



Informatica®  
10.4.0

# Glosario

Este software y la documentación se proporcionan exclusivamente en virtud de un acuerdo de licencia independiente que contiene restricciones de uso y divulgación. Ninguna parte de este documento puede ser reproducida o transmitida de cualquier forma o manera (electrónica, fotocopia, grabación o mediante otros métodos) sin el consentimiento previo de Informatica LLC.

Informatica y el logotipo de Informatica son marcas comerciales o marcas comerciales registradas de Informatica LLC en Estados Unidos y en las diversas jurisdicciones de todo el mundo. La lista actual de marcas comerciales de Informatica está disponible en Internet en <https://www.informatica.com/trademarks.html>. Otros nombres de productos y empresas pueden ser nombres o marcas comerciales de sus respectivos titulares.

Las bases de datos, el software y los programas de DERECHOS DEL GOBIERNO DE LOS ESTADOS UNIDOS, y la documentación e información técnica relacionadas entregadas a los clientes del Gobierno de los Estados Unidos constituyen "software informático comercial" o "datos técnicos comerciales" de acuerdo con el Reglamento de Adquisición Federal y las regulaciones complementarias específicas del organismo que correspondan. Como tales, el uso, la duplicación, la divulgación, la modificación y la adaptación están sujetos a las restricciones y los términos de licencia establecidos en el contrato gubernamental aplicable, y hasta donde sea aplicable en función de los términos del contrato gubernamental, a los derechos adicionales establecidos en FAR 52.227-19, Licencia de Software Informático Comercial.

El producto incluye software ACE(TM) y TAO(TM) con copyright de Douglas C. Schmidt y su grupo de investigación de la Washington University, University of California, Irvine y Vanderbilt University, Copyright (©) 1993-2006, todos los derechos reservados.

Este producto incluye software Curl con Copyright 1996-2013, Daniel Stenberg, <daniel@haxx.se>. Todos los derechos reservados. Los permisos y las limitaciones relativos a este software están sujetos a los términos disponibles en la dirección <http://curl.haxx.se/docs/copyright.html>. La autorización para utilizar, copiar, modificar y distribuir este software para cualquier propósito con o sin tasas se concede por el presente, siempre que el aviso de copyright anterior y este aviso de permiso aparezcan en todas las copias. El producto incluye copyright de software 2001-2005 (©) MetaStuff, Ltd. Todos los derechos reservados. Los permisos y las limitaciones relativos a este software están sujetos a los términos disponibles en la dirección <http://www.dom4j.org/license.html>.

El producto incluye copyright de software © 2004-2007, The Dojo Foundation. Todos los derechos reservados. Los permisos y las limitaciones relativos a este software están sujetos a los términos disponibles en la dirección <http://dojotoolkit.org/license>.

Este producto incluye software ICU con copyright de International Business Machines Corporation y otros. Todos los derechos reservados. Los permisos y las limitaciones relativos a este software están sujetos a los términos disponibles en la dirección <http://source.icu-project.org/repos/icu/icu/trunk/license.html>.

Este producto incluye software OSSP UUID con Copyright © 2002 Ralf S. Engelschall, Copyright © 2002 The OSSP Project Copyright © 2002 Cable & Wireless Deutschland. Los permisos y las limitaciones relativas a este software están sujetos a los términos disponibles en la dirección <http://www.opensource.org/licenses/mit-license.php>.

La información contenida en esta documentación está sujeta a cambios sin previo aviso. Si encuentra algún problema en esta documentación, escribanos a [infa\\_documentation@informatica.com](mailto:infa_documentation@informatica.com) para notificarnoslo.

Los productos de Informatica gozan de garantía en función de los términos y condiciones de los acuerdos conforme a los cuales se proporcionen. INFORMATICA PROPORCIONA LA INFORMACIÓN DE ESTE DOCUMENTO "TAL CUAL" SIN GARANTÍA DE NINGÚN TIPO, EXPRESA O IMPLÍCITA, INCLUIDAS LAS GARANTÍAS DE COMERCIALIZACIÓN, ADAPTACIÓN A UN FIN PARTICULAR Y CUALQUIER GARANTÍA O CONDICIÓN DE NO INCUMPLIMIENTO.

Algunas partes de este software o la documentación están sujetas a derechos de autor de terceros. Se incluyen con el producto los avisos obligatorios de terceros.

Fecha de publicación: 2020-03-02

# Tabla de contenido

<b>Prefacio .....</b>	<b>4</b>
Recursos de Informatica . . . . .	4
Informatica Network. . . . .	4
Base de conocimiento de Informatica. . . . .	4
Documentación de Informatica . . . . .	4
Matrices de disponibilidad de producto de Informatica. . . . .	5
Informatica Velocity. . . . .	5
Catálogo de soluciones de Informatica. . . . .	5
Servicio internacional de atención al cliente de Informatica. . . . .	5
 <b>Apéndice A: Glosario.....</b>	 <b>6</b>

# Prefacio

Consulte el *glosario de Informatica®* para aprender la terminología propia de los productos de Informatica que se utiliza con Informatica Administrator, Informatica Analyst e Informatica Developer.

## Recursos de Informatica

Informatica proporciona una variedad de recursos de productos a través de Informatica Network y otros portales en línea. Use los recursos para sacar el mayor provecho de los productos y las soluciones de Informatica y aprender de otros expertos en la materia y usuarios de Informatica.

### Informatica Network

Informatica Network es la puerta de entrada a muchos recursos, entre ellos, la base de conocimientos de Informatica y el servicio internacional de atención al cliente de Informatica. Para entrar en Informatica Network, visite <https://network.informatica.com>.

Como miembro de Informatica Network, tiene las siguientes opciones:

- Buscar recursos de productos en la base de conocimientos
- Ver la información de disponibilidad del producto
- Crear y revisar casos de soporte
- Buscar su red de grupos de usuarios de Informatica locales y colaborar con sus pares

### Base de conocimiento de Informatica

Use la base de conocimientos de Informatica para encontrar recursos de productos como artículos prácticos, procedimientos recomendados, tutoriales de video y respuestas a preguntas frecuentes.

