

# SQL Server Integration Services

SEK 3 – Mai 2009  
Bernd Jungbluth

[www.berndjungbluth.de](http://www.berndjungbluth.de)



# SQL Server Integration Services

## Vorstellung

- Bernd Jungbluth
  
- IT-Erfahrung
  - SQL und Datenbanken seit 1991
  - SQL Server seit Version 7.0
  
- Freiberuflicher Berater und Entwickler
  - Administration, Entwicklung, Optimierung von und mit SQL Server
  - SQL Server Reporting Services
  - SQL Server Integration Services



# SQL Server Integration Services

## Agenda

- Einsatzmöglichkeiten der SSIS
- Konzept und Funktionsweise der SSIS
- Entwicklung von SSIS-Projekten
- Installation und Betrieb von SSIS-Projekten
- Migration von DTS nach SSIS



# SQL Server Integration Services

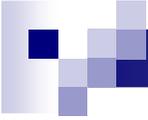
**Demo**



# SQL Server Integration Services

## Einsatzmöglichkeiten

- Wartungspläne
- Workflowlösungen
  
- Import-/Export-Assistent
- Import oder Export von Daten
  
- SQL Server Integration Services
- Workflow- und Datenintegrationslösungen



# SQL Server Integration Services

## Einsatzmöglichkeiten – Datenintegrationslösungen

- **Datenkonsolidierung**
  - Zusammenfassen von Daten aus unterschiedlichen Datenquellen in einer zentralen Datensammlung
  - Datawarehouse
  - „Unternehmensweite Datenkonsistenz“
- **Datentransformation**
  - Daten bereinigen
  - Daten aufbereiten
  - Daten veredeln
  - ETL



# SQL Server Integration Services

## Exkurs – Datawarehouse

### ■ Definitionen

- Bill Inmon (Vater des Datawarehousing):  
„eine themenorientierte, integrierte, chronologisierte und persistente Sammlung von Daten zur Unterstützung des Managements bei Entscheidungsprozessen“
- Ralph Kimball:  
„eine für Auswertungen speziell strukturierte Kopie transaktionaler Daten“
- Thomas Zeh:  
„ein physischer Datenbestand, der eine integrierte Sicht auf die zugrundeliegenden Datenquellen ermöglicht“

### ■ Wikipedia

- „zentrale Datensammlung deren Inhalt sich aus Daten unterschiedlicher Quellen zusammensetzt“



# SQL Server Integration Services

## Exkurs – ETL

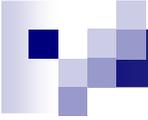
- Extraktion
  - Lesen der Daten aus den Datenquellen
  - Konvertieren der Daten in ein eigenes Datenformat
- Transformation
  - Syntaktische Transformation zur Datendarstellung
  - Semantische Transformation für den Datengehalt
- Laden
  - Konvertieren der Daten in die Struktur des Datenziels
  - Laden der transformierten Daten in die zentrale Datensammlung



# SQL Server Integration Services

## Agenda

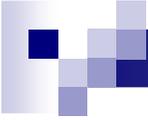
- Einsatzmöglichkeiten der SSIS
- ***Konzept und Funktionsweise der SSIS***
- Entwicklung von SSIS-Projekten
- Installation und Betrieb von SSIS-Projekten
- Migration von DTS nach SSIS



# SQL Server Integration Services

## Von DTS nach SSIS – DTS

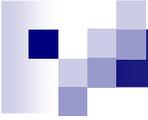
- DTS – Data Transformation Services
  - Eingeführt mit den OLAP-Diensten in SQL Server 7.0
  - Tool zum Füllen von Datawarehouse
  - Tool zur Transformation von Daten
  - Bewährtes Tool mit vielen bestehenden Installationen
- Nachteile
  - Keine moderne Entwicklungsumgebung
  - Unübersichtliche Verwaltung der DTS-Pakete
  - Unübersichtliche Datentransformationen
  - Umständliche Installation



# SQL Server Integration Services

## Von DTS nach SSIS – SSIS

- SQL Server Integration Services
  - Eingeführt mit SQL Server 2005
  - Neues Format mit SQL Server 2008
- Entwicklung
  - Keine Weiterentwicklung der DTS
  - Komplette Neuentwicklung
- Architektur
  - Integration Services Runtime Engine
  - Integration Services Data Flow Engine
  - Integration Services Service



# SQL Server Integration Services

## Von DTS nach SSIS – SSIS – Architektur

- Integration Services Runtime Engine
  - Prozess zur Steuerung des Arbeitsablaufs in SSIS
  - Beinhaltet den Start der Data Flow Engine
- Integration Services Data Flow Engine
  - Eigenständiger Prozess für den Datenfluss
  - Beinhaltet den eigentlichen ETL-Prozess
- Integration Services Service
  - Nicht zwingend erforderlich
  - Bietet nur unterstützende Funktionen
  - Liefert Informationen über aktuell laufende Pakete



# SQL Server Integration Services

## Von DTS nach SSIS – SSIS – Vorteile

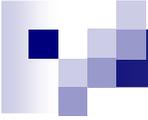
- Bessere Performance
- Neuer Entwicklungsansatz
  - Ausführen des ETL-Prozesses primär im Arbeitsspeicher
  - Reduzierung der Festplattenzugriffe
  - Schnelles Lesen und Schreiben von Daten
- Moderne Entwicklungsumgebung
  - Fehlerbehandlung bis hin zur Zeilenebene
  - Debugging und Source-Safe
  - Entwicklung von SSIS-Projekten
- Einfaches Setup und einfache Installation



# SQL Server Integration Services

## Begriffe

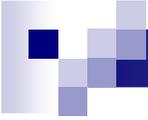
- Paket
- Ablaufsteuerung
- Datenfluss
- Ereignishandler
- Verbindungs-Manager
- Variablen
- Ausdrücke



# SQL Server Integration Services

## Begriffe – Paket

- Ausführbare Datei mit SSIS
  - DTSX-Datei im XML-Format
  - Ablage im Dateisystem oder in SQL Server
  - Entwicklung im Business Intelligence Development Studio
  - Ausführung mit Paketausführungs-Assistent
- Inhalt
  - Arbeitsablauf
  - Konfiguration der Datenverbindungen
  - Schemadaten der Datenquellen und Datenziele
  - Transformationsaufgaben



# SQL Server Integration Services

## Begriffe – Ablaufsteuerung

- **Arbeitsablauf des Pakets**
  - Abbildung einzelner Aufgaben zur Realisierung der Gesamtaufgabe
  - Startpunkt des Pakets
  - Endpunkt des Pakets
- **Inhalt**
  - Ausführung einzelner Aufgaben
  - Verbindung der einzelnen Aufgaben zu einem sequentiellen Ablauf
  - Eigenschaften, Variablen und Ereignishandler



# SQL Server Integration Services

## Begriffe – Ablaufsteuerung – Ablaufsteuerungselement

### ■ Aufgabe

- Ablaufsteuerungselement („Task“)
- Kann aus einer oder mehreren Teilaufgaben bestehen

