



© JMA 2009. All rights reserved

1

INTRODUCCIÓN AL SQL SERVER

© JMA 2009. All rights reserved

3

Características

- Microsoft SQL Server es un sistema para la gestión de bases de datos producido por Microsoft basado en el modelo relacional cuya principal función es la de almacenar y consultar datos solicitados por otras aplicaciones, sin importar si están en el mismo equipo, si están conectadas a una red local o si están conectadas a través de internet.
- SQL Server es algo mas que un SGBDR, es una plataforma de datos de misión critica, ofreciendo desarrollo dinámico e inteligencia de negocio, y va mas allá de la estructura de datos relacional.

© JMA 2009. All rights reserved

4

Modelo relacional

- Basado en la teoría álgebra y la teoría de conjuntos
- Visión de la organización de los datos muy sencilla
 - Filas independientes
 - Filas únicas
 - Columnas independientes
 - Valores de las columnas unitarios
- Los usuarios no necesitan tener un conocimiento de la posición física de los datos, la navegación por la base de datos se realiza de forma automática
- Lenguaje de acceso a los datos muy sencillo (SQL)
- Sencillez del lenguaje, similar al lenguaje natural

© JMA 2009. All rights reserved

5

Soporte a otros modelos

- Jerárquico
 - La posibilidad de almacenar datos en formato XML le permite trabajar con datos semiestructurados siguiendo un modelo jerárquico.
 - El uso de esquema XSD impone unas restricciones tan fuertes como las del modelo relacionas.
 - Los datos se pueden consultar mediante XPath y Xquery.
- Objetos
 - La posibilidad de crear nuevos tipos de datos CLR permite el almacenamiento de objetos
- Documental
 - La búsqueda de texto completo permite el tratamiento de los campos de texto muy largo como si fueran documentos

© JMA 2009. All rights reserved

6

Características habituales

- Almacenamiento, extracción y actualización de datos
- Catálogo accesible por el usuario
- Servicios para mejorar la independencia de datos (metadatos)
- Soporte para la tramitación de datos (comunicaciones)
- Soporte multiusuario y servicios de autorización
- Servicios de control de concurrencia
- Soporte de transacciones
- Servicios de integridad
- Servicios de recuperación (y restauración)
- Servicios de utilidad: Importación/Exportación, Monitorización y Análisis de rendimientos, Automatización

© JMA 2009. All rights reserved

7

Características

- Fusiona el procesamiento transaccional en línea (OLTP) y el de procesamiento analítico en línea (OLAP)
- Facilidad de uso
 - SQL Server Management Studio
 - SQL Management Objects (SMO)
- Disponibilidad
 - Reflejo de base de datos (1 BD espejo)
 - Clústeres de conmutación por error
 - Instantáneas de base de datos
 - Recuperación rápida
 - Conexión del administrador no compartida
 - Operaciones en línea (Índices, Operaciones y Restauración)
 - Réplica (incluida entre iguales)

© JMA 2009. All rights reserved

8

Características

- Escalabilidad
 - Partición de tablas e índices
 - Aislamiento de instantáneas
 - Supervisor de réplicas
 - Compatible con Itanium 2 de 64 bits y x64
- Seguridad
 - Autenticación
 - Autorización
 - Cifrado nativo y Cifrado de base de datos transparente
 - SQL Server y Trustworthy Computing (Seguro por diseño, inicio e implementación)

© JMA 2009. All rights reserved

9

Características

- Rendimiento
 - Compresión de datos y de copia de seguridad
 - Partición de Tablas e Índices
 - Índices filtrados, Columnas dispersas
 - Regulador de recursos
 - Copias de seguridad reflejadas
- Administración
 - Gestión de la administración basada en directivas
 - Servidores de administración central
 - Recopilador de datos y Seguimiento de cambios
 - Auditoría

© JMA 2009. All rights reserved

10

Características

- Extensibilidad
 - Business Intelligence Development Studio
 - Integración CLR/.NET Framework
 - Tipos y agregados definidos por el usuario
 - Servicios Web y XML
 - Service Broker

© JMA 2009. All rights reserved

11

Inteligencia empresarial

- La plataforma de datos SQL Server incluye las siguientes herramientas:
 - Motor de base de datos
 - Analysis Services.
 - Reporting Services.
 - Integration Services
 - Servicios de réplica.
 - Master Data Services
 - Service Broker

© JMA 2009. All rights reserved

12

Motor de base de datos

- El Motor de base de datos es el servicio principal para almacenar, procesar y proteger los datos.
- El Motor de base de datos proporciona acceso controlado y procesamiento rápido de transacciones para cumplir los requisitos de las aplicaciones consumidoras de datos más exigentes de su empresa.
- El Motor de base de datos también proporciona una completa compatibilidad para mantener una gran disponibilidad.

© JMA 2009. All rights reserved

13

Analysis Services

- Datos multidimensionales:
 - Analysis Services admite OLAP al permitir al usuario diseñar, crear y administrar estructuras multidimensionales que contienen datos agregados de otros orígenes tales como bases de datos relacionales.
- Minería de datos:
 - Analysis Services permite al usuario diseñar, crear y visualizar modelos de minería de datos.
 - Estos modelos de minería de datos se pueden construir a partir de otros orígenes de datos empleando una amplia variedad de algoritmos de minería de datos estándar.

© JMA 2009. All rights reserved

14

Integration Services

- Integration Services es una plataforma para generar soluciones de integración de datos de alto rendimiento, lo que incluye paquetes que proporcionan procesamiento de extracción, transformación y carga (ETL) para almacenamiento de datos.

© JMA 2009. All rights reserved

15

Reporting Services

- Reporting Services ofrece funcionalidad empresarial de informes habilitados para Web con el fin de poder crear informes que extraigan contenido a partir de una variedad de orígenes de datos, publicar informes con distintos formatos y administrar centralmente la seguridad y las suscripciones.
- Una solución global para crear, administrar y proporcionar tanto informes tradicionales orientados al papel como informes interactivos basados en la Web.

© JMA 2009. All rights reserved

16

Replicación

- La replicación es un conjunto de tecnologías destinadas a la copia y distribución de datos y objetos de base de datos de una base de datos a otra, para luego sincronizar ambas bases de datos con el fin de mantener su coherencia.
- La replicación permite distribuir datos a diferentes ubicaciones y a usuarios remotos o móviles mediante redes de área local y de área extensa, conexiones de acceso telefónico, conexiones inalámbricas e Internet.

© JMA 2009. All rights reserved

17

Master Data Services

- Master Data Services es el origen de datos maestros para la organización.
- Mediante la integración de diferentes sistemas de análisis y operaciones con Master Data Services, puede asegurarse de que todas las aplicaciones de la organización dependen de una fuente de información central y precisa.
- Con Master Data Services, se crea un origen único de datos maestros y se mantiene un registro de los datos, que cambian en el tiempo y que se puede controlar.

© JMA 2009. All rights reserved

18

Service Broker

- Service Broker ayuda a los desarrolladores de software a crear aplicaciones de base de datos escalables y seguras.
- Esta nueva tecnología de Motor de base de datos proporciona una plataforma de comunicación basada en mensajes que permite a los componentes de aplicación independientes trabajar como un conjunto funcional.
- Service Broker incluye infraestructura para programación asíncrona que se puede utilizar para aplicaciones en una base de datos única o instancia única, y también para aplicaciones distribuidas.

© JMA 2009. All rights reserved

19

Herramientas de administración

- SQL Server incluye herramientas integradas de administración para administración y optimización avanzadas de bases de datos, así como también integración directa con herramientas tales como Microsoft Operations Manager (MOM) y Microsoft Systems Management Server (SMS).
- Los protocolos de acceso de datos estándar reducen drásticamente el tiempo que demanda integrar los datos en SQL Server con los sistemas existentes.
- Asimismo, el soporte del servicio Web nativo está incorporado en SQL Server para garantizar la interoperabilidad con otras aplicaciones y plataformas.

© JMA 2009. All rights reserved

20

Herramientas de desarrollo

- SQL Server ofrece herramientas integradas de desarrollo para el motor de base de datos, extracción, transformación y carga de datos, minería de datos, OLAP e informes que están directamente integrados con Microsoft Visual Studio para ofrecer capacidades de desarrollo de aplicación de extremo a extremo.
- Cada subsistema principal en SQL Server se entrega con su propio modelo de objeto y conjunto de interfaces del programa de aplicación (API) para ampliar el sistema de datos en cualquier dirección que sea específica de su negocio.

© JMA 2009. All rights reserved

21

Transact-SQL

- Es el lenguaje que se utiliza para administrar instancias del SQL Server Database Engine, para crear y administrar objetos de base de datos, y para insertar, recuperar, modificar y eliminar datos.
- Transact-SQL es una extensión (supraconjunto) del lenguaje definido en los estándares de SQL publicados por International Standards Organization (ISO) y American National Standards Institute (ANSI).
- Cuenta con:
 - Instrucciones SQL estándar ampliadas
 - Instrucciones procedurales (declaración de variables, expresiones, control, excepciones, ...)
 - Instrucciones para administración, monitorización y optimización
 - Tablas, vistas, procedimientos y constantes del sistema

© JMA 2009. All rights reserved

22

Estructura del Lenguaje SQL

- El lenguaje de Consultas Estructurado SQL se puede dividir en:
 - Lenguaje DQL (Lenguaje de Consulta de Datos)
 - Sentencia SELECT
 - Lenguaje DML (Lenguaje de Manipulación de Datos)
 - Sentencias INSERT, DELETE, UPDATE y MERGE
 - Lenguaje DDL (Lenguaje de Definición de Datos)
 - Sentencias CREATE, DROP y ALTER
 - Lenguaje DCL (Lenguaje de Control de Datos)
 - Sentencias GRANT, REVOKE y DENY
 - Lenguaje DVL (Lenguaje de Validación de Datos)
 - Sentencias COMMIT, ROLLBACK y SET TRANSACTION

© JMA 2009. All rights reserved

23

Competencias de Administración de bases de datos

- Administrar un sistema de bases de datos, interpretando su diseño y estructura, y realizando la adaptación del modelo a los requerimientos del sistema gestor de bases de datos (SGBD), así como la configuración y administración del mismo a nivel físico y lógico, a fin de asegurar la integridad, disponibilidad y confidencialidad de la información almacenada.
- Unidades de competencia:
 - Instalación e implementación de SGDB
 - Implementación e instalación de BD
 - Mantenimiento del SGDB y sus BD

© JMA 2009. All rights reserved

24

INSTALACIÓN

© JMA 2009. All rights reserved

25

Ediciones del SQL Server

- Express: Edición gratuita para producción; limitado a 4Gb de datos como máximo.
- Workgroup: licencia por procesador y estación. Para producción.
- Web: licencia por procesador. Para producción. Requiere versiones server del Windows.
- Standard : licencia por procesador y estación. Para producción. Requiere versiones server del Windows.
- Enterprise: licencia por procesador y estación. Para producción. Requiere versiones server del Windows.
- Evaluation: Edición Enterprise gratuita. Para de evaluación únicamente (180 días). No requiere versiones server del Windows.
- Developer: Edición Enterprise semi gratuita. Para desarrollo únicamente. No requiere versiones server del Windows.

© JMA 2009. All rights reserved

26

SQL Server Express Edition

- SQL Server Express es una base de datos gratuita, fácil de usar y sencilla de administrar. Facilita el desarrollo de aplicaciones controladas por datos que sean de gran capacidad, seguras en almacenamiento y rápidas de implementar.
- 1 procesador físico (4 núcleos), 1 GB de memoria y 10 GB de almacenamiento.
- Informes simples
- Réplica y Cliente SSB (Solo suscriptor)
- SQL Server Express Edition with Tools
 - Management Studio
- SQL Server Express Edition with Advanced Services
 - Agente SQL Server
 - Búsqueda de texto completo
- LocalDB es una versión ligera de Express que cuenta con todas las características de programación, pero se ejecuta en modo usuario.

© JMA 2009. All rights reserved

27

SQL Server Web Edition

- SQL Server Web es una opción con un costo total de propiedad bajo para los hosts de web y los sitios web que proporciona capacidades de administración y escalabilidad para propiedades web, tanto de pequeña como de gran escala.
- 4 CPU, RAM ilimitada (32 bits: 4Gb, 64Gb con AWE; 64 bits: 32 terabytes)
- Management Studio
- Importación/Exportación
- Publicación limitada de réplica (Solo suscriptor)
- Transmisión de registros de seguridad

© JMA 2009. All rights reserved

29

SQL Server Standard Edition

- Standard Edition es la plataforma de análisis y administración de datos para organizaciones pequeñas y medianas. Incluye la funcionalidad básica necesaria para soluciones de línea de negocio, almacenamiento de datos y comercio electrónico.
- Versiones nativas de 32 bits o 64 bits
- 1 a 4 CPU, RAM ilimitada (32 bits: 4Gb, 64Gb con AWE; 64 bits: 32 terabytes)
- Reflejo de bases de datos
- ETL básica
- Servidor OLAP estándar con Analysis Services
- Informes estándar con Reporting Services
- Minería de datos
- Réplica completa y Publicación SSB
- Admite Itanium 2 y x64

© JMA 2009. All rights reserved

30

SQL Server Enterprise Edition

- Enterprise Edition es la edición más completa de SQL Server y es idónea para las organizaciones más grandes y para los requisitos más complejos.
- Versiones nativas de 32 bits o 64 bits
- Escala y particionamiento ilimitados (32 bits: 4Gb, 64Gb con AWE)
- Particiones, Reflejo de bases de datos avanzados, operaciones completas en línea y paralelas e instantáneas de bases de datos
- Herramientas de análisis avanzadas con OLAP completo y minería de datos
- Creación avanzada de informes con informes especiales, de muy alto nivel y personalizados
- ETL avanzada con enrutamiento de datos y capacidades de transformación

© JMA 2009. All rights reserved

31

Otras ediciones

- SQL Server Evaluation Edition
 - SQL Server Enterprise Edition, pero se autoriza su uso como sistema de evaluación, no como servidor de producción.
- SQL Server Developer Edition
 - SQL Server Enterprise Edition, pero se autoriza su uso como sistema de desarrollo y prueba, no como servidor de producción.
- SQL Server Compact Edition 3.1 y 3.5 SP1 (Mobile Edition)
 - Microsoft SQL Server Compact Edition ofrece las funciones básicas de base de datos relacional en un formato compacto, ideal para integrarlo en aplicaciones para móviles o desktop, como por ejemplo la nueva generación de aplicaciones dinámicas de conexión ocasional. Monousuario, con API de ejecución sin motor de base de datos.

© JMA 2009. All rights reserved

32

Requerimientos del Sistema

- CPU
 - **32 bits** 500 MHz o superior (min 1 GHz)
 - **64 bits** IA64 como mínimo: procesador Itanium de 733 MHz o más veloz o X64 superior como mínimo: AMD Opteron de 1 GHz o más veloz, AMD Athlon 64, Intel Xeon con soporte Intel EM64T, Intel Pentium IV con soporte EM64T
- Memoria
 - **32 bits** Mínimo 512 Mb (1 Gb o mas, max 4Gb: SO)
 - **64 bits** Mínimo 512 Mb (1 Gb o mas, 32 terabytes como máximo: SO)
- Disco Duro
 - 350 MB de espacio disponible en el disco duro para la instalación completa
 - 390 MB para bases de datos de muestra

© JMA 2009. All rights reserved

33

Ediciones de SQL Server

	Express	Standard	Workgroup	Enterprise	Developer
Windows Server 2003 SP1 (all PC editions)	✓	✓	✓	✓	✓
Windows XP SP2 (Home Edition)	✓				✓
Windows XP SP2 (Professional, Media and Tablet Editions)	✓	✓	✓		✓
Windows 2000 Professional SP4	✓	✓	✓		✓
Windows 2000 Server SP4 (all editions)	✓	✓	✓	✓	✓

© JMA 2009. All rights reserved

34

Instancias

- SQL Server denomina instancias a cada instalación concurrente en el mismo servidor. El concepto de instancia está disociado del de base de datos.
- SQL Server se puede instalar hasta 16 veces en el mismo servidor (50 veces en la Enterprise).
- Las diferentes instancias no comparten nada, compiten por los recursos entre si.
- La ejecución de varias instancias de SQL Server en un único equipo es necesaria cuando:
 - Se alcanza el número máximo de bases soportado por la instancia (32.000), como es el caso de los ISP.
 - Debe admitir diferentes sistemas que tienen que estar aislados de manera segura el uno del otro; por ejemplo, cuando pertenecen a clientes diferentes.
 - Necesita admitir varias bases de datos de prueba y desarrollo con configuraciones y entornos diferentes.
 - Se tiene que ejecutar varias aplicaciones en un escritorio y cada aplicación instala una instancia diferente de SQL Server Express.