Para buscar en la base de conocimiento, visite <https://search.informatica.com>. Si tiene preguntas, comentarios o ideas relacionadas con la base de conocimiento de Informatica, póngase en contacto con el equipo de la base de conocimiento de Informatica en [KB\\_Feedback@informatica.com](mailto:KB_Feedback@informatica.com).

### Documentación de Informatica

Use el portal de documentación de Informatica para recorrer una extensa biblioteca de documentación para las versiones de productos actuales y recientes. Para recorrer el portal de documentación, visite <https://docs.informatica.com>.

Si tiene preguntas, comentarios o ideas acerca de la documentación de los productos, póngase en contacto con el equipo de la documentación de Informatica en [infa\\_documentation@informatica.com](mailto:infa_documentation@informatica.com).

## Matrices de disponibilidad de producto de Informatica

Las matrices de disponibilidad de producto (PAM, Product Availability Matrixes) indican las versiones de sistemas operativos, bases de datos y otros tipos de orígenes y destinos de datos admitidos por la versión de un producto. Puede recorrer las PAM de Informatica en

<https://network.informatica.com/community/informatica-network/product-availability-matrices>.

## Informatica Velocity

Informatica Velocity es una colección de consejos y procedimientos recomendados desarrollados por los servicios profesionales de Informatica que se basan en experiencias reales de cientos de proyectos de administración de datos. Informatica Velocity representa el conocimiento colectivo de los consultores de Informatica que trabajan con organizaciones de todo el mundo para planificar, desarrollar, implementar y dar mantenimiento a soluciones de administración de datos exitosas.

Puede encontrar recursos de Informatica Velocity en <http://velocity.informatica.com>. Si tiene alguna pregunta, comentario o idea acerca de Informatica Velocity, póngase en contacto con los servicios profesionales de Informatica en [ips@informatica.com](mailto:ips@informatica.com).

## Catálogo de soluciones de Informatica

El catálogo de soluciones de Informatica es un foro donde puede buscar soluciones que aumenten, amplíen o mejoren sus implementaciones de Informatica. Aproveche cualquiera de los cientos de soluciones de socios y desarrolladores de Informatica que se encuentran en el catálogo para mejorar su productividad y acelerar la implementación de los proyectos. Puede encontrar el catálogo de soluciones de Informatica en

<https://marketplace.informatica.com>.

## Servicio internacional de atención al cliente de Informatica

Puede ponerse en contacto con un centro de atención global por teléfono o a través del Informatica Network.

Para encontrar el número de teléfono local del servicio internacional de atención al cliente de Informatica, visite el sitio web de Informatica en el siguiente vínculo:

<https://www.informatica.com/services-and-training/customer-success-services/contact-us.html>.

Para encontrar recursos de soporte en línea en Informatica Network, visite <https://network.informatica.com> y seleccione la opción eSupport.

# APÉNDICE A

## Glosario

### **Administrador de Data Transformation lógico (LDTM)**

Componente de servicio del servicio de integración de datos que optimiza y compila trabajos y después los envía al Administrador de Data Transformation (DTM).

### **almacén de creación de perfiles**

Una base de datos relacional que almacena información relacionada con la creación de perfiles, como resultados de perfil y resultados de cuadro de mando.

### **aplicación**

Un objeto que se puede implementar en un servicio de integración de datos. Puede contener objetos de datos, asignaciones, servicios de datos SQL, servicios web y flujos de trabajo.

### **aplicación completa**

Una aplicación que se debe volver a implementar en el servicio de integración de datos después de editar los objetos de aplicación.

### **aplicación en tiempo de diseño**

Una aplicación que se edita en Developer tool. La aplicación contiene instancias de tiempo de diseño de los objetos de aplicación.

### **aplicación en tiempo de ejecución**

Una aplicación que se implementa en un servicio de integración de datos y contiene instancias en tiempo de ejecución de objetos de aplicación.

### **aplicación incremental**

Una aplicación que se puede actualizar mediante la implementación de una revisión de aplicación en el servicio de integración de datos.

### **archivo de esquema XSD**

Archivo \* .xsd que contiene un esquema XML que define los elementos, atributos y estructura de los documentos XML.

### **asignación**

Un conjunto de entradas y salidas vinculadas por objetos de transformación que definen las reglas de la transformación de datos.

**asignación de escritura de objeto de datos lógicos**

Una asignación que escribe datos en los destinos usando un objeto de datos lógicos como entrada. Contiene uno o varios objetos de datos lógicos como entrada y un objeto de datos físicos como destino.

**asignación de lectura de objeto de datos lógicos**

Una asignación que ofrece una vista de los datos a través de un objeto de datos lógicos. Contiene uno o varios objetos de datos físicos como orígenes y un objeto de datos lógicos como la salida de asignación.

**asignación de malla**

Asignación de Informática que el motor de Blaze compila y distribuye en un clúster de nodos.

**asignación de objeto de datos lógicos**

Una asignación que vincula un objeto de datos lógicos a uno o más objetos de datos físicos. Puede incluir una lógica de transformación.

**asignación de operación**

Una asignación que realiza la operación de servicio web para el cliente del servicio web. Una asignación de operación puede contener una transformación de entrada, una transformación de salida y varias transformaciones de fallo.

**asignación de tablas virtuales**

Una asignación que contiene una tabla virtual como destino.

**asignación dinámica**

Asignación en la que se pueden cambiar orígenes, destinos y la lógica de transformación en tiempo de ejecución en función de los parámetros y las reglas que se hayan definido. Puede configurar asignaciones dinámicas para permitir cambios de metadatos en orígenes y destinos. Puede determinar qué puertos recibe una transformación, qué puertos desea utilizar en la lógica de transformación y qué vínculos quiere establecer entre los grupos de transformación.