### ■ Ablaufsteuerungselemente

- Beinhaltet die Beschreibung zur Ausführung einer Aufgabe
- Konfiguration mit Eigenschaften, Variablen und Ereignishandler
- Ausführung einer einzelnen Aufgabe
- Gruppierte Ausführung einer oder mehrerer Aufgaben im Sequenzcontainer
- Wiederholte Ausführung einer oder mehrerer Aufgaben im Schleifencontainer



# SQL Server Integration Services

## Begriffe – Ablaufsteuerung – Container

### ■ Definition

- Enthält mindestens ein Ablaufsteuerungselement
- Konfiguration mit Eigenschaften, Variablen und Ereignishandler
- Erbt die Konfiguration vorheriger Container
- Vererbt die eigene Konfiguration an nachfolgende Container

### ■ Einheitliche Struktur

- Paket und Ablaufsteuerung
- Ablaufsteuerungselement (= Taskhost-Container)
- Sequenz- und Schleifencontainer
- Ereignishandler



# SQL Server Integration Services

## Begriffe – Ablaufsteuerung – Rangfolgeeinschränkung

### ■ Definition

- Verbindung der Ablaufsteuerungselemente zu einem sequentiellen Ablauf
- Abhängig von Status und/oder Bedingungen

### ■ Farbliche Darstellung abhängig vom Status

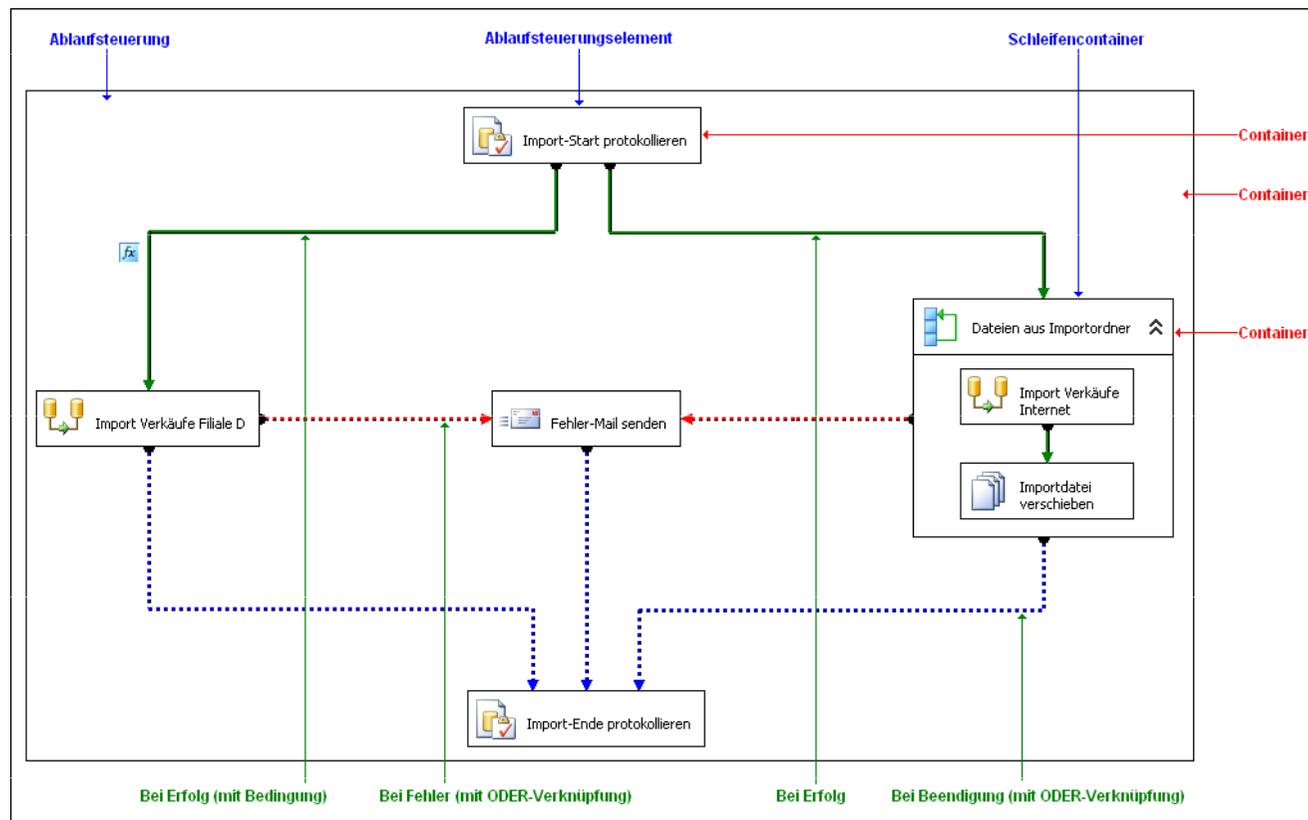
- Grün = Erfolg
- Rot = Fehler
- Blau = Beendigung

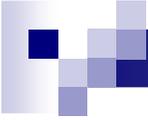
### ■ Darstellung abhängig der Bedingung

