management Console. Si necesita hacer una conexión de base de datos nueva o modificar o eliminar una conexión de base de datos existente, haga clic en Administrar.</p> <p>Si va a agregar o modificar una conexión de base de datos, complete los siguientes campos:</p> <table border="0"> <tr> <td>Nombre de conexión</td> <td>Ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.</td> </tr> <tr> <td>Controlador de base de datos</td> <td>Seleccione el tipo de base de datos correspondiente.</td> </tr> <tr> <td>Opciones de conexión</td> <td>Especifique el host, el puerto, la instancia, el nombre de usuario y la contraseña que se utilizarán para conectarse con la base de datos.</td> </tr> </table>	Nombre de conexión	Ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.	Controlador de base de datos	Seleccione el tipo de base de datos correspondiente.	Opciones de conexión	Especifique el host, el puerto, la instancia, el nombre de usuario y la contraseña que se utilizarán para conectarse con la base de datos.
Nombre de conexión	Ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.						
Controlador de base de datos	Seleccione el tipo de base de datos correspondiente.						
Opciones de conexión	Especifique el host, el puerto, la instancia, el nombre de usuario y la contraseña que se utilizarán para conectarse con la base de datos.						

Nombre de la opción	Descripción
Tabla/Vista	Haga clic en el botón de examinar ([...]) para navegar hasta la tabla o la vista que desea utilizar.
Campos de base de datos	En esta columna, se enumeran los nombres de campo en la base de datos. No puede modificar estos nombres de campos.
Campos de etapa	Esta columna lista los nombres de campo que se usan en el flujo de datos. No puede modificar estos nombres de campos.
Tipos	Esta columna muestra el tipo de dato de cada campo.

Ficha Tiempo de ejecución

Nombre de la opción	Descripción
Método de carga	<p>Especifica el método de escritura de datos en las tablas de la base de datos Teradata.</p> <ul style="list-style-type: none"> • Seleccione Insertar para escribir datos en una tabla vacía. • Seleccione Anexar para escribir datos en una tabla de datos existente. • Seleccione Truncar e insertar para truncar los datos antes de escribirlos en la tabla de la base de datos.
Carpeta de archivo de registro	Seleccione la ubicación de la carpeta para guardar los archivos de registro del proceso de carga.
Generar archivo malo	Marque esta casilla de verificación para generar un registro de los registros que presentaron error al cargarse.
Carpeta de archivo malo	Seleccione la ubicación de la carpeta para guardar los archivos de registro de los registros con error. El registro entrega detalles, como códigos de error y nombres de campo de error de los registros con error.
Máximo de errores permitidos	Especifique el límite de errores permitidos para la sesión de carga. Si la cantidad de errores excede este valor, el proceso de carga se detiene.

Nota: No se admiten múltiples instancias de ejecución para Teradata Loader.

Field Parser

La etapa **Field Parser** permite extraer campos desde datos XML y delimitados presentes en la columna de entrada especificada. Para configurar las opciones de **Field Parser**, realice las siguientes tareas.

1. En el campo **Origen**, seleccione la columna que contiene los datos XML o delimitados que se desean analizar.

Nota: La lista desplegable muestra todas las columnas de entrada de cadena.

2. Seleccione el **Formato** XML o Delimitado en función de los tipos de datos que desea analizar y seleccione las opciones que se describen a continuación, según corresponda.

Opciones de Field Parser para datos XML

Nombre de la opción	Descripción
Server name (Nombre de servidor)	Indica si el archivo seleccionado para inferir el esquema se encuentra ubicado en la computadora que ejecuta Enterprise Designer o en el servidor. Si selecciona un archivo en la computadora local, el nombre del servidor será Mi computadora. Si selecciona un archivo que se encuentra en el servidor, el nombre del servidor será Spectrum™ Technology Platform.
Archivo de esquema	<p>Especifica la ruta a un archivo de esquema XSD. Haga clic en el botón de puntos suspensivos (...) para navegar a la ubicación del archivo. El archivo de esquema puede encontrarse en el servidor o en el sistema local.</p> <p>De forma alternativa, también puede especificar un archivo XML en vez de un archivo XSD. Si especifica un archivo XML, el esquema se inferirá a partir de la estructura del archivo XML. Usar un archivo XML en vez de un XSD tiene las siguientes limitaciones:</p> <ul style="list-style-type: none"> • El archivo XML no puede pesar más de 1 MB. Si el archivo XML pesa más de 1 MB, intente eliminar algunos de los datos sin afectar la estructura del archivo. • El archivo de datos no se validará con el esquema inferido. <p>Nota: Si el servidor de Spectrum™ Technology Platform está ejecutándose en Unix o Linux, recuerde que estas plataformas distinguen las mayúsculas de las minúsculas en los nombres de los archivos y las rutas.</p>

Nombre de la opción	Descripción
Campos de salida	<p>En esta sección se muestran detalles del esquema seleccionado. Incluye el elemento raíz seguido de los elementos secundarios, junto con sus atributos.</p> <p>De forma predeterminada, todos los nodos del esquema permanecen seleccionados. Sin embargo, puede desmarcar la casilla de verificación de los nodos que no desea que pasen a la siguiente etapa.</p> <ul style="list-style-type: none"> • Buscar nodo: escriba el nombre del nodo al cual desea navegar en el árbol de esquemas. El nodo ingresado se resalta en el panel de vista previa bajo el campo. • XPath: haga clic en cualquier lugar de este campo para ver la ruta XML (XPath) de los elementos y atributos del nodo resaltado en el árbol de esquema. Para ver todas las XPaths anteriores que ha visto, haga clic en la flecha hacia abajo en el extremo inferior del campo. <p>Nota: XPath es un lenguaje para buscar información en un documento XML. Para obtener más detalles, consulte www.w3schools.com/xpath/.</p>

Opciones de Field Parser para datos delimitados

Nombre de la opción	Descripción
Separador de campo	<p>En la lista desplegable, seleccione el separador de campo utilizado en la columna delimitada que se desea analizar.</p> <p>Si en la columna delimitada se utiliza un carácter distinto como separador de campo, haga clic en el botón de los puntos suspensivos para seleccionar otro carácter como separador.</p>
Text qualifier (Calificador de texto)	<p>En la lista desplegable, seleccione el calificador de texto utilizado en la columna delimitada que se desea analizar.</p> <p>Nota: Los calificadores de texto corresponden al carácter que se utiliza para demarcar los valores de texto en los datos delimitado.</p> <p>Si en la columna delimitada se utiliza un calificador de texto distinto, haga clic en el botón de los puntos suspensivos para seleccionar otro carácter como calificador de texto.</p>
Tipo de salida	<p>Elija si desea que los valores de salida analizados se muestren en forma de Lista (presentación jerárquica de los valores) o de Campos.</p> <p>Nota: En el caso de las listas como tipo de salida, puede agregar solo un campo de salida, mientras que la opción Campos le permite agregar varios campos en los que puede ver los valores segregados durante el análisis.</p>

Nombre de la opción	Descripción
---------------------	-------------

Campos de salida

En esta sección, puede agregar/modificar los diversos campos en los que desea obtener detalles sobre la columna delimitada que se desea segregar. También puede eliminar cualquiera de los campos de salida agregados.

Para agregar un nuevo campo para la presentación de los valores de salida analizados, haga clic en el botón **Agregar** y siga los pasos de la ventana emergente **Configuración de campo** que se muestra:

1. Ingrese el **Nombre** del campo.
2. En la lista desplegable **Tipo**, seleccione el tipo de datos para el campo que se está agregando. De acuerdo con el tipo seleccionado, se pueden definir otros campos. Por ejemplo, en el caso de la *fecha*, puede definir su formato como *M/d/yy*, *MMM d. yyyy* o *MMMM d. yyyy*. Para obtener detalles sobre los tipos de datos y definirlos, consulte [Definición de los campos en un archivo de entrada delimitado](#) en la página 158.

Nota: Si selecciona *Cadena* como el tipo de datos, se analizará cualquier tipo de datos delimitados. Sin embargo, también puede usar el tipo específico, según los datos que desea analizar en el campo.

3. En el campo **Posición**, ingrese la posición del tipo de datos (en el archivo de entrada) que se debe analizar en este campo. Por ejemplo, en el siguiente fragmento de archivo, si desea analizar los valores de fecha y hora en el campo que se está agregando, ingrese el valor de **Posición** como 3.

```
true;"02/02/2022";"10/2/92 5:05
AM";598985994665542.25634;1;
"Arjun";74785.155;5:05PM,1,Deepak,65152
false;"15/03/1923";"3/23/90 11:55
AM";3425699466554.2563;2;
"sharma";5.1;5:45AM,2,Arjun,365273
```

4. Haga clic en **Agregar campo y Cerrar**.

El campo agregado y los detalles respectivos se muestran en el cuadro.

Nota: Si desea eliminar los caracteres que ocupan demasiado espacio al principio y al final de la cadena de valores de un campo, seleccione la casilla **Recortar**.

Modificar: haga clic en este botón para cambiar los detalles de cualquiera de los campos de salida agregados.

Quitar: haga clic en este botón para eliminar cualquiera de los campos de salida agregados.

Tiempo de ejecución: use este botón para especificar varias instancias de tiempo de ejecución. Esto da como resultado una mejora significativa en el rendimiento.

Aceptar: haga clic en este botón para guardar todos los detalles ingresados en esta etapa.

Cancelar: haga clic en este botón para cancelar todas las actualizaciones que realizó.

Ayuda: haga clic en este botón para consultar el archivo de ayuda para esta etapa.

Field Combiner

La etapa **Field Combiner** combina campos de la etapa anterior en un flujo de datos para crear una cadena XML.

Opciones de Field Combiner

Nombre de la opción	Descripción
Nombre de la columna de salida	Especifique el nombre de la columna que tendrá los campos combinados como cadena XML.
Campos de salida	<p>Esta sección le ayuda a seleccionar los campos que se combinarán y realizar diversas acciones en dichos campos.</p> <p>Nota: La lista desplegable Aplicar espacio de nombre y los botones Elemento, Atributo y Eliminar se activan cuando selecciona uno o más elementos o atributos en el panel del esquema.</p> <ul style="list-style-type: none"> • Agregado rápido: haga clic en este botón para abrir la ventana emergente Agregar/Quitar campos. Esta ventana muestra una lista de todos los campos que provienen de la etapa anterior. Seleccione los campos que desea combinar. <p>Nota: Los campos seleccionados se muestran en el panel de esquemas bajo el campo Buscar nodo.</p> • Buscar nodo: escriba el nombre del nodo al cual desea navegar en el panel de esquemas. El nodo ingresado se resalta y su ruta XML aparece en el campo XPath bajo el panel. • Elemento: haga clic en este botón para convertir un atributo en un elemento en la cadena XML. • Atributo: haga clic en este botón para convertir un elemento en un atributo en la cadena XML. • Aplicar espacio de nombre: si desea especificar un espacio de nombre XML para un elemento o un atributo, selecciónelo aquí. <p>Nota: Puede crear espacios de nombre haciendo clic en el botón Espacio de nombre más arriba. Para obtener detalles sobre esto, consulte Definición de espacios de nombre.</p> • Eliminar: haga clic en este botón para eliminar los elementos o atributos que no desea en la cadena XML.

Nombre de la opción	Descripción
XPath	Haga clic en cualquier lugar de este campo para ver la ruta XML (XPath) del nodo resaltado en el panel del esquema. Para ver todas las XPaths anteriores que ha visto, haga clic en la flecha hacia abajo en el extremo inferior del campo. Nota: XPath es un lenguaje para buscar información en un documento XML. Para obtener más detalles, consulte www.w3schools.com/xpath/ .

- **Tiempo de ejecución:** use este botón para especificar varias instancias de tiempo de ejecución de Field Combiner. Esto da como resultado una mejora significativa en el rendimiento.
- **Aceptar:** haga clic en este botón para guardar todos los detalles ingresados en esta etapa.
- **Cancelar:** haga clic en este botón para cancelar todas las actualizaciones que realizó.
- **Ayuda:** haga clic en este botón para consultar el archivo de ayuda para esta etapa.

Definición de espacio de nombre

Los espacios de nombre le permiten tener nombres de atributo y elemento duplicados en su salida mediante la asignación de cada elemento o atributo a un espacio de nombre XML.

Para definir uno o más espacios de nombre:

1. En la pantalla **Opciones de Field Combiner**, haga clic en el botón **Espacio de nombre**. Aparecerá la ventana emergente **Detalles de espacio de nombre**.
2. En la columna **Prefijo**, ingrese el prefijo que se asociará con un elemento o atributo.
3. En la columna **Namespace**, especifique la dirección URL del espacio de nombres.
4. Repita este paso para definir todos los espacios de nombre que desee usar.

Field Selector

Field Selector permite escoger cuáles campos pasarán a la siguiente etapa en el flujo de datos. Puede usar Field Selector para eliminar los campos no deseados en un flujo de datos. Por ejemplo, si ha creado un nuevo campo al combinar los datos de dos campos y ya no necesita los dos campos de origen, puede utilizar Field Selector para conservar solamente el campo nuevo y extraer los dos campos de origen del flujo de datos.

Opciones

Opción	Descripción
Seleccione los campos que se enviarán a la siguiente etapa	Active la casilla situada junto a cada uno de los campos que desee enviar a la siguiente etapa en el flujo de datos. A fin de evitar que un campo se envíe a la etapa siguiente, desactive la casilla.
Seleccionar todo	Active esta casilla para seleccionar todos los campos en el flujo de datos. Desactive esta casilla para eliminar la selección de todos los campos.

Generate Time Dimension

Generate Time Dimension genera registros de fecha, uno para cada día del rango de fechas especificado. Entonces puede escribir dichos registros en una tabla de dimensión de tiempo en una base de datos usando la etapa Write to DB. La tabla de dimensión de tiempo puede utilizarse entonces para realizar cálculos exactos basados en un periodo de tiempo. Por ejemplo, ventas por trimestre, presupuesto invertido por trimestre e ingresos diarios son todos análisis que precisan de una dimensión de tiempo. Las tabla de dimensión de tiempo le permiten además dar cuenta en el análisis de los años fiscales o trimestres atípicos.

Ejemplo del uso de una tabla de dimensiones de tiempo

Las tablas de dimensiones de tiempo son necesarias para realizar cálculos precisos basados en tiempo porque a veces no es fácil extraer los datos de fecha necesarios de los registros. Por ejemplo, los siguientes registros se encuentran en una base de datos de ventas. Tenga en cuenta que hay diferencias de tiempo entre registros. Por ejemplo, no hay ningún registro para el día 1/4/2012.

Fecha	Producto	Monto
03/01/2012	Red Shirt	\$10.00
05/01/2012	Red Shirt	\$5.00
07/01/2012	Red Shirt	\$15.00

Si consulta estos registros y calcula el promedio de ventas por día, la respuesta sería \$10.00 ($\$30 / 3$ registros). Sin embargo, esto no es correcto porque los tres registros en realidad abarcan un período de cinco días. Si tiene una tabla de dimensiones de tiempo con un registro para cada día, puede combinar esta tabla con la tabla anterior para obtener:

Fecha	Producto	Monto
03/01/2012	Red Shirt	\$10.00
04/01/2012		
05/01/2012	Red Shirt	\$5.00
06/01/2012		
07/01/2012	Red Shirt	\$15.00

Al calcular el promedio de ventas por día utilizando estos registros, obtendrá la respuesta correcta: \$6.00 ($\$30 / 5$ días).

Además, podría tener en cuenta los atributos de tiempo arbitrarios, tales como días festivos, fines de semana y trimestres en el cálculo. Por ejemplo, si 1/6/2012 fue un día festivo y a usted solo le interesan las ventas promedio por día laboral, entonces la respuesta sería \$7.50.

Opciones

Generate Time Dimension tiene las siguientes opciones.

Opción	Descripción
Fecha de inicio	El primer día del intervalo de fechas para el que se va a generar registros de dimensión de tiempo.
Fecha de término	Seleccione esta opción si desea especificar una fecha de término determinada para la dimensión de tiempo. Se generará un registro para cada día que transcurre entre la fecha de inicio y la fecha que especifique aquí.

Opción	Descripción
Duración	Seleccione esta opción si desea que la dimensión de tiempo abarque un número determinado de días, meses o años. Se generará un registro por cada día de duración. Por ejemplo, si especifica una semana, se generarán siete registros, a partir de la fecha especificada en el campo Fecha de inicio .

Opción	Descripción
Atributo de tiempo	

Opción

Descripción

Opción	Descripción
	<p>Especifica el tipo de información de tiempo que desea incluir en la dimensión de tiempo. Cada atributo será un campo en los informes diarios. Una de las siguientes:</p>
Fecha	El día de la fecha con el formato <Fecha> <Mes> <Año>, con el nombre del mes en el idioma de la configuración regional del servidor. Por ejemplo, si el servidor se ejecuta en una configuración regional en inglés, la fecha tendría la siguiente apariencia: 30 October 2012.
Día del mes	Un número que representa el día del mes. Por ejemplo, 10 significa que es el día 10 del mes.
Día del año	Un número que representa el día del año. Por ejemplo, 304 significa que es el día 304 del año.
Es año bisiesto	Valor booleano que indica si el día está en un año bisiesto, en cuyo caso el valor sería <i>verdadero</i> , o no, en cuyo caso el valor sería <i>falso</i> .
Es día laborable	Valor booleano que indica si el día es día de semana (lunes a viernes), en cuyo caso el valor sería <i>verdadero</i> , o no, en cuyo caso el valor sería <i>falso</i> .
Es fin de semana	Valor booleano que indica si el día es día de fin de semana (sábado o domingo), en cuyo caso el valor sería <i>verdadero</i> , o no, en cuyo caso el valor sería <i>falso</i> .
Día juliano	Un número único del día El sistema de fecha juliana cuenta 1 Enero, 4713 a. C. como el día 1 y enumera secuencialmente todos los días desde entonces. Por ejemplo, 2456231 significa que el día es el 2.456.231 desde el 1 de enero, 4713 a. C.
Semana juliana	Un número único de la semana El sistema de fecha juliana cuenta la primera semana de 4713 a. C. como la semana 1 y enumera todas las semanas siguientes secuencialmente desde entonces. Por ejemplo, 350890 significa que el día está en la semana 350.890 desde la primera semana de 4713 a. C.
Año juliano	Un número único del año El sistema de fecha juliana cuenta 4713 a. C. como el año 1 y enumera secuencialmente todos los años desde entonces. Por ejemplo, 6725 significa que el día está dentro del año 2012 A. D.
Nombre del mes	El nombre del mes en español.
Número de mes	El número del mes del año. Por ejemplo, 3 significa que el día está en el tercer mes del año.
Trimestre	Un número que representa el trimestre del año. Por ejemplo, 1 significa que el día está en el primer trimestre del año.
Semana del año	Un número que representa la semana del año. Por ejemplo, 43 significa que el día está en la semana 43 del año.
Nombre de día	El nombre del día en español. Por ejemplo, lunes.

Opción	Descripción
	<p>laborable</p> <p>Número de día laborable Un número que representa el día de la semana, con 1 como lunes. Así, por ejemplo el martes es el día 2, el miércoles es el día 3 y así sucesivamente.</p> <p>Año El día de año según el calendario Gregoriano. Por ejemplo, 2012 significa que el día está dentro del año 2012.</p>
Campo	El nombre del campo que se crea en el flujo de datos para contener el atributo tiempo. Se proporciona un nombre predeterminado del campo, pero puede modificar el nombre del campo.
Tipo	El tipo de datos del campo. Generate Time Dimension selecciona automáticamente el tipo de datos adecuado para cada atributo tiempo.
Calendario	Especifica si se debe usar un calendario personalizado para el cálculo del tiempo predeterminado o el atributo calendario Gregoriano predeterminado. Para definir un calendario personalizado, haga clic en Calendario .

Creación de un calendario

Un calendario define características importantes del año, de una manera que es apropiada para el tipo de análisis que desee realizar. De manera predeterminada, el calendario es el gregoriano estándar. Sin embargo, el calendario gregoriano estándar no siempre es adecuado. Por ejemplo, si el ejercicio contable de su empresa comienza el 1 de junio en lugar del 1 de enero, puede definir el primer mes del año como junio y el atributo de tiempo "mes del año" tendrá junio como mes 1, julio como mes 2, etc., en lugar de enero como mes 1, febrero como mes 2, etc.

- Realice una de estas acciones:
 - Si está configurando la etapa Generate Time Dimension en Enterprise Designer, haga clic en **Calendarios**.
 - En Management Console, vaya a **Recursos > Calendarios**.
- Haga clic en **Agregar**.
- En el campo **Calendar name** asigne un nombre al calendario, que sea significativo para usted.
- En el campo **Start month** seleccione el primer mes del año.
Por ejemplo, si selecciona julio, julio es el mes 1, agosto es el mes 2, etc.

Nota: El cambio del mes de inicio afecta en la manera de calcular los valores en los siguientes campos: DayOfYear, MonthNumber, Quarter y WeekOfYear.

5. Haga clic en **Guardar**.

Salida

Generate Time Dimension produce los siguientes campos de salida.

Tabla 1: Salida de Generate Time Dimension

Nombre de campo	Descripción
Date	El día de la fecha con el formato <Fecha> <Mes> <Año>, con el nombre del mes en el idioma de la configuración regional del servidor. Por ejemplo, si el servidor se ejecuta en una configuración regional en inglés, la fecha tendría la siguiente apariencia: 30 October 2012.
DayOfMonth	Un número que representa el día del mes. Por ejemplo, 10 significa que es el día 10 del mes.
DayOfYear	Un número que representa el día del año. Por ejemplo, 304 significa que es el día 304 del año.
IsLeapYear	Valor booleano que indica si el día está en un año bisiesto, en cuyo caso el valor sería <i>verdadero</i> , o no, en cuyo caso el valor sería <i>falso</i> .
IsWeekday	Valor booleano que indica si el día es día de semana (lunes a viernes), en cuyo caso el valor sería <i>verdadero</i> , o no, en cuyo caso el valor sería <i>falso</i> .
IsWeekend	Valor booleano que indica si el día es día de fin de semana (sábado o domingo), en cuyo caso el valor sería <i>verdadero</i> , o no, en cuyo caso el valor sería <i>falso</i> .
JulianDay	Un número único del día El sistema de fecha juliana cuenta 1 Enero, 4713 a. C. como el día 1 y enumera secuencialmente todos los días desde entonces. Por ejemplo, 2456231 significa que el día es el 2.456.231 desde el 1 de enero, 4713 a. C.

Nombre de campo	Descripción
JulianWeek	Un número único de la semana El sistema de fecha juliana cuenta la primera semana de 4713 a. C. como la semana 1 y enumera todas las semanas siguientes secuencialmente desde entonces. Por ejemplo, 350890 significa que el día está en la semana 350.890 desde la primera semana de 4713 a. C.
JulianYear	Un número único del año El sistema de fecha juliana cuenta 4713 a. C. como el año 1 y enumera secuencialmente todos los años desde entonces. Por ejemplo, 6725 significa que el día está dentro del año 2012 A. D.
MonthName	El nombre del mes en español.
MonthNumber	El número del mes del año. Por ejemplo, 3 significa que el día está en el tercer mes del año.
Quarter	Un número que representa el trimestre del año. Por ejemplo, 1 significa que el día está en el primer trimestre del año.
WeekOfYear	Un número que representa la semana del año. Por ejemplo, 43 significa que el día está en la semana 43 del año.
WeekdayName	El nombre del día en español. Por ejemplo, lunes.
WeekdayNumber	Un número que representa el día de la semana, con 1 como lunes. Así, por ejemplo el martes es el día 2, el miércoles es el día 3 y así sucesivamente.
Year	El día de año según el calendario Gregoriano. Por ejemplo, 2012 significa que el día está dentro del año 2012.

Query Cache

Query Cache busca datos en una caché basándose en los valores de uno o más campos de flujos de datos y devuelve datos a partir de registros coincidentes en la caché, añadiendo los datos del registro de la caché al registro en el flujo de datos. La búsqueda de datos en una caché puede mejorar el rendimiento en comparación con la búsqueda de datos en una base de datos.

Hay dos tipos de caché: global y local.

Opciones de caché global

Una caché global abarca todo el sistema, se comparte y reside en la memoria. Elija una caché global si desea que esté disponible para varios flujos de datos, si los datos no cambian con frecuencia o si permanecen relativamente estáticos y si el almacenamiento no es limitado. Una caché global es estática; es decir, solo puede escribir en ella una vez. La caché no se puede actualizar una vez que se ha creado.

Una caché global se crea en la etapa Write to Cache. Antes de poder usar una caché global, debe llenar la caché con los datos que desea buscar. Para ello, cree un flujo de datos que contenga la etapa **Write to Cache**.

Nombre de la opción	Descripción
Tipo de caché	Seleccione la opción Global cache.
Nombre de caché	Especifica la caché que desea consultar. Para crear una caché, use la etapa Write to Cache.
Campos de caché	En esta columna aparecen los campos de la caché. No puede modificar estos nombres de campos.
Campos de etapa	Esta columna lista los nombres de campo que se usan en el flujo de datos. Si desea cambiar un nombre de campo, haga clic en el nombre de campo e ingrese un nuevo nombre.
Tipo	En esta columna aparece el tipo de datos de cada campo del flujo de datos.

Nombre de la opción	Descripción
Incluir	Active la casilla de esta columna para que la consulta devuelva el valor del campo de la caché. Desactive la casilla si no desea que la consulta devuelva el campo de la caché.

Nombre de la opción	Descripción
---------------------	-------------

Valor de error predeterminado	
-------------------------------	--

Nombre de la opción

Descripción

Especifica el valor que se va a mostrar en el campo de flujo de datos si se produce un error en la consulta. La lista desplegable muestra los valores válidos que corresponden a los tipos de datos del campo consultado. Por ejemplo, en el caso de un **número entero** la opción que se muestra es **-1**.

También puede ingresar un valor para este campo. Consulte la tabla que se muestra a continuación para ver una lista de los valores válidos de errores predeterminados correspondientes a varios tipos de datos.

Tipo de datos Valor de error predeterminado válido junto con un tipo de datos (en paréntesis)

Tipo de datos	Nulo	-1	1899-12-30	1899-12-30	12:00:00	Falso	Vacío
Fecha							
Integer		✓					
Long Entero largo		✓					
Float		✓					
Grande Decimal		✓					
Double		✓					
Cadena	✓	✓	✓	✓	✓	✓	✓
Hora					✓		
Fecha y							

Nombre de la opción	Descripción																																										
	<p>Tipo de datos</p> <p>Valor de error predeterminado válido junto con un tipo de datos (en paréntesis)</p>																																										
	<table border="1"> <thead> <tr> <th>Nulo</th> <th>-1</th> <th>1899-</th> <th>1899-</th> <th>12:00:00</th> <th>Falso</th> <th>Vacío</th> </tr> </thead> <tbody> <tr> <td></td> <td>(Entero)</td> <td>12-30</td> <td>12-30</td> <td>(Hora)</td> <td></td> <td></td> </tr> <tr> <td></td> <td></td> <td>12:00:00</td> <td>(Fecha)</td> <td></td> <td></td> <td></td> </tr> <tr> <td></td> <td></td> <td>(Fecha)</td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>hora</td> <td>✓</td> <td></td> <td></td> <td></td> <td></td> <td></td> </tr> <tr> <td>Boolean</td> <td></td> <td></td> <td></td> <td></td> <td>✓</td> <td></td> </tr> </tbody> </table>	Nulo	-1	1899-	1899-	12:00:00	Falso	Vacío		(Entero)	12-30	12-30	(Hora)					12:00:00	(Fecha)						(Fecha)					hora	✓						Boolean					✓	
Nulo	-1	1899-	1899-	12:00:00	Falso	Vacío																																					
	(Entero)	12-30	12-30	(Hora)																																							
		12:00:00	(Fecha)																																								
		(Fecha)																																									
hora	✓																																										
Boolean					✓																																						
Campo de clave	Especifica el campo de la caché que se utilizará como una clave de búsqueda. Si el valor del campo en la columna Input Field coincide con el valor en el campo clave de la caché, la consulta devuelve los datos de ese registro en la caché.																																										
Campo de entrada	Especifica el campo de flujo de datos cuyo valor se usará como clave. Si el valor de este campo coincide con el valor en el campo clave de la caché, la consulta devuelve los datos de ese registro en la caché.																																										

Opciones de caché local

Una caché local es una caché provisoria que se usa únicamente durante la ejecución de una etapa Query Cache. La Query Cache construye la caché a partir de la tabla de base de datos que elija. Luego, busca datos en la caché basándose en los campos clave y las condiciones de búsqueda y devuelve datos a partir de registros de cruce en la caché, y agrega los datos del registro de la caché al registro en el flujo de datos.

Una caché local es dinámica, ya que se crea durante la ejecución del trabajo de la Query Cache. Una vez que la Query Cache completa la lectura de los datos, la caché se borra automáticamente de la memoria. Una caché local se vuelve a crear cada vez que se ejecuta la etapa Query Cache. Elija una caché local si solo se va a utilizar en un flujo de datos o si la tabla de búsqueda cambia con frecuencia.

Nombre de la opción	Descripción
Tipo de caché	Especifica la opción Caché local.
Conexión	<p>Seleccione la conexión de base de datos que quiere utilizar. Las opciones pueden variar según las conexiones definidas en el Administrador de conexión de base de datos de Management Console. Si necesita hacer una conexión de base de datos nueva o modificar o eliminar una conexión de base de datos existente, haga clic en Administrar.</p> <p>Si va a agregar o modificar una conexión de base de datos, complete los siguientes campos:</p> <p>Nombre de conexión Ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.</p> <p>Controlador de base de datos Seleccione el tipo de base de datos correspondiente.</p> <p>Opciones de conexión Especifique el host, el puerto, la instancia, el nombre de usuario y la contraseña que se utilizarán para conectarse con la base de datos.</p>
Tabla/Vista	Especifique la tabla o vista en la base de datos que se desea consultar.
Campos de base de datos	En esta columna aparecen los campos de la base de datos. No puede modificar estos nombres de campos.
Campos de etapa	Esta columna lista los nombres de campo que se usan en el flujo de datos. Si desea cambiar el nombre de un campo, haga clic en el nombre del campo y escriba un nuevo nombre.
Tipo	En esta columna aparece el tipo de datos de cada campo del flujo de datos.
Incluir	Active la casilla de esta columna para que la consulta devuelva el valor del campo de la caché. Desactive la casilla si no desea que la consulta devuelva el campo de la caché.

Nombre de la opción	Descripción
---------------------	-------------

Valor de error predeterminado	
----------------------------------	--


Nombre de la opción Descripción

Especifica el valor que se va a mostrar en el campo de flujo de datos si se produce un error en la consulta. La lista desplegable muestra los valores válidos que corresponden a los tipos de datos del campo consultado. Por ejemplo, en el caso de un **número entero** la opción que se muestra es **-1**.

También puede ingresar un valor para este campo. Consulte la tabla que se muestra a continuación para ver una lista de los valores válidos de errores predeterminados correspondientes a varios tipos de datos.

Tipo de datos Valor de error predeterminado válido junto con un tipo de datos (en paréntesis)

Tipo de datos	Nulo	-1	1899-12-30	1899-12-30 12:00:00	12:00:00	Falso	Vacío
Fecha							
Integer		✓					
Long Entero largo		✓					
Float		✓					
Grande Decimal		✓					
Double		✓					
Cadena	✓	✓	✓	✓	✓	✓	✓
Hora					✓		
Fecha y hora			✓				

Nombre de la opción	Descripción
	<p>Tipo de datos Valor de error predeterminado válido junto con un tipo de datos (en paréntesis)</p> <p>Nulo -1 1899- 1899- 12:00:00 Falso Vacío (Entero) 12-30 12-30 (Hora) 12:00:00 (Fecha) (Fecha-hora)</p> <p>Boolean </p>
Campo de clave	Especifica el campo de la base de datos que se utilizará como una clave de búsqueda. Si el valor del campo en la columna Input Field coincide con el valor en el Campo clave de la base de datos, la consulta devuelve los datos de ese registro en la base de datos.
Campo de entrada	Especifica el campo del flujo de datos cuyo valor se usará como la clave. Si el valor de este campo coincide con el valor en el Campo clave de la base de datos, la consulta devuelve los datos de ese registro en la base de datos.

Opciones de caché avanzada

Una caché avanzada es una caché temporal similar a la caché local. Se utiliza durante la ejecución de la etapa Query Cache. Crea la caché en función de la consulta SQL que lee los datos desde las tablas que me mencionan en la consulta. Luego, busca datos en la caché basándose en las claves de búsqueda que se mencionan en la cláusula *where* y devuelve datos a partir de registros de cruce en la caché, y agrega los datos del registro de la caché al registro en el flujo de datos.

Una caché avanzada es dinámica, ya que se crea durante la ejecución del trabajo de la Query Cache. Una vez que la Query Cache completa la lectura de los datos, la caché se borra automáticamente de la memoria. Una caché avanzada se vuelve a crear cada vez que se ejecuta la Query Cache. Elija una opción de caché avanzada si va a leer los datos de varias tablas y existen algunas consultas complejas que deben ejecutarse para la creación de la caché.

Nombre de la opción	Descripción
Tipo de caché	Especifica la opción de caché avanzada.

Nombre de la opción	Descripción
Conexión	<p>Seleccione la conexión de base de datos que quiere utilizar. Las opciones pueden variar según las conexiones definidas en el Administrador de conexión de base de datos de Management Console. Si necesita hacer una conexión de base de datos nueva o modificar o eliminar una conexión de base de datos existente, haga clic en Administrar.</p> <p>Si va a agregar o modificar una conexión de base de datos, complete los siguientes campos:</p> <p>Nombre de conexión Ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.</p> <p>Controlador de base de datos Seleccione el tipo de base de datos correspondiente.</p> <p>Opciones de conexión Especifique el host, el puerto, la instancia, el nombre de usuario y la contraseña que se utilizarán para conectarse con la base de datos.</p>
Consulta	Proporciona la consulta SQL para leer los datos desde la base de datos. La consulta puede leer datos desde múltiples tablas.
Where	Este texto se utiliza como la cláusula <i>where</i> para buscar la caché que se creó en función de Query. El usuario puede especificar un campo de entrada en la Query con el operador \$ como prefijo. Por ejemplo, una expresión <code>_id = \${_inputId}</code> , en la que <code>_inputId</code> es el campo de entrada e <code>_id</code> es la columna de búsqueda en la caché.
Obtener campos	Esto completa la cuadrícula con los campos que se seleccionan para la creación de la caché con la consulta SQL.
Campos de base de datos	En esta columna aparecen los campos que se recuperan en la base de datos. No puede modificar estos nombres de campos.
Campos de etapa	Esta columna lista los nombres de campo que se usan en el flujo de datos. Si desea cambiar el nombre de un campo, haga clic en el nombre del campo y escriba un nuevo nombre.
Tipo	En esta columna aparece el tipo de datos de cada campo del flujo de datos.

Nombre de la opción	Descripción
---------------------	-------------

Valor de error predeterminado	
----------------------------------	--


Nombre de la opción	Descripción
---------------------	-------------

Especifica el valor que se va a mostrar en el campo de flujo de datos si se produce un error en la consulta. La lista desplegable muestra los valores válidos que corresponden a los tipos de datos del campo consultado. Por ejemplo, en el caso de un **número entero** la opción que se muestra es **-1**.

También puede ingresar un valor para este campo. Consulte la tabla que se muestra a continuación para ver una lista de los valores válidos de errores predeterminados correspondientes a varios tipos de datos.

Tipo de datos	Valor de error predeterminado válido junto con un tipo de datos (en paréntesis)
---------------	---

Tipo de datos	Nulo	-1 (Entero)	1899-12-30 12:00:00 (Fecha)	1899-12-30 12:00:00 (Fecha)	12:00:00 (Hora)	Falso	Vacío
Fecha							
Integer		✓					
Long Entero largo		✓					
Float		✓					
Grande Decimal		✓					
Double		✓					
Cadena	✓	✓	✓	✓	✓	✓	✓
Hora					✓		
Fecha y hora			✓				

Nombre de la opción	Descripción
	Tipo de datos
	Valor de error predeterminado (en paréntesis)
	Nulo -1 (Entero) 1899-12-30 1899-12-30 12:00:00 (Hora) Falso Vacío
	12:00:00 (Fecha)
	(Fecha-ha)
	Boolean 

Ficha Tiempo de ejecución

Las opciones disponibles en la ficha Runtime son comunes para las cachés globales, locales y avanzadas.

Nombre de la opción	Descripción
Opciones de coincidencia	Especifica qué se debe hacer si hay más de un registro en la caché que coincide con la consulta. Una de las siguientes: <ul style="list-style-type: none"> Devolver todas las coincidencias Devolver los datos de todos los registros de la caché que tengan un valor coincidente en el campo clave o en los campos clave. Devolver el primer registro coincidente Devolver solo los datos del primer registro de la caché que tenga un valor coincidente en el campo clave o en los campos clave. Devolver el último registro coincidente Devolver solo los datos del último registro de la caché que tenga un valor coincidente en el campo clave o en los campos clave.

Nombre de la opción	Descripción
Opciones de etapa	<p>En esta sección, se enumeran las opciones de flujo de datos que se utilizan en la consulta SQL de esta etapa, y le permite brindar un valor predeterminado para todas estas opciones. En la columna Nombre, se enumeran las opciones, y en la columna Valor correspondiente puede ingresar los valores predeterminados.</p> <p>Nota: El valor predeterminado que se brinda aquí también se muestra en la sección Opciones de flujo de datos de mapa para las etapas del cuadro de diálogo Opciones de flujo de datos. El cuadro de diálogo también le permite cambiar el valor predeterminado. En caso de conflicto entre los valores predeterminados que se proporcionan a través de Opciones de etapa, Opciones de flujo de datos y Ejecutor de trabajos, el orden de prioridad es: el valor que se proporciona a través de Ejecutor de trabajos > el valor que se define a través del cuadro de diálogo Opciones de flujo de datos > el valor que se ingresa a través de Opciones de etapa.</p>

Query DB

La etapa Query DB le permite utilizar los campos como parámetros en una consulta de base de datos y devolver los resultados de la consulta como nuevos campos en el flujo de datos.

Nota: Si desea realizar consultas en una base de datos espacial, utilice la etapa Query Spatial Data en vez de Query DB.

Ficha General

Opción	Descripción
Conexión	<p>Seleccione la conexión de base de datos que quiere utilizar. Las opciones pueden variar según las conexiones definidas en el Administrador de conexión de base de datos de Management Console. Si necesita hacer una conexión de base de datos nueva o modificar o eliminar una conexión de base de datos existente, haga clic en Administrar.</p> <p>Si va a agregar o modificar una conexión de base de datos, complete los siguientes campos:</p> <p>Nombre de conexión Ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.</p> <p>Controlador de base de datos Seleccione el tipo de base de datos correspondiente.</p> <p>Opciones de conexión Especifique el host, el puerto, la instancia, el nombre de usuario y la contraseña que se utilizarán para conectarse con la base de datos.</p>
Tabla/Vista	Especifique la tabla o vista en la base de datos que se desea consultar.

Opción	Descripción
Where	<p>Si desea utilizar una instrucción <code>WHERE</code>, ingrésela aquí. Tenga en cuenta que no debe incluir la palabra <code>WHERE</code> en la instrucción. El propósito de la instrucción <code>WHERE</code> es devolver solo los datos de los registros que coincidan con la condición que se especifique.</p> <p>Para especificar un valor desde un campo del flujo de datos, use esta sintaxis:</p> <pre> \${<field name>} </pre> <p>Donde <code><field name></code> es el nombre de un campo en el flujo de datos.</p> <p>Por ejemplo:</p> <pre> account_number=\${customer_key} </pre> <p>En este ejemplo, la consulta devolverá los datos de los registros en los que el valor en la columna de la tabla <code>account_number</code> coincida con el valor en el campo del flujo de datos <code>customer_key</code>.</p> <p>Nota: Si está consultando una base de datos que distingue entre mayúsculas y minúsculas, asegúrese de ingresar el nombre del campo en el mismo formato que se utilizó en la tabla de la base de datos. En otras palabras, encierre el nombre del campo entre comillas (") si los nombres de campo se encerraron entre comillas durante la creación de la tabla.</p> <p>Haga clic en Vista previa para previsualizar los datos (los primeros 50 registros) según los criterios que haya definido.</p> <p>Nota: La función de vista previa en Query DB no funciona si se utiliza un campo de flujo de datos en la instrucción <code>WHERE</code>. En su lugar, se puede obtener una vista previa del resultado utilizando la herramienta de inspección de flujo de datos en Enterprise Designer.</p>
Arrojar registros sin resultados	<p>Active esta casilla si desea que los registros cuya consulta no devuelva resultados de todos modos aparezcan en Query DB. Si desactiva esta casilla, el registro no se devolverá. Se recomienda dejar esta opción seleccionada.</p>
Incluir	<p>En la tabla campos, seleccione los campos que desea incluir haciendo clic en el cuadro Incluir ubicado al lado del campo.</p>

Ficha Ordenar

Si desea ordenar los registros según el valor de un campo, especifique los campos que desea ordenar.

Establecimiento de parámetros de Query DB en tiempo de ejecución

Puede configurar la etapa Query DB de modo que los valores de la cláusula `WHERE` se especifiquen en el tiempo de ejecución. Esto resulta útil en casos donde desea hacer que el nombre de la columna en la cláusula `WHERE` se pueda configurar utilizando las opciones de flujo de datos.

1. Abra el flujo de datos en Enterprise Designer.
2. Configure los campos de nombre **Conexión** y **Tabla/Vista** para apuntar a la base de datos que desea consultar.
3. En el campo **Donde**, ingrese una instrucción `WHERE` utilizando el siguiente formato para los valores para los que desea establecer parámetros: `${parameter}`.

Por ejemplo:

```
${COL}=${EmployeeID}
```

En este caso, COL representa una opción de flujo de datos que se completará con el nombre de la columna para la tabla en el tiempo de ejecución.

4. Cierre la ventana de opciones de Query DB.
5. Haga clic en el icono Opciones de flujo de datos en la barra de herramientas o haga clic en **Editar > Opciones de flujo de datos**. Aparecerá la ventana **Opciones de flujo de datos**.
6. Haga clic en **Agregar**. Aparecerá la ventana **Definir opciones de flujo de datos**.
7. Seleccione la etapa Query DB.
8. Especifique **Nombre de opción** y **Rótulo de opción**.

El valor en el campo **Nombre de opción** debe ser igual al valor ingresado en el formato `${parameter}` en la cláusula `WHERE`. En el campo **Rótulo de opción** puede especificar un rótulo diferente o mantenerlo igual que el **Nombre de opción**.

Por ejemplo: COL

9. Especifique el **Valor predeterminado**. Por ejemplo: "EmpID".
10. Haga clic en **Aceptar**.

Este procedimiento asigna el nombre de columna de base de datos real, es decir EmpID, al nombre de opción de tiempo de ejecución "COL". El nombre de columna de la base de datos debe citarse correctamente con el identificador de cita específico para la base de datos en particular.

Query NoSQL DB

La etapa **Query NoSQL DB** permite búsquedas de datos requeridos de una base de datos NoSQL. La etapa admite bases de datos MongoDB.

Pestaña General

Nombre de campo	Descripción
Conexión	<p>Seleccione la conexión de base de datos obligatoria en la lista desplegable. Las opciones que se muestran varían según las conexiones especificadas en Management Console.</p> <p>Para agregar una nueva conexión, consulte Conexión con NoSQL en la página 54.</p> <p>Para modificar una conexión existente, selecciónela de la lista de conexiones en la página Fuentes de datos de Management Console, ábrala, haga los cambios necesarios y, a continuación, haga clic en el botón Guardar.</p>
Tabla/Vista	<p>Especifica la colección o vista en la base de datos que se desea consultar.</p> <p>Nota: En una base de datos MongoDB, una tabla/vista se llama <i>colección</i>.</p>
Archivo de esquema	<p>Haga clic en el botón Examinar (...) para seleccionar un archivo de esquema JSON. Este archivo es opcional. Los campos de la pestaña de campos se pueden generar con el archivo de esquema o con la tabla/visualización de la base de datos.</p> <p>Para borrar la ruta del archivo seleccionado, haga clic en Borrar.</p> <p>Nota: Los campos siempre se generarán con el archivo de esquema si se selecciona uno.</p>
Where	<p>Ingrese la cláusula <code>where</code> para definir los criterios para la búsqueda.</p> <p>Para ver una lista de los operadores admitidos en la cláusula <code>WHERE</code>, consulte http://docs.mongodb.org/manual/reference/operator/query/.</p>
Vista previa	<p>Muestra los registros desde la tabla o vista seleccionada.</p> <p>Nota: Si hace clic en Vista previa, se recopilan los primeros 50 registros de la base de datos seleccionada, sin aplicar los criterios de filtro especificados en el campo Dónde.</p>

Nombre de campo	Descripción
Expandir todo	Expande los elementos en el árbol de vista previa.
Cerrar todo	Cierra los elementos en el árbol de vista previa.

Pestaña Campos

La pestaña Campos le permite seleccionar los datos que desea pasar a la siguiente etapa. Para obtener más información, consulte [Definición de campos - Query NoSQL DB](#) en la página 141.

Definición de campos - Query NoSQL DB

La ficha **Campos** muestra el campo y su tipo tal como se define en el archivo NoSQL DB/Schema.

1. En la **ficha Campos** haga clic en **Regenerar**.

Esto genera los datos agregados a partir de los primeros 50 registros. Los datos se muestran en el siguiente formato: `Fieldname (datatype)`

Nota: Si se explora el archivo de esquema, se generan los campos utilizando el archivo de esquema. Se omite la tabla o vista. Para restablecer el archivo de esquema, haga clic en **Borrar**.

2. Para modificar el nombre y tipo de un campo, resalte el campo y haga clic en **Modificar**.
3. En el campo **Nombre**, elija el campo que desea agregar o ingrese el nombre del campo.
4. En el campo **Tipo** puede dejar el tipo de datos como cadena, si no se ha propuesto realizar ninguna operación matemática con los datos. Sin embargo, si se propone realizar ese tipo de operaciones, seleccione un tipo de datos apropiado. Esto convertirá los datos de cadena del archivo en un tipo de datos que permitirá el manejo adecuado de los datos en el flujo de datos.

La etapa admite los siguientes tipos de datos:

double	Un tipo de dato numérico que contiene números de doble precisión tanto negativos como positivos entre 2^{-1074} y $(2-2^{-52}) \times 2^{1023}$. En notación científica, el rango de valores es - 1.79769313486232E+308 a 1.79769313486232E+308.
float	Un tipo de dato numérico que contiene números de precisión simple tanto negativos como positivos entre 2^{-149} y $(2-2^{-23}) \times 2^{127}$. En notación científica, el rango de valores es - 3.402823E+38 a 3.402823E+38.
integer	Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{31} (-2,147,483,648) y $2^{31}-1$ (2,147,483,647).
long	Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{63} (-9,223,372,036,854,775,808) y $2^{63}-1$ (9,223,372,036,854,775,807).
string	Secuencia de caracteres.

5. También puede agregar campos adicionales que no estén presentes en el archivo de tabla o esquema. Haga clic en **Agregar** para agregar un nuevo campo. Para eliminar un campo, haga clic en **Eliminar**.

Nota: Solo puede agregar un nuevo campo en Tipo de lista.

6. Haga clic en **Aceptar**.

Configuración de opciones de flujo de datos - Query NoSQL DB

Este procedimiento describe la forma de configurar un flujo de datos compatible con las opciones de ejecución para la base de datos **Query NoSQL DB**.

1. Abra el flujo de datos en Enterprise Designer.
2. Si desea configurar las opciones de tiempo de ejecución para una etapa en un flujo de datos incrustado, abra este último.
3. Haga clic en el ícono Opciones de flujo de datos en la barra de herramientas o haga clic en **Editar > Opciones de flujo de datos**. Aparecerá el cuadro de diálogo **Opciones de flujo de datos**.
4. Haga clic en **Agregar**. Aparecerá el cuadro de diálogo **Definir opción de flujo de datos**.
5. Expanda la etapa **Query NoSQL DB**.
6. Se exponen las siguientes opciones de flujo de datos para consultar una base de datos Mongo DB:
 - Conexión
 - Tabla

El nombre de la opción **Query NoSQL DB** seleccionada aparece en los campos **Nombre de opción** y **Etiqueta de opción**. Este es el nombre de opción que deberá especificarse durante el tiempo de ejecución para definir esta opción.

7. Ingrese una descripción de la opción en el campo **Descripción**.
8. En el campo **Destino**, seleccione la opción **Etapas seleccionadas**.
9. Si desea limitar los valores que pueden especificarse durante la ejecución, edite las opciones en el campo **Valores válidos** haciendo clic en el icono a la derecha del campo.
10. Si desea cambiar el valor predeterminado, especifique un valor diferente en el campo **Valor predeterminado**.

Nota: En el caso de un servicio, solo puede modificar los valores predeterminados antes de exponer el servicio por primera vez. Después de exponer el servicio, ya no es posible modificar los valores predeterminados con Enterprise Designer. En cambio, debe usar Management Console. Para obtener más información, consulte [Especificación de las opciones de servicio predeterminadas](#).

11. Haga clic en **Aceptar**.
12. Continúe agregando opciones, según sus preferencias.
13. Haga clic en **ACEPTAR** en el cuadro de diálogo **Opciones de flujo de datos** cuando haya terminado de agregar opciones.
14. Guarde y exponga el flujo de datos.

Read from DB

La etapa **Read From DB** lee los datos desde una tabla o vista de base de datos como entrada a un flujo de datos. La etapa está disponible para los trabajos, servicios y subflujos, pero no para los flujos de procesos.