**base de datos virtual**

Un servicio de datos SQL que se puede consultar. Contiene objetos virtuales y proporciona una vista uniforme de los datos procedentes de orígenes diferentes y heterogéneos.

**biblioteca de definiciones de tipos**

Objeto del repositorio de modelos que almacena definiciones de tipo de datos complejo para una asignación o un mapplet.

**big data**

Conjunto de datos tan grande y complejo que no se puede procesar mediante herramientas de administración de bases de datos estándar.

**búsqueda de detección**

Un tipo de búsqueda en la Herramienta del analista que identifica los activos en función de correlaciones directas a la consulta de búsqueda, así como relaciones con otros objetos que coinciden con la consulta de búsqueda.

**canal**

Un origen y todas las transformaciones y destinos que reciben datos de ese origen. Cada asignación contiene uno o varios canales.

**carpeta**

Un contenedor de objetos en el repositorio de modelos. Utilice carpetas para organizar objetos en un proyecto y crear carpetas para agrupar objetos basados en las necesidades del negocio.

**clave candidata**

Una columna o un conjunto de columnas que identifican de forma exclusiva cada fila de origen en una tabla de base de datos.

**clave deducida**

Una columna o un conjunto de columnas que la Herramienta del analista o Developer tool deducen como clave candidata según los datos de la columna.

**clave documentada**

Una clave principal declarada en la base de datos de origen.

**Cloudera's Distribution Including Apache Hadoop (CDH)**

Versión de Cloudera del marco de software de Hadoop de código abierto.

**clúster de Hadoop**

Clúster de equipos configurado para ejecutar aplicaciones y servicios de Hadoop. Un clúster de Hadoop típico incluye un nodo principal y varios nodos de trabajo. El nodo principal ejecuta los daemons principales JobTracker y NameNode. Un nodo esclavo o de trabajo ejecuta los daemons DataNode y TaskTracker. En los clústeres pequeños, el nodo principal puede ejecutar también los daemons esclavos.

**coincidencia directa**

En una búsqueda global, una coincidencia directa es un activo que coincide con toda la consulta de búsqueda. En la búsqueda de detección, una coincidencia directa es una coincidencia en la que algunos o todos los metadatos del activo coinciden con la consulta de búsqueda.

**coincidencia indirecta**

Coincidencia de los resultados de una búsqueda de detección que está vinculada con el activo que directamente coincide con algunas o todas las consultas de búsqueda.

**columna de dependientes**

En una dependencia funcional, la columna que contiene valores que vienen determinados por una columna de determinantes.

**columna de determinantes**

En una dependencia funcional, un conjunto de columnas que determina el valor de la columna de dependientes. Si el determinante tiene cero columnas, el dependiente es constante.



**componente de inicio**

El componente ejecutable que Data Transformation inicia en primer lugar cuando ejecuta una transformación del procesador de datos.

**CompressionCodec**

Interfaz de compresión de Hadoop. Un códec es la implementación de un algoritmo de compresión y descompresión. En Hadoop, un códec se representa mediante una implementación de la interfaz CompressionCodec.

**configuración de tipo**

Conjunto de propiedades de puerto complejo que especifican el tipo de datos de los elementos de tipo de datos complejo o el esquema de los datos.

**conservación**

El proceso de validar y administrar metadatos detectados de un origen de datos para que los metadatos se puedan usar y sean adecuados para crear informes.

**contenedor**

Asignación de memoria y recursos de CPU en un nodo con la función de cálculo. Los servicios de aplicación utilizan los contenedores para realizar cálculos en los nodos de forma remota. Por ejemplo, un servicio de integración de datos que se ejecuta en una malla puede ejecutar de modo remoto una asignación dentro de un contenedor en un nodo con la función de cálculo.

**contenido de correo electrónico dinámico**

Contenido de correo electrónico definido en un parámetro o variable de flujo de trabajo.

**creación de particiones**

Proceso de dividir los datos subyacentes en subconjuntos que pueden ejecutarse en varios subprocesos de procesamiento. Cuando los administradores habilitan el servicio de integración de datos para maximizar el paralelismo, el servicio aumenta el número de subprocesos de procesamiento, que pueden optimizar el rendimiento de las asignaciones y la creación de perfiles.

**cuadro de mando**

Representación gráfica de los valores válidos de una columna de origen o resultado de una regla de resultados del perfil. Emplee cuadros de mando para medir el progreso de la calidad de los datos.

**DataNode**

Nodo HDFS que almacena datos en el sistema de archivos de Hadoop. Un clúster HDFS puede tener varios DataNode, con datos replicados entre ellos.

**datos jerárquicos**

Conjunto de datos relacionados de manera jerárquica. La relación jerárquica se representa como un esquema. Las transformaciones de Informatica utilizan tipos de datos complejos para representar datos jerárquicos.

**datos virtuales**

La información obtenida cuando se consultan tablas virtuales o se ejecutan procedimientos almacenados en un servicio de datos SQL.

**definición de tipo de datos anidado**

Definición de tipo de datos complejo que hace referencia a otras definiciones de tipo de datos complejo.

**definición de tipo de datos complejo**

Representación reutilizable del esquema de los datos a los que se hace referencia en un puerto struct o en un puerto complejo que contiene elementos de tipo struct. Uno o varios puertos complejos pueden utilizar la definición de tipo de datos complejo.

**definición de tipo de datos recursivo**

Definición de tipo de datos anidado en la que una de las definiciones de tipo de datos complejo en cualquier nivel es la misma que la de un elemento principal.

**dependencia funcional**

La relación entre un conjunto de columnas en una tabla determinada, en la cual la columna de determinantes determina de forma funcional la columna de dependientes.

**desarrollo basado en equipos**

La colaboración de los miembros del equipo en un proyecto de desarrollo. La colaboración incluye funciones como el control de versiones mediante la desprotección y protección de objetos del repositorio.

**destinatario**

Un usuario o grupo en el dominio de Informática que recibe una notificación durante un flujo de trabajo.