- Punkte = mehrere Oder-Verbindungen
- Symbol = Bedingung

# SQL Server Integration Services

## Begriffe – Ablaufsteuerung





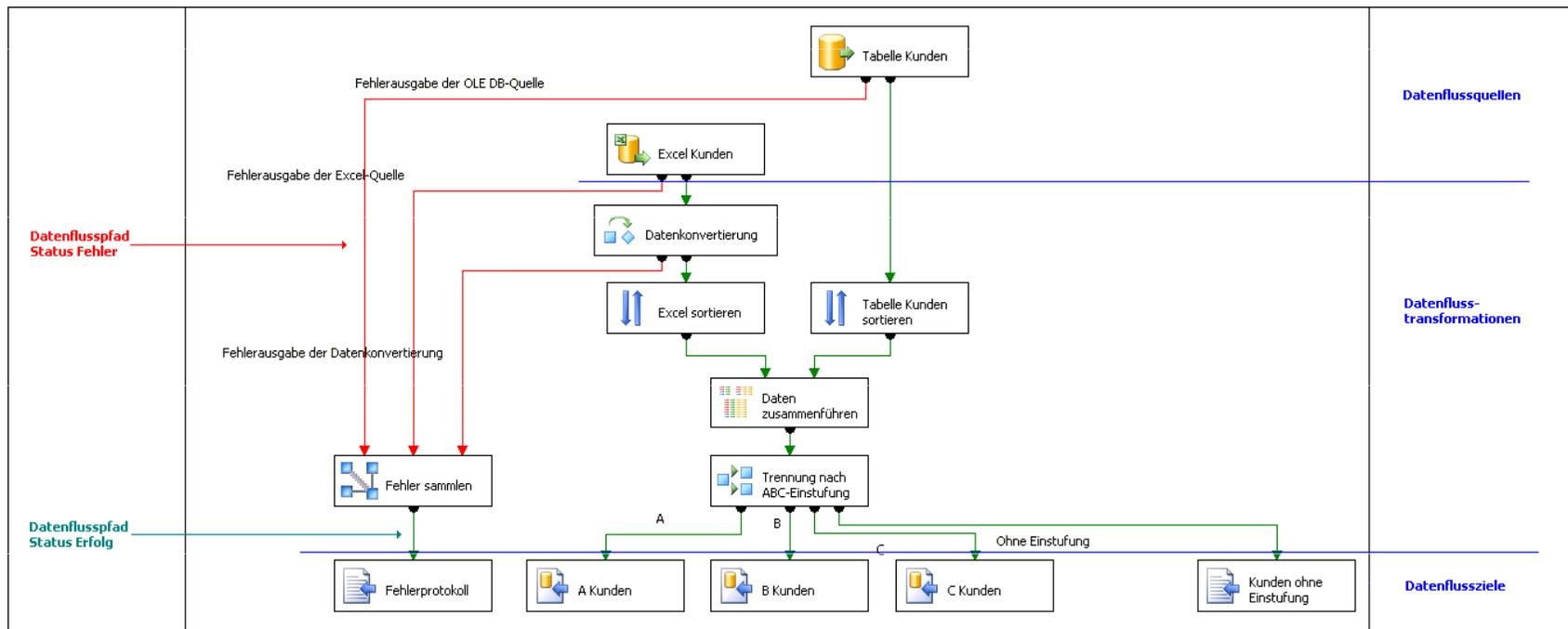
# SQL Server Integration Services

## Begriffe – Datenfluss

- ETL-Prozess in SSIS
  - Ein Ablaufsteuerungselement
  - Konfiguration in eigener Registerkarte
- Datenflusselemente
  - Datenflussquellen
  - Datenflusstransformationen
  - Datenflussziele
- Datenflusspfade
  - Verbindung der Datenflusselemente zu einem Datenfluss
  - Status: Erfolg, Fehler

# SQL Server Integration Services

## Begriffe – Datenfluss





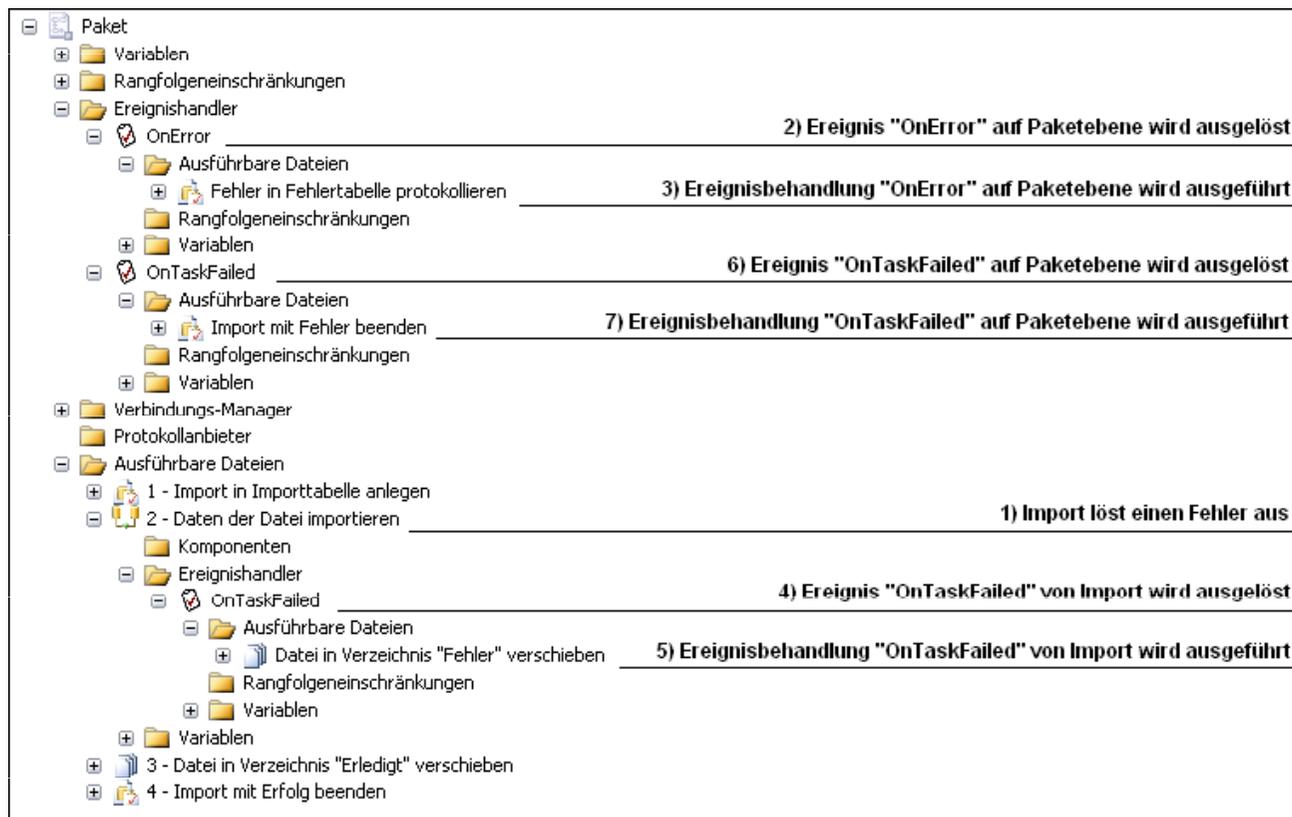
# SQL Server Integration Services

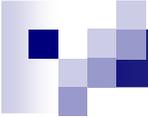
## Begriffe – Ereignishandler

- Ereignisse
  - Ereignishandler für zwölf verschiedene Ereignisse
  - Pro Container (Ablaufsteuerung bzw. Ablaufsteuerungselement)
- Ereignisbehandlung
  - Pro Ereignis in einem eigenen Container
  - Eigene Ablaufsteuerung mit Ablaufsteuerungselementen
  - Konfiguration mit Eigenschaften, Variablen und Ereignissen
- Verschachtelung
  - Gezielte Behandlung pro Ebene möglich
  - Einheitliche Behandlung pro Paket möglich

# SQL Server Integration Services

## Begriffe – Ereignishandler





# SQL Server Integration Services

## Begriffe – Verbindungs-Manager

### ■ Definition

- Konfiguration der Datenverbindung
- Beinhaltet nur den Connection-String
- Beinhaltet keinen Datenzugriff
- Speichert bei erster Ausführung bzw. beim Erstellen des Pakets die Verbindung im binären Format im Paket

### ■ Verwendung

- Ablaufsteuerungselemente
- Datenflussquellen, Datenflussziele, Datenflusstransformationen
- Erstellt Verbindung nur zur Ausführung eines Elements