Pestaña General

Nombre de campo	Descripción						
Conexión	<p>Seleccione la conexión de base de datos que quiere utilizar. Las opciones pueden variar según las conexiones definidas en el Administrador de conexión de base de datos de Management Console. Si necesita hacer una conexión de base de datos nueva o modificar o eliminar una conexión de base de datos existente, haga clic en Administrar.</p> <p>Si va a agregar o modificar una conexión de base de datos, complete los siguientes campos:</p> <table border="0"> <tr> <td>Nombre de conexión</td> <td>Ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.</td> </tr> <tr> <td>Controlador de base de datos</td> <td>Seleccione el tipo de base de datos correspondiente.</td> </tr> <tr> <td>Opciones de conexión</td> <td>Especifique el host, el puerto, la instancia, el nombre de usuario y la contraseña que se utilizarán para conectarse con la base de datos.</td> </tr> </table>	Nombre de conexión	Ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.	Controlador de base de datos	Seleccione el tipo de base de datos correspondiente.	Opciones de conexión	Especifique el host, el puerto, la instancia, el nombre de usuario y la contraseña que se utilizarán para conectarse con la base de datos.
Nombre de conexión	Ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.						
Controlador de base de datos	Seleccione el tipo de base de datos correspondiente.						
Opciones de conexión	Especifique el host, el puerto, la instancia, el nombre de usuario y la contraseña que se utilizarán para conectarse con la base de datos.						
SQL	<p>Escriba la consulta SQL para especificar los registros que se deben leer desde el origen de datos en el flujo de datos que se está ejecutando. Puede escribir manualmente la consulta SQL en este campo. De manera alternativa, utilice el Generador de consultas visual para elaborar la consulta mediante un clic en Elaborar SQL....</p> <p>La consulta SQL puede incluir variables en lugar de nombres de columnas. El uso de variables permite personalizar la consulta durante la ejecución. Para obtener más información, consulte Variables de las consultas en la página 149.</p>						

Nombre de campo	Descripción
Elaborar SQL	<p>Seleccione varias columnas para crear una consulta compleja, y cree combinaciones y consultas anidadas con un clic en Elaborar SQL. Se abre el Generador de consultas visual. Para obtener más información, consulte Generador de consultas visual en la página 145.</p> <p>Nota: Toda consulta creada mediante el Generador de consultas visual aparece con los nombres completamente calificados de las columnas y las tablas en el campo SQL.</p>
Regenerar campos	<p>Para ver el esquema de los datos que se deben recopilar mediante la consulta, haga clic en Regenerar campos.</p> <p>Si edita una consulta existente, haga clic en Regenerar campos para recopilar el esquema modificado.</p> <p>Nota: Después de hacer clic en Regenerar campos, los nombres de entidad de la consulta SQL se mantienen, no se reemplazan con sus nombres calificados completos.</p>
Vista previa	<p>Para ver un ejemplo de los registros recopilados con la consulta SQL, haga clic en Vista previa.</p>

Nota: La etapa **Read From DB** permite modificar el tipo de un campo de entrada.

Nota: La etapa **Read From DB** lee todos los valores del tipo de datos `date` como valores `String`. Este es el comportamiento del *controlador jTDS*, que es el controlador predeterminado que usa Spectrum. Para manejar todos los valores de tipos de datos `date` tal como están, use el controlador JDBC de Microsoft.

Pestaña Tiempo de ejecución

Nombre de campo	Descripción
Tamaño de búsqueda:	<p>Seleccione esta opción para especificar el número de registros que se van a leer desde la tabla de base de datos al mismo tiempo. Por ejemplo, si el valor del Tamaño de recuperación es 100 y el número total de registros que se van a leer es 1000, haría 10 recorridos a la base de datos para leer todos los registros.</p> <p>La configuración de un tamaño de recuperación óptimo puede mejorar significativamente el rendimiento.</p> <p>Nota: Puede calcular el tamaño de recuperación óptimo para su entorno, al probar los tiempos de ejecución entre una etapa Read from DB y una etapa Write to Null. Para obtener más información, consulte Determinar un tamaño de recuperación óptimo en la página 322.</p>
Opciones de etapa	<p>En esta sección, se enumeran las opciones de flujo de datos que se utilizan en la consulta SQL de esta etapa, y le permite brindar un valor predeterminado para todas estas opciones. En la columna Nombre, se enumeran las opciones, y en la columna Valor correspondiente puede ingresar los valores predeterminados.</p> <p>Nota: El valor predeterminado que se brinda aquí también se muestra en la sección Opciones de flujo de datos de mapa para las etapas del cuadro de diálogo Opciones de flujo de datos. El cuadro de diálogo también le permite cambiar el valor predeterminado. En caso de conflicto entre los valores predeterminados que se proporcionan a través de Opciones de etapa, Opciones de flujo de datos y Ejecutor de trabajos, el orden de prioridad es: el valor que se proporciona a través de Ejecutor de trabajos > el valor que se define a través del cuadro de diálogo Opciones de flujo de datos > el valor que se ingresa a través de Opciones de etapa.</p>

Generador de consultas visual

El Generador de consultas visual ofrece una interfaz visual para construir consultas complejas de SQL en la etapa Read from DB (Lectura desde base de datos). Para trabajar con el Generador de consultas visual, necesita tener conocimientos básicos sobre SQL.

Para acceder al Generador de consultas visual, haga clic en el botón **Elaborar SQL** en Read from DB.

La ventana principal del generador de consultas está dividida en las siguientes partes:

- El **área de generación de consultas** es el área principal en el que se muestra la representación visual de las consultas. Esta área permite definir objetos de base de datos de origen, definir vínculos entre ellos y configurar las propiedades de las tablas y los enlaces.

- El **panel de columnas** está ubicado debajo del área de generación de consultas. Se utiliza para realizar todas las operaciones necesarias con columnas y expresiones de salida de consulta. En este panel, se pueden definir alias de campos, tareas de ordenación y agrupamiento, y definir criterios.
- El control de la página ubicado arriba del área de generación de consultas permite alternar entre consulta principal y subconsultas.

Adición de objetos a una consulta

Para agregar un objeto a la consulta, utilice el árbol **Objetos de la base de datos**. Los objetos que se encuentran dentro del árbol están agrupados por base de datos, esquema y tipo. Arrastre el objeto que desea agregar a la consulta y déjelo en el área de creación de consultas. De forma alternativa, también puede hacer doble clic en el objeto para agregarlo a la consulta.

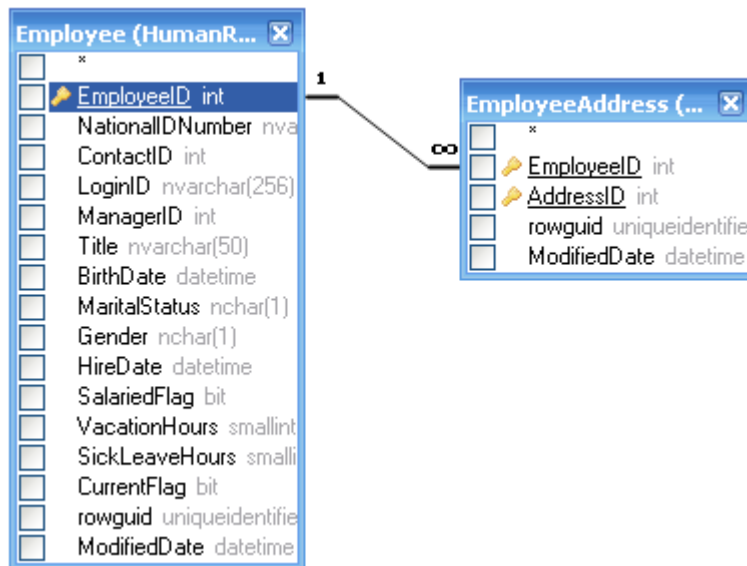
Configuración del alias de un objeto

Para configurar el alias de un objeto o una tabla derivada en la consulta, haga doble clic en el objeto y seleccione **Propiedades**. Aparecerá el cuadro de diálogo **Propiedades de origen de datos**. Es posible que contenga otras opciones de origen de datos específicas del servidor, pero la propiedad del alias es igual en todos los servidores de bases de datos.

Unión de tablas

Cuando se agregan dos objetos referenciados con una relación de clave externa a la consulta, se unen automáticamente mediante una cláusula INNER JOIN. Para aquellos servidores que no admiten la cláusula JOIN, el generador de consultas agrega una condición a la parte WHERE de la consulta.

Para unir dos objetos de forma manual, seleccione el campo mediante el cual desea enlazar un objeto con otro y arrástrelo al campo correspondiente del otro objeto. Después de que haya arrastrado el objeto, aparecerá una línea que conecta los campos enlazados. En los extremos del enlace se ubican los principales símbolos de cardinalidad cuando existe una relación correspondiente en la base de datos.



Para quitar un enlace entre objetos, haga doble clic en la línea de enlace y seleccione **Eliminar**.

Para cambiar el tipo de combinación, haga doble clic en la línea de enlace.

Selección de los campos de salida

Para agregar un campo a la lista de campos de salida de la consulta, marque la casilla que se encuentra a la izquierda del nombre del campo en la lista de campos de fuentes de datos del área de **Generación de consultas**. Para incluir todos los campos del objeto, marque la casilla que se encuentra a la izquierda del asterisco en la lista de campos de origen de datos. También es posible arrastrar campos desde el área de **Generación de consultas** al panel **Columnas** para obtener el mismo resultado.

Si no selecciona ningún campo de los orígenes de datos de la consulta, se agregará un asterisco a la lista de selección de la consulta resultante ("Select * From ..."). Ello se debe a que una consulta SELECT sin columnas producirá un error en la mayoría de los servidores de base de datos. El asterisco es eliminado de la consulta si selecciona cualquier campo o agrega cualquier expresión de salida a la consulta.

Consejo: Otra forma de agregar un campo es seleccionar un nombre de campo desde la lista desplegable de la columna Expresión en el panel **Columnas**. También es posible ingresar cualquier expresión válida en la columna **Expresión** del panel **Columnas**. Para insertar una línea vacía en el panel **Columnas**, presione las teclas Alt+Insertar.

Para eliminar un campo del panel **Columnas**, desmarque la casilla que se encuentra a la izquierda del nombre del campo en el área de **Generación de consultas** o presione las teclas Alt+Suprimir en el panel **Columnas**.

Para subir una línea, presione las teclas Alt+Arriba. Para bajar una línea, presione las teclas Alt+Abajo.

Para eliminar una expresión de la lista SELECT de la consulta, desmarque la casilla de la columna **Salida**.

A fin de configurar un alias para una expresión, ingrese el alias en la columna **Alias**. Los alias se convierten en encabezados de columnas en el conjunto de datos resultante.

Ordenación de un conjunto de datos

Para ordenar el conjunto resultante de datos, utilice las columnas Tipo de ordenación y Criterio de ordenación del panel **Columnas**. La columna Tipo de ordenación permite ordenar los datos de forma ascendente o descendente. La columna Criterio de ordenación permite configurar el criterio con el que se ordenarán los campos, en caso de que se ordenara más de un campo.

Para desactivar la ordenación por campo, desmarque la columna Tipo de ordenación del campo.

Definición de criterios

Para definir criterios, utilice la columna **Criterios** y las columnas **O** del panel **Columnas**. Al escribir las condiciones en estas columnas, omita la expresión en sí misma. Por ejemplo, para que los siguientes criterios se incluyan en su consulta:

```
WHERE (Field1 >= 10) AND (Field1 <= 20)
```

Ingrese lo siguiente en la celda Criterios de la expresión Field1:

```
>= 10 AND <= 20
```

Los criterios ubicados en las columnas **O** se agruparán por columna utilizando el operador AND y luego se concatenan en la cláusula WHERE (o HAVING) utilizando el operador OR. Por ejemplo, los criterios que se muestran a continuación producirán la siguiente instrucción SQL. Cabe destacar que los criterios para Field1 están ubicados tanto en la columna Criterios como en la columna O.

Output	Expression	Aggregate	Alias	Sort Type	Sort Order	Grouping	Criteria	Or...	Or...
<input checked="" type="checkbox"/>	Field1					<input type="checkbox"/>	= 10	= 10	
<input checked="" type="checkbox"/>	Field2					<input type="checkbox"/>	< 0	> 10	
<input type="checkbox"/>						<input type="checkbox"/>			

```
WHERE (Field1= 10) AND ((Field2 < 0) OR (Field2 > 10))
```

Algunas expresiones pueden ser del tipo boolean, por ejemplo, la cláusula EXISTS. En este caso, debería ingresar "= True" en la columna Criterios de esas expresiones o "= False" si quiere colocar un operador NOT antes de la expresión.

Agrupamiento de campos de salida

Para generar una consulta con agrupamiento, marque las expresiones de agrupamiento con la casilla **Agrupamiento**.

Una consulta con agrupamiento puede tener solo expresiones agregadas o de agrupamiento en la lista SELECT. Por lo tanto, el generador de consultas permite configurar la casilla Salida para expresiones agregadas o de agrupamiento. Si intenta configurar esta casilla para una columna que no tiene la función de agregado o agrupamiento, se configurará automáticamente una casilla de Agrupamiento para mantener la validez de la consulta SQL resultante.

Cuando el panel **Columnas** contiene columnas que tienen marcada la casilla **Agrupamiento**, aparece en la cuadrícula una columna nueva denominada **Criterios para**. Esta columna aplica criterios a los grupos de expresión o a sus valores.

Por ejemplo, si tiene una columna "Cantidad" con la función de agregado "Avg" en su consulta e ingresa > 10 en la columna Criterios. Si el valor "for groups" está configurado en los criterios de columnas, la consulta resultante contendrá solo grupos con una cantidad promedio superior a 10 y su consulta tendrá la condición "Avg(Quantity) > 10" en una cláusula HAVING. Si el valor "for values" está configurado en los Criterios de columna, la consulta resultante calculará la función de agregado promedio solo para los registros cuyo valor de Cantidad sea superior a 10 y su consulta tendrá la condición "Quantity > 10" en una cláusula WHERE.

Definición de las propiedades de una consulta SQL

Es posible definir las opciones específicas de su servidor de base de datos utilizando el menú emergente contextual del área de **Generación de consultas**.

Variables de las consultas

En la etapa **Read From DB**, mientras se define la consulta que se ejecutará, puede incluir las variables en lugar de los nombres reales de las columnas. Usar variables en la consulta le permite personalizar las condiciones de la consulta en el tiempo de ejecución (con el uso de **Opciones de flujo de datos**) o a través del **Ejecutor de trabajos**.

Sin embargo, también puede proporcionar el valor Opciones de etapa en la ficha de **Tiempo de ejecución** y ver el esquema y la muestra de los registros que la consulta va a recopilar mediante el uso de los botones **Regenerar campos** y **Vista previa** respectivamente.

Se define una variable al utilizar el formato `#{variable}` e insertarla en la cláusula `select` o `where` de una consulta SQL.

Nota: Puede editar una consulta generada mediante el **Generador de consultas visual** para incluir las variables. Sin embargo, la consulta editada ya no podrá ser leída por el Generador de consultas visual. El botón **Elaborar SQL** se desactiva cuando incluye una variable en una consulta SQL escrita o generada de forma manual.

Inserción de una variable a la consulta

1. Abra el trabajo requerido, que incluye una etapa de **Read From DB**. De manera alternativa, agregue una etapa **Read From DB** al trabajo.

2. Abra las **Opciones de Read From DB** de la etapa **Read From DB**.
3. Cree la consulta SQL en el campo **SQL**, ya sea manualmente o mediante el Generador de consultas visual. Para obtener más información, consulte [Generador de consultas visual](#) en la página 145.
4. Agregue las condiciones deseadas en la cláusula `where` de la consulta mediante variables con la sintaxis `#{variable}`.
Por ejemplo, en la tabla `CUSTOMERS`, que tiene una columna `AGE` con valores como 28, 32, 30 y así sucesivamente; y una columna `SALARY` con valores como 1000, 1500, 2200, etc., formule una consulta SQL de la siguiente manera:

```
select * from CUSTOMERS where #{condition1} > 28 and #{condition2} > 1200
```

Nota: Mediante la inserción de una variable en la cláusula `where` de la consulta SQL, el botón **Elaborar SQL...** se desactiva.

5. Para ver el esquema y los registros de muestra que se obtendrán mediante la consulta, ingrese el valor Opciones de etapa en la ficha **Tiempo de ejecución** y, a continuación, haga clic en los botones **Regenerar campos** y **Vista previa** respectivamente.
6. Haga clic en **Aceptar**.

La cláusula `where` de la consulta SQL ahora se puede personalizar en el tiempo de ejecución utilizando las **Opciones del flujo de datos**, o bien durante la ejecución del trabajo mediante JobExecutor.

Nota: También se puede ubicar una variable en la cláusula `select` de una consulta SQL. Sin embargo, el nombre de dicha variable debe coincidir con el nombre de una de las columnas de la tabla que se consulta.

Configuración de una variable de consulta como una opción de flujo de datos

1. Abra el trabajo requerido para la consulta que contenga las variables que se han definido en una etapa **Read From DB**.
2. Abra **Editar > Opciones del flujo de datos**.
3. Haga clic en **Agregar**.
4. En la sección **Aplicar opciones de flujo de datos a etapas**, amplíe la entrada **Read From DB**.
Las variables definidas en la consulta SQL durante la etapa **Read From DB** se encuentran detalladas junto con los demás atributos de la etapa.
5. Seleccione la variable que desea personalizar mediante la casilla de verificación correspondiente.
6. Ingrese un nombre pertinente para la variable en el campo **Rótulo de opción**.
7. En el campo **Valor predeterminado**, ingrese el nombre de la columna que se utilizará en lugar de la variable seleccionada en la cláusula `where` de la consulta SQL. De manera alternativa, ingrese un valor constante para que sea utilizado en lugar de la variable en la cláusula `where`.

Por ejemplo, para la siguiente consulta SQL, definida en la etapa **Read From DB**:

```
select * from CUSTOMERS where #{condition1} > 28 and #{condition2} > 1200
```

puede seleccionar la columna `AGE` de la tabla `CUSTOMERS` como el **Valor predeterminado** para la variable `condition1`, y la columna `SALARY` como el **Valor predeterminado** para la variable `condition2`.

Durante el tiempo de ejecución, la consulta es interpretada como:

```
select * from CUSTOMERS where AGE > 28 and SALARY > 1200
```

8. Repetir los pasos 5 a 7 para todas las variables ubicadas en la consulta SQL durante la etapa **Read From DB**.
9. Haga clic en **Aceptar**.

Durante la ejecución del flujo de datos, se emplea la consulta personalizada para obtener los datos necesarios.

Configuración de una variable de consulta para Job Executor

Nota: Asegúrese de haber descargado Spectrum™ Job Executor en su servidor.

1. Cree un archivo de texto que defina los valores predeterminados de las variables incluidas en la consulta SQL en la etapa **Read From DB** del trabajo.

Para asignar un nombre de columna `AGE` como valor predeterminado para una variable `condition1`, cree un archivo de texto, por ejemplo `variables.txt`, e incluya la siguiente línea en el archivo:

```
condition1=AGE
```

Para asignar un valor constante, por ejemplo 20, como valor predeterminado para una variable `condition1`, incluya la siguiente línea en el archivo:

```
condition1=20
```

2. Mientras ejecuta el trabajo por medio de Job Executor en el símbolo del sistema, use el argumento `-o` seguido de la ruta al archivo de texto creado.
Por ejemplo, para ejecutar el trabajo `ReadCustomerDataJob`, para el cual se definieron los valores predeterminados de sus variables en el archivo de texto `variables.txt`, ejecute el siguiente comando en un símbolo del sistema:

```
java -jar jobexecutor.jar -h "localhost" -u "admin" -p "admin" -s "8080" -j "ReadCustomerDataJob" -o "variables.txt"
```

Durante la ejecución del trabajo mediante Job Executor, se emplea la consulta personalizada para obtener los datos necesarios.

Nota: Para obtener instrucciones y sintaxis de línea de comandos, consulte *Ejecución de un trabajo* en la Guía del diseñador de flujos de datos.

Read From File

La etapa Read from File (Lectura desde archivo) especifica un archivo de entrada para un trabajo o subflujo. No está disponible para servicios.

Nota: Si desea utilizar un archivo XML como entrada de su flujo de datos, utiliza la etapa Read from XML en lugar de la etapa Read from File. Si desea utilizar un archivo de formato variable, use Read from Variable Format File.

Ficha Propiedades del archivo

Nombre de campo	Descripción
Server name (Nombre de servidor)	Indica si el archivo que seleccionó como entrada está ubicado en la computadora que está ejecutando Enterprise Designer o en el servidor de Spectrum™ Technology Platform. Si selecciona un archivo en la computadora local, el nombre del servidor será Mi computadora. Si selecciona un archivo que se encuentra en el servidor, el nombre del servidor será Spectrum™ Technology Platform.

Nombre de campo	Descripción
Nombre de archivo	<p>Especifica la ruta al archivo. Haga clic en el botón de puntos suspensivos (...) para dirigirse al archivo que desea.</p> <p>Puede leer múltiples archivos al usar el carácter comodín para leer datos desde múltiples archivos en el directorio. Los caracteres comodines * y ? son compatibles. Por ejemplo, puede especificar *.csv para leer todos los archivos con extensión .csv que se encuentran en el directorio. Para leer correctamente múltiples archivos, cada uno debe tener la misma disposición (los mismos campos en las mismas posiciones). Cualquier registro que no coincida con la disposición especificada en la ficha Campos se tratará como un registro malformado.</p> <p>Cuando lee un archivo del servidor de archivos HDFS, los formatos de compresión compatibles son:</p> <ol style="list-style-type: none"> 1. GZIP (.gz) 2. BZIP2 (.bz2) <p>Nota: La extensión del archivo indica el formato de compresión que se utiliza para descomprimir el archivo.</p> <p>Atención: Si el servidor de Spectrum™ Technology Platform está ejecutándose en Unix o Linux, recuerde que estas plataformas distinguen las mayúsculas de las minúsculas en los nombres de los archivos y las rutas.</p>
Tipo de registro	<p>Formato de los registros en el archivo. Seleccione uno de:</p> <p>En líneas secuenciales Archivo de texto en el que los registros se separan con un carácter de final de línea (EOL), como un retorno de carro o un salto de línea (CR o LF) y cada campo tiene una posición fija para los caracteres de inicio y final.</p> <p>De ancho fijo Archivo de texto en el que cada registro es un número específico de caracteres de longitud y cada campo tiene una posición fija para los caracteres de inicio y final.</p> <p>Delimitado Archivo de texto en el que los registros se separan con un carácter de final de línea (EOL), como un retorno de carro o un salto de línea (CR o LF) y cada campo se separa con un carácter designado; por ejemplo, una coma.</p>

Nombre de campo	Descripción
Character encoding (Codificación de caracteres)	<p data-bbox="535 331 1430 367">Codificación del archivo de texto. Seleccione uno de estos:</p> <p data-bbox="535 380 1430 478">UTF-8 Admite todos los caracteres Unicode y es compatible con las versiones anteriores de ASCII. Para obtener más información sobre UTF, consulte unicode.org/faq/utf_bom.html.</p> <p data-bbox="535 491 1430 619">UTF-16 Admite todos los caracteres Unicode, pero no es compatible con las versiones anteriores de ASCII. Para obtener más información sobre UTF, consulte unicode.org/faq/utf_bom.html.</p> <p data-bbox="535 632 1430 697">US-ASCII Codificación de caracteres basada en el orden del alfabeto inglés.</p> <p data-bbox="535 709 1430 774">UTF-16BE Codificación UTF-16 con serialización de bytes big endian (byte más significativo al principio).</p> <p data-bbox="535 787 1430 852">UTF-16LE Codificación UTF-16 con serialización de bytes little endian (byte menos significativo al principio).</p> <p data-bbox="535 865 1430 963">ISO-8859-1 Codificación de caracteres ASCII, que suele utilizarse para los idiomas de Europa occidental. También se la conoce como Latin-1.</p> <p data-bbox="535 976 1430 1075">ISO-8859-3 Codificación de caracteres ASCII, que suele utilizarse para los idiomas de Europa meridional. También se la conoce como Latin-3.</p> <p data-bbox="535 1087 1430 1152">ISO-8859-9 Codificación de caracteres ASCII, que suele utilizarse para el idioma turco. También se la conoce como Latin-5.</p> <p data-bbox="535 1165 1430 1230">CP850 Página de códigos ASCII que se utiliza para escribir idiomas de Europa occidental.</p> <p data-bbox="535 1243 1430 1308">CP500 Página de códigos EBCDIC que se utiliza para escribir idiomas de Europa occidental.</p> <p data-bbox="535 1320 1430 1354">Shift_JIS Codificación de caracteres para el idioma japonés.</p> <p data-bbox="535 1367 1430 1474">MS932 Una extensión de Microsoft de Shift_JIS para incluir caracteres especiales NEC, selección NEC de extensiones IBM y extensiones IBM.</p> <p data-bbox="535 1486 1430 1554">CP1047 Una página de códigos EBCDIC con el conjunto de caracteres completo del alfabeto latino 1.</p>

Nombre de campo	Descripción
Separador de campo	<p data-bbox="535 325 1433 388">Especifica el carácter que se utiliza para separar los campos en un archivo delimitado.</p> <p data-bbox="535 409 1433 441">Por ejemplo, este registro utiliza una barra vertical () como separador de campo:</p> <pre data-bbox="552 462 1421 514">7200 13TH ST MIAMI FL 33144</pre> <p data-bbox="535 535 1433 567">Estos caracteres disponibles para definir como separadores de campo son:</p> <ul data-bbox="552 588 779 787" style="list-style-type: none">• Espacio de claves• Tab• Coma• Punto• Punto y coma• Barra vertical <p data-bbox="535 808 1433 871">Si el archivo utiliza un carácter distinto como separador de campo, haga clic en el botón de los puntos suspensivos para seleccionar otro carácter como delimitador.</p>
Text qualifier (Calificador de texto)	<p data-bbox="535 934 1433 976">El carácter que se utiliza para demarcar los valores de texto en un archivo delimitado.</p> <p data-bbox="535 987 1433 1018">Por ejemplo, este registro utiliza comillas dobles (") como calificador de texto.</p> <pre data-bbox="552 1039 1421 1092">"7200 13TH ST" "MIAMI" "FL" "33144"</pre> <p data-bbox="535 1113 1433 1144">Los caracteres disponibles para definir como calificadores de texto son:</p> <ul data-bbox="552 1165 779 1228" style="list-style-type: none">• Comilla simple (')• Comilla doble (") <p data-bbox="535 1249 1433 1312">Si el archivo utiliza un calificador de texto distinto, haga clic en el botón de los puntos suspensivos para seleccionar otro carácter como calificador de texto.</p>

Nombre de campo	Descripción
Record separator (Separador de registro)	<p>Especifica el carácter que se utiliza para separar registros en archivos secuenciales o delimitados. Este campo no está disponible si selecciona la casilla Utilizar final de línea predeterminado .</p> <p>Los parámetros de separador de registro disponibles son:</p> <p>Unix (U+000A) Un carácter de salto de línea separa los registros. Este es el separador de registros estándar para los sistemas Unix.</p> <p>Macintosh (U+000D) Un carácter de retorno de carro separa los registros. Este es el separador de registros estándar para los sistemas Macintosh.</p> <p>Windows (U+000D U+000A) Un carácter de retorno de carro seguido por un salto de línea separa los registros. Este es el separador de registros estándar para los sistemas Windows.</p> <p>Si su archivo utiliza un separador de registro distinto, haga clic en el botón de los puntos suspensivos para seleccionar otro carácter como separador de registro.</p>
Use default EOL (Utilizar final de línea predeterminado)	<p>Especifica que el separador de registro del archivo es el carácter de final de línea (EOL) predeterminado que utiliza el sistema operativo en el que se ejecuta el servidor Spectrum™ Technology Platform.</p> <p>No seleccione esta opción si el archivo utiliza un carácter de final de línea distinto del que utiliza de forma predeterminada el sistema operativo del servidor. Por ejemplo, si el archivo utiliza un final de línea de Windows pero el servidor está ejecutándose en Linux, no seleccione esta opción. En su lugar, seleccione la opción Windows en el campo Separador de registro.</p>
Record length (Longitud de registro)	<p>Para los archivos de ancho fijo, especifica la cantidad exacta de caracteres en cada registro.</p> <p>Para los archivos secuenciales, especifica la longitud, en caracteres, del registro más largo del archivo.</p>
First row is header record (La primera fila es el registro de encabezado)	<p>Especifica si el primer registro de un archivo delimitado contiene información del encabezado pero no datos.</p> <p>Por ejemplo, este fragmento de archivo muestra una fila de encabezado en el primer registro.</p> <pre>"AddressLine1" "City" "StateProvince" "PostalCode" "7200 13TH ST" "MIAMI" "FL" "33144" "One Global View" "Troy" "NY" "12180"</pre>

Nombre de campo	Descripción
Considerar malformados los registros con menos campos que los especificados	Se considerarán malformados los registros de archivos delimitados que contengan menos campos que los especificados en la ficha Campos .
Importar	Importa la definición de diseño de un archivo, los ajustes de codificación y las opciones de ordenamiento desde un archivo de configuración. El archivo de configuración se crea al exportar la configuración de otra etapa Read from File o Write to File que utilizó el mismo archivo de entrada o un archivo con el mismo diseño que aquel en el que usted está trabajando.
Exportar	<p>Guarda la definición de diseño de un archivo, los ajustes de codificación y las opciones de ordenamiento en un archivo de configuración. Luego, puede importar estas configuraciones a otra etapa Read from File o Write from File que utilice el mismo archivo de entrada o un archivo que tenga las mismas características que aquel en que usted esté trabajando actualmente. También puede usar el archivo de configuración con el ejecutor de trabajos para especificar la configuración del archivo durante la ejecución.</p> <p>Para obtener información acerca de los comandos disponibles, consulte El archivo de configuración de definición de archivos en la página 165.</p>

Ficha Campos

La ficha Campos define los nombres, posiciones y, para los archivos secuenciales y de ancho fijo, la longitud de los campos en el archivo. Para obtener más información, consulte estos temas:

[Definición de los campos en un archivo de entrada delimitado](#) en la página 158

[Definición de los campos en un archivo secuencial o de ancho fijo](#) en la página 161

Ficha Ordenar campos

La ficha Ordenar campos define los campos según los cuales se ordenan los registros de entrada antes de ser enviados al flujo de datos. La ordenación es opcional. Para obtener más información, consulte [Ordenación de los registros de entrada](#) en la página 163.

Ficha Tiempo de ejecución

Nombre de campo	Descripción
Nombre de archivo	Muestra el nombre del archivo seleccionado en la primera ficha.

Nombre de campo	Descripción
Starting record (Registro de inicio)	Si desea omitir los registros ubicados al comienzo del archivo cuando lee registros en el flujo de datos, especifique el primer registro que desea leer. Por ejemplo, si desea omitir los primeros 50 registros de un archivo, especifique 51. El registro número 51 será el primero que se leerá en el flujo de datos.
Todos los registros	Seleccione esta opción si desea leer todos los registros, comenzando desde el registro especificado en el campo Registro de inicio hasta el final del archivo.
Máximo de registros	Seleccione esta opción si desea leer solamente cierta cantidad de registros, comenzando desde el registro especificado en el campo Registro de inicio . Por ejemplo, si desea leer los primeros 100 registros, seleccione esta opción e ingrese 100.

Definición de los campos en un archivo de entrada delimitado

La ficha **Campos** define los nombres, las posiciones y, para algunos tipos de archivo, las longitudes de los campos del archivo. Una vez que haya definido un archivo de entrada en la ficha **Propiedades del archivo**, podrá definir los campos.

Si el archivo de entrada no contiene un registro de encabezado o si desea definir los campos manualmente, siga estos pasos en la ficha **Campos**:

1. Para definir los campos que ya existen en el archivo de entrada, haga clic en **Regenerar**. Después haga clic en **Detectar tipo**. El tipo de datos de cada campo se configurará automáticamente de acuerdo con los primeros 50 registros del archivo.
2. Para agregar campos adicionales en la salida, haga clic en **Agregar**.
3. En el campo **Nombre**, elija el campo que desea agregar o ingrese el nombre del campo.
4. En el campo **Tipo** puede dejar el tipo de datos como `string` (cadena) si no pretende realizar ninguna operación matemática o de fecha y hora con los datos. Sin embargo, si se propone realizar ese tipo de operaciones, seleccione un tipo de datos apropiado. Esto convertirá los datos de cadena del archivo en un tipo de datos que permitirá el manejo adecuado de los datos en el flujo de datos.

Spectrum™ Technology Platform admite estos tipos de datos:

- bigdecimal** Un tipo de dato numérico que admite 38 puntos decimales de precisión. Utilice este tipo de datos para datos que se usarán en cálculos matemáticos que requieren un alto grado de precisión, especialmente aquellos que implican datos financieros. El tipo de datos `bigdecimal` es compatible con cálculos más precisos que el tipo de datos `double`.

boolean	Tipo de dato lógico con dos valores: verdadero y falso
bytearray	Una matriz (lista) de bytes. Nota: Bytearray no se admite como datos de entrada para un servicio REST.
fecha	Un tipo de dato que contiene un mes, día y año. Por ejemplo, 2012-01-30 o 30 de enero, 2012. Puede especificar un formato de fecha predeterminado en Management Console.
fecha y hora	Un tipo de dato que contiene un mes, día, año y horas, minutos y segundos. Por ejemplo, 2012/01/30 6:15 PM.
double	Un tipo de dato numérico que contiene números de doble precisión tanto negativos como positivos entre 2^{-1074} y $(2-2^{-52}) \times 2^{1023}$. En notación científica, el rango de valores es - 1.79769313486232E+308 a 1.79769313486232E+308.
float	Un tipo de dato numérico que contiene números de precisión simple tanto negativos como positivos entre 2^{-149} y $(2-2^{-23}) \times 2^{127}$. En notación científica, el rango de valores es - 3.402823E+38 a 3.402823E+38.
integer	Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{31} (-2,147,483,648) y $2^{31}-1$ (2,147,483,647).
list	En sentido estricto, una lista no es un tipo de dato. Sin embargo, cuando un campo contiene datos jerárquicos, es tratado como un campo "lista". En Spectrum™ Technology Platform una lista es una colección de datos que consiste en múltiples valores. Por ejemplo, un campo Nombres puede contener una lista de valores de nombre, lo que puede representarse mediante una estructura XML de la siguiente forma: <pre><Names> <Name>John Smith</Name> <Name>Ann Fowler</Name> </Names></pre> <p>Cabe destacar que el tipo de dato lista de Spectrum™ Technology Platform es diferente del tipo de dato lista del esquema XML puesto que este último es un tipo de dato simple que está compuesto de valores múltiples, mientras que el tipo de dato lista de Spectrum™ Technology Platform es similar a un tipo de dato complejo del esquema XML.</p>
long	Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{63} (-9,223,372,036,854,775,808) y $2^{63}-1$ (9,223,372,036,854,775,807).
string	Secuencia de caracteres.
hora	Un tipo de dato que contiene la hora del día. Por ejemplo, 21:15:59 o 9:15:59 PM.

5. Si ha seleccionado un tipo de datos de fecha, hora, o numérico, puede usar el formato de fecha/hora predeterminado o puede indicar un formato diferente para este campo específico. El formato predeterminado es, o bien el formato de sistema predeterminado que se haya definido en las opciones de conversión de tipos en Management Console, o bien es el formato predeterminado del flujo de datos especificado en la opciones de conversión de tipos de Enterprise Designer. Se mostrará el formato que se esté usando. Para usar el formato predeterminado, deje seleccionado **Predeterminado**. Para indicar un formato diferente, seleccione **Personalizado** y siga los siguientes pasos:

Nota: Es importante que seleccione un formato de fecha y hora que refleje con exactitud los datos que está leyendo desde el archivo. Por ejemplo, si el archivo contiene datos con el formato de Mes/Día/Año, pero usted selecciona Día/Mes/Año, todos los cálculos de fecha que usted realice en el flujo de datos, como por ejemplo, clasificar según fecha, no reflejarán la fecha correcta. Por otro lado, los registros pueden fallar en el tipo de conversión, en cuyo caso se hará efectivo el comportamiento ante errores que se especifique en las opciones de conversión de tipos en la Consola de administración o en Enterprise Designer.

- a) En el campo **Configuración regional** seleccione el país cuya convención acerca de formatos usted desea utilizar. Su selección determinará los valores predeterminados en el campo **Formato**. En cuanto a los datos referidos a fechas, su selección determinará además el idioma que se usará al enunciar un mes. Por ejemplo, si ha indicado Inglés, el primer mes del año será "January" pero si ha especificado Francés, este será "Janvier."
- b) En el campo **Formato**, seleccione el formato de los datos. El formato depende del tipo de datos en el campo. Se proporciona una lista de los formatos de uso más común en la configuración regional seleccionada.

A la derecha del campo **Formato** se muestra un ejemplo del formato seleccionado.

Además puede especificar sus propios formatos de fecha, hora, y número en caso de que los que están disponibles para seleccionar no se ajusten a sus necesidades. Para especificar su propio formato de fecha u hora, ingrese el formato en el campo usando la notación que se describe en [Patrones de fecha y hora](#) en la página 312. Para especificar su propio formato de número, ingrese el formato en el archivo usando la notación que se describe en [Patrones de números](#) en la página 315.

6. En el campo **Posición**, ingrese la posición de este campo dentro del registro.

Por ejemplo, en este archivo de entrada, AddressLine1 está en posición 1; City, en posición 2; StateProvince, en posición 3; y PostalCode, en posición 4.

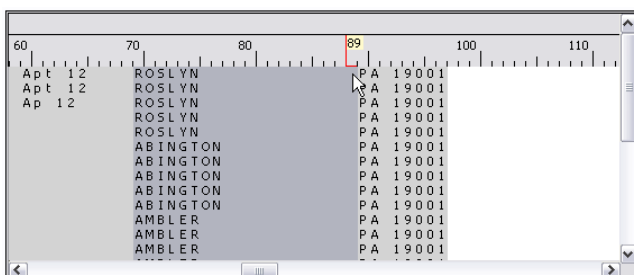
```
"AddressLine1"|"City"|"StateProvince"|"PostalCode"
"7200 13TH ST"|"MIAMI"|"FL"|"33144"
"One Global View"|"Troy"|"NY"|12180
```

7. Si desea eliminar los caracteres que ocupan demasiado espacio al principio y al final de la cadena de valores de un campo, seleccione la casilla **Recortar**.

Definición de los campos en un archivo secuencial o de ancho fijo

En la etapa Read from File, la ficha **Campos** define los nombres, las posiciones y, para algunos tipos de archivos, las longitudes de los campos en el archivo. Una vez que haya definido un archivo de entrada en la ficha **Propiedades del archivo**, podrá definir los campos.

1. En la ficha **Campos**, en **Vista previa**, haga clic al comienzo de un campo y arrastre hacia la izquierda para resaltar el campo deseado, como se muestra a continuación:



2. En el campo **Nombre**, ingrese el campo que desea agregar.
3. En el campo **Tipo** puede dejar el tipo de datos como cadena si no se ha propuesto realizar ninguna operación matemática o de fecha y hora con los datos. Sin embargo, si se propone realizar ese tipo de operaciones, seleccione un tipo de datos apropiado. Esto convertirá los datos de cadena del archivo en un tipo de datos que permitirá el manejo adecuado de los datos en el flujo de datos.

Spectrum™ Technology Platform admite los siguientes tipos de datos:

bigdecimal Un tipo de dato numérico que admite 38 puntos decimales de precisión. Utilice este tipo de datos para datos que se usarán en cálculos matemáticos que requieren un alto grado de precisión, especialmente aquellos que implican datos financieros. El tipo de datos bigdecimal es compatible con cálculos más precisos que el tipo de datos double.

boolean Tipo de dato lógico con dos valores: verdadero y falso

bytearray Una matriz (lista) de bytes.

Nota: Bytearray no se admite como datos de entrada para un servicio REST.

fecha Un tipo de dato que contiene un mes, día y año. Por ejemplo, 2012-01-30 o 30 de enero, 2012. Puede especificar un formato de fecha predeterminado en Management Console.

fecha y hora Un tipo de dato que contiene un mes, día, año y horas, minutos y segundos. Por ejemplo, 2012/01/30 6:15 PM.

double	Un tipo de dato numérico que contiene números de doble precisión tanto negativos como positivos entre 2^{-1074} y $(2-2^{-52}) \times 2^{1023}$. En notación científica, el rango de valores es - 1.79769313486232E+308 a 1.79769313486232E+308.
float	Un tipo de dato numérico que contiene números de precisión simple tanto negativos como positivos entre 2^{-149} y $(2-2^{-23}) \times 2^{127}$. En notación científica, el rango de valores es - 3.402823E+38 a 3.402823E+38.
integer	Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{31} (-2,147,483,648) y $2^{31}-1$ (2,147,483,647).
list	En sentido estricto, una lista no es un tipo de dato. Sin embargo, cuando un campo contiene datos jerárquicos, es tratado como un campo "lista". En Spectrum™ Technology Platform una lista es una colección de datos que consiste en múltiples valores. Por ejemplo, un campo Nombres puede contener una lista de valores de nombre, lo que puede representarse mediante una estructura XML de la siguiente forma: <div data-bbox="423 814 1424 970" data-label="Code-Block"> <pre><Names> <Name>John Smith</Name> <Name>Ann Fowler</Name> </Names></pre> </div> <p>Cabe destacar que el tipo de dato lista de Spectrum™ Technology Platform es diferente del tipo de dato lista del esquema XML puesto que este último es un tipo de dato simple que está compuesto de valores múltiples, mientras que el tipo de dato lista de Spectrum™ Technology Platform es similar a un tipo de dato complejo del esquema XML.</p>
long	Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{63} (-9,223,372,036,854,775,808) y $2^{63}-1$ (9,223,372,036,854,775,807).
string	Secuencia de caracteres.
hora	Un tipo de dato que contiene la hora del día. Por ejemplo, 21:15:59 o 9:15:59 PM.

- Para aplicar el formato de almacenamiento `packed decimal` al campo, marque la casilla de verificación **Decimal empaquetado**. El tipo `packed decimal` usa 4 bits, en comparación con el tipo `integer`, que usa 8 bits.

Nota: Este formato de almacenamiento está disponible solo al seleccionar los tipos de datos `double`, `integer` y `long` mientras se leen los archivos secuenciales y de ancho fijo.

- Si ha seleccionado un tipo de datos de fecha, hora, o numérico, puede usar el formato de fecha/hora predeterminado o puede indicar un formato diferente para este campo específico. El formato predeterminado es, o bien el formato de sistema predeterminado que se haya definido en las opciones de conversión de tipos en Management Console, o bien es el formato predeterminado del flujo de datos especificado en la opciones de conversión de tipos de

Enterprise Designer. Se mostrará el formato que se esté usando. Para usar el formato predeterminado, deje seleccionado **Predeterminado**. Para indicar un formato diferente, seleccione **Personalizado** y siga los siguientes pasos:

Nota: Es importante que seleccione un formato de fecha y hora que refleje con exactitud los datos que está leyendo desde el archivo. Por ejemplo, si el archivo contiene datos con el formato de Mes/Día/Año, pero usted selecciona Día/Mes/Año, todos los cálculos de fecha que usted realice en el flujo de datos, como por ejemplo, clasificar según fecha, no reflejarán la fecha correcta. Por otro lado, los registros pueden fallar en el tipo de conversión, en cuyo caso se hará efectivo el comportamiento ante errores que se especifique en las opciones de conversión de tipos en la Consola de administración o en Enterprise Designer.

- a) En el campo **Configuración regional** seleccione el país cuya convención acerca de formatos usted desea utilizar. Su selección determinará los valores predeterminados en el campo **Formato**. En cuanto a los datos referidos a fechas, su selección determinará además el idioma que se usará al enunciar un mes. Por ejemplo, si ha indicado Inglés, el primer mes del año será "January" pero si ha especificado Francés, este será "Janvier."
- b) En el campo **Formato**, seleccione el formato de los datos. El formato depende del tipo de datos en el campo. Se proporciona una lista de los formatos de uso más común en la configuración regional seleccionada.

A la derecha del campo **Formato** se muestra un ejemplo del formato seleccionado.

Además puede especificar sus propios formatos de fecha, hora, y número en caso de que los que están disponibles para seleccionar no se ajusten a sus necesidades. Para especificar su propio formato de fecha u hora, ingrese el formato en el campo usando la notación que se describe en [Patrones de fecha y hora](#) en la página 312. Para especificar su propio formato de número, ingrese el formato en el archivo usando la notación que se describe en [Patrones de números](#) en la página 315.

6. Los campos **Posición de inicio** y **Longitud** se completan automáticamente sobre la base de la selección que usted hizo en la vista previa del archivo.
7. Si desea eliminar los caracteres que ocupan demasiado espacio al principio y al final de la cadena de caracteres de un campo, seleccione la casilla **Recortar espacios**.
8. Haga clic en **Aceptar**.

Ordenación de los registros de entrada

En la etapa Read from File, la ficha **Ordenar campos** define los campos según los cuales se ordenan los registros de entrada antes de ser enviados al flujo de datos. La ordenación es opcional.

1. En la ficha **Ordenar campos**, haga clic en **Agregar**.

2. Haga clic en la flecha desplegable de la columna **Nombre del campo** y seleccione el campo según el cual quiere ordenar. Los campos que pueden seleccionarse dependen de los campos definidos en este archivo de entrada.
3. En la columna **Criterio de ordenación**, seleccione Ascending o Descending.
4. Repita hasta haber agregado todos los campos de entrada que desea utilizar para la ordenación. Cambie el criterio de ordenación destacando la fila del campo que desea desplazar y haciendo clic en **Arriba** o **Abajo**.
5. Las opciones predeterminadas de rendimiento de orden para su sistema se definen en Management Console. Si desea anular las opciones de rendimiento de orden predeterminadas de su sistema, haga clic en **Opciones avanzadas**. El cuadro de diálogo **Opciones avanzadas** contiene las siguientes opciones de rendimiento de orden:

Límite de registros en memoria

Especifica la cantidad máxima de filas de datos que un ordenador guardará en la memoria antes de empezar la paginación al disco. De forma predeterminada, la ordenación de 10.000 registros o menos se realizará en la memoria y la de más de 10.000 registros se realizará como una ordenación en disco. El límite máximo es de 100.000 registros. Normalmente, la ordenación que se realiza en la memoria es más rápida que la que se realiza en disco; por lo tanto, este valor se debe establecer lo suficientemente alto para que la mayor parte de la ordenación sea en la memoria y solo los conjuntos de datos grandes se escriban en el disco.

Nota: Tenga cuidado en entornos en los que se ejecuten trabajos de forma concurrente, porque aumentar la configuración del **Límite de registros en memoria** aumenta las posibilidades de agotar la memoria.

Número máximo de archivos temporales

Especifica el número máximo de archivos temporales que puede utilizar el proceso de orden. El uso de una gran cantidad de archivos temporales puede producir un mejor rendimiento. Sin embargo, la cantidad óptima depende en gran medida de la configuración en el servidor que ejecuta Spectrum™ Technology Platform. Debe experimentar distintas configuraciones y observar el efecto en el rendimiento de usar una mayor o menor cantidad de archivos temporales. Para calcular la cantidad aproximada de archivos temporales que se pueda necesitar, use la siguiente ecuación:

$$\frac{(NumberOfRecords \times 2)}{InMemoryRecordLimit} = NumberOfTempFiles$$

Tenga en cuenta que la cantidad máxima de archivos temporales no puede ser superior a 1.000.

Activar compresión

Especifica que los archivos temporales se comprimen cuando se los escribe en el disco.

Nota: La configuración del rendimiento óptimo de orden depende de la configuración del hardware de su servidor. No obstante, la siguiente ecuación generalmente produce un buen rendimiento de orden:

$$(InMemoryRecordLimit \times MaxNumberOfTempFiles \div 2) \geq TotalNumberOfRecords$$

El archivo de configuración de definición de archivos

Un archivo de configuración de definición de archivos contiene el diseño de archivo, la codificación y las opciones de orden que se han exportado de una etapa Read from File o Write to File. Este archivo puede importarse a Read from File o Write to File para configurar rápidamente las opciones de la etapa en lugar de tener que especificar las opciones manualmente.

La forma más fácil de crear un archivo de configuración de definición de archivos es usar Read from File o Write to File para especificar la configuración del archivo, y luego hacer clic en el botón **Exportar** para generar el archivo.

No obstante, a título informativo, a continuación se muestra el esquema del archivo de configuración de definición de archivos. Cada elemento de un archivo XML tiene un tipo de dato, y, si ese tipo no es string o integer, se muestran los valores aceptables. Estos valores se corresponden directamente con las opciones del cuadro de diálogo de la etapa. Por ejemplo, el elemento FileTypeEnum corresponde al campo Tipo de registro de la ficha Propiedades del archivo, y los siguientes tres valores se muestran en el esquema: linesequential, fixedwidth, y delimited.