**destinatario dinámico**

Un destinatario de notificación definido en un parámetro o variable de flujo de trabajo.

**destino dinámico**

Archivo sin formato o destino relacional de una asignación que puede cambiar en tiempo de ejecución. Las transformaciones de escritura pueden definir columnas de destino en tiempo de ejecución en función del flujo de asignación o a partir de un destino asociado. Las transformaciones de escritura también pueden eliminar y sustituir la tabla de destino en tiempo de ejecución.

**detección de clave externa**

El proceso que busca columnas en un origen de datos que coincide con las columnas de clave principal en el origen de datos primario.

**detección de clave principal**

El proceso para identificar una columna o una combinación de columnas que identifica de manera inequívoca una fila en un origen de datos.

**detección de dependencia funcional**

El proceso que busca relaciones de dependencia funcional entre columnas en un origen de datos.

**detección del dominio de datos**

Proceso que identifica todos los dominios de datos asociados a una columna según los valores o el nombre de la columna.

**detección empresarial**

Proceso que busca las estadísticas de perfil de columna, los dominios de datos, las claves principales y las claves externas en un gran número de orígenes de datos diseminados en varias conexiones o esquemas.

**dirección de correo electrónico dinámica**

Una dirección de correo electrónico definida en un parámetro o variable de flujo de trabajo.

**documento de origen**

Documento que representa la entrada de una transformación del procesador de datos.

**documento de origen de ejemplo**

Ejemplo de los documentos que procesa una transformación del procesador de datos.

**documento de salida**

Documento que es el resultado de una transformación del procesador de datos.

**dominio de datos**

Un objeto del repositorio de modelos predefinido o definido por el usuario que representa el significado funcional de una columna según los datos o del nombre de la columna. Son ejemplos de ello el número de la seguridad social, el de la tarjeta de crédito y el ID de correo electrónico.

**ejecución de Administrador de Data Transformation (DTM)**

Componente de cálculo del servicio de integración de datos que extrae, transforma y carga datos para finalizar un trabajo de transformación de datos.

**ejecutor de Blaze**

Componente de DTM que puede simplificar y convertir una asignación en un plan de ejecución de Blaze que se ejecuta en un clúster de Hadoop.

**ejecutor de Hive**

Componente de DTM que puede simplificar y convertir una correlación o un perfil en un plan de ejecución de Hive que se ejecuta en un clúster de Hadoop.

**entorno de tiempo de ejecución**

Entorno configurado por el usuario para ejecutar una asignación o un perfil. El entorno de tiempo de ejecución puede ser nativo o de Hive.

**entorno de validación**

Entorno que se configura para validar una asignación o un perfil. Valide una asignación o un perfil para asegurarse de que se puede ejecutar en un entorno de tiempo de ejecución. El entorno de validación puede ser de Hive, nativo, o ambos.

**entorno Hadoop**

Entorno que puede configurarse para ejecutar una asignación o un perfil en un clúster de Hadoop. Debe establecer Hadoop como el entorno de validación y tiempo de ejecución.

**entorno nativo**

Entorno predeterminado del dominio de Informática que ejecuta una asignación, un flujo de trabajo o un perfil. El servicio de integración de datos realiza la extracción, transformación y carga de datos. De forma alternativa, el servicio de integración de datos podría insertar el procesamiento en un entorno no nativo, que es un clúster distribuido fuera del dominio de Informática, como Hadoop o Databricks.

**entorno no nativo**

Un clúster distribuido fuera del dominio de Informática, como Hadoop o Databricks, donde el servicio de integración de datos puede insertar el procesamiento en tiempo de ejecución. Si el servicio de integración de datos no inserta el procesamiento en un entorno no nativo, procesará el trabajo en el entorno nativo del dominio de Informática.

**entrada de tarea**

Datos que pasan a una tarea desde los parámetros y las variables del flujo de trabajo. La tarea utiliza los datos de entrada para completar una unidad de trabajo.

**especificación de ingesta masiva**

Una configuración que determina de qué forma se ingestan los datos de origen de una base de datos relacional en una ubicación específica de Data Lake o un clúster de Hadoop.

**esquema**

Definición de la estructura de los datos. Los puertos complejos de tipo de datos struct utilizan definiciones de tipo de datos complejo para representar esquemas.

**esquema virtual**

Un esquema de una base de datos virtual que define la estructura de la base de datos.

**esquema XML**

Definición de los elementos, atributos y estructura utilizados en documentos XML. El esquema cumple el estándar XML Schema del World Wide Web Consortium y se almacena como un archivo \*.xsd.

**estructura inteligente**

Un modelo del patrón y los tipos de campos de datos que Intelligent Structure Discovery identifica en un archivo de entrada. El modelo incluye tanto elementos sencillos como elementos complejos, según la organización de datos del archivo.

**evento**

Un objeto de flujo de trabajo que inicia o finaliza el flujo de trabajo. Un evento representa algo que ocurre cuando se ejecuta el flujo de trabajo. El editor muestra los eventos como círculos.

**evento de procesador de datos**

Suceso durante la ejecución de una transformación del procesador de datos.

**expresión de predicado**

Una expresión que filtra los datos de una asignación. Una expresión de predicado devuelve verdadero o falso.

**flujo de secuencia**

Un conector entre objetos de flujo de trabajo que especifica el orden en el que el servicio de integración de datos ejecuta los objetos. El editor muestra flujos de secuencia como flechas.

**flujo de secuencia condicional**

Un flujo de secuencia que incluye una expresión que el servicio de integración de datos evalúa como TRUE o FALSE. Si la expresión tiene un resultado TRUE, el servicio de integración de datos ejecuta el siguiente objeto en el flujo de trabajo. Si la expresión tiene como resultado FALSE, el servicio de integración de datos no ejecuta el siguiente objeto en el flujo de trabajo.

