

# DAT324 SQL Server 2005 Integration Service开发与部署

黄春雷  
高级顾问  
顾问咨询部  
微软（中国）有限公司



创新 · 远见 · 分享 · 协作

## Microsoft SQL Server 2005 商业智能平台



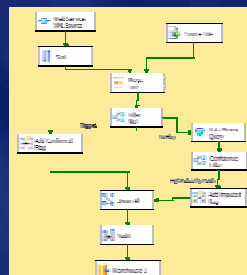
## 议程

目标：对SSIS有一个全面的了解，并能够开始轻松上手！

- 什么是Integration Service (SSIS)
- 与DTS的对比
- SSIS的主要工具集
- 实战SSIS新功能

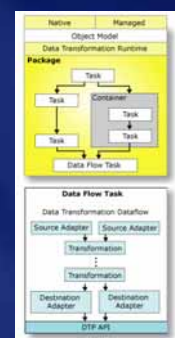
## 什么是SSIS?

- 一个全新的SQL Server商业智能应用组件
- 上一个版本的SQL Server中DTS的替代应用
- 新一代的高性能数据整合技术平台

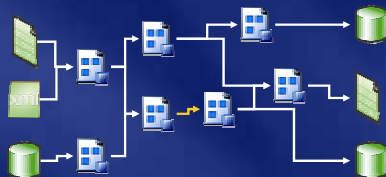


## SSIS的技术架构

- 控制流 (Control Flow)
  - 并行的工作流引擎
  - 执行“容器”与“任务”
- 数据流 (Data Flow)
  - SSIS的核心组件
  - 高性能的数据管道
  - 图形化的数据转移流程设计
  - 包括“源”，“目标”和“数据转换”等组件
  - 智能并行操作



## SSIS的运作方式



- 支持各种数据源，用户甚至可以自定义开发数据源适配器
- 数据转换组件通过各种方式对数据进行加工和处理
- 在数据转换过程中，数据流可以根据规则进行条件判断和路由
- 数据处理流程可以按照真实的业务规则进行复杂的定义
- 最后，数据可以被并行地加载到各种不同的目标数据源中

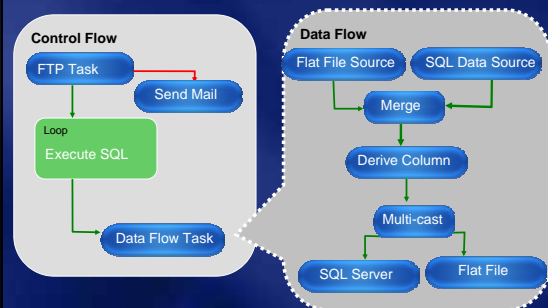
## 一些其他的说明.....

- 通过SQL Agent来进行调度
- “Package”是工作的基础单元
- 一个SSIS的Package是一个XML
  - Package可以保存在文件系统或者SQL Server中
- SSIS是一个独立的功能组件
- 与SQL Server的其他功能组件有很好的集成

## SSIS的特点

- 性能
  - 能够高效地处理海量并且复杂的数据流程
- 功能
  - 具有许多内建的数据源适配器和转换工具
  - 灵活的对象模型允许用户通过定制开发对其进行扩展
  - 可视化的开发与调试环境
- 智能
  - 智能处理复杂无规则的数据

## SSIS Package的主要组成部分



## 演示

### SSIS开发环境概貌

- 快速浏览SSIS的开发环境——BI Studio
- Control Flow & Data Flow

## 议程

目标：对SSIS有一个全面的了解，并能够开始轻松上手！

- 什么是Integration Service (SSIS)
- 与DTS的对比
- SSIS的主要工具集
- 实战SSIS新功能

DTS与SSIS的比较

	DTS	SSIS
数据源与数据目标	<ul style="list-style-type: none"><li>OLEDB</li><li>Access</li><li>Excel</li><li>dBase 5</li><li>HTML File (Source)</li><li>Paradox 5.X</li><li>Text File (Source)</li><li>Text File (Destination)</li><li>ODBC for Oracle</li></ul>	<ul style="list-style-type: none"><li>DataReader Source</li><li>Excel Source</li><li>Flat File Source</li><li>OLE DB Source</li><li>Raw File Source</li><li>Script Component Source</li><li>XML Source</li><li>Data Mining Model Training Destination</li><li>DataReader Destination</li><li>Dimension Processing Destination</li><li>Excel Destination</li><li>Flat File Destination</li><li>OLE DB Destination</li><li>Partition Processing Destination</li><li>Raw File Destination</li><li>Recordset Destination</li><li>Script Component Destination</li><li>SQL Server Mobile Destination</li><li>SQL Server Destination</li></ul>