Nota: Si usted ingresa "custom" en los campos LineSeparator, FieldSeparator o TextQualifier, también debe incluir un elemento custom correspondiente (por ejemplo, "CustomLineSeparator", "CustomFieldSeparator" o "CustomTextQualifier") con un número hexadecimal que represente el carácter, o secuencia de caracteres, que se utilizará.

```
<?xml version="1.0" encoding="utf-8"?>
<xs:schema elementFormDefault="qualified"
  xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="FileSchema" nillable="true" type="FileSchema"/>
  <xs:complexType name="FileSchema">
    <xs:sequence>
      <xs:element
        minOccurs="0"
        maxOccurs="1"
        default="linesequential"
        name="Type"
        type="FileTypeEnum"/>
      <xs:element
        minOccurs="0"
        maxOccurs="1"
        default="UTF-8" name="Encoding" type="xs:string"/>
    </xs:sequence>
  </xs:complexType>
</xs:schema>
```

```

<xs:element
  minOccurs="0"
  maxOccurs="1"
  name="RecordLength"
  type="xs:int"/>
<xs:element
  minOccurs="0"
  maxOccurs="1"
  default="default"
  name="LineSeparator"
  type="LineSeparatorEnum"/>
<xs:element
  minOccurs="0"
  maxOccurs="1"
  name="CustomLineSeparator"
  type="xs:string"/>
<xs:element
  minOccurs="0"
  maxOccurs="1"
  default="comma"
  name="FieldSeparator"
  type="FieldSeparatorEnum"/>
<xs:element
  minOccurs="0"
  maxOccurs="1"
  name="CustomFieldSeparator"
  type="xs:string"/>
<xs:element
  minOccurs="0"
  maxOccurs="1"
  default="none"
  name="TextQualifier"
  type="TextQualifierEnum"/>
<xs:element
  minOccurs="0"
  maxOccurs="1"
  name="CustomTextQualifier"
  type="xs:string"/>
<xs:element
  minOccurs="0"
  maxOccurs="1"
  default="false"
  name="HasHeader"
  type="xs:boolean"/>
<xs:element
  minOccurs="0"
  maxOccurs="1"
  default="true"
  name="EnforceColumnCount"
  type="xs:boolean"/>
<xs:element
  minOccurs="0"
  maxOccurs="1"

```

```

        name="Fields"
        type="ArrayOfFieldSchema"/>
    </xs:sequence>
</xs:complexType>
<xs:simpleType name="FileTypeEnum">
    <xs:restriction base="xs:string">
        <xs:enumeration value="linesequential"/>
        <xs:enumeration value="fixedwidth"/>
        <xs:enumeration value="delimited"/>
    </xs:restriction>
</xs:simpleType>
<xs:simpleType name="LineSeparatorEnum">
    <xs:restriction base="xs:string">
        <xs:enumeration value="default"/>
        <xs:enumeration value="windows"/>
        <xs:enumeration value="unix"/>
        <xs:enumeration value="mac"/>
        <xs:enumeration value="custom"/>
    </xs:restriction>
</xs:simpleType>
<xs:simpleType name="FieldSeparatorEnum">
    <xs:restriction base="xs:string">
        <xs:enumeration value="comma"/>
        <xs:enumeration value="tab"/>
        <xs:enumeration value="space"/>
        <xs:enumeration value="semicolon"/>
        <xs:enumeration value="period"/>
        <xs:enumeration value="pipe"/>
        <xs:enumeration value="custom"/>
    </xs:restriction>
</xs:simpleType>
<xs:simpleType name="TextQualifierEnum">
    <xs:restriction base="xs:string">
        <xs:enumeration value="none"/>
        <xs:enumeration value="single"/>
        <xs:enumeration value="double"/>
        <xs:enumeration value="custom"/>
    </xs:restriction>
</xs:simpleType>
<xs:complexType name="ArrayOfFieldSchema">
    <xs:sequence>
        <xs:element
            minOccurs="0"
            maxOccurs="unbounded"
            name="Field"
            nillable="true"
            type="FieldSchema"/>
    </xs:sequence>
</xs:complexType>
<xs:complexType name="FieldSchema">
    <xs:sequence>
        <xs:element
            minOccurs="0"

```

```

        maxOccurs="1"
        name="Name"
        type="xs:string"/>
<xs:element
  minOccurs="0"
  maxOccurs="1"
  default="string"
  name="Type"
  type="xs:string"/>
<xs:element
  minOccurs="1"
  maxOccurs="1"
  name="Position"
  type="xs:int"/>
<xs:element
  minOccurs="0"
  maxOccurs="1"
  name="Length"
  type="xs:int"/>
<xs:element
  minOccurs="0"
  maxOccurs="1"
  default="false"
  name="Trim"
  type="xs:boolean"/>
<xs:element
  minOccurs="0"
  maxOccurs="1"
  name="Locale"
  type="Locale"/>
<xs:element
  minOccurs="0"
  maxOccurs="1"
  name="Pattern"
  type="xs:string"/>
<xs:element
  minOccurs="0"
  maxOccurs="1"
  default="none"
  name="Order"
  type="SortOrderEnum"/>
</xs:sequence>
</xs:complexType>
<xs:complexType name="Locale">
  <xs:sequence>
    <xs:element
      minOccurs="0"
      maxOccurs="1"
      name="Country"
      type="xs:string"/>
    <xs:element
      minOccurs="0"
      maxOccurs="1"

```



```

        name="Language"
        type="xs:string"/>
    <xs:element
        minOccurs="0"
        maxOccurs="1"
        name="Variant"
        type="xs:string"/>
    </xs:sequence>
</xs:complexType>
<xs:simpleType name="SortOrderEnum">
    <xs:restriction base="xs:string">
        <xs:enumeration value="none"/>
        <xs:enumeration value="ascending"/>
        <xs:enumeration value="descending"/>
    </xs:restriction>
</xs:simpleType>
</xs:schema>

```

Configuración de opciones de flujo de datos

Este procedimiento describe cómo configurar un flujo de datos compatible con las opciones de tiempo de ejecución para la etapa Read From File.

1. Abra el flujo de datos en Enterprise Designer.
2. Si desea configurar las opciones de tiempo de ejecución para una etapa en un flujo de datos incrustado, abra este último.
3. Haga clic en el ícono Opciones de flujo de datos en la barra de herramientas o haga clic en **Editar > Opciones de flujo de datos**. Aparecerá el cuadro de diálogo **Opciones de flujo de datos**.
4. Haga clic en **Agregar**. Aparecerá el cuadro de diálogo **Definir opción de flujo de datos**.
5. Expanda la etapa **Read from File**.

Las opciones de flujo de datos que se exponen son:

1. Codificación de caracteres
 2. Separador de campo
 3. Calificador de texto
 4. Longitud de registro
 5. La primera fila es el registro de encabezado
 6. Registro de inicio
 7. Máximo de registros
6. El nombre de la opción Read From File seleccionada se muestra en los campos **Nombre de opción** y **Etiqueta de opción**. Este es el nombre de opción que deberá especificarse durante el tiempo de ejecución para definir esta opción.
 7. Ingrese una descripción de la opción en el campo **Descripción**.

8. En el campo **Destino**, seleccione la opción **Etapas seleccionadas**.
9. Si desea limitar los valores que pueden especificarse durante la ejecución, edite las opciones en el campo **Valores válidos** haciendo clic en el icono a la derecha del campo.
10. Si desea cambiar el valor predeterminado, especifique un valor diferente en el campo **Valor predeterminado**.

Nota: En el caso de un servicio, solo puede modificar los valores predeterminados antes de exponer el servicio por primera vez. Después de exponer el servicio, ya no es posible modificar los valores predeterminados con Enterprise Designer. En cambio, debe usar Management Console. Para obtener más información, consulte [Especificación de las opciones de servicio predeterminadas](#).

11. Haga clic en **Aceptar**.
12. Continúe agregando opciones, según sus preferencias.
13. Haga clic en **ACEPTAR** en el cuadro de diálogo **Opciones de flujo de datos** cuando haya terminado de agregar opciones.
14. Guarde y exponga el flujo de datos.

Reglas de las opciones de flujo de datos

1. *Codificación de carácter:* se aceptan todos los tipos de codificación que son válidos para la JVM subyacente. Esta opción no puede estar vacía.
2. *Separador de campo:* se acepta cualquier delimitador de carácter individual. Actualmente, no se admiten los valores ni los espacios HEX.
3. *Calificador de texto:* se acepta cualquier calificador de carácter individual. No se admiten los valores HEX.
4. *Longitud de registro:* solo se aceptan números enteros. No puede haber espacios en blanco ni caracteres no numéricos.
5. *Registro de inicio:* solo se aceptan números enteros. No pueden ser no numéricos.
6. *Máximo de registros:* solo se aceptan números enteros. No pueden ser no numéricos.
7. *La primera fila es el encabezado:* solo se aceptan los valores booleanos de `verdadero` y `falso`. No puede quedar vacío.

Read from Hadoop Sequence File

La etapa Read from Hadoop Sequence File lee datos desde un archivo de secuencia como entrada a un flujo de datos. Un archivo de secuencia es un archivo simple que consta de pares de clave/valor binarios. Para obtener más información, visite <http://wiki.apache.org/hadoop/SequenceFile>.

Nota: La etapa Read from Hadoop Sequence File solo admite archivos delimitados sin comprimir, ubicados en el Sistema de archivos distribuido Hadoop (Hadoop Distributed File System, HDFS).

Ficha Propiedades del archivo

Campos	Descripción
Servidor	Indica que el archivo que seleccionó en el campo Nombre de archivo está ubicado en el sistema Hadoop. Debe crear una conexión al servidor de archivos Hadoop en Management Console, antes de usarlo en esta etapa. Si selecciona un archivo en el sistema Hadoop, el nombre del servidor será el nombre que especifique en la Management Console mientras se crea un servidor de archivos.
Nombre de archivo	Especifica la ruta al archivo. Haga clic en el botón de los puntos suspensivos (...) para buscar el archivo que desea.
Separador de campo	<p>Especifica el carácter que se utiliza para separar los campos en un archivo delimitado. Por ejemplo, este registro utiliza una barra vertical () como separador de campo:</p> <pre>7200 13TH ST MIAMI FL 33144</pre> <p>Estos caracteres disponibles para definir como separadores de campo son:</p> <ul style="list-style-type: none"> • Espacio de claves • Tab • Coma • Punto • Punto y coma • Barra vertical <p>Si el archivo utiliza un carácter distinto como separador de campo, haga clic en el botón de los puntos suspensivos para seleccionar otro carácter como delimitador.</p>
Text qualifier (Calificador de texto)	<p>El carácter que se utiliza para demarcar los valores de texto en un archivo delimitado. Por ejemplo, este registro utiliza comillas dobles (") como calificador de texto.</p> <pre>"7200 13TH ST" "MIAMI" "FL" "33144"</pre> <p>Los caracteres disponibles para definir como calificadores de texto son:</p> <ul style="list-style-type: none"> • Comilla simple (') • Comilla doble (") <p>Si el archivo utiliza un calificador de texto distinto, haga clic en el botón de los puntos suspensivos para seleccionar otro carácter como calificador de texto.</p>

Ficha Campos

La ficha Campos define los nombres, las posiciones y los tipos de campos en el archivo. Para obtener más información, consulte [Definición de campos en un archivo de secuencia de entrada](#) en la página 172.

Ficha Ordenar campos

La ficha Ordenar campos define los campos según los cuales se ordenan los registros de entrada antes de ser enviados al flujo de datos. La ordenación es opcional. Para obtener más información, consulte [Ordenación de los registros de entrada](#) en la página 173.

Ficha Filtro

La ficha Campos define los campos según los cuales se filtran los registros de entrada antes de ser enviados al flujo de datos. Para obtener más información, consulte [Filtrado de registros de entrada](#) en la página 173.

Definición de campos en un archivo de secuencia de entrada

En la etapa Read from Hadoop Sequence File, la ficha **Campos** define los nombres, las posiciones y los tipos de campos en el archivo. Una vez que haya definido un archivo de entrada en la ficha **Propiedades del archivo**, podrá definir los campos.

Si el archivo de entrada no contiene un registro de encabezado o si desea definir los campos manualmente, siga estos pasos en la ficha **Campos**:

1. Para definir los campos que ya existen en el archivo de entrada, haga clic en **Regenerar**. Después haga clic en **Detectar tipo**. El tipo de datos de cada campo se configurará automáticamente de acuerdo con los primeros 50 registros del archivo.
2. Para agregar campos adicionales en la salida, haga clic en **Agregar**.
3. En el campo **Nombre**, elija el campo que desea agregar o ingrese el nombre del campo.
4. En el campo **Tipo** puede dejar el tipo de datos como cadena, si no se ha propuesto realizar ninguna operación matemática con los datos. Sin embargo, si se propone realizar ese tipo de operaciones, seleccione un tipo de datos apropiado. Esto convertirá los datos de cadena del archivo en un tipo de datos que permitirá el manejo adecuado de los datos en el flujo de datos.

La etapa admite los siguientes tipos de datos:

double	Un tipo de dato numérico que contiene números de doble precisión tanto negativos como positivos entre 2^{-1074} y $(2-2^{-52}) \times 2^{1023}$. En notación científica, el rango de valores es - 1.79769313486232E+308 a 1.79769313486232E+308.
float	Un tipo de dato numérico que contiene números de precisión simple tanto negativos como positivos entre 2^{-149} y $(2-2^{-23}) \times 2^{127}$. En notación científica, el rango de valores es - 3.402823E+38 a 3.402823E+38.

integer	Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{31} (-2,147,483,648) y $2^{31}-1$ (2,147,483,647).
long	Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{63} (-9,223,372,036,854,775,808) y $2^{63}-1$ (9,223,372,036,854,775,807).
string	Secuencia de caracteres.

5. En el campo **Posición**, ingrese la posición de este campo dentro del registro.

Por ejemplo, en este archivo de entrada, AddressLine1 está en posición 1; City, en posición 2; StateProvince, en posición 3; y PostalCode, en posición 4.

```
"AddressLine1"|"City"|"StateProvince"|"PostalCode"
"7200 13TH ST"|"MIAMI"|"FL"|"33144"
"One Global View"|"Troy"|"NY"|12180
```

6. Si desea eliminar los caracteres que ocupan demasiado espacio al principio y al final de la cadena de valores de un campo, seleccione la casilla **Recortar**.

Ordenación de los registros de entrada

En la etapa Read from Hadoop Sequence File, la ficha **Ordenar campos** define los campos según los cuales se ordenan los registros de entrada antes de ser enviados al flujo de datos. La ordenación es opcional.

1. En Read from Hadoop Sequence File, haga clic en la ficha **Ordenar campos**.
2. En la ficha **Ordenar campos**, haga clic en **Agregar**.
3. Haga clic en la flecha desplegable de la columna **Nombre del campo** y seleccione el campo según el cual quiere ordenar. Los campos que pueden seleccionarse dependen de los campos definidos en este archivo de entrada.
4. En la columna **Criterio de ordenación**, seleccione Ascending o Descending.
5. Repita hasta haber agregado todos los campos de entrada que desea utilizar para la ordenación. Cambie el criterio de ordenación destacando la fila del campo que desea desplazar y haciendo clic en **Arriba** o **Abajo**.

Filtrado de registros de entrada

En la etapa Read from Hadoop Sequence File, en la pestaña **Filtrar**, se definen los campos según los cuales se filtran los registros de entrada antes de ser enviados al flujo de datos. El filtrado es opcional.

1. En Read from Hadoop Sequence File, haga clic en la pestaña **Filtrar**.

- En el campo **Combinar método de expresión**, seleccione **Todo** si desea que todas las expresiones se evalúen como verdaderas para que el registro se envíe a este puerto; seleccione **Cualquiera** si desea que los registros se envíen a este puerto en caso de que una o más de las expresiones sean verdaderas.
- Haga clic en **Agregar** y especifique el campo por evaluar, el operador y un valor. Los operadores se enumeran en la tabla siguiente.

Operador	Descripción
Igual	Comprueba si el valor en el campo coincide con el valor especificado.
No es igual	Comprueba si el valor del campo no coincide con el valor especificado.
Es mayor que	Comprueba si el campo tiene un valor numérico superior al valor especificado. Este operador funciona en tipos de datos numéricos y en campos que contienen números.
Es superior o igual a	Comprueba si el campo tiene un valor numérico superior o igual al valor especificado. Este operador funciona en tipos de datos numéricos y en campos que contienen números.
Es menor que	Comprueba si el campo tiene un valor numérico inferior al valor especificado. Este operador funciona en tipos de datos numéricos y en campos que contienen números.
Es inferior o igual a	Comprueba si el campo tiene un valor numérico inferior o igual al valor especificado. Este operador funciona en tipos de datos numéricos y en campos que contienen números.
Es nulo	Comprueba si el campo es un valor nulo.
No es nulo	Comprueba si el campo no es un valor nulo.

- Seleccione la opción **Recortar** según lo desee. Esta opción, primero recorta todos los espacios en blanco que se pueden presentar antes y después del valor del campo, antes de filtrar los datos en el campo.
- Repita hasta haber agregado todos los campos de entrada que desea utilizar para el filtro.

Read From Hive File

La etapa **Read from Hive File** lee datos del archivo seleccionado, que puede estar en cualquiera de los siguientes formatos:

- ORC
- RC
- Parquet
- Avro

Ficha *Propiedades del archivo*

Campos	Descripción
Servidor	Indica que el archivo que seleccionó en el campo Nombre de archivo está ubicado en el sistema Hadoop. Debe crear una conexión al servidor de archivos Hadoop en Management Console, antes de usarlo en esta etapa. Si selecciona un archivo en el sistema Hadoop, el nombre del servidor será el nombre que especifique en la Management Console mientras se crea un servidor de archivos.
Nombre de archivo	<p>Especifica la ruta al archivo. Haga clic en el botón de los puntos suspensivos (...) para buscar el archivo que desea.</p> <p>Nota: El esquema de un archivo de entrada se importa tan pronto navega hasta la ubicación correcta y selecciona el archivo. Este esquema importado no puede editarse.</p> <p>Sin embargo, puede cambiar el nombre de las columnas del esquema según sea necesario.</p> <p>Se recopilan los primeros 50 registros del archivo en la cuadrícula Vista previa al seleccionar el archivo.</p>
Tipo de archivo	<p>Seleccione el tipo de archivo que se lee:</p> <ul style="list-style-type: none"> • ORC • RC • Parquet • Avro

Nota: En el caso de los archivos RC, para generar la Vista previa, defina el esquema en la ficha **Campos** y luego haga clic en **Vista previa** en la ficha **Propiedades del archivo**.

Ficha Campos

La ficha **Campos** define los nombres, tipos de datos y posiciones de los campos como se encuentran en el archivo de entrada, así como también los nombres asignados por el usuario a los campos. Para obtener más información, consulte [Definición de campos para la lectura de Hive File](#) en la página 176.

Definición de campos para la lectura de Hive File

En la ficha **Campos** de la etapa **Read from Hive File**, se detallan los nombres de esquema, los tipos de datos, las posiciones y los nombres de los campos del archivo.

1. Haga clic en **Regenerar**.

En el caso de los archivos ORC, Avro y Parquet, esto genera un esquema basado en los metadatos de un archivo existente. En el caso de los archivos RC, se elimina cualquier campo que se haya agregado antes de hacer clic en **Vista previa**.

La cuadrícula muestra las columnas **Nombre**, **Tipo**, **Campo de etapa** e **Incluir**.

La columna **Nombre** muestra el nombre del campo, como se derivó del registro de encabezado del archivo.

La columna **Tipo** detalla los tipos de datos de cada respectivo campo del archivo.

La etapa admite los siguientes tipos de datos:

boolean	Tipo de dato lógico con dos valores: verdadero y falso
fecha	Un tipo de dato que contiene un mes, día y año. Por ejemplo, 2012-01-30 o 30 de enero, 2012. Puede especificar un formato de fecha predeterminado en Management Console.
fecha y hora	Un tipo de dato que contiene un mes, día, año y horas, minutos y segundos. Por ejemplo, 2012/01/30 6:15 PM. Nota: El tipo de dato <code>datetime</code> en los mapas Spectrum para el tipo de dato <code>timestamp</code> de los archivos Hive.
double	Un tipo de dato numérico que contiene números de doble precisión tanto negativos como positivos entre 2^{-1074} y $(2-2^{-52}) \times 2^{1023}$. En notación científica, el rango de valores es - 1.79769313486232E+308 a 1.79769313486232E+308.
bigdecimal	Un tipo de dato numérico que admite 38 puntos decimales de precisión. Utilice este tipo de datos para datos que se usarán en cálculos matemáticos que requieren un alto grado de precisión, especialmente aquellos que implican datos financieros. El tipo de datos bigdecimal es compatible con cálculos más precisos que el tipo de datos double.

Nota: Para los archivos RC, Avro y Parquet Hive, los campos del tipo de dato `decimal` en el archivo de entrada se convierten a tipo de `bigdecimal`.

long Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{63} (-9,223,372,036,854,775,808) y $2^{63}-1$ (9,223,372,036,854,775,807).

Nota: El tipo de dato `long` en los mapas Spectrum para el tipo de `bigint` de los archivos Hive.

integer Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{31} (-2,147,483,648) y $2^{31}-1$ (2,147,483,647).

float Un tipo de dato numérico que contiene números de precisión simple tanto negativos como positivos entre 2^{-149} y $(2-2^{23})\times 2^{127}$. En notación científica, el rango de valores es - 3.402823E+38 a 3.402823E+38.

string Secuencia de caracteres.

Nota: En el caso de los archivos RC, no se admiten `smallint` ni tipos de datos complejos.

La columna **Posición** muestra la posición inicial del campo respectivo de un registro.

2. En la columna **Campo de etapa**, edite el nombre existente y cambie el nombre que desea para cada campo.

De forma predeterminada, esta columna muestra los nombres de campos como se leen del archivo.

3. En la columna **Incluir**, seleccione las casillas de verificación según los campos que desea incluir en los resultados de la etapa.

De forma predeterminada, todos los campos están seleccionados en esta columna.

4. Para los archivos RC, puede agregar y eliminar campos, y modificar las secuencias de las columnas seleccionadas en la salida con los botones que figuran a continuación:

Nombre de la opción	Descripción
Agregar	Agrega un campo a la salida.
Modificar	Modifica el nombre y el tipo de dato del campo seleccionado.
Quitar	Elimina el campo seleccionado de la salida.

Nombre de la opción	Descripción
Subir/Bajar	Reordena la posición del campo seleccionado en la salida.

Nota: Esta función solo está disponible para los archivos RC.

5. Haga clic en **Aceptar**.

Archivo Read From HL7

La etapa **Read from HL7 File** lee datos de Health Level Seven (HL7) desde un archivo de texto como entrada a un flujo de datos. HL7 es un estándar de mensajería que se utiliza en la industria de la salud para el intercambio de datos entre sistemas. Para obtener más información acerca de HL7, vaya a www.hl7.org.

Formato de mensaje HL7

Los datos de un mensaje HL7 se organizan jerárquicamente de la siguiente manera:

- mensaje
 - segmento
 - campo
 - componente
 - Subcomponente

Cada una de las líneas de un mensaje HL7 es un segmento. Un *segmento* es una agrupación lógica de campos. Los primeros tres caracteres en un segmento identifican el tipo de segmento. En el mensaje anterior, hay cinco segmentos `MSH` (encabezado del mensaje), `PID` (identificación de paciente), dos segmentos `NK1` (familiares) e `IN1` (seguro).

Cada segmento está formado por una serie de campos. Los *campos* contienen información relacionada con el objetivo del segmento; por ejemplo, el nombre de la compañía de seguros en el segmento `IN1` (seguro). Los campos normalmente (pero no siempre) están delimitados por un carácter `|`.

Los campos se pueden dividir en *componentes*. Normalmente, los componentes se indican con un carácter `^`. En el ejemplo anterior, el segmento `PID` (identificación del paciente) contiene un campo

de nombre del paciente con LEVERKUHN^ADRIAN^C que tiene tres partes, el apellido (LEVERKUHN), el nombre (ADRIAN) y la inicial del segundo nombre (C). Los componentes se pueden dividir en *subcomponentes*. Normalmente, los subcomponentes se indican con un carácter &.

El siguiente es un ejemplo de un mensaje HL7:

```
MSH|^~\&||.|||199908180016||ADT^A04|ADT.1.1698593|P|2.7
PID|1||000395122||LEVERKUHN^ADRIAN^C||19880517180606|M|||6 66TH AVE
NE^WEIMAR^DL^98052||(157)983-3296|||S||12354768|87654321
NK1|1|TALLIS^THOMAS^C|GRANDFATHER|12914 SPEM
ST^ALIUM^IN^98052|(157)883-6176
NK1|2|WEBERN^ANTON|SON|12 STRASSE MUSIK^^VIENNA^AUS^11212|(123)456-7890
IN1|1|PRE2||LIFE PRUDENT BUYER|PO BOX
23523^WELLINGTON^ON^98111|||19601|||||THOMAS^JAMES^M|F|||||||||||||||ZKA535529776
```

Nota: Para crear un archivo HL7 usando el texto de muestra proporcionado:

1. Copie y pegue el texto de muestra en un documento nuevo por medio de cualquier software para la edición de texto, como Notepad++.
2. Realice los cambios de contenido que sean necesarios.
3. Configure para ver el final de línea (EOL) en el texto. En Notepad++, vaya a **Ver > Mostrar símbolo > Mostrar final de línea**.
4. Cambie el formato de conversión de final de línea a CR (retorno de carro). En Notepad++, vaya a **Editar > Conversión de EOL > Formato antiguo de Mac**.
5. Guarde el archivo HL7 después de realizar este cambio de formato.

Pestaña Propiedades del archivo

Nombre de campo	Descripción
Server name (Nombre de servidor)	Indica si el archivo que seleccionó como entrada está ubicado en la computadora que está ejecutando Enterprise Designer o en el servidor de Spectrum™ Technology Platform. Si selecciona un archivo en la computadora local, el nombre del servidor será Mi computadora. Si selecciona un archivo que se encuentra en el servidor, el nombre del servidor será Spectrum™ Technology Platform.
Nombre de archivo	Especifica la ruta al archivo. Haga clic en el botón de los puntos suspensivos (...) para buscar el archivo que desea. Nota: Si el servidor de Spectrum™ Technology Platform está ejecutándose en Unix o Linux, recuerde que estas plataformas distinguen las mayúsculas de las minúsculas en los nombres de los archivos y las rutas.

Nombre de campo	Descripción
Versión de HL7	La versión de HL7 estándar utilizada en el archivo especificado. Por ejemplo, "2.7" significa que el archivo utiliza HL7 versión 2.7. La versión de HL7 se indica en el 12.º campo del segmento <code>MSH</code> .
Character encoding (Codificación de caracteres)	<p>Codificación del archivo de texto. Seleccione uno de estos:</p> <p>UTF-8 Admite todos los caracteres Unicode y es compatible con las versiones anteriores de ASCII. Para obtener más información sobre UTF, consulte unicode.org/faq/utf_bom.html.</p> <p>UTF-16 Admite todos los caracteres Unicode, pero no es compatible con las versiones anteriores de ASCII. Para obtener más información sobre UTF, consulte unicode.org/faq/utf_bom.html.</p> <p>US-ASCII Codificación de caracteres basada en el orden del alfabeto inglés.</p> <p>UTF-16BE Codificación UTF-16 con serialización de bytes big endian (byte más significativo al principio).</p> <p>UTF-16LE Codificación UTF-16 con serialización de bytes little endian (byte menos significativo al principio).</p> <p>ISO-8859-1 Codificación de caracteres ASCII, que suele utilizarse para los idiomas de Europa occidental. También se la conoce como Latin-1.</p> <p>ISO-8859-3 Codificación de caracteres ASCII, que suele utilizarse para los idiomas de Europa meridional. También se la conoce como Latin-3.</p> <p>ISO-8859-9 Codificación de caracteres ASCII, que suele utilizarse para el idioma turco. También se la conoce como Latin-5.</p> <p>CP850 Página de códigos ASCII que se utiliza para escribir idiomas de Europa occidental.</p> <p>CP500 Página de códigos EBCDIC que se utiliza para escribir idiomas de Europa occidental.</p> <p>Shift_JIS Codificación de caracteres para el idioma japonés.</p> <p>MS932 Una extensión de Microsoft de Shift_JIS para incluir caracteres especiales NEC, selección NEC de extensiones IBM y extensiones IBM.</p> <p>CP1047 Una página de códigos EBCDIC con el conjunto de caracteres completo del alfabeto latino 1.</p>

Nombre de campo	Descripción
Validar	<p>Estas opciones especifican si se debe revisar el archivo para garantizar que se cumpla el estándar HL7 2.7. Si se encuentra un error de validación para cualquiera de los mensajes del archivo, se considerará como un registro malformado y entrarán en vigor las opciones de registro malformado especificadas para el trabajo (en Enterprise Designer, en Editar > Opciones de trabajo) o para el sistema (en Management Console).</p> <p>Campos obligatorios Marque esta casilla de verificación si desea asegurarse de que cada segmento, campo, componente y subcomponente contienen los elementos que requeridos por el estándar HL7 2.7.</p> <p>Longitud Marque esta casilla de verificación si desea asegurarse de que cada elemento cumple la longitud mínima y máxima para el elemento, según se define en el estándar HL7 2.7.</p>

Nombre de campo	Descripción
-----------------	-------------

Ignorar los elementos inesperados	
-----------------------------------	--

Nombre de campo

Descripción

Seleccione estas opciones si desea permitir mensajes que contengan segmentos, campos, componentes y subcomponentes que no estén en la posición esperada. Las posiciones esperadas se definen en el HL7 estándar o, en el caso de tipos de mensajes predeterminados, en la herramienta Administración de esquema de HL7 en Enterprise Designer.

Por ejemplo, considere el siguiente esquema de mensaje predeterminado:

```
MSH
[PID]
{ZSS}
PV1
NK1
{[DG1]}
```

Y estos datos:

```
MSH|^~\&|Pharm|GenHosp|CIS|GenHosp|198807050000||RAS^O17|RAS1234|P|2.7
ZSS|100|abc
PID|1234||PATID1234^5^M11^ADT1^MR^GOOD HEALTH
HOSPITAL~123456789^^^USSSA^SS|
PV1||O|O/R|||0148^ADDISON, JAMES|0148^ADDISON, JAMES
NK1|Son|John
```

En este caso, el segmento `PID` es inesperado porque está antes del segmento `ZSS`.

Los mensajes que contienen elementos en posiciones inesperadas se considerarán como registros malformados y entrarán en vigor las opciones de registro malformado especificadas para el trabajo (en Enterprise Designer, en **Editar > Opciones de trabajo**) o para el sistema (en Management Console).

De forma predeterminada, todas las opciones **Ignorar elemento inesperado** están habilitadas para permitir tantos registros como se pueda procesar con éxito.

- | | |
|-----------------------|---|
| Segmentos | Marque esta casilla de verificación para permitir que los mensajes contengan segmentos que no están definidos en el estándar HL7 2.7. Los segmentos inesperados se ignoran y los otros segmentos en el mensaje se procesan. |
| Campos | Marque esta casilla de verificación para permitir que los segmentos contengan campos que no están definidos en el estándar HL7 2.7. Los campos inesperados se ignoran y los otros campos en el segmento se procesan. |
| Componentes | Marque esta casilla de verificación para permitir que los campos contengan componentes que no están definidos en el estándar HL7 2.7. Los componentes inesperados se ignoran y los otros componentes en el campo se procesan. |
| Subcomponentes | Marque esta casilla de verificación para permitir que los componentes contengan subcomponentes que no están definidos en el estándar HL7 2.7. Los subcomponentes inesperados del componente se ignoran y los otros |

Nombre de campo

Descripción

 subcomponentes se procesan.

Pestaña Campos

En la pestaña **Campos** se muestran los segmentos, campos, componentes y subcomponentes. Utilice la pestaña **Campos** para seleccionar los datos que desea leer en el flujo de datos.

Los grupos de segmento, que son recolecciones de segmentos usados juntos para incluir una categoría de datos, se muestran con un sistema de numeración que muestra dónde aparece el grupo en el esquema del mensaje. Cada grupo de segmento recibe un rótulo "Group_n" donde "n" es un número que corresponde a la posición del grupo en el esquema del mensaje. Para ilustrar cómo funciona el sistema de numeración, considere este ejemplo:

```

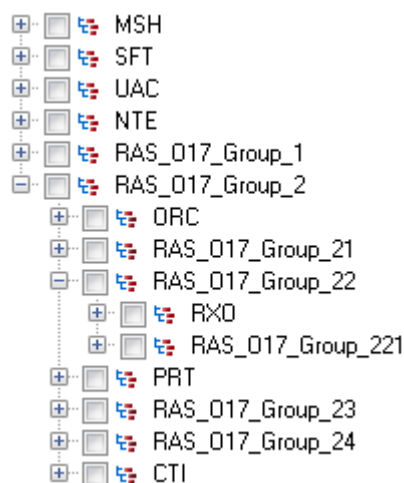
+ MSH
+ SFT
+ UAC
+ NTE
+ RAS_017_Group_1
+ RAS_017_Group_2
  
```

Este ejemplo muestra la lista de campo para el mensaje RAS^017. Este mensaje tiene dos grupos de segmento: RAS_017_Group_1 y RAS_017_Group_2. El grupo de segmento "Group_1" hace referencia al primer grupo de segmento en el esquema RAS^017 y el segundo grupo, "Group_2", hace referencia al segundo grupo indicado en el esquema RAS^017.

Para determinar qué grupo de segmento se representa con "Group_1" y "Group_2", encuentre la descripción del mensaje RAS^017 en el documento *HL7 versión 2.7 estándar*. Puede descargar una copia de este documento desde www.hl7.org.

En la descripción del mensaje, encuentre el primer grupo, el cual, en el caso de RAS^017, es el grupo PATIENT (Paciente). El segundo grupo en el esquema es el grupo ORDEN.

Los grupos de segmento que se anidan bajo un grupo de segmento tienen un número adicional adjunto a su número de grupo. Por ejemplo, el Group_21 representa el primer grupo anidado bajo el segundo grupo. Los posteriores subgrupos tienen números adicionales adjuntos a ellos, como Group_221, el cual, para el mensaje RAS^017, representa el grupo de segmento ORDER_DETAIL_SUPPLEMENT. A continuación se muestra un ejemplo de grupos anidados:



Los controles de la pestaña **Campos** se describen en la siguiente tabla.

Tabla 2: Pestaña Campos

Nombre de la opción	Descripción
Regenerar	Haga clic en este botón para rellenar la pestaña Campos con una lista de todos los segmentos, campos, componentes y subcomponentes para el tipo de mensaje contenido en el archivo de entrada. Todos los elementos para el tipo de mensaje se mostrarán de acuerdo con el esquema HL7, independientemente de si el archivo de entrada contiene todos los elementos. Por ejemplo, si el archivo contiene un mensaje RAS, se mostrará el esquema del tipo de mensaje RAS completo independientemente de si el archivo de entrada contiene los datos de todos los segmentos, campos, componentes y subcomponentes. Si definió cualquier elemento personalizado con la herramienta Administración de esquema de HL7 en Enterprise Designer, aquellos elementos también se enumeran.
Expandir todo	Expande todos los elementos en la pestaña campos para ver todos los segmentos, campos, componentes y subcomponentes de los tipos de mensaje incluidos en el archivo.
Cerrar todo	Cierra todos los nodos de la vista, de manera que se muestran solo los segmentos. Utilice esta opción para ver fácilmente los segmentos para los tipos de mensajes en el archivo. Después, puede ampliar cada uno de los segmentos con el fin de ver los campos, componentes y subcomponentes en un segmento.
Seleccionar todo	Marque esta casilla de verificación para crear campos de flujo de datos para todos los segmentos, campos, componentes y subcomponentes para todos los tipos de mensaje incluidos en el archivo.

Cómo comprimir datos HL7

Los datos HL7 se organizan jerárquicamente. Dado que muchas etapas requieren que los datos estén en formato plano, es posible que tenga que comprimir los datos para poder utilizarlos en etapas posteriores para acciones como la validación de direcciones o geocodificación.

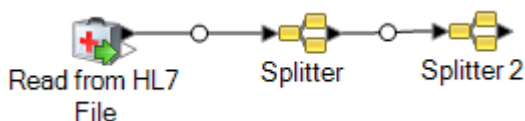
El siguiente procedimiento describe cómo utilizar la etapa Splitter para comprimir datos HL7.

1. Agregue una etapa Read From HL7 File a su flujo de datos y configúrela.
2. Agregue una etapa Splitter y conéctela a Read From HL7 File.
3. Agregue etapas Splitter adicionales según sea necesario para tener una etapa Splitter por segmento, campo o componente que desea comprimir.

Nota: Solo debe comprimir los datos de los que desea procesar en una etapa posterior. Los otros datos pueden permanecer en formato jerárquico. Por ejemplo, si solo desea procesar los datos de la dirección, solo necesita comprimir los datos de la dirección.

4. Conecte todas las etapas Splitter.

Ahora, su flujo de datos debería tener el siguiente aspecto:



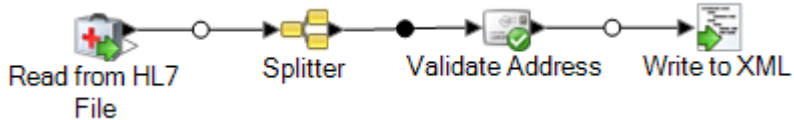
5. Haga doble clic en la primera etapa Splitter para abrir las opciones de la etapa.
6. En el campo **División en**, seleccione el segmento, campo o componente que desee comprimir.
7. Haga clic en **Aceptar**.
8. Configure cada etapa Splitter adicional seleccionando un segmento o componente en cada campo **Split at** de Splitter.
9. Agregue etapas adicionales según sea necesario después de la última Splitter para completar su flujo de datos.

Ejemplo

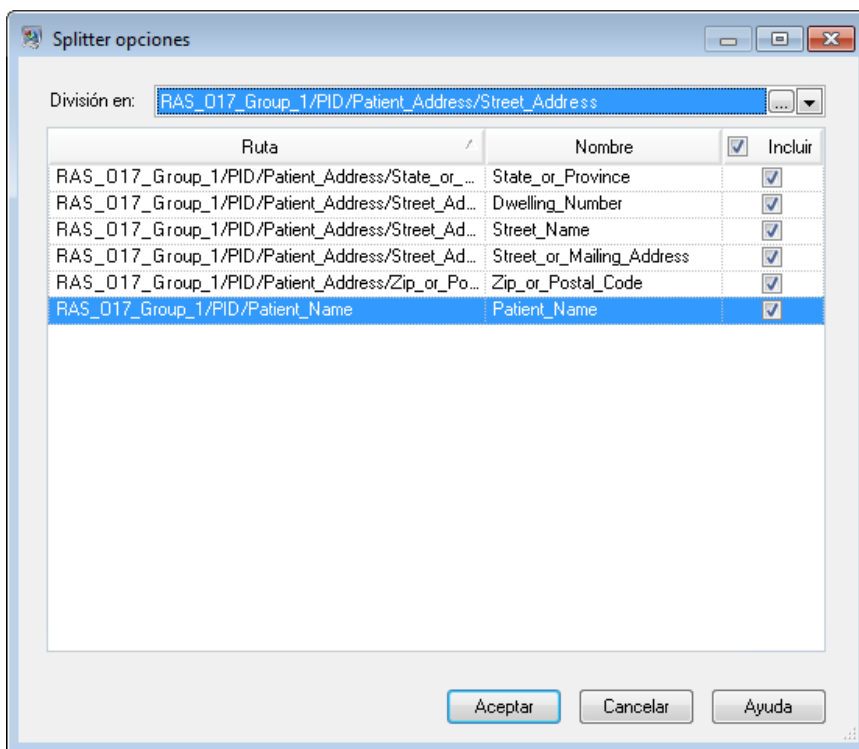
Tiene los siguientes datos HL7 y desea validar la dirección en el segmento PID.

```
MSH|^~\&|||.||199908180016||RAS^O17|ADT.1.1698594|P|2.7
PID|1||000395122||SMITH^JOHN^D||19880517180606|M||One Global
View^^Troy^NY^12180||(630)123-4567||S||12354768|87654321
```

Para realizar esto, debe convertir los datos de dirección a datos planos para que se puedan procesar mediante la etapa Validate Address. Por eso, cree un flujo de datos que contenga una etapa Splitter seguida de una etapa Validate Address, como esto:



La etapa Splitter está configurada para dividir al componente PID/Patient_Address/Street_Address, que convierte estos datos planos.



El canal que conecta la etapa Splitter a la etapa Validate Address cambia el nombre de los campos para utilizar los nombres de campo requeridos por Validate Address: Street_or_Mailing_Address is renamed to AddressLine1, State_or_Province cambia a StateProvince y Zip_or_Postal_Code cambia a PostalCode.

En este ejemplo, el resultado se escribe en un archivo XML que contiene los datos.

```

<?xml version='1.0' encoding='UTF-8'?>
<XmlRoot xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"

  xmlns:xsd="http://www.w3.org/2001/XMLSchema">
  <PatientInformation>
    <Confidence>95</Confidence>
    <RecordType>Normal</RecordType>
    <CountryLevel>A</CountryLevel>
    <ProcessedBy>USA</ProcessedBy>
    <MatchScore>0</MatchScore>
    <AddressLine1>1 Global Vw</AddressLine1>
  
```

```

<City>Troy</City>
<StateProvince>NY</StateProvince>
<PostalCode>12180-8371</PostalCode>
<PostalCode.Base>12180</PostalCode.Base>
<PostalCode.AddOn>8371</PostalCode.AddOn>
<Country>United States Of America</Country>
<Patient_Name>
  <Family_Name>
    <Surname>SMITH</Surname>
  </Family_Name>
  <Given_Name>JOHN</Given_Name>

<Second_and_Further_Given_Names_or_Initials_Thereof>
  D

</Second_and_Further_Given_Names_or_Initials_Thereof>
</Patient_Name>
</PatientInformation>
</XmlRoot>

```

Cómo agregar un mensaje HL7 personalizado

La etapa Read From HL7 File valida los mensajes mediante el uso del esquema HL7 2.7. Sin embargo, es posible que sus datos HL7 contengan mensajes que no forman parte del estándar HL7. Si desea que la etapa Read from HL7 File valide sus datos HL7 personalizados, debe crear un esquema HL7 personalizado. Esta sección describe la forma de crear un esquema HL7 personalizado mediante el uso de la herramienta Administración de esquema de HL7. Para obtener más información acerca de HL7, vaya a www.hl7.org.

1. En Enterprise Designer, vaya a **Herramientas > Administración de esquema de HL7**.

Esto abrirá la ventana Administración de esquema de HL7 que contiene una lista de los mensajes admitidos. Estos mensajes están predefinidos mediante HL7.

2. En la ventana **Administración de esquema de HL7**, haga clic en **Agregar**.
3. En el campo **Tipo de mensaje**, especifique un tipo para un mensaje de HL7 personalizado.

El tipo de mensaje indica qué información relacionada con la salud se proporciona en el mensaje. Por ejemplo, se usa un mensaje ADT (Admin Discharge Transfer) para intercambiar el estado del paciente al interior de un centro de salud y se usa un mensaje tipo ORU (Observation Result) para transmitir las observaciones y los resultados desde LIS (Lab Information System) hacia HIS (Hospital Information System).

4. En el campo **Evento desencadenador**, especifique un código de evento.

El evento desencadenador es un evento real que inicia la comunicación y el envío de un mensaje. Tanto el tipo de mensaje como el evento desencadenador se encuentran en el campo

MSH-9 del mensaje. Por ejemplo, el campo MSH-9 puede contener el valor ADT^A01. Esto quiere decir que ADT es el tipo de mensaje de HL7, y A01 es el evento desencadenador.

5. En el campo **Descripción**, escriba una descripción para un mensaje de HL7 personalizado.

Este campo le ayuda a entender más acerca de un tipo de mensaje. Por ejemplo, si va a agregar un mensaje tipo XYZ, puede proporcionar una descripción que se use para intercambiar el estado de un paciente al interior de un centro de salud.

Ahora verá el mensaje recién creado en la **Definición**. Haga clic en el símbolo más para expandir el mensaje. Puede ver que el segmento MSH se agrega automáticamente.

6. Para agregar un segmento existente a un mensaje
 - a) Haga clic en **Seleccionar segmento**.
 - b) Seleccione los segmentos que desee agregar al mensaje y haga clic en **Aceptar**.

En la cuadrícula **Esquema de segmentos** aparece un esquema de un segmento seleccionado y los mensajes marcados se agregan en el esquema de mensajes.

7. Para agregar un segmento personalizado a un mensaje
 - a) Haga clic en **Seleccionar segmento**.
 - b) Haga clic en **Agregar segmento**.
 - c) En el campo **Nombre**, especifique un nombre para el segmento y haga clic en **Aceptar**.

El segmento recién agregado se muestra en la parte inferior de la lista **Segmentos compatibles**.

- d) Seleccione el segmento personalizado que se agregó y luego haga clic en el botón **Agregar campo**.
 - e) En el campo **Nombre**, especifique un nombre de campo del segmento seleccionado.

Por ejemplo, un segmento PID (Información del paciente) contiene nombres de campo tales como ID del paciente, Nombre del paciente, Dirección del paciente, Código del país, etc.

- f) En el campo **Tipo**, seleccione un tipo de datos apropiado.

Los tipos de datos HL7 definen la clase de datos que se puede incluir en un campo, y que se usa en toda la estructura del mensaje de HL7. Por ejemplo, ST para cadena, TX para datos de texto y FT para datos con formato.

- g) En el campo **Longitud normativa**, especifique la longitud máxima y mínima del campo, con el siguiente formato: m..n. Del mismo modo, puede especificar una lista de valores posibles para una longitud del campo, con el siguiente formato: x,y,z.

Por ejemplo, la longitud de 1..3 significa que la longitud de un elemento puede ser 1, 2 o 3 y la longitud de 1, 3, 4 quiere decir que la longitud de un elemento puede ser 1, 3 o 4, pero no 2. Un valor distinto a 1, 3, y 4 se consideraría inválido.

- h) En el campo **Opcionalidad**, especifique si un campo es opcional u obligatorio.

O

El campo es opcional.

R

El campo es obligatorio.

- i) En el campo **Repetición**, si desea permitir que el campo aparezca más de una vez en un segmento, marque la casilla **Repetición** y especifique la cantidad de veces que este se puede utilizar.

Por ejemplo, un valor de 3 significaría que el campo puede tener tres ocurrencias. Si no se especifica, solo habrá una ocurrencia, lo que indica que este campo no se repetirá.

8. Haga clic en **Aceptar**.

También puede elegir las alternativas **Opcional** y **Repetitivo** en las propiedades del segmento.

9. Elija **Opcional** para dejar el segmento como opcional y elija **Repetitivo** para permitir una repetición de un segmento seleccionado en un mensaje.

10. Haga clic en **Aceptar**.

Se muestra el mensaje agregado recientemente en la parte inferior de la lista.

Read from NoSQL DB

La etapa Read from NoSQL DB lee los datos desde una tabla de base de datos como entrada a un flujo de datos. La etapa admite los tipos de base de datos MongoDB y Couchbase.

Pestaña General

Nombre de campo	Descripción
Conexión	<p>Seleccione la conexión de base de datos obligatoria en la lista desplegable. Las opciones que se muestran varían según las conexiones especificadas en Management Console.</p> <p>Para agregar una nueva conexión, consulte Conexión con NoSQL en la página 54.</p> <p>Para modificar una conexión existente, selecciónela de la lista de conexiones en la página Fuentes de datos de Management Console, ábrala, haga los cambios necesarios y, a continuación, haga clic en el botón Guardar.</p>

Nombre de campo	Descripción
Tabla/Vista	<p>Especifica la colección o vista en la base de datos que se desea consultar.</p> <p>Nota: Mientras que en la interfaz de usuario se usa el término Tabla/Vista, en MongoDB se llama <i>colección</i>, y en Couchbase se llama <i>vista</i>.</p>
Archivo de esquema	<p>Haga clic en el botón Examinar (...) para seleccionar un archivo de esquema JSON. Este archivo es opcional. Los campos de la pestaña de campos se pueden generar con el archivo de esquema o con la tabla/visualización de la base de datos.</p> <p>Para borrar la ruta del archivo seleccionado, haga clic en Borrar.</p> <p>Nota: Los campos siempre se generarán con el archivo de esquema si se selecciona uno.</p>
Where	<p>Ingrese los criterios de filtro requeridos, de existir, utilizando la sintaxis MongoDB para recopilar registros específicos. Deje el campo en blanco si no se requieren criterios de filtro.</p> <p>La siguiente sintaxis es para una cláusula con un operador <i>igual a</i>:</p> <pre style="background-color: #f0f0f0; padding: 5px;">{"<column name>" : "<filter value>"}</pre> <p>Puede combinar varias cláusulas utilizando los operadores requeridos. Para ver una lista de los operadores admitidos en la cláusula <code>where</code>, consulte http://docs.mongodb.org/manual/reference/operator/query/.</p> <p>Por ejemplo, para capturar registros en los cuales el valor de la columna <code>customer_name</code> se cruza con el valor <code>John</code>, y el valor de la columna <code>customer_age</code> es mayor o igual que <code>45</code>, ingrese lo siguiente:</p> <pre style="background-color: #f0f0f0; padding: 5px;">{\$and: [{"customer_name": "John"}, {\$gte: ["customer_age", "45"]}]}</pre> <p>Atención: Asegúrese de no incluir la palabra clave <code>where</code> en este campo.</p> <p>Nota: Actualmente, este campo es visible únicamente al seleccionar una conexión MongoDB.</p>
Ignore los campos ausentes	<p>Los campos definidos en el esquema, si no están presentes en el registro real no fluirán hasta la próxima etapa si se selecciona esta opción.</p> <p>Nota: Si no habilita esta opción, los campos que no están presentes en la tabla o visualización de base de datos, se agregan y se procesan con el valor NULO.</p>

Nombre de campo	Descripción
Vista previa	<p>Muestra los registros desde la tabla seleccionada.</p> <p>Nota: Para las fuentes de datos MongoDB, si hace clic en Vista previa se mostrarán los registros filtrados, si se han ingresado una o más cláusulas <code>where</code> en el campo Donde. Si no se ha ingresado ninguna cláusula "dónde", la vista previa muestra todos los registros.</p> <p>Nota: Para las fuentes de datos Couchbase, si hace clic en Vista previa también se mostrará el campo agregado <code>_id</code> que contiene la clave. Si el registro ya tiene un campo <code>_id</code>, el campo <code>_id</code> agregado sobrescribirá el preexistente en el momento de previsualizar los campos.</p>
Expandir todo	Expande los elementos en el árbol de vista previa.
Cerrar todo	Cierra los elementos en el árbol de vista previa.

Pestaña Campos

La pestaña Campos le permite seleccionar los datos que desea pasar a la siguiente etapa. Para obtener más información, consulte [Definición de campos en una base de datos NoSQL](#) en la página 192

Definición de campos en una base de datos NoSQL

La ficha **Campos** muestra el campo y su tipo tal como se define en el archivo NoSQL DB/Schema.

1. En la **ficha Campos** haga clic en **Regenerar**.

Esto genera los datos agregados a partir de los primeros 50 registros. Los datos se muestran en el siguiente formato: `Fieldname (datatype)`

Nota: Si se explora el archivo de esquema, se generan los campos utilizando el archivo de esquema y se omite la tabla o la visualización. Para restablecer el archivo de esquema, haga clic en **Borrar**.