© JMA 2009. All rights reserved

35

Reglas para nombres de instancias

- Un nombre de instancia no distingue entre mayúsculas y minúsculas.
- El nombre de instancia no puede ser Default o MSSQLServer.
- Los nombres de instancias deben seguir las reglas de los identificadores de SQL Server y no pueden ser palabras reservadas.
- Los nombres de instancias están limitados a 16 caracteres.
- El primer carácter del nombre de la instancia debe ser una letra, &, _ o #. Las letras aceptables son las que define el estándar Unicode 2.0, que incluye los caracteres latinos a-z y A-Z, además de caracteres de letras de otros idiomas.
- Los caracteres subsiguientes pueden ser: Letras, Números, \$, # o _.
- Los caracteres especiales o los espacios incrustados no se permiten en los nombres de instancia. Tampoco la barra diagonal inversa (\), la coma (,), los dos puntos (:) o el carácter de arroba (@).

© JMA 2009. All rights reserved

36

Nombres de Servidores

- La instancia predeterminada no requiere nombre (para desambiguación internamente se la denomina MSSQLServer).
- Dado que lo habitual es tener una sola instancia por servidor, suelen ser sinónimos instancia y servidor.
- Referencia a las instancias:
 - Instancia Predeterminada: servername
 - Otras instancias: servername\instancename
- Alternativas para el servidor local:
 - .\instancename
 - (local)\instancename
 - localhost\instancename
- En caso de no usar el puerto predeterminado 1433:
 - servername\instancename,12345

© JMA 2009. All rights reserved

37

Instalación

- Configurar cuentas de servicio de Windows
- Preparar ubicaciones de archivos para las instancias predeterminadas y con nombre de SQL Server
- Configurar Protocolos de red y bibliotecas de red
- Configurar Firewall de Windows para permitir el acceso a SQL Server
- Configurar el arranque de los servicios del SQL Server.
- Ajustar la superficie de la instancia

© JMA 2009. All rights reserved

38

Servicios de SQL Server

- Servicios de bases de datos de SQL Server: servicio del Database Engine (Motor de base de datos) relacional de SQL Server.
- Agente SQL Server: ejecuta trabajos, supervisa SQL Server, activa alertas y habilita la automatización de algunas tareas administrativas.
- Analysis Services: proporciona funciones de procesamiento analítico en línea (OLAP) y minería de datos para aplicaciones de Business Intelligence.
- Reporting Services: administra, ejecuta, crea, programa y envía informes.
- Búsqueda de texto completo: crea rápidamente índices de texto completo del contenido y de las propiedades de los datos estructurados y semiestructurados para permitir el filtrado de documentos y la separación de palabras en SQL Server.
- Integration Services: proporciona compatibilidad de administración para el almacenamiento y la ejecución de paquetes de Integration Services.
- Explorador de SQL Server: servicio de resolución de nombres que proporciona información de conexión de SQL Server a los equipos cliente.
- Servicio auxiliar de Active Directory de SQL Server: publica y administra los servicios de SQL Server en Active Directory.
- Objeto de escritura de SQL: permite que las aplicaciones de copias de seguridad y restauración funcionen en el marco del Servicio de instantáneas de volumen (VSS).

© JMA 2009. All rights reserved

39

SQL Server Management Studio

- SQL Server Management Studio es un entorno integrado para obtener acceso a todos los componentes de SQL Server, configurarlos, administrarlos y desarrollarlos.
- SQL Server Management Studio combina un amplio grupo de herramientas gráficas con una serie de editores de script enriquecidos para ofrecer acceso a SQL Server a desarrolladores y administradores de todos los niveles de especialización.
- SQL Server Management Studio combina las características del Administrador corporativo, el Analizador de consultas y Analysis Manager, herramientas incluidas en versiones anteriores de SQL Server, en un único entorno.
- Además, SQL Server Management Studio funciona con todos los componentes de SQL Server, como Reporting Services, Integration Services y SQL Server Compact 3.5 SP2.
- Los administradores de bases de datos obtienen una única herramienta completa que combina herramientas gráficas fáciles de usar con funcionalidad de scripting enriquecida.

© JMA 2009. All rights reserved

40

Herramientas de datos de SQL Server

- SQL Server Data Tools (SSDT) transforma el desarrollo de bases de datos al presentar un modelo ubicuo y declarativo que abarca todas las fases del desarrollo en Visual Studio. Las funcionalidades de diseño de Transact-SQL de SSDT se pueden usar para compilar, depurar, mantener y refactorizar bases de datos. Puede trabajar con un proyecto de base de datos o directamente con una instancia de una base de datos conectada ubicada dentro o fuera de las instalaciones locales.
- Los desarrolladores pueden usar las conocidas herramientas de navegación por el código, IntelliSense, compatibilidad con lenguajes similar al disponible para C# y Visual Basic, validación específica de la plataforma, depuración y edición declarativa en el editor de Transact-SQL. SSDT también proporciona un Diseñador de tablas visual para crear y editar tablas en proyectos de base de datos o en instancias de bases de datos conectadas. Se puede usar el control de versiones para todos los archivos.
- Dispone de proyectos para:
 - Bases de datos relacionales y Base de datos SQL de Azure
 - Modelos de Analysis Services
 - Informes de Reporting Services
 - paquetes de Integration Services

© JMA 2009. All rights reserved

41

Herramientas

- Herramientas de configuración
 - Administrador de configuración de SQL Server
 - Configuración de Reporting Services
 - Configuración de superficie de SQL Server
 - Informes de uso y errores de SQL Server
 - Símbolo del sistema de Notification Services
- Herramientas de rendimiento
 - Asistente para la optimización de motor de base de datos
 - SQL Server Profiler
- Analysis Services
 - Asistente para la implementación
- Documentación y tutoriales
 - Libros en pantalla de SQL Server
- Utilidades del símbolo del sistema

© JMA 2009. All rights reserved

42

Instalación de MySQL

- Descargar e instalar:
 - <https://mariadb.org/download/>
- Incluir en la sección [mysqld] de %MYSQL_ROOT%/data/my.ini
 - default_time_zone='+01:00'
- Descargar bases de datos de ejemplos:
 - <https://dev.mysql.com/doc/index-other.html>
- Instalar bases de datos de ejemplos:
 - mysql -u root -p < employees.sql
 - mysql -u root -p < sakila-schema.sql
 - mysql -u root -p < sakila-data.sql
- Instalar MySQL Connectors
 - <https://dev.mysql.com/downloads/connector/>

© JMA 2009. All rights reserved

43

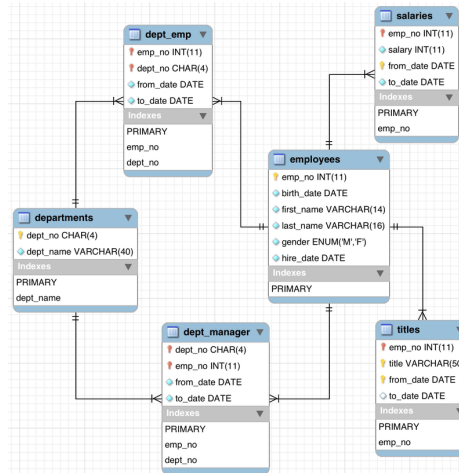
Conexión

- Proveedores .Net\MySQL Data Provider
 - Connection:
 - Database: sakila
 - Port: 3306
 - Server: localhost
 - Security:
 - Password: ****
 - User Id: root
 - ConnectionString: server=localhost;user id=root;password:****;port=3306;database=sakila;
- ODBC:
 - Driver={MySQL ODBC 8.0 Unicode Driver};Server=127.0.0.1;Database=sakila;UID=root;PWD=*****

© JMA 2009. All rights reserved

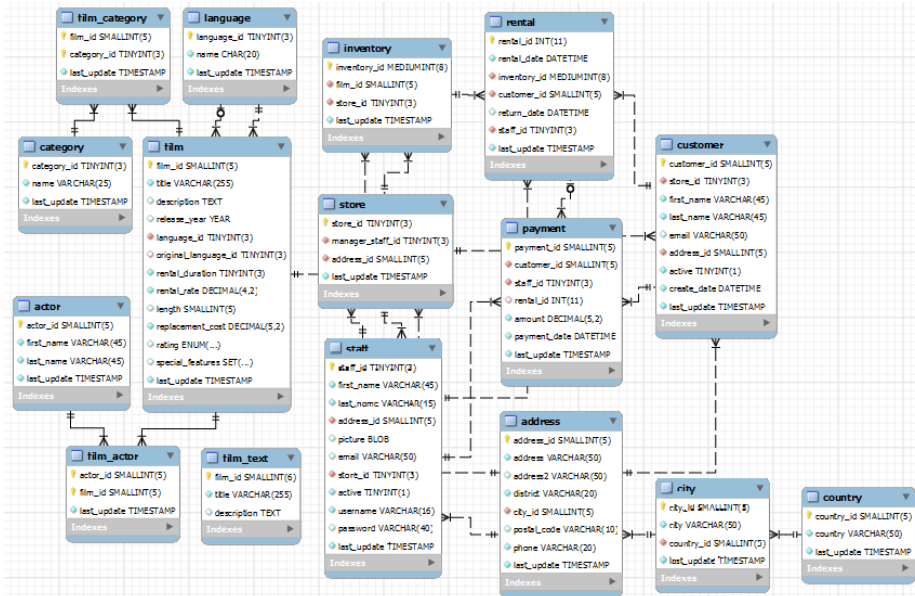
44

Modelos de datos



© JMA 2009. All rights reserved

45



© JMA 2009. All rights reserved

46

ARQUITECTURA LÓGICA

© JMA 2009. All rights reserved

47

Bases de Datos

- Bases de Datos de sistema
 - Almacenan información necesaria para el funcionamiento del servidor SQL Server
 - SQL Server utiliza estas BBDD para manejar y gestionar el sistema
 - Bases de datos de usuarios
 - Bases de Datos donde se implementan los modelos de datos de los usuarios (OLTP)
 - Almacén de datos
 - Organizar grandes cantidades de datos estables para facilitar el análisis y la recuperación con OLAP
-

© JMA 2009. All rights reserved

48

Bases de Datos de sistema

- master
 - Controla las BBDD de los usuarios
 - Guarda información del servidor de SQL Server: Inicios de sesión, Variables de entorno configurables, Mensajes de error del sistema, ...
- model
 - Suministra un patrón para la creación de nuevas Bases de Datos de usuario
- tempdb
 - Proporciona un área de almacenamiento para tablas temporales y otras necesidades de almacenamiento temporal del sistema. Se vacía automáticamente al arrancar la instancia.
- msdb
 - Almacena la información relativa a la planificación de trabajos e información histórica de las BBDD

© JMA 2009. All rights reserved

49

Bases de Datos de sistema

- resource (Mssqlsystemresource.mdf)
 - Base de datos de sólo lectura que contiene objetos del sistema que se incluyen con SQL Server. Los objetos del sistema persisten físicamente en la base de datos Resource, pero aparecen lógicamente en el esquema sys de cada base de datos. No se muestra en el SSMS.
- distribution
 - Se crea cuando se configura SQL Server como integrante de un modelo de replicación
 - Almacena datos históricos y de transacciones usadas en el modelo de replicación
- reportserver y reportservertempdb
 - Repositorios del Reporting Service

© JMA 2009. All rights reserved

50

Objetos de la BBDD

- Diagramas de Bases de Datos
 - Herramienta visual que permite diseñar y ver una base de datos completa o parcialmente
- Tablas
 - Colección de datos organizados en filas y columnas
- Índices
 - Estructura de datos que proporciona una manera rápida de acceder a la información
- Restricciones
 - Sirven para definir reglas relativas a los valores permitidos en las columnas
 - Es el mecanismo utilizado para reforzar la integridad referencial en las Bases de Datos que SQL Server exige automáticamente
 - Las restricciones establecen los valores aceptados por las columnas: nulos, rangos de valores, duplicados, claves principales o ajenas.

© JMA 2009. All rights reserved

51

Objetos de la BBDD

- Estadísticas
 - Índices de distribución de valores que utiliza el optimizador de consultas para determinar el plan de consultas óptimo.
- Vistas
 - Proporciona una forma de ver los datos provenientes de una o más tablas
 - Sirven para facilitar el acceso a los datos, y garantizar la seguridad de los mismos
- Sinónimos
 - Es un nombre alternativo para un objeto de ámbito de esquema
- Service Broker
 - Tipos y contratos, colas, servicios, rutas, enlaces.
- Almacenamiento
 - Catálogos e índices de texto
 - Permiten las búsquedas de texto complejas en datos de cadenas de caracteres.
 - Esquemas de partición
 - Funciones de partición

© JMA 2009. All rights reserved

52

Objetos para programación

- **Procedimientos Almacenados**
 - Conjunto de sentencias de Transact-SQL precompiladas con nombre que se ejecutan como una unidad
 - Mejoran el rendimiento del sistema
 - Se pueden implementar en .NET CLR
- **Funciones (SQL definidas por el usuario)**
 - Funciones definidas por el usuario
 - Conjunto de sentencias de Transact-SQL o CLR que suelen utilizarse para encapsular una transformación lógica
 - Pueden devolver un valor escalar o una tabla
- **Desencadenadores**
 - También llamados triggers o disparadores
 - Procedimiento almacenado de ejecución automática ante una operación DML o DDL
 - Se crean sobre una tabla
 - Ayudan a mantener la integridad

© JMA 2009. All rights reserved

53

Objetos para programación

- **Ensamblados**
 - Son archivos DLL que se utilizan en una instancia de SQL Server para implementar funciones, procedimientos almacenados, desencadenadores, agregados definidos por el usuario y tipos definidos por el usuario que están escritos en uno de los lenguajes de código administrado que se alojan en Microsoft .NET Framework.
- **Tipos de Datos**
 - Definen los valores permitidos para una columna o variable
 - SQL Server suministra un conjunto de datos definidos por el sistema, aunque el usuario puede crearse tipos de datos propios
- **Reglas**
 - Contiene información que define los valores permitidos que serán almacenados en una columna o para un tipo de dato determinado
- **Valores predeterminados**
 - Valor por defecto para cuando no se suministre valor en una columna determinada

© JMA 2009. All rights reserved

54

Objetos para la seguridad

- **Inicios de sesión**
 - Cuentas que controlan el acceso al sistema SQL Server.
- **Usuarios**
 - Identificador que asocia un inicio de sesión a un usuario dentro de una base de datos para otorgar los permisos.
- **Esquemas**
 - Es una colección de entidades de base de datos que forma un solo espacio de nombres
- **Funciones (Roles)**
 - Agrupan los permisos necesarios para los diferentes tipos de usuarios.
- **Claves simétricas y asimétricas**
 - Encriptación y desencriptación, autenticaciones, ..
- **Certificados**
 - Autenticación ante terceros (comunicaciones seguras) y cifrado
- **Credenciales**
 - Una credencial es un registro que contiene la información de autenticación necesaria para conectarse a un recurso fuera de SQL Server. La mayoría de las credenciales constan de un inicio de sesión de Windows y una contraseña.

© JMA 2009. All rights reserved

55

Otros objetos

- **Propiedades extendidas**
 - Propiedades definidas por los usuarios para almacenar información adicional o específica sobre los objetos de una base de datos.
- **Intercalaciones**
 - Controlan el juego de caracteres utilizado en el almacenamiento físico de las cadenas, especificado las combinaciones de bits que representa cada carácter y las reglas por las que los caracteres se ordenan y comparan.