DTS与SSIS的比较

	DTS	SSIS
数据传输与数据库维护	<ul style="list-style-type: none"><li>Copy SQL Server Objects</li><li>Transfer DB</li><li>Transfer Master SP</li><li>Transfer Job</li><li>Transfer Login</li><li>Transfer Error Messages</li></ul>	<ul style="list-style-type: none"><li>Transfer Objects</li><li>Transfer DB</li><li>Transfer Master SP</li><li>Transfer Job</li><li>Transfer Login</li><li>Transfer Error Messages</li><li>Back Up Database Task</li><li>Check Database Integrity Task</li><li>Execute SQL Server Agent Job Task</li><li>Execute T-SQL Statement Task</li><li>History Cleanup Task</li><li>Notify Operator Task</li><li>Rebuild Index Task</li><li>Reorganize Index Task</li><li>Shrink Database Task</li><li>Update Statistics Task</li></ul>

DTS与SSIS的比较

	DTS	SSIS						
数据转换	<ul style="list-style-type: none"><li>Transform Data</li><li>Data Driven</li></ul>	<table><tr><td>Business Intelligence<ul style="list-style-type: none"><li>Fuzzy Grouping Transformation</li><li>Fuzzy Lookup Transformation</li><li>Term Extraction Transformation</li><li>Term Lookup Transformation</li><li>Data Mining Query Transformation</li></ul></td><td>Row<ul style="list-style-type: none"><li>Character Map Transformation</li><li>Copy Column Transformation</li><li>Copy Column Transformation</li><li>Data Conversion Transformation</li><li>Derived Column Transformation</li><li>Script Component</li><li>OLE DB Command Transformation</li></ul></td></tr><tr><td>Rowset<ul style="list-style-type: none"><li>Aggregate Transformation</li><li>Sort Transformation</li><li>Percentage Sampling Transformation</li><li>Row Sampling Transformation</li><li>Pivot Transformation</li><li>Unpivot Transformation</li></ul></td><td>Split and Join<ul style="list-style-type: none"><li>Conditional Split Transformation</li><li>Multicast Transformation</li><li>Union All Transformation</li><li>Merge Transformation</li><li>Merge Join Transformation</li><li>Lookup Transformation</li></ul></td></tr><tr><td>Other<ul style="list-style-type: none"><li>Export Column Transformation</li><li>Import Column Transformation</li><li>Audio Transformation</li><li>Row Count Transformation</li><li>Slowly Changing Dimension Transformation</li></ul></td><td>Custom Transformations</td></tr></table>	Business Intelligence <ul style="list-style-type: none"><li>Fuzzy Grouping Transformation</li><li>Fuzzy Lookup Transformation</li><li>Term Extraction Transformation</li><li>Term Lookup Transformation</li><li>Data Mining Query Transformation</li></ul>	Row <ul style="list-style-type: none"><li>Character Map Transformation</li><li>Copy Column Transformation</li><li>Copy Column Transformation</li><li>Data Conversion Transformation</li><li>Derived Column Transformation</li><li>Script Component</li><li>OLE DB Command Transformation</li></ul>	Rowset <ul style="list-style-type: none"><li>Aggregate Transformation</li><li>Sort Transformation</li><li>Percentage Sampling Transformation</li><li>Row Sampling Transformation</li><li>Pivot Transformation</li><li>Unpivot Transformation</li></ul>	Split and Join <ul style="list-style-type: none"><li>Conditional Split Transformation</li><li>Multicast Transformation</li><li>Union All Transformation</li><li>Merge Transformation</li><li>Merge Join Transformation</li><li>Lookup Transformation</li></ul>	Other <ul style="list-style-type: none"><li>Export Column Transformation</li><li>Import Column Transformation</li><li>Audio Transformation</li><li>Row Count Transformation</li><li>Slowly Changing Dimension Transformation</li></ul>	Custom Transformations
Business Intelligence <ul style="list-style-type: none"><li>Fuzzy Grouping Transformation</li><li>Fuzzy Lookup Transformation</li><li>Term Extraction Transformation</li><li>Term Lookup Transformation</li><li>Data Mining Query Transformation</li></ul>	Row <ul style="list-style-type: none"><li>Character Map Transformation</li><li>Copy Column Transformation</li><li>Copy Column Transformation</li><li>Data Conversion Transformation</li><li>Derived Column Transformation</li><li>Script Component</li><li>OLE DB Command Transformation</li></ul>							
Rowset <ul style="list-style-type: none"><li>Aggregate Transformation</li><li>Sort Transformation</li><li>Percentage Sampling Transformation</li><li>Row Sampling Transformation</li><li>Pivot Transformation</li><li>Unpivot Transformation</li></ul>	Split and Join <ul style="list-style-type: none"><li>Conditional Split Transformation</li><li>Multicast Transformation</li><li>Union All Transformation</li><li>Merge Transformation</li><li>Merge Join Transformation</li><li>Lookup Transformation</li></ul>							
Other <ul style="list-style-type: none"><li>Export Column Transformation</li><li>Import Column Transformation</li><li>Audio Transformation</li><li>Row Count Transformation</li><li>Slowly Changing Dimension Transformation</li></ul>	Custom Transformations							