# SQL Server Integration Services

## Begriffe – Variablen

### ■ Systemvariablen

- Beinhalten Informationen des Pakets zur Laufzeit
- Schreibgeschützte Werte
- Namespace „System“

### ■ Benutzervariablen

- Variablen zum Speichern von Ergebnissen
- Variablen zum Steuern von Eigenschaften
- Verschiedene Datentypen möglich
- Initialisierung bei Deklaration möglich
- Namespace „User“ oder eigene Namespaces



# SQL Server Integration Services

## Begriffe – Variablen

- Verfügbarkeit
  - Verbindungs-Manager
  - Ablaufsteuerungselemente
  - Rangfolgeneinschränkungen
  - Datenflussquellen, Datenflusstransformationen, Datenflussziele
  - Eigenschaften – direkt oder indirekt über Ausdrücke
- Gültigkeit einer Variablen
  - In der Ebene der Definition und den darauf folgenden Ebenen
  - Nicht in Ebenen, die über der Ebene der Definition liegen

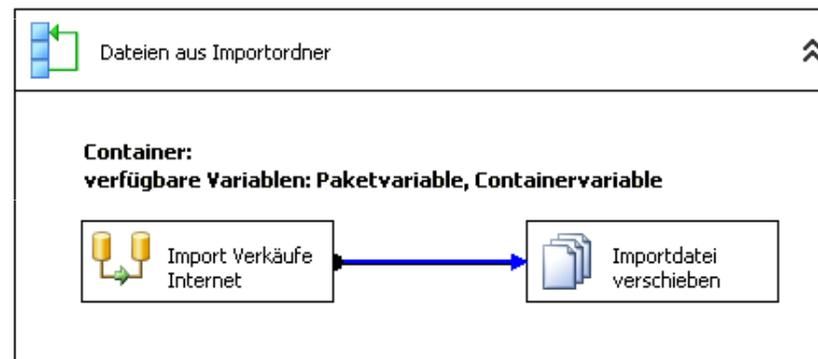
# SQL Server Integration Services

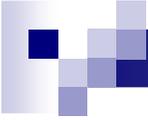
## Begriffe – Variablen



Name	Bereich	Datentyp	Wert
Containervariable	Dateien aus Importordner	Int32	0
Paketvariable	Test_Variablen	Int32	0

**Ablaufsteuerung:**  
**verfügbare Variablen: Paketvariable**





# SQL Server Integration Services

## Begriffe – Ausdrücke

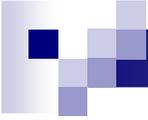
- Einzeilige Programmierung zur Ermittlung eines Werts
- Verwendung
  - Flexible Konfiguration von Eigenschaften zur Laufzeit
  - Erzeugen neuer Werte innerhalb des Datenflusses
- Ausdruckseditor
  - Entwicklungsoberfläche für Ausdrücke
  - Bietet viele Funktionen wie Typumwandlungen  
Mathematische Funktionen, Zeichenfolgenfunktionen, ...
  - Bietet Zugriff auf System- und Benutzervariablen
  - Bietet Zugriff auf Daten des Datenflusses



# SQL Server Integration Services

## Agenda

- Einsatzmöglichkeiten der SSIS
- Konzept und Funktionsweise der SSIS
- ***Entwicklung von SSIS-Projekten***
- Installation und Betrieb von SSIS-Projekten
- Migration von DTS nach SSIS



# SQL Server Integration Services

## Verbindungs-Manager

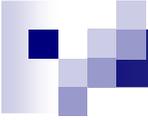
- Konfiguration
  - Verwaltung in eigener Registerkarte
  - 17 Verbindungs-Manager
- Primäre Verbindungs-Manager
  - Verfügbar über Kontextmenü der Registerkarte
  - Verfügbar über Dialog
- Sekundäre Verbindungs-Manager
  - Nur über Dialog verfügbar
  - Erweiterbar mit Verbindungs-Managern von Drittanbietern
  - Erweiterbar mit eigenen Verbindungs-Managern



# SQL Server Integration Services

## Verbindungs-Manager

- Datenbankschnittstellen
  - ADO.NET, OLE DB, ODBC, ADO
  - Zugriff auf Access, DB2, Informix, AS400, ...
- Dateien
  - Flatfile, Multiflatfile, Excel
  - XML über Verbindungs-Manager „Dateiverbindung“
- Weitere
  - Dateiverbindung, Multifile, Analysis Services, Cache
  - FTP, HTTP, SMTP
  - MSMQ, WMI, SMO



# SQL Server Integration Services

## Ablaufsteuerung

### ■ Definition

- Steuerung des Ablaufs rund um den ETL-Prozess
- Steuerung des Ablaufs einer Workflow-Lösung
- Verwaltung in eigener Registerkarte

### ■ Ablaufsteuerungselemente

- 27 Ablaufsteuerungselemente
- 3 Container
- 11 Wartungsplantasks
- Erweiterbar mit Ablaufsteuerungselementen von Drittanbietern
- Erweiterbar mit eigenen Ablaufsteuerungselementen



# SQL Server Integration Services

## Ablaufsteuerungselemente – Container

### ■ Sequenzcontainer

- Ablaufsteuerung innerhalb einer Ablaufsteuerung
- Enthält mindestens ein Ablaufsteuerungselement
- Definition eigener Eigenschaften, Variablen und Ereignishandler

### ■ For-Schleifencontainer

- Ausführung bis zum Eintreten einer definierten Bedingung
- Definition der Bedingung durch Variablen

### ■ For-Each-Schleifencontainer

- Ausführung anhand verschiedener Enumeratoren
- Dateien eines Verzeichnisses, Datensätze eines Datasets, ...



# SQL Server Integration Services

## Ablaufsteuerungselemente – Programmierung

- Skripttask
  - Visual Studio Tools for Applications (VSTA)
  - Visual Basic.NET und C#
  - Verwenden von Assemblies möglich
  - Einbinden von Webservices möglich
- Task ‚SQL ausführen‘
  - SQL-Anweisungen über verschiedene Datenverbindungen
  - Zur Datenmanipulation oder zur Datenermittlung
  - Ausführen von Gespeicherten Prozeduren möglich
  - Parameterübergabe möglich



# SQL Server Integration Services

## Ablaufsteuerungselemente

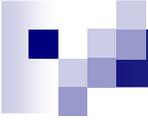
- Demo



# SQL Server Integration Services

## Ablaufsteuerungselemente – Daten und Dateien

- Task ‚Dateisystem‘
  - Verwalten von Dateien und Ordnern
  - Verschieben, Kopieren, Löschen, Erstellen und Umbenennen von Dateien und Ordnern, sowie das Festlegen von Attributen
- FTP-Task
  - Senden von Dateien
  - Empfangen von Dateien
- Task ‚Webdienst‘
  - Ausführen von Methoden eines Webservice
  - Ermitteln von Daten per Webservice



# SQL Server Integration Services

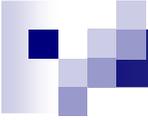
## Ablaufsteuerungselemente – Daten und Dateien