Nota: Mientras lee los datos de una base de datos Couchbase, la clave de cada registro también se lee. Esta clave se almacena como una parte del registro que utiliza un campo `_id` agregado en el momento de la regeneración de los campos, y también se incluye en los datos enviados a la siguiente etapa. Si el registro ya tiene un campo `_id`, se sobrescribirá con el campo `_id` agregado en el momento de volver a generar los campos.

2. Para modificar el nombre y tipo de un campo, resalte el campo y haga clic en **Modificar**.
3. En el campo **Nombre**, elija el campo que desea agregar o ingrese el nombre del campo.

- En el campo **Tipo** puede dejar el tipo de datos como cadena, si no se ha propuesto realizar ninguna operación matemática con los datos. Sin embargo, si se propone realizar ese tipo de operaciones, seleccione un tipo de datos apropiado. Esto convertirá los datos de cadena del archivo en un tipo de datos que permitirá el manejo adecuado de los datos en el flujo de datos.

La etapa admite los siguientes tipos de datos:

double	Un tipo de dato numérico que contiene números de doble precisión tanto negativos como positivos entre 2^{-1074} y $(2-2^{-52})\times 2^{1023}$. En notación científica, el rango de valores es - 1.79769313486232E+308 a 1.79769313486232E+308.
float	Un tipo de dato numérico que contiene números de precisión simple tanto negativos como positivos entre 2^{-149} y $(2-2^{-23})\times 2^{127}$. En notación científica, el rango de valores es - 3.402823E+38 a 3.402823E+38.
integer	Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{31} (-2,147,483,648) y $2^{31}-1$ (2,147,483,647).
long	Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{63} (-9,223,372,036,854,775,808) y $2^{63}-1$ (9,223,372,036,854,775,807).
string	Secuencia de caracteres.

- También puede agregar campos adicionales que no estén presentes en el archivo de tabla o esquema. Haga clic en **Agregar** para agregar un nuevo campo. Para eliminar un campo, haga clic en **Eliminar**.

Nota: Solo puede agregar un nuevo campo en Tipo de lista.

- Haga clic en **Aceptar**.

Opciones de flujo de datos para la base de datos NoSQL

Este procedimiento describe la forma de configurar un flujo de datos compatible con las opciones de ejecución para la base de datos NoSQL.

- Abra el flujo de datos en Enterprise Designer.
- Si desea configurar las opciones de tiempo de ejecución para una etapa en un flujo de datos incrustado, abra este último.
- Haga clic en el ícono Opciones de flujo de datos en la barra de herramientas o haga clic en **Editar > Opciones de flujo de datos**. Aparecerá el cuadro de diálogo **Opciones de flujo de datos**.
- Haga clic en **Agregar**. Aparecerá el cuadro de diálogo **Definir opción de flujo de datos**.
- Expanda la etapa de NoSQLDB.
- Las opciones de flujo de datos se exponen tal como se describen en la siguiente tabla:

Base de datos	Leer	Escribir
Mongo DB	Conexión	Conexión
	Tabla	Tabla
Base de datos Couchbase	Conexión	Conexión
	Mostrar	
	Nombre de documento de diseño	

El nombre de la opción de base de datos NoSQL seleccionada aparece en los campos **Nombre de opción** y **Etiqueta de opción**. Este es el nombre de opción que deberá especificarse durante el tiempo de ejecución para definir esta opción.

7. Ingrese una descripción de la opción en el campo **Descripción**.
8. En el campo **Destino**, seleccione la opción **Etapas seleccionadas**.
9. Si desea limitar los valores que pueden especificarse durante la ejecución, edite las opciones en el campo **Valores válidos** haciendo clic en el icono a la derecha del campo.
10. Si desea cambiar el valor predeterminado, especifique un valor diferente en el campo **Valor predeterminado**.

Nota: En el caso de un servicio, solo puede modificar los valores predeterminados antes de exponer el servicio por primera vez. Después de exponer el servicio, ya no es posible modificar los valores predeterminados con Enterprise Designer. En cambio, debe usar Management Console. Para obtener más información, consulte [Especificación de las opciones de servicio predeterminadas](#).

11. Haga clic en **Aceptar**.
12. Continúe agregando opciones, según sus preferencias.
13. Haga clic en **ACEPTAR** en el cuadro de diálogo **Opciones de flujo de datos** cuando haya terminado de agregar opciones.
14. Guarde y exponga el flujo de datos.

Read from SAP

La etapa Read from SAP lee los datos desde base de datos SAP como entrada a un flujo de datos. Puede leer los datos desde una o varias tablas. Cuando lee los datos desde varias tablas, la etapa realiza una operación de combinación para determinar qué registros leer en el flujo de datos.

Cómo conectarse a SAP

Para leer datos desde SAP en un flujo de datos mediante Read from SAP, debe crear una conexión entre Spectrum™ Technology Platform y su sistema SAP.

1. Abra el administrador de conexión SAP. Puede hacerlo en Enterprise Designer en **Herramientas > Administración de conexión SAP** o en la etapa Read from SAP al hacer clic en el botón **Administrar** junto al campo **Conexión**.
2. Haga clic en **Agregar**.
3. En el campo **Nombre de conexión** asigne un nombre a la conexión.
4. Complete los otros campos con la información acerca del servidor SAP al cual desee conectarse. Solicite la información necesaria al administrador de SAP Basis.

Importante: La ID de usuario y la contraseña deben ser de una cuenta SAP con privilegios de administrador.

5. Haga clic en **Probar** para verificar la conexión.
6. Haga clic en **Aceptar**.

No creó una conexión que la etapa Read from SAP pueda usar para leer datos desde SAP en un flujo de datos.

Leer datos de una tabla SAP única

La etapa Read From SAP puede configurarse para que lea datos de una sola tabla en la base de datos SAP o de múltiples tablas. Este procedimiento explica cómo configurar Read from SAP para que lea datos de una sola tabla.

1. En Enterprise Designer, arrastre Read From SAP al lienzo.
2. Haga doble clic en la etapa Read from SAP en el lienzo.
3. En el campo **Conexión**, seleccione el servidor SAP que contiene los datos que quiere leer en el flujo de datos. Si no hay ninguna conexión determinada para el servidor SAP, deberá crear la conexión. Para ello, haga clic en **Administrar**.
4. En el campo **Tipo de fuente**, elija **Única**.
5. Haga clic en **Seleccionar**.
6. Seleccione la tabla que quiere leer en el flujo de datos y, luego, haga clic en **Aceptar**.

Nota: Solo se muestran las primeras 200 tablas. Use la función de búsqueda si quiere ver alguna tabla que no está en la lista de las primeras 200. El campo de búsqueda solo analiza los valores en las columnas **Nombre** y **Rótulo**.

7. Para ver los nombres de los campos que se usarán en el flujo de datos, marque la casilla **Mostrar nombre técnico**.

Los campos en SAP tienen un nombre sencillo (intuitivo) que se usa para visualización y un nombre único que puede ser menos legible. Por ejemplo, un campo puede tener como nombre sencillo "Distribution Channel" y como nombre técnico "DIS_CHANNEL". Para garantizar que el nombre del campo sea válido en el flujo de datos, se usa el nombre técnico como nombre de campo.

8. Marque la casilla en la columna **Incluir** para cada campo que quiera que se lea en el flujo de datos.
9. Haga clic en **Aceptar**.
10. Si desea leer solo ciertos registros, puede especificar las condiciones de filtro en la ficha **Filtro**. Para que un registro se pueda leer en el flujo de datos, debe cumplir todas las condiciones que usted defina.
11. Puede mejorar el rendimiento si especifica un tamaño de recuperación apropiado en la ficha **Tiempo de ejecución**.

Seleccione esta opción para especificar el número de registros que se van a leer desde la tabla de base de datos al mismo tiempo. Por ejemplo, si el valor del **Tamaño de recuperación** es 100 y el número total de registros que se van a leer es 1000, haría 10 recorridos a la base de datos para leer todos los registros.

La configuración de un tamaño de recuperación óptimo puede mejorar significativamente el rendimiento.

Nota: Puede calcular el tamaño de recuperación óptimo para su entorno, al probar los tiempos de ejecución entre una etapa Read from DB y una etapa Write to Null. Para obtener más información, consulte [Determinar un tamaño de recuperación óptimo](#) en la página 322.

El tamaño de recuperación predeterminado de Read from SAP es 10.000.

12. Haga clic en **Aceptar**.

La etapa Read From SAP ahora está configurada para leer datos de una sola tabla de la base de datos SAP en el flujo de datos.

Leer datos de varias tablas SAP

La etapa Read From SAP puede configurarse para que lea datos de una sola tabla en la base de datos SAP o de múltiples tablas. Este procedimiento explica cómo configurar Read from SAP para que lea datos de varias tablas. Para leer datos de varias tablas, defina una declaración JOIN para combinar datos en una sola cadena.

1. En Enterprise Designer, arrastre Read From SAP al lienzo.

2. Haga doble clic en la etapa Read from SAP en el lienzo.
3. En el campo **Conexión**, seleccione el servidor SAP que contiene los datos que quiere leer en el flujo de datos. Si no hay ninguna conexión determinada para el servidor SAP, deberá crear la conexión. Para ello, haga clic en **Administrar**.
4. En el campo **Tipo de fuente**, elija **Múltiple**.
5. Haga clic en **Agregar**.
6. Seleccione las tablas que quiere leer en el flujo de datos y, luego, haga clic en **Aceptar**.

Nota: Solo se muestran las primeras 200 tablas. Use la función de búsqueda si quiere ver alguna tabla que no está en la lista de las primeras 200. El campo de búsqueda solo analiza los valores en las columnas **Nombre** y **Rótulo**.

7. Seleccione la primera tabla en la lista y haga clic en **Crear relación**. Esta es la tabla de origen.
8. En el campo **Clave origen** seleccione la columna desde la tabla de origen cuyo valor se utilizará para comparar los registros con los registros de otra tabla.
9. En el campo **Tipo de unión** seleccione una de las siguientes opciones:

INNER JOIN Devuelve solo los registros que tienen una coincidencia entre las tablas de origen y destino.

LEFT JOIN Devuelve todos los registros desde la tabla de origen, si no hay coincidencias entre las tablas de origen y destino. Esta opción devuelve todos los registros desde la tabla de origen, más todos los registros que coincidan en la tabla de destino.

10. En el campo **Tabla** seleccione la tabla de destino.
11. En el campo **Clave de tabla** seleccione la columna en la tabla de destino que contiene los datos que desea comparar con los del campo **Clave origen**, para determinar si el registro cumple la condición de unión.
12. Haga clic en **Aceptar**.
13. Haga clic en **Seleccionar esquema**.
14. Elija los campos que desea leer en el flujo de datos. Para ver los nombres de los campos que se usarán en el flujo de datos, marque la casilla **Mostrar nombre técnico**.

Los campos en SAP tienen un nombre sencillo (intuitivo) que se usa para visualización y un nombre único que puede ser menos legible. Por ejemplo, un campo puede tener como nombre sencillo "Distribution Channel" y como nombre técnico "DIS_CHANNEL". Para garantizar que el nombre del campo sea válido en el flujo de datos, se usa el nombre técnico como nombre de campo.

15. Haga clic en **Aceptar**.
16. Si desea leer solo ciertos registros, puede especificar las condiciones de filtro en la ficha **Filtro**. Para que un registro se pueda leer en el flujo de datos, debe cumplir todas las condiciones que usted defina.

17. Puede mejorar el rendimiento si especifica un tamaño de recuperación apropiado en la ficha **Tiempo de ejecución**.

Seleccione esta opción para especificar el número de registros que se van a leer desde la tabla de base de datos al mismo tiempo. Por ejemplo, si el valor del **Tamaño de recuperación** es 100 y el número total de registros que se van a leer es 1000, haría 10 recorridos a la base de datos para leer todos los registros.

La configuración de un tamaño de recuperación óptimo puede mejorar significativamente el rendimiento.

Nota: Puede calcular el tamaño de recuperación óptimo para su entorno, al probar los tiempos de ejecución entre una etapa Read from DB y una etapa Write to Null. Para obtener más información, consulte [Determinar un tamaño de recuperación óptimo](#) en la página 322.

El tamaño de recuperación predeterminado de Read from SAP es 10.000.

La etapa Read From SAP ahora está configurada para leer datos de varias tablas de la base de datos SAP en el flujo de datos.

Cómo filtrar registros en Read from SAP

La configuración de filtro en Read from SAP le permite leer un subconjunto de registros desde una tabla SAP, en lugar de todos los registros en la tabla. Para filtrar los registros debe especificar los valores que estos deben incluir, a fin de que se puedan leer en el flujo de datos. Si no especifica las condiciones de filtro, todos los registros de la tabla se leen en el flujo de datos. El uso de las condiciones de filtro es opcional.

Nota: Si la etapa Read from SAP se configura para leer datos desde varias tablas SAP, el filtro se aplica después de realizar la operación JOIN.

1. En la etapa Read from SAP, haga clic en la ficha **Filtro**.
2. Haga clic en **Agregar**.
3. En el campo **Nombre de tabla** seleccione la tabla que incluye los registros que desea filtrar.
4. En el campo **Filtrar por** seleccione el campo que incluye los datos que desea utilizar como base para el filtro.
5. Elija uno de los siguientes operadores:

Nota: Los operadores que están disponibles varían según el tipo de datos del campo en el cual se aplica el filtro.

Operador	Descripción
Contiene	Comprueba si la cadena contiene el valor especificado.
Es igual a	Comprueba si el valor en el campo coincide con el valor especificado.
No es igual a	Comprueba si el valor del campo no coincide con el valor especificado.
Mayor que	Comprueba si el campo tiene un valor numérico superior al valor especificado. Este operador funciona en tipos de datos numéricos y en campos que contienen números.
Es mayor o igual que	Comprueba si el campo tiene un valor numérico superior o igual al valor especificado. Este operador funciona en tipos de datos numéricos y en campos que contienen números.
Menor que	Comprueba si el campo tiene un valor numérico inferior al valor especificado. Este operador funciona en tipos de datos numéricos y en campos que contienen números.
Menor o igual que	Comprueba si el campo tiene un valor numérico inferior o igual al valor especificado. Este operador funciona en tipos de datos numéricos y en campos que contienen números.
Es nulo	Comprueba si el campo es un valor nulo.
No es nulo	Comprueba si el campo no es un valor nulo.
Comienza con	Comprueba si el campo comienza con el valor especificado.
Finaliza con	Comprueba si el campo finaliza con el valor especificado.

6. Ingrese el valor con el cual desea comparar el valor del campo seleccionado.
7. Haga clic en **Aceptar**.
8. Agregue condiciones de filtro adicionales si es necesario.

Nota: Si especifica varias condiciones de filtro, todas estas deben ser verdaderas para que el registro se pueda leer en el flujo de datos. Si alguna de las condiciones no es verdadera, el registro no se leerá en el flujo de datos.

Leer de la hoja de cálculo

La función Leer de la hoja de cálculo lee datos de una hoja de cálculo de Excel como de entrada a un flujo de datos. Admite formatos de archivo .xls y .xlsx.

Pestaña Propiedades del archivo

La pestaña **Propiedades de archivo** contiene opciones que se utilizan para especificar la hoja de cálculo y los datos que se leerán en el flujo de datos.

Nombre de campo	Descripción
Server name (Nombre de servidor)	Indica si el archivo que seleccionó como entrada está ubicado en la computadora que está ejecutando Enterprise Designer o en el servidor de Spectrum™ Technology Platform. Si selecciona un archivo en la computadora local, el nombre del servidor será Mi computadora. Si selecciona un archivo que se encuentra en el servidor, el nombre del servidor será Spectrum™ Technology Platform.
Nombre de archivo	Especifica la ruta al archivo. Haga clic en el botón de los puntos suspensivos (...) para buscar el archivo que desea. Nota: Si el servidor de Spectrum™ Technology Platform está ejecutándose en Unix o Linux, recuerde que estas plataformas distinguen las mayúsculas de las minúsculas en los nombres de los archivos y las rutas.
Selección de datos	Especifica cómo desea seleccionar los datos de la hoja de cálculo para introducirlos en el flujo de datos. Una de las siguientes: <ul style="list-style-type: none"> Datos de hoja Seleccione esta opción para leer todos los datos de una hoja en la hoja de cálculo. Datos de rango Seleccione esta opción para introducir un subconjunto de datos de una hoja mediante la especificación de un rango de celdas a leer. Rango con nombre Seleccione esta opción para introducir un subconjunto de datos de una hoja mediante la especificación de un rango con nombre de la hoja de cálculo.
Selección de hoja	Si elige Datos de hoja o Datos de rango en el campo Selección de datos , use esta opción para elegir la hoja desde la que desea introducir datos en el flujo de datos.

Nombre de campo	Descripción
Rango	Si elige Datos de rango en el campo Selección de datos , use esta opción para especificar la celda que inicia el rango y la celda que lo finaliza.
Rango con nombre	Si elige Rango con nombre en el campo Selección de datos , use esta opción para especificar el nombre del rango que desea introducir en el flujo de datos. Los rangos se encuentran definidos en la hoja de cálculo. Si no hay rangos detallados, esto significa que no se definió ningún rango en la hoja de cálculo.
Fila de encabezado	<p>Marque esta casilla para especificar una fila que contenga encabezados de columna. Los encabezados de columna pasan a ser los nombres de campo del flujo de datos, aunque pueden cambiarse en la pestaña Campos. Si no marca esta casilla, se les asignan nombres predeterminados genéricos a los campos de flujo de datos, como Column1, Column2, etc.</p> <p>La fila de encabezado que especifique es en relación con la selección de datos. Por ejemplo, si elige Datos de rango en el campo Selección de datos y el rango comienza en la quinta fila y especifica 1 como la fila de encabezado, entonces la quinta fila en la hoja de cálculo se usará como el encabezado ya que la quinta fila de la hoja de cálculo es la primera fila del rango.</p>
Desplazamiento de datos del encabezado	Si especifica una fila de encabezado, este campo especifica la primera fila que contiene datos, en relación con el encabezado. Por ejemplo, si especifica 1, la primera fila debajo del encabezado será la primera fila de los datos que se van a leer en el flujo de datos. Si especifica 2, la segunda fila debajo del encabezado será la primera fila que se leerá en el flujo de datos y así sucesivamente.
Primera fila con datos	Si no especifica una fila de encabezado, este campo especifica qué fila dentro de la selección de datos contiene la primera fila de los datos que se van a leer en el flujo de datos. La fila que especifique es en relación con la selección de datos. Por ejemplo, si elige Datos de rango en el campo Selección de datos y el rango comienza en la quinta fila y especifica 1 como la primera fila con datos, entonces la primera fila de datos que se va a leer en el flujo de datos será la quinta fila.
Ignorar filas vacías	<p>Seleccione esta opción si desea excluir las filas vacías del flujo de datos. Si no selecciona esta opción, las filas vacías en la hoja de datos pasarán a ser registros vacíos en el flujo de datos.</p> <p>Nota: Esta opción no afecta los datos que se muestran en la vista previa. las filas vacías siempre se muestran en la vista previa, aunque esta opción esté seleccionada.</p>

Pestaña Campos

La pestaña **Campos** contiene opciones para asignar los datos de la hoja de cálculo a los campos en el flujo de datos.

Opción	Descripción
Regenerar	Haga clic en este botón para completar la pestaña de campos con los campos del archivo de entrada especificado en la pestaña Propiedades del archivo .
Detectar tipo	Haga clic en este botón para determinar automáticamente el tipo de datos de todos los campos. Puede cambiar manualmente el tipo de datos de un campo; para ello, debe seleccionar el campo y hacer clic en Modificar .
Modificar	Seleccione un campo y luego haga clic en este botón para modificar el nombre de campo o el tipo de datos.

Read From Variable Format File

Read from Variable Format File lee los datos desde un archivo que contiene registros de diseño variable. Cada registro se lee como un campo de lista. Puede especificar la etiqueta que indica el tipo de registro principal, y todos los demás tipos de registro se convertirán en campos de lista bajo ese elemento principal.

Los archivos de formato variable poseen estas características:

- Los registros del archivo pueden tener campos diferentes y distinta cantidad de campos.
- Todos los registros deben contener una etiqueta (normalmente un número) que identifique el tipo de registro.
- Se admiten las relaciones jerárquicas.

Ejemplo de archivo de formato variable

En este ejemplo se muestra un archivo de formato variable que contiene información sobre la actividad de la cuenta corriente de dos clientes: Joe Smith y Anne Johnson.

En este ejemplo, el archivo es un archivo delimitado que usa la coma como delimitador de campo.

```
001 Joe,Smith,M,100 Main St,555-234-1290
100 CHK12904567,12/2/2007,6/1/2012,CHK
200 1000567,1/5/2012,Fashion Shoes,323.12
001 Anne,Johnson,F,1202 Lake St,555-222-4932
100 CHK238193875,1/21/2001,4/12/2012,CHK
200 1000232,3/5/2012,Blue Goose Grocery,132.11
200 1000232,3/8/2012,Trailway Bikes,540.00
```

El primer campo de cada registro contiene la etiqueta que identifica el tipo de registro y, por lo tanto, el formato del registro:

- 001: registro del cliente
- 100: registro de la cuenta
- 200: registro de la transacción de la cuenta

Para los archivos delimitados, es común que el valor de la etiqueta (001, 100, 200) sea un número fijo de bytes ubicado al principio del registro, como se muestra en el ejemplo anterior.

Cada registro tiene su propio formato:

- 001: Nombre,Apellido,Sexo,Dirección,Número_de_teléfono
- 100: ID_de_la_cuenta,Fecha_de_apertura,Fecha_de_vencimiento,Tipo_de_cuenta
- 200: ID_de_la_transacción,Fecha_de_la_transacción,Proveedor,Monto

El formato de registro 100 (registro de cuenta) es un campo secundario del registro anterior 001, y el formato de registro 200 (registro de transacción de cuenta) es un campo secundario del registro anterior 100 (registro de cuenta). En el archivo de ejemplo, en la cuenta CHK12904567 de Joe Smith, se realizó una transacción el 5/1/12 por un importe de 323,12 en Fashion Shoes. A su vez, en la cuenta CHK238193875 de Anne Johnson, se realizaron dos transacciones, una el 5/3/12 en Blue Goose Grocery y la otra el 8/3/12 en Trailway Bikes.

Pestaña Propiedades del archivo

Nombre de la opción	Descripción
Server name (Nombre de servidor)	Indica si el archivo que seleccionó como entrada está ubicado en la computadora que está ejecutando Enterprise Designer o en el servidor de Spectrum™ Technology Platform. Si selecciona un archivo en la computadora local, el nombre del servidor será Mi computadora. Si selecciona un archivo que se encuentra en el servidor, el nombre del servidor será Spectrum™ Technology Platform.

Nombre de la opción	Descripción
Nombre de archivo	<p>Especifica la ruta al archivo. Haga clic en el botón de puntos suspensivos (...) para dirigirse al archivo que desea.</p> <p>Puede leer múltiples archivos al usar el carácter comodín para leer datos desde múltiples archivos en el directorio. Los caracteres comodines * y ? son compatibles. Por ejemplo, puede especificar *.CSV para leer todos los archivos con extensión .CSV que se encuentran en el directorio. Para leer correctamente múltiples archivos, cada uno debe tener la misma disposición (los mismos campos en las mismas posiciones). Cualquier registro que no coincida con la disposición especificada en la ficha Campos se tratará como un registro malformado.</p> <p>Cuando lee un archivo del servidor de archivos HDFS, los formatos de compresión compatibles son:</p> <ol style="list-style-type: none"> 1. GZIP (.gz) 2. BZIP2 (.bz2) <p>Nota: La extensión del archivo indica el formato de compresión que se utiliza para descomprimir el archivo.</p> <p>Atención: Si el servidor de Spectrum™ Technology Platform está ejecutándose en Unix o Linux, recuerde que estas plataformas distinguen las mayúsculas de las minúsculas en los nombres de los archivos y las rutas.</p>
Tipo de registro	<p>Formato de los registros en el archivo. Seleccione uno de:</p> <p>En líneas secuenciales Archivo de texto en el que los registros se separan con un carácter de final de línea (EOL), como un retorno de carro o un salto de línea (CR o LF) y cada campo tiene una posición fija para los caracteres de inicio y final.</p> <p>De ancho fijo Archivo de texto en el que cada registro es un número específico de caracteres de longitud y cada campo tiene una posición fija para los caracteres de inicio y final.</p> <p>Delimitado Archivo de texto en el que los registros se separan con un carácter de final de línea (EOL), como un retorno de carro o un salto de línea (CR o LF) y cada campo se separa con un carácter designado; por ejemplo, una coma.</p>

Nombre de la opción	Descripción
Character encoding (Codificación de caracteres)	Codificación del archivo de texto. Seleccione uno de estos:
	UTF-8 Admite todos los caracteres Unicode y es compatible con las versiones anteriores de ASCII. Para obtener más información sobre UTF, consulte unicode.org/faq/utf_bom.html .
	UTF-16 Admite todos los caracteres Unicode, pero no es compatible con las versiones anteriores de ASCII. Para obtener más información sobre UTF, consulte unicode.org/faq/utf_bom.html .
	US-ASCII Codificación de caracteres basada en el orden del alfabeto inglés.
	UTF-16BE Codificación UTF-16 con serialización de bytes big endian (byte más significativo al principio).
	UTF-16LE Codificación UTF-16 con serialización de bytes little endian (byte menos significativo al principio).
	ISO-8859-1 Codificación de caracteres ASCII, que suele utilizarse para los idiomas de Europa occidental. También se la conoce como Latin-1.
	ISO-8859-3 Codificación de caracteres ASCII, que suele utilizarse para los idiomas de Europa meridional. También se la conoce como Latin-3.
	ISO-8859-9 Codificación de caracteres ASCII, que suele utilizarse para el idioma turco. También se la conoce como Latin-5.
	CP850 Página de códigos ASCII que se utiliza para escribir idiomas de Europa occidental.
	CP500 Página de códigos EBCDIC que se utiliza para escribir idiomas de Europa occidental.
	Shift_JIS Codificación de caracteres para el idioma japonés.
	MS932 Una extensión de Microsoft de Shift_JIS para incluir caracteres especiales NEC, selección NEC de extensiones IBM y extensiones IBM.
CP1047 Una página de códigos EBCDIC con el conjunto de caracteres completo del alfabeto latino 1.	
Longitud de registro	Para los archivos de ancho fijo, especifica la cantidad exacta de caracteres en cada registro.

Nombre de la opción	Descripción
Separador de campo	<p data-bbox="535 325 1429 388">Especifica el carácter que se utiliza para separar los campos en un archivo delimitado.</p> <p data-bbox="535 399 1429 441">Por ejemplo, este registro utiliza una barra vertical () como separador de campo:</p> <pre data-bbox="552 451 1412 525">7200 13TH ST MIAMI FL 33144</pre> <p data-bbox="535 535 1429 577">Estos caracteres disponibles para definir como separadores de campo son:</p> <ul data-bbox="535 588 779 787" style="list-style-type: none"> • Espacio de claves • Tab • Coma • Punto • Punto y coma • Barra vertical <p data-bbox="535 798 1429 871">Si el archivo utiliza un carácter distinto como separador de campo, haga clic en el botón de los puntos suspensivos para seleccionar otro carácter como delimitador.</p>
Separador de etiqueta	<p data-bbox="535 913 1429 1018">Especifica el carácter ubicado después del campo de etiqueta para demarcar el campo de identificación de cada registro en un archivo delimitado. El separador de etiqueta debe ser un solo carácter.</p> <p data-bbox="535 1029 1429 1092">De forma predeterminada, estos caracteres están disponibles para ser elegidos como separadores de etiqueta:</p> <ul data-bbox="535 1102 779 1312" style="list-style-type: none"> • Espacio de claves • Tab • Coma • Punto • Punto y coma • Barra vertical <p data-bbox="535 1323 1429 1417">Si el archivo utiliza un carácter distinto como separador de etiqueta, haga clic en el botón de los puntos suspensivos para agregar y seleccionar un separador de etiqueta personalizado.</p> <p data-bbox="633 1428 1429 1564">Nota: De forma predeterminada, el carácter Separador de registro es el mismo que el carácter Separador de campo seleccionado. Para activar este campo y seleccionar un carácter diferente, desmarque la casilla Igual que separador de campo.</p>
Igual que separador de campo	<p data-bbox="535 1617 1429 1690">Indica si el separador de etiqueta es el mismo que el separador de campo. Desmárquelo para seleccionar un carácter diferente como separador de etiqueta.</p> <p data-bbox="633 1701 1429 1774">Nota: De manera predeterminada, esta casilla de verificación está marcada y el campo Separador de etiqueta está desactivado.</p>

Nombre de la opción	Descripción
Text qualifier (Calificador de texto)	<p>El carácter que se utiliza para demarcar los valores de texto en un archivo delimitado. Por ejemplo, este registro utiliza comillas dobles (") como calificador de texto.</p> <pre>"7200 13TH ST" "MIAMI" "FL" "33144"</pre> <p>Los caracteres disponibles para definir como calificadores de texto son:</p> <ul style="list-style-type: none"> • Comilla simple (') • Comilla doble (") <p>Si el archivo utiliza un calificador de texto distinto, haga clic en el botón de los puntos suspensivos para seleccionar otro carácter como calificador de texto.</p>
Record separator (Separador de registro)	<p>Especifica el carácter que se utiliza para separar registros en archivos secuenciales o delimitados. Este campo no está disponible si selecciona la casilla Utilizar final de línea predeterminado.</p> <p>Los parámetros de separador de registro disponibles son:</p> <p>Unix (U+000A) Un carácter de salto de línea separa los registros. Este es el separador de registros estándar para los sistemas Unix.</p> <p>Macintosh (U+000D) Un carácter de retorno de carro separa los registros. Este es el separador de registros estándar para los sistemas Macintosh.</p> <p>Windows (U+000D U+000A) Un carácter de retorno de carro seguido por un salto de línea separa los registros. Este el separador de registros estándar para los sistemas Windows.</p> <p>Si su archivo utiliza un separador de registro distinto, haga clic en el botón de los puntos suspensivos para seleccionar otro carácter como separador de registro.</p>
Nombre de etiqueta de raíz	<p>La etiqueta que se usa para registros principales; es decir, de los que dependen otros tipos de registros. Por ejemplo, si usted tiene los tipos de registro 001, 100 y 200, y los tipos de registro 100 y 200 son elementos secundarios de 001, entonces 001 es la etiqueta de raíz.</p>
Usar etiquetas de ancho fijo	<p>Especifica si se asigna una cantidad fija de espacio al principio de cada registro para ubicar la etiqueta de registro. Por ejemplo, a continuación se muestra un archivo que posee las etiquetas 001, 100 y 200 en un campo de ancho fijo:</p> <pre>001 Joe, Smith, M, 100 Main St, 555-234-1290 100 CHK12904567, 12/2/2007, 6/1/2012, CHK 200 1000567, 1/5/2012, Mike's Shoes, 323.12</pre>

Nombre de la opción	Descripción
Posición de inicio de etiqueta	<p>Si marca la casilla Usar etiquetas de ancho fijo, esta opción especifica la posición dentro de cada registro donde comienza la etiqueta. Por ejemplo, si la etiqueta comienza en el cuarto carácter del registro, usted especificará 4.</p>
Ancho de etiqueta	<p>Si marca la casilla Usar etiquetas de ancho fijo esta opción especifica la cantidad de espacios designados para las etiquetas a partir de la posición que se especifica en el campo Posición de inicio de etiqueta. Por ejemplo, si especifica 3 en el campo Posición de inicio de etiqueta y 7 en el campo Ancho de etiqueta, las posiciones 4 a 10 se reservarán para la etiqueta de registro. El valor que especifique debe ser suficiente para incluir todos los caracteres del nombre de etiqueta más largo.</p> <p>El valor en el campo Ancho de etiqueta aumenta automáticamente si usted alarga el nombre de etiqueta en el campo Nombre de etiqueta de raíz.</p> <p>El ancho de etiqueta máximo es 1024.</p>
Use default EOL (Utilizar final de línea predeterminado)	<p>Especifica que el separador de registro del archivo es el carácter de final de línea (EOL) predeterminado que utiliza el sistema operativo en el que se ejecuta el servidor Spectrum™ Technology Platform.</p> <p>No seleccione esta opción si el archivo utiliza un carácter de final de línea distinto del que utiliza de forma predeterminada el sistema operativo del servidor. Por ejemplo, si el archivo utiliza un final de línea de Windows pero el servidor está ejecutándose en Linux, no seleccione esta opción. En su lugar, seleccione la opción Windows en el campo Separador de registro.</p>
Considerar malformados los registros con menos campos que los especificados	<p>Si usted habilita esta opción, se considerarán malformados los registros secundarios que contengan menos campos que un registro completo. Cuando se encuentra un registro malformado, el procesamiento avanza hasta la siguiente etiqueta de raíz, ignorando todas las etiquetas secundarias que haya en el medio. Se escribe una excepción en el registro, con información sobre registros secundarios malformados junto con un número de línea.</p> <p>Nota: Siempre se consideran malformados los registros en las siguientes situaciones, independientemente de si activó esta opción o no.</p> <ul style="list-style-type: none"> • La etiqueta es desconocida • La línea está vacía • Hay una etiqueta sin datos • Si un registro tiene una etiqueta que depende de otra etiqueta, ese registro aparecerá inmediatamente después del registro con la etiqueta de raíz

Pestaña Campos

En la pestaña **Campos** se especifican las características de cada campo leído desde el archivo.

[Definición de los campos en un archivo delimitado de formato variable](#) en la página 209
[Definición de los campos en un archivo secuencial o de ancho fijo de formato variable](#) en la página 213

Ficha Tiempo de ejecución

Nombre de campo	Descripción
Nombre de archivo	Muestra el nombre del archivo seleccionado en la primera ficha.
Starting record (Registro de inicio)	Si desea omitir los registros ubicados al comienzo del archivo cuando lee registros en el flujo de datos, especifique el primer registro que desea leer. Por ejemplo, si desea omitir los primeros 50 registros de un archivo, especifique 51. El registro número 51 será el primero que se leerá en el flujo de datos.
Todos los registros	Seleccione esta opción si desea leer todos los registros, comenzando desde el registro especificado en el campo Registro de inicio hasta el final del archivo.
Máximo de registros	Seleccione esta opción si desea leer solamente cierta cantidad de registros, comenzando desde el registro especificado en el campo Registro de inicio . Por ejemplo, si desea leer los primeros 100 registros, seleccione esta opción e ingrese 100.

Definición de los campos en un archivo delimitado de formato variable

Este procedimiento indica cómo definir campos en la etapa Read from Variable Format File para los archivos delimitados.

1. En la etapa Read from Variable Format File, haga clic en la pestaña **Campos**.
2. Haga clic en **Regenerar**.

Se muestra una lista de todos los campos para cada tipo de registro. Se muestra la siguiente información para cada uno de los campos:

Elemento principal	La etiqueta del archivo de entrada que indica el tipo de registro en el que aparece el campo. Si la etiqueta comienza con un número, se le agrega el prefijo "NumericTag_". Por ejemplo, una etiqueta denominada 100 pasa a ser NumericTag_100. El prefijo es necesario porque los nombres de los campos del flujo de datos no pueden comenzar con números.
---------------------------	---

Campo	El nombre que se utilizará para el campo en el flujo de datos. De forma predeterminada, los campos se nombran con el formato: <TagName>_<Column n>. Por ejemplo, el primer campo del tipo de registro Owner sería Owner_Column1, el segundo sería Owner_Column2, y así sucesivamente.
Tipo	El tipo de datos del campo.

Nota: Los primeros 50 registros se utilizan para generar la lista de campos. El archivo de entrada debe contener al menos dos etiquetas de raíz para poder generar una lista de campos.

3. Si desea modificar las relaciones de elemento principal/ secundario entre las etiquetas:
 - a) Haga clic en **Modificar jerarquía de etiquetas**.
 - b) Haga clic sobre las etiquetas y arrástrelas para definir la jerarquía que desea.
 - c) Haga clic en **Aceptar**.
4. Si desea modificar el nombre o el tipo de datos de un campo, seleccione el campo y haga clic en **Modificar**.
5. En el campo **Nombre**, elija el campo que desea agregar o ingrese el nombre del campo.

Por lo general, es conveniente que reemplace los nombres predeterminados por nombres relevantes para representar los datos en el campo. Por ejemplo, considere estos datos de entrada:

```
001 Joe,Smith,M,100 Main St,555-234-1290
```

Este registro tiene una etiqueta principal de 001 y, de forma predeterminada, se le crearían los siguientes campos:

```
NumericTag_001_Column1: Joe
NumericTag_001_Column2: Smith
NumericTag_001_Column3: M
NumericTag_001_Column4: 100 Main St
NumericTag_001_Column5: 555-234-1290
```

Es conveniente que cambie el nombre de los campos para que describan los datos. Por ejemplo:

```
FirstName: Joe
LastName: Smith
Sexo: M
AddressLine1: 100 Main St
PhoneNumber: 555-234-1290
```

Nota: No puede cambiar el nombre de los campos de lista. Los campos de lista, que contienen todos los campos de un determinado tipo de registro, siempre usan el nombre de etiqueta del archivo de entrada como nombre de campo.

6. Para modificar el tipo de datos de un campo, seleccione el tipo de datos que desea en el campo **Tipo**.

Están disponibles los siguientes tipos de datos:

bigdecimal Un tipo de dato numérico que admite 38 puntos decimales de precisión. Utilice este tipo de datos para datos que se usarán en cálculos matemáticos que requieren un alto grado de precisión, especialmente aquellos que implican datos financieros. El tipo de datos bigdecimal es compatible con cálculos más precisos que el tipo de datos double.

boolean Tipo de dato lógico con dos valores: verdadero y falso

bytearray Una matriz (lista) de bytes.

Nota: Bytearray no se admite como datos de entrada para un servicio REST.

fecha Un tipo de dato que contiene un mes, día y año. Por ejemplo, 2012-01-30 o 30 de enero, 2012. Puede especificar un formato de fecha predeterminado en Management Console.

fecha y hora Un tipo de dato que contiene un mes, día, año y horas, minutos y segundos. Por ejemplo, 2012/01/30 6:15 PM.

double Un tipo de dato numérico que contiene números de doble precisión tanto negativos como positivos entre 2^{-1074} y $(2-2^{-52}) \times 2^{1023}$. En notación científica, el rango de valores es - 1.79769313486232E+308 a 1.79769313486232E+308.

float Un tipo de dato numérico que contiene números de precisión simple tanto negativos como positivos entre 2^{-149} y $(2-2^{-23}) \times 2^{127}$. En notación científica, el rango de valores es - 3.402823E+38 a 3.402823E+38.

integer Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{31} (-2,147,483,648) y $2^{31}-1$ (2,147,483,647).

list En sentido estricto, una lista no es un tipo de dato. Sin embargo, cuando un campo contiene datos jerárquicos, es tratado como un campo "lista". En Spectrum™ Technology Platform una lista es una colección de datos que consiste en múltiples valores. Por ejemplo, un campo Nombres puede contener una lista de valores de nombre, lo que puede representarse mediante una estructura XML de la siguiente forma:

```
<Names>
  <Name>John Smith</Name>
  <Name>Ann Fowler</Name>
</Names>
```

Cabe destacar que el tipo de dato lista de Spectrum™ Technology Platform es diferente del tipo de dato lista del esquema XML puesto que este último es un tipo de dato simple que está compuesto de valores múltiples, mientras que el

tipo de dato lista de Spectrum™ Technology Platform es similar a un tipo de dato complejo del esquema XML.

long	Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{63} (-9,223,372,036,854,775,808) y $2^{63}-1$ (9,223,372,036,854,775,807).
string	Secuencia de caracteres.
hora	Un tipo de dato que contiene la hora del día. Por ejemplo, 21:15:59 o 9:15:59 PM.

7. Si ha seleccionado un tipo de datos de fecha, hora, o numérico, puede usar el formato de fecha/hora predeterminado o puede indicar un formato diferente para este campo específico. El formato predeterminado es, o bien el formato de sistema predeterminado que se haya definido en las opciones de conversión de tipos en Management Console, o bien es el formato predeterminado del flujo de datos especificado en la opciones de conversión de tipos de Enterprise Designer. Se mostrará el formato que se esté usando. Para usar el formato predeterminado, deje seleccionado **Predeterminado**. Para indicar un formato diferente, seleccione **Personalizado** y siga los siguientes pasos:

Nota: Es importante que seleccione un formato de fecha y hora que refleje con exactitud los datos que está leyendo desde el archivo. Por ejemplo, si el archivo contiene datos con el formato de Mes/Día/Año, pero usted selecciona Día/Mes/Año, todos los cálculos de fecha que usted realice en el flujo de datos, como por ejemplo, clasificar según fecha, no reflejarán la fecha correcta. Por otro lado, los registros pueden fallar en el tipo de conversión, en cuyo caso se hará efectivo el comportamiento ante errores que se especifique en las opciones de conversión de tipos en la Consola de administración o en Enterprise Designer.

- En el campo **Configuración regional** seleccione el país cuya convención acerca de formatos usted desea utilizar. Su selección determinará los valores predeterminados en el campo **Formato**. En cuanto a los datos referidos a fechas, su selección determinará además el idioma que se usará al enunciar un mes. Por ejemplo, si ha indicado Inglés, el primer mes del año será "January" pero si ha especificado Francés, este será "Janvier."
- En el campo **Formato**, seleccione el formato de los datos. El formato depende del tipo de datos en el campo. Se proporciona una lista de los formatos de uso más común en la configuración regional seleccionada.

A la derecha del campo **Formato** se muestra un ejemplo del formato seleccionado.

Además puede especificar sus propios formatos de fecha, hora, y número en caso de que los que están disponibles para seleccionar no se ajusten a sus necesidades. Para especificar su propio formato de fecha u hora, ingrese el formato en el campo usando la notación que se describe en [Patrones de fecha y hora](#) en la página 312. Para especificar su propio formato de número, ingrese el formato en el archivo usando la notación que se describe en [Patrones de números](#) en la página 315.

8. Haga clic en **Aceptar**.

Definición de los campos en un archivo secuencial o de ancho fijo de formato variable

Este procedimiento indica cómo definir campos en la etapa Read from Variable Format File para los archivos secuenciales y de ancho fijo.

1. En la etapa Read from Variable Format File, haga clic en la ficha **Campos**.
2. Haga clic en **Obtener etiquetas**.

Se muestra una lista de todos los campos para cada tipo de registro. Se muestra la siguiente información para cada uno de los campos:

Elemento principal	La etiqueta del archivo de entrada que indica el tipo de registro en el que aparece el campo. Si la etiqueta comienza con un número, se le agrega el prefijo "NumericTag_". Por ejemplo, una etiqueta denominada 100 pasa a ser NumericTag_100. El prefijo es necesario porque los nombres de los campos del flujo de datos no pueden comenzar con números.
Campo	El nombre que se utilizará para el campo en el flujo de datos. De forma predeterminada, los campos se nombran con el formato: <TagName>_<Column n>. Por ejemplo, el primer campo del tipo de registro Owner sería Owner_Column1, el segundo sería Owner_Column2, y así sucesivamente.
Tipo	El tipo de datos del campo.

Nota: Los primeros 50 registros se utilizan para generar la lista de campos. El archivo de entrada debe contener al menos dos etiquetas de raíz para poder generar una lista de campos.

3. En el campo **Filtro**, seleccione la etiqueta del tipo de registro cuyos campos desea definir y luego haga clic en **Agregar**.

Nota: El filtro no afecta de modo alguno qué datos se leerán en el flujo de datos. Solamente filtra la lista de campos para que sea más fácil de examinar.

4. En el campo **Nombre**, elija el campo que desea agregar o ingrese el nombre del campo.
5. En el campo **Tipo** puede dejar el tipo de datos como `string` (cadena) si no pretende realizar ninguna operación matemática o de fecha y hora con los datos. Sin embargo, si se propone realizar ese tipo de operaciones, seleccione un tipo de datos apropiado. Esto convertirá los datos de cadena del archivo en un tipo de datos que permitirá el manejo adecuado de los datos en el flujo de datos.

Spectrum™ Technology Platform admite estos tipos de datos:

bigdecimal	Un tipo de dato numérico que admite 38 puntos decimales de precisión. Utilice este tipo de datos para datos que se usarán en cálculos matemáticos que
-------------------	---

requieren un alto grado de precisión, especialmente aquellos que implican datos financieros. El tipo de datos bigdecimal es compatible con cálculos más precisos que el tipo de datos double.

boolean	Tipo de dato lógico con dos valores: verdadero y falso
bytearray	Una matriz (lista) de bytes. Nota: Bytearray no se admite como datos de entrada para un servicio REST.
fecha	Un tipo de dato que contiene un mes, día y año. Por ejemplo, 2012-01-30 o 30 de enero, 2012. Puede especificar un formato de fecha predeterminado en Management Console.
fecha y hora	Un tipo de dato que contiene un mes, día, año y horas, minutos y segundos. Por ejemplo, 2012/01/30 6:15 PM.
double	Un tipo de dato numérico que contiene números de doble precisión tanto negativos como positivos entre 2^{-1074} y $(2-2^{-52}) \times 2^{1023}$. En notación científica, el rango de valores es - 1.79769313486232E+308 a 1.79769313486232E+308.
float	Un tipo de dato numérico que contiene números de precisión simple tanto negativos como positivos entre 2^{-149} y $(2-2^{-23}) \times 2^{127}$. En notación científica, el rango de valores es - 3.402823E+38 a 3.402823E+38.
integer	Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{31} (-2,147,483,648) y $2^{31}-1$ (2,147,483,647).
list	En sentido estricto, una lista no es un tipo de dato. Sin embargo, cuando un campo contiene datos jerárquicos, es tratado como un campo "lista". En Spectrum™ Technology Platform una lista es una colección de datos que consiste en múltiples valores. Por ejemplo, un campo Nombres puede contener una lista de valores de nombre, lo que puede representarse mediante una estructura XML de la siguiente forma: <pre><Names> <Name>John Smith</Name> <Name>Ann Fowler</Name> </Names></pre>
	Cabe destacar que el tipo de dato lista de Spectrum™ Technology Platform es diferente del tipo de dato lista del esquema XML puesto que este último es un tipo de dato simple que está compuesto de valores múltiples, mientras que el tipo de dato lista de Spectrum™ Technology Platform es similar a un tipo de dato complejo del esquema XML.
long	Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{63} (-9,223,372,036,854,775,808) y $2^{63}-1$ (9,223,372,036,854,775,807).
string	Secuencia de caracteres.

hora Un tipo de dato que contiene la hora del día. Por ejemplo, 21:15:59 o 9:15:59 PM.

6. Si ha seleccionado un tipo de datos de fecha, hora, o numérico, puede usar el formato de fecha/hora predeterminado o puede indicar un formato diferente para este campo específico. El formato predeterminado es, o bien el formato de sistema predeterminado que se haya definido en las opciones de conversión de tipos en Management Console, o bien es el formato predeterminado del flujo de datos especificado en la opciones de conversión de tipos de Enterprise Designer. Se mostrará el formato que se esté usando. Para usar el formato predeterminado, deje seleccionado **Predeterminado**. Para indicar un formato diferente, seleccione **Personalizado** y siga los siguientes pasos:

Nota: Es importante que seleccione un formato de fecha y hora que refleje con exactitud los datos que está leyendo desde el archivo. Por ejemplo, si el archivo contiene datos con el formato de Mes/Día/Año, pero usted selecciona Día/Mes/Año, todos los cálculos de fecha que usted realice en el flujo de datos, como por ejemplo, clasificar según fecha, no reflejarán la fecha correcta. Por otro lado, los registros pueden fallar en el tipo de conversión, en cuyo caso se hará efectivo el comportamiento ante errores que se especifique en las opciones de conversión de tipos en la Consola de administración o en Enterprise Designer.

- a) En el campo **Configuración regional** seleccione el país cuya convención acerca de formatos usted desea utilizar. Su selección determinará los valores predeterminados en el campo **Formato**. En cuanto a los datos referidos a fechas, su selección determinará además el idioma que se usará al enunciar un mes. Por ejemplo, si ha indicado Inglés, el primer mes del año será "January" pero si ha especificado Francés, este será "Janvier."
- b) En el campo **Formato**, seleccione el formato de los datos. El formato depende del tipo de datos en el campo. Se proporciona una lista de los formatos de uso más común en la configuración regional seleccionada.

A la derecha del campo **Formato** se muestra un ejemplo del formato seleccionado.

Además puede especificar sus propios formatos de fecha, hora, y número en caso de que los que están disponibles para seleccionar no se ajusten a sus necesidades. Para especificar su propio formato de fecha u hora, ingrese el formato en el campo usando la notación que se describe en [Patrones de fecha y hora](#) en la página 312. Para especificar su propio formato de número, ingrese el formato en el archivo usando la notación que se describe en [Patrones de números](#) en la página 315.

7. En el campo **Posición de inicio**, ingrese la posición del primer carácter del campo, y en el campo **Longitud** ingrese el número de caracteres que tiene el campo.

Por ejemplo, si el campo comienza en el décimo carácter del registro y tiene 5 caracteres de largo, ingresará 10 como posición de inicio y 5 como longitud.

8. Haga clic en **Agregar**.
9. Repita este procedimiento para agregar campos adicionales al tipo de registro o haga clic en **Cerrar** si no desea agregar más campos.