DTS与SSIS技术架构的不同

DTS	SSIS
数据转移→数据转换 (ELT)	数据转换→数据加载 (ETL)
串行/顺序处理	并行/异步处理
工作流驱动	数据流驱动
磁盘IO压力密集型	内存压力密集型
全部执行/全部未执行	事务管理下的断点续延
脚本编写	内置的灵活性
1:1的源:目标	n:n的源:目标 (合并/拆分/复制)
个人开发模式	团队开发模式

SSIS的新功能

- 流程控制
  - 可视化设计复杂的数据处理流程
  - 根据规则和条件进行分支判断
- 配置信息
  - XML
  - 数据库表
  - 环境变量
- 断点续延
  - Checkpoints
  - 事务
  - 分离级别

- 开发
  - Source Safe
  - Solution / Project
- 数据源
  - 在Package之间共享
- 部署
  - 文件系统
  - 数据库
- 功能扩展
  - 自定义适配器, 任务和转换功能

.....

DTS To SSIS

- 原则
  - Migrate what we can, wrap what we can't.
- 策略
  - 通过迁移工具将DTS Package导入SSIS中, 并保存成SSIS的格式, 允许用户修改和编辑
  - 通过迁移工具将DTS Package导入SSIS中, 但无法保存成SSIS的格式, 以Intermediate SQL Server 2000 Packages的形式保存在SSIS中, 用户不能直接修改和编辑
  - 无法导入SSIS中, 仍然维持原有的DTS格式和状态, 通过SQL Server 2000 DTS Runtime来运行

有时, 在SSIS中体验重新设计开发也十分简单有趣!

## 议程

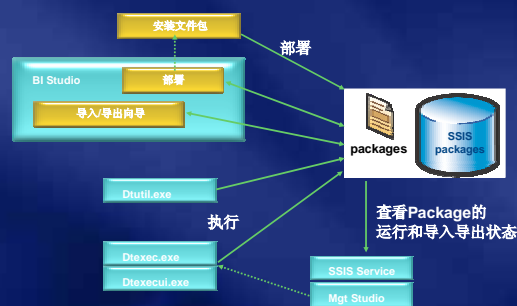
目标：对**SSIS**有一个全面的了解，并能够开始轻松上手！

- 什么是Integration Service (SSIS)
- 与DTS的对比
- **SSIS的主要工具集**
- 实战**SSIS**新功能

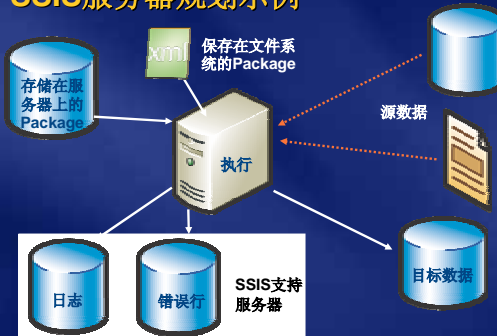
## SSIS生命周期管理工具/功能

- 设计SSIS Package
  - Business Intelligence Studio (BI Studio)
  - Package迁移向导
  - Visual Source Safe集成
- 部署与执行
  - 部署工具
  - 命令行执行工具 (dtutil.exe/dtutilui.exe)
  - 配置信息管理
- 运行与支持
  - 日志管理
  - 状态监控 (SQL Server Management Studio)
  - 断点管理