### ■ XML-Task

- Auswerten einer XML-Datei mit XPATH
- Zusammenführen zweier XML-Dateien
- Ermitteln der Unterschiede zweier XML-Dateien
- Prüfen der Daten einer XML-Datei anhand DTD oder XSD
- Aufbereiten der Daten einer XML-Datei mittels XSLT

### ■ Datenprofilerstellungs-Task

- Prüfen der Datenqualität einer Datenquelle
- Ergebnisse innerhalb der Ablaufsteuerung verfügbar
- Externe Auswertung der Ergebnisse mit dem Data Profile Viewer



# SQL Server Integration Services

## Ablaufsteuerungselemente – Workflow-Tasks

- Task ‚Paket ausführen‘
  - Ausführen eines Pakets in einem Paket
  - Kapselung von Funktionalität in Paketen
  - Wiederverwendung bestehender Pakete
- Task ‚Prozess ausführen‘
  - Ausführen eines Prozesses
  - Beispiel: Entpacken einer gepackten Datei
- Task ‚Mail senden‘
  - Senden von Mails per SMTP
  - Unterstützt nur anonyme und Windows-Authentifizierung



# SQL Server Integration Services

## Ablaufsteuerungselemente – Workflow-Tasks

- Task ‚Nachrichtenwarteschlange‘
  - Lesen und Abarbeiten von Nachrichtenwarteschlangen
  - Verteiltes Ausführen von Paketen
- Task ‚WMI-Datenleser‘
  - Lesen des Windows-Ereignisprotokolls
  - Ermitteln von Hardwarekomponenten und Softwareinstallationen
  - Per WMI-Query-Language (WQL)
- Task ‚WMI-Ereignisüberwachung‘
  - Protokollieren von Ereignissen des Betriebssystems
  - Reagieren auf Ereignisse des Betriebssystems



# SQL Server Integration Services

## Ablaufsteuerungselemente

- Demo



# SQL Server Integration Services

## Ablaufsteuerungselemente – Abwärtskompatibilität

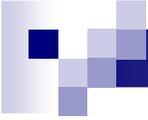
- Task ‚DTS 2000-Paket ausführen‘
  - Ausführen von DTS-Paketen eines SQL Server 2000
  - Ausführen von strukturierten Speicherdateien
  - Integration von DTS 2000-Paketen in SSIS
  - Übergabe von Variablen möglich
- Active X-Skript-Task
  - VB-Script und Java-Script
  - Keine Entwicklungsunterstützung
  - Einfaches Eingabefeld
- Laufzeitunterstützung für DTS erforderlich



# SQL Server Integration Services

## Ablaufsteuerungselemente – SQL Server-Tasks

- Tasks zur Übertragung von SQL Server-Objekten
  - Task ‚Anmeldungen übertragen‘
  - Task ‚Aufträge übertragen‘
  - Task ‚Datenbanken übertragen‘
  - Task ‚Fehlermeldungen übertragen‘
  - Task ‚In Master gespeicherte Prozeduren übertragen‘
  - Task ‚SQL Server-Objekte übertragen‘
- Konfigurationsmöglichkeiten
  - Dedizierte Auswahl der einzelnen Objekte
  - Optionales Überschreiben bestehender Objekte möglich



# SQL Server Integration Services

## Ablaufsteuerungselemente – Wartungsplantasks

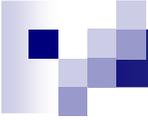
- Tasks zur Wartung von SQL Server-Datenbanken
  - Verwendung nicht nur in Wartungsplänen sinnvoll
  - 11 Wartungsplantasks
- Beispiele
  - Sichern der Ziel-Datenbank vor Datenimport
  - Reorganisieren der Indizes nach Datenimport
  - Überprüfen der Datenbankintegrität nach Datenimport
  - Aktualisieren der Statistiken nach Datenimport
  - Benachrichtigen des Operators per „Net Send“, Pager, eMail
  - Ausführen von Aufträgen des SQL Server-Agenten



# SQL Server Integration Services

## Ablaufsteuerungselemente – Analysis Services

- Analysis Services-Task ‚DDL ausführen‘
  - Erstellen, Ändern und Löschen von AS-Strukturen
  - Cubes, Partitionen, Dimensionen, Data Mining-Modellen
  - DDL mit ASSL über XMLA
- Analysis Services-Verarbeitungstasks
  - Erstellen und Aufbereiten von AS-Komponenten
  - Cubes, Partitionen, Dimensionen, Data Mining-Modellen
- Data Mining-Abfragetask
  - „Vorhersehe-Abfragen“ auf Data Mining-Modellen
  - Abfrage per DMX mittels Abfragedesigner



# SQL Server Integration Services

## Ablaufsteuerungselemente – Daten verarbeiten

### ■ Masseneinfügungstask

- bcp mit SSIS
- Schnelles Einlesen von Daten aus Textdateien in SQL Server
- Keine Datentransformation möglich
- Keine Datenprüfung möglich
- Übereinstimmung der Struktur von Quelle und Ziel erforderlich

### ■ Datenflusstask

- Der ETL-Prozess in SSIS
- Kontrollierte Datentransformation
- Eigene Entwicklungsoberfläche in eigener Registerkarte



# SQL Server Integration Services

## Werbung

- Workshop „SQL Server Integration Services“
- Inhalt:
  - Vertiefung dieses Vortrags
  - Beispiele und Übungen
- Erster Termin:
  - 17. Juni 2009
  - Hotel Birkenhof, 55469 Klosterkumbd
- Infos und Anmeldung unter
  - [workshop@berndjungbluth.de](mailto:workshop@berndjungbluth.de)
  - [www.sqlfaq.de/blog](http://www.sqlfaq.de/blog)



# SQL Server Integration Services

## Datenfluss

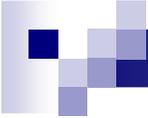
- Demo



# SQL Server Integration Services

## Datenfluss

- Der ETL-Prozess in SSIS
- Komponenten
  - Datenflussquellen
  - Datenflusstransformationen
  - Datenflussziele
- Funktionsweise
  - Die Daten der Datenquelle werden blockweise extrahiert.
  - Die Daten werden blockweise in SSIS-Datentypen konvertiert.
  - Die Daten werden blockweise transformiert.
  - Die Daten werden blockweise in das Datenziel geladen.



