

Flexible Schnitt- stelle für Flat Files in das DWH

DOAG Konferenz 2012

Thomas Mauch
Consultant, Trivadis GmbH

20.11.2012, Nürnberg

BASEL BERN LAUSANNE ZÜRICH DÜSSELDORF FRANKFURT A.M. FREIBURG I.BR. HAMBURG MÜNCHEN STUTTGART WIEN

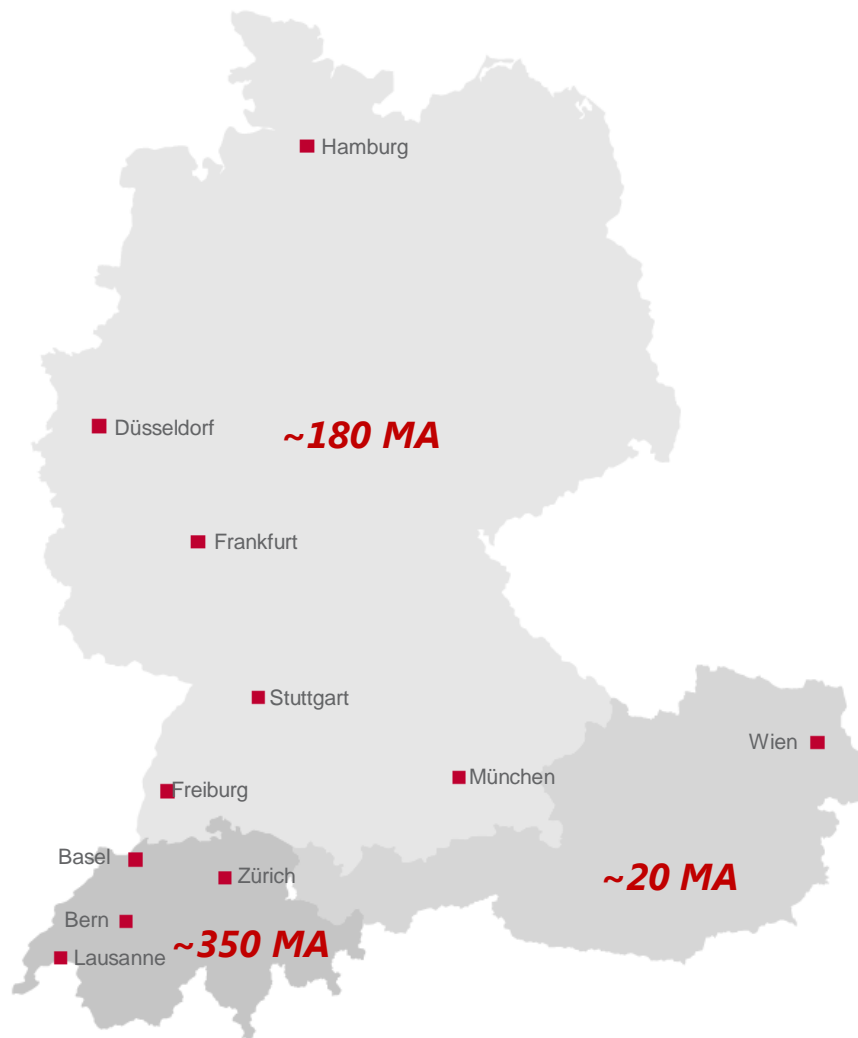
1

2012 © Trivadis

Flexible Schnittstelle für Flat Files in das DWH
20.11.2012

trivadis
makes IT easier. ■ ■ ■

Trivadis Facts & Figures



11 Trivadis Niederlassungen mit über 600 Mitarbeitern

Finanziell unabhängig und nachhaltig profitabel

Kennzahlen 2011

- Umsatz EUR 84 Mio.
- Dienstleistungen für über 700 Kunden in mehr als 1'800 Projekten
- Über 170 Service Level Agreements
- Mehr als 5'000 Trainingsteilnehmer
- Forschungs- und Entwicklungsbudget: EUR 3.6 Mio.