## SSIS工具集



## SSIS服务器规划示例



## 议程

目标：对**SSIS**有一个全面的了解，并能够开始轻松上手！

- 什么是Integration Service (SSIS)
- 与DTS的对比
- **SSIS的主要工具集**
- 实战**SSIS**新功能

## 实战SSIS新功能

- .Net功能组件集成
- 日志
- 断点续延
- 条件判断
- 配置管理
- **SSIS项目部署**

## 演示

### .Net功能组件集成

- 真正的.Net开发环境 —— Script Task
- .Net功能组件集成

## SSIS新功能介绍 —— 日志

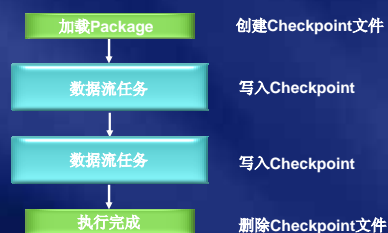
- 记录Package运行的详细的状态及结果
- 可以明细到每个Package，每个流程对象
- 允许选择日志字段
- 定义多个日志提供者（Log Provider）
- 在调试状态有“日志窗口”供开发人员随时观察

## 演示

### 日志功能

- 定义多个日志提供者（Log Provider）
- 记录日志
- 日志窗口

## 断点续延



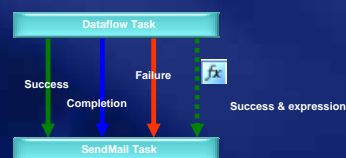
## 演示

### 断点续延

- 执行Package并强制出错
- 从断点处重新启动

## 条件判断

- 控制执行的顺序
- 状态：成功，失败，完成
- 属性表达式
- AND / OR



# 演示

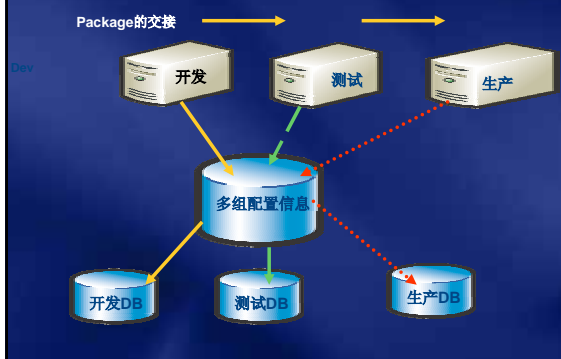
## 条件判断

- 定义不同的分支流程
- 根据不同的条件执行不同的流程

## 配置信息管理

- 对Packages进行动态的初始化
- 配置信息源的种类
  - SQL, XML文件, 注册表, 环境变量, 父Package
- 一个Package允许有多个配置信息
- 多个Package可以共享同一组配置信息

## 部署与配置信息



# 演示

## 配置信息管理

- 为Package定义配置信息
- 更改配置信息察看Package运行的变化

## 部署流程

- 用于方便地组织和拷贝Package及相关的文件的工具



# 演示

## SSIS项目部署

- 为SSIS项目制作安装文件
- 将SSIS项目部署到指定位置

## SQL Management Studio

- 使用SSIS service
- 可以监控正在执行的Package
- 维护Package的存储结构
- 随时手动执行Package

## 演示

### SQL Server Management Studio

- 监控Package的运行
- 观察Package的保存情况

## 总结

- 丰富的开发和管理功能
- 兼具灵活性和可扩展性
- 内建的安全机制和部署功能
- SQL Server功能组件的重要组成部分

## 社区资源

- SQL Server Integration Services public site – Links to blogs, whitepapers, etc.:  
<http://msdn.microsoft.com/SQL/sqlwarehouse/SSIS/default.aspx>
- SQL Server Business Intelligence public site:  
<http://www.microsoft.com/sql/evaluation/bi/default.asp>
- SSIS MVP community site:  
<http://www.sqlis.com>
- Newsgroups:  
Microsoft.private.sqlserver2005.dts
- TechNet and MSDN Webcasts
- 我的邮件地址:  
alberthu@microsoft.com

**Microsoft®**

您的潜力。我们的动力