Cómo comprimir datos de formato variable

Los datos de los archivos de formato variable a menudo contienen registros que presentan relaciones jerárquicas, en los que un tipo de registro es el principal del que dependen los demás tipos de registros. Dado que muchas etapas requieren que los datos estén en formato plano, es posible que tenga que comprimir los datos para poder utilizarlos en etapas posteriores. Por ejemplo, considere estos datos de entrada:

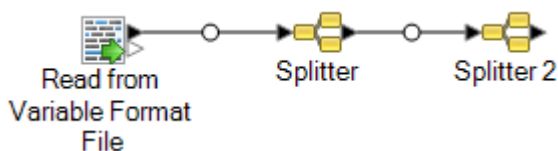
```
001 Joe,Smith,M,100 Main St,555-234-1290
100 CHK12904567,12/2/2007,6/1/2012,CHK
200 1000567,1/5/2012,Fashion Shoes,323.12
001 Anne,Johnson,F,1202 Lake St,555-222-4932
100 CHK238193875,1/21/2001,4/12/2012,CHK
200 1000232,3/5/2012,Blue Goose Grocery,132.11
200 1000232,3/8/2012,Trailway Bikes,540.00
```

Es recomendable que comprima los registros para tener un solo registro por transacción. En el ejemplo anterior, eso significaría tomar los registros de transacción (los que tienen la etiqueta 200) y comprimirlos para que incluyan la información del titular de la cuenta (registros con la etiqueta 001) y los detalles de cuenta (registros con la etiqueta 100).

El siguiente procedimiento describe cómo utilizar las etapas Splitter para comprimir registros.

1. Agregue una etapa Read from Variable Format File a su flujo de datos y configúrela. Para obtener más información, consulte [Read From Variable Format File](#) en la página 202.
2. Agregue una etapa Splitter y conéctela a Read from Variable Format File.
3. Agregue cuantas etapas Splitter sean necesarias hasta tener una etapa Splitter para cada tipo de registro secundario en sus datos de entrada.
4. Conecte todas las etapas Splitter.

Ahora, su flujo de datos debería tener el siguiente aspecto:



5. Haga doble clic en la primera etapa Splitter para abrir las opciones de la etapa.
6. En el campo **División en**, seleccione uno de los tipos de registro secundario.
7. Haga clic en **Aceptar**.
8. Configure cada etapa Splitter adicional al seleccionar un tipo de registro secundario diferente para cada campo **División en** de Splitter.

Read from XML

La etapa Read from XML (Lectura desde XML) lee archivos XML en un trabajo o subflujo. Define la ruta del archivo y el formato de los datos, incluidos el esquema de XML y los detalles de los elementos de datos.

Los elementos XML simples se convierten en archivos planos y se envían a la etapa siguiente. Los datos XML simples consisten en registros compuestos por elementos XML que contienen solamente datos, sin elementos secundarios. Por ejemplo, a continuación se muestra un archivo de datos XML simple:

```
<customers>
  <customer>
    <name>Sam</name>
    <gender>M</gender>
    <age>43</age>
    <country>United States</country>
  </customer>
  <customer>
    <name>Jeff</name>
    <gender>M</gender>
    <age>32</age>
    <country>Canada</country>
  </customer>
  <customer>
    <name>Mary</name>
    <gender>F</gender>
    <age>61</age>
    <country>Australia</country>
  </customer>
</customers>
```

Tenga en cuenta que en este ejemplo, cada registro contiene elementos de XML simples, como <name>, <gender>, <age> y <country>. Ninguno de los elementos contiene elementos secundarios.

La etapa Read from XML comprime automáticamente datos simples como estos porque la mayoría de las etapas requieren que los datos estén en formato plano. Si desea conservar la estructura jerárquica, utilice una etapa Aggregator después de Read from XML para convertir los datos en datos jerárquicos.

Los elementos XML complejos permanecen en formato jerárquico y son enviados como campos de lista. Dado que muchas etapas requieren que los datos estén en formato plano, es posible que tenga que comprimir los elementos XML complejos para poder utilizarlos en etapas posteriores. Para obtener más información, consulte [Compresión de elementos XML complejos](#) en la página 222.

Nota: Read From XML no admite tipos de XML `xs:anyType` y `xs:anySimpleType`.

Ficha Propiedades del archivo

Tabla 3: Ficha Propiedades del archivo

Nombre de la opción	Descripción
Archivo de esquema	<p>Especifica la ruta a un archivo de esquema XSD. Haga clic en el botón de los puntos suspensivos (...) para buscar el archivo que desea. Cabe destacar que el archivo de esquema debe estar en el servidor para que el archivo de datos pueda validarse con el esquema. Si el archivo de esquema no está en el servidor, se deshabilita la validación.</p> <p>De forma alternativa, puede especificar un archivo XML en vez de un archivo XSD. Si especifica un archivo XML, el esquema se inferirá a partir de la estructura del archivo XML. Usar un archivo XML en vez de un XSD tiene las siguientes limitaciones:</p> <ul style="list-style-type: none"> • El archivo XML no puede pesar más de 1 MB. Si el archivo XML pesa más de 1 MB, intente eliminar algunos de los datos sin afectar la estructura del archivo. • El archivo de datos no se validará con el esquema inferido. <p>Nota: Si el servidor de Spectrum™ Technology Platform está ejecutándose en Unix o Linux, recuerde que estas plataformas distinguen las mayúsculas de las minúsculas en los nombres de los archivos y las rutas.</p>
Archivo de datos	<p>Especifica la ruta al archivo de datos XML. Haga clic en el botón de los puntos suspensivos (...) para buscar el archivo que desea.</p> <p>Nota: Si el servidor de Spectrum™ Technology Platform está ejecutándose en Unix o Linux, recuerde que estas plataformas distinguen las mayúsculas de las minúsculas en los nombres de los archivos y las rutas.</p>
Vista previa	<p>Muestra una vista previa del esquema o archivo XML. Cuando especifica un archivo XSD, la estructura de árbol refleja el archivo XSD seleccionado. Una vez que haya especificado tanto el archivo de esquema como el archivo de datos, haga clic en los elementos del esquema que están en negrita para previsualizar los datos que contiene ese elemento.</p>

Ficha Campos

Tabla 4: Ficha Campos

Nombre de la opción	Descripción
Filtro	Filtra la lista de elementos y atributos para que sea más fácil de examinar. El filtro no tiene ningún impacto sobre qué campos se incluirán en la salida. Solamente filtra la lista de elementos y atributos para que sea más fácil de examinar.
XPath	La columna XPath muestra la expresión XPath del elemento o atributo. Se muestra únicamente con fines informativos. Para obtener más información sobre XPath, consulte www.w3schools.com/xpath/ .
Campo	El nombre que se utilizará para el elemento o atributo en el flujo de datos. Para cambiar el nombre del campo, haga doble clic y escriba el nombre que desee.

Nombre de la opción	Descripción
---------------------	-------------

Tipo	
------	--

Nombre de la opción

Descripción

El tipo de datos que se utilizará en el campo.

bigdecimal Un tipo de dato numérico que admite 38 puntos decimales de precisión. Utilice este tipo de datos para datos que se usarán en cálculos matemáticos que requieren un alto grado de precisión, especialmente aquellos que implican datos financieros. El tipo de datos bigdecimal es compatible con cálculos más precisos que el tipo de datos double.

boolean Tipo de dato lógico con dos valores: verdadero y falso

fecha Un tipo de dato que contiene un mes, día y año. Las fechas deben estar en el formato *aaaa-MM-dd*. Por ejemplo, 2012-01-30.

fecha y hora Un tipo de dato que contiene un mes, día, año y horas, minutos y segundos. Datetime debe estar en el formato *aaaa-MM-dd'THH:mm:ss*. Por ejemplo, 2012-01-30T06:15:30

double Un tipo de dato numérico que contiene números de doble precisión tanto negativos como positivos entre 2^{-1074} y $(2-2^{-52}) \times 2^{1023}$. En notación científica, el rango de valores es - 1.79769313486232E+308 a 1.79769313486232E+308.

float Un tipo de dato numérico que contiene números de precisión simple tanto negativos como positivos entre 2^{-149} y $(2-2^{23}) \times 2^{127}$. En notación científica, el rango de valores es - 3.402823E+38 a 3.402823E+38.

integer Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{31} (-2,147,483,648) y $2^{31}-1$ (2,147,483,647).

list En sentido estricto, una lista no es un tipo de dato. Sin embargo, cuando un campo contiene datos jerárquicos, es tratado como un campo "lista". En Spectrum™ Technology Platform una lista es una colección de datos que consiste en múltiples valores. Por ejemplo, un campo Nombres puede contener una lista de valores de nombre, lo que puede representarse mediante una estructura XML de la siguiente forma:

```
<Names>
  <Name>John Smith</Name>
  <Name>Ann Fowler</Name>
</Names>
```

Cabe destacar que el tipo de dato lista de Spectrum™ Technology Platform es diferente del tipo de dato lista del esquema XML puesto que este último es un tipo de dato simple que está compuesto de valores múltiples, mientras que el tipo de dato lista de Spectrum™ Technology Platform es similar a un tipo de dato complejo del esquema XML.

long Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{63} (-9223372036854775808) y $2^{63}-1$

Nombre de la opción	Descripción
	(9223372036854775807).
string	Secuencia de caracteres.
hora	Un tipo de dato que contiene la hora del día. La hora debe estar en el formato <i>HH:mm:ss</i> . Por ejemplo, 21:15:59.
Incluir	Especifica si este campo estará disponible en el flujo de datos o si se lo excluirá.

Ejemplo: Archivo XML simple

En este ejemplo, usted desea leer el siguiente archivo en un flujo de datos:

```
<addresses>
  <address>
    <addressline1>One Global View</addressline1>
    <city>Troy</city>
    <state>NY</state>
    <postalcode>12128</postalcode>
  </address>
  <address>
    <addressline1>1825B Kramer Lane</addressline1>
    <city>Austin</city>
    <state>TX</state>
    <postalcode>78758</postalcode>
  </address>
</addresses>
```

En este ejemplo, podría incluir `<addressline1>`, `<city>`, `<state>` y `<postalcode>`. Como resultado, se crearía un registro para cada elemento `<address>` dado que `<address>` es el elemento padre común para `<addressline1>`, `<city>`, `<state>` y `<postalcode>`.

Compresión de elementos XML complejos

La mayoría de las etapas en un flujo de datos requieren que los datos estén en un formato plano. Esto significa que cuando lea datos jerárquicos de un archivo XML en un flujo de datos, tendrá que aplanarlo si los datos contienen elementos XML complejos. Un elemento XML complejo es un

elemento que contiene otros elementos o atributos. Por ejemplo, en el siguiente archivo de datos, el elemento `<address>` y el elemento `<account>` son elementos XML complejos:

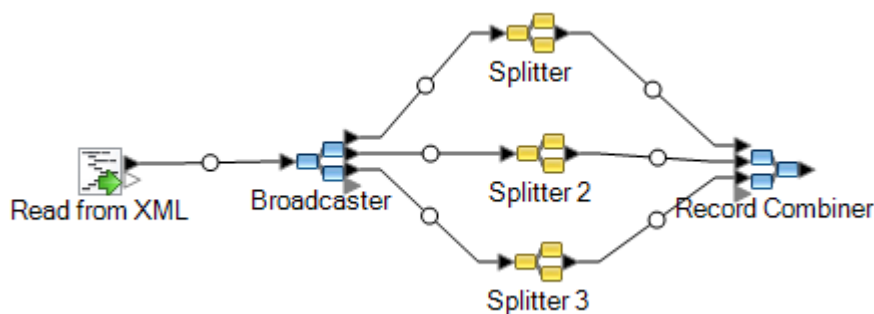
```
<customers>
  <customer>
    <name>Sam</name>
    <gender>M</gender>
    <age>43</age>
    <country>United States</country>
    <address>
      <addressline1>1253 Summer St.</addressline1>
      <city>Boston</city>
      <stateprovince>MA</stateprovince>
      <postalcode>02110</postalcode>
    </address>
    <account>
      <type>Savings</type>
      <number>019922</number>
    </account>
  </customer>
  <customer>
    <name>Jeff</name>
    <gender>M</gender>
    <age>32</age>
    <country>Canada</country>
    <address>
      <addressline1>26 Wellington St.</addressline1>
      <city>Toronto</city>
      <stateprovince>ON</stateprovince>
      <postalcode>M5E 1S2</postalcode>
    </address>
    <account>
      <type>Checking</type>
      <number>238832</number>
    </account>
  </customer>
  <customer>
    <name>Mary</name>
    <gender>F</gender>
    <age>61</age>
    <country>Australia</country>
    <address>
      <addressline1>Level 7, 1 Elizabeth Plaza</addressline1>
      <city>North Sydney</city>
      <stateprovince>NSW</stateprovince>
      <postalcode>2060</postalcode>
    </address>
    <account>
      <type>Savings</type>
      <number>839938</number>
    </account>
  </customer>
</customers>
```

El siguiente procedimiento describe cómo usar las etapas Splitter (Divisor) para comprimir datos XML que contengan varios elementos XML complejos.

Nota: Si sus datos contienen un solo elemento XML complejo, puede utilizar una sola etapa Splitter para comprimirlos, simplemente conectando la etapa Read from XML a la etapa Splitter. No es necesario utilizar las etapas Broadcaster o Record Combiner que se describen en este procedimiento para los archivos de datos que contengan un solo elemento XML complejo.

1. Agregue una etapa Read from XML a su flujo de datos y configúrela. Para obtener más información, consulte [Read from XML](#) en la página 217.
2. Agregue una etapa Broadcaster y conecte a ella la etapa Read from XML.
3. Agregue una etapa Splitter por cada elemento XML complejo que haya en sus datos.
4. Conecte la etapa Broadcaster a cada Splitter.
5. Agregue una etapa Record Combiner y conecte a ella cada Splitter.

Ahora, su flujo de datos debería tener el siguiente aspecto:



6. Haga doble clic en la primera etapa Splitter para abrir las opciones de la etapa.
7. En el campo **Split at**, seleccione uno de los campos complejos. En el archivo de datos del ejemplo anterior, podría ser el campo de dirección.
8. Haga clic en **Aceptar**.
9. Configure cada etapa Splitter adicional seleccionando un elemento XML complejo diferente para cada campo **Split at** de Splitter.

Ahora, el flujo de datos está configurado para tomar entradas XML que contengan registros con elementos XML complejos y comprimir los datos. Los registros que resultan de la etapa Record Combiner pueden enviarse a cualquier etapa que requiera datos planos. Por ejemplo, podría conectar la etapa Record Combiner a una etapa Validate Address para validar la dirección.

SQL Command

SQL Command ejecuta uno o varios comandos SQL para cada registro en el flujo de datos. Puede utilizar el SQL Command para lo siguiente:

- Ejecutar las instrucciones INSERTAR/ACTUALIZAR complejas, como por ejemplo instrucciones que contengan subconsultas/uniones con otras tablas.
- Actualizar tablas luego de insertar/actualizar datos para mantener la integridad de la referencia.
- Actualizar o eliminar un registro en una base de datos antes de cargar un reemplazo de registro.
- Actualizar múltiples tablas en una única transacción.

Puede ejecutar SQL Commands adicionales antes y después de ejecutar los principales SQL Commands, y puede invocar procedimientos almacenados.

Nota: Los procedimientos almacenados invocados desde SQL Command no deben utilizar parámetros OUT.

Nota: Se pueden lograr mejoras significativas del rendimiento mediante el uso de múltiples instancias de ejecución de SQL Command. Para especificar las múltiples instancias de ejecución, haga clic en el botón **Tiempo de ejecución**.

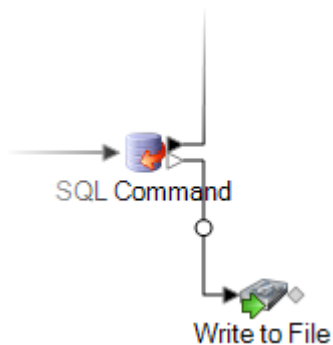
General

La pestaña **General** es donde usted especifica las instrucciones SQL dinámicas que desea ejecutar una vez para cada registro. La siguiente tabla describe las opciones disponibles en la pestaña **General**.

Opción	Descripción
Conexión	<p>Seleccione la conexión de base de datos que quiere utilizar. Las opciones pueden variar según las conexiones definidas en el Administrador de conexión de base de datos de Management Console. Si necesita hacer una conexión de base de datos nueva o modificar o eliminar una conexión de base de datos existente, haga clic en Administrar.</p> <p>Si va a agregar o modificar una conexión de base de datos, complete los siguientes campos:</p> <p>Nombre de conexión Ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.</p> <p>Controlador de base de datos Seleccione el tipo de base de datos correspondiente.</p> <p>Opciones de conexión Especifique el host, el puerto, la instancia, el nombre de usuario y la contraseña que se utilizarán para conectarse con la base de datos.</p>
Instrucciones SQL	<p>Ingrese las instrucciones SQL que desea ejecutar para cada uno de los registros en el flujo de datos. A medida que escriba, se abrirá una ventana emergente con los SQL Commands válidos. Separe varias instrucciones SQL con un punto y coma (;).</p> <p>Para especificar un valor desde un campo del flujo de datos, use esta sintaxis:</p> <pre> \${<field name>} </pre> <p>Donde <field name> es el nombre de un campo en el flujo de datos.</p> <p>Por ejemplo:</p> <pre> UPDATE MyDatabase.dbo.customer SET name=\${Name} WHERE id=\${ID}; </pre> <p>En este ejemplo, \${Name} se reemplazará con el valor del campo Nombre del flujo de datos y \${ID} se reemplazará con el valor del campo ID del flujo de datos.</p> <p>Nota: Las consultas deben usar el nombre cualificado completo. Por ejemplo, MyDatabase.dbo.customer.</p>
Procesamiento de la transacción	<p>Especifica si se van a procesar los registros en los lotes o si se van a procesar todos los registros al mismo tiempo. Una de las siguientes:</p> <p>Tamaño de lote Agrupa los registros en lotes del tamaño que especifique y procesa un lote a la vez.</p> <p>Ejecución total Crea un gran lote para todos los registros y procesa todas las transacciones al mismo tiempo.</p>

Opción	Descripción
Procesamiento de error	<p>Especifica lo que se va a hacer si se encuentra un error al ejecutar los SQL Commands. Una de las siguientes:</p> <p>No terminar el flujo de datos en caso de error El flujo continúa ejecutándose si la base de datos arroja un error al ejecutar los SQL commands.</p> <p>Terminar el flujo de datos después de encontrar esta cantidad de errores El flujo de datos detendrá la ejecución después de que la base de datos arroje el número especificado de errores.</p> <p>Nota: Si hay un error de sintaxis en SQL, el flujo de datos se terminará siempre, independientemente de la configuración que seleccione aquí.</p>

Además, si lo desea, puede escribir registros de error a un receptor al conectar el puerto de error SQL Command al tipo de receptor que desea. El puerto de error es el triángulo blanco en el lado derecho del ícono de la etapa en el flujo de datos. Por ejemplo, para escribir registros de error en un archivo sin formato, debe conectar el puerto de error de SQL Command a una etapa Write to File, como se muestra aquí:



Antes/después de SQL

La pestaña **Pre/Post SQL** es donde especifica las instrucciones SQL que desea ejecutar una vez por flujo de datos, en lugar de una vez por registro como es el caso con el SQL que especifica en la pestaña **General**. La siguiente tabla describe las opciones disponibles en la pestaña **Pre/Post SQL**.

Opción	Descripción
Antes de SQL	<p>Escriba una o más instrucciones SQL que desee ejecutar antes de que los registros que llegan a la etapa se procesen. Las instrucciones SQL que se introduce aquí se ejecutan una vez ejecución, después de que comienza a ejecutarse el flujo de datos, pero antes de que la etapa SQL Command procese los primeros registros.</p> <p>Un ejemplo de uso de Antes de SQL sería crear una tabla para los registros que se procesarán.</p>
Autoconfirmar antes de SQL	<p>Marque esta casilla para confirmar las instrucciones Antes de SQL antes de ejecutar las instrucciones SQL en la pestaña General.</p> <p>Si no marca esta casilla, las instrucciones Antes de SQL se confirmarán en la misma transacción que las instrucciones SQL en la pestaña General.</p> <p>Nota: Si no marca las casillas Autoconfirmar antes de SQL ni Autoconfirmar después de SQL, entonces todas las instrucciones SQL para la etapa se confirman en una sola transacción.</p>
Después de SQL	<p>Escriba una o más instrucciones SQL que desee ejecutar después de que todos los registros se procesen. Las instrucciones SQL que introduce aquí se ejecutan una vez por ejecución después de que la etapa SQL Command haya terminado, pero antes de que termine el flujo de datos.</p> <p>Un ejemplo del uso Antes de SQL sería crear un índice después de procesar los registros.</p>
Autoconfirmar después de SQL	<p>Marque esta casilla para confirmar las instrucciones Después de SQL en su propia transacción después de que se confirmen los SQL commands en la pestaña General.</p> <p>Si no marca esta casilla, las instrucciones Después de SQL se confirmarán en la misma transacción que las instrucciones SQL en la pestaña General.</p> <p>Nota: Si no marca las casillas Autoconfirmar antes de SQL ni Autoconfirmar después de SQL, entonces todas las instrucciones SQL para la etapa se confirman en una sola transacción.</p>

Pestaña *Tiempo de ejecución*

En la pestaña **Tiempo de ejecución** se muestran las **Opciones de etapa** y se le brinda la flexibilidad de definir los valores predeterminados para las opciones de etapa.

Nombre de campo	Descripción
Opciones de etapa	<p>En esta sección, se enumeran las opciones de flujo de datos que se utilizan en la consulta SQL de esta etapa, y le permite brindar un valor predeterminado para todas estas opciones. En la columna Nombre, se enumeran las opciones, y en la columna Valor correspondiente puede ingresar los valores predeterminados.</p> <p>Nota: El valor predeterminado que se brinda aquí también se muestra en la sección Opciones de flujo de datos de mapa para las etapas del cuadro de diálogo Opciones de flujo de datos. El cuadro de diálogo también le permite cambiar el valor predeterminado. En caso de conflicto entre los valores predeterminados que se proporcionan a través de Opciones de etapa, Opciones de flujo de datos y Ejecutor de trabajos, el orden de prioridad es: el valor que se proporciona a través de Ejecutor de trabajos > el valor que se define a través del cuadro de diálogo Opciones de flujo de datos > el valor que se ingresa a través de Opciones de etapa.</p>

Especificación de SQL Command en el Tiempo de ejecución

Este procedimiento describe cómo configurar un flujo de datos para que admita opciones de tiempo de ejecución para SQL Command, y también cómo especificar los argumentos del ejecutor de trabajos para hacerlo.

1. Abra el flujo de datos en Enterprise Designer.
2. Si desea configurar las opciones de tiempo de ejecución para una etapa en un flujo de datos incrustado, abra este último.
3. Haga clic en el ícono Opciones de flujo de datos en la barra de herramientas o haga clic en **Editar > Opciones de flujo de datos**. Aparecerá el cuadro de diálogo **Opciones de flujo de datos**.
4. Haga clic en **Agregar**. Aparecerá el cuadro de diálogo **Definir opción de flujo de datos**.
5. Expanda la etapa de SQL Command.
6. Seleccione una opción SQL Command. Puede ser **PreSqlCommand**, **SqlCommand** o **PostSqlCommand**.

PreSqlCommand Instrucciones SQL que desee ejecutar antes de que los registros que llegan a la etapa se procesen. Estas instrucciones SQL se ejecutan una vez por ejecución, después de que comienza a ejecutarse el flujo de datos, pero antes de que la etapa SQL Command procese el primer registro.

Un ejemplo de uso de Antes de SQL sería crear una tabla para los registros que se procesarán.

SqlCommand Las instrucciones SQL que desee ejecutar para cada uno de los registros en el flujo de datos.

PostSqlCommand Las instrucciones SQL que desee ejecutar después de que se procesen todos los registros. Estas instrucciones SQL se ejecutan una vez por ejecución después de que la etapa SQL Command haya terminado, pero antes de que termine el flujo de datos.

Un ejemplo del uso Posterior de SQL sería crear un índice después de procesar los registros.

El nombre de la opción de SQL Command seleccionada aparece en los campos **Nombre de opción** y **Rótulo de opción**. Este es el nombre de opción que deberá especificarse durante el tiempo de ejecución para definir esta opción.

7. Ingrese una descripción de la opción en el campo **Descripción**.
8. En el campo **Destino**, seleccione la opción **Etapas seleccionadas**.
9. Si desea limitar los valores que pueden especificarse durante la ejecución, edite las opciones en el campo **Valores válidos** haciendo clic en el icono a la derecha del campo.
10. Si desea cambiar el valor predeterminado, especifique un valor diferente en el campo **Valor predeterminado**.

Nota: En el caso de un servicio, solo puede modificar los valores predeterminados antes de exponer el servicio por primera vez. Después de exponer el servicio, ya no es posible modificar los valores predeterminados con Enterprise Designer. En cambio, debe usar Management Console. Para obtener más información, consulte [Especificación de las opciones de servicio predeterminadas](#).

11. Haga clic en **Aceptar**.
12. Continúe agregando opciones, según sus preferencias.
13. Haga clic en **ACEPTAR** en el cuadro de diálogo **Opciones de flujo de datos** cuando haya terminado de agregar opciones.
14. Guarde y exponga el flujo de datos.
15. Cree un archivo de texto que contenga la Instrucción SQL que desea usar en el tiempo de ejecución.

Posiblemente el archivo de texto tenga la siguiente apariencia:

```
SqlCommand = UPDATE CustomersSET
ContactName='Alfred Schmidt'
City='Hamburg'
WHERE CustomerName='Alfreds Futterkiste';
```

En este ejemplo, SqlCommand es uno de los nombres de opción de la etapa de SQL Command.

16. Use el argumento -o cuando ejecute un ejecutor de trabajo desde la línea de comando.

```
java -jar jobexecutor.jar -h "noipa019sh-11" -u "admin" -p "admin" -s
"8080" -o "options.txt" -j "FetchOracleData" -w
```

El nombre de archivo (options.txt) especifica un nombre del archivo de texto que creó en el paso 14.

Para obtener más información, consulte [Ejecución de un trabajo desde la línea de comandos](#) en la página 231

Ejecución de un trabajo desde la línea de comandos

Antes de que pueda ejecutar un trabajo desde la línea de comandos, debe estar expuesto. Para exponer un trabajo, abra el trabajo en Enterprise Designer y seleccione **Archivo > Exponer/Anular exposición y Guardar**.

Para ejecutar un trabajo desde la línea de comandos, debe instalar la utilidad del ejecutor de trabajos del sistema en que desea ejecutar el trabajo. Job Executor está disponible en la página de bienvenida de Spectrum™ Technology Platform en el servidor Spectrum™ Technology Platform (por ejemplo, <http://myserver:8080>).

Uso

```
java -jar jobexecutor.jar -u UserID -p Password -j Job [Optional Arguments]
```

Requerido	Argumento	Descripción
No	-?	Imprime información de uso.
No	-d <i>delimiter</i>	Configura un delimitador de instancia/estado. Solo aparece en la salida sincronizada.
No	-e	Utiliza una conexión HTTPS segura para comunicarse con el servidor Spectrum™ Technology Platform.
No	-f <i>property file</i>	Especifica una ruta a un archivo de propiedades de un trabajo. Un archivo de propiedades del trabajo contiene argumentos del ejecutor de trabajos. Para obtener más información sobre archivos de propiedades de trabajos, consulte Utilización de un archivo de propiedades de trabajo .
No	-h <i>host name</i>	Especifica el nombre o la dirección IP del servidor Spectrum™ Technology Platform.
No	-i <i>poll interval</i>	Especifica la frecuencia con que se verifican los trabajos completos, en segundos. Solo se aplica en el modo sincronizado.
Sí	-j <i>job name</i>	Lista de trabajos por ser ejecutados, separados por comas. Los nombres de los trabajos distinguen mayúsculas de

Requerido	Argumento	Descripción
		minúsculas. Los trabajos comienzan según el orden de la lista.
No	-n <i>email list</i>	Especifica una lista de direcciones de correo electrónico adicionales separadas por comas para las notificaciones de trabajos configuradas.
No	-o <i>property file</i>	<p>Especifica una ruta a un archivo de propiedades de opciones de flujo de datos. Use un archivo de propiedades de opciones de flujo de datos para configurar las opciones de las etapas del flujo de datos. Para establecer las opciones de flujo de datos con un archivo de propiedades, debe configurar el flujo de datos para exponer las opciones de etapa durante el tiempo de ejecución. Para obtener más información, consulte Incorporación de opciones de tiempo de ejecución para el flujo de datos.</p> <p>Por ejemplo, un archivo de propiedades de opciones de flujo de datos para un flujo de datos que contiene la etapa Assign GeoTAX Info puede tener el siguiente aspecto:</p> <pre>OutputCasing=U UseStreetLevelMatching=N TaxKey=T Database.GTX=gsl</pre>
Sí	-p <i>password</i>	La contraseña del usuario.
No	-r	<p>Especifique este argumento para devolver un informe detallado sobre el trabajo. Esta opción solo funciona si también especifica --w. El informe ofrece la siguiente información:</p> <ul style="list-style-type: none"> • Posición 1: nombre del trabajo • Posición 2: ID de procesamiento del trabajo • Posición 3: estado • Posición 4: fecha/hora de comienzo (MM/DD/AAAA HH:MM:SS) • Posición 5: fecha/hora de finalización (MM/DD/AAAA HH:MM:SS) • Posición 6: número de registros exitosos • Posición 7: número de registros fallidos

Requerido	Argumento	Descripción
		<ul style="list-style-type: none"> • Posición 8: número de registros malformados • Posición 9: registros que actualmente no están en uso <p>Por ejemplo:</p> <pre>MySimpleJob 4 succeeded 04/09/2010 14:50:47 04/09/2010 14:50:47 100 0 0 </pre> <p>La información se delimita utilizando un delimitador especificado en el argumento <code>-d</code>.</p>
No	<code>-s port</code>	Socket (puerto) en el que se ejecuta el servidor Spectrum™ Technology Platform. El valor predeterminado es 8080.
No	<code>-t timeout</code>	Establece el límite de espera (en segundos) para el modo sincronizado. El valor predeterminado es 3600. El valor máximo es 2147483. El límite de espera es global y agregado, y representa el tiempo máximo que se esperará para que se completen todos los trabajos generados.
Sí	<code>-u user name</code>	Nombre de inicio de sesión del usuario.
No	<code>-v</code>	Devuelve una salida detallada.
No	<code>-w</code>	<p>Ejecuta el ejecutor de trabajos en modo sincronizado. Esto significa que el ejecutor de trabajos permanece en ejecución hasta que el trabajo finaliza.</p> <p>Si no especifica <code>-w</code>, el ejecutor de trabajos deja de funcionar después de empezar el trabajo, a menos que el trabajo lea desde los archivos o les escriba en el servidor. En este caso, el ejecutor de trabajos se ejecutará hasta que todos los archivos locales se procesen y luego dejará de funcionar.</p>
No	<code>StageName=Protocol:FileName</code>	Anula el archivo de entrada o de salida especificado en Read from File o Write to File. Para obtener más información, consulte Sobrescribir ubicaciones de archivos de trabajo .

Requerido	Argumento	Descripción
No	<code>StageName:schema=Protocol:SchemaFile</code>	Anula la definición de diseño de un archivo especificada en Read from File o Write to File y la reemplaza por una definida en un archivo de esquema. Para obtener más información, consulte Reemplazo del formato de archivo en la línea de comandos .

Ejemplo de utilización de Job Executor

Los siguientes ejemplos muestran la salida e invocación de una línea de comandos:

```
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
```

En este ejemplo, la salida indica que el trabajo denominado 'validateAddressJob1' se ejecutó (con el identificador 105) sin errores. Otro posible resultado es "falló" o "en ejecución".

Ejecución de comandos SQL antes o después de un flujo de datos

La actividad **Execute SQL** realiza operaciones en la base de datos en cualquier punto durante un flujo de proceso. Esta actividad le permite ejecutar las instrucciones SQL antes y después de la ejecución del Spectrum™ Technology Platform flujo de datos o un programa externo. Por ejemplo, la actividad **Execute SQL** se puede utilizar para eliminar índices antes de la ejecución de un flujo de datos Spectrum™ Technology Platform y para volver a crear índices después de la ejecución del flujo de datos. Para ejecutar instrucciones SQL utilizando la actividad **Execute SQL**, debe crear un flujo de proceso.

Nota: Consulte la *Guía de Dataflow Designer* para obtener instrucciones acerca de cómo crear y programar flujos de procesos.

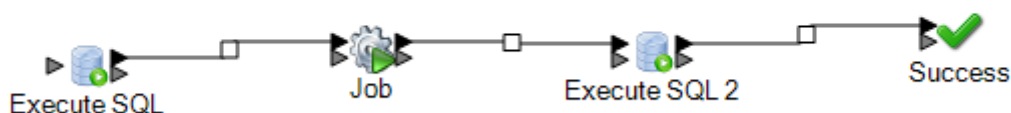
1. Arrastre la actividad **Ejecutar SQL** al lienzo.
2. Haga doble clic en la actividad **Execute SQL**.
3. Seleccione una conexión de base de datos que quiere utilizar.

Si necesita hacer una conexión de base de datos nueva o modificar o eliminar una conexión de base de datos existente, haga clic en **Administrar**.

Si va a agregar o modificar una conexión de base de datos, complete los siguientes campos:

Nombre de conexión	Ingrese un nombre para la conexión. El nombre puede ser cualquiera que elija.
Controlador de base de datos	Seleccione el tipo de base de datos correspondiente.
Opciones de conexión	Especifique el host, el puerto, la instancia, el nombre de usuario y la contraseña que se utilizarán para conectarse con la base de datos.

4. Escriba la instrucción SQL en la casilla **Instrucciones SQL**.
De manera predeterminada, la opción **Finalizar flujo en caso de error** está marcada, lo que significa que el proceso se finalizará si se produce una excepción. Si la opción **Finalizar flujo en caso de error** no está marcada y se produce una excepción, el flujo de proceso no se detendrá y la excepción se registrará en los registros del servidor.
5. Agregue la acción que desee que ejecute un flujo de proceso.
Puede agregar un trabajo al arrastrar el icono del trabajo al lienzo, o agregar un programa externo al arrastrar el icono Ejecutar programa al lienzo.
6. Conecte las dos actividades.
7. Agregue una actividad **Execute SQL** adicional si es necesario.
Consulte el paso 2 al 5 para realizar acciones en **Execute SQL**.
8. Cuando haya agregado todos los trabajos y las actividades de Ejecutar programa y Execute SQL que desee ejecutar en el flujo de proceso, arrastre una actividad Sin errores en el lienzo y conéctela a la última actividad en el flujo de



~~procc~~

9. Ejecute el flujo de proceso.

Transposer

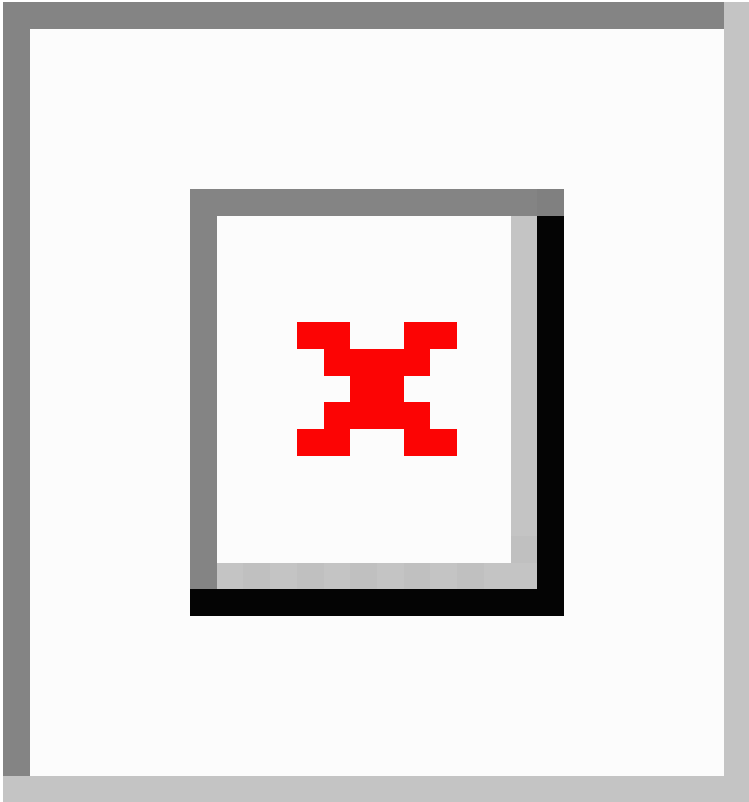
Transposer convierte columnas a filas. La transposición de datos es lo opuesto a crear tablas con la etapa Group Statistics, que transforma datos de filas en columnas.

Para entender Transposer, considere el siguiente ejemplo. Una tabla contiene datos de ventas de cuatro trimestres y desea agregar todos los ingresos generados y analizar el crecimiento alcanzado en los tres primeros trimestres. Para lograr esto, use Transposer para crear una columna que

contenga todos los ingresos de tres trimestres transpuestos. Usar Transposer para agregar todos los ingresos generados de distintas columnas en una sola, puede mejorar el rendimiento en lugar de agregarlos en columnas diferentes.

En la siguiente tabla se explican las opciones del cuadro de diálogo de Transposer.

Opción	Descripción
Encabezado de campos de Transpuesto	Escriba un encabezado para la columna que va a contener las columnas que se van a transponer. Esta nueva columna se añade de manera automática al flujo de datos.
Encabezado de los valores transpuestos	Escriba un encabezado para la columna que va a contener los valores transpuestos de las columnas. Esta nueva columna se añade de manera automática al flujo de datos.
Conservar los campos transpuestos	Seleccione esta opción para conservar todos los campos transpuestos como columnas en la salida.
Nombre de campo	Muestra todos los encabezados de columnas del archivo de entrada.

Opción	Descripción
Tipo	<p>Muestra el tipo de datos de los campos correspondientes (encabezados de columna).</p> <p>Las columnas que se van a transponer deben ser compatibles con el tipo de datos en el archivo de entrada de origen. A continuación, se encuentra la matriz de compatibilidad. Las cuadrículas con la marca de verificación corresponden a los tipos de datos compatibles.</p> 
Transpuesto	<p>Active la casilla situada junto a cada uno de los campos que desea convertir en una columna. Con el fin de evitar que una columna se transponga y conservarla en la salida, desactive la casilla.</p>

Ejemplo del uso de Transposer

Los siguientes datos de entrada contienen datos de ventas de cuatro trimestres por tienda. Tenga en cuenta que Q1, Q2, Q3 y Q4 representan cuatro trimestres de ventas (en millones).

Almacénar (EE.UU.)	Q1	Q2	Q3	Q4
Nueva York	100.00	200.10	300.00	400.00
California	250.10	450.00	550.00	650.00
Illinois	150.00	250.10	350.00	450.00

Los casos que se mencionan a continuación ilustran el comportamiento de Transposer mediante las opciones que existen en la etapa. Tenga en cuenta que Trimestre es el nombre de la columna para el encabezado de los campos transpuestos y que Ingresos es el nombre de la columna para los valores de los campos transpuestos.

Caso 1

Supongamos que desea que las columnas Q1, Q2 y Q3 se transpongan y que Q4 se conserve en la salida. Para ello, active la casilla situada debajo del encabezado **Transpuesto** junto a cada columna que se transpondrá. Ahora verá Q1, Q2 y Q3 como filas, mientras que Q4 seguirá siendo una columna en la salida.

Almacénar (EE.UU.)	Trimestre	Ingresos	Q4
Nueva York	Q1	100.00	400.00
Nueva York	Q2	200.10	400.00
Nueva York	Q3	300.00	400.00
California	Q1	250.10	650.00
California	Q2	450.00	650.00
California	Q3	550.00	650.00
Illinois	Q1	150.00	450.00
Illinois	Q2	250.10	450.00
Illinois	Q3	350.00	450.00

Caso 2

Supongamos que desea que las columnas Q1 y Q2 se transpongan y que Q3 y Q4 se conserven en la salida. Además, también desea conservar todos los campos transpuestos (Q1 y Q2) en forma de columnas en la salida. Para ello, marque la opción **Conservar los campos transpuestos** y el cuadro bajo el encabezado **Transpuesto** junto a cada columna que se transpondrá. Ahora puede ver Q1 y Q2 como filas, mientras que Q3 y Q4 seguirán siendo columnas en la salida junto con Q1 y Q2.

Almacén (EE.UU.)	Trimestre	Ingresos	Q1	Q2	Q3	Q4
Nueva York	Q1	100.00	100.00	200.10	300.00	400.00
Nueva York	Q2	200.10	100.00	200.10	300.00	400.00
California	Q1	250.10	250.10	450.00	550.00	650.00
California	Q2	450.00	250.10	450.00	550.00	650.00
Illinois	Q1	150.00	150.00	250.10	350.00	450.00
Illinois	Q2	250.10	150.00	250.10	350.00	450.00

Unique ID Generator

La etapa Unique ID Generator crea una clave única que se utiliza para identificar un registro específico. La ID única es esencial para las iniciativas de almacenamiento de datos en las que las transacciones no incluyan todos los datos de nombre y dirección, pero deben estar atribuidas al mismo registro o contacto. Es posible implementar una ID única a nivel individual, de hogares, de empresas o de lugar. Unique ID Generator ofrece una variedad de algoritmos para crear ID únicas.

La ID única está basada en un número secuencial o una marca de fecha y hora. Además, se puede utilizar de manera opcional una variedad de algoritmos para generar datos para anexar a la ID, lo que aumenta la probabilidad de que la ID sea única. Las ID que están constituidas por un número secuencial o una marca de fecha y hora son necesarias y no pueden eliminarse de la ID generada.

Unique ID Generator se puede utilizar para generar una clave no exclusiva usando alguno de los algoritmos de generación de claves. En el modo no exclusivo, puede crear claves para utilizar en

los cruces. Esto puede ser útil en un almacén de datos en el que se han agregado claves a una dimensión y se desea generar una clave para los nuevos registros con el fin de ver si los nuevos registros coinciden con un registro existente.

El siguiente ejemplo muestra que cada registro de la entrada está asignado a una ID secuencial de registro en la salida.

Inicio de sesión	RecordID
John Smith	0
Mary Smith	1
Jane Doe	2
John Doe	3

La etapa Unique ID produce un campo llamado RecordID que contiene el identificador exclusivo. Puede cambiar el nombre del campo RecordID según sea necesario.

Definición de una ID única

De forma predeterminada, la etapa Unique ID Generator crea una ID secuencial, con el primer registro que tiene una ID de 0, el segundo con una ID de 1, el tercer registro con una ID de 2, y así sucesivamente. Si desea cambiar la forma en que se genera la ID única, siga este procedimiento.

1. En la etapa Unique ID Generator, en la ficha **Reglas**, haga clic en **Modificar**.
2. Elija el método que desea utilizar para generar la ID única.

Opciones	Descripción
La etiqueta numérica secuencial comenzará en	Asigna un valor numérico a cada registro que comienza con el número que usted especifique. Si se especifica 0 , el primer registro tendrá una ID de 0, el segundo tendrá una ID de 1 y así sucesivamente.

Opciones

Descripción

La etiqueta secuencial numérica comienza en el valor de un campo en una base de datos

Opciones

Descripción

Asigna un valor numérico incremental a cada uno de los registros que comiencen con la lectura del número máximo del campo de la base de datos. Este número se incrementa en 1 y se asigna al primer registro. Por ejemplo, si el número leído desde el campo de la base de datos es 30, el primer registro tendrá una ID de 31, el segundo tendrá una ID de 32, y así sucesivamente.

Conexión Seleccione la conexión de base de datos que quiere utilizar. Las opciones pueden variar según las conexiones definidas en el Administrador de conexión de base de datos de Management Console. Si necesita hacer una nueva base de datos o modificar o eliminar una conexión existente, haga clic en **Administrar**. Si va a agregar o modificar una conexión de base de datos, complete los siguientes campos:

Nombre de conexión

Ingrese un nombre para la conexión. Puede ser cualquier nombre que usted elija.

Controlador de base de datos

Seleccione el tipo de base de datos correspondiente.

Opciones de conexión

Especifique el host, el puerto, la instancia, el nombre de usuario y la contraseña que se utilizarán para conectarse con la base de datos.

Vista de tabla Especifique la tabla o vista en la base de datos que se desea consultar.

Campo Base de datos Seleccione una columna de la lista para generar una clave única. Los tipos de datos admitidos para la generación de ID única son:

long Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{63} ($-9,223,372,036,854,775,808$) y $2^{63}-1$ ($9,223,372,036,854,775,807$).

integer Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{31} ($-2,147,483,648$) y $2^{31}-1$ ($2,147,483,647$).

bigdecimal Un tipo de dato numérico que admite 38 puntos decimales de precisión. Utilice este tipo de datos para datos que se usarán en cálculos matemáticos que requieren un alto grado de precisión, especialmente aquellos que implican datos financieros. El tipo de datos bigdecimal es compatible con cálculos más precisos que el tipo de datos double.

double Un tipo de dato numérico que contiene números de doble precisión tanto negativos como positivos entre 2^{-1074} y $(2-2^{-52}) \times 2^{1023}$. En notación científica, el rango de valores es $-1.79769313486232E+308$

Opciones	Descripción
	a 1.79769313486232E+308.
float	Un tipo de dato numérico que contiene números de precisión simple tanto negativos como positivos entre 2^{-149} y $(2-2^{23}) \times 2^{127}$. En notación científica, el rango de valores es - 3.402823E+38 a 3.402823E+38.
Marca de fecha/hora	Crea una clave única basada en la fecha y la marca de hora en lugar de una numeración secuencial.
UUID	Crea una clave única de identificador de 32 dígitos de forma universal para cada registro. Estos dígitos en la clave se muestran en cinco grupos separados por guiones, en la forma 8-4-4-4-12 para un total de 36 caracteres (32 caracteres alfanuméricos y cuatro guiones). Por ejemplo: 123e4567-e89b-12d3-a456-432255330000
Desactivar	Seleccione esta opción solo si desea generar una clave no única mediante un algoritmo.

- Haga clic en **Aceptar**.

Uso de algoritmos para expandir una ID única

Unique ID Generator genera una ID única para cada registro con una numeración secuencial para cada registro o una marca de fecha y hora para cada registro. Opcionalmente, puede utilizar algoritmos para agregar información adicional a la ID única secuencial o de fecha y hora, creando así una ID única más compleja que es más probable que sea realmente irreplicable.

- En Unique ID Generator, haga clic en **Agregar**.
- En el campo **Algoritmo**, seleccione el algoritmo que desea utilizar para generar información adicional en la ID. Una de las siguientes:

Consonante	Devuelve campo especificados con consonantes eliminadas.
Doble metaphone	Devuelve un código basado en una representación fonética de sus caracteres. Double Metaphone es una versión mejorada del algoritmo Metaphone e intenta explicar la gran cantidad de irregularidades encontradas en los distintos idiomas.
Koeln	Indexa los nombres por sonido, tal como se los pronuncia en alemán. Permite que los nombres que tienen la misma pronunciación se codifiquen con la misma representación, de modo tal que puedan cruzarse a pesar de las

pequeñas diferencias en el modo en que se escriben. El resultado es siempre una secuencia de números; los caracteres especiales y los espacios en blanco se ignoran. Esta opción se desarrolló para responder a las limitaciones de Soundex.

MD5	Un algoritmo de síntesis del mensaje que produce un valor hash de 128 bits. Este algoritmo suele usarse para comprobar la integridad de los datos.
Metaphone	Devuelve una clave codificada Metaphone de campos seleccionados. Metaphone es un algoritmo para codificar palabras utilizando su pronunciación del inglés.
Metaphone (español)	Devuelve una clave codificada Metaphone de campos seleccionados para el idioma español. El algoritmo Metaphone codifica palabras con base en su pronunciación en español.
Metaphone 3	Mejora los algoritmos de Metaphone y Double Metaphone con parámetros más exactos de consonantes y vocales internas que permiten pronunciar palabras y nombres que, fonéticamente, coinciden en mayor o menor medida con los términos de búsqueda. Metaphone 3 aumenta la precisión de la codificación fonética en un 98%. Esta opción se desarrolló para responder a las limitaciones de Soundex.
Nysiis	Algoritmo de codificación fonética que cruza una pronunciación aproximada con una palabra escrita exacta e indexa palabras que se pronuncian de manera similar. Parte del sistema de inteligencia e identificación del estado de Nueva York (New York State Identification and Intelligence System). Supongamos que está buscando información sobre alguien en una base de datos de personas. Cree que el nombre de esa persona suena algo así como "John Smith", pero en realidad se escribe "Jon Smyth". Si busca una coincidencia exacta para "John Smith" no obtendrá resultados. Sin embargo, si indexa la base de datos con el algoritmo de NYSIIS y realiza la búsqueda usando ese mismo algoritmo, obtendrá una coincidencia correcta porque el algoritmo indexa "John Smith" y "Jon Smyth" como "JAN SNATH".
Phonix	Preprocesa cadenas de nombres aplicando más de 100 reglas de transformación a caracteres únicos o secuencias de varios caracteres. Diecinueve de esas reglas se aplican solo si los caracteres se encuentran al comienzo de la cadena, mientras que 12 de ellas se aplican solo cuando los caracteres se encuentran en medio de la cadena, y 28 se aplican únicamente si están al final de la cadena. La cadena de nombre transformada se codifica en un código compuesto por la primera letra seguida de tres dígitos (sin ceros ni números duplicados). Esta opción se desarrolló como respuesta a las limitaciones de Soundex; es más compleja y, por lo tanto, más lenta que esta última.
Soundex	Devuelve un código Soundex de campos seleccionados. El algoritmo Soundex produce un código de longitud fija basado en la pronunciación en inglés de las palabras.

Subcadena de caracteres Devuelve una porción específica del campo seleccionado.

3. En el campo **Campos de entrada**, seleccione el campo al que desee aplicar el algoritmo. Por ejemplo, si selecciona el algoritmo soundex y elige un campo denominado Ciudad, el código se generaría aplicando el algoritmo soundex para los datos en el campo Ciudad.
4. Si seleccionó el algoritmo de subcadena, especifique la parte del campo que desea utilizar en la subcadena:
 - a) En el campo **Posición de inicio**, especifique la posición en el campo en el que desea que comience la subcadena.
 - b) En el campo **Longitud**, seleccione el número de caracteres desde la posición de inicio que desea incluir en la subcadena.