**flujo de secuencia predeterminado**

El flujo de secuencia saliente desde una puerta de enlace exclusiva que siempre tienen como resultado TRUE. Cuando todos los demás flujos de secuencia condicionales dan como resultado FALSE, el servicio de integración de datos ejecuta el objeto conectado al flujo de secuencia saliente predeterminado.

**flujo de trabajo**

Una representación gráfica de un conjunto de eventos, tareas y decisiones que definen un proceso de negocio. Puede utilizar Developer tool para añadir objetos a un flujo de trabajo y conectar los objetos con flujos de secuencia. El servicio de integración de datos utiliza las instrucciones configuradas en el flujo de trabajo para ejecutar los objetos.

**función compleja**

Tipo de función predefinida en la que la entrada o la devolución es un tipo de datos complejo.

**función del nodo**

El propósito de un nodo. Un nodo con la función de servicio puede ejecutar servicios de aplicación. Un nodo con la función de cálculo puede realizar cálculos solicitados por servicios de aplicación remotos. Un nodo con ambas funciones puede ejecutar servicios de aplicación y realizar cálculos de forma local para esos servicios.

**función de usuario**

Colección de privilegios que se asigna a un usuario o grupo. Asigne funciones a los usuarios y grupos del dominio y a algunos de los servicios de aplicación en el dominio.

**glosario de dominio de datos**

Contenedor donde están todos los dominios de datos y grupos de dominios de datos de la Herramienta del analista o de Developer tool.

**grupo de dominio de datos**

Conjunto de dominios de datos bajo una categoría de dominio de datos concreta.

**grupo de mediciones**

Un grupo de mediciones definido por el usuario.

## **Hadoop Distributed File System (HDFS)**

Sistema de almacenamiento de archivos distribuido que utilizan las aplicaciones de Hadoop.

## **Herramienta Informatica Monitoring**

La Herramienta Informatica Monitoring es una aplicación que proporciona un vínculo directo a la ficha Supervisar de la Herramienta del administrador. La ficha Supervisar muestra propiedades, además de estadísticas e informes de tiempo de ejecución acerca de los objetos de integración que se ejecutan en un servicio de integración de datos.

## **Hive**

Infraestructura de almacén de datos creada sobre Hadoop. Hive admite un lenguaje similar a SQL denominado HiveQL para realizar resúmenes, consultas y análisis de datos.

## **ID de instancia de flujo de trabajo**

Un número que identifica de forma exclusiva una instancia de flujo de trabajo que se ha ejecutado.

## **implementar**

Hacer que los objetos de una aplicación estén accesibles para los usuarios finales: Dependiendo de los tipos de objetos de la aplicación, los usuarios finales pueden realizar consultas para los objetos, acceder a los servicios web, ejecutar asignaciones o ejecutar flujos de trabajo.

## **Informatica Administrator**

Informatica Administrator (Herramienta del administrador) es una aplicación que consolida las tareas administrativas de los objetos de dominio como servicios, nodos, licencias y mallas. Puede administrar el dominio y la seguridad del dominio mediante la Herramienta del administrador.

## **Informatica Developer**

Informatica Developer (la herramienta Developer) es una aplicación que se utiliza para diseñar soluciones de integración de datos. El repositorio de modelos almacena los objetos creados en la herramienta Developer.

## **Informatica Mass Ingestion**

Informatica Mass Ingestion (la herramienta de ingesta masiva) es una aplicación que se puede usar para configurar, implementar, ejecutar y supervisar las especificaciones de ingesta masiva.

## **ingesta masiva**

El movimiento de grandes cantidades de datos entre una base de datos relacional y Data Lake o un clúster de Hadoop.

## **instancia de DTM**

Representación lógica y específica de la ejecución del Administrador de Data Transformation (DTM) que el servicio de integración de datos crea para ejecutar un trabajo. En función de cómo esté configurado el servicio de integración de datos, las instancias de DTM pueden ejecutarse en el proceso del servicio de integración de datos, en un proceso DTM independiente en el nodo local, o en un proceso DTM independiente en un nodo remoto.

**instancia de flujo de trabajo**

La representación de tiempo de ejecución de un flujo de trabajo. Cuando ejecuta un flujo de trabajo desde una aplicación implementada, debe ejecutar una instancia del flujo de trabajo. Puede simultáneamente ejecutar varias instancias del mismo flujo de trabajo.

**Intelligent Structure Discovery**

Un servicio en los servicios de nube inteligentes de Informática que utiliza algoritmos de aprendizaje automático para identificar de forma automática los tipos de información de un archivo de entrada. Intelligent Structure Discovery crea una estructura inteligente, un modelo del patrón y los tipos de campos de datos detectados en el archivo. El modelo incluye tanto elementos sencillos como elementos complejos, según la organización de datos del archivo.

**JobTracker**

Servicio de Hadoop que coordina tareas de asignación y de reducción y las programa para ejecutarse en los TaskTrackers.

**linaje de cuadro de mando**

Un diagrama que muestra el origen de los datos, describe la ruta y muestra cómo fluyen los datos para una métrica o grupo de métricas en un cuadro de mando. En el análisis del linaje de cuadro de mando, los cuadros o nodos representan objetos. Las flechas representan las relaciones en el flujo de datos.

**mapplet**

Un objeto reutilizable que contiene un conjunto de transformaciones que puede utilizar en varias asignaciones o validar como una regla.

**MapReduce**

Modelo de programación para procesar grandes volúmenes de datos en paralelo.

**medición**

Una columna de un origen o resultado de datos de una regla que forma parte de un cuadro de mando.

**memoria caché del conjunto de resultados**

Una caché que contiene los resultados de cada consulta del servicio de datos SQL o de cada solicitud de servicio web. Con la memoria caché del conjunto de resultados, el servicio de integración de datos devuelve los resultados almacenados en la memoria caché cuando los usuarios ejecutan consultas idénticas. La memoria caché del conjunto de resultados reduce el tiempo de ejecución en el caso de consultas idénticas.