# SQL Server Integration Services

## Datenfluss – Datenflusspfade

- Definition
  - Verbindung der Datenflusselemente zu einem Datenfluss
  - In Abhängigkeit des Status „Erfolg“ oder „Fehler“
- Status „Erfolg“
  - Darstellung als grüner Pfeil
  - Mehrere Datenflusspfade bei Datenflusselementen möglich
- Status „Fehler“
  - Darstellung als roter Pfeil
  - Nur bei entsprechend konfigurierter Fehlerbehandlung verfügbar
  - Nicht bei allen Datenflusselementen verfügbar



# SQL Server Integration Services

## Datenflussquellen

### ■ Definition

- Das „E“ von ETL
- Extraktion von Daten einer Datenquelle
- Verwendet entsprechenden Verbindungs-Manager
- 7 Datenflussquellen

### ■ Konfiguration

- Zugriff über Tabelle, Sicht oder SQL-Anweisung
- Per SQL-Anweisung auch Zugriff über Gespeicherte Prozedur
- Auswahl der einzelnen Spalten für den Datenfluss



# SQL Server Integration Services

## Datenflussquellen

- Datenbankschnittstellen
  - ADO.NET-Quelle
  - OLE DB-Quelle
- Dateien
  - Excel-Quelle, XML-Quelle
  - Flatfilequelle, Rohdatendatei-Quelle
- Weitere
  - Skriptkomponente
  - Drittanbieter wie DB2, Oracle, SAP BI, Sybase, Teradata, ...
  - Erweiterbar mit eigenen Datenflussquellen



# SQL Server Integration Services

## Datenflussziele

### ■ Definition

- Das „L“ von ETL
- Laden der Daten in die Struktur des Datenziels
- Verwendet entsprechenden Verbindungs-Manager
- 13 Datenflussziele

### ■ Konfiguration

- Auswahl des Datenziels über eine Tabelle oder Sicht
- Zuweisung der einzelnen Spalten des Datenflusses zu den Spalten des Datenflussziels
- Abhängig vom Datenflussziel weitere Optionen verfügbar



# SQL Server Integration Services

## Datenflussziele

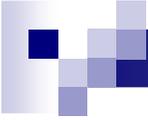
- Datenbankschnittstellen
  - ADO.NET Destination
  - OLE DB-Ziel
- Dateien
  - Excel-Ziel
  - Flatfileziel
  - Rohdatendatei-Ziel
- Arbeitsspeicher
  - DataReader-Ziel
  - Recordsetziel



# SQL Server Integration Services

## Datenflussziele

- SQL Server
  - SQL Server Compact-Ziel
  - SQL Server-Ziel
- Analysis Services
  - Data Mining-Modeltraining
  - Dimensions- und Partitionsverarbeitung
- Weitere
  - Skriptkomponente
  - Drittanbieter wie Trash Destination, DB2, Oracle, SAP BI, ...
  - Erweiterbar mit eigenen Datenflusszielen



# SQL Server Integration Services

## Datenflusstransformationen

### ■ Definition

- Das „T“ von ETL
- Transformation der Daten im Datenfluss
- Semantische wie syntaktische Transformationen
- 29 Datenflusstransformationen
- Erweiterbar mit Datenflusstransformationen von Drittanbietern
- Erweiterbar mit eigenen Datenflusstransformationen

### ■ Konfiguration

- Über entsprechende Konfigurations-Editoren
- Über zusätzliche „erweiterte Editoren“



# SQL Server Integration Services

## Datenflusstransformationen – Datenaufbereitung

### ■ Aggregat

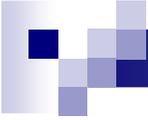
- Liefert Aggregationen von Datenspalten
- Mehrere parallele Aggregationen in einem Task möglich
- Liefert pro Aggregation einen Datenflusspfad

### ■ Pivot

- Aufbereiten von Spalten des Datenfluss zu einer Pivot-Tabelle
- Drehen von Zeilenwerten zu Spaltenwerten

### ■ Unpivot

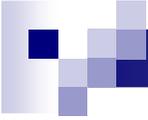
- Aufbereiten von Spalten einer Pivot-Tabelle zu Datensätzen
- Drehen von Spaltenwerten zu Zeilenwerten



# SQL Server Integration Services

## Datenflusstransformationen – Datenaufbereitung

- Datenkonvertierung
  - Konvertierung der Spalten in andere Datentypen
  - Datentyp, Länge, Genauigkeit, Dezimalstellen, Codepage
- Zeichenzuordnung
  - Mehrere kombinierbare Konvertierungsoptionen
  - Großbuchstaben, Kleinbuchstaben, Hiranga nach Katanga, ...
- Sortieren
  - Aufsteigend und absteigend über eine oder mehrere Spalten
  - Mehrere kombinierbare Vergleichsoptionen



# SQL Server Integration Services

## Datenflusstransformationen – Datenaufbereitung

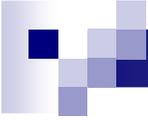
- Spalte importieren
  - Importieren von Daten aus dem Dateisystem in den Datenfluss
  - Funktioniert nur mit Spalten vom Typ „image“, „text“, „ntext“
- Spalte exportieren
  - Exportieren von Daten des Datenfluss ins Dateisystem
  - Funktioniert nur mit Spalten vom Typ „image“, „text“, „ntext“
- Spalte kopieren
  - Ergänzen des Datenfluss mit der Kopie einer Spalte
  - Beispiel: Kopie wird transformiert; Original bleibt erhalten



# SQL Server Integration Services

## Datenflusstransformationen – Datenaufbereitung

- Abgeleitete Spalte
  - Ergänzen des Datenfluss mit weiteren Spalten
  - Wertermittlung über Ausdruck aus Datenspalten bzw. Variablen
- Überwachung
  - Ergänzen des Datenfluss mit Laufzeitinformationen
  - Paket- und Task-Name, Startzeit der Ausführung, ...
- Suche
  - Lookup-Funktion
  - Ersetzen von Daten des Datenfluss mit dem Suchergebnis
  - Ergänzen des Datenfluss mit Spalten des Suchergebnis



# SQL Server Integration Services

## Datenflusstransformationen – Programmfluss

- Cachetransformation
  - Zwischenspeichern der Lookup-Daten in einem Cache
  - Optionales Zwischenspeichern des Caches im Dateisystem
- Bedingtes Teilen
  - Der CASE-Befehl im Datenfluss
  - Pro CASE-Zweig ein Datenflusspfad
  - Prüfen der CASE-Bedingung mittels Ausdruck
- Multicast
  - Aufteilung des Datenfluss in mehrere Datenflusspfade
  - Beispiel: Aufteilung in Originaldaten und zur Aggregation



# SQL Server Integration Services

## Datenflusstransformationen – Programmfluss

- Union All
  - Zusammenfassen mehrerer Datenflusspfade
  - Beinhaltet keine Sortierung der Daten
- Zusammenführen
  - Zusammenführen von Daten unterschiedlicher Datenquellen
  - Entspricht einem „Inner Join“
- Zusammenführungsverknüpfung
  - Zusammenführen von Daten unterschiedlicher Datenquellen
  - Entspricht einem „Inner Join“, „Left Join“ oder „Full Join“



# SQL Server Integration Services

## Datenfluss

- Demo



# SQL Server Integration Services

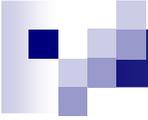
## Datenflusstransformationen – Programmierung

### ■ OLE DB-Befehl

- Ausführung einer Anweisung über eine OLE DB-Verbindung
- Verarbeitung der Anweisung pro Datensatz
- Zeilenweise, langsame Verarbeitung
- Nur für Datenmanipulationen