Por ejemplo, supongamos que tenemos los siguientes datos en un campo llamado LastName:

Augustine

Si especifica 3 como la posición de inicio y 6 como la posición final, la subcadena sería:

gustin

5. Marque la casilla **Eliminar el ruido de caracteres** para eliminar todos los caracteres no numéricos ni alfanuméricos, como guiones, espacios en blanco y otros caracteres especiales del campo antes de aplicar el algoritmo.
6. Para algoritmos de subcadena y consonantes, puede ordenar los datos en el campo antes de aplicar el algoritmo de control de la casilla **Ordenar entrada**. Después, puede optar por ordenar los caracteres o los términos en el campo en orden alfabético.
7. Haga clic en **Aceptar** para guardar la configuración.
8. Repita los pasos las veces que sea necesario si desea agregar algoritmos para producir una ID más compleja.

Nota: La definición de clave única siempre se muestra en un color distinto y no se puede borrar.

Definición de una ID no exclusiva

Unique ID Generator se puede utilizar para generar una clave no exclusiva usando alguno de los algoritmos de generación de claves. En el modo no exclusivo, puede crear claves para utilizar en los cruces. Esto puede ser útil en un almacén de datos en el que se han agregado claves a una dimensión y se desea generar una clave para los nuevos registros con el fin de ver si los nuevos registros coinciden con un registro existente.

1. En la etapa Unique ID Generator, en la ficha **Reglas**, haga clic en **Modificar**.

2. Seleccione **Desactivar**.

Esto desactiva la porción de ID única de las reglas de generación de ID. Con esta porción desactivada, solo se usará el algoritmo que elija en los siguientes pasos para crear la ID. esto quiere decir que cualquier registro que tenga los mismos datos en los campos que use para generar la ID, tendrán la misma ID. Posteriormente, puede usar la ID para tareas de comparación.

3. Haga clic en **Aceptar**.4. En el aviso de advertencia, haga clic en **Sí**.5. En Unique ID Generator, haga clic en **Agregar**.6. En el campo **Algoritmo**, seleccione el algoritmo que desea utilizar para generar información adicional en la ID. Una de las siguientes:

Consonante	Devuelve campo especificados con consonantes eliminadas.
Doble metaphone	Devuelve un código basado en una representación fonética de sus caracteres. Double Metaphone es una versión mejorada del algoritmo Metaphone e intenta explicar la gran cantidad de irregularidades encontradas en los distintos idiomas.
Koeln	Indexa los nombres por sonido, tal como se los pronuncia en alemán. Permite que los nombres que tienen la misma pronunciación se codifiquen con la misma representación, de modo tal que puedan cruzarse a pesar de las pequeñas diferencias en el modo en que se escriben. El resultado es siempre una secuencia de números; los caracteres especiales y los espacios en blanco se ignoran. Esta opción se desarrolló para responder a las limitaciones de Soundex.
MD5	Un algoritmo de síntesis del mensaje que produce un valor hash de 128 bits. Este algoritmo suele usarse para comprobar la integridad de los datos.
Metaphone	Devuelve una clave codificada Metaphone de campos seleccionados. Metaphone es un algoritmo para codificar palabras utilizando su pronunciación del inglés.
Metaphone (español)	Devuelve una clave codificada Metaphone de campos seleccionados para el idioma español. El algoritmo Metaphone codifica palabras con base en su pronunciación en español.
Metaphone 3	Mejora los algoritmos de Metaphone y Double Metaphone con parámetros más exactos de consonantes y vocales internas que permiten pronunciar palabras y nombres que, fonéticamente, coinciden en mayor o menor medida con los términos de búsqueda. Metaphone 3 aumenta la precisión de la codificación fonética en un 98%. Esta opción se desarrolló para responder a las limitaciones de Soundex.
Nysiis	Algoritmo de codificación fonética que cruza una pronunciación aproximada con una palabra escrita exacta e indexa palabras que se pronuncian de manera similar. Parte del sistema de inteligencia e identificación del estado de Nueva

York (New York State Identification and Intelligence System). Supongamos que está buscando información sobre alguien en una base de datos de personas. Cree que el nombre de esa persona suena algo así como "John Smith", pero en realidad se escribe "Jon Smyth". Si busca una coincidencia exacta para "John Smith" no obtendrá resultados. Sin embargo, si indexa la base de datos con el algoritmo de NYSIIS y realiza la búsqueda usando ese mismo algoritmo, obtendrá una coincidencia correcta porque el algoritmo indexa "John Smith" y "Jon Smyth" como "JAN SNATH".

Phonix Preprocesa cadenas de nombres aplicando más de 100 reglas de transformación a caracteres únicos o secuencias de varios caracteres. Diecinueve de esas reglas se aplican solo si los caracteres se encuentran al comienzo de la cadena, mientras que 12 de ellas se aplican solo cuando los caracteres se encuentran en medio de la cadena, y 28 se aplican únicamente si están al final de la cadena. La cadena de nombre transformada se codifica en un código compuesto por la primera letra seguida de tres dígitos (sin ceros ni números duplicados). Esta opción se desarrolló como respuesta a las limitaciones de Soundex; es más compleja y, por lo tanto, más lenta que esta última.

Soundex Devuelve un código Soundex de campos seleccionados. El algoritmo Soundex produce un código de longitud fija basado en la pronunciación en inglés de las palabras.

Subcadena de caracteres Devuelve una porción específica del campo seleccionado.

7. En el campo **Campos de entrada**, seleccione el campo al que desee aplicar el algoritmo. Por ejemplo, si selecciona el algoritmo soundex y elige un campo denominado Ciudad, el código se generaría aplicando el algoritmo soundex para los datos en el campo Ciudad.
8. Si seleccionó el algoritmo de subcadena, especifique la parte del campo que desea utilizar en la subcadena:
 - a) En el campo **Posición de inicio**, especifique la posición en el campo en el que desea que comience la subcadena.
 - b) En el campo **Longitud**, seleccione el número de caracteres desde la posición de inicio que desea incluir en la subcadena.

Por ejemplo, supongamos que tenemos los siguientes datos en un campo llamado LastName:

Augustine

Si especifica 3 como la posición de inicio y 6 como la posición final, la subcadena sería:

gustin

9. Marque la casilla **Eliminar el ruido de caracteres** para eliminar todos los caracteres no numéricos ni alfanuméricos, como guiones, espacios en blanco y otros caracteres especiales del campo antes de aplicar el algoritmo.

10. Para algoritmos de subcadena y consonantes, puede ordenar los datos en el campo antes de aplicar el algoritmo de control de la casilla **Ordenar entrada**. Después, puede optar por ordenar los caracteres o los términos en el campo en orden alfabético.
11. Haga clic en **Aceptar** para guardar la configuración.
12. Repita los pasos las veces que sea necesario si desea agregar algoritmos para producir una ID más compleja.

Nota: La definición de clave única siempre se muestra en un color distinto y no se puede borrar.

Write to Cache

Write to Cache carga resultados de un flujo de salida en un caché global, de modo que los datos estén disponibles para búsqueda en la etapa de Caché de consultas. Con una caché global, las búsquedas de datos mejoran el rendimiento con respecto a las búsquedas en bases de datos.

Una caché global abarca todo el sistema, se comparte y reside en la memoria. Elija una caché global si desea que esté disponible para varios flujos de datos, si los datos no cambian con frecuencia o si permanecen relativamente estáticos y si el almacenamiento no es limitado. Una caché global es estática; es decir, solo puede escribir en ella una vez. La caché no se puede actualizar una vez que se ha creado.

Nota: Write to Cache sobrescribe la caché cada vez que el flujo de datos se ejecuta.

General

Nombre de la opción	Descripción
Nombre de caché	Especifica el nombre que desea dar a la caché. Si ya hay cachés en el sistema, se incluyen en una lista y puede elegir si desea rellenar la caché existente con los nuevos datos. Para crear una nueva caché, escriba el nombre que desea para la nueva caché. El nombre debe comenzar con una letra. Puede tener subrayado pero no otros caracteres especiales. El nombre puede tener valores numéricos.
Campos de caché	Esta columna crea una lista con los nombres de campo que se usan en la caché. Si desea cambiar un nombre de campo, haga clic en el nombre de campo e ingrese un nuevo nombre.

Nombre de la opción	Descripción
Campos de etapa	Esta columna lista los nombres de campo que se usan en el flujo de datos. No puede modificar estos nombres de campos.
Tipo	Esta columna muestra el tipo de dato de cada campo.
Incluir	Marque la casilla en esta columna para incluir el campo escrito en la caché. Desactive la casilla si no desea que el campo se escriba en la caché.
Campo de clave	<p>Active la casilla de esta columna si desea que el campo se utilice como clave en la etapa de caché de consultas. Por ejemplo, si tiene un campo de flujo de datos denominado AccountNumber y desea que el caché de consultas busque datos mediante la consulta de un valor coincidente en el campo AccountNumber, active la casilla de la columna de campo clave en el campo AccountNumber.</p> <p>Los campos que especifique como campos clave están disponibles para su selección en la caché de consultas como campos clave.</p>

Cómo eliminar una memoria caché global

Para eliminar una memoria caché global debe crear y ejecutar un flujo de proceso. El flujo de proceso debe incluir una actividad Eliminar memoria caché. La actividad Eliminar memoria caché elimina la memoria caché global, pero no la borra. También puede eliminar la memoria caché de manera automática, al programar un flujo de proceso.

Nota: Consulte la *Guía de Dataflow Designer* para obtener instrucciones acerca de cómo crear y programar flujos de procesos.



Para eliminar manualmente los datos de la memoria caché global, siga los pasos a continuación:

1. Arrastre la actividad **Eliminar memoria caché** al lienzo
2. Arrastre la actividad **Sin error** al lienzo.

3. Conecte las dos actividades.
4. Haga doble clic en la actividad **Eliminar memoria caché**.
5. Seleccione la memoria caché. También puede seleccionar varias memorias caché para eliminar sus datos.
Las memorias caché que crea en la etapa Write To Cache se muestran en la actividad Eliminar memoria caché.
6. Ejecute el flujo de proceso.

Write to DB

La etapa Write to DB escribe la salida de un flujo de datos en una base de datos.

Nota: Se pueden lograr mejoras significativas del rendimiento usando múltiples instancias de ejecución de Write to DB. Para especificar las múltiples instancias de ejecución, haga clic en el botón **Runtime** (Ejecución).

Ficha General

Nombre de la opción	Descripción
Conexión	<p>Seleccione la conexión para la base de datos que quiere usar en el campo Conexión. Para realizar una nueva conexión de base de datos, haga clic en Manage (Administrar). Para obtener más información sobre la creación de conexiones de bases de datos, consulte Administrador de conexión de base de datos.</p>
Tabla/Vista	<p>Después de seleccionar una conexión, especifique la tabla o la vista donde se escribirá. Haga clic en el botón de examinar ([...]) para navegar hasta la tabla o la vista que quiere utilizar o haga clic en Create Table (Crear tabla) para crear una nueva tabla en la base de datos.</p> <p>Nota: Si escribe en una base de datos SQL, no puede escribir en vistas que hagan referencia a más de una tabla. Ello se debe a una limitación en el servidor SQL.</p>

Nombre de la opción	Descripción
Create Table (Crear tabla)	<p>Crea una nueva tabla en la base de datos seleccionada. Elija el titular de la tabla en el campo Table owner (Titular de tabla) y especifique el nombre de la nueva tabla en el campo Table name (Nombre de tabla). Los nombres de las tablas distinguen mayúsculas de minúsculas. Especifique una clave principal al seleccionar una casilla de verificación en la columna Clave principal. Además, especifique los campos que quiere escribir en la nueva tabla seleccionando el casillero de la columna Include (Incluir). La columna Ancho especifica la longitud del campo para el tipo de datos de cadena de caracteres. De manera predeterminada es 512. Si se marca la columna Permitir nulos y los campos de entrada contienen un valor nulo, el flujo de datos escribirá el valor nulo en la base de datos. Cabe señalar que es posible editar el nombre de la columna cambiando el valor de la columna Campos de salida.</p> <p>El botón Create Table (Crear tabla) admite la creación de tablas en las siguientes bases de datos:</p> <ul style="list-style-type: none"> • Axion • DB2 • Derby/Cloudscape • Firebird • HSQLDB • Interbase • MaxDB/SapDB • McKoi • MySQL • Oracle • PostgreSQL • SQL Server • Sybase <p>Nota: En las bases de datos DB2, si intenta crear una tabla y el tamaño de página es más pequeño que el largo total de todas las columnas, aparecerá un error que indica que no es posible crear un elemento del contenido y que la clase serializable no está disponible.</p>
Campos de etapa	<p>En la columna Campos de etapa, puede especificar el campo que quiere escribir en el campo de base de datos que se muestra en la columna Campo de base de datos.</p>
Incluir	<p>La columna Incluir le permite seleccionar los campos en los que quiere escribir.</p> <p>Nota: Para evitar un rendimiento deficiente, debe contar con un índice ordenado o una clave ordenada en la tabla de base de datos.</p>

Nota: La etapa **Write to DB** permite escribir todos los valores del tipo de datos `date` como valores `String`. Este es el comportamiento del *controlador jTDS*, que es el controlador

predeterminado que usa Spectrum. Para manejar todos los valores de tipos de datos `date` tal como están, use el controlador JDBC de Microsoft.

Administrador de conexión de base de datos

El Administrador de conexión de base de datos le permite administrar las conexiones registradas de bases de datos. Para agregar, modificar, eliminar y probar las conexiones:

1. En el cuadro de diálogo **Write To DB Options** (Opciones de Write To DB), haga clic en **Manage** (Administrar).
2. Haga clic en **Add** (Agregar), **Modify** (Modificar) o **Delete** (Eliminar).
3. Si va a agregar o modificar una conexión de base de datos, complete los siguientes campos:
 - Nombre de conexión: ingrese el nombre de la conexión nueva..
 - Controlador de base de datos: seleccione el tipo de base de datos correspondiente.
 - Opciones de conexión: especifique todas las opciones, por ejemplo, host, puerto, instancia, nombre de usuario y contraseña.

Nota: Puede probar la conexión haciendo clic en **Test** (Probar).

4. Si va a borrar una conexión de base de datos, seleccione la conexión que quiere quitar y haga clic en Eliminar.

Ficha Tiempo de ejecución

Nombre de la opción	Descripción
Modo de escritura	<p>Especifica el tipo de acción que se realizará cuando se escriba en la base de datos, Una de las siguientes:</p> <p>Insertar Inserta registros nuevos en la base de datos pero no actualiza los registros existentes. Esta es la configuración predeterminada.</p> <p>Actualizar Actualiza los registros existentes en la base de datos pero no inserta nuevos registros.</p> <p>Nota: Si selecciona Update (Actualizar), el nombre de columna de clave principal que se utiliza en la tabla de entrada debe coincidir con el nombre de columna de clave principal de la tabla de salida. Si trata de actualizar una tabla cuya clave principal del nombre de la columna no coincide con la entrada o no está definida, la actualización no se realizará.</p> <p>Insertar si no es posible actualizar Inserta nuevos registros en la base de datos si el registro no existe, de lo contrario, actualiza el registro existente.</p>

Nombre de la opción	Descripción
Confirmación de lote	<p>Seleccione esta opción para confirmar cambios en la base de datos tras procesar un número específico de registros. Esta opción no está seleccionada de forma predeterminada, por lo tanto, los cambios se confirman luego de que se procesa cada registro. Si se selecciona esta opción, es posible mejorar significativamente el rendimiento de la etapa Write to DB.</p>
Tamaño de lote	<p>Si activa la opción Batch commit (Confirmación de lote), se especifica el número de registros a confirmar en la base de datos dentro de cada lote. El valor predeterminado es 1,000. Para los flujos de datos creados en Spectrum™ Technology Platform 7.0 y versiones anteriores, el valor predeterminado es 100.</p> <p>No siempre los lotes de gran tamaño ofrecen mejor rendimiento de carga. Tenga en cuenta los siguientes factores al elegir el tamaño del lote:</p> <ul style="list-style-type: none"> • Data arrival rate to Write To DB stage (Velocidad de llegada de los datos a la etapa Write To DB): Si los datos llegan a un ritmo más lento de lo que puede procesar la base de datos, entonces la modificación del tamaño del lote no mejorará el rendimiento general del flujo de datos. Por ejemplo, es posible que los flujos de datos con validación de direcciones o geocodificación no mejoren el rendimiento si se aumenta el tamaño del lote. • Network traffic (Tráfico de la red): En el caso de redes lentas, el aumento del lote a un tamaño mediano (de 1000 a 10.000) mejora el rendimiento. • Database load and/or processing speed (Velocidad de procesamiento o carga de la base de datos): En el caso de las bases de datos que tienen una gran capacidad de procesamiento, el aumento del tamaño del lote mejorará el rendimiento. • Multiple runtime instances (Utilización de múltiples instancias): Si utiliza múltiples instancias de ejecución de la etapa Write to DB, los lotes de gran tamaño consumirán mucha memoria, por lo tanto, es recomendable usar un tamaño pequeño o mediano (de 100 a 10.000). • Database roll backs (Reversión de la base de datos): Cada vez que falla una instrucción, todo el lote se revierte. Mientras más grande sea el tamaño del lote, más tiempo tardará en revertirse.
Confirmación al final	<p>Seleccione esta opción para garantizar que la operación de confirmación en la base de datos ocurra después de transferir todos los registros a la base de datos.</p>
Conteo por lotes para confirmar	<p>Especifica un valor después del cual los registros se deben confirmar. Los registros se confirman en la base de datos después de que se transfieren una cantidad de registros (conteo por lotes para confirmar * tamaño del lote) a la base de datos. Por ejemplo, si Tamaño del lote se configura como 1000 y Conteo por lotes para confirmar se establece como 3, la confirmación ocurre después de que se transfieren 3000 registros a la base de datos.</p>
Truncar tabla antes de insertar datos	<p>Seleccione esta opción si desea borrar todos los datos de la tabla antes de escribir la base de datos.</p>

Nombre de la opción	Descripción
Descartar la tabla y recrearla si ya existe	<p data-bbox="535 367 1433 493">Seleccione esta opción para eliminar y volver a crear la tabla antes de copiarle la salida del flujo de datos. Esta opción resulta útil si se desea que el esquema de tabla coincida con los campos del flujo de datos y no contenga información de esquema no relacionada.</p> <p data-bbox="535 504 1433 716">La tabla que se elimina y se crea nuevamente es la que se especifica en el campo Table/View (Tabla/Vista) de la ficha General. Por ejemplo, si se especifica la tabla Customers (Clientes) en el campo Table/View (Tabla/Vista) y se selecciona la opción Drop and recreate the table if it already exists (Descartar la tabla y recrearla si ya existe), la tabla Customers (Clientes) se eliminará de la base de datos y se creará una nueva tabla con el mismo nombre, con un esquema que coincide con los campos reales copiados en la tabla.</p>

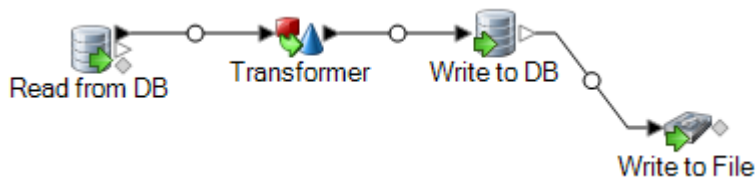
Configuración de Manejo de errores en Write to DB

La etapa Write to DB tiene ahora un puerto de error que permite filtrar aquellos registros que causan errores en la base de datos cuando se escribe el registro en una base de datos, tales como una infracción de restricción de la clave principal o una violación de restricción única. Estos registros pueden ser redirigidos hacia otra ruta en el flujo de datos, mientras se confirman otros registros. Por ejemplo, si está procesando 100 registros y los registros 4, 23, y 56 provocan un error en la base de datos, estos tres registros serán redirigidos por medio del puerto de error, en tanto los otros 97 registros serán confirmados en la base de datos.

Nota: El uso del puerto de error es opcional. Si no usas el puerto de error, el trabajo fallará si cualquier registro causa un error.

- Desde la paleta, seleccione la etapa del tipo en la cual usted desea manejar errores (por ejemplo, Write to File) y arrástrela hacia el lienzo. Hay un par de opciones para seleccionar una etapa:
 - Para escribir registros fallidos en un archivo, arrastre alguno de los siguientes hacia el lienzo: Write to File, Write to XML, o Write to Variable Format File.
 - Para sencillamente descartar los registros fallidos, arrastre Write to Null hacia el lienzo.
- Conecte el puerto de error de Write to DB con la etapa en la que desea manejar los registros fallidos.

El siguiente ejemplo muestra el puerto de error en Write to DB conectado con una etapa Write to File. En este ejemplo, los registros que causan un error cuando se escriben en la base de datos son en cambio escritos en el archivo especificado en la etapa Write to File.



Cuando ejecuta el flujo de datos, los registros que causan un error son redirigidos a través del puerto de error. Los registros del puerto de error contienen los campos especificados en Write to DB además de los siguientes campos:

- Código de error** Este campo contiene el código de error numérico devuelto desde la base de datos. Por ejemplo, dado el error `ORA-00001: unique constraint ANKUSH.SYS_C0010018) violated`, el valor en el campo `Error.code` sería 1. Consulte la documentación de la base de datos de su software para encontrar una lista de los códigos de error.
- Error.Message** Este campo contiene el mensaje de error devuelto desde la base de datos. Por ejemplo: `ORA-01034 ORACLE not available`. En este caso, `ORACLE not available` sería el valor en el campo `Error.Message`. Consulte la documentación de la base de datos de su software para encontrar una lista de los mensajes de error.
- Error.SQLState** Este campo contiene el código SQLSTATE que entrega información pormenorizada acerca de la causa del error. Para obtener una lista de códigos SQLSTATE, consulte la documentación de la base de datos de su software.
- Timestamp** La fecha y hora en el servidor Spectrum™ Technology Platform en el momento en que se produjo el error.
- UserName** El nombre del usuario de Spectrum™ Technology Platform que ejecutó el flujo de datos.

Write to File

Write to File escribe salidas de flujo de datos en un archivo plano. Todos los registros incluirán los mismos campos. Si desea escribir registros de formato variable, consulte [Write to Variable Format File](#) en la página 292. Si desea escribir registros en un archivo XML, consulte [Write to XML](#) en la página 302.

Consejo: Puede copiar el origen y pegarlo como receptor en su flujo de datos para configurar rápidamente el archivo, y usar los mismos campos que definió en el origen.

Pestaña Propiedades del archivo

Nombre de campo	Descripción
Server name (Nombre de servidor)	Indica si el archivo que seleccionó como entrada está ubicado en la computadora que está ejecutando Enterprise Designer o en el servidor de Spectrum™ Technology Platform. Si selecciona un archivo en la computadora local, el nombre del servidor será Mi computadora. Si selecciona un archivo que se encuentra en el servidor, el nombre del servidor será Spectrum™ Technology Platform.
Nombre de archivo	<p>Especifica la ruta al archivo. Haga clic en el botón de los puntos suspensivos (...) para buscar el archivo que desea.</p> <p>Cuando escribe un archivo en un servidor de archivos HDFS, los formatos de compresión compatibles son:</p> <ol style="list-style-type: none"> 1. GZIP (.gz) 2. BZIP2 (.bz2) <p>Nota: Incluya la extensión apropiada en el nombre de archivo para indicar el formato de compresión deseado que se debe utilizar mientras escribe en el archivo.</p> <p>Atención: Si el servidor de Spectrum™ Technology Platform está ejecutándose en Unix o Linux, recuerde que estas plataformas distinguen las mayúsculas de las minúsculas en los nombres de los archivos y las rutas.</p>
Tipo de registro	<p>Formato de los registros en el archivo. Seleccione uno de:</p> <p>En líneas secuenciales Archivo de texto en el que los registros se separan con un carácter de final de línea (EOL), como un retorno de carro o un salto de línea (CR o LF) y cada campo tiene una posición fija para los caracteres de inicio y final.</p> <p>De ancho fijo Archivo de texto en el que cada registro es un número específico de caracteres de longitud y cada campo tiene una posición fija para los caracteres de inicio y final.</p> <p>Delimitado Archivo de texto en el que los registros se separan con un carácter de final de línea (EOL), como un retorno de carro o un salto de línea (CR o LF) y cada campo se separa con un carácter designado; por ejemplo, una coma.</p>

Nombre de campo	Descripción
Character encoding (Codificación de caracteres)	Codificación del archivo de texto. Seleccione uno de estos:
UTF-8	Admite todos los caracteres Unicode y es compatible con las versiones anteriores de ASCII. Para obtener más información sobre UTF, consulte unicode.org/faq/utf_bom.html .
UTF-16	Admite todos los caracteres Unicode, pero no es compatible con las versiones anteriores de ASCII. Para obtener más información sobre UTF, consulte unicode.org/faq/utf_bom.html .
US-ASCII	Codificación de caracteres basada en el orden del alfabeto inglés.
UTF-16BE	Codificación UTF-16 con serialización de bytes big endian (byte más significativo al principio).
UTF-16LE	Codificación UTF-16 con serialización de bytes little endian (byte menos significativo al principio).
ISO-8859-1	Codificación de caracteres ASCII, que suele utilizarse para los idiomas de Europa occidental. También se la conoce como Latin-1.
ISO-8859-3	Codificación de caracteres ASCII, que suele utilizarse para los idiomas de Europa meridional. También se la conoce como Latin-3.
ISO-8859-9	Codificación de caracteres ASCII, que suele utilizarse para el idioma turco. También se la conoce como Latin-5.
CP850	Página de códigos ASCII que se utiliza para escribir idiomas de Europa occidental.
CP500	Página de códigos EBCDIC que se utiliza para escribir idiomas de Europa occidental.
Shift_JIS	Codificación de caracteres para el idioma japonés.
MS932	Una extensión de Microsoft de Shift_JIS para incluir caracteres especiales NEC, selección NEC de extensiones IBM y extensiones IBM.
CP1047	Una página de códigos EBCDIC con el conjunto de caracteres completo del alfabeto latino 1.

Nombre de campo	Descripción
Separador de campo	<p data-bbox="535 367 1433 430">Especifica el carácter que se utiliza para separar los campos en un archivo delimitado.</p> <p data-bbox="535 441 1433 483">Por ejemplo, este registro utiliza una barra vertical () como separador de campo:</p> <pre data-bbox="552 493 1421 556">7200 13TH ST MIAMI FL 33144</pre> <p data-bbox="535 567 1433 609">Estos caracteres disponibles para definir como separadores de campo son:</p> <ul data-bbox="535 619 779 829" style="list-style-type: none">• Espacio de claves• Tab• Coma• Punto• Punto y coma• Barra vertical <p data-bbox="535 840 1433 903">Si el archivo utiliza un carácter distinto como separador de campo, haga clic en el botón de los puntos suspensivos para seleccionar otro carácter como delimitador.</p>
Text qualifier (Calificador de texto)	<p data-bbox="535 976 1433 1018">El carácter que se utiliza para demarcar los valores de texto en un archivo delimitado.</p> <p data-bbox="535 1029 1433 1071">Por ejemplo, este registro utiliza comillas dobles (") como calificador de texto.</p> <pre data-bbox="552 1081 1421 1144">"7200 13TH ST" "MIAMI" "FL" "33144"</pre> <p data-bbox="535 1155 1433 1197">Los caracteres disponibles para definir como calificadores de texto son:</p> <ul data-bbox="535 1207 779 1270" style="list-style-type: none">• Comilla simple (')• Comilla doble (") <p data-bbox="535 1281 1433 1375">Si el archivo utiliza un calificador de texto distinto, haga clic en el botón de los puntos suspensivos para seleccionar otro carácter como calificador de texto.</p>

Nombre de campo	Descripción
Record separator (Separador de registro)	<p>Especifica el carácter que se utiliza para separar registros en archivos secuenciales o delimitados. Este campo no está disponible si selecciona la casilla Utilizar final de línea predeterminado .</p> <p>Los parámetros de separador de registro disponibles son:</p> <p>Unix (U+000A) Un carácter de salto de línea separa los registros. Este es el separador de registros estándar para los sistemas Unix.</p> <p>Macintosh (U+000D) Un carácter de retorno de carro separa los registros. Este es el separador de registros estándar para los sistemas Macintosh.</p> <p>Windows (U+000D U+000A) Un carácter de retorno de carro seguido por un salto de línea separa los registros. Este el separador de registros estándar para los sistemas Windows.</p> <p>Si su archivo utiliza un separador de registro distinto, haga clic en el botón de los puntos suspensivos para seleccionar otro carácter como separador de registro.</p>
Use default EOL (Utilizar final de línea predeterminado)	<p>Especifica que el separador de registro del archivo es el carácter de final de línea (EOL) predeterminado que utiliza el sistema operativo en el que se ejecuta el servidor Spectrum™ Technology Platform.</p> <p>No seleccione esta opción si el archivo utiliza un carácter de final de línea distinto del que utiliza de forma predeterminada el sistema operativo del servidor. Por ejemplo, si el archivo utiliza un final de línea de Windows pero el servidor está ejecutándose en Linux, no seleccione esta opción. En su lugar, seleccione la opción Windows en el campo Separador de registro.</p>
Record length (Longitud de registro)	<p>Para los archivos de ancho fijo, especifica la cantidad exacta de caracteres en cada registro.</p> <p>Para los archivos secuenciales, especifica la longitud, en caracteres, del registro más largo del archivo.</p>
First row is header record (La primera fila es el registro de encabezado)	<p>Especifica si el primer registro de un archivo delimitado contiene información del encabezado pero no datos.</p> <p>Por ejemplo, este fragmento de archivo muestra una fila de encabezado en el primer registro.</p> <pre>"AddressLine1" "City" "StateProvince" "PostalCode" "7200 13TH ST" "MIAMI" "FL" "33144" "One Global View" "Troy" "NY" 12180</pre>

Nombre de campo	Descripción
Considerar malformados los registros con menos campos que los especificados	Se considerarán malformados los registros de archivos delimitados que contengan menos campos que los especificados en la ficha Campos .
Importar	Importa la definición de diseño de un archivo, los ajustes de codificación y las opciones de ordenamiento desde un archivo de configuración. El archivo de configuración se crea al exportar la configuración de otra etapa Read from File o Write to File que utilizó el mismo archivo de entrada o un archivo con el mismo diseño que aquel en el que usted está trabajando.
Exportar	<p>Guarda la definición de diseño de un archivo, los ajustes de codificación y las opciones de ordenamiento en un archivo de configuración. Luego, puede importar estas configuraciones a otra etapa Read from File o Write from File que utilice el mismo archivo de entrada o un archivo que tenga las mismas características que aquel en que usted esté trabajando actualmente. También puede usar el archivo de configuración con el ejecutor de trabajos para especificar la configuración del archivo durante la ejecución.</p> <p>Para obtener información acerca de los comandos disponibles, consulte El archivo de configuración de definición de archivos en la página 165.</p>

Pestaña Campos

En la pestaña Campos se definen los nombres, las posiciones y, para los archivos secuenciales y de ancho fijo, la longitud de los campos en el archivo. Para obtener más información, consulte los temas siguientes:

[Definición de los campos en un archivo de salida delimitado](#) en la página 261

[Definición de los campos en un archivo secuencial o de ancho fijo](#) en la página 264

Pestaña Ordenar campos

En la pestaña Ordenar campos se definen los campos según los cuales se ordenan los registros de salida antes de escribirlos en el archivo de salida. La ordenación es opcional. Para obtener más información, consulte [Ordenación de registros de salida](#) en la página 267.

Pestaña Tiempo de ejecución

Nombre de la opción	Descripción
Nombre de archivo	Muestra el archivo definido en la pestaña Propiedades del archivo .

Nombre de la opción	Descripción
Generar varios archivos	<p>Seleccione esta opción para escribir registros en diferentes archivos en lugar de escribirlos todos en un único archivo. El archivo en el cual se escribe cada registro está especificado en el propio registro. Cada registro debe contener un campo que especifique, ya sea un nombre de archivo o la ruta completa del archivo en el cual desea que se escriba el registro. Por ejemplo, si desea enviar el precio de las acciones de diferentes empresas (de diversos grupos) a todos los clientes por separado, esta función escribe el precio de las acciones de las diferentes empresas en archivos independientes, que se pueden enviar a cada uno de los clientes, si así lo desea. Si se activa la opción Generar múltiples archivos, debe especificar un archivo de salida en el servidor Spectrum™ Technology Platform o en un servidor FTP. Si desea escribir datos desde archivos en un servidor FTP, debe definir una conexión con el servidor de archivos mediante Management Console.</p> <p>Nota: Los registros en la columna que seleccione en Campo de ruta de archivo deben estar ordenados. Utilice esta función cuando el registro contenga ya sea un nombre de archivo o la ruta de acceso del archivo completa.</p>
Campo de ruta de archivo	<p>Selecciona el campo que contiene la ruta (ya sea un nombre de archivo o la ruta completa del archivo) del archivo en el cual desea que se escriba el registro. Este campo está habilitado solamente si selecciona Generar múltiples archivos.</p>
Modo de escritura	<p>Especifica si se agrega la salida del flujo de datos al final del archivo o si se eliminan los datos ya existentes en el archivo antes de escribir la salida. Una de las siguientes:</p> <p>Sobrescribir Reemplaza los datos ya existentes en el archivo de salida cada vez que se ejecuta el flujo de datos.</p> <p>Adjuntar Agrega la salida del flujo de datos al final del archivo sin eliminar los datos ya existentes del archivo.</p>

Definición de los campos en un archivo de salida delimitado

En la etapa Write To File, la pestaña **Campos** define los nombres, las posiciones y, para algunos tipos de archivos, las longitudes de los campos del archivo. Una vez que haya definido un archivo de salida en la ficha **Propiedades del archivo**, puede definir los campos.

Si el archivo de salida contiene un registro de encabezado, usted puede definir los campos rápidamente haciendo clic en **Regenerar**.

Para definir campos con valores predeterminados de posición, longitud y tipo de datos, haga clic en **Agregado rápido** y seleccione los campos que desea agregar.

Si el archivo de entrada no contiene un registro de encabezado o desea definir los campos manualmente, siga estos pasos:

1. Haga clic en **Agregar**.
2. En el campo **Nombre**, elija el campo que desea agregar.
3. En el campo **Tipo**, seleccione el tipo de datos del campo proveniente del flujo de datos.

Spectrum™ Technology Platform admite los siguientes tipos de datos:

bigdecimal Un tipo de dato numérico que admite 38 puntos decimales de precisión. Utilice este tipo de datos para datos que se usarán en cálculos matemáticos que requieren un alto grado de precisión, especialmente aquellos que implican datos financieros. El tipo de datos bigdecimal es compatible con cálculos más precisos que el tipo de datos double.

boolean Tipo de dato lógico con dos valores: verdadero y falso

bytearray Una matriz (lista) de bytes.

Nota: Bytearray no se admite como datos de entrada para un servicio REST.

fecha Un tipo de dato que contiene un mes, día y año. Por ejemplo, 2012-01-30 o 30 de enero, 2012. Puede especificar un formato de fecha predeterminado en Management Console.

fecha y hora Un tipo de dato que contiene un mes, día, año y horas, minutos y segundos. Por ejemplo, 2012/01/30 6:15 PM.

double Un tipo de dato numérico que contiene números de doble precisión tanto negativos como positivos entre 2^{-1074} y $(2-2^{-52}) \times 2^{1023}$. En notación científica, el rango de valores es - 1.79769313486232E+308 a 1.79769313486232E+308.

float Un tipo de dato numérico que contiene números de precisión simple tanto negativos como positivos entre 2^{-149} y $(2-2^{-23}) \times 2^{127}$. En notación científica, el rango de valores es - 3.402823E+38 a 3.402823E+38.

integer Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{31} (-2,147,483,648) y $2^{31}-1$ (2,147,483,647).

list En sentido estricto, una lista no es un tipo de dato. Sin embargo, cuando un campo contiene datos jerárquicos, es tratado como un campo "lista". En Spectrum™ Technology Platform una lista es una colección de datos que consiste en múltiples valores. Por ejemplo, un campo Nombres puede contener una lista de valores de nombre, lo que puede representarse mediante una estructura XML de la siguiente forma:

```
<Names>
  <Name>John Smith</Name>
  <Name>Ann Fowler</Name>
</Names>
```

Cabe destacar que el tipo de dato lista de Spectrum™ Technology Platform es diferente del tipo de dato lista del esquema XML puesto que este último es un tipo de dato simple que está compuesto de valores múltiples, mientras que el tipo de dato lista de Spectrum™ Technology Platform es similar a un tipo de dato complejo del esquema XML.

long	Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{63} (-9,223,372,036,854,775,808) y $2^{63}-1$ (9,223,372,036,854,775,807).
string	Secuencia de caracteres.
hora	Un tipo de dato que contiene la hora del día. Por ejemplo, 21:15:59 o 9:15:59 PM.

4. Si ha seleccionado un tipo de datos de fecha, hora, o numérico, puede usar el formato de fecha/hora predeterminado o puede indicar un formato diferente para este campo específico. El formato predeterminado es, o bien el formato de sistema predeterminado que se haya definido en las opciones de conversión de tipos en Management Console, o bien es el formato predeterminado del flujo de datos especificado en la opciones de conversión de tipos de Enterprise Designer. Se mostrará el formato que se esté usando. Para usar el formato predeterminado, deje seleccionado **Predeterminado**. Para indicar un formato diferente, seleccione **Personalizado** y siga los siguientes pasos:

- a) En el campo **Configuración regional** seleccione el país cuya convención acerca de formatos usted desea utilizar. Su selección determinará los valores predeterminados en el campo **Formato**. En cuanto a los datos referidos a fechas, su selección determinará además el idioma que se usará al enunciar un mes. Por ejemplo, si ha indicado Inglés, el primer mes del año será "January" pero si ha especificado Francés, este será "Janvier."
- b) En el campo **Formato**, seleccione el formato de los datos. El formato depende del tipo de datos en el campo. Se proporciona una lista de los formatos de uso más común en la configuración regional seleccionada.

A la derecha del campo **Formato** se muestra un ejemplo del formato seleccionado.

Además puede especificar sus propios formatos de fecha, hora, y número en caso de que los que están disponibles para seleccionar no se ajusten a sus necesidades. Para especificar su propio formato de fecha u hora, ingrese el formato en el campo usando la notación que se describe en [Patrones de fecha y hora](#) en la página 312. Para especificar su propio formato de número, ingrese el formato en el archivo usando la notación que se describe en [Patrones de números](#) en la página 315.

5. Haga clic en **Agregar**.

Después de definir los campos de su archivo de salida, puede editar su contenido y diseño.

Nombre de la opción	Descripción
Agregar	Agrega un campo a la salida. Puede anexas un campo al final del diseño existente o insertar un campo en una posición existente y la posición de los campos restantes se ajustará según corresponda.
Modificar	Modifica el nombre y el tipo del campo.
Quitar	Elimina el campo seleccionado de la salida.
Subir/Bajar	Reordena el campo seleccionado.

Definición de los campos en un archivo secuencial o de ancho fijo

En la etapa Write To File, la ficha **Campos** define los nombres, las posiciones y, para algunos tipos de archivos, las longitudes de los campos del archivo. Una vez que haya definido un archivo de salida en la ficha **Propiedades del archivo**, puede definir los campos.

Para definir campos con valores predeterminados de posición, longitud y tipo de datos, haga clic en **Agregado rápido** y seleccione los campos que desea agregar.

Para agregar campos manualmente desde una lista de campos utilizados en el flujo de datos, siga el procedimiento que se indica a continuación.

1. Haga clic en **Agregar**.
2. En el campo **Nombre**, elija el campo que desea agregar.
3. En el campo **Tipo**, seleccione el tipo de datos del campo proveniente del flujo de datos.

Spectrum™ Technology Platform admite los siguientes tipos de datos:

bigdecimal	Un tipo de dato numérico que admite 38 puntos decimales de precisión. Utilice este tipo de datos para datos que se usarán en cálculos matemáticos que requieren un alto grado de precisión, especialmente aquellos que implican datos financieros. El tipo de datos bigdecimal es compatible con cálculos más precisos que el tipo de datos double.
boolean	Tipo de dato lógico con dos valores: verdadero y falso
bytearray	Una matriz (lista) de bytes.

Nota: Bytearray no se admite como datos de entrada para un servicio REST.

fecha	Un tipo de dato que contiene un mes, día y año. Por ejemplo, 2012-01-30 o 30 de enero, 2012. Puede especificar un formato de fecha predeterminado en Management Console.
fecha y hora	Un tipo de dato que contiene un mes, día, año y horas, minutos y segundos. Por ejemplo, 2012/01/30 6:15 PM.
double	Un tipo de dato numérico que contiene números de doble precisión tanto negativos como positivos entre 2^{-1074} y $(2-2^{-52}) \times 2^{1023}$. En notación científica, el rango de valores es - 1.79769313486232E+308 a 1.79769313486232E+308.
float	Un tipo de dato numérico que contiene números de precisión simple tanto negativos como positivos entre 2^{-149} y $(2-2^{-23}) \times 2^{127}$. En notación científica, el rango de valores es - 3.402823E+38 a 3.402823E+38.
integer	Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{31} (-2,147,483,648) y $2^{31}-1$ (2,147,483,647).
list	En sentido estricto, una lista no es un tipo de dato. Sin embargo, cuando un campo contiene datos jerárquicos, es tratado como un campo "lista". En Spectrum™ Technology Platform una lista es una colección de datos que consiste en múltiples valores. Por ejemplo, un campo Nombres puede contener una lista de valores de nombre, lo que puede representarse mediante una estructura XML de la siguiente forma: <div data-bbox="422 1029 1424 1186" data-label="Code-Block"> <pre><Names> <Name>John Smith</Name> <Name>Ann Fowler</Name> </Names></pre> </div> <p>Cabe destacar que el tipo de dato lista de Spectrum™ Technology Platform es diferente del tipo de dato lista del esquema XML puesto que este último es un tipo de dato simple que está compuesto de valores múltiples, mientras que el tipo de dato lista de Spectrum™ Technology Platform es similar a un tipo de dato complejo del esquema XML.</p>
long	Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{63} (-9,223,372,036,854,775,808) y $2^{63}-1$ (9,223,372,036,854,775,807).
string	Secuencia de caracteres.
hora	Un tipo de dato que contiene la hora del día. Por ejemplo, 21:15:59 o 9:15:59 PM.

- Si ha seleccionado un tipo de datos de fecha, hora, o numérico, puede usar el formato de fecha/hora predeterminado o puede indicar un formato diferente para este campo específico. El formato predeterminado es, o bien el formato de sistema predeterminado que se haya definido en las opciones de conversión de tipos en Management Console, o bien es el formato predeterminado del flujo de datos especificado en la opciones de conversión de tipos de Enterprise Designer. Se mostrará el formato que se esté usando. Para usar el formato

predeterminado, deje seleccionado **Predeterminado**. Para indicar un formato diferente, seleccione **Personalizado** y siga los siguientes pasos:

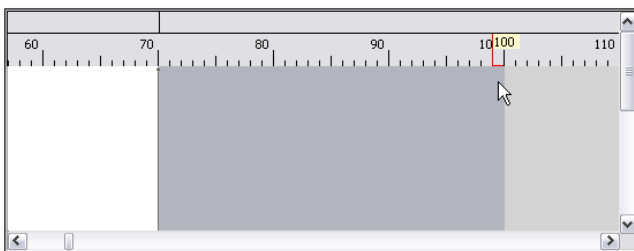
- a) En el campo **Configuración regional** seleccione el país cuya convención acerca de formatos usted desea utilizar. Su selección determinará los valores predeterminados en el campo **Formato**. En cuanto a los datos referidos a fechas, su selección determinará además el idioma que se usará al enunciar un mes. Por ejemplo, si ha indicado Inglés, el primer mes del año será "January" pero si ha especificado Francés, este será "Janvier."
- b) En el campo **Formato**, seleccione el formato de los datos. El formato depende del tipo de datos en el campo. Se proporciona una lista de los formatos de uso más común en la configuración regional seleccionada.

A la derecha del campo **Formato** se muestra un ejemplo del formato seleccionado.

Además puede especificar sus propios formatos de fecha, hora, y número en caso de que los que están disponibles para seleccionar no se ajusten a sus necesidades. Para especificar su propio formato de fecha u hora, ingrese el formato en el campo usando la notación que se describe en [Patrones de fecha y hora](#) en la página 312. Para especificar su propio formato de número, ingrese el formato en el archivo usando la notación que se describe en [Patrones de números](#) en la página 315.

5. Los campos **Posición de inicio** y **Longitud** se completan automáticamente en función de los datos del flujo de datos y la cantidad de campos que ya se han agregado.
6. Haga clic en **Agregar**.

Como alternativa, también puede agregar un campo definiendo primero la posición inicial y la longitud del campo. Para hacer esto, en **Archivo de muestra**, haga clic en la posición donde desea comenzar un campo y arrastre hacia la izquierda para resaltar el campo deseado, como se muestra a continuación:



Después de definir los campos de su archivo de salida, puede editar su contenido y diseño. La opción **Volver a calcular posición de inicio** le pide a la etapa Write to File que vuelva a calcular la posición de los campos cuando usted modifique, mueva, o elimine un campo del archivo de salida. Desmarque la casilla si no desea que se vuelvan a calcular las posiciones y desea que los campos queden en la posición en que estaban cuando termine de editar el archivo de salida.

Nombre de la opción	Descripción
Agregar	Agrega un campo a la salida.
Modificar	Modifica el nombre, el tipo, la posición de inicio y la longitud del campo.
Quitar	Elimina el campo seleccionado de la salida.
Subir/Bajar	Reordena el campo seleccionado.

Ordenación de registros de salida

En la etapa Write to File, la ficha **Ordenar campos** define los campos según los cuales se ordenan los registros de salida antes de escribirlos en el archivo de salida.. La ordenación es opcional.

1. En Write to File, haga clic en la ficha **Ordenar campos**.
2. Haga clic en **Agregar**.
3. Haga clic en la flecha desplegable de la columna **Nombre del campo** y seleccione el campo según el cual quiere ordenar. Los campos que se pueden seleccionar dependen de los campos del flujo de datos.
4. En la columna **Criterio de ordenación**, seleccione Ascending o Descending.
5. Repita hasta terminar de agregar todos los campos de salida que desea utilizar para la ordenación. Cambie el criterio de ordenación destacando la fila del campo que desea desplazar y haciendo clic en **Arriba** o **Abajo**.
6. Las opciones predeterminadas de rendimiento de orden para su sistema están definidas en Management Console. Si desea anular las opciones de rendimiento de orden predeterminadas de su sistema, haga clic en **Opciones avanzadas**. El cuadro de diálogo **Opciones avanzadas** contiene las siguientes opciones de rendimiento de orden:

Límite de registros en memoria	Especifica la cantidad máxima de filas de datos que un ordenador guardará en la memoria antes de empezar la paginación al disco. De forma predeterminada, la ordenación de 10.000 registros o menos se realizará en la memoria y la de más de 10.000 registros se realizará como una ordenación en disco. El límite máximo es de 100.000 registros. Normalmente, la ordenación que se realiza en la memoria es más rápida que la que se realiza en disco; por lo tanto, este valor se debe establecer lo suficientemente alto para que la mayor parte de la ordenación sea en la memoria y solo los conjuntos de datos grandes se escriban en el disco.
---------------------------------------	---

Nota: Tenga cuidado en entornos en los que se ejecuten trabajos de forma concurrente, porque aumentar la configuración del **Límite de registros en memoria** aumenta las posibilidades de agotar la memoria.

Número máximo de archivos temporales Especifica el número máximo de archivos temporales que puede utilizar el proceso de orden. El uso de una gran cantidad de archivos temporales puede producir un mejor rendimiento. Sin embargo, la cantidad óptima depende en gran medida de la configuración en el servidor que ejecuta Spectrum™ Technology Platform. Debe experimentar distintas configuraciones y observar el efecto en el rendimiento de usar una mayor o menor cantidad de archivos temporales. Para calcular la cantidad aproximada de archivos temporales que se pueda necesitar, use la siguiente ecuación:

$$\frac{(NumberOfRecords \times 2)}{InMemoryRecordLimit} = NumberOfTempFiles$$

Tenga en cuenta que la cantidad máxima de archivos temporales no puede ser superior a 1.000.

Activar compresión Especifica que los archivos temporales se comprimen cuando se los escribe en el disco.

Nota: La configuración del rendimiento óptimo de orden depende de la configuración del hardware de su servidor. No obstante, la siguiente ecuación generalmente produce un buen rendimiento de orden:

$$\frac{(InMemoryRecordLimit \times MaxNumberOfTempFiles \div 2)}{TotalNumberOfRecords} \geq$$

El archivo de configuración de definición de archivos

Un archivo de configuración de definición de archivos contiene el diseño de archivo, la codificación y las opciones de orden que se han exportado de una etapa Read from File o Write to File. Este archivo puede importarse a Read from File o Write to File para configurar rápidamente las opciones de la etapa en lugar de tener que especificar las opciones manualmente.

La forma más fácil de crear un archivo de configuración de definición de archivos es usar Read from File o Write to File para especificar la configuración del archivo, y luego hacer clic en el botón **Exportar** para generar el archivo.

No obstante, a título informativo, a continuación se muestra el esquema del archivo de configuración de definición de archivos. Cada elemento de un archivo XML tiene un tipo de dato, y, si ese tipo no es string o integer, se muestran los valores aceptables. Estos valores se corresponden directamente

con las opciones del cuadro de diálogo de la etapa. Por ejemplo, el elemento FileTypeEnum corresponde al campo Tipo de registro de la ficha Propiedades del archivo, y los siguientes tres valores se muestran en el esquema: linesequential, fixedwidth, y delimited.