© JMA 2009. All rights reserved

56

MODELO LÓGICO

© JMA 2009. All rights reserved

57

Nomenclatura

- **Identificadores estándar**
 - El primer carácter debe ser un carácter alfabético
 - De uno a 128 caracteres.
 - Otros caracteres pueden incluir letras, números o símbolos: @, \$, # o _.
 - Los identificadores que comienzan con un símbolo tienen usos especiales: @, # o ##.
- **Identificadores delimitados**
 - Se utilizan cuando los nombres contienen espacios incrustados
 - Se utilizan cuando partes de los nombres incluyen palabras reservadas
 - Deben encerrarse entre corchetes ([])
 - o dobles comillas (" ") si SET QUOTED_IDENTIFIER ON
- **Nombre completos:**
 - Instancia.BaseDeDatos.Eschema.Objeto.SubElemento

© JMA 2009. All rights reserved

58

Bases de datos

- Creación de una BBDD
 - Mediante el SQL Server Management Studio
 - Mediante Transact-SQL
 - Con la instrucción CREATE DATABASE
 - Existe una plantilla para la creación en el SSMS
 - Se puede generar el script de creación en el SSMS, y ejecutarlo posteriormente como sentencia SQL
- Las Bases de Datos heredan las características de la Base de Datos model
 - Algunas propiedades pueden ser determinadas durante la creación de la BBDD
- Se pueden modificar posteriormente
 - Mediante el SSMS
 - Mediante Transact-SQL
 - Exec sp_dboption nombre_bd, opcion, TRUE | FALSE
 - Mediante el comando ALTER DATABASE
- Para ver las propiedades actuales de una Base de Datos
 - Exec sp_dboption nombre_bd

© JMA 2009. All rights reserved

59

Valores predeterminados

- Los valores predeterminados especifican el valor por defecto de una columna si no se especifica un valor al insertar las filas
- Para crear un nuevo valor predeterminado es necesario asignar:
 - Nombre del valor predeterminado
 - Valor que puede ser cualquier expresión cuyo resultado sea una constante: constantes, funciones integradas o expresiones matemáticas

Hoy: getdate()

Mañana: DATEADD(day, 1, getdate())

Cero: 0

© JMA 2009. All rights reserved

60

Reglas

- Las reglas contienen la condición que define los valores permitidos para una columna o para un tipo de dato determinado (por compatibilidad con versiones anteriores)
- Para crear una nueva regla es necesario asignar:
 - Nombre de la regla
 - Expresión condicional, válida en una cláusula WHERE, que incluye una variable local que representa el valor a comprobar (@valor)
- Una regla se puede utilizar en tantos objetos como sea necesario.
- Se vinculan mediante Transact-SQL

```
CREATE RULE id_chk AS @id BETWEEN 0 and 10000
sp_bindrule id_chk, 'cust_sample.cust_id'
```

© JMA 2009. All rights reserved

61

Números de secuencia (2012)

- Una secuencia es un objeto enlazado a un esquema definido por el usuario que genera una secuencia de valores numéricos según la especificación con la que se creó la secuencia.
- A diferencia de los valores de columnas de identidad que se generan cuando se insertan filas, una aplicación puede obtener el número de secuencia siguiente sin insertar la fila llamando a la función NEXT VALUE FOR .
- Se asigna el número de secuencia cuando se llama a NEXT VALUE FOR aun cuando el número nunca se inserta en una tabla.
- La función NEXT VALUE FOR se puede utilizar como valor predeterminado para una columna en una definición de tabla.
- Se utiliza sp_sequence_get_range para obtener un intervalo de varios números de secuencia enseguida.

© JMA 2009. All rights reserved

62

Tipos de datos

- Numéricos exactos
- Numéricos aproximados
- Fecha y hora
- Cadenas de caracteres
- Cadenas de caracteres Unicode
- Cadenas binarias
- Otros tipos de datos

© JMA 2009. All rights reserved

63

Tipos de datos

Tipo	Bytes	Descripción
bigint	8	Datos enteros (números enteros) comprendidos entre -2^{63} (-9223372036854775808) y $2^{63}-1$ (9223372036854775807).
int	4	Datos enteros (números enteros) comprendidos entre -2^{31} (-2.147.483.648) y $2^{31}-1$ (2.147.483.647).
smallint	2	Datos enteros comprendidos entre -2^{15} (-32.768) y $2^{15}-1$ (32.767).
tinyint	1	Datos enteros comprendidos 0 y 255.
bit	1*	Datos enteros con valor 1 ó 0.
decimal	5, 9, 13, 17	Datos de precisión y escala numérica fijas con 38 dígitos significativos (escala + precisión). [vardecimal con SP2]
numeric		Funcionalmente equivalente a decimal.

© JMA 2009. All rights reserved

64

Tipos de datos

Tipo	Bytes	Descripción
money	8	Valores de moneda comprendidos entre -922.337.203.685.477,5808 y +922.337.203.685.477,5807, con una precisión de una diezmilésima de la unidad monetaria. (escala 4)
smallmoney	4	Valores de moneda comprendidos entre -214.748,3648 y +214.748,3647, con una precisión de una diezmilésima de la unidad monetaria.
Numéricos aproximados		
float	4, 8	Números con precisión de coma flotante comprendidos entre -1,79E + 308 y 1,79E + 308.
real	4	Números con precisión de coma flotante comprendidos entre -3,40E + 38 y 3,40E + 38. float(24)

© JMA 2009. All rights reserved

65

Tipos de datos

Tipo	Bytes	Descripción
datetime	8	Datos de fecha y hora comprendidos entre el 1 de enero de 1753 y el 31 de diciembre de 9999, con una precisión de 3,33 milisegundos.
smalldatetime	4	Datos de fecha y hora comprendidos entre el 1 de enero de 1900 y el 6 de junio de 2079, con una precisión de un minuto.
datetime2	6,7,8	Datos de fecha y hora comprendidos entre el Del 1 de enero del año 1 después de Cristo al 31 de diciembre de 9999. De 00:00:00 a 23:59:59.9999999
date	3	De 0001-01-01 a 9999-12-31
time	3 a 5	De 00:00:00.0000000 a 23:59:59.9999999
datetimeoffset	8..10	De 0001-01-01 00:00:00.0000000 a 9999-12-31 23:59:59.9999999 (en UTC)

© JMA 2009. All rights reserved

66

Tipos de datos

Tipo	Bytes	Descripción
char	1	Datos de caracteres ANSI de longitud fija con una longitud máxima de 8.000 caracteres.
varchar	1...	Datos ANSI de longitud variable con un máximo de 8.000 caracteres.
text varchar(max)	8K...	Datos ANSI de longitud variable con una longitud máxima de 231 - 1 (2.147.483.647) caracteres.
nchar	2	Datos Unicode de longitud variable con una longitud máxima de 4.000 caracteres.
nvarchar	2...	Datos Unicode de longitud variable con una longitud máxima de 4.000 caracteres. sysname es el tipo de datos suministrado por el sistema y definido por el usuario que es funcionalmente equivalente a nvarchar(128) y que se utiliza para hacer referencia a nombres de objetos de bases de datos.
ntext nvarchar(max)	8K...	Datos Unicode de longitud variable con una longitud máxima de 230 - 1 (1.073.741.823) caracteres.

© JMA 2009. All rights reserved

67

Tipos de datos

Tipo	Bytes	Descripción
binary	1...	Datos binarios de longitud fija con una longitud máxima de 8.000 bytes.
varbinary	1...	Datos binarios de longitud variable con una longitud máxima de 8.000 bytes.
image varbinary(max)	8K...	Datos binarios de longitud variable con una longitud máxima de 231 - 1 (2.147.483.647) bytes.
xml	8K...	Datos XML de longitud variable con una longitud máxima de 2 Gb (2.147.483.647) caracteres.
sql_variant	1...	Un tipo de datos que almacena valores de varios tipos de datos aceptados en SQL Server, excepto text, ntext, timestamp y sql_variant.
timestamp rowversion	8	Un número único para toda la base de datos que se actualiza cada vez que se actualiza una fila.
uniqueidentifier	16	Un identificador exclusivo global (GUID).

© JMA 2009. All rights reserved

68

Tipos de datos

Tipo	Bytes	Descripción
hierarchyid	1..5	Dato utilizado para representar la posición en una jerarquía
geography	2...	Datos elípticos (globo), como las coordenadas de latitud y longitud del sistema GPS. Pueden contener uno o mas puntos.
geometry	2...	Datos en un sistema de coordenadas euclídeo (plano). Pueden contener uno o mas puntos.
Tipos programáticos		
cursor		Una referencia a un cursor.
table		Un tipo de datos especial que se utiliza para almacenar un conjunto de resultados para un proceso posterior.

© JMA 2009. All rights reserved

69

Tipos de Datos definidos por el usuario

- Los tipos de datos definidos por el usuario están basados en un tipo integrado y permiten establecer su longitud, si acepta nulos, qué regla restringe sus valores y cuál es su valor por defecto.
- Para crear un nuevo tipo es necesario asignar:
 - Nombre del nuevo tipo
 - Tipo de datos del sistema en el que se basa el nuevo tipo de datos
 - Longitud (si el tipo de datos lo permite)
 - Aceptación de valores NULL
- Opcionalmente se le pueden asignar:
 - Regla que definen los valores permitidos
 - Valor predeterminado
- Una vez creados solo se puede modificar la regla y el valor predeterminado.

© JMA 2009. All rights reserved

70

Tipos de Tablas

- Definición:
 - Nombre y ubicación
 - Columnas
 - Claves y restricciones
- Tipos de Tablas
 - Tablas del sistema
 - Tablas de datos
 - Tablas temporales
 - Locales (#) y Globales (##)
 - Tablas anchas (1.024 → 30.000 columnas)
 - Tablas internas (variables)

© JMA 2009. All rights reserved

71

Definición de las columnas

- Nombre de columna
- Tipo de datos
- Longitud de los datos
 - Sólo puede cambiar la longitud de los tipo de datos sin longitud fija
 - Para los tipos decimal y numéricos
 - Precisión: Muestra el número máximo de dígitos
 - Escala: Muestra el número máximo de dígitos decimales
- Aceptación de valores NULL (Restricción)
- Descripción
- Valor predeterminado
- Intercalación (sólo para columnas de texto)

© JMA 2009. All rights reserved

72

Columnas especiales

- Columnas de numeración automática (Autonúmericas)
 - Identidad, una sola columna por tabla
 - Inicialización de identidad: valor de inicial de una columna de identidad
 - Incremento de identidad: valor en que se incrementa la columna de identidad
 - Vista `sys.identity_columns` columna `last_value`
- Columnas de identificadores exclusivos globales (GUID)
 - Es RowGuid: Muestra si SQL Server utiliza la columna como una columna ROWGUID. Este valor se puede establecer como Sí sólo para una columna de identidad
 - Valor predeterminado: `newid()`
- Columnas calculadas
 - Fórmula: Muestra la fórmula para una columna calculada
 - Ignora tipo y longitud

© JMA 2009. All rights reserved

73

Restricciones

- Sirven para definir reglas relativas a los valores permitidos en las columnas
- Es el mecanismo utilizado para reforzar la integridad en las Bases de Datos que SQL Server exige automáticamente
- Las restricciones establecen los valores aceptados por las columnas: nulos, rangos de valores, duplicados, claves principales o ajenas.
- Se dividen en restricciones de tabla y de columna.
- Admite cinco clases de restricciones.
 - NOT NULL especifica que la columna no acepta valores NULL (Obligatorio).
 - Las restricciones CHECK exigen la integridad del dominio mediante la limitación de los valores que se pueden asignar a una columna.
 - Las restricciones UNIQUE exigen la unicidad de los valores de un conjunto de columnas (conlleva un índice).
 - Las restricciones PRIMARY KEY identifican la columna o el conjunto de columnas cuyos valores identifican de forma unívoca cada una de las filas de una tabla (NOT NULL + UNIQUE) (conlleva un índice).
 - Las restricciones FOREIGN KEY identifican las relaciones entre las tablas (REFERENCES).

© JMA 2009. All rights reserved

74

Relaciones

- Se definen a nivel de tabla.
- Puede crear una relación entre las tablas en un diagrama de base de datos para mostrar cómo se vinculan las columnas de una tabla a las columnas de otra tabla.
- Permiten exigir la integridad referencial:
 - No se puede especificar un valor en la columna de clave externa de la tabla relacionada si ese valor no existe en la clave principal de la tabla relacionada (o NULL si la columna lo acepta).
 - No se puede eliminar una fila de una tabla de claves principales si existen filas que coinciden con ella en una tabla relacionada.
 - No se puede cambiar un valor de clave principal en la tabla de clave principal si esa fila tiene filas relacionadas.
- Pueden automatizar el proceso en cascada.

© JMA 2009. All rights reserved

75

Seguimiento de cambios de datos

- A partir del SQL Server 2008 se proporciona dos características que realizan el seguimiento de los cambios en los datos de una base de datos: captura de datos modificados y seguimiento de cambios.
- Estas características permiten a las aplicaciones determinar los cambios de DML (operaciones de inserción, actualización y eliminación) que se realizaron en las tablas de usuario de una base de datos.
- La captura de datos modificados y el seguimiento de cambios pueden habilitarse en la misma base de datos; no se requiere ninguna consideración especial.

© JMA 2009. All rights reserved

76

Diagramas de Bases de Datos

- Herramienta visual que permite diseñar y ver una base de datos completa o parcialmente
- Utiliza como repositorio la tabla de sistema: `dbo.sysdiagrams`.

© JMA 2009. All rights reserved

77

Vistas

- Una vista es una tabla virtual cuyo contenido está definido por una consulta.
 - Combinar columnas de varias tablas de forma que parezcan una sola tabla.
 - Agregar información agrupada o calculada en lugar de presentar los detalles.
 - Unir información de varios resultados
 - Restringir el acceso del usuario a filas concretas de una tabla.
 - Restringir el acceso del usuario a columnas específicas.
- Tipos:
 - Vistas estándar
 - Vistas indexadas o materializada
 - Vistas con particiones o distribuidas
 - Vistas parametrizadas (funciones de tipo tabla)

© JMA 2009. All rights reserved

78

Vistas

- Se puede modificar los datos mediante una vista.
 - Son actualizables (UPDATE, DELETE o INSERT) mientras la modificación afecte sólo a una de las tablas base.
 - Las columnas a modificar en la vista deben hacer referencia directa a los datos subyacentes de las columnas de la tabla (no puede ser resultado de una función de agregado o cálculo).
 - Deben aparecer todas las columnas requeridas que no dispongan de valor predeterminado.
 - Las columnas que se modifican no pueden verse afectadas por cláusulas GROUP BY, HAVING o DISTINCT.
 - Permite el uso de desencadenadores INSTEAD OF
 - La cláusula WITH CHECK OPTION exige que todas las instrucciones de modificación de datos ejecutadas en la vista se ajusten a los criterios especificados en la instrucción SELECT que define la vista.

© JMA 2009. All rights reserved

79

Vistas

- Una vista distribuida combina los datos procedentes de un conjunto de tablas miembro en uno o más servidores, y hace que los datos parezcan proceder todos de una sola tabla que combinado con desencadenadores INSTEAD OF simulan tablas particionadas o distribuidas.
- La cláusula WITH ENCRYPTION convertirá el texto original en un formato ofuscado.
- Las vistas, las tablas o las funciones que participan en una vista creada con la cláusula SCHEMABINDING no se pueden quitar, a menos que se quite o cambie esa vista de forma que deje de tener un enlace de esquema.
- Además, las instrucciones ALTER TABLE sobre tablas que participan en vistas que tienen enlaces de esquemas provocarán un error si estas instrucciones afectan a la definición de la vista.