### ■ Skriptkomponente

- Visual Studio Tools for Applications (VSTA)
- Visual Basic.NET und C#
- Verwenden von Assemblies möglich
- Einbinden von Webservices möglich



# SQL Server Integration Services

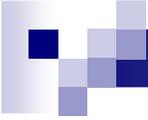
## Datenflusstransformationen – Analysis Services

### ■ Data Mining-Abfrage

- „Vorhersehe-Abfragen“ auf Data Mining-Modellen
- Abfrage per DMX mittels Abfragedesigner
- „Lookup“ auf Data Mining Modelle

### ■ Langsam veränderliche Dimensionen

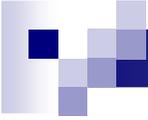
- Steuert die Verwaltung historischer Daten in Dimensionstabellen
- Überschreibt vorherigen Wert mit neuem Wert
- Speichert neuen Wert in neuem Datensatz und kennzeichnet vorherigen Wert als veraltet
- Konfiguration nur über Assistenten möglich



# SQL Server Integration Services

## Datenflusstransformationen – Mengenermittlung

- Prozentwert-Stichprobe
  - Zufällige Auswahl von Datensätzen über Prozentsatz
  - Datenflusspfade für Stichprobe und nicht ausgewählte Daten
- Zeilenstichprobe
  - Zufällige Auswahl von Datensätzen über definierte Anzahl
  - Datenflusspfade für Stichprobe und nicht ausgewählte Daten
- Zeilenanzahl
  - Ermitteln der Anzahl Datensätze im Datenfluss
  - Variable für Ergebnis erforderlich



# SQL Server Integration Services

## Datenflusstransformationen – Weitere

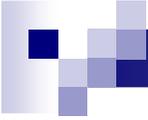
- Ausdrucksextrahierung und Ausdruckssuche
  - Extraktion von Ausdrücken in Spalten
  - Suche von Ausdrücken in Spalten
  - Funktioniert nur mit Unicode
  - Funktioniert nur mit englischen Texten
- Fuzzygruppierung und Fuzzysuche
  - Gruppierung nach ähnlichen Werten
  - Suche nach ähnlichen Werten
  - Liefert Aufstellung über Ähnlichkeit und Häufigkeit
  - Nur in Enterprise Edition verfügbar



# SQL Server Integration Services

## Datenfluss

- Demo



# SQL Server Integration Services

## Entwicklungsunterstützung

- Optische Darstellung bei Ausführung
  - Gelb = aktive Ausführung
  - Grün = erfolgreiche Ausführung
  - Rot = fehlerhafte Ausführung
  - Anzahl Datensätze im Datenfluss pro Datenflusspfad
- Informationen zur Ausführung
  - Registerkarte „Status“
  - Registerkarte „Ausführungsergebnisse“
  - Fenster „Ausgabe“
  - Beinhalten alle Meldungen zur Ausführung



# SQL Server Integration Services

## Entwicklungsunterstützung

### ■ Haltepunkte

- Haltepunkte an Ereignissen von Containern
- Werte der Variablen in den Fenster „Lokal“ oder „Überwachen“
- In Skripten des Skripttasks und der Skriptkomponente

### ■ Ausführungssteuerung

- Deaktivieren von Ablaufsteuerungselementen
- Ausführen einzelner Ablaufsteuerungselementen

### ■ Datenviewer

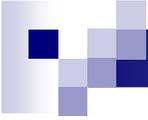
- Liefert eine Ansicht der Daten des Datenfluss
- Ausgabe in Tabellenform oder als Diagramm



# SQL Server Integration Services

## Entwicklungsunterstützung

- Demo



# SQL Server Integration Services

## Fehlerbehandlung

### ■ Datenfluss

- Aktivierung der Fehlerbehandlung am Datenflusselement
- Definition der Fehlerbehandlung pro Spalte möglich
- Fehlerbehandlung für „allgemeine Fehler“
- Fehlerbehandlung für „Fehler beim Abschneiden“
- Optionen „Fehler ignorieren“ oder „Zeile umleiten“
- Verfügbar bei fast allen Datenflusselementen

### ■ Ablaufsteuerung

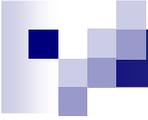
- Mittels der Ereignishandler „OnError“ und „OnTaskFailed“
- Verfügbar pro Container (Ablaufsteuerung bzw. Ablaufsteuerungselement)



# SQL Server Integration Services

## Fehlerbehandlung

- Demo



# SQL Server Integration Services

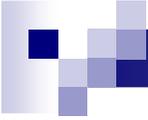
## Ereignishandler

### ■ Ereignisse

- Behandlung von zwölf verschiedenen Ereignissen je Container
- Ereignisse der Ablaufsteuerung
- Ereignisse der Ablaufsteuerungselemente
- Ereignisse für Datenflusstask, nicht für Datenflusselemente

### ■ Konfiguration

- Entwicklung in eigener Registerkarte
- Aktivierung des jeweiligen Ereignishandler erforderlich
- Ereignisbehandlung mit eigener Ablaufsteuerung
- Konfiguration mit Eigenschaften, Variablen und Ereignissen



# SQL Server Integration Services

## Ereignishandler

Ereignishandler	Ereignis
OnError	Bei einem Fehler
OnExecStatusChanged	Bei Änderung des Ausführungsstatus
OnInformation	Bei Ausgabe einer Informationsmeldung
OnPostExecute	Nach Ausführung eines Ablaufsteuerungselements
OnPostValidate	Nach Prüfung eines Ablaufsteuerungselements
OnPreExecute	Vor Ausführung eines Ablaufsteuerungselements
OnPreValidate	Vor Prüfung eines Ablaufsteuerungselements
OnProgress	Bei Ausführung eines Ablaufsteuerungselements
OnQueryCancel	Bei Abbruch eines Ablaufsteuerungselements
OnTaskFailed	Bei fehlerhafter Beendigung eines Ablaufsteuerungselements
OnVariableValueChanged	Bei Änderung des Werts einer Variablen
OnWarning	Bei Ausgabe einer Warnungsmeldung



# SQL Server Integration Services

## Ereignishandler

### ■ Funktionsweise

- Die Ereignisbehandlung ist ein „Paket im Paket“.
- Jeder Ereignishandler beinhaltet eine eigene Ablaufsteuerung.
- Jede Ablaufsteuerung verwendet Ablaufsteuerungselemente.
- Für jedes Ablaufsteuerungselement gibt es Ereignishandler.
- Jeder Ereignishandler beinhaltet eine eigene Ablaufsteuerung ...