**metastore**

Base de datos que Hive utiliza para almacenar los metadatos de las tablas de Hive almacenadas en HDFS. Los metastores pueden ser locales, incorporados o remotos.

**Minigráfico**

Un minigráfico es un gráfico de líneas que muestra la variación de un valor nulo, un valor exclusivo o un valor no exclusivo en las cinco últimas ejecuciones de perfil consecutivas.

**modelo de objeto de datos lógicos**

Un modelo de datos que describe los datos de una organización y la relación entre ellos. Contiene objetos de datos lógicos y define las relaciones entre ellos.

**Módulo del servicio de flujo de trabajo**

Un componente del servicio de integración de datos que administra las solicitudes para ejecutar flujos de trabajo.

**Módulo del servicio SQL**

El servicio de componentes del servicio de integración de datos que administra las consultas SQL enviadas a un servicio de datos SQL desde una herramienta cliente de terceros.

**Módulo del servicio web**

Un componente del servicio de integración de datos que administra las solicitudes de operación de servicio web enviadas al servicio web desde un cliente de servicio web.

**NameNode**

Nodo del clúster de Hadoop que gestiona el espacio de nombres del sistema de archivos, mantiene el árbol del sistema de archivos y los metadatos de todos los archivos y directorios del árbol.

**nodo**

Una representación de un nivel de la jerarquía de un mensaje de servicio web.

**objeto de datos físicos**

Una representación física de datos que se utiliza para leer recursos, buscarlos y escribir en ellos.

**objeto de datos lógicos**

Un objeto que describe una entidad lógica de una organización. Tiene atributos y claves y describe las relaciones entre los atributos.

**objeto de datos personalizados**

Un objeto de datos físicos que utiliza como orígenes uno o varios recursos relacionales relacionados u objetos de datos relacionales. Utilice objetos de datos personalizados para realizar tareas tales como unir datos de recursos relacionados o filtrar filas. Los objetos de datos personalizados utilizan una única conexión y una única instrucción SQL para las tablas de origen.

**operación de escritura de objeto de datos**

Objeto de repositorio que contiene las propiedades necesarias para realizar ciertas operaciones en tiempo de ejecución en destinos. Una operación de escritura de objeto de datos está asociada a un objeto de datos de destino.

**operación de lectura de objeto de datos**

Objeto de repositorio que contiene las propiedades necesarias para realizar ciertas operaciones en tiempo de ejecución en orígenes. Una operación de lectura de objeto de datos está asociada a un objeto de datos de origen.



**operador complejo**

Tipo de operador que permite hacer referencia a nombres de elementos o acceder a ellos en un tipo de datos complejo.

**optimización basada en el coste**

Método de optimización que reduce el tiempo de ejecución en asignaciones que realizan operaciones join. Con la optimización basada en el coste, el servicio de integración de datos crea planes diferentes para ejecutar una asignación y calcula el coste de cada plan. El servicio de integración de datos ejecuta el plan con el coste más pequeño. Además, calcula el coste en función de las estadísticas de la base de datos y de la información de E/S, CPU, red y memoria.

**optimización de inserción**

Método de optimización que inserta la lógica de transformación en una base de datos de origen o de destino. Con la optimización de inserción, el servicio de integración de datos convierte la lógica de transformación en consultas SQL y las envía a la base de datos. La base de datos ejecuta las consultas SQL para procesar los datos.

**optimización del predicado**

Método de optimización que simplifica o reescribe las expresiones de predicado de una asignación. Con la optimización del predicado, el servicio de integración de datos intenta aplicar expresiones de predicado tan pronto como es posible para mejorar el rendimiento de la asignación.

**optimización de primera proyección**

Método de optimización que reduce la cantidad de datos que se mueven entre transformaciones en la asignación. Con la optimización de primera proyección, el servicio de integración de datos identifica los puertos sin utilizar y elimina los vínculos entre los puertos en una asignación.

**optimización de primera selección**

Método de optimización que reduce el número de filas que pasan por la asignación. Con la optimización de primera selección, el servicio de integración de datos acerca los filtros al origen de asignación en la canalización.

**optimización de semi-join**

Método de optimización que reduce el número de filas extraídas del origen. Con la optimización de semi-join, el servicio de integración de datos modifica las operaciones join de una asignación. El servicio de integración de datos aplica el método de optimización de semi-join a una transformación Joiner cuando un grupo de entrada grande tiene filas que no coinciden con un grupo de entrada más pequeño en la condición join. El servicio de integración de datos lee las filas de los grupos más pequeños, encuentra las filas coincidentes en el grupo más grande y ejecuta la operación de unión.

**origen dinámico**

Archivo sin formato u origen relacional de una asignación que puede cambiar en tiempo de ejecución. Las transformaciones de lectura y de búsqueda pueden obtener cambios de definiciones o metadatos directamente desde el origen. Si utiliza un parámetro para el origen, puede cambiar el origen en tiempo de ejecución.

**parámetro de flujo de trabajo**

Un valor constante que define antes de ejecutar el flujo de trabajo. Los parámetros conservan el mismo valor durante toda la ejecución del flujo de trabajo. Defina el valor del parámetro en un archivo de parámetros. Todos los parámetros de flujo de trabajo están definidos por el usuario.

**perfil**

Un objeto que contiene reglas para detectar patrones en datos de origen. Puede ejecutar un perfil para evaluar la estructura de datos y para verificar que las columnas de datos contengan el tipo de información esperada.

**perfil de columna**

Un tipo de perfil que determina las características de las columnas en un origen de datos, tales como frecuencias de valor, porcentajes, patrones y tipos de datos.

**perfil de detección empresarial**

Un tipo de perfil que se utiliza para realizar la detección empresarial.

**perfil de objeto de datos**

Un objeto del repositorio que define el tipo de análisis que realiza en un origen de datos.