© JMA 2009. All rights reserved

80

Índices

- Compuestos por una o varias columnas ordenadas ascendente o descendentemente
- Guardan clave indexación, clave en índice agrupado o marcador fila real y columnas incluidas.
- Se pueden crear sobre Tablas, Vistas y columnas calculadas.
- Los dos tipos:
 - Agrupado
 - No agrupado
- Un índice único garantiza que la columna indexada no contiene valores no nulos duplicados.
- Índices y Restricciones:
 - Las restricciones PRIMARY KEY crean índices agrupados de forma automática de tipo UNIQUE.
 - Las restricciones UNIQUE crean índices UNIQUE de forma automática.
 - Las restricciones FOREIGN KEY no crean índices, pero suelen ser convenientes.
- Los no agrupados permiten columnas INCLUIDAS

© JMA 2009. All rights reserved

81

Índices

- El factor de relleno determina la densidad de ocupación de páginas por parte del índice.
- El número de índices afecta al rendimiento y ocupación de la base de datos.
- El Asistente para optimización de índices suministra recomendaciones en función a la monitorización del sistema.
- Limitaciones
 - Sólo el propietario de la tabla puede crear índices en la misma tabla.
 - Un índice agrupado por tabla.
 - 249 índices como máximo (incluidos los creados por restricciones PRIMARY KEY o UNIQUE).
 - 16 columnas como máximo.
 - El tamaño máximo de la clave de indexación es 900 bytes.
- Índices especiales:
 - Texto completo, XML y Espaciales

© JMA 2009. All rights reserved

82

Estadísticas

- SQL Server (2005...) permite crear información estadística acerca de la distribución de valores en una columna.
- El optimizador de consultas utiliza esta información estadística para determinar el plan de consultas óptimo realizando una estimación del costo de usar un índice para evaluar la consulta.
- Se almacenan con la misma estructura que los índices, sustituyendo el localizador por un contador.
- Se pueden crear automáticamente o manualmente. Es necesario actualizarlas periódicamente.

© JMA 2009. All rights reserved

83

Procedimientos almacenados

- Un procedimiento almacenado es un grupo de instrucciones Transact-SQL compiladas en un único plan de ejecución, se almacena en `sys.sql_modules` y en cache.
- Los procedimientos almacenados:
 - Permiten una programación modular.
 - Permiten una ejecución más rápida.
 - Pueden reducir el tráfico de red.
 - Pueden utilizarse como mecanismo de seguridad.
 - Un máximo de 2.100 parámetros.

© JMA 2009. All rights reserved

84

Procedimientos almacenados

- Tipos:
 - Procedimientos almacenados del sistema (sp_)
 - Procedimientos almacenados definidos por el usuario
 - Transact-SQL y CLR
 - Procedimientos almacenados extendidos (xp_)
 - Procedimientos almacenados temporales
 - Locales (#) y Globales (##)
- Los procedimientos almacenados extendidos permiten crear rutinas externas propias, en un lenguaje de programación como C.
- Se pueden crear procedimientos almacenados temporales (agregando el prefijo # o ##) que dejan de existir cuando se cierra SQL Server.
- Los procedimientos de inicio deben estar en la base de datos master y no pueden contener parámetros de entrada (INPUT) ni salida (OUTPUT). La ejecución de los procedimientos almacenados se inicia cuando se recupera la base de datos master en el inicio.
 - sp_procoption 'procedure', 'startup', 'on'

© JMA 2009. All rights reserved

85

Procedimientos almacenados

- Devuelven datos de cuatro formas distintas:
 - Parámetros de salida: datos o cursores.
 - Códigos de retorno, que siempre son un valor entero.
 - Un conjunto de resultados por una o varias instrucciones SELECT.
 - Un cursor global al que se puede hacer referencia desde fuera del procedimiento almacenado.
- Cláusulas:
 - WITH RECOMPILE (en el procedimiento o en el EXEC)
 - WITH ENCRYPTION (convertirá el texto original en un formato ofuscado)
 - { EXEC | EXECUTE } AS { CALLER | SELF | OWNER | 'user_name' }

© JMA 2009. All rights reserved

86

Funciones definidas por el usuario

- Las funciones son subrutinas formadas por una o varias instrucciones Transact-SQL o CLR que se pueden utilizar para encapsular un código con el fin de utilizarlo de nuevo posteriormente.
- Son subrutinas que se utilizan para encapsular lógica que se ejecuta frecuentemente que no modifique la base de datos.
- Dos tipos de funciones: integradas y definidas por el usuario.
- Las funciones definidas por el usuario pueden tener parámetros de entrada y devuelven un único valor (dato o tabla interna)
- SQL Server admite cuatro tipos de funciones definidas por el usuario:
 - Funciones escalares
 - Funciones de valores de tabla en línea
 - Funciones de valores de tabla de múltiples instrucciones
 - Funciones de agregado (solo CLR)
- Todas las funciones son deterministas o no deterministas:
 - Las funciones deterministas siempre devuelven el mismo resultado cada vez que son llamadas con un conjunto específico de valores de entrada.
 - Las funciones no deterministas podrían devolver resultados diferentes cada vez que son llamadas con un conjunto específico de valores de entrada.

© JMA 2009. All rights reserved

87

Desencadenadores DML

- Un desencadenador es un tipo especial de procedimiento almacenado que entra en vigor cuando se modifican datos en una tabla especificada utilizando una o más operaciones de modificación de datos: UPDATE (actualización), INSERT (inserción) o DELETE (eliminación).
- Los desencadenadores pueden exigir restricciones más complejas que las definidas con restricciones CHECK y permiten implementar las reglas de negocio.
- Pueden realizar cambios en cascada a través de tablas relacionadas de la base de datos.
- Los desencadenadores son automáticos: se activan inmediatamente después de que se efectúen modificaciones en los datos de la tabla, como una entrada manual o una acción de la aplicación.
 - Se pueden habilitar y deshabilitar.
 - Se pueden definir varios por tabla, cada uno para una o varias operaciones.
 - Se puede fijar cuál es el primero y el último en ejecutarse: `sp_settriggerorder 'First|None|Last'`

© JMA 2009. All rights reserved

88

Desencadenadores DML

- Se pueden implementar en Transact-SQL o CLR
- Los desencadenadores también pueden evaluar el estado de una tabla antes y después de realizar una modificación de datos y actuar en función de la diferencia.
- Seudotablas:
 - INSERTED: nuevos valores
 - DELETED: valores antiguos
- Funciones:
 - UPDATE(column) y COLUMNS_UPDATED()
- Momentos:
 - AFTER
 - INSTEAD OF

© JMA 2009. All rights reserved

89

Desencadenadores DDL/Logon

- Pueden utilizarse para tareas administrativas como auditar y regular las operaciones de base de datos: CREATE, ALTER y DROP.
 - ámbito del servidor
 - ámbito de la base de datos
- La información acerca de un evento que activa un desencadenador DDL se captura mediante la función EVENTDATA (formato XML)

© JMA 2009. All rights reserved

90

Sinónimos

- Un sinónimo es un nombre alternativo para un objeto (local o remoto) de ámbito de esquema.
- Proporciona una capa de abstracción que protege una aplicación cliente de cambios hechos en el nombre o la ubicación del objeto base.
- Se pueden proporcionar sinónimos a:
 - Tablas, Vistas, Procedimientos Almacenados y Funciones.

© JMA 2009. All rights reserved

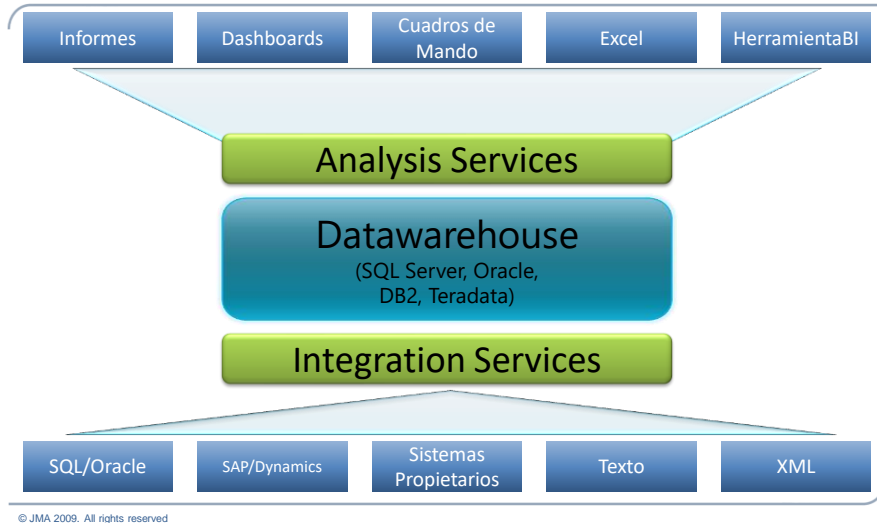
91

IMPORTACIÓN Y EXPORTACIÓN

© JMA 2009. All rights reserved

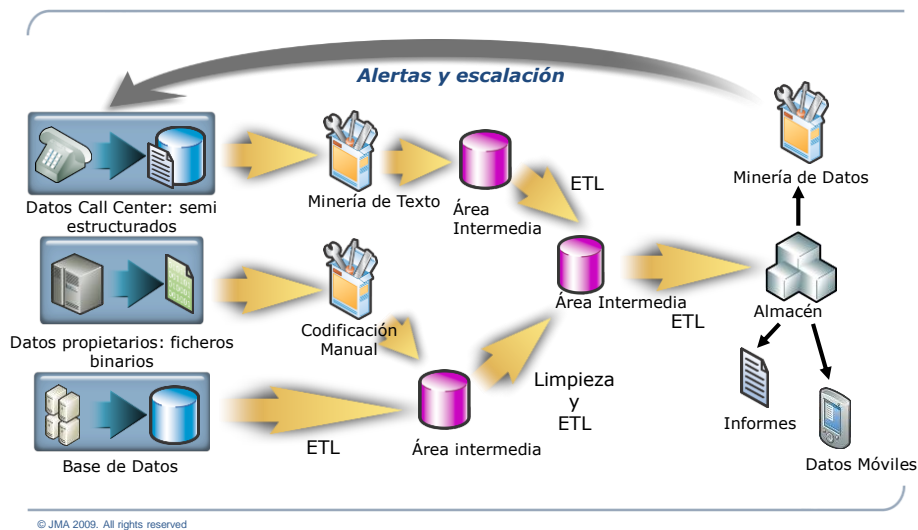
92

Arquitectura B.I.



93

Antes de Integration Services



94

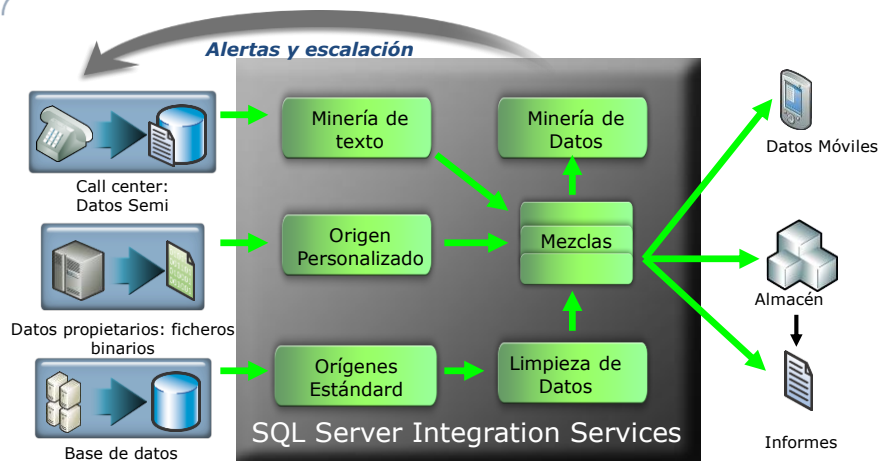
Antes de Integration Services

- Integración y almacenamiento requieren operaciones y almacenamiento intermedio.
- La preparación de datos necesita de herramientas diferentes, en muchos casos incompatibles.
- Reporting y escalación son procesos lentos, retrasando las respuestas.
- Con grandes volúmenes de datos, este escenario poco manejable.

© JMA 2009. All rights reserved

95

Con Integration Services



© JMA 2009. All rights reserved

96

Con Integration Services

- Integración y almacenamiento en una operación sencilla y fácilmente administrable.
- Origen, preparación y carga de datos en un único proceso auditable.
- Reporting y escalación pueden paralelizarse con la carga del almacén.
- Puede escalar a requisitos muy altos de complejidad y carga.

© JMA 2009. All rights reserved

97

Métodos para la transferencia masiva de datos

- BCP
 - Comando de sistema operativo que permite:
 - Exportar datos de una tabla o vista a un fichero.
 - Importar datos de un fichero a una tabla.
- BULK INSERT
 - Instrucción Transact-SQL para
- SELECT INTO/INSERT SELECT
- Funciones de conjuntos de filas
- Replicación/Duplicación
- PowerShell
- Integration Services (DTS)

© JMA 2009. All rights reserved

98

BCP

- La utilidad bcp hace copias masivas de los datos entre una instancia de Microsoft SQL Server y un archivo de datos en un formato especificado por el usuario.
- La utilidad bcp se puede usar para importar un número elevado de filas nuevas en tablas de SQL Server o para exportar datos de tablas a archivos de datos.
- Excepto cuando se usa con la opción queryout, la utilidad no exige ningún conocimiento de Transact-SQL.
- Para importar datos en una tabla, debe usar un archivo de formato creado para esa tabla o comprender la estructura de la tabla y los tipos de datos que son válidos para sus columnas.
 bcp AdventureWorks2012.Sales.Currency out Currency.dat -c -U<login_id> -S<server_name\instance_name>
 bcp AdventureWorks2012.Sales.Currency2 in Currency.dat -T -c

© JMA 2009. All rights reserved

99

BULK INSERT

- Importa un archivo de datos en una tabla o vista de base de datos con un formato especificado por el usuario.
 BULK INSERT AdventureWorks2012.Sales.SalesOrderDetail
 FROM 'f:\orders\lineitem.tbl'
 WITH
 (
 FIELDTERMINATOR = '|',
 ROWTERMINATOR = ':\n',
 FIRE_TRIGGERS
);

© JMA 2009. All rights reserved

100

Funciones de conjuntos de filas

- **OPENQUERY (linked_server , 'query')**
 - Ejecuta la consulta de paso a través especificada en el servidor vinculado especificado. (OLEDB)
- **OPENDATASOURCE (provider_name, init_string)**
 - Proporciona información de conexión ad hoc como parte de un nombre de objeto de cuatro partes sin utilizar un nombre de servidor vinculado.
- **OPENROWSET('provider_name' , 'datasource' ; 'user_id' ; 'password', 'query')**
 - Contiene toda la información de conexión necesaria para tener acceso a datos remotos desde un origen de datos OLE DB.
- **OPENXML(idoc int [in] , rowpattern nvarchar [in] , [flags byte [in]])**
 - Proporciona una vista de un conjunto de filas en un documento XML.
- **OPENJSON(jsonExpression [, path])**
 - Es una función con valores de tabla que analiza el texto JSON

```
SELECT * FROM OPENDATASOURCE('Microsoft.Jet.OLEDB.4.0',
'Data Source=C:\Docs\TestExcel.xls;Extended Properties=EXCEL 5.0')...[Sheet1$];
```

© JMA 2009. All rights reserved

101

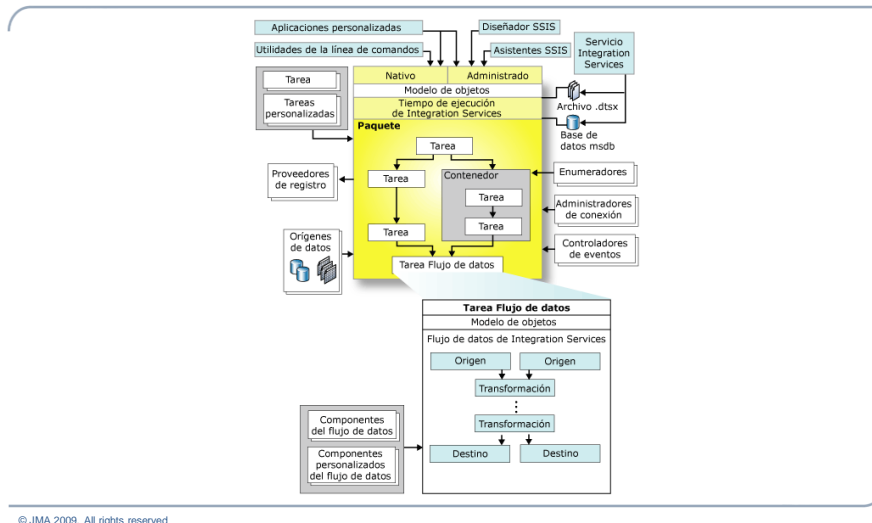
SQL Server Integration Services

- Microsoft Integration Services es una plataforma para compilar soluciones de transformaciones de datos e integración de datos de nivel empresarial.
- Integration Services sirve para resolver complejos problemas empresariales mediante la copia o descarga de archivos, el envío de mensajes de correo electrónico como respuesta a eventos, la actualización de almacenamientos de datos, la limpieza y minería de datos, y la administración de objetos y datos de SQL Server.
- Los paquetes pueden funcionar en solitario o junto con otros paquetes para hacer frente a las complejas necesidades de la empresa. Integration Services puede extraer y transformar datos de diversos orígenes como archivos de datos XML, archivos planos y orígenes de datos relacionales y, después, cargar los datos en uno o varios destinos.
- Integration Services incluye una amplia gama de tareas integradas y transformaciones; herramientas de creación de paquetes y el Integration Services service para ejecutar y administrar paquetes.
- Las herramientas gráficas de Integration Services se pueden usar para crear soluciones sin escribir una sola línea de código. También se puede programar el amplio modelo de objetos de Integration Services para crear paquetes mediante programación y codificar tareas personalizadas y otros objetos de paquete.