Nota: Si usted ingresa "custom" en los campos LineSeparator, FieldSeparator o TextQualifier, también debe incluir un elemento custom correspondiente (por ejemplo, "CustomLineSeparator", "CustomFieldSeparator" o "CustomTextQualifier") con un número hexadecimal que represente el carácter, o secuencia de caracteres, que se utilizará.

```
<?xml version="1.0" encoding="utf-8"?>
<xs:schema elementFormDefault="qualified"
  xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="FileSchema" nillable="true" type="FileSchema"/>
  <xs:complexType name="FileSchema">
    <xs:sequence>
      <xs:element
        minOccurs="0"
        maxOccurs="1"
        default="linesequential"
        name="Type"
        type="FileTypeEnum"/>
      <xs:element
        minOccurs="0"
        maxOccurs="1"
        default="UTF-8" name="Encoding" type="xs:string"/>
      <xs:element
        minOccurs="0"
        maxOccurs="1"
        name="RecordLength"
        type="xs:int"/>
      <xs:element
        minOccurs="0"
        maxOccurs="1"
        default="default"
        name="LineSeparator"
        type="LineSeparatorEnum"/>
      <xs:element
        minOccurs="0"
        maxOccurs="1"
        name="CustomLineSeparator"
        type="xs:string"/>
      <xs:element
        minOccurs="0"
        maxOccurs="1"
        default="comma"
        name="FieldSeparator"
        type="FieldSeparatorEnum"/>
      <xs:element
        minOccurs="0"
        maxOccurs="1"
        name="CustomFieldSeparator"
        type="xs:string"/>
    </xs:sequence>
  </xs:complexType>
</xs:schema>
```

```

<xs:element
  minOccurs="0"
  maxOccurs="1"
  default="none"
  name="TextQualifier"
  type="TextQualifierEnum"/>
<xs:element
  minOccurs="0"
  maxOccurs="1"
  name="CustomTextQualifier"
  type="xs:string"/>
<xs:element
  minOccurs="0"
  maxOccurs="1"
  default="false"
  name="HasHeader"
  type="xs:boolean"/>
<xs:element
  minOccurs="0"
  maxOccurs="1"
  default="true"
  name="EnforceColumnCount"
  type="xs:boolean"/>
<xs:element
  minOccurs="0"
  maxOccurs="1"
  name="Fields"
  type="ArrayOfFieldSchema"/>
</xs:sequence>
</xs:complexType>
<xs:simpleType name="FileTypeEnum">
  <xs:restriction base="xs:string">
    <xs:enumeration value="linesequential"/>
    <xs:enumeration value="fixedwidth"/>
    <xs:enumeration value="delimited"/>
  </xs:restriction>
</xs:simpleType>
<xs:simpleType name="LineSeparatorEnum">
  <xs:restriction base="xs:string">
    <xs:enumeration value="default"/>
    <xs:enumeration value="windows"/>
    <xs:enumeration value="unix"/>
    <xs:enumeration value="mac"/>
    <xs:enumeration value="custom"/>
  </xs:restriction>
</xs:simpleType>
<xs:simpleType name="FieldSeparatorEnum">
  <xs:restriction base="xs:string">
    <xs:enumeration value="comma"/>
    <xs:enumeration value="tab"/>
    <xs:enumeration value="space"/>
    <xs:enumeration value="semicolon"/>
    <xs:enumeration value="period"/>
  </xs:restriction>
</xs:simpleType>

```

```

    <xs:enumeration value="pipe"/>
    <xs:enumeration value="custom"/>
  </xs:restriction>
</xs:simpleType>
<xs:simpleType name="TextQualifierEnum">
  <xs:restriction base="xs:string">
    <xs:enumeration value="none"/>
    <xs:enumeration value="single"/>
    <xs:enumeration value="double"/>
    <xs:enumeration value="custom"/>
  </xs:restriction>
</xs:simpleType>
<xs:complexType name="ArrayOfFieldSchema">
  <xs:sequence>
    <xs:element
      minOccurs="0"
      maxOccurs="unbounded"
      name="Field"
      nillable="true"
      type="FieldSchema"/>
  </xs:sequence>
</xs:complexType>
<xs:complexType name="FieldSchema">
  <xs:sequence>
    <xs:element
      minOccurs="0"
      maxOccurs="1"
      name="Name"
      type="xs:string"/>
    <xs:element
      minOccurs="0"
      maxOccurs="1"
      default="string"
      name="Type"
      type="xs:string"/>
    <xs:element
      minOccurs="1"
      maxOccurs="1"
      name="Position"
      type="xs:int"/>
    <xs:element
      minOccurs="0"
      maxOccurs="1"
      name="Length"
      type="xs:int"/>
    <xs:element
      minOccurs="0"
      maxOccurs="1"
      default="false"
      name="Trim"
      type="xs:boolean"/>
    <xs:element
      minOccurs="0"

```

```

        maxOccurs="1"
        name="Locale"
        type="Locale"/>
<xs:element
  minOccurs="0"
  maxOccurs="1"
  name="Pattern"
  type="xs:string"/>
<xs:element
  minOccurs="0"
  maxOccurs="1"
  default="none"
  name="Order"
  type="SortOrderEnum"/>
</xs:sequence>
</xs:complexType>
<xs:complexType name="Locale">
  <xs:sequence>
    <xs:element
      minOccurs="0"
      maxOccurs="1"
      name="Country"
      type="xs:string"/>
    <xs:element
      minOccurs="0"
      maxOccurs="1"
      name="Language"
      type="xs:string"/>
    <xs:element
      minOccurs="0"
      maxOccurs="1"
      name="Variant"
      type="xs:string"/>
  </xs:sequence>
</xs:complexType>
<xs:simpleType name="SortOrderEnum">
  <xs:restriction base="xs:string">
    <xs:enumeration value="none"/>
    <xs:enumeration value="ascending"/>
    <xs:enumeration value="descending"/>
  </xs:restriction>
</xs:simpleType>
</xs:schema>

```

Configuración de opciones de flujo de datos

Este procedimiento describe cómo configurar un flujo de datos compatible con las opciones de tiempo de ejecución para la etapa Write To File.

1. Abra el flujo de datos en Enterprise Designer.

2. Si desea configurar las opciones de tiempo de ejecución para una etapa en un flujo de datos incrustado, abra este último.
3. Haga clic en el ícono Opciones de flujo de datos en la barra de herramientas o haga clic en **Editar > Opciones de flujo de datos**. Aparecerá el cuadro de diálogo **Opciones de flujo de datos**.
4. Haga clic en **Agregar**. Aparecerá el cuadro de diálogo **Definir opción de flujo de datos**.
5. Expanda la etapa **Write to File**.

Las opciones de flujo de datos que se exponen son:

1. Codificación de caracteres
 2. Separador de campo
 3. Calificador de texto
 4. Longitud de registro
 5. La primera fila es el registro de encabezado
6. El nombre de la opción Write to File seleccionada se muestra en los campos **Nombre de opción** y **Etiqueta de opción**. Este es el nombre de opción que deberá especificarse durante el tiempo de ejecución para definir esta opción.
 7. Ingrese una descripción de la opción en el campo **Descripción**.
 8. En el campo **Destino**, seleccione la opción **Etapas seleccionadas**.
 9. Si desea limitar los valores que pueden especificarse durante la ejecución, edite las opciones en el campo **Valores válidos** haciendo clic en el icono a la derecha del campo.
 10. Si desea cambiar el valor predeterminado, especifique un valor diferente en el campo **Valor predeterminado**.

Nota: En el caso de un servicio, solo puede modificar los valores predeterminados antes de exponer el servicio por primera vez. Después de exponer el servicio, ya no es posible modificar los valores predeterminados con Enterprise Designer. En cambio, debe usar Management Console. Para obtener más información, consulte [Especificación de las opciones de servicio predeterminadas](#).

11. Haga clic en **Aceptar**.
12. Continúe agregando opciones, según sus preferencias.
13. Haga clic en **ACEPTAR** en el cuadro de diálogo **Opciones de flujo de datos** cuando haya terminado de agregar opciones.
14. Guarde y exponga el flujo de datos.

Reglas de las opciones de flujo de datos

1. *Codificación de carácter:* se aceptan todos los tipos de codificación que son válidos para la JVM subyacente. Esta opción no puede estar vacía.
2. *Separador de campo:* se acepta cualquier delimitador de carácter individual. Actualmente, no se admiten los valores ni los espacios HEX.
3. *Calificador de texto:* se acepta cualquier calificador de carácter individual. No se admiten los valores HEX.

4. *Longitud de registro*: solo se aceptan números enteros. No puede haber espacios en blanco ni caracteres no numéricos.
5. *Registro de inicio*: solo se aceptan números enteros. No pueden ser no numéricos.
6. *Máximo de registros*: solo se aceptan números enteros. No pueden ser no numéricos.
7. *La primera fila es el encabezado*: solo se aceptan los valores booleanos de `verdadero` y `falso`. No puede quedar vacío.

Write to Hadoop Sequence File

La etapa Write to Hadoop Sequence File escribe datos en un archivo de secuencia como una salida desde un flujo de datos. Un archivo de secuencia es un archivo simple que consta de pares de clave/valor binarios. Para obtener más información, visite

<http://wiki.apache.org/hadoop/SequenceFile>.

Nota: La etapa Write to Hadoop Sequence File solo admite archivos delimitados sin comprimir, ubicados en el Sistema de archivos distribuido Hadoop (Hadoop Distributed File System, HDFS).

Ficha Propiedades del archivo

Campos	Descripción
Servidor	Indica que el archivo que seleccionó en el campo Nombre de archivo está ubicado en el sistema Hadoop. Debe crear una conexión al servidor de archivos Hadoop en Management Console, antes de usarlo en esta etapa. Si selecciona un archivo en el sistema Hadoop, el nombre del servidor será el nombre que especifique en la Management Console mientras se crea un servidor de archivos.
Nombre de archivo	Especifica la ruta al archivo. Haga clic en el botón de los puntos suspensivos (...) para buscar el archivo que desea.

Campos	Descripción
Separador de campo	<p>Especifica el carácter que se utiliza para separar los campos en un archivo delimitado. Por ejemplo, este registro utiliza una barra vertical () como separador de campo:</p> <pre>7200 13TH ST MIAMI FL 33144</pre> <p>Estos caracteres disponibles para definir como separadores de campo son:</p> <ul style="list-style-type: none"> • Espacio de claves • Tab • Coma • Punto • Punto y coma • Barra vertical <p>Si el archivo utiliza un carácter distinto como separador de campo, haga clic en el botón de los puntos suspensivos para seleccionar otro carácter como delimitador.</p>
Text qualifier (Calificador de texto)	<p>El carácter que se utiliza para demarcar los valores de texto en un archivo delimitado. Por ejemplo, este registro utiliza comillas dobles (") como calificador de texto.</p> <pre>"7200 13TH ST" "MIAMI" "FL" "33144"</pre> <p>Los caracteres disponibles para definir como calificadores de texto son:</p> <ul style="list-style-type: none"> • Comilla simple (') • Comilla doble (") <p>Si el archivo utiliza un calificador de texto distinto, haga clic en el botón de los puntos suspensivos para seleccionar otro carácter como calificador de texto.</p>

Ficha Campos

La ficha Campos define los nombres, las posiciones y los tipos de campos en el archivo. Para obtener más información, consulte [Definición de campos en un archivo de secuencia de salida](#) en la página 275

Definición de campos en un archivo de secuencia de salida

En la etapa Write to Hadoop Sequence File, la ficha **Campos** define los nombres, las posiciones y los tipos de campos en el archivo. Una vez que haya definido un archivo de entrada en la ficha **Propiedades del archivo**, podrá definir los campos.

1. Para seleccionar los campos deseados en los datos de entrada, o un archivo existente, haga clic en **Agregado rápido**.

- a) Seleccione los campos específicos de los datos de entrada.
 - b) Haga clic en **Aceptar**.
2. Para agregar nuevos campos, haga clic en **Agregar**.
- a) Ingrese el **Nombre** del campo.
 - b) Seleccione el **Tipo** del campo. La etapa admite los siguientes tipos de datos:

boolean	Tipo de dato lógico con dos valores: verdadero y falso
fecha	Un tipo de dato que contiene un mes, día y año. Por ejemplo, 2012-01-30 o 30 de enero, 2012. Puede especificar un formato de fecha predeterminado en Management Console.
fecha y hora	Un tipo de dato que contiene un mes, día, año y horas, minutos y segundos. Por ejemplo, 2012/01/30 6:15 PM.
Nota: En los archivos Parquet, los tipos de datos <code>datetime</code> y <code>time</code> se asignan como <code>String</code> . En los archivos RC, el tipo de datos <code>datetime</code> se asigna como <code>timestamp</code> .	
double	Un tipo de dato numérico que contiene números de doble precisión tanto negativos como positivos entre 2^{-1074} y $(2-2^{-52}) \times 2^{1023}$. En notación científica, el rango de valores es - 1.79769313486232E+308 a 1.79769313486232E+308.
float	Un tipo de dato numérico que contiene números de precisión simple tanto negativos como positivos entre 2^{-149} y $(2-2^{-23}) \times 2^{127}$. En notación científica, el rango de valores es - 3.402823E+38 a 3.402823E+38.
integer	Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{31} (-2,147,483,648) y $2^{31}-1$ (2,147,483,647).
bigdecimal	Un tipo de dato numérico que admite 38 puntos decimales de precisión. Utilice este tipo de datos para datos que se usarán en cálculos matemáticos que requieren un alto grado de precisión, especialmente aquellos que implican datos financieros. El tipo de datos <code>bigdecimal</code> es compatible con cálculos más precisos que el tipo de datos <code>double</code> .
Nota: Para los archivos RC, Avro y Parquet Hive, el tipo de dato <code>bigdecimal</code> se convierte en un tipo de dato <code>decimal</code> con precisión de 38 y una escala 10.;	
long	Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{63} (-9,223,372,036,854,775,808) y $2^{63}-1$ (9,223,372,036,854,775,807).
Nota: En los archivos RC, el tipo de datos <code>long</code> se asigna como tipo de datos <code>bigint</code> .	
string	Secuencia de caracteres.

- c) En el campo **Posición**, ingrese la posición de este campo dentro del registro.

Por ejemplo, en este archivo de entrada, AddressLine1 está en posición 1; City, en posición 2; StateProvince, en posición 3; y PostalCode, en posición 4.

```
"AddressLine1"|"City"|"StateProvince"|"PostalCode"
"7200 13TH ST"|"MIAMI"|"FL"|"33144"
"One Global View"|"Troy"|"NY"|12180
```

- Si está sobrescribiendo un archivo actual, haga clic en **Regenerar** para elegir el esquema del archivo actual y, luego, modifíquelo.
Esto genera el esquema a partir de los primeros 50 registros en los datos de entrada de esta etapa.
- Si desea eliminar los caracteres que ocupan demasiado espacio al principio y al final de la cadena de caracteres de un campo, seleccione la casilla **Recortar espacios**.
- Especifique una de las siguientes opciones para generar la clave:

Generar automáticamente	El clúster Hadoop genera la clave automáticamente. Para la generación automática, todos los campos en la cuadrícula se consideran campos de valor. El tipo de datos de la clave es largo.
Definido por el usuario	De manera predeterminada, el primer campo en la cuadrícula se selecciona como el campo clave. Se muestra un icono para indicar que el campo es el campo clave. Puede seleccionar cualquier otro campo como el campo clave.

Después de definir los campos de su archivo de salida, puede editar su contenido y diseño.

Nombre de la opción	Descripción
Agregar	Agrega un campo a la salida. Puede anexar un campo al final del diseño existente o insertar un campo en una posición existente y la posición de los campos restantes se ajustará según corresponda.
Modificar	Modifica el nombre y el tipo del campo.
Quitar	Elimina el campo seleccionado de la salida.
Subir/Bajar	Reordena el campo seleccionado.

Escribir en archivo Hive

La etapa **Write to Hive File** escribe la entrada del flujo de datos para el archivo de salida Hive específico.

Puede seleccionar cualquiera de los formatos de archivo Hive admitidos para el archivo de salida:

- ORC
- RC
- Parquet
- Avro

Ficha Propiedades del archivo

Tabla 5: Propiedades de archivo comunes

Campos	Descripción
Server name (Nombre de servidor)	Indica que el archivo seleccionado en el campo Nombre de archivo está ubicado en el sistema Hadoop. Una vez que selecciona un archivo ubicado en un sistema Hadoop, el Nombre del servidor refleja el nombre del servidor de archivo correspondiente, según se encuentra especificado en Management Console.
Nombre de archivo	Haga clic en el botón de los puntos suspensivos (...) para navegar al archivo de salida Hive que se creará en el servidor de archivos Hadoop definido. Los datos de salida de esta etapa se escriben en el archivo seleccionado. Nota: Debe crear una conexión con el servidor de archivos Hadoop, requerido en Management Console, antes de usarlo en esta etapa.
Tipo de archivo	Seleccione uno de los cuatro formatos de archivo Hive admitidos: <ul style="list-style-type: none"> • ORC • RC • Parquet • Avro

Tabla 6: Propiedades del archivo ORC

Campos	Descripción
Tamaño de área	<p>Define el tamaño de área que se asignará mientras se escribe a un archivo ORC. Esto se especifica en kilobytes.</p> <p>Nota: El tamaño de área predeterminado es 256 KB.</p>
Tamaño de zancada	<p>Define el tamaño de las zancadas que se crearán mientras se escribe a un archivo ORC. Esto se especifica en megabytes.</p> <p>Nota: El tamaño de zancada predeterminado es 64 MB.</p>
Intervalo de índice de fila	<p>Define la cantidad de filas que se escribirán entre dos entradas de índice de fila consecutivas.</p> <p>Nota: El intervalo de índice de fila predeterminado es 10000 filas.</p>
Tipo de compresión	<p>Define el tipo de compresión que se utilizará mientras se escribe a un archivo ORC. Los tipos de compresiones disponibles son <i>ZLIB</i> y <i>SNAPPY</i>.</p> <p>Nota: El tipo de compresión predeterminada es <i>ZLIB</i>.</p>
Relleno	<p>Indica si las zancadas tienen relleno para minimizar las zancadas que cruzan los límites de bloque HDFS, mientras se escribe a un archivo ORC.</p> <p>Nota: De forma predeterminada, la casilla de verificación Relleno está marcada.</p>
Vista previa	<p>Se recopilan los primeros 50 registros del archivo escrito y se muestran en la cuadrícula Vista previa, luego de que se ejecute el flujo de datos al menos una vez y que se hayan escrito los datos al archivo seleccionado.</p>

Tabla 7: Propiedades del archivo RC

Campos	Descripción
Tamaño de área	<p>Define el tamaño de área que se asignará mientras se escribe a un archivo RC. Esto se especifica en kilobytes.</p> <p>Nota: El tamaño de área predeterminado es 256 KB.</p>

Campos	Descripción
Tamaño del bloque	<p>Define el tamaño de los bloques que se crearán mientras se escribe a un archivo RC. Esto se especifica en megabytes.</p> <p>Nota: El tamaño de bloque predeterminado es 64 MB.</p>
Tipo de compresión	<p>Define el tipo de compresión que se utilizará mientras se escribe a un archivo RC. Los tipos de compresiones disponibles son NONE y DEFLATE.</p> <p>Nota: El tipo de compresión predeterminada es NONE.</p>
Vista previa	<p>Se recopilan los primeros 50 registros del archivo escrito y se muestran en la cuadrícula Vista previa, luego de que se ejecute el flujo de datos al menos una vez y que se hayan escrito los datos al archivo seleccionado.</p> <p>La ficha Campos se utiliza para definir la secuencia y el tipo de datos de los campos obligatorios.</p> <p>Nota: Para el tipo de archivo RC, debe definir los metadatos del archivo de salida antes de hacer clic en Vista previa para cargar la cuadrícula Vista previa.</p>

Tabla 8: Propiedades del archivo Parquet

Campos	Descripción
Tipo de compresión	<p>Define el tipo de compresión que se utilizará mientras se escribe a un archivo PARQUET. Los tipos de compresiones disponibles son UNCOMPRESSED, GZIP y SNAPPY.</p> <p>Nota: El tipo de compresión predeterminada es UNCOMPRESSED.</p>
Tamaño del bloque	<p>Define el tamaño del bloque que se creará mientras se escribe a un archivo PARQUET. Esto se especifica en megabytes.</p> <p>Nota: El tamaño de bloque predeterminado es 128 MB.</p>
Tamaño de la página	<p>El tamaño de la página es para la compresión. Cuando lee, cada página puede descomprimirse de forma independiente. Esto se especifica en kilobytes.</p> <p>Nota: El tamaño de la página predeterminada es 1024 KB.</p>

Campos	Descripción
Activar diccionario	<p>Para activar/desactivar la codificación del diccionario.</p> <p>Atención: El diccionario debe activarse para que el Tamaño de la página del diccionario se habilite.</p> <p>Nota: El valor predeterminado es <code>true</code>.</p>
Tamaño de la página del diccionario	<p>Hay una página del diccionario por columna por grupo de filas cuando se utiliza la codificación del diccionario. El tamaño de la página del diccionario funciona como el tamaño de página. Esto se especifica en kilobytes.</p> <p>Nota: El tamaño de la página del diccionario predeterminado es 1024 KB.</p>
Versión de escritor	<p>Parquet admite dos versiones API de escritor: <code>PARQUET_1_0</code> y <code>PARQUET_2_0</code>.</p> <p>Nota: El valor predeterminado es <code>PARQUET_1_0</code>.</p>
Vista previa	<p>Se recopilan los primeros 50 registros del archivo escrito y se muestran en la cuadrícula Vista previa, luego de que se ejecute el flujo de datos al menos una vez y que se hayan escrito los datos al archivo seleccionado.</p>

Tabla 9: Propiedades del archivo Avro

Campos	Descripción
Intervalo de sincronización (en bytes)	<p>Especifica la cantidad aproximada de bytes sin comprimir que se van a escribir en cada bloque. Los valores válidos oscilan entre 32 y 2^{30}. Sin embargo, se sugiere conservar el intervalo de sincronización entre 2K y 2M.</p> <p>Nota: El intervalo de sincronización predeterminado es 16000.</p>
Compresión	<p>Define el tipo de compresión que se utilizará mientras se escribe a un archivo Avro. Los tipos de compresiones disponibles son NONE, SNAPPY y DEFLATE. Elegir la compresión DEFLATE le brinda la opción adicional de seleccionar el nivel de compresión (que se describe a continuación).</p> <p>Nota: El tipo de compresión predeterminada es <code>NONE</code>.</p>

Campos	Descripción
Nivel de compresión	<p>Este campo se muestra si selecciona la opción <code>DEFLATE</code> en el campo Compresión que está arriba.</p> <p>Puede tener valores que oscilen entre 0 y 9, donde 0 indica que no existe compresión. El nivel de compresión aumenta de 1 a 9, con un aumento simultáneo en el tiempo que se utiliza para comprimir los datos.</p> <p>Nota: El nivel de compresión predeterminado es 1.</p>
Vista previa	Se recopilan los primeros 50 registros del archivo escrito, y se muestran en esta cuadrícula, luego de que se ejecute el flujo de datos al menos una vez y que se escriban los datos al archivo seleccionado.

Ficha Campos

La ficha **Campos** define los nombres y tipos de campos como se encuentran en el archivo de origen de esta etapa, y que se seleccionarán para que se escriban en el archivo de salida.

Para obtener más información, consulte [Definición de campos para la escritura de Hive File](#) en la página 282.

Ficha Tiempo de ejecución

La ficha **Tiempo de ejecución** brinda la opción de **Sobrescribir** un archivo existente del mismo nombre en el servidor de archivos Hadoop configurado. Si marca la casilla de verificación **Sobrescribir**, cuando ejecute el flujo de datos, el nuevo archivo de salida Hive sobrescribirá cualquier archivo existente que tenga el mismo nombre en el mismo servidor de archivos Hadoop.

De forma predeterminada, la casilla de verificación **Sobrescribir** está marcada.

Nota: Si no selecciona **Sobrescribir**, se arrojará una excepción cuando se ejecute el flujo de datos si el archivo que debe escribirse tiene el mismo nombre que un archivo existente en el mismo servidor de archivos Hadoop.

Definición de campos para la escritura de Hive File

En la ficha **Campos** de la etapa **Write to Hive File**, se detallan los nombres de esquema y tipos de datos de los campos en los datos de entrada de la etapa.

- Para seleccionar los campos deseados en los datos de entrada, o un archivo existente, haga clic en **Agregado rápido**.
 - Seleccione los campos específicos de los datos de entrada.
 - Haga clic en **Aceptar**.

2. Para agregar nuevos campos, haga clic en **Agregar**.

a) Ingrese el **Nombre** del campo.

b) Seleccione el **Tipo** del campo. La etapa admite los siguientes tipos de datos:

boolean Tipo de dato lógico con dos valores: verdadero y falso

fecha Un tipo de dato que contiene un mes, día y año. Por ejemplo, 2012-01-30 o 30 de enero, 2012. Puede especificar un formato de fecha predeterminado en Management Console.

fecha y hora Un tipo de dato que contiene un mes, día, año y horas, minutos y segundos. Por ejemplo, 2012/01/30 6:15 PM.

Nota: En los archivos Parquet, los tipos de datos `datetime` y `time` se asignan como `String`. En los archivos RC, el tipo de datos `datetime` se asigna como `timestamp`.

double Un tipo de dato numérico que contiene números de doble precisión tanto negativos como positivos entre 2^{-1074} y $(2-2^{-52}) \times 2^{1023}$. En notación científica, el rango de valores es - 1.79769313486232E+308 a 1.79769313486232E+308.

float Un tipo de dato numérico que contiene números de precisión simple tanto negativos como positivos entre 2^{-149} y $(2-2^{-23}) \times 2^{127}$. En notación científica, el rango de valores es - 3.402823E+38 a 3.402823E+38.

integer Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{31} (-2,147,483,648) y $2^{31}-1$ (2,147,483,647).

bigdecimal Un tipo de dato numérico que admite 38 puntos decimales de precisión. Utilice este tipo de datos para datos que se usarán en cálculos matemáticos que requieren un alto grado de precisión, especialmente aquellos que implican datos financieros. El tipo de datos `bigdecimal` es compatible con cálculos más precisos que el tipo de datos `double`.

Nota: Para los archivos RC, Avro y Parquet Hive, el tipo de dato `bigdecimal` se convierte en un tipo de dato `decimal` con precisión de 38 y una escala 10.;

long Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{63} (-9,223,372,036,854,775,808) y $2^{63}-1$ (9,223,372,036,854,775,807).

Nota: En los archivos RC, el tipo de datos `long` se asigna como tipo de datos `bigint`.

string Secuencia de caracteres.

c) En el campo **Posición**, ingrese la posición de este campo dentro del registro.

Por ejemplo, en este archivo de entrada, AddressLine1 está en posición 1; City, en posición 2; StateProvince, en posición 3; y PostalCode, en posición 4.

```
"AddressLine1"|"City"|"StateProvince"|"PostalCode"
"7200 13TH ST"|"MIAMI"|"FL"|"33144"
"One Global View"|"Troy"|"NY"|"12180"
```

3. Si sobrescribe un archivo existente, haga clic en **Regenerar** para seleccionar el esquema del archivo existente, y luego modifíquelo.

Esto genera un esquema basado en los metadatos de un archivo existente, en el caso de los archivos de salida ORC y Parquet. Para los archivos de salida RC, debe agregar los campos de forma explícita para sobrescribir los campos existentes.

La columna **Nombre** detalla los nombres de las diversas columnas de los datos de entrada. La columna **Tipo** detalla los tipos de datos de cada campo de los datos de entrada respectivamente.

Nota: En el caso del tipo de archivo *Parquet*, otra columna **Nulo** indica si el campo es nulo o no. Puede marcar esta casilla de verificación para un campo particular a fin de hacer que el campo sea nulo o desmarcarla.

4. Puede modificar los nombres, tipos de datos y secuencia de las columnas seleccionadas en la salida utilizando los siguientes botones:

Nombre de la opción	Descripción
Agregar	Agrega un campo a la salida.
Modificar	Modifica el nombre y el tipo de dato del campo seleccionado.
Quitar	Elimina el campo seleccionado de la salida.
Subir/Bajar	Reordena la posición del campo seleccionado en la salida.

5. Haga clic en **Aceptar**.

Write to NoSQL DB

En la etapa Write to NoSQL DB se escribe la salida de un flujo de datos en una base de datos. La etapa admite los tipos de base de datos MongoDB y Couchbase.

Nota: Se pueden lograr mejoras significativas del rendimiento usando múltiples instancias de ejecución de Write to NoSQLDB. Para especificar las múltiples instancias de ejecución, haga clic en el botón **Tiempo de ejecución**.

Pestaña General

Campos	Descripción
Conexión	<p>Seleccione la conexión de base de datos obligatoria en la lista desplegable. Las opciones que se muestran varían según las conexiones especificadas en Management Console.</p> <p>Para agregar una nueva conexión, consulte Conexión con NoSQL en la página 54.</p> <p>Para modificar una conexión existente, selecciónela de la lista de conexiones en la página Fuentes de datos de Management Console, ábrala, haga los cambios necesarios y, a continuación, haga clic en el botón Guardar.</p>
Tabla/Vista	<p>Especifica el nombre de la colección en la que desea escribir.</p> <p>Puede crear una nueva colección en la base de datos NoSQL ingresando el nombre de la colección en el cuadro de Tabla/Visualización y haciendo clic en Crear tabla.</p> <p>Nota: Para Couchbase se desactiva el botón 'desplegable de tabla/visualización' y 'crear tabla', ya que lo grabamos en un depósito en lugar de una vista. Además, el botón Vista previa también se desactiva.</p>
Ignorar los valores NULOS	<p>Si se activa esta opción, se ignoran los campos con valores NULOS.</p> <p>Nota: Si no activa esta opción, los campos con valor NULO se grabarán también en la base de datos.</p>

Campos	Descripción
Vista previa	<p>Muestra los registros desde la tabla seleccionada.</p> <p>Nota: Para las fuentes de datos MongoDB, si hace clic en Vista previa se mostrarán los registros filtrados, si se han ingresado una o más cláusulas <i>where</i> en el campo Donde. Si no se ha ingresado ninguna cláusula "dónde", la vista previa muestra todos los registros.</p> <p>Nota: Para las fuentes de datos Couchbase, si hace clic en Vista previa también se mostrará el campo agregado <code>_id</code> que contiene la clave. Si el registro ya tiene un campo <code>_id</code>, el campo <code>_id</code> agregado sobrescribirá el preexistente en el momento de previsualizar los campos.</p>
Expandir todo	Expande los elementos en el árbol de vista previa.
Cerrar todo	Cierra los elementos en el árbol de vista previa.

Pestaña Campos

La pestaña Campos le permite seleccionar los datos que desea escribir en la base de datos. Para obtener más información, consulte [Definición de campos en una base de datos NoSQL](#) en la página 286.

Definición de campos en una base de datos NoSQL

En la etapa Write to NoSQL DB, la ficha **Campos** define los nombres y los tipos de campos en el archivo recuperados de la etapa anterior.

1. En la **ficha Campos** haga clic en **Regenerar**.

Esto muestra los campos que surgen luego de la etapa anterior.

Los datos se muestran en el siguiente formato: `Fieldname (datatype)`.

Nota: Para Couchbase, si el registro tiene un campo `_id`, este campo se usará como clave para escribir el registro en la base de datos. Sino, la clave para el registro, se autogenerará y se escribirá en la base de datos.

2. Para modificar el nombre y tipo de un campo, resalte el campo y haga clic en **Modificar**.
3. En el campo **Nombre**, elija el campo que desea agregar o ingrese el nombre del campo.
4. En el campo **Tipo** puede dejar el tipo de datos como cadena, si no se ha propuesto realizar ninguna operación matemática con los datos. Sin embargo, si se propone realizar ese tipo de operaciones, seleccione un tipo de datos apropiado. Esto convertirá los datos de cadena del archivo en un tipo de datos que permitirá el manejo adecuado de los datos en el flujo de datos.

La etapa admite los siguientes tipos de datos:

double	Un tipo de dato numérico que contiene números de doble precisión tanto negativos como positivos entre 2^{-1074} y $(2-2^{-52})\times 2^{1023}$. En notación científica, el rango de valores es - 1.79769313486232E+308 a 1.79769313486232E+308.
float	Un tipo de dato numérico que contiene números de precisión simple tanto negativos como positivos entre 2^{-149} y $(2-2^{-23})\times 2^{127}$. En notación científica, el rango de valores es - 3.402823E+38 a 3.402823E+38.
integer	Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{31} (-2,147,483,648) y $2^{31}-1$ (2,147,483,647).
long	Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{63} (-9,223,372,036,854,775,808) y $2^{63}-1$ (9,223,372,036,854,775,807).
string	Secuencia de caracteres.

- Desde la ficha **Campos**, puede seleccionar de manera individual cada campo para que se escriba en la base de datos o hacer clic en **Seleccionar todo** para seleccionar todos los campos.
- De manera opcional, desde la ficha **Tiempo de ejecución**, puede establecer el tamaño del lote. El tamaño del lote denota la cantidad de registros que se escriben en la base de datos de una vez.
- Haga clic en **Aceptar**.

Opciones de flujo de datos para la base de datos NoSQL

Este procedimiento describe la forma de configurar un flujo de datos compatible con las opciones de ejecución para la base de datos NoSQL.

- Abra el flujo de datos en Enterprise Designer.
- Si desea configurar las opciones de tiempo de ejecución para una etapa en un flujo de datos incrustado, abra este último.
- Haga clic en el ícono Opciones de flujo de datos en la barra de herramientas o haga clic en **Editar > Opciones de flujo de datos**. Aparecerá el cuadro de diálogo **Opciones de flujo de datos**.
- Haga clic en **Agregar**. Aparecerá el cuadro de diálogo **Definir opción de flujo de datos**.
- Expanda la etapa de NoSQLDB.
- Las opciones de flujo de datos se exponen tal como se describen en la siguiente tabla:

Base de datos	Leer	Escribir
Mongo DB	Conexión	Conexión
	Tabla	Tabla
Base de datos Couchbase	Conexión	Conexión

Base de datos	Leer	Escribir
	Mostrar	
	Nombre de documento de diseño	

El nombre de la opción de base de datos NoSQL seleccionada aparece en los campos **Nombre de opción** y **Etiqueta de opción**. Este es el nombre de opción que deberá especificarse durante el tiempo de ejecución para definir esta opción.

- Ingrese una descripción de la opción en el campo **Descripción**.
- En el campo **Destino**, seleccione la opción **Etapas seleccionadas**.
- Si desea limitar los valores que pueden especificarse durante la ejecución, edite las opciones en el campo **Valores válidos** haciendo clic en el icono a la derecha del campo.
- Si desea cambiar el valor predeterminado, especifique un valor diferente en el campo **Valor predeterminado**.

Nota: En el caso de un servicio, solo puede modificar los valores predeterminados antes de exponer el servicio por primera vez. Después de exponer el servicio, ya no es posible modificar los valores predeterminados con Enterprise Designer. En cambio, debe usar Management Console. Para obtener más información, consulte [Especificación de las opciones de servicio predeterminadas](#).

- Haga clic en **Aceptar**.
- Continúe agregando opciones, según sus preferencias.
- Haga clic en **ACEPTAR** en el cuadro de diálogo **Opciones de flujo de datos** cuando haya terminado de agregar opciones.
- Guarde y exponga el flujo de datos.

Write to Spreadsheet

Write to Spreadsheet escribe datos en una hoja de cálculo de Excel como resultado de un flujo de datos.

Pestaña Propiedades del archivo

La pestaña **Propiedades del archivo** contiene opciones que se utilizan para especificar la hoja de cálculo y los datos que se escribirán desde un flujo de datos.

Nombre de campo	Descripción
Nombre de servidor	Indica que el archivo que seleccionó en el campo Nombre de archivo está ubicado en el servidor Spectrum™ Technology Platform.
Nombre de archivo	<p>Especifica la ruta al archivo. Haga clic en el botón de los puntos suspensivos (...) para buscar el archivo que desea.</p> <p>Cuando escribe un archivo en un servidor de archivos HDFS, los formatos de compresión compatibles son:</p> <ol style="list-style-type: none"> 1. GZIP (.gz) 2. BZIP2 (.bz2) <p>Nota: Incluya la extensión apropiada en el nombre de archivo para indicar el formato de compresión deseado que se debe utilizar mientras escribe en el archivo.</p> <p>Atención: Si el servidor de Spectrum™ Technology Platform está ejecutándose en Unix o Linux, recuerde que estas plataformas distinguen las mayúsculas de las minúsculas en los nombres de los archivos y las rutas.</p>
Modo de escritura	<p>Especifica cómo desea escribir datos en la hoja de cálculo para también escribirlos en el flujo de datos. Puede crear una hoja de cálculo durante el tiempo de ejecución con cualquiera de las siguientes opciones. Las opciones son las siguientes:</p> <p>Crear o sobrescribir Crea un nuevo archivo y reemplaza los datos ya existentes en el archivo de salida cada vez que se ejecuta el flujo de datos.</p> <p>Insertar Agrega el resultado del flujo de datos al área asignada y cambia los datos, si los hubiese.</p> <p>Adjuntar Agrega la salida del flujo de datos al final del archivo sin eliminar los datos ya existentes del archivo.</p>
Nombre de hoja	Especifica un nombre de hoja a la hoja de cálculo en la que desea escribir datos, en el flujo de datos.
Comenzar a escribir desde	Especifica una combinación de columna-fila (A1 o B2) o una columna desde donde desea comenzar a escribir datos. Para la opción Insertar debe proporcionar la fila y la columna. Para la opción Agregar solo puede proporcionar la columna, ya que ignora el valor de fila.
Primera fila como encabezado de columna	Especifica que la primera fila de un archivo contiene la información del encabezado, pero no datos.

Pestaña Campos

En la pestaña Campos se definen las columnas, posiciones, los tipos y valores nulos para los campos del archivo. Para obtener más información, consulte el siguiente tema:

[Definición de campos en un archivo de salida](#) en la página 290

Definición de campos en un archivo de salida

En Write to Spreadsheet, la ficha **Campos** define los nombres, posición y tipos de datos de los campos del archivo. Una vez que haya definido un archivo de salida en la ficha **Propiedades del archivo**, puede definir los campos. Si la opción **Nulo** está marcada y el campo **Nombre** contiene un valor nulo, entonces el flujo de datos escribirá el valor nulo en la hora de cálculo.

Si el archivo de salida contiene un registro de encabezado, usted puede definir los campos rápidamente haciendo clic en **Regenerar**.

Para definir campos con valores predeterminados de posición, longitud y tipo de datos, haga clic en **Agregado rápido** y seleccione los campos que desea agregar.

Si el archivo de entrada no contiene un registro de encabezado o desea definir los campos manualmente, siga estos pasos:

1. Haga clic en **Agregar**.
2. En el campo **Nombre**, elija el campo que desea agregar.
3. En el campo **Tipo**, seleccione el tipo de datos del campo proveniente del flujo de datos.

Spectrum™ Technology Platform admite los siguientes tipos de datos:

bigdecimal Un tipo de dato numérico que admite 38 puntos decimales de precisión. Utilice este tipo de datos para datos que se usarán en cálculos matemáticos que requieren un alto grado de precisión, especialmente aquellos que implican datos financieros. El tipo de datos bigdecimal es compatible con cálculos más precisos que el tipo de datos double.

boolean Tipo de dato lógico con dos valores: verdadero y falso

bytearray Una matriz (lista) de bytes.

Nota: Bytearray no se admite como datos de entrada para un servicio REST.

fecha Un tipo de dato que contiene un mes, día y año. Por ejemplo, 2012-01-30 o 30 de enero, 2012. Puede especificar un formato de fecha predeterminado en Management Console.

fecha y hora Un tipo de dato que contiene un mes, día, año y horas, minutos y segundos. Por ejemplo, 2012/01/30 6:15 PM.

double	Un tipo de dato numérico que contiene números de doble precisión tanto negativos como positivos entre 2^{-1074} y $(2-2^{-52}) \times 2^{1023}$. En notación científica, el rango de valores es - 1.79769313486232E+308 a 1.79769313486232E+308.
float	Un tipo de dato numérico que contiene números de precisión simple tanto negativos como positivos entre 2^{-149} y $(2-2^{-23}) \times 2^{127}$. En notación científica, el rango de valores es - 3.402823E+38 a 3.402823E+38.
integer	Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{31} (-2,147,483,648) y $2^{31}-1$ (2,147,483,647).
long	Un tipo de datos numérico que contiene números enteros positivos y negativos entre -2^{63} (-9,223,372,036,854,775,808) y $2^{63}-1$ (9,223,372,036,854,775,807).
string	Secuencia de caracteres.
hora	Un tipo de dato que contiene la hora del día. Por ejemplo, 21:15:59 o 9:15:59 PM.

4. Haga clic en **Agregar**.

Después de definir los campos de su archivo de salida, puede editar su contenido y diseño.

Nombre de la opción	Descripción
Agregar	Agrega un campo a la salida. Puede anexas un campo al final del diseño existente o insertar un campo en una posición existente y Write to Spreadsheet ajustará los campos restantes según corresponda.
Modificar	Modifica el nombre y el tipo del campo.
Quitar	Elimina el campo seleccionado de la salida.
Subir/Bajar	Reordena el campo seleccionado.

Write to Variable Format File

Write to Variable Format File escribe registros de diseño variable en un archivo.

Los archivos de formato variable poseen estas características:

- Los registros del archivo pueden tener campos diferentes y distinta cantidad de campos.
- Todos los registros deben contener una etiqueta (normalmente un número) que identifique el tipo de registro.
- Se admiten las relaciones jerárquicas.

Ejemplo de archivo de formato variable

En este ejemplo se muestra un archivo de formato variable que contiene información sobre la actividad de la cuenta corriente de dos clientes: Joe Smith y Anne Johnson. En este ejemplo, el archivo es un archivo delimitado que usa la coma como delimitador de campo.

```
001   Joe,Smith,M,100 Main St,555-234-1290
100   CHK12904567,12/2/2007,6/1/2012,CHK
200   1000567,1/5/2012,Fashion Shoes,323.12
001   Anne,Johnson,F,1202 Lake St,555-222-4932
100   CHK238193875,1/21/2001,4/12/2012,CHK
200   1000232,3/5/2012,Blue Goose Grocery,132.11
200   1000232,3/8/2012,Trailway Bikes,540.00
```

El primer campo de cada registro contiene la etiqueta que identifica el tipo de registro y, por lo tanto, el formato del registro:

- 001: registro del cliente
- 100: registro de la cuenta
- 200: registro de la transacción de la cuenta

Para los archivos delimitados, es común que el valor de la etiqueta (001, 100, 200) sea un número fijo de bytes ubicado al principio del registro, como se muestra en el ejemplo anterior.

Cada registro tiene su propio formato:

- 001: Nombre,Apellido,Sexo,Dirección,Número_de_teléfono
- 100: ID_de_la_cuenta,Fecha_de_apertura,Fecha_de_vencimiento,Tipo_de_cuenta
- 200: ID_de_la_transacción,Fecha_de_la_transacción,Proveedor,Monto

El formato de registro 100 (registro de cuenta) es un campo secundario del registro anterior 001, y el formato de registro 200 (registro de transacción de cuenta) es un campo secundario del registro anterior 100 (registro de cuenta). En el archivo de

ejemplo, en la cuenta CHK12904567 de Joe Smith, se realizó una transacción el 5/1/12 por un importe de 323,12 en Fashion Shoes. A su vez, en la cuenta CHK238193875 de Anne Johnson, se realizaron dos transacciones, una el 5/3/12 en Blue Goose Grocery y la otra el 8/3/12 en Trailway Bikes.

Pestaña Propiedades del archivo

Nombre de la opción	Descripción
Server name (Nombre de servidor)	Indica si el archivo que seleccionó como entrada está ubicado en la computadora que está ejecutando Enterprise Designer o en el servidor de Spectrum™ Technology Platform. Si selecciona un archivo en la computadora local, el nombre del servidor será Mi computadora. Si selecciona un archivo que se encuentra en el servidor, el nombre del servidor será Spectrum™ Technology Platform.
Nombre de archivo	Especifica la ruta al archivo. Haga clic en el botón de los puntos suspensivos (...) para buscar el archivo que desea. Nota: Si el servidor de Spectrum™ Technology Platform está ejecutándose en Unix o Linux, recuerde que estas plataformas distinguen las mayúsculas de las minúsculas en los nombres de los archivos y las rutas.
Nombre de etiqueta de raíz	La etiqueta que se usa para registros principales; es decir, de los que dependen otros tipos de registros. Por ejemplo, si usted tiene los tipos de registro 001, 100 y 200, y los tipos de registro 100 y 200 son elementos secundarios de 001, entonces 001 es la etiqueta de raíz.
Usar etiquetas de ancho fijo	Especifica si se asigna una cantidad fija de espacio al principio de cada registro para ubicar la etiqueta de registro. Por ejemplo, a continuación se muestra un archivo que posee las etiquetas 001, 100 y 200 en un campo de ancho fijo: <div style="background-color: #f0f0f0; padding: 5px; margin-top: 10px;"> <pre>001 Joe, Smith, M, 100 Main St, 555-234-1290 100 CHK12904567, 12/2/2007, 6/1/2012, CHK 200 1000567, 1/5/2012, Mike 's Shoes, 323.12</pre> </div>

Nombre de la opción	Descripción
Ancho de etiqueta	<p>Si usted marca la casilla Usar etiquetas de ancho fijo, esta opción especifica la cantidad de espacios designados para las etiquetas al principio de cada registro. Por ejemplo, si especifica 7, las primeras 7 posiciones de cada registro quedarán reservadas para la etiqueta. El valor que especifique tiene que ser superior o igual al tamaño en caracteres del nombre de etiqueta más largo. Para obtener información sobre los nombres de etiqueta, consulte Nombres de etiqueta en archivos de formato variable en la página 301.</p> <p>El valor en el campo Ancho de etiqueta aumenta automáticamente si agrega, en la pestaña Campos, campos cuyos nombres sean más largos que el valor especificado.</p> <p>El ancho de etiqueta máximo es 1024.</p>
Eliminar prefijo de etiqueta numérica	<p>Elimina el texto "NumericTag_" del nombre del campo antes de escribir la etiqueta en el archivo. La etapa Read from Variable Format File agrega el prefijo "NumericTag_" a todos los nombres de etiqueta que comienzan con un número. Esto se debe a que el nombre de etiqueta se usa como nombre de un campo de lista del flujo de datos que contiene los datos del registro, y los nombres de los campos de flujo de datos no pueden comenzar con un número. Por ejemplo, la etiqueta 100 pasará a ser el campo de lista "NumericTag_100". Si activa esta opción, el campo se escribirá en el archivo de salida como un registro con una etiqueta "100", en lugar de "NumericTag_100".</p>

Nombre de la opción	Descripción
Character encoding (Codificación de caracteres)	Codificación del archivo de texto. Seleccione uno de estos:
UTF-8	Admite todos los caracteres Unicode y es compatible con las versiones anteriores de ASCII. Para obtener más información sobre UTF, consulte unicode.org/faq/utf_bom.html .
UTF-16	Admite todos los caracteres Unicode, pero no es compatible con las versiones anteriores de ASCII. Para obtener más información sobre UTF, consulte unicode.org/faq/utf_bom.html .
US-ASCII	Codificación de caracteres basada en el orden del alfabeto inglés.
UTF-16BE	Codificación UTF-16 con serialización de bytes big endian (byte más significativo al principio).
UTF-16LE	Codificación UTF-16 con serialización de bytes little endian (byte menos significativo al principio).
ISO-8859-1	Codificación de caracteres ASCII, que suele utilizarse para los idiomas de Europa occidental. También se la conoce como Latin-1.
ISO-8859-3	Codificación de caracteres ASCII, que suele utilizarse para los idiomas de Europa meridional. También se la conoce como Latin-3.
ISO-8859-9	Codificación de caracteres ASCII, que suele utilizarse para el idioma turco. También se la conoce como Latin-5.
CP850	Página de códigos ASCII que se utiliza para escribir idiomas de Europa occidental.
CP500	Página de códigos EBCDIC que se utiliza para escribir idiomas de Europa occidental.
Shift_JIS	Codificación de caracteres para el idioma japonés.
MS932	Una extensión de Microsoft de Shift_JIS para incluir caracteres especiales NEC, selección NEC de extensiones IBM y extensiones IBM.
CP1047	Una página de códigos EBCDIC con el conjunto de caracteres completo del alfabeto latino 1.