**perfil de sistema operativo**

Un tipo de seguridad que los servicios de integración de datos en UNIX o Linux utilizan para aislar el entorno de usuario de tiempo de ejecución. El perfil de sistema operativo contiene las variables del proceso de servicio, las variables de entorno, los permisos y el nombre de usuario del sistema operativo. El servicio de integración de datos ejecuta asignaciones, flujos de trabajo, trabajos de creación de perfiles y cuadros de mando con los permisos del sistema del usuario del sistema operativo y las propiedades que se han definido en el perfil de sistema operativo.

**perfil de unión**

Un tipo de perfil que determina el grado de superposición entre un conjunto de una o más columnas en un origen de datos y un conjunto similar en el mismo o en diferentes orígenes de datos.

**plan de ejecución Hive**

Una serie de tareas de Hive que el ejecutor de Hive genera después de procesar una asignación o un perfil. Plan de ejecución de Hive que también se puede denominar flujo de trabajo de Hive.

**ponderación de la medición**

Entero mayor o igual a 0 asignado a una medición. La ponderación de la medición define la contribución de esta a la puntuación del grupo de mediciones.

**preprocesador**

Procesador de documentos que se utiliza para realizar modificaciones generales en un documento de origen antes de ejecutar la transformación principal.

**procedimiento almacenado virtual**

Un conjunto de instrucciones de procedimientos o de flujo de datos en un servicio de datos SQL.

**procesador de documentos**

Componente que actúa en un documento como un todo, por lo general, realizando conversiones preliminares antes del análisis.

**proceso DTM**

Proceso de sistema operativo que el servicio de integración de datos inicia para ejecutar instancias de DTM. En función de la configuración del servicio de integración de datos, el servicio puede ejecutar cada instancia de DTM en un proceso DTM independiente en un nodo local o remoto.

**proyecto**

El contenedor de nivel superior para almacenar los objetos creados en Informatica Analyst e Informatica Developer. Cree proyectos en función de sus objetivos o requisitos empresariales. Los proyectos aparecen tanto en Informatica Analyst como en Informatica Developer.

**puerta de enlace**

Un objeto de flujo de trabajo que divide y fusiona rutas de acceso del flujo de trabajo según cómo el servicio de integración de datos evalúe las expresiones en flujos de secuencia condicionales. El editor muestra puertas de enlace como diamantes.

**Puerta de enlace exclusiva**

Una puerta de enlace que representa una decisión tomada en un flujo de trabajo. Cuando una puerta de enlace exclusiva divide el flujo de trabajo, el servicio de integración de datos toma una decisión para tomar una de las ramas salientes. Cuando una puerta de enlace exclusiva fusiona el flujo de trabajo, el servicio de integración de datos espera a una ramificación entrante antes de desencadenar la rama saliente.

**puerto complejo**

Tipo de puerto al que se le asigna un tipo de datos complejo (como array, struct o map) para pasar datos jerárquicos.

**puerto complejo anidado**

Puerto complejo que contiene una definición de tipo de datos complejo anidado.

**puerto dinámico**

Puerto que puede recibir una o varias columnas procedentes de una transformación de nivel superior y crear un puerto generado para cada columna.

**puerto generado**

Puerto dentro de un puerto dinámico que representa una única columna. Developer tool genera puertos en función de una o varias reglas de entrada.

**puerto variable con estado**

Puerto variable que hace referencia a los valores de las filas anteriores.

**punto de partición**

Límite entre etapas de un canal de asignación. Cuando se habilita la creación de particiones, el servicio de integración de datos puede redistribuir filas de datos en puntos de partición.

**puntuación de la medición**

El porcentaje de valores válidos de la medición.

**puntuación del grupo de mediciones**

Media ponderada computada de todas las puntuaciones de las mediciones del grupo de mediciones.

**regla**

Lógica empresarial reutilizable que define las condiciones que se aplican a los datos de origen al ejecutarse un perfil. Emplee reglas para validar los datos de un perfil y para medir el progreso de la calidad de los datos. Puede crear una regla en Informatica Analyst o en Informatica Developer.

**regla de datos**

Lógica empresarial reutilizable que identifica una columna por sus valores como perteneciente a un dominio de datos en concreto.

**regla de entrada**

Regla que determina qué puertos generados se crearán dentro de un puerto dinámico.

**regla de nombre de columna**

Lógica empresarial reutilizable que identifica una columna por su nombre como perteneciente a un dominio de datos concreto.

**revisión de aplicación**

Un objeto que se puede implementar en un servicio de integración de datos. Contiene los metadatos que se pueden usar para actualizar una aplicación incremental implementada.

**salida de tarea**

Datos que pasan desde una tarea a las variables del flujo de trabajo. Cuando configure una tarea, especifique los valores de salida de la tarea que desea asignar a variables del flujo de trabajo. El servicio de integración de datos copia los valores de salida de la tarea a las variables del flujo de trabajo cuando la tarea finaliza. El servicio de integración de datos puede acceder a estos valores desde las variables del flujo de trabajo cuando evalúa las expresiones de los flujos de secuencia condicionales y cuando ejecuta más objetos en el flujo de trabajo.

**secuencias de comandos de Hive**

Secuencia de comandos en lenguaje de consultas de Hive que contiene consultas y comandos de Hive para ejecutar la asignación.

**segmento de malla**

Fragmento de una asignación de malla que está incluido en una tarea de malla.

**Servicio de administrador de recursos**

Servicio del sistema que gestiona recursos de cálculo en el dominio y distribuye trabajos para alcanzar un rendimiento y una escalabilidad óptimos. El servicio de administrador de recursos recopila información acerca de los nodos con la función de cálculo. El servicio empareja los requisitos de los trabajos con la disponibilidad de los recursos para identificar el nodo de cálculo idóneo para ejecutar un trabajo. El servicio de administrador de recursos se comunica con los nodos de cálculo en una malla del servicio de integración

de datos. Habilite el servicio de administrador de recursos cuando configure una malla del servicio de integración de datos para ejecutar trabajos en procesos remotos independientes.