© JMA 2009. All rights reserved

102

Componentes



© JMA 2009. All rights reserved

103

Paquetes

- Un paquete es una colección organizada de conexiones, elementos de flujo de control, elementos de flujo de datos, controladores de eventos, variables, parámetros y configuraciones que se pueden ensamblar con la ayuda de las herramientas gráficas de diseño proporcionadas por SQL Server Integration Services o mediante programación.
- Un paquete se puede guardar en SQL Server, en el Almacén de paquetes SSIS o en el sistema de archivos, o se puede implementar el proyecto de ssISnoversion en el servidor SSIS.
- El paquete es la unidad de trabajo que se recupera, ejecuta y guarda.
- Cada paquete contiene uno o más pasos que se ejecutan de forma secuencial o en paralelo cuando se ejecuta el paquete
- El flujo de trabajo permite vincular dos tareas en función de si se ejecuta la primera, si se ejecuta correctamente o si se produce un error
- La Conectividad se basa en la arquitectura OLE DB, que permite copiar y transformar datos de distintos orígenes

© JMA 2009. All rights reserved

104

Flujos

- Los elementos de flujo de control (tareas y contenedores) para generar el flujo de control en un paquete. Los elementos de flujo de control preparan o copian datos, interactúan con otros procesos o implementan flujo de trabajo repetido. Las restricciones de precedencia ordenan en una secuencia los elementos de flujo de control en un flujo de control ordenado y especifican las condiciones para ejecutar tareas o contenedores.
- Los componentes de flujo de datos (orígenes, transformaciones y destinos) para generar flujos de datos en un paquete que extrae, transforma y carga datos. Las rutas ordenan los componentes del flujo de datos en un flujo de datos ordenado.

© JMA 2009. All rights reserved

105

Otros componentes

- Los **administradores de conexión** que se conectan a diferentes tipos de orígenes de datos para extraer y cargar datos.
- Las **variables** que se pueden usar en expresiones para actualizar de forma dinámica los valores de columna y expresiones de propiedad, controlar la ejecución de flujos de control repetidos y definir las condiciones a las que se aplican las restricciones de precedencia.
- Los **parámetros** permiten asignar valores a propiedades de los paquetes en el momento de la ejecución de los mismos.
- Los **controladores de eventos** que se ejecutan como respuesta ante los eventos de tiempo de ejecución que activan los paquetes, tareas y contenedores.
- Los **proveedores de registro** que admiten el registro de información de tiempo de ejecución de paquetes, como la hora de inicio y de detención del paquete y sus tareas y contenedores.

© JMA 2009. All rights reserved

106

Flujo de control

- Un paquete consta de un flujo de control y, opcionalmente, uno o varios flujos de datos.
- SQL Server Integration Services proporciona tres tipos diferentes de elementos de flujo de control:
 - los contenedores que proporcionan las estructuras de los paquetes.
 - las tareas que proporcionan la funcionalidad .
 - las restricciones de precedencia que conectan los ejecutables, los contenedores y las tareas en un flujo de control ordenado.

© JMA 2009. All rights reserved

107

Contenedores

- Los contenedores son objetos que proporcionan estructura a los paquetes y servicios a las tareas. Permiten repetir flujos de control en paquetes y agrupan tareas y contenedores en unidades de trabajo significativas que deben completarse correctamente o no completarse como una unidad. Los contenedores pueden incluir otros contenedores, además de tareas.
 - Foreach Loop: Ejecuta un flujo de control repetidamente mediante un enumerador.
 - de bucles For: Ejecuta un flujo de control repetidamente probando una condición.
 - de secuencias: Agrupa tareas y contenedores en flujos de control que son subconjuntos del flujo de control del paquete.
 - de tarea: Proporciona servicios a una tarea individual.

© JMA 2009. All rights reserved

108

Tareas

- Las tareas son elementos de flujo de control que definen las unidades de trabajo que se realizan en un flujo de control de paquetes.
- Un paquete consta de una o más tareas.
- Si el paquete contiene más de una tarea, las tareas se conectan y ordenan en el flujo de control mediante restricciones de precedencia.
- **Tarea Flujo de datos:**
 - Tarea que ejecuta flujos de datos para extraer datos, aplicar transformaciones de nivel de columna y cargar datos.
 - Tarea Flujo de datos

© JMA 2009. All rights reserved

109

Tareas de preparación de datos

- Tareas que copian archivos y directorios, descargan archivos y datos, ejecutan métodos Web y aplican operaciones a documentos XML.
 - Tarea Sistema de archivos
 - Tarea FTP
 - Tarea Servicio Web
 - Tarea XML

© JMA 2009. All rights reserved

110

Tareas de flujo de trabajo

- Tareas que se comunican con otros procesos para ejecutar paquetes, ejecutar programas o archivos de proceso por lotes, enviar y recibir mensajes entre paquetes, enviar mensajes de correo electrónico, leer datos de Instrumental de administración de Windows (WMI) y detectar eventos de WMI.
 - Tarea Ejecutar paquete
 - Tarea Ejecutar paquete DTS 2000
 - Tarea Ejecutar proceso
 - Tarea Cola de mensajes
 - Tarea Enviar correo
 - Tarea Lector de datos WMI
 - Tarea Monitor de eventos WMI

© JMA 2009. All rights reserved

111

Tareas de SQL Server

- Tareas de acceso, copia, inserción, eliminación y modificación de objetos y datos de SQL Server.
 - Tarea Inserción masiva
 - Tarea Ejecutar SQL
 - Tarea Transferir bases de datos
 - Tarea Transferir mensajes de error
 - Tarea Transferir trabajos
 - Tarea Transferir inicios de sesión
 - Tarea Transferir procedimientos almacenados principales
 - Tarea Transferir objetos de SQL Server

© JMA 2009. All rights reserved

112

Tareas de mantenimiento

- Tareas que realizan funciones administrativas como crear copias de seguridad y reducir bases de datos de SQL Server, volver a generar y reorganizar índices, y ejecutar trabajos del Agente SQL Server.
 - Tarea Copia de seguridad de la base de datos
 - Tarea Comprobar la integridad de la base de datos
 - Tarea Ejecutar trabajo del Agente SQL Server
 - Tarea Ejecutar instrucción T-SQL
 - Tarea Limpieza de historial
 - Tarea Notificar al operador
 - Tarea Volver a generar índice
 - Tarea Reorganizar índice
 - Tarea Reducir base de datos
 - Tarea Actualizar estadísticas

© JMA 2009. All rights reserved

113

Tareas de Analysis Services

- Tareas de creación, modificación, eliminación y procesamiento de objetos de Analysis Services.
 - Tarea Procesamiento de Analysis Services
 - Tarea Ejecutar DDL de Analysis Services
 - Tarea Consulta de minería de datos

© JMA 2009. All rights reserved

114

Tareas de secuencias de comandos

- Tareas que amplían la funcionalidad de los paquetes mediante secuencias de comandos.
 - Tarea Secuencia de comandos ActiveX
 - Tarea Secuencia de comandos
- **Tareas personalizadas**
 - Además, también puede escribir tareas personalizadas mediante un lenguaje de programación compatible con COM, como Visual Basic, o un lenguaje de programación .NET, como C#. Si desea tener acceso a una tarea personalizada en el Diseñador SSIS, puede crear y registrar

© JMA 2009. All rights reserved

115

Tarea Flujo de datos

- **Orígenes:**
 - es el componente de flujo de datos que pone datos de diferentes orígenes de datos externos a disposición de los demás componentes en el flujo de datos.
- **Destinos:**
 - es el componente de flujo de datos que escribe los datos desde un flujo de datos en un almacén de datos específico, o crea un conjunto de datos almacenado en la memoria.
- **Transformaciones:**
 - Las transformaciones son los componentes en el flujo de datos de un paquete que agregan, mezclan, distribuyen y modifican datos.
 - Las transformaciones también pueden realizar operaciones de búsqueda y generar conjuntos de datos de ejemplo.

© JMA 2009. All rights reserved

117

Orígenes

- Origen de OLE DB: Consume datos de un proveedor OLE DB.
- Origen de Excel: Extrae datos de un archivo de Excel.
- Origen de archivo plano: Extrae datos de un archivo plano.
- Origen de archivo sin procesar: Extrae datos sin procesar de un archivo.
- Origen de DataReader: Consume datos de un proveedor de datos .NET Framework.
- Componente de secuencia de comandos: Usa secuencias de comandos para extraer, transformar o cargar datos.
- Origen XML: Extrae datos de un archivo XML.

© JMA 2009. All rights reserved

118

Destinos

- OLE DB: Carga datos mediante un proveedor OLE DB.
- Excel: Escribe datos en un libro de Excel.
- Archivo plano: Escribe datos en un archivo plano. Destino de archivo sin procesar: Escribe datos sin procesar en un archivo.
- Destinos SQL Server Analysis Services .
 - Entrenamiento del modelo de minería de datos: Entrena modelos de minería de datos.
 - Procesamiento de dimensiones: Carga y procesa una dimensión de SQL Server Analysis Services .
 - Procesamiento de particiones: Carga y procesa una partición de Analysis Services.
- Conjunto de registros: Crea un conjunto de registros ADO.
- DataReader: Expone los datos en un flujo de datos mediante la interfaz DataReader ADO.NET.
- Componente de secuencia de comandos: Usa secuencias de comandos para extraer, transformar o cargar datos.
- SQL Server Mobile: Inserta filas en una base de datos de SQL Server Mobile.
- SQL Server: Realiza una inserción masiva de datos en una tabla o vista de SQL Server.

© JMA 2009. All rights reserved

119

Transformaciones de fila

- Las siguientes transformaciones actualizan los valores de columna y crean nuevas columnas. La transformación se aplica a cada fila en la entrada de transformación.
 - Mapa de caracteres: Transformación que aplica funciones de cadena a caracteres.
 - Copiar columna: Transformación que agrega copias de columnas de entrada a la salida de transformación.
 - Conversión de datos: Transformación que convierte el tipo de datos de una columna en un tipo de datos diferente.
 - Columna derivada: Transformación que llena las columnas con los resultados de las expresiones.
 - Componente de secuencia de comandos: Transformación que usa secuencias de comandos para extraer, transformar o cargar datos.
 - Comando de OLE DB: Transformación que ejecuta comandos SQL para cada fila en un flujo de datos.

© JMA 2009. All rights reserved

120

Transformaciones de conjunto de filas

- Las siguientes transformaciones crean nuevos conjuntos de filas. El conjunto de filas puede incluir valores agregados y ordenados, conjuntos de filas de ejemplo y conjuntos de filas dinamizados y de anulación de dinamización.
 - Agregado: Transformación que realiza agregaciones tales como AVERAGE, SUM y COUNT.
 - Ordenar: Transformación que ordena datos.
 - Muestreo de porcentaje: Transformación que crea un conjunto de datos de ejemplo mediante un porcentaje para especificar el tamaño del ejemplo.
 - Muestreo de fila: Transformación que crea un conjunto de datos de ejemplo especificando la cantidad de filas en la muestra.
 - Dinámica: Transformación que crea una versión menos normalizada de una tabla normalizada.
 - Anulación de dinamización: Transformación que crea una versión más normalizada de una tabla no normalizada.

© JMA 2009. All rights reserved

121

Transformaciones de división y combinación

- Las siguientes transformaciones distribuyen filas a diferentes salidas, crean copias de las entradas de transformación, combinan varias entradas en una salida y realizan operaciones de búsqueda.
 - División condicional: Transformación que enruta las filas de datos a diferentes salidas.
 - Multidifusión: Transformación que distribuye conjuntos de datos a varias salidas.
 - Unión de todo: Transformación que mezcla varios conjuntos de datos.
 - Mezclar: Transformación que mezcla dos conjuntos de datos ordenados.
 - Combinación de mezcla: Transformación que mezcla dos conjuntos de datos mediante una combinación FULL, LEFT o INNER.
 - Búsqueda: Transformación que busca valores en una tabla de referencia mediante una coincidencia exacta.

© JMA 2009. All rights reserved

122

Transformaciones de inteligencia empresarial

- Las siguientes transformaciones realizan operaciones de inteligencia empresarial tales como limpiar datos, realizar minería de texto y ejecutar consultas de predicción de minería de datos.
 - Agrupación aproximada: Transformación que estandariza los valores en los datos de columnas.
 - Búsqueda aproximada: Transformación que busca valores en una tabla de referencia mediante una coincidencia aproximada.
 - Extracción de términos: Transformación que extrae términos del texto.
 - Búsqueda de términos: Transformación que busca términos en una tabla de referencia y cuenta términos extraídos del texto.
 - Consulta de minería de datos: Transformación que ejecuta consultas de predicción de minería de datos.

© JMA 2009. All rights reserved

123

Otras transformaciones

- Integration Services incluye las siguientes transformaciones para exportar e importar datos, agregar información de auditoría, contar filas y trabajar con dimensiones de variación lenta.
 - Exportar columna: Transformación que inserta datos desde un flujo de datos en un archivo.
 - Importar columna: Transformación que lee datos desde un archivo y los agrega a un flujo de datos.
 - Auditar: Transformación que coloca información sobre el entorno a disposición del flujo de datos en un paquete.
 - Recuento de filas: Transformación que cuenta filas a medida que se mueven por ella y almacena el recuento final en una variable.
 - Dimensión de variación lenta: Transformación que configura la actualización de una dimensión de variación lenta.
- Transformaciones personalizadas
 - También puede escribir transformaciones personalizadas.