Nombre de la opción	Descripción
Separador de campo	<p data-bbox="535 325 1429 388">Especifica el carácter que se utiliza para separar los campos en un archivo delimitado.</p> <p data-bbox="535 399 1429 441">Por ejemplo, este registro utiliza una barra vertical () como separador de campo:</p> <pre data-bbox="552 451 1412 525">7200 13TH ST MIAMI FL 33144</pre> <p data-bbox="535 535 1429 577">Estos caracteres disponibles para definir como separadores de campo son:</p> <ul data-bbox="535 588 779 787" style="list-style-type: none"> • Espacio de claves • Tab • Coma • Punto • Punto y coma • Barra vertical <p data-bbox="535 798 1429 871">Si el archivo utiliza un carácter distinto como separador de campo, haga clic en el botón de los puntos suspensivos para seleccionar otro carácter como delimitador.</p>
Separador de etiqueta	<p data-bbox="535 924 1429 1029">Especifica el carácter ubicado después del campo de etiqueta para demarcar el campo de identificación de cada registro en un archivo delimitado. El separador de etiqueta debe ser un solo carácter.</p> <p data-bbox="535 1039 1429 1102">De forma predeterminada, estos caracteres están disponibles para ser elegidos como separadores de etiqueta:</p> <ul data-bbox="535 1113 779 1312" style="list-style-type: none"> • Espacio de claves • Tab • Coma • Punto • Punto y coma • Barra vertical <p data-bbox="535 1323 1429 1428">Si el archivo utiliza un carácter distinto como separador de etiqueta, haga clic en el botón de los puntos suspensivos para agregar y seleccionar un separador de etiqueta personalizado.</p> <p data-bbox="633 1438 1429 1564">Nota: De forma predeterminada, el carácter Separador de registro es el mismo que el carácter Separador de campo seleccionado. Para activar este campo y seleccionar un carácter diferente, desmarque la casilla Igual que separador de campo.</p>
Igual que separador de campo	<p data-bbox="535 1627 1429 1690">Indica si el separador de etiqueta es el mismo que el separador de campo. Desmárquelo para seleccionar un carácter diferente como separador de etiqueta.</p> <p data-bbox="633 1701 1429 1774">Nota: De manera predeterminada, esta casilla de verificación está marcada y el campo Separador de etiqueta está desactivado.</p>

Nombre de la opción	Descripción
Text qualifier (Calificador de texto)	<p>El carácter que se utiliza para demarcar los valores de texto en un archivo delimitado. Por ejemplo, este registro utiliza comillas dobles (") como calificador de texto.</p> <pre>"7200 13TH ST" "MIAMI" "FL" "33144"</pre> <p>Los caracteres disponibles para definir como calificadores de texto son:</p> <ul style="list-style-type: none"> • Comilla simple (') • Comilla doble (") <p>Si el archivo utiliza un calificador de texto distinto, haga clic en el botón de los puntos suspensivos para seleccionar otro carácter como calificador de texto.</p>
Record separator (Separador de registro)	<p>Especifica el carácter que se utiliza para separar registros en archivos secuenciales o delimitados. Este campo no está disponible si selecciona la casilla Utilizar final de línea predeterminado.</p> <p>Los parámetros de separador de registro disponibles son:</p> <p>Unix (U+000A) Un carácter de salto de línea separa los registros. Este es el separador de registros estándar para los sistemas Unix.</p> <p>Macintosh (U+000D) Un carácter de retorno de carro separa los registros. Este es el separador de registros estándar para los sistemas Macintosh.</p> <p>Windows (U+000D U+000A) Un carácter de retorno de carro seguido por un salto de línea separa los registros. Este es el separador de registros estándar para los sistemas Windows.</p> <p>Si su archivo utiliza un separador de registro distinto, haga clic en el botón de los puntos suspensivos para seleccionar otro carácter como separador de registro.</p>
Use default EOL (Utilizar final de línea predeterminado)	<p>Especifica que el separador de registro del archivo es el carácter de final de línea (EOL) predeterminado que utiliza el sistema operativo en el que se ejecuta el servidor Spectrum™ Technology Platform.</p> <p>No seleccione esta opción si el archivo utiliza un carácter de final de línea distinto del que utiliza de forma predeterminada el sistema operativo del servidor. Por ejemplo, si el archivo utiliza un final de línea de Windows pero el servidor está ejecutándose en Linux, no seleccione esta opción. En su lugar, seleccione la opción Windows en el campo Separador de registro.</p>

Pestaña Campos

Con la pestaña **Campos** se controla qué campos del flujo de datos se incluyen en el archivo de salida.

Nombre de la opción	Descripción
Agregar	Haga clic para agregar un campo a la salida. Para obtener información sobre la construcción de campos de flujo de datos para utilizar con Write to Variable Format File, consulte Escritura de datos planos en un archivo de formato variable en la página 299.
Modificar	Haga clic para modificar el nombre de la etiqueta. Este botón se activa solamente cuando hay una etiqueta seleccionada. Cuando la opción Usar etiquetas de ancho fijo está habilitada en la pestaña Propiedades del archivo , el ancho de etiqueta se ajustará automáticamente si usted ingresa un nombre de etiqueta más largo. Nota: El uso de este botón para modificar el nombre de etiqueta de raíz equivale a modificar el valor del campo Nombre de etiqueta de raíz , en la pestaña Propiedades del archivo .
Quitar	Elimina el campo seleccionado de la salida. Si elimina un campo de lista, también elimina todos sus campos secundarios. Si elimina un campo secundario, únicamente se elimina el campo secundario seleccionado del campo de lista.
xx Eliminar todos	Elimina todos los campos de la salida.
Subir/Bajar	Reordena el campo seleccionado.

Pestaña Tiempo de ejecución

Nombre de la opción	Descripción
Nombre de archivo	Muestra el archivo definido en la pestaña Propiedades del archivo .

Nombre de la opción	Descripción
Generar varios archivos	<p>Seleccione esta opción para escribir registros en diferentes archivos en lugar de escribirlos todos en un único archivo. El archivo en el cual se escribe cada registro está especificado en el propio registro. Cada registro debe contener un campo que especifique, ya sea un nombre de archivo o la ruta completa del archivo en el cual desea que se escriba el registro. Por ejemplo, si desea enviar el precio de las acciones de diferentes empresas (de diversos grupos) a todos los clientes por separado, esta función escribe el precio de las acciones de las diferentes empresas en archivos independientes, que se pueden enviar a cada uno de los clientes, si así lo desea. Si se activa la opción Generar múltiples archivos, debe especificar un archivo de salida en el servidor Spectrum™ Technology Platform o en un servidor FTP. Si desea escribir datos desde archivos en un servidor FTP, debe definir una conexión con el servidor de archivos mediante Management Console.</p> <p>Nota: Los registros en la columna que seleccione en Campo de ruta de archivo deben estar ordenados. Utilice esta función cuando el registro contenga ya sea un nombre de archivo o la ruta de acceso del archivo completa.</p>
Campo de ruta de archivo	<p>Selecciona el campo que contiene la ruta (ya sea un nombre de archivo o la ruta completa del archivo) del archivo en el cual desea que se escriba el registro. Tenga en cuenta que en el Campo de ruta de archivo solo se enumeran los elementos sencillos de tipo asignados directamente a una etiqueta raíz. Este campo está habilitado solamente si selecciona Generar múltiples archivos.</p>
Modo de escritura	<p>Especifica si se agrega la salida del flujo de datos al final del archivo o si se eliminan los datos ya existentes en el archivo antes de escribir la salida. Una de las siguientes:</p> <p>Sobrescribir Reemplaza los datos ya existentes en el archivo de salida cada vez que se ejecuta el flujo de datos.</p> <p>Adjuntar Agrega la salida del flujo de datos al final del archivo sin eliminar los datos ya existentes del archivo.</p>

Escritura de datos planos en un archivo de formato variable

En un flujo de datos de Spectrum™ Technology Platform, cada registro tiene los mismos campos. No obstante, en un archivo de formato variable, no todos los registros contienen los mismos campos. Para escribir datos planos desde un flujo de datos en un archivo de formato variable, es necesario dividir cada registro en el flujo de datos, agrupando los campos de cada registro en campos de lista que correspondan al tipo de registro que quiera utilizar para el archivo de formato variable. Un campo de lista es una colección de campos. Por ejemplo, los campos FirstName (Nombre), LastName (Apellido), Género, Dirección y Teléfono podrían agruparse en un campo de lista denominado AccountOwner (Titular de cuenta).

Para escribir datos planos en un archivo de formato variable, utilice la etapa Aggregator (Agregador) a fin de agrupar campos en los campos de lista que correspondan a los tipos de registros que desea escribir en el archivo de formato variable. Para esto, haga lo siguiente:

1. Coloque una etapa Aggregator en su flujo de datos en cualquier sitio precedente a la etapa Write to Variable Format File.
2. Haga doble clic en la etapa Aggregator para abrir la ventana de opciones.
3. Seleccione **Group By** y haga clic en **Agregar**.
4. En el campo **Group By**, seleccione el campo que contenga un identificador único que pueda utilizarse para identificar datos relacionados. El valor de este campo debe ser único en todos los registros de los datos planos. Por ejemplo, un número de cuenta, un número de seguro social o un número de teléfono.

Nota: El campo que seleccione debe estar ordenado. Si no lo está, utilice una etapa Sorter (Ordenador) para ordenar los registros por campo.

5. Haga clic en **Aceptar**.
6. Seleccione **Listas de salida** y haga clic en **Agregar**.
Cada lista de salida representará un tipo de registro del archivo en formato variable.
7. Seleccione **New data type** y, en el campo **Type name**, especifique el tipo de información que se incluirá en este tipo de datos. Esto se convertirá en un tipo de registro en el archivo de formato variable. Por ejemplo, este tipo de datos contendrá registros relacionados con transacciones de cuenta, por lo que podría denominarlo "AccountTransaction" (Transacciones de cuenta).
8. En el campo **Name**, ingrese el nombre con el que desea designar este campo. Puede ser el mismo nombre que especifique en el campo **Type name**.
9. Haga clic en **Aceptar**.
10. Seleccione el tipo de datos que acaba de crear y haga clic en **Agregar**.
11. Deje seleccionada la opción **Existing field**, seleccione uno de los campos que desee incluir en este tipo de datos y haga clic en **OK**. Recuerde que esto se convertirá en un tipo de registro en el archivo de formato variable. Repita el proceso para agregar otros campos a este tipo de registro.
12. Cree listas de salida adicionales para cada tipo de registro que desee incluir en el archivo de formato variable. Cuando haya terminado, haga clic en **OK** para cerrar las opciones de la etapa Aggregator.

Los campos provenientes de la etapa Aggregator se han agrupado en campos de lista que corresponden a los tipos de registros que desea incluir en la salida del archivo de formato variable.

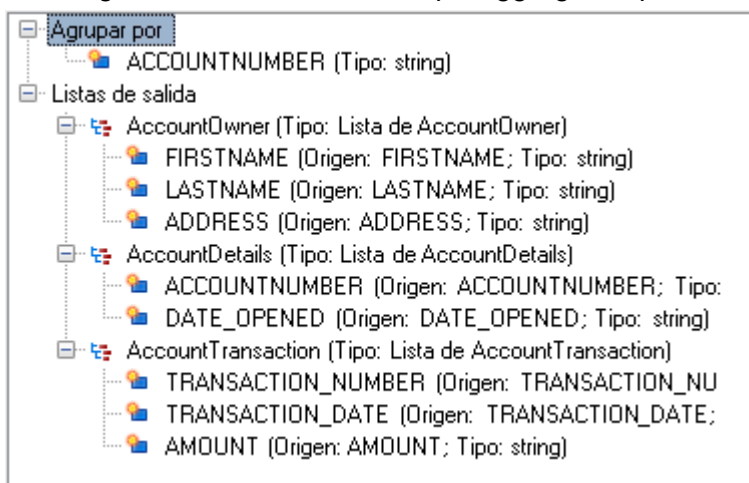
Por ejemplo, dados estos datos planos:

```
FIRSTNAME, LASTNAME, ADDRESS, ACCOUNTNUMBER, DATE_OPENED, TRANSACTION NUMBER, TRANSACTION DATE, AMOUNT
Joe, Smith, 100 Main St, CHK12904567, 12/2/2007, 1000567, 1/5/2012, 323.12
```

Sería conveniente que, en el archivo de formato variable, los convirtiera en algo similar a lo que se muestra a continuación:

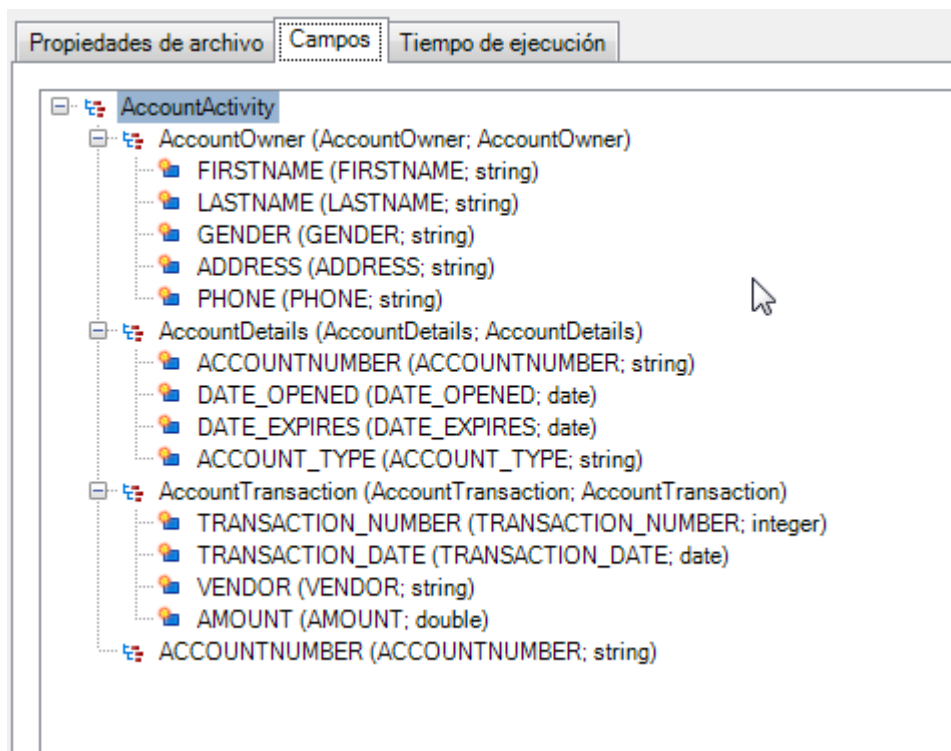
```
AccountOwner      Joe,Smith,100 Main St
AccountInformation  CHK12904567,12/2/2007
Transaction       1000567,1/5/2012,323.12
```

Para lograrlo, debe crear una etapa Aggregator que esté configurada de la siguiente manera:



Nombres de etiqueta en archivos de formato variable

En un archivo de formato variable, cada registro del archivo de salida tiene una etiqueta que indica el tipo de registro. En Write To Variable Format File (Escritura en archivo de formato variable), el nombre del campo se usa como nombre de etiqueta en el archivo de salida. Por ejemplo, observe los siguientes campos:



Estos campos se escribirían en el archivo como se muestra a continuación. Cabe destacar que, en este ejemplo, la cuenta tiene dos registros AccountTransaction.

AccountOwner	Anne, Johnson, F, 1202 Lake St, 555-222-4932
AccountDetails	CHK238193875, 1/21/2001, 4/12/2012, CHK
AccountTransaction	1000232, 3/5/2012, Blue Goose Grocery, 132.11
AccountTransaction	1000232, 3/8/2012, Trailway Bikes, 540.00

Nota: Solamente los campos de lista que contengan campos simples, como cadenas, se escriben en el archivo de salida. Si un campo de lista consiste únicamente en otros campos de lista, no se escribe en el archivo de salida. En el ejemplo anterior, ninguno de los registros con la etiqueta AccountActivity se escribirían en el archivo de salida, porque AccountActivity consiste únicamente en otros campos de lista (AccountOwner, AccountDetails y AccountTransaction).

Write to XML

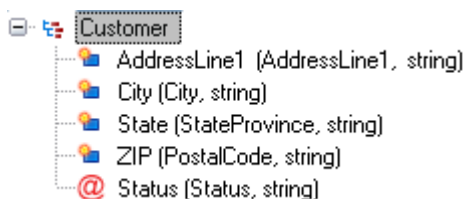
La etapa Write to XML (Escritura en XML) escribe la salida de un trabajo o subflujo en un archivo XML.

Ficha Propiedades del archivo

Nombre de campo	Descripción
Archivo de datos	<p>Especifica la ruta al archivo XML de salida. Haga clic en el botón de los puntos suspensivos (...) para buscar el archivo que desea.</p> <p>Nota: Si el servidor de Spectrum™ Technology Platform está ejecutándose en Unix o Linux, recuerde que estas plataformas distinguen las mayúsculas de las minúsculas en los nombres de los archivos y las rutas.</p>
Vista previa	<p>Muestra la estructura especificada en la ficha Campos. Si hace clic en un elemento y el archivo especificado en el campo Archivo de datos contiene dicho elemento, se mostrará una vista previa de los datos. Cabe destacar que solo pueden mostrarse en la vista previa los datos de elementos simples.</p>
Export Schema (Exportar esquema)	<p>Haga clic en este botón para guardar un archivo XSD que representa el esquema que se muestra en la vista Actual File. El archivo de esquema se guarda inmediatamente en la ubicación que especifique.</p>

Ficha Campos

La ficha **Campos** define los campos que desea incluir en el archivo XML de salida. Cuando agrega campos, se los visualiza en una estructura de árbol. El árbol muestra el nombre del elemento o atributo que se escribirá en el archivo XML. A continuación del nombre del elemento o atributo, se encuentra, entre paréntesis, el nombre del campo de flujo de datos, seguido del tipo de datos. Por ejemplo:



Esto indica que, en el archivo XML, se escribirán cuatro elementos y un atributo. El atributo se indica con el signo "@" en rojo.

Cabe destacar que el elemento Estado contendrá los datos del campo StateProvince (Estado/Provincia) y será del tipo datos en cadena. A su vez, el elemento Código postal contendrá

los datos del campo PostalCode (Código postal) y será del tipo datos en cadena. Posiblemente el archivo XML tenga la siguiente apariencia:

```
<XmlRoot>
  <Customer Status="0">
    <AddressLine1>7713 Mullen Dr</AddressLine1>
    <City>Austin</City>
    <State>TX</State>
    <ZIP>78757-1346</ZIP>
  </Customer>
  <Customer Status="0">
    <AddressLine1>1825B Kramer Ln</AddressLine1>
    <City>Austin</City>
    <State>TX</State>
    <ZIP>78758-4260</ZIP>
  </Customer>
</XmlRoot>
```

Nota: El nombre del elemento raíz (en este ejemplo <XmlRoot>) está especificado en la ficha **Propiedades del archivo**.

La siguiente tabla describe las opciones de la ficha **Campos**.

Nombre de la opción	Descripción
Agregar	Agrega un campo a la salida.

Nombre de la opción	Descripción
---------------------	-------------

Modificar	
-----------	--

Nombre de la opción

Descripción

Modifica el modo en que el campo se escribe en XML. Puede especificar lo siguiente:

Tipo de salida Esta opción está disponible si se está modificando un campo sencillo. Especifica si el campo del flujo de datos debe escribirse a un elemento o a un atributo XML.

Elemento Seleccione esta opción para escribir los datos del campo en un elemento XML. Especifique el nombre del elemento que desea utilizar en el campo **Nombre de elemento**.

Atributo Escribe los datos del campo en un atributo del elemento principal. Especifique el nombre del atributo que desea utilizar en el campo **Nombre de atributo**.

Nombre de elemento/Nombre de atributo Especifica el nombre del elemento o atributo que se escribirá en el archivo XML. El nombre predeterminado es el nombre del campo de flujo de datos.

Cambiar todos los elementos secundarios a Esta opción está disponible si se está modificando un elemento complejo. Especifica el tipo de XML que quiere que contenga el elemento complejo. Una de las siguientes:

Sin cambios Los tipos secundarios se mantienen tal como están definidos actualmente, ya sean elementos o atributos. Puede especificar individualmente el tipo de cada uno de los campos, seleccionando el campo y haciendo clic en **Modificar**.

Elementos Todos los campos sencillos bajo el elemento se escriben como elementos XML.

Atributos Todos los campos sencillos bajo el elemento se escriben como atributos XML.

Espacio de nombres: Si desea especificar un espacio de nombres XML para usar para el elemento o el atributo, puede seleccionarlo aquí. Puede crear espacios de nombres en la ficha **Campos** de la etapa Write to XML.

Incluir campos vacíos Marque esta casilla para incluir en el archivo de salida elementos XML con valores nulos o sin datos. Si no marca esta casilla, los elementos vacíos no se incluirán en la salida.

Por ejemplo, si define un elemento denominado <City> pero hay un registro que no tiene ningún dato en el campo City, el archivo de salida XML contendrá lo siguiente en caso de que marque **Incluir campos vacíos**:

```
<City xs:nil="true"></City>
```

Si no marca esta casilla, el elemento <City> no se escribirá

Nombre de la opción	Descripción
	<p data-bbox="764 338 1011 380">en el archivo de salida.</p> <p data-bbox="639 407 1419 533">Nota: Campo de flujo de datos muestra el campo cuyos datos se escribirán en el elemento o atributo. Esto se muestra de forma que, si se cambia el nombre del elemento o atributo, se pueda seguir viendo de qué campo son los datos contenidos en el elemento o atributo.</p>
Quitar	<p data-bbox="553 617 1430 709">Elimina el campo seleccionado de la salida. Si elimina un campo de lista, también elimina todos sus campos secundarios. Si elimina un campo secundario, únicamente se elimina el campo secundario seleccionado del campo de lista.</p>
Eliminar todos	<p data-bbox="553 779 963 812">Elimina todos los campos de la salida.</p>
Subir/Bajar	<p data-bbox="553 884 919 917">Reordena el campo seleccionado.</p> <p data-bbox="553 932 1430 1087">Cabe destacar que no puede mover elementos simples a elementos complejos. Si desea modificar los elementos de un elemento complejo, debe modificar la etapa Aggregator de su flujo de datos para incluir los campos de flujo de datos que desee en el elemento complejo. Para obtener más información, consulte Creación de elementos XML complejos a partir de datos planos en la página 309.</p>
Regenerar	<p data-bbox="553 1157 1430 1224">Reemplaza los campos ya definidos por los campos que ingresan en Write to XML desde el canal de carga de datos.</p>

Ficha Tiempo de ejecución

Nombre de la opción	Descripción
Generar varios archivos	<p>Seleccione esta opción para escribir registros en diferentes archivos en lugar de escribirlos todos en un único archivo. El archivo en el cual se escribe cada registro está especificado en el propio registro. Cada registro debe contener un campo que especifica, ya sea un nombre de archivo o la ruta completa del archivo en el cual desea que se escriba el registro. Por ejemplo, si desea enviar el precio de las acciones de diferentes empresas (de diversos grupos) a todos los clientes por separado, esta función escribe el precio de las acciones de las diferentes empresas en archivos independientes, que se pueden enviar a cada uno de los clientes, si así lo desea. Si se activa la opción Generar múltiples archivos, debe especificar un archivo de salida en el servidor Spectrum™ Technology Platform o en un servidor FTP. Si desea escribir datos desde archivos en un servidor FTP, debe definir una conexión con el servidor de archivos mediante Management Console.</p> <p>Nota: Los registros en la columna que seleccione en Campo de ruta de archivo deben estar ordenados. Utilice esta función cuando el registro contenga ya sea un nombre de archivo o la ruta de acceso del archivo completa.</p>
Campo de ruta de archivo	<p>Selecciona el campo que contiene la ruta (ya sea un nombre de archivo o la ruta completa del archivo) del archivo en el cual desea que se escriba el registro. Tenga en cuenta que en el campo Campo de ruta de archivo sólo se enumeran los elementos sencillos de tipo asignado directamente a una raíz. Este campo está habilitado solamente si selecciona Generar múltiples archivos.</p>
Generar esquema en el tiempo de ejecución	<p>Seleccione esta opción para generar un XSD en el tiempo de ejecución e insertar una referencia <code>noNamespaceSchemaLocation</code> al esquema en el archivo XML. El valor del atributo <code>noNamespaceSchemaLocation</code> es el nombre del archivo XSD en el archivo XML. Si exporta el esquema mientras edita el flujo de datos, no habrá referencias al XSD en el archivo XML de salida y el usuario tendrá que agregar manualmente la referencia al XSD.</p>
Ruta del esquema	<p>Especifica la ruta para guardar el archivo XSD que contiene el esquema del archivo XML de salida. Haga clic en el botón de los puntos suspensivos (...) para buscar el archivo que desea. El archivo de esquema se guarda en la ubicación que especifica cuando ejecuta el flujo de datos.</p>

Uso de espacios de nombres en un archivo XML de salida

Los espacios de nombre le permiten tener nombres de atributo y elemento duplicados en su salida mediante la asignación de cada elemento o atributo a un espacio de nombre XML.

1. En Enterprise Designer, abra el flujo de datos.
2. Haga doble clic en la etapa Write to XML en el lienzo.

3. Haga clic en la ficha **Campos**.
4. Defina uno o más espacios de nombres:
 - a) En la columna **Prefijo**, ingrese el prefijo que desea usar para asociar un elemento o atributo con el espacio de nombres.
 - b) En la columna **Namespace**, especifique la dirección URL del espacio de nombres.
 - c) Repita este paso para definir todos los espacios de nombres que desee usar para el archivo XML de salida.
5. Asocie uno o más elementos o atributos al espacio de nombres.
 - a) En la ficha **Campos**, seleccione el elemento o atributo que desea asociar con un espacio de nombres y luego haga clic en **Modificar**, o cree un nuevo elemento o atributo haciendo clic en **Agregar**.
 - b) En el campo **Namespace**, seleccione el prefijo de espacio de nombres que desee asociar con el elemento o atributo.
 - c) Haga clic en **Aceptar**.

Creación de elementos XML complejos a partir de datos planos

A menudo, los flujos de datos producen registros que contienen campos planos que son escritos en XML como elementos XML simples. Si desea organizar campos planos en elementos XML complejos para producir datos jerárquicos, puede hacerlo usando una o más etapas Aggregator.

Por ejemplo, dados los siguientes datos planos cuya primera línea es un registro de encabezado:

```
addressline1,age,city,country,gender,name,number,postalcode,stateprovince,type
1253 Summer St.,43,Boston,United States,M,Sam,019922,02110,MA,Savings
```

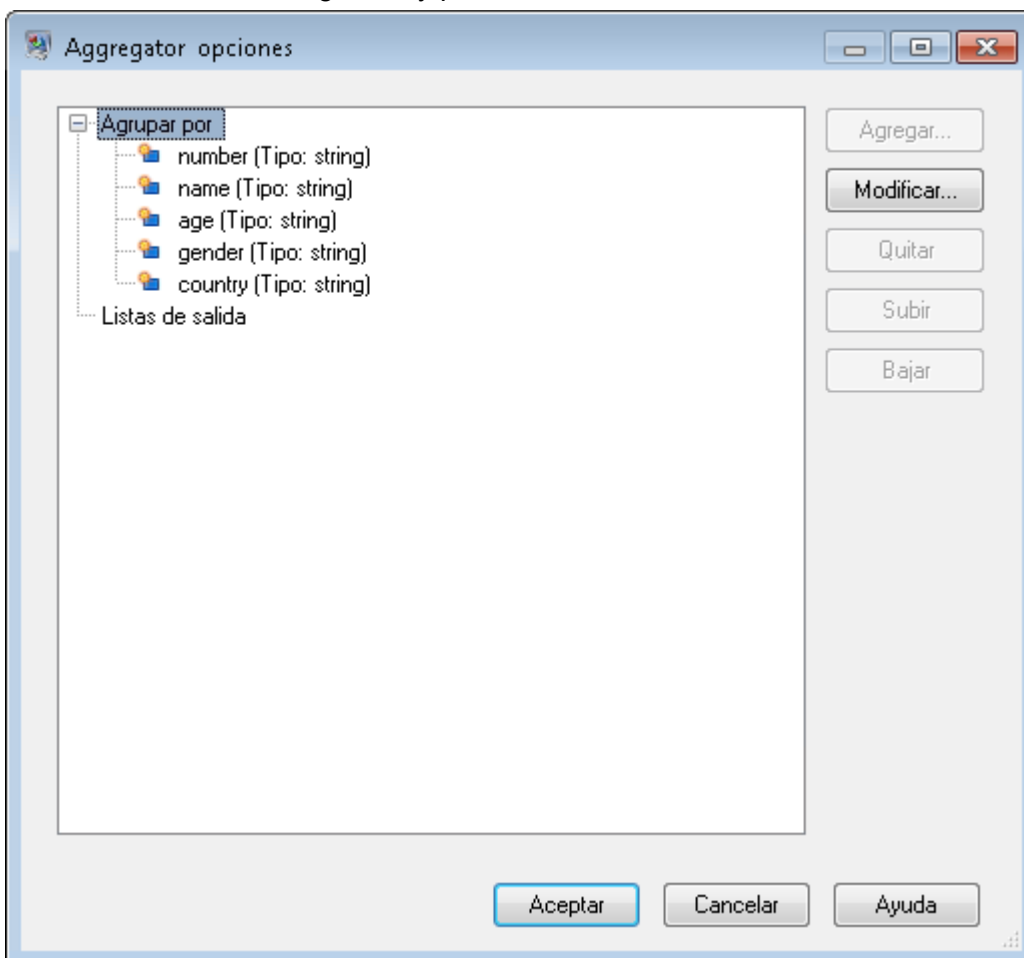
Es conveniente que agrupe los campos de datos relacionados con la dirección y los campos relacionados con la cuenta en elementos XML complejos con los nombres `<Address>` y `<Account>` asignados, como se muestra aquí:

```
<CustomerRecord>
  <name>Sam</name>
  <age>43</age>
  <gender>M</gender>
  <country>United States</country>
  <Address>
    <addressline1>1253 Summer St.</addressline1>
    <city>Boston</city>
    <stateprovince>MA</stateprovince>
    <postalcode>02110</postalcode>
  </Address>
  <Account>
    <number>019922</number>
    <type>Savings</type>
```

```
</Account>
</CustomerRecord>
```

1. Agregue una etapa Aggregator al punto en el flujo de datos en que desee construir elementos complejos.
2. Haga doble clic en la etapa Aggregator para abrir las opciones de la etapa.
3. Seleccione **Agrupar por** y haga clic en **Agregar**.
4. Seleccione el campo que contenga un valor único para cada registro, como un número de cuenta y haga clic en **Aceptar**.
5. Si hay otros campos simples que desee atravesar, vuelva a seleccionar **Agrupar por** y a hacer clic en **Agregar**, y agregue todos los campos simples que desee incluir.

Por ejemplo, en este caso, hay cinco campos simples que serán incluidos en cada registro: número, nombre, edad, género y país.

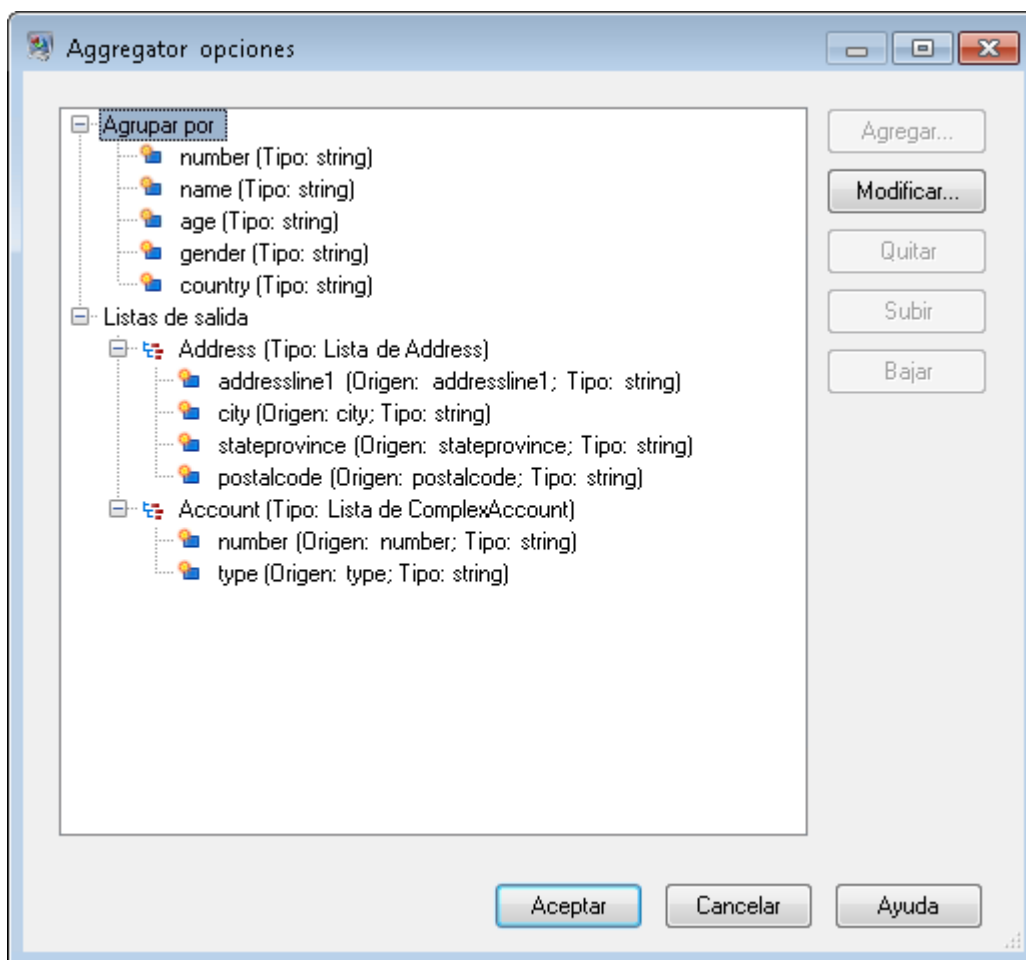


6. Seleccione **Listas de salida** y haga clic en **Agregar**.
7. Seleccione **Nuevo tipo de datos**. Esto tendrá el efecto de definir un nuevo elemento complejo. Ingrese una descripción para la clase de datos que contendrá este elemento complejo. Por

ejemplo, podría ingresar "Complejo" dado que está construyendo un elemento XML complejo. El nombre del tipo de datos queda a su elección.

8. En el campo **Nombre** ingrese el nombre que se utilizará para el campo. Este también será el nombre del elemento XML.
9. Haga clic en **Aceptar**.
10. Seleccione el campo que acaba de crear y haga clic en **Agregar**.
11. Después de seleccionar **Campo existente**, elija un campo que desee agregar como campo secundario del elemento complejo y haga clic en **Aceptar**.
12. Repita los dos pasos anteriores para agregar más campos al elemento complejo.
13. Agregue cuantos campos complejos adicionales considere necesario.

Cuando haya terminado, debería tener una etapa Aggregator que enumere cada campo simple y complejo que desee incluir en cada registro. Por ejemplo:



14. Haga clic en **Aceptar**.

Patrones de fecha y número

Patrones de fecha y hora

Al definir las opciones de tipo de dato de los datos de fecha y hora, puede crear su propio patrón personalizado de fecha u hora en caso de que los patrones predeterminados no se ajusten a sus necesidades. Para crear un patrón de fecha y hora, use la notación que se describe en la siguiente tabla. Por ejemplo, el siguiente patrón:

dd MMMM aaaa

Producirá una fecha como la que se muestra:

14 de diciembre de 2012

Letra	Descripción	Ejemplo
G	Diseñador para Era	AD
aa	Año de dos dígitos	96
aaaa	Año de cuatro dígitos	1996
M	Mes numérico del año.	7
MM	Mes numérico del año. Si el número es inferior a 10, se agrega un cero para convertirlo en un número de dos dígitos.	07
MMM	Nombre abreviado del mes	Jul
MMMM	Nombre completo del mes	Julio
w	Semana del año	27
ww	Semana del año de dos dígitos. Si la semana es inferior a 10 se agrega un cero adicional.	06

Letra	Descripción	Ejemplo
W	Semana del mes	2
D	Día del año	189
DDD	Día del año de tres dígitos. Si el número contiene menos de tres dígitos, se agregan ceros.	006
d	Día del mes	10
dd	Día del mes de dos dígitos. A los números inferiores a 10 se les agrega un cero.	09
F	Día de la semana en el mes	2
AND	Nombre abreviado del día de la semana	Mar
EEEE	Nombre completo del día de la semana	Martes
a	Marcador AM/PM	PM
H	Hora del día, siendo 0 la primera hora y 23 la última.	0
HH	Hora del día de dos dígitos, siendo 0 la primera hora y 23 la última. A los números inferiores a 10 se les agrega un cero.	08
k	Hora del día, siendo 1 la primera hora y 24 la última.	24
kk	Hora del día de dos dígitos, siendo 1 la primera hora y 24 la última. A los números inferiores a 10 se les agrega un cero.	02
k	Hora de la mañana (a.m.) o de la tarde (p.m.), siendo 0 la primera hora y 11 la última.	0
KK	Hora del día de dos dígitos, siendo 1 la primera hora y 24 la última. A los números inferiores a 10 se les agrega un cero.	02
h	Hora de la mañana (a.m.) o de la tarde (p.m.), siendo 1 la primera hora y 12 la última.	12

Letra	Descripción	Ejemplo
hh	Hora de dos dígitos de la mañana (a.m.) o de la tarde (p.m.), siendo 1 la primera hora y 12 la última. A los números inferiores a 10 se les agrega un cero.	09
m	Minuto de la hora	30
mm	Minuto de la hora de dos dígitos. A los números inferiores a 10 se les agrega un cero.	05
s	Segundo del minuto	55
ss	Segundo del minuto de dos dígitos. A los números inferiores a 10 se les agrega un cero.	02
S	Milisegundo del segundo	978
SSS	Milisegundo del segundo de tres dígitos. A los números con menos de tres dígitos se les agregarán ceros, para convertirlos en tres dígitos.	978 078 008
z	Abreviatura de tiempo del nombre de la zona horaria. Si el huso horario no cuenta con un nombre, el GMT se desplaza.	PST GMT-08:00
zzzz	El nombre completo de la zona horaria. Si el huso horario no cuenta con un nombre, el GMT se desplaza.	Hora del Pacífico GMT-08:00
Z	La zona horaria RFC 822.	-0800
X	La zona horaria ISO 8601.	-08Z
XX	La zona horaria ISO 8601 con minutos.	-0800Z
XXX	La zona horaria ISO 8601 con minutos y un separador de dos puntos entre las horas y los minutos.	-08:00Z

Patrones de números

Al definir las opciones de tipo de datos para los datos de fecha y hora, puede crear su propio patrón de números personalizado en caso de que los patrones predeterminados no se ajusten a sus necesidades. Un patrón de número básico consta de los siguientes elementos:

- Un prefijo, como un símbolo de moneda por ejemplo (opcional)
- Un patrón de números que contenga un carácter de agrupación opcional (por ejemplo una coma como separador de miles)
- Un sufijo (opcional)

Por ejemplo, el siguiente patrón:

\$ ###,###.00

Produciría un número con formato similar a este (observe el uso de un separador de miles después de los primeros tres dígitos):

\$232.998,60

Patrones para los números negativos

De forma predeterminada, los números negativos se formatean igual que números positivos, pero han añadido como prefijo el signo negativo. El carácter que se utiliza para el signo de número se basa en la configuración regional. El signo negativo es "-" en la mayoría de las configuraciones regionales. Por ejemplo, si especifica el siguiente patrón de número:

0.00

El número negativo diez tendría el siguiente formato en la mayoría de las configuraciones regionales:

-10.00

Sin embargo, si se desea definir un prefijo o sufijo diferente para utilizar en los números negativos, especifique un segundo patrón, separándolo del primer patrón con un punto y coma (";"). Por ejemplo:

0.00; (0.00)

En este modelo, los números negativos se muestran entre paréntesis:

(10.00)

Notación científica

Si desea dar formato a un número en notación científica, utilice el carácter `E` seguido por el número mínimo de dígitos que desee incluir en el exponente. Por ejemplo, dado el siguiente patrón:

0.###E0

El número 1234 tendría el siguiente formato:

1.234E3

En otras palabras, $1,234 \times 10^3$.

Cabe destacar que:

- El número de caracteres de dígitos después del exponente da el recuento de dígitos del mínimo exponente. No hay un máximo.
- Los exponentes negativos usan el formato con el signo menos localizado, no usan los prefijo y sufijos del patrón.
- La notación científica no puede contener separadores de grupos (por ejemplo, un separador de miles).

Caracteres de patrones de números especiales

Los siguientes caracteres son utilizados para producir otros caracteres, al contrario de ser reproducidos literalmente en el número resultante. Si desea utilizar alguno de estos caracteres especiales como caracteres literales en el prefijo o sufijo de su patrón de número, ponga el carácter especial entre comillas.

Símbolo	Descripción
0	Representa un dígito en el patrón incluyendo los ceros cuando es necesario completar en el patrón. Por ejemplo, el número veintisiete, cuando se aplica a este patrón: 0000 Sería: 0027
#	Representa un dígito pero se omiten los ceros. Por ejemplo, el número veintisiete, cuando se aplica a este patrón: #### Sería: 27
.	El separador de decimales o el separador decimal de la moneda utilizado en la configuración regional seleccionada. Por ejemplo, en Estados Unidos se usa el punto (.) se utiliza como separador decimal, pero en Francia, se usa una coma (,) como separador decimal.
-	El signo negativo en la configuración regional seleccionada. Para la mayoría de las configuraciones regionales es el signo menos (-).

Símbolo	Descripción
,	<p>El carácter de grupo utilizado en la configuración regional seleccionada. Se usará el carácter apropiado para la configuración regional seleccionada. Por ejemplo, en EE.UU. se usa la coma (,) como separador.</p> <p>El separador de grupo se utiliza por lo general para los miles, pero en algunos países separa las cifras de diez mil. El tamaño de grupo es un número constante de dígitos entre los caracteres de grupo, como 3 para 100.000.000 o 4 para 1.0000.0000. Si se proporciona un patrón con varios caracteres de grupo, se usa el intervalo entre el último y el extremo del número entero es el que se utiliza. Por ejemplo, todos los siguientes patrones producen el mismo resultado:</p> <pre>#, ##, ###, #### #####, ##### ##, ####, #####</pre>
E	<p>Separa la mantisa y el exponente en la notación científica. No es necesario poner la E entre comillas en su patrón. Consulte Notación científica en la página 315.</p>
;	<p>Separa los patrones secundarios positivos y negativos. Consulte Patrones para los números negativos en la página 315.</p>
%	<p>Multiplica el número por 100 y muestra el número en términos de porcentaje. Por ejemplo, el número 35, cuando se aplica a este patrón:</p> <pre>##%</pre> <p>Produciría el siguiente resultado:</p> <pre>35%</pre>
¤	<p>El símbolo de la moneda para la configuración regional seleccionada. Si se duplicó, se usa el símbolo de moneda internacional. Si está presente en un patrón monetario, se usa el separador de decimales monetarios en lugar del separador de decimales.</p>
'	<p>Se usa para poner entre comillas caracteres especiales en un prefijo o sufijo. Por ejemplo:</p> <pre>"' #' #"</pre> <p>Formatos 123 para:</p> <pre>"#123"</pre> <p>Para crear una comilla sencilla en sí, se utilizan dos seguidas:</p> <pre>"# o' 'clock"</pre>

6 - Configuraciones

In this section

Configuraciones de Oracle LogMiner

319

Configuraciones de Oracle LogMiner

Oracle LogMiner es una utilidad back-end, creada para una base de datos Oracle, que permite que Spectrum consulte registros y acceda a ellos .

La utilidad permite que Spectrum™ lea los registros de fuentes de datos de Oracle para controlar los cambios realizados en las columnas de sus tablas, como parte de la etapa **DB Change Data Reader**.

Tipos de datos y atributos de almacenamiento de tablas no compatibles

Oracle LogMiner no es compatible con los siguientes tipos de datos y atributos de almacenamiento de tablas:

- Tipo de datos **BFILE**
- Tipos de datos abstractos (**ADT**) simples y anidados
- Colecciones (tablas anidadas y VARRAY)
- Referencias a objetos
- Tablas con compresión activada
- SecureFiles

Bases de datos compatibles y versiones del archivo de registro Rehacer

LogMiner funciona en bases de datos Oracle versión 8.1 o posteriores.

También puede usar LogMiner para analizar archivos de registro Rehacer de bases de datos Oracle 8.0. Sin embargo, la información recolectada depende de la versión del registro y no de la versión de la base de datos que se utilice.

Por ejemplo, para usar LogMiner de manera óptima, los archivos de registro Rehacer de Oracle9i pueden aumentarse para capturar información adicional cuando se activa el registro suplementario. Los archivos de registro Rehacer creados con versiones anteriores de Oracle no tienen los datos adicionales y, por lo tanto, son limitados en cuanto a las operaciones y los tipos de datos compatibles con LogMiner.

*Restricciones de SQL*Loader*

Spectrum CDC puede capturar datos que se hayan cargado en las tablas Oracle mediante la utilidad SQL*Loader. Sin embargo, tiene las siguientes restricciones:

1. La carga de datos debe realizarse a través de una *ruta convencional*. Spectrum CDC no puede capturar datos que se hayan cargado mediante una *carga de ruta directa*, ya que Oracle LogMiner no es compatible con este tipo de carga.
2. El método de carga puede ser INSERT, APPEND o REPLACE.

TRUNCATE no es compatible, ya que el comando TRUNCATE hace que SQL*Loader emita el comando TRUNCATE TABLE DDL. Debido a que la función Spectrum CDC no captura el

DDL que se menciona, no se capturan las eliminaciones de filas que se generan gracias al uso del comando TRUNCATE TABLE DDL.

Privilegios de usuario necesarios

La siguiente tabla identifica los privilegios de sistema mínimos con los que deben contar los usuarios de Oracle CDC:

Privilegio de sistema	Versión de Oracle
ALTER ANY TABLE	TODOS
CONNECT	TODOS
LOCK ANY TABLE	TODOS
SELECT ANY TRANSACTION	10g o posterior

La siguiente tabla identifica los privilegios de objeto mínimos con los que deben contar los usuarios de Oracle CDC:

Nombre de objeto	Privilegio
Tablas de origen	LOCK ANY TABLE OR SELECT
PUBLIC.V\$DATABASE	SELECT
PUBLIC.V\$LOGMNR_CONTENTS	SELECT
SYS.DBMS_LOGMNR	EXECUTE
SYS.DBMS_LOGMNR_D	EXECUTE

Para obtener más información sobre Oracle LogMiner, consulte [aquí](#).

7 - Cómo optimizar el rendimiento

In this section

Determinar un tamaño de recuperación óptimo

322

Determinar un tamaño de recuperación óptimo

En la etapa **Read from DB**, el tamaño de recuperación óptimo se calcula tomando lecturas del tiempo de ejecución entre una etapa **Read from DB** y una etapa **Write to Null**.

Asegúrese de tomar los tiempos de ejecución de su prueba con los diferentes valores de tamaños de recuperación, usando su propia aplicación con el módulo Spectrum™ Data Integration.

1. Crear un trabajo en Enterprise Designer.
2. Arrastre una etapa **Read from DB** al lienzo.
3. Arrastre una etapa **Write to Null** al lienzo.
4. Cree un canal entre las dos etapas.
5. Haga doble clic en la etapa **Read from DB** para configurarla de manera tal que lea los datos de una tabla que contenga los datos de prueba.
 - a) En la pestaña **General**, en el campo **Conexión**, seleccione la base de datos que contiene los datos de prueba.
 - b) Haga clic en **Elaborar SQL** para crear la consulta SQL usando el esquema y las tablas seleccionadas de las que se deben leer los datos.
Asegúrese de que la tabla seleccionada tenga al menos 1000 registros para permitir que los valores de prueba sean óptimos.
 - c) En la pestaña **Ejecución**, marque la casilla de verificación **Tamaño de recuperación**.
 - d) En el campo adjunto, ingrese la cantidad de registros que quiere leer en una instancia.
Se ha verificado que Spectrum™ Technology Platform funciona de manera óptima con un tamaño de recuperación de hasta 1000 registros.
 - e) Haga clic en **Aceptar**.
6. Guarde el trabajo
7. Ejecute el trabajo.
Se abre la ventana **Detalles de ejecución**.
8. Haga clic en **Actualizar**.
9. Registre el tiempo **Started** (Comienzo) y el tiempo **Finished** (Finalizado).
10. Repita los pasos 7 a 9 de manera gradual aumentando el tamaño de recuperación hasta identificar la configuración óptima para su servidor.

Con esto, habrá identificado la configuración de tamaño de recuperación que le brinda un rendimiento óptimo a su entorno.

Notices

© 2018 Pitney Bowes Software Inc. Todos los derechos reservados. MapInfo y Group 1 Software son marcas comerciales de Pitney Bowes Software Inc. El resto de marcas comerciales son propiedad de sus respectivos propietarios.

Avisos de USPS®

Pitney Bowes Inc. posee una licencia no exclusiva para publicar y vender bases de datos ZIP + 4® en medios magnéticos y ópticos. Las siguientes marcas comerciales son propiedad del Servicio Postal de los Estados Unidos: CASS, CASS Certified, DPV, eLOT, FASTforward, First-Class Mail, Intelligent Mail, LACS^{Link}, NCOA^{Link}, PAVE, PLANET Code, Postal Service, POSTNET, Post Office, RDI, Suite^{Link}, United States Postal Service, Standard Mail, United States Post Office, USPS, ZIP Code, y ZIP + 4. Esta lista no es exhaustiva de todas las marcas comerciales que pertenecen al servicio postal.

Pitney Bowes Inc. es titular de una licencia no exclusiva de USPS® para el procesamiento NCOA^{Link}®.

Los precios de los productos, las opciones y los servicios del software de Pitney Bowes no los establece, controla ni aprueba USPS® o el gobierno de Estados Unidos. Al utilizar los datos RDI™ para determinar los costos del envío de paquetes, la decisión comercial sobre qué empresa de entrega de paquetes se va a usar, no la toma USPS® ni el gobierno de Estados Unidos.

Proveedor de datos y avisos relacionados

Los productos de datos que se incluyen en este medio y que se usan en las aplicaciones del software de Pitney Bowes Software, están protegidas mediante distintas marcas comerciales, además de un o más de los siguientes derechos de autor:

© Derechos de autor, Servicio Postal de los Estados Unidos. Todos los derechos reservados.

© 2014 TomTom. Todos los derechos reservados. TomTom y el logotipo de TomTom son marcas comerciales registradas de TomTom N.V.

© 2016 HERE

Fuente: INEGI (Instituto Nacional de Estadística y Geografía)

Basado en los datos electrónicos de © National Land Survey Sweden.

© Derechos de autor Oficina del Censo de los Estados Unidos

© Derechos de autor Nova Marketing Group, Inc.

Algunas partes de este programa tienen © Derechos de autor 1993-2007 de Nova Marketing Group Inc. Todos los derechos reservados

© Copyright Second Decimal, LLC

© Derechos de autor Servicio de correo de Canadá

Este CD-ROM contiene datos de una compilación cuyos derechos de autor son propiedad del servicio de correo de Canadá.

© 2007 Claritas, Inc.

El conjunto de datos Geocode Address World contiene datos con licencia de GeoNames Project (www.geonames.org) suministrados en virtud de la licencia de atribución de Creative Commons (la “Licencia de atribución”) que se encuentra en <http://creativecommons.org/licenses/by/3.0/legalcode>. El uso de los datos de GeoNames (según se describe en el manual de usuario de Spectrum™ Technology Platform) se rige por los términos de la Licencia de atribución. Todo conflicto entre el acuerdo establecido con Pitney Bowes Software, Inc. y la Licencia de atribución se resolverá a favor de la Licencia de atribución exclusivamente en cuanto a lo relacionado con el uso de los datos de GeoNames.



3001 Summer Street
Stamford CT 06926-0700
USA

www.pitneybowes.com