### **servicio de aplicación**

Servicio que se ejecuta en uno o varios nodos en el dominio de Informatica. Puede crear y administrar servicios de aplicación en Informatica Administrator o mediante el programa de comandos infacmd. Los servicios de aplicación incluyen servicios que pueden tener varias instancias en el dominio y servicios del sistema que pueden tener una única instancia en el dominio. Configure cada servicio de aplicación en función de los requisitos de su entorno.

### **servicio de datos**

Un conjunto de operaciones reutilizables que puede ejecutar para acceder a los datos y transformarlos. Un servicio de datos proporciona un modelo de datos unificado al que puede acceder a través de un servicio web o contra el que puede ejecutar una consulta SQL.

### **servicio de datos SQL**

Una base de datos virtual que puede consultar. Contiene objetos virtuales y proporciona una vista uniforme de los datos procedentes de orígenes diferentes y heterogéneos.

### **Servicio de ingesta masiva**

Un servicio de aplicación que administra las especificaciones de ingesta masiva que se crean en la herramienta de ingesta masiva. El servicio de ingesta masiva valida las especificaciones, programa las especificaciones que se van a ejecutar en un servicio de integración de datos y supervisa las estadísticas de la ingesta.

### **Servicio de integración de datos**

Un servicio de aplicación que realiza trabajos de integración de datos para Informatica Analyst, Informatica Developer y clientes externos. Los trabajos de integración de datos incluyen la vista previa de datos y la ejecución de asignaciones, perfiles, servicios de datos SQL, servicios web y flujos de trabajo.

### **servicio de repositorio de modelos**

Un servicio de aplicación en el dominio de Informatica que ejecuta y administra el repositorio de modelos. El repositorio de modelos almacena los metadatos creados por los productos de Informatica en una base de datos relacional para habilitar la colaboración entre productos.

### **servicio de sistema**

Servicio de aplicación que puede tener una sola instancia en el dominio. Al crear el dominio, los servicios del sistema se crean automáticamente. Puede habilitar, deshabilitar y configurar servicios de sistema.

### **tabla virtual**

Una tabla de una base de datos virtual.

### **tarea**

Un objeto de flujo de trabajo que ejecuta una unidad de trabajo en el flujo de trabajo, tal como ejecutar una asignación, enviar un correo electrónico o ejecutar un comando de shell. Una tarea representa algo que se debe realizar durante el flujo de trabajo. El editor muestra las tareas como cuadrados.

**tarea de comando**

Tarea de preprocesamiento o posprocesamiento de los datos locales de un flujo de trabajo de motor de Blaze.

**tarea de Hive**

Tarea del plan de ejecución de Hive. Un plan de ejecución de Hive contiene muchas tareas de Hive. Las tareas de Hive contienen secuencias de comandos de Hive.

**tarea de malla**

Solicitud de trabajo de procesamiento en paralelo. Cuando la asignación se ejecuta en el entorno Hadoop, el ejecutor del motor de Blaze envía la solicitud al administrador de mallas. Cuando la asignación se ejecuta en el entorno nativo y el servicio de integración de datos se ejecuta en modo remoto, el servicio de integración de datos envía la solicitud al administrador de servicios del nodo de cálculo principal.

**tasklet**

Partición de un segmento de malla que se ejecuta en un DTM independiente.

**TaskTracker**

Nodo del clúster de Hadoop que ejecuta tareas, como las de asignación o de reducción, por ejemplo. Los TaskTrackers envían informes de progreso al JobTracker.

**tipo de datos anidado**

Tipo de datos complejo que contiene al menos un elemento de un tipo de datos complejo. Por ejemplo, un tipo de datos struct que contiene un elemento de tipo array.

**tipo de datos complejo**

Tipo de datos que permite representar varios valores de datos en una sola posición de columna. Los valores de datos se denominan elementos.

**tipo de datos primitivo**

Tipo de datos que permite representar un único valor de datos en una sola posición de columna.

**trabajo MapReduce**

Unidad de trabajo que consta de los datos de entrada, el programa MapReduce e información de configuración. Hadoop ejecuta el trabajo MapReduce dividiéndolo en tareas de asignación y tareas de reducción.

**transformación**

Un objeto de repositorio en una asignación que genera, modifica o transfiere datos. Cada transformación realiza una función diferente.

**transformación de servicio web**

Una transformación que procesa solicitudes o respuestas de servicio web. Ejemplos de transformaciones de servicios web son transformación de entrada, transformación de salida, transformación de fallo y la transformación de consumidor de servicio web.

**Valor atípico**

Un valor atípico es un patrón, un valor o una frecuencia de una columna en los resultados de perfil que se sale del rango de valores esperado.

**variable de flujo de trabajo**

Un valor que puede cambiar durante la ejecución de un flujo de trabajo. Utilice variables del flujo de trabajo para hacer referencia a variables y registrar información de tiempo de ejecución. Puede utilizar variables del flujo de trabajo definidas por el usuario o del sistema.

**variable de flujo de trabajo definida por el usuario**

Una variable de flujo de trabajo que captura la salida de tarea o captura los criterios que especifique. Después de crear una variable de flujo de trabajo definida por el usuario, configure el flujo de trabajo para asignar un valor de tiempo de ejecución a la variable.

**variable de flujo de trabajo del sistema**

Una variable de flujo de trabajo que devuelve información de tiempo de ejecución del sistema, tal como el ID de instancia del flujo de trabajo, el usuario que ha iniciado el flujo de trabajo o la hora de inicio del flujo de trabajo.

**vínculo de tiempo de ejecución**

Enlace de grupo a grupo que utiliza una directiva o un parámetro, o ambos, para determinar qué puertos deben vincularse entre los grupos en tiempo de ejecución.

**vista virtual de los datos**

Una base de datos virtual definida por un servicio de datos SQL que puede consultar como si fuera una base de datos física.

**XMap**

Objeto de transformación del procesador de datos que asigna un documento XML de entrada a otro documento XML.

**XPath**

Lenguaje de consulta utilizado para seleccionar nodos en un documento XML y realizar cálculos.