© JMA 2009. All rights reserved

124

Las Herramientas SSIS

- Asistente de importación de archivos planos
- Asistente de Importación / Exportación
- SSIS y el Management Studio
- SQL Server Data Tools
 - Solution Explorer
 - SSIS Package Designer

© JMA 2009. All rights reserved

127

Asistente de importación de archivos planos

- El Asistente para importación de archivos planos es una forma muy sencilla de copiar datos de un archivo plano (.csv, .txt) en un destino determinado.
- Para acceder al Asistente para importación de archivos planos, siga estos pasos:
 - Abrir SQL Server Management Studio.
 - Conéctare a una instancia de Motor de base de datos de SQL Server o el localhost.
 - Expandir Bases de datos, hacer clic con el botón derecho en una base de datos, seleccionar Tareas y hacer clic en Importar archivo plano:
 - Especificar el archivo de entrada
 - Vista previa de los datos
 - Modificar las columnas
 - Ejecutar

© JMA 2009. All rights reserved

128

Asistente de Importación / Exportación

- El Asistente para importación y exportación de SQL Server ofrece una manera sencilla de copiar datos de un origen a un destino.
- Iniciar el Asistente para importación y exportación de SQL Server desde el menú Inicio
 - En el menú Inicio, busque y expanda Microsoft SQL Server 2017.
 - Haga clic en una de las opciones siguientes.
 - Importar y exportar datos de SQL Server 2017 (64 bits)
 - Importar y exportar datos de SQL Server 2017 (32 bits)
- Iniciar el asistente para importación y exportación de SQL Server desde SQL Server Management Studio (SSMS)
 - Expanda Bases de datos.
 - Haga clic con el botón derecho en una base de datos.
 - Seleccione Tareas y haga clic en una de las opciones siguientes.
 - Importar datos
 - Exportar datos

© JMA 2009. All rights reserved

129

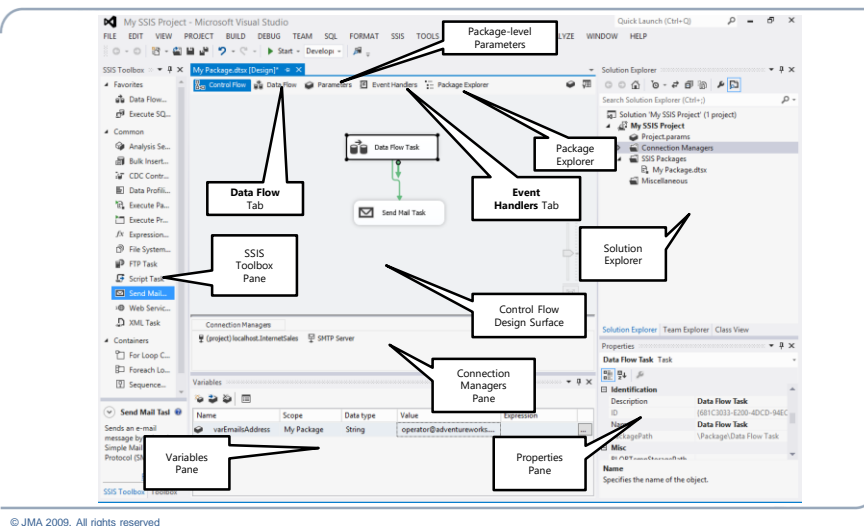
SQL Server Data Tools

- SQL Server proporciona SQL Server Data Tools (SSDT) para el desarrollo de paquetes de Integration Services .
- Los paquetes de Integration Services residen en los proyectos.
- Los proyectos se almacenan en soluciones.
- Primero se crea una solución y luego se agrega un proyecto de Integration Services a la solución.
- Si no existe solución, SQL Server Data Tools (SSDT) crea automáticamente una cuando se crea el proyecto.
- Una solución puede contener varios proyectos de tipos diferentes.

© JMA 2009. All rights reserved

130

SQL Server Data Tools



131

EL FLUJO DE DATOS

© JMA 2009. All rights reserved

132

Tarea Flujo de datos

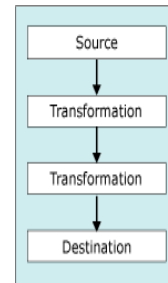
- La tarea Flujo de datos encapsula el motor de flujo de datos que mueve datos entre orígenes y destinos, y permite al usuario transformar, limpiar y modificar datos a medida que se mueven. Agregar una tarea Flujo de datos a un flujo de control de paquetes permite que el paquete extraiga, transforme y cargue datos.
 - Un flujo de datos se compone de por lo menos un componente de flujo de datos, pero normalmente es un conjunto de componentes de flujo de datos conectados: orígenes que extraen datos, transformaciones que modifican, enrutan o resumen datos, y destinos que cargan datos.
 - En el tiempo de ejecución, la tarea Flujo de datos genera un plan de ejecución a partir del flujo de datos y el motor de flujo de datos ejecuta el plan. Se puede crear una tarea Flujo de datos que no tenga flujo de datos, pero la tarea solo se ejecuta si incluye por lo menos un flujo de datos.
 - Para realizar una inserción masiva de datos de archivos de texto en una base de datos de SQL Server, se puede usar la tarea Inserción masiva en lugar de la tarea Flujo de datos y un flujo de datos.
-

© JMA 2009. All rights reserved

134

Tarea Flujo de datos

- **Orígenes:**
 - es el componente de flujo de datos que pone datos de diferentes orígenes de datos externos a disposición de los demás componentes en el flujo de datos.
- **Destinos:**
 - es el componente de flujo de datos que escribe los datos desde un flujo de datos en un almacén de datos específico, o crea un conjunto de datos almacenado en la memoria.
- **Transformaciones:**
 - Las transformaciones son los componentes en el flujo de datos de un paquete que agregan, mezclan, distribuyen y modifican datos.
 - Las transformaciones también pueden realizar operaciones de búsqueda y generar conjuntos de datos de ejemplo.



© JMA 2009. All rights reserved

135

Administradores de conexiones (conexiones).

- Un paquete suele incluir al menos un administrador de conexiones.
- Un administrador de conexiones es un vínculo entre un paquete y un origen de datos que define la cadena de conexión para acceder a los datos que las tareas, transformaciones y controladores de eventos del paquete usan.
- Integration Services incluye tipos de conexiones para orígenes de datos tales como archivos de texto y XML, bases de datos relacionales y proyectos y bases de datos de Analysis Services.
- Se puede crear un administrador de conexiones en el nivel de paquete o en el nivel de proyecto.
 - El administrador de conexiones creado en el nivel de proyecto está disponible para todos los paquetes del proyecto.
 - El administrador de conexiones creado en el nivel de paquete está disponible para ese paquete específico.

© JMA 2009. All rights reserved

136

Orígenes

- Origen de OLE DB: Consume datos de un proveedor OLE DB.
- Origen de Excel: Extrae datos de un archivo de Excel.
- Origen de archivo plano: Extrae datos de un archivo plano.
- Origen de archivo sin procesar: Extrae datos sin procesar de un archivo.
- Origen de DataReader: Consume datos de un proveedor de datos .NET Framework.
- Componente de secuencia de comandos: Usa secuencias de comandos para extraer, transformar o cargar datos.
- Origen XML: Extrae datos de un archivo XML.

© JMA 2009. All rights reserved

137

Destinos

- OLE DB: Carga datos mediante un proveedor OLE DB.
- Excel: Escribe datos en un libro de Excel.
- Archivo plano: Escribe datos en un archivo plano. Destino de archivo sin procesar: Escribe datos sin procesar en un archivo.
- Destinos SQL Server Analysis Services .
 - Entrenamiento del modelo de minería de datos: Entrena modelos de minería de datos.
 - Procesamiento de dimensiones: Carga y procesa una dimensión de SQL Server Analysis Services .
 - Procesamiento de particiones: Carga y procesa una partición de Analysis Services.
- Conjunto de registros: Crea un conjunto de registros ADO.
- DataReader: Expone los datos en un flujo de datos mediante la interfaz DataReader ADO.NET.
- Componente de secuencia de comandos: Usa secuencias de comandos para extraer, transformar o cargar datos.
- SQL Server Mobile: Inserta filas en una base de datos de SQL Server Mobile.
- SQL Server: Realiza una inserción masiva de datos en una tabla o vista de SQL Server.

© JMA 2009. All rights reserved

138

Transformar datos con transformaciones

- Integration Services ofrecen la siguiente funcionalidad:
 - Dividir, copiar y combinar conjuntos de filas y realizar operaciones de búsqueda.
 - Actualizar valores de columnas y crear nuevas columnas aplicando transformaciones tales como cambio de minúsculas por mayúsculas.
 - Operaciones de inteligencia empresarial tales como limpiar datos, realizar minería de texto y ejecutar consultas de predicción de minería de datos.
 - Crear nuevos conjuntos de filas que se componen de valores agregados u ordenados, datos de muestra o datos dinamizados y de anulación de dinamización.
 - Realizar tareas tales como exportar e importar datos, proporcionar información de auditoría y trabajar con dimensiones de variación lenta.
- También se pueden escribir transformaciones personalizadas.
- Después de agregar la transformación al diseñador de flujo de datos, pero antes de configurar la transformación, se debe conectar la transformación al flujo de datos conectando la salida de otra transformación u origen del flujo de datos a la entrada de esta transformación. El conector entre dos componentes de flujo de datos se denomina ruta.

© JMA 2009. All rights reserved

139

Transformaciones de fila

- Las siguientes transformaciones actualizan los valores de columna y crean nuevas columnas. La transformación se aplica a cada fila en la entrada de transformación.
 - Mapa de caracteres: Transformación que aplica funciones de cadena a caracteres.
 - Copiar columna: Transformación que agrega copias de columnas de entrada a la salida de transformación.
 - Conversión de datos: Transformación que convierte el tipo de datos de una columna en un tipo de datos diferente.
 - Columna derivada: Transformación que llena las columnas con los resultados de las expresiones.
 - Componente de secuencia de comandos: Transformación que usa secuencias de comandos para extraer, transformar o cargar datos.
 - Comando de OLE DB: Transformación que ejecuta comandos SQL para cada fila en un flujo de datos.

© JMA 2009. All rights reserved

140

Transformación Mapa de caracteres

- La transformación Mapa de caracteres se aplica a funciones de cadena que operan sobre datos de caracteres, como la conversión de minúsculas a mayúsculas. Esta transformación solo opera en datos de columnas con un tipo de datos de cadena.
- La transformación Mapa de caracteres puede convertir datos de columna in situ o agregar una columna a la salida de transformación y colocar los datos convertidos en la nueva columna. Puede aplicar distintos conjuntos de operaciones de asignación a la misma columna de entrada y colocar los resultados en columnas diferentes. Por ejemplo, puede convertir la misma columna a mayúsculas y minúsculas, y almacenar el resultado en dos columnas diferentes.
- En algunas situaciones, la asignación puede provocar un truncamiento de datos. Por ejemplo, se puede producir un truncamiento cuando se asignan caracteres de un byte a caracteres representados con varios bytes. La transformación Mapa de caracteres incluye una salida de error que se puede usar para dirigir los datos truncados a otra salida distinta.

© JMA 2009. All rights reserved

141

Transformación Copiar columna

- La transformación crea columnas nuevas copiando columnas de entrada y agregando las columnas nuevas a la salida de transformación.
- En una fase posterior del flujo de datos se pueden aplicar distintas transformaciones a las copias de columnas.
- Por ejemplo, puede usar la transformación Copiar columna para crear una copia de una columna y después convertir los datos copiados a mayúsculas mediante la transformación Mapa de caracteres, o aplicar agregaciones a la nueva columna mediante la transformación Agregado.
- Se puede crear varias copias de una columna o crear copias de varias columnas en una operación.
- Esta transformación tiene una entrada y una salida. No admite una salida de error.

© JMA 2009. All rights reserved

142

Transformación Conversión de datos

- La transformación Conversión de datos convierte los datos de una columna de entrada a otro tipo de datos diferente y después los copia a una nueva columna de salida.
- Por ejemplo, un paquete puede extraer los datos de diferentes orígenes y después usar esta transformación para convertir las columnas al tipo de datos necesario para el almacén de datos de destino.
- Se puede aplicar múltiples conversiones a una sola columna de entrada.
- Un paquete puede utilizar esta transformación para realizar las siguientes conversiones de tipos de datos:
 - Cambiar el tipo de datos.
 - Establecer la longitud de la columna de los datos de cadena así como la precisión y la escala de los datos numéricos.
 - Especificar una página de códigos.
- Si la longitud de una columna de salida de datos de tipo cadena es menor que la longitud de la columna de entrada correspondiente, se truncarán los datos de salida. Para más información, vea Control de errores en los datos.
- Esta transformación tiene una entrada, una salida y una salida de error.

© JMA 2009. All rights reserved

143

Tipos de datos

Tipo de datos	Descripción
DT_BOOL	Valor booleano.
DT_BYTES	Valor de datos binarios. La longitud es variable y la longitud máxima es de 8000 bytes.
DT_CY	Un valor de moneda. Este tipo de datos es un entero con signo de 8 bytes con una escala de 4 y una precisión máxima de 19 dígitos.
DT_DATE	Una estructura de fecha formada por año, mes, día, hora, minuto, segundo y fracciones de segundo. Las fracciones de segundo tienen una escala fija de 7 dígitos.
DT_DBDATE	Una estructura de fecha compuesta por año, mes y día.
DT_DBTIME	Una estructura de hora compuesta por horas, minutos y segundos.
DT_DBTIME2	Una estructura de hora formada por hora, minuto, segundo y fracciones de segundo. Las fracciones de segundo tienen una escala máxima de 7 dígitos.
DT_DBTIMESTAMP	Una estructura de marca de tiempo formada por año, mes, día, hora, minuto, segundo y fracciones de segundo. Las fracciones de segundo tienen una escala máxima de 3 dígitos.
DT_DBTIMESTAMP2	Una estructura de marca de tiempo formada por año, mes, día, hora, minuto, segundo y fracciones de segundo. Las fracciones de segundo tienen una escala máxima de 7 dígitos.
DT_DBTIMESTAMPOFFSET	Una estructura de marca de tiempo formada por año, mes, día, hora, minuto, segundo y fracciones de segundo. Las fracciones de segundo tienen una escala máxima de 7 dígitos.

© JMA 2009. All rights reserved

144

Tipos de datos

Tipo de datos	Descripción
DT_DECIMAL	Un valor numérico exacto con una precisión fija y una escala fija. Este tipo de dato es un entero sin signo de 12 bytes con un signo aparte, una escala de 0 a 28 y una precisión máxima de 29.
DT_FILETIME	Un valor de 64 bits que representa la cantidad de intervalos de 100 nanosegundos desde el 1 de enero de 1601. Las fracciones de segundo tienen una escala máxima de 3 dígitos.
DT_GUID	Identificador único global (GUID).
DT_I1	Un entero con signo de un byte.
DT_I2	Un entero con signo de dos bytes.
DT_I4	Un entero con signo de cuatro bytes.
DT_I8	Un entero con signo de ocho bytes.
DT_NUMERIC	Un valor numérico exacto con una precisión y escala fijas. Este tipo de dato es un entero sin signo de 16 bytes con un signo aparte, una escala de 0 a 38 y una precisión máxima de 38.
DT_R4	Un valor de punto flotante y precisión simple.
DT_R8	Un valor de punto flotante y precisión doble.
DT_STR	Una cadena de caracteres ANSI/MBCS terminada en NULL, con una longitud máxima de 8000 caracteres. (Si un valor de columna contiene otros terminadores NULL, la cadena se truncará cuando se encuentre el primer NULL.)

© JMA 2009. All rights reserved

145

Tipos de datos

Tipo de datos	Descripción
DT_UI1	Un entero sin signo de un byte.
DT_UI2	Un entero sin signo de dos bytes.
DT_UI4	Un entero sin signo de cuatro bytes.
DT_UI8	Un entero sin signo de ocho bytes.
DT_WSTR	Una cadena de caracteres Unicode terminada en NULL, con una longitud máxima de 4000 caracteres. (Si un valor de columna contiene otros terminadores NULL, la cadena se truncará cuando se encuentre el primer NULL.)
DT_IMAGE	Valor binario con un tamaño máximo de $2^{31}-1$ (2 147 483 647) bytes. .
DT_NTEXT	Cadena de caracteres Unicode con una longitud máxima de $2^{30}-1$ (1 073 741 823) caracteres.
DT_TEXT	Una cadena de caracteres ANSI/MBCS con una longitud máxima de $2^{31}-1$ (2 147 483 647) caracteres.