### ■ Ausführungsreihenfolge

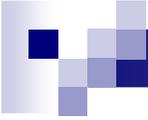
- Hierarchische Ausführung von der tiefsten zur höchsten Ebene
- Gezielte Ereignisbehandlung pro Ablaufsteuerungselement
- Einheitliche Ereignisbehandlung pro Ebene



# SQL Server Integration Services

## Agenda

- Einsatzmöglichkeiten der SSIS
- Konzept und Funktionsweise der SSIS
- Entwicklung von SSIS-Projekten
- ***Installation und Betrieb von SSIS-Projekten***
- Migration von DTS nach SSIS



# SQL Server Integration Services

## Setup und Installation

- Deployment
  - Zusammenfassen und Liefern aller notwendigen Dateien
  - Installation des SSIS-Projekts auf dem Zielsystem
- Herausforderung
  - Speichern aller notwendigen Projekt-Dateien im Zielsystem
  - Anpassen der Datenverbindungen an das Zielsystem
- Vorgehensweise
  - Erstellen einer Paketkonfiguration
  - Erstellen einer Manifest-Datei und eines Deployment-Ordnerns
  - Installation des SSIS-Projekts auf Zielsystem



# SQL Server Integration Services

## Setup und Installation – Paketkonfiguration

- Definition
  - Initialisierung bestimmter Eigenschaften beim Paketstart
  - Auswahl fast aller Eigenschaften des Pakets möglich
  - Bestandteil des Deployments
- Speichermöglichkeit für mehrere Eigenschaften
  - XML-Konfigurationsdatei
  - SQL Server-Datenbank
- Speichermöglichkeit für einzelne Eigenschaften
  - Umgebungsvariable
  - Registry-Eintrag



# SQL Server Integration Services

## Setup und Installation – Deployment

### ■ Konfiguration

- In Gruppe „Bereitstellungsprogramm“ der Projekteigenschaften
- Definition Manifest-Datei in Option „CreateDeploymentUtility“
- Definition Deployment-Ordner in Option „DeploymentOutputPath“

### ■ Erstellen

- Aufruf über Kontextmenü der Projektmappe
- Erzeugt Deployment-Ordner und Manifest-Datei

### ■ Installation

- Kopieren des Deployment-Ordners zum Zielsystem
- Starten der Installation mit der Manifest-Datei



# SQL Server Integration Services

## Setup und Installation – Installation

- Installation mit Paketinstallations-Assistent
- Schritt 1
  - Speichern der Pakete
  - Im Dateisystem oder im SQL Server
- Schritt 2
  - Speichern der Paketabhängigkeiten im Dateisystem
  - Speichern abhängiger Pakete im Dateisystem nicht möglich
- Schritt 3
  - Anpassen der Paketkonfiguration
  - Änderung freigegebener Eigenschaften an Zielumgebung



# SQL Server Integration Services

## Setup und Installation

- Demo



# SQL Server Integration Services

## Betrieb eines SSIS-Projekts

- Paketausführungs-Assistent
  - Konfigurationsoberfläche über „dtExecUI“
  - Eigentliche Ausführung per „dtExec“
- Manuell
  - Aus dem Dateisystem
  - Aus SQL Server
  - Aus Applikationen
- Automatisch
  - Per SQL Server Agent
  - Per Geplante Tasks



# SQL Server Integration Services

## Betrieb eines SSIS-Projekts

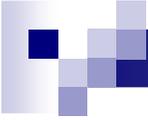
- Demo



# SQL Server Integration Services

## Agenda

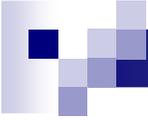
- Einsatzmöglichkeiten der SSIS
- Konzept und Funktionsweise der SSIS
- Entwicklung von SSIS-Projekten
- Installation und Betrieb von SSIS-Projekten
- ***Migration von DTS nach SSIS***



# SQL Server Integration Services

## Migration DTS nach SSIS

- Möglichkeiten
  - Migration von DTS-Paketen zu SSIS-Paketen
  - Ablaufsteuerungselement Task ‚DTS 2000-Paket ausführen‘
  - Verwendung von DTS
- Migration von DTS-Paketen zu SSIS-Paketen
  - Paketmigrations-Assistent verfügbar
  - Mehr Integration als Migration
  - Integration des DTS 2000 Paket in Ablaufsteuerungselement Task ‚DTS 2000-Paket ausführen‘
  - Migration nur bei DTS-Paketen mit simplem Datenimport
  - Manuelle Nacharbeit erforderlich



# SQL Server Integration Services

## Migration DTS nach SSIS

- Ablaufsteuerungselement ‚DTS 2000-Paket ausführen‘
  - Aufruf eines DTS 2000-Pakets in einem SSIS-Paket
  - Assimilation des DTS 2000-Pakets in SSIS möglich
  - Entwicklungsumgebung über DTS 2000 Designer verfügbar
- Verwendung von DTS
  - Auch unter SQL Server 2008 weiterhin verfügbar
  - Zu finden im SSMS unter „Verwaltung/Legacy“
  - Entwicklungsumgebung über DTS 2000 Designer verfügbar
  - DTS 2000 Runtime und DTS 2000 Designer notwendig
- Die beste „Migration“ ist die Neuentwicklung des Pakets.



# SQL Server Integration Services

## Das war noch nicht alles ...

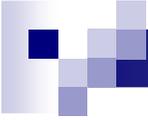
### ■ Entwicklung

- Beeinflussen der Ablaufsteuerung durch Eigenschaften
- Verwenden von Transaktionen in SSIS
- Verwenden von Change Data Capture und MERGE

### ■ Betrieb

- Paketsicherheit und Paketverschlüsselung
- Wiederanlauf von Paketen mittels Checkpoints
- Protokollieren von Ereignissen in Text- oder XML-Dateien, Windows-Ereignisprotokoll, SQL Server-Datenbank o. Profiler

### ■ Entwickeln eigener Komponenten



# SQL Server Integration Services

## Fazit

- Leistungsstarkes ETL-Tool
  - Schnell, stabil und skalierbar
  - Ernsthafte Konkurrenz zu anderen ETL-Tools
- Vielfältige Einsatzmöglichkeiten
  - Mehr als ein einfaches Import-/Export-Tool
  - Für Datenintegrationslösungen
  - Für Workflowlösungen
- Moderne Entwicklungsumgebung
  - Integration in Visual Studio und Source Safe
  - Einfache Installation auf Zielsystemen



# SQL Server Integration Services

## Links

- SQLServer2005\_BC.msi
  - Auf Installations-DVD
  - Beinhaltet DTS 2000 Runtime und DTS 2000 Designer
- Websites
  - [www.sqlis.de](http://www.sqlis.de)
  - [www.sql-ssis.de](http://www.sql-ssis.de)
  - [www.sqlis.com](http://www.sqlis.com) (Trash Destination)
  - MSDN: [msdn.microsoft.com/de-de/sqlserver/cc511477.aspx](http://msdn.microsoft.com/de-de/sqlserver/cc511477.aspx)
- PASS – die SQL Server Community: [www.sqlpass.de](http://www.sqlpass.de)
- Workshop-Termine unter [www.sqlfaq.de/blog](http://www.sqlfaq.de/blog)

# SQL Server Integration Services

Vielen Dank für die Aufmerksamkeit