© JMA 2009. All rights reserved

146

Transformación Columna derivada

- La transformación Columna derivada crea nuevos valores de columna aplicando expresiones a las columnas de entrada de la transformación. Una expresión puede contener cualquier combinación de variables, funciones, operadores y columnas de la entrada de transformación. El resultado puede agregarse como una nueva columna o insertarse en una columna existente como un valor de reemplazo. La transformación Columna derivada puede definir varias columnas derivadas, y cualquier variable o columna de entrada puede aparecer en varias expresiones.
- Puede utilizar esta transformación para realizar las siguientes tareas:
 - Concatenar datos de distintas columnas en una columna derivada. Por ejemplo, puede combinar valores de las columnas FirstName y LastName en una sola columna derivada, denominada FullName, mediante la expresión `FirstName + " " + LastName`.
 - Extraer caracteres de datos de cadena mediante funciones como `SUBSTRING` y después almacenar el resultado en una columna derivada. Por ejemplo, puede extraer de la columna FirstName la inicial del nombre de una persona mediante la expresión `SUBSTRING(FirstName,1,1)`.
 - Aplicar funciones matemáticas a datos numéricos y almacenar el resultado en una columna derivada. Por ejemplo, puede cambiar la longitud y la precisión de una columna numérica, SalesTax, a un número con dos cifras decimales mediante la expresión `ROUND(SalesTax, 2)`.
 - Crear expresiones que comparen columnas de entrada y variables. Por ejemplo, puede comparar la variable Version con los datos de la columna ProductVersion, en función del resultado de la comparación, usar el valor de Version o ProductVersion mediante la expresión `ProductVersion == @Version? ProductVersion : @Version`.
 - Extraer partes de un valor datetime. Por ejemplo, puede utilizar las funciones `GETDATE` y `DATEPART` para extraer el año actual mediante la expresión `DATEPART("year",GETDATE())`.
 - Convierta las cadenas de fecha a un formato específico mediante una expresión.

© JMA 2009. All rights reserved

147

Transformación Comando de OLE DB

- La transformación Comando de OLE DB ejecuta una instrucción SQL para cada fila en un flujo de datos. Por ejemplo, puede ejecutar una instrucción SQL que inserte, actualice o elimine filas en una tabla de base de datos.
- Puede configurar el administrador de conexiones OLE DB de las maneras siguientes:
 - Proporcionar la instrucción SQL que la transformación ejecuta para cada fila.
 - Especifica la cantidad de segundos que tienen que transcurrir antes de que la instrucción SQL agote el tiempo de espera.
 - Especificar la página de códigos predeterminada.
- Normalmente, la instrucción SQL incluye parámetros. Los valores de parámetro se almacenan en columnas externas en la entrada de transformación y al asignar una columna de entrada a una columna externa se asigna una columna de entrada a un parámetro.
- La transformación Comando de OLE DB proporciona los nombres de parámetro y no puede modificarlos. Los nombres de parámetro son Param_0, Param_1 y así sucesivamente.

© JMA 2009. All rights reserved

148

Transformación Limpieza de DQS

- La transformación Limpieza de DQS usa Data Quality Services (DQS) para corregir datos de un origen de datos conectado aplicando reglas aprobadas que se crearon para el origen de datos conectado o un origen de datos similar.
- Para determinar si es necesario corregir los datos, la transformación Limpieza de DQS procesa los datos de una columna de entrada cuando se cumplen las condiciones siguientes:
 - La columna está seleccionada para la corrección de datos.
 - El tipo de datos de la columna se admite para la corrección de datos.
 - La columna está asignada a un dominio que tiene un tipo de datos compatible.

© JMA 2009. All rights reserved

149

Transformaciones de conjunto de filas

- Las siguientes transformaciones crean nuevos conjuntos de filas. El conjunto de filas puede incluir valores agregados y ordenados, conjuntos de filas de ejemplo y conjuntos de filas dinamizados y de anulación de dinamización.
 - Agregado: Transformación que realiza agregaciones tales como AVERAGE, SUM y COUNT.
 - Ordenar: Transformación que ordena datos.
 - Muestreo de porcentaje: Transformación que crea un conjunto de datos de ejemplo mediante un porcentaje para especificar el tamaño del ejemplo.
 - Muestreo de fila: Transformación que crea un conjunto de datos de ejemplo especificando la cantidad de filas en la muestra.
 - Dinámica: Transformación que crea una versión menos normalizada de una tabla normalizada.
 - Anulación de dinamización: Transformación que crea una versión más normalizada de una tabla no normalizada.

© JMA 2009. All rights reserved

150

Transformación Agregado

- La transformación Agregado aplica funciones de agregado, como Average, a los valores de columnas y copia los resultados en la salida de transformación.
- Además de las funciones de agregado, la transformación proporciona la cláusula GROUP BY, que se puede usar para especificar los grupos en los que se debe realizar el agregado.
 - GROUP BY: Divide los conjuntos de datos en grupos. Se pueden usar columnas de cualquier tipo de datos para la agrupación.
 - SUM: Suma los valores de una columna. Solo podrán sumarse las columnas con tipos de datos numéricos.
 - AVG: Devuelve la media de los valores de columna de una columna. Solo podrá calcularse la media de las columnas con tipos de datos numéricos.
 - COUNT: Devuelve el número de elementos de un grupo.
 - COUNT DISTINCT: Devuelve el número de valores únicos distintos de NULL de un grupo.
 - MIN: Devuelve el valor mínimo en un grupo.
 - MAX : Devuelve el valor máximo en un grupo.

© JMA 2009. All rights reserved

151

Transformación Ordenar

- La transformación Ordenar ordena los datos de entrada en sentido ascendente o descendente, y copia los datos ordenados a la salida de transformación. Puede aplicar varias ordenaciones a una entrada; cada ordenación se identifica mediante un numeral que determina el criterio de ordenación. La columna con el número más bajo se ordenará primero, la columna con el segundo número más bajo se ordena a continuación, etc.
- Las columnas que no están seleccionadas para ordenar se copian automáticamente a la salida de transformación junto con las columnas ordenadas.
- La transformación Ordenar incluye un conjunto de opciones de comparación para definir cómo controlará la transformación los datos de cadena de una columna.
- La transformación Ordenar también puede quitar filas duplicadas como parte de la ordenación. Las filas duplicadas son filas con los mismos criterios de ordenación. El valor del criterio de ordenación se genera a partir de las opciones de comparación de cadenas usadas, lo que implica que cadenas literales diferentes pueden tener los mismos criterios de ordenación. La transformación identifica filas en las columnas de entrada que tienen valores distintos pero un mismo criterio de ordenación que los duplicados.

© JMA 2009. All rights reserved

152

Transformaciones de división y combinación

- Las siguientes transformaciones distribuyen filas a diferentes salidas, crean copias de las entradas de transformación, combinan varias entradas en una salida y realizan operaciones de búsqueda.
 - División condicional: Transformación que enruta las filas de datos a diferentes salidas.
 - Multidifusión: Transformación que distribuye conjuntos de datos a varias salidas.
 - Unión de todo: Transformación que mezcla varios conjuntos de datos.
 - Mezclar: Transformación que mezcla dos conjuntos de datos ordenados.
 - Combinación de mezcla: Transformación que mezcla dos conjuntos de datos mediante una combinación FULL, LEFT o INNER.
 - Búsqueda: Transformación que busca valores en una tabla de referencia mediante una coincidencia exacta.

© JMA 2009. All rights reserved

153

Transformación Multidifusión

- La transformación Multidifusión distribuye su entrada en una o más salidas. Esta transformación es similar a la transformación División condicional. Ambas transformaciones dirigen una entrada a varias salidas. La diferencia entre las dos es que la transformación Multidifusión dirige cada fila a cada salida, mientras que la División condicional dirige una fila a una sola salida.
- Para configurar la transformación Multidifusión, debe agregar salidas.
- Un paquete puede crear copias lógicas de datos mediante la transformación Multidifusión. Esta capacidad resulta útil cuando el paquete tiene que aplicar varios conjuntos de transformaciones a los mismos datos. Por ejemplo, cuando se agrega una copia de los datos y solo se carga la información de resumen en su destino, mientras otra copia de los datos se extiende con valores de búsqueda y columnas derivadas antes de cargarla en su destino.

© JMA 2009. All rights reserved

154

Transformación División condicional

- La transformación División condicional puede dirigir filas de datos a salidas diferentes en función del contenido de los datos. La implementación de la transformación División condicional es similar a una estructura de decisión CASE de un lenguaje de programación. Evalúa expresiones y, en función de los resultados, dirige la fila de datos a la salida especificada. Esta transformación también proporciona una salida predeterminada, de modo que si una fila no coincide con ninguna expresión, se dirige a la salida predeterminada.
- Se puede configurar la transformación División condicional de las siguientes maneras :
 - Proporcionar una expresión cuya evaluación devuelva un valor booleano para cada condición que desee probar con la transformación.
 - Especificar el orden de evaluación de las condiciones. El orden es importante, ya que una fila se envía a la salida correspondiente a la primera condición que dé como resultado True.
 - Especificar la salida predeterminada para la transformación. La transformación requiere que se especifique una salida predeterminada.

© JMA 2009. All rights reserved

155

Transformación Unión de todo

- La transformación Unión de todo combina varias entradas en una salida. Por ejemplo, las salidas de cinco orígenes de archivos planos distintos pueden ser entradas de la transformación Unión de todo y combinarse en una salida.
- Las entradas de la transformación se agregan una detrás de otra a la salida de transformación; las filas no se reordenan. Si el paquete requiere una salida ordenada, debe usar la transformación Combinar en lugar de Unión de todo.
- La primera entrada que se puede conectar a la transformación Unión de todo es la entrada a partir de la cual la transformación crea su salida. Las columnas de las entradas que se conecten posteriormente a la transformación se asignarán a las columnas de la salida de transformación.
- Para combinar entradas, debe asignar columnas de las entradas a columnas de la salida. Se debe asignar una columna con al menos una entrada a cada columna de salida. La asignación entre dos columnas requiere que los metadatos de las columnas coincidan.
- Si las columnas asignadas contienen datos de cadena y la columna de salida es de menor longitud que la columna de entrada, se aumenta automáticamente la longitud de la columna de salida para que pueda contener la columna de entrada. Las columnas de entrada que no se asignan a columnas de salida se establecen en valores NULL en las columnas de salida.

© JMA 2009. All rights reserved

156

Transformación Mezclar

- La transformación Mezclar combina dos conjuntos de datos ordenados en un solo conjunto de datos. Las filas de cada conjunto de datos se insertan en la salida en función de los valores de sus columnas de clave.
- Si incluye la transformación Mezclar en un flujo de datos, podrá realizar las siguientes tareas:
 - Combinar datos de dos orígenes de datos, como tablas y archivos.
 - Crear conjuntos de datos complejos anidando transformaciones de combinación.
 - Volver a combinar filas después de corregir errores en los datos.
- La transformación Mezclar es similar a las transformaciones Unión de todo. Use la transformación Unión de todo en lugar de la transformación Mezclar en las siguientes situaciones:
 - Las entradas de la transformación no están ordenadas.
 - La salida combinada no tiene que ordenarse.
 - La transformación tiene más de dos entradas.
- La transformación Mezclar requiere datos ordenados para sus entradas.
- La transformación Mezclar también requiere que las columnas combinadas en sus entradas tengan metadatos coincidentes. Por ejemplo, no puede combinar una columna que tenga un tipo de datos numérico con una columna que tenga un tipo de datos de carácter.

© JMA 2009. All rights reserved

157

Transformación Combinación de mezcla

- La transformación Combinación de mezcla proporciona una salida que se genera combinando dos conjuntos de datos ordenados mediante una combinación FULL, LEFT o INNER. Por ejemplo, puede utilizar una combinación LEFT para combinar una tabla que incluye información de productos con una tabla que incluye el país o la región en que se fabricó un producto. El resultado es una tabla que muestra todos los productos y su país o región de origen.
- Puede configurar la transformación Combinación de mezcla de las siguientes maneras:
 - Especificar que la combinación es una combinación FULL, LEFT o INNER.
 - Especificar las columnas utilizadas por la combinación.
 - Especificar si la transformación controla valores NULL como iguales a otros valores NULL.
- La transformación Combinación de mezcla requiere datos ordenados para sus entradas.
- La transformación Combinación de mezcla requiere que las columnas combinadas tengan metadatos coincidentes.

© JMA 2009. All rights reserved

158

Transformación Búsqueda

- La transformación Búsqueda realiza búsquedas mediante la combinación de datos de columnas de entrada con columnas de un conjunto de datos de referencia. La búsqueda se utiliza para tener acceso a información adicional en una tabla relacionada que está basada en valores de columnas comunes.
- El conjunto de datos de referencia puede ser un archivo caché, una tabla o una vista existente, una tabla nueva o el resultado de una consulta SQL. La transformación Búsqueda utiliza un administrador de conexiones OLE DB o un administrador de conexiones de caché para conectar con el conjunto de datos de referencia.
- Se puede configurar la transformación Búsqueda de las siguientes maneras:
 - Seleccionar el administrador de conexiones que desea utilizar. Si desea conectarse a una base de datos, seleccione un administrador de conexiones OLE DB. Si desea conectarse a un archivo caché, seleccione un administrador de conexiones de caché.
 - Especificar la tabla o vista que contiene el conjunto de datos de referencia.
 - Generar una base de datos de referencia al especificar una instrucción SQL.
 - Especificar las combinaciones entre los datos de entrada y el conjunto de datos de referencia.
 - Agregar columnas del conjunto de datos de referencia a la salida de transformación Búsqueda.
 - Configurar las opciones de almacenamiento en caché.
- La transformación Búsqueda admite los siguientes proveedores de bases de datos para el administrador de conexiones OLE DB:
 - SQL Server
 - Oracle
 - DB2

© JMA 2009. All rights reserved

159

Transformación Muestreo de fila

- La transformación Muestreo de fila se usa para obtener un subconjunto seleccionado aleatoriamente de un conjunto de datos de entrada. Puede especificar el tamaño exacto del ejemplo de salida y especificar un valor de inicialización para el generador de números aleatorios.
- Hay muchas aplicaciones para el muestreo aleatorio. Por ejemplo, una compañía que desea seleccionar aleatoriamente 50 empleados que recibirán los premios de una lotería podría aplicar la transformación Muestreo de fila a la base de datos de empleados para generar el número exacto de ganadores.
- La transformación Muestreo de fila también es útil durante el desarrollo de paquetes para crear un conjunto de datos pequeño pero representativo. Puede probar la ejecución del paquete y la transformación de datos con datos representativos. Tardará menos tiempo, ya que usará una muestra aleatoria en lugar del conjunto de datos completo. Como el conjunto de datos de ejemplo utilizado por el paquete de prueba tiene siempre el mismo tamaño, al usar el subconjunto de ejemplo también resulta más fácil identificar posibles problemas de rendimiento del paquete.
- Esta transformación es similar a la transformación Muestreo de porcentaje, que crea un conjunto de datos de ejemplo seleccionando un porcentaje de las filas de entrada.

© JMA 2009. All rights reserved

160

Transformación Muestreo de porcentaje

- La transformación Muestreo de porcentaje crea un conjunto de datos de muestra seleccionando un porcentaje de las filas de entrada de transformación. El conjunto de datos de muestra es una selección aleatoria de filas de la entrada de transformación, de forma que la muestra resultante sea representativa de la entrada.
- Además del porcentaje especificado, la transformación Muestreo de porcentaje utiliza un algoritmo para determinar si se debe incluir una fila en la salida de ejemplo. Esto significa que el número de filas de la salida de ejemplo podría no reflejar exactamente el porcentaje especificado.
- La transformación Muestreo de porcentaje es especialmente útil para la minería de datos. Utilizando esta transformación, puede dividir de forma aleatoria un conjunto de datos en dos conjuntos de datos: uno para el entrenamiento del modelo de minería de datos y otro para probar el modelo.
- La transformación Muestreo de porcentaje también es útil para crear conjuntos de datos de ejemplo de desarrollo de paquetes. Si aplica la transformación Muestreo de porcentaje a un flujo de datos, puede reducir uniformemente el tamaño de los conjuntos de datos conservando sus características. El paquete de prueba podrá ejecutarse más rápido porque utilizará un conjunto de datos pequeño, pero representativo.

© JMA 2009. All rights reserved

161

Transformación Dinámica

- La transformación Dinámica transforma un conjunto de datos normalizado en una versión menos normalizada pero más compacta dinamizando los datos de entrada en un valor de columna.
- Cuando se dinamiza un conjunto de datos, las columnas de entrada ejecutan diferentes roles en el proceso de dinamización. Una columna puede participar de las siguientes maneras:
 - La columna se pasa sin cambios a la salida. Dado que varias filas de entrada pueden obtener una sola fila de salida, la transformación copia solamente el primer valor de entrada para la columna.
 - La columna funciona como la clave o parte de la clave que identifica un conjunto de registros.
 - La columna define la dinamización. Los valores de esta columna se asocian con columnas en el conjunto de datos dinamizado.
 - La columna contiene valores que se colocan en las columnas creadas por la dinamización.

© JMA 2009. All rights reserved

162

Transformación Anulación de dinamización

- La transformación Anulación de dinamización transforma un conjunto de datos sin normalizar en una versión más normalizada ampliando los valores de varias columnas de un solo registro en varios registros con los mismos valores en una sola columna.
- Por ejemplo, un conjunto de datos que enumera nombres de clientes tiene una fila para cada cliente, con los productos y la cantidad comprada en columnas dentro de la fila.
- Después de que la transformación Anulación de dinamización normaliza el conjunto de datos, el conjunto de datos contiene una fila diferente para cada producto que compró el cliente.

© JMA 2009. All rights reserved

163

Transformaciones de inteligencia empresarial

- Las siguientes transformaciones realizan operaciones de inteligencia empresarial tales como limpiar datos, realizar minería de texto y ejecutar consultas de predicción de minería de datos.
 - Agrupación aproximada: Transformación que estandariza los valores en los datos de columnas.
 - Búsqueda aproximada: Transformación que busca valores en una tabla de referencia mediante una coincidencia aproximada.
 - Extracción de términos: Transformación que extrae términos del texto.
 - Búsqueda de términos: Transformación que busca términos en una tabla de referencia y cuenta términos extraídos del texto.
 - Consulta de minería de datos: Transformación que ejecuta consultas de predicción de minería de datos.

© JMA 2009. All rights reserved

164

Transformación Dimensión de variación lenta

- La transformación Dimensión de variación lenta coordina la actualización e inserción de registros en las tablas de dimensiones de almacenamiento de datos. Por ejemplo, puede usar esta transformación para configurar las salidas de transformación que insertan y actualizan registros en la tabla DimProduct de la base de datos AdventureWorksDW2012 con datos de la tabla Production.Products de la base de datos OLTP AdventureWorks.
- La transformación Dimensión de variación lenta proporciona la siguiente funcionalidad para administrar dimensiones variables lentas:
 - Hacer coincidir filas entrantes con filas de la tabla de búsqueda para identificar filas nuevas y existentes.
 - Identificar las filas entrantes que contienen cambios cuando no se permiten cambios.
 - Identificar registros de miembros deducidos que requieren actualización.
 - Identificar filas entrantes que contienen cambios históricos que requieren la inserción de nuevos registros y la actualización de registros expirados.
 - Detección de filas entrantes que contienen cambios que requieren la actualización de los registros existentes, incluyendo los expirados.
- La transformación Dimensión de variación lenta admite cuatro tipos de cambios: atributo variable, atributo histórico, atributo fijo y miembro deducido.
 - Los cambios de atributo variable sobrescriben los registros existentes. Este tipo de cambio es equivalente a un cambio del Tipo 1. La transformación Dimensión de variación lenta dirige estas filas a una salida llamada Salida de actualizaciones de atributos variables.
 - Los cambios de atributo histórico crean nuevos registros en lugar de actualizar registros existentes. El único cambio que se permite en un registro existente es una actualización a una columna que indica si el registro está actualizado o expirado. Este tipo de cambio es equivalente a un cambio del tipo 2. La transformación Dimensión de variación lenta dirige estas filas a dos salidas: Salida de inserciones de atributos históricos y Nueva salida.
 - Los cambios de atributo fijo indican que el valor de la columna no debe cambiar. La transformación Dimensión de variación lenta detecta cambios y puede dirigir las filas con cambios a una salida llamada Salida de atributo fijo.
 - Un miembro deducido indica que la fila es un registro de miembro deducido en la tabla de dimensiones. Existe un miembro deducido cuando una tabla de hechos hace referencia a un miembro de dimensión que todavía no se han cargado. Un registro de miembro deducido mínimo se crea antes de los datos de dimensiones relevantes, que se proporcionan en una carga posterior de los datos de dimensiones. La transformación dirige estas filas a una salida llamada Salida de actualizaciones de miembros deducidos. Cuando se cargan datos del miembro deducido, se puede actualizar el registro existente en lugar de crear uno nuevo.

© JMA 2009. All rights reserved

165

Transformación Búsqueda de términos

- La transformación Búsqueda de términos compara los términos extraídos del texto en una columna de entrada de transformación con los términos de una tabla de referencia. A continuación, cuenta la cantidad de veces que un término aparece en la tabla de búsqueda en el conjunto de datos de entrada y escribe el recuento junto con el término de la tabla de referencia en columnas en la salida de transformación. Esta transformación resulta útil para crear una lista personalizada de palabras basada en el texto de entrada, que incluye estadísticas de frecuencia de aparición de palabras.
- Antes de que la transformación Búsqueda de términos realice una búsqueda, extrae palabras del texto en una columna de entrada aplicando el mismo método que la transformación Extracción de términos:
 - El texto se divide en frases.
 - Las frases se dividen en palabras.
 - Las palabras se normalizan.
- Para personalizar de forma adicional cuáles son los términos que deben coincidir, la transformación Búsqueda de términos puede configurarse para obtener coincidencias que distingan mayúsculas de minúsculas.
- La transformación Búsqueda de términos realiza una búsqueda y devuelve un valor aplicando las siguientes reglas:
 - Si la transformación se configura para obtener coincidencias que distingan mayúsculas de minúsculas, se descartan las coincidencias que no pasan la comparación con distinción de mayúsculas y minúsculas. Por ejemplo, estudiante y ESTUDIANTE se consideran palabras diferentes.
 - Si la forma plural del nombre o frase existe en la tabla de referencia, la búsqueda coincide solo con la forma plural del nombre o frase. Por ejemplo, se contarían todas las instancias de estudiantes independientemente de las instancias de estudiante.
 - Si solo se encuentra la forma singular de la palabra en la tabla de referencia, las formas singular y plural de la palabra o frase coinciden con la forma singular. Por ejemplo, si la tabla de búsqueda contiene la palabra estudiantey la transformación encuentra las palabras estudiante y estudiantes, ambas palabras se contarían como coincidencia del término buscado estudiante.
 - Si el texto de la columna de entrada es una frase lematizada, solo la última palabra en la frase se ve afectada por la normalización. Por ejemplo, la versión lematizada de citas con los médicos es cita con los médicos.

© JMA 2009. All rights reserved

166

Transformación Consulta de minería de datos

- La transformación Consulta de minería de datos realiza consultas de predicción en modelos de minería de datos. Esta transformación contiene un generador de consultas para crear consultas de Extensiones de minería de datos (DMX). El generador de consultas permite crear instrucciones personalizadas para evaluar los datos de entrada de la transformación en un modelo de minería existente mediante el lenguaje DMX.
- Una transformación puede ejecutar múltiples consultas de predicción si los modelos se generan a partir de la misma estructura de minería de datos.
- La transformación Consulta de minería de datos utiliza un administrador de conexiones de Analysis Services para conectar con el proyecto de Analysis Services o la instancia de Analysis Services que contiene la estructura de minería y los modelos de minería.
- Esta transformación tiene una entrada y una salida. No admite una salida de error.

© JMA 2009. All rights reserved

167

Transformación Búsqueda aproximada

- La Transformación Búsqueda aproximada realiza tareas de limpieza de datos como normalizar datos, corregir datos y proporcionar valores que faltan.
- La transformación Búsqueda aproximada difiere de la Búsqueda en su uso de coincidencia aproximada. La transformación Búsqueda utiliza una combinación de igualdad para localizar los registros que coinciden en la tabla de referencia. Devuelve los registros que tienen al menos un registro coincidente y devuelve registros que no tienen registros coincidentes. En cambio, la transformación Búsqueda aproximada emplea la coincidencia aproximada para devolver una o más coincidencias similares en la tabla de referencia.
- La transformación Búsqueda aproximada suele ir después de una transformación Búsqueda en un flujo de datos de paquete. Primero, la transformación Búsqueda intenta encontrar una coincidencia exacta. Si no lo consigue, la transformación Búsqueda aproximada proporciona coincidencias similares de la tabla de referencia.
- La transformación necesita acceso a un origen de datos de referencia que contenga los valores que se utilizan para limpiar y ampliar los datos de entrada. El origen de los datos de referencia debe ser una tabla de una base de datos de SQL Server. La coincidencia entre el valor en una columna de entrada y el valor en la tabla de referencia puede ser exacta o aproximada. Sin embargo, la transformación requiere que se configure al menos una coincidencia de columna para la coincidencia aproximada. Si desea buscar únicamente coincidencias exactas, utilice la transformación Búsqueda.

© JMA 2009. All rights reserved

168

Transformación Agrupación aproximada

- La transformación Agrupación aproximada realiza tareas de limpieza de datos, identificando filas de datos que probablemente se van a duplicar y seleccionando una fila de datos canónica para utilizarla en la normalización de los datos.
- La transformación de Búsqueda aproximada requiere una conexión a una instancia de SQL Server para crear las tablas temporales de SQL Server que necesita el algoritmo de la transformación para realizar su trabajo. La conexión debe establecerla un usuario que tenga permiso para crear tablas en la base de datos.

© JMA 2009. All rights reserved

169

Otras transformaciones

- Integration Services incluye las siguientes transformaciones para exportar e importar datos, agregar información de auditoría, contar filas y trabajar con dimensiones de variación lenta.
 - Exportar columna: Transformación que inserta datos desde un flujo de datos en un archivo.
 - Importar columna: Transformación que lee datos desde un archivo y los agrega a un flujo de datos.
 - Auditar: Transformación que coloca información sobre el entorno a disposición del flujo de datos en un paquete.
 - Recuento de filas: Transformación que cuenta filas a medida que se mueven por ella y almacena el recuento final en una variable.
 - Dimensión de variación lenta: Transformación que configura la actualización de una dimensión de variación lenta.
- Transformaciones personalizadas
 - También puede escribir transformaciones personalizadas.

© JMA 2009. All rights reserved

170

Transformación Auditar

- La transformación Auditar habilita el flujo de datos en un paquete para incluir datos sobre el entorno en el que se ejecuta el paquete. Por ejemplo, el nombre del paquete, el equipo y el operador se pueden agregar al flujo de datos.
 - El GUID que identifica la instancia de ejecución del paquete.
 - Identificador único del paquete.
 - Nombre del paquete.
 - La versión del paquete.
 - La hora a la que se inició la ejecución del paquete.
 - El nombre del equipo.
 - El nombre de inicio de sesión de la persona que inició el paquete.
 - El nombre de la tarea Flujos de datos a la que está asociada la transformación Auditar.
 - El identificador único de la tarea Flujo de datos.

© JMA 2009. All rights reserved

171

Transformación Recuento de filas

- La transformación Recuento de filas cuenta las filas a medida que pasan por un flujo de datos y almacena el recuento final en una variable.
- Se puede usar recuentos de filas para actualizar las variables utilizadas en scripts, expresiones y expresiones de propiedades. (Por ejemplo, la variable que almacena el número de filas puede actualizar el texto de un mensaje de correo electrónico para incluir el número de filas). La variable que se va a utilizar en la transformación Recuento de filas debe existir ya y debe estar dentro del ámbito de la tarea Flujo de datos a la que pertenece el flujo de datos con la transformación Recuento de filas.
- La transformación almacena el valor del recuento de filas en la variable únicamente después de que la última fila ya ha pasado a través de ella. Por lo tanto, el valor de la variable no está actualizado en el momento en que el valor actualizado se va a usar en el flujo de datos que contiene la transformación Recuento de filas. Puede usar la variable actualizada en un flujo de datos independiente.

© JMA 2009. All rights reserved

172

Transformación Exportar/Importar

- La transformación Exportar columna lee datos de un flujo de datos e inserta dichos datos en un archivo. Por ejemplo, si el flujo de datos contiene información de productos, como una foto de cada producto, puede usar la transformación Exportar columna para guardar las imágenes en archivos.
- La transformación Importar columna lee datos de archivos y agrega los datos a columnas de un flujo de datos. Un paquete puede utilizar esta transformación para agregar texto e imágenes almacenadas en archivos distintos a un flujo de datos. Por ejemplo, un flujo de datos que carga datos en una tabla que almacena información de productos puede incluir la transformación Importar columna para importar revisiones de clientes de cada producto desde archivos y agregar las revisiones al flujo de datos.

© JMA 2009. All rights reserved

173

Transformación Balanced Data Distributor (BDD)

- La transformación Balanced Data Distributor (BDD) aprovecha la capacidad de procesamiento simultáneo de las CPU actuales. Distribuye uniformemente los búferes de filas entrantes entre las salidas de subprocesos independientes. Al usar subprocesos independientes para cada ruta de acceso de los resultados, el componente BDD mejora el rendimiento de un paquete de SSIS en equipos con varios núcleos o varios procesadores.
- La transformación Balanced Data Distributor ayuda a mejorar el rendimiento de un paquete en un escenario que cumple las condiciones siguientes:
 - Entra una gran cantidad de datos en la transformación BDD. Si el tamaño de los datos es pequeño y solo un búfer puede contener los datos, no vale la pena usar la transformación BDD. Si el tamaño de los datos es grande y se necesitan varios búferes para almacenar los datos, BDD puede procesar eficazmente los búferes de datos en paralelo mediante subprocesos independientes.
 - Los datos se pueden leer a una velocidad mayor que la velocidad de proceso del resto del flujo de datos. En este escenario, las transformaciones que se realizan en los datos se ejecutan despacio en comparación con la velocidad a la que llegan los datos. Si el cuello de botella está en el destino, el destino debe ser paralelizable.
 - No es necesario que los datos estén ordenados. Por ejemplo, si los datos deben estar ordenados, no debe dividirlos mediante la transformación BDD.
- Hay que tener en cuenta que si el cuello de botella en un paquete de SSIS se debe a la velocidad a la que se pueden leer los datos del origen, el componente BDD no ayuda a mejorar el rendimiento. Si el cuello de botella en un paquete de SSIS se debe a que el destino no admite paralelismo, el BDD no es útil; sin embargo, puede realizar todas las transformaciones en paralelo y usar la transformación Unión de todo para combinar los datos de salida que salen de las diferentes rutas de acceso de los resultados de la transformación BDD antes de enviar los datos al destino.

© JMA 2009. All rights reserved

174

Transformación Componente de script

- El componente de script hospeda el script y permite a un paquete incluir y ejecutar código personalizado de script. Puede usar el componente de script en paquetes para los siguientes fines:
 - Aplicar varias transformaciones a los datos en lugar de usar varias transformaciones en el flujo de datos. Por ejemplo, un script puede sumar los valores de dos columnas y luego calcular el promedio de la suma.
 - Tener acceso a las reglas de negocios en un ensamblado .NET existente. Por ejemplo, un script puede aplicar una norma empresarial que especifica el intervalo de valores que son válidos en la columna Income .
 - Usar fórmulas personalizadas y funciones, además de las funciones y operadores que proporciona la gramática de la expresión Integration Services . Por ejemplo, validar números de tarjeta de crédito mediante la fórmula LUHN.
 - Validar datos de columna y omitir registros que contienen datos no válidos. Por ejemplo, un script puede evaluar si un valor de franqueo es razonable y omitir registros con valores extremadamente altos o bajos.
- El componente de script proporciona una manera fácil y rápida de incluir funciones personalizadas en un flujo de datos. Sin embargo, si planea reutilizar el código de script en varios paquetes, quizás resulte más conveniente programar un componente personalizado en lugar de usar el componente de script.

© JMA 2009. All rights reserved