Das Besondere

Kundenindividuelle Lösungskompetenz und Herstellerunabhängigkeit

- bietet fundierte Methodenkenntnisse und eigenentwickelte Vorgehensweisen
- garantiert wiederholbare Qualität und Realisierungssicherheit

Technologiekompetenz

- hat über 17 Jahre Expertise in Oracle und Microsoft
- verfügt über ein eigenes Technology Center und setzt auf technologische Exzellenz

Lösungs- und Integrations-Know-how

- hat eine breite, branchenübergreifende Kundenbasis und jährlich über 1800 Projekte
- verbindet technologisches Spezialistenwissen mit dem Verständnis für die Business-Spezifika des Kunden

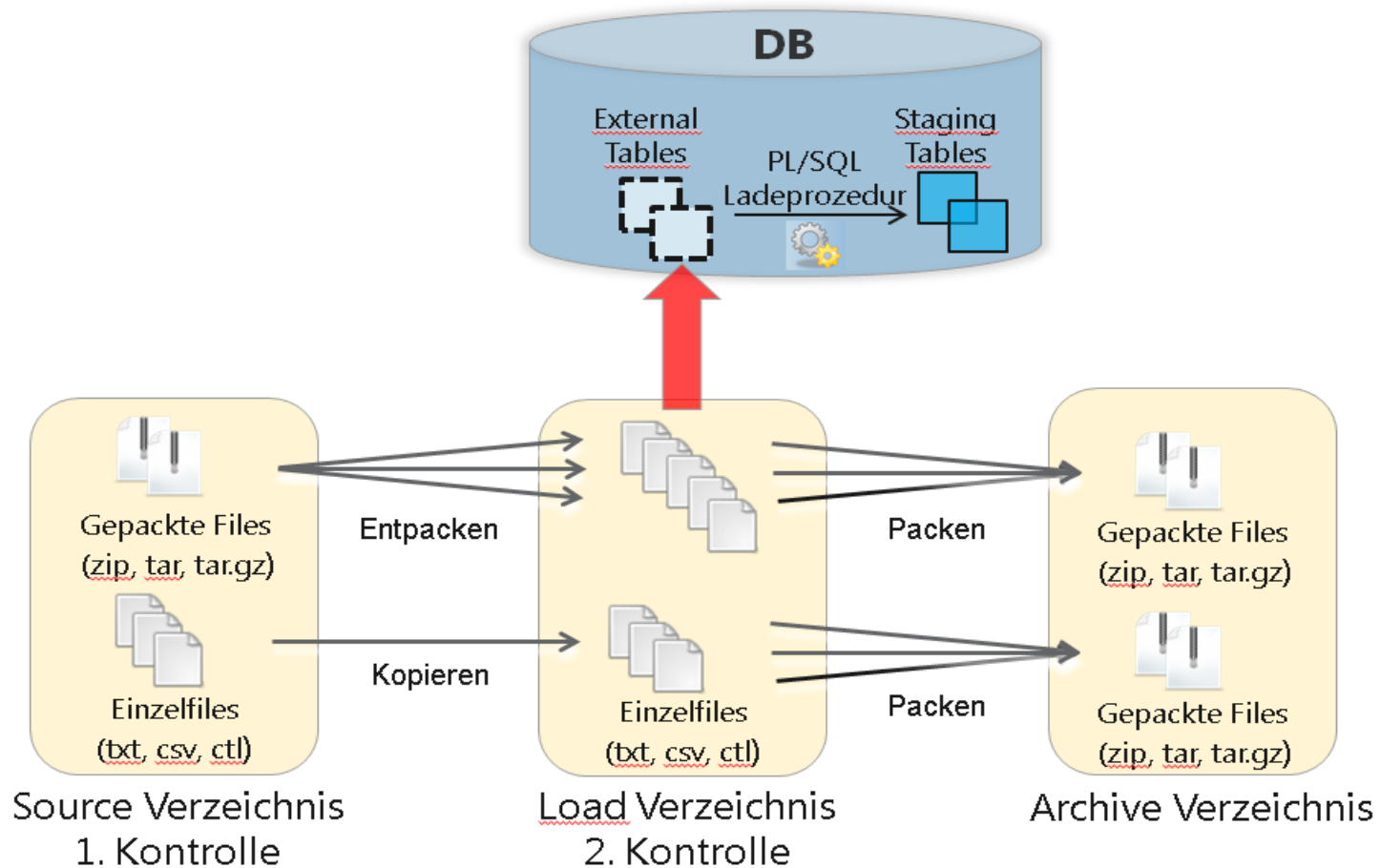
Begleitung über den gesamten IT-Projekt- Lifecycle

- begleitet den gesamten IT-Projekt-Lifecycle mit einem modularen Dienstleistungsportfolio
- bietet für jeden „Reifegrad“ die passende Dienstleistungs- und Lösungskombination

AGENDA

1. Architektur
2. Vorbereiten der Files
3. Laden der Files
4. Ergebniskontrolle
5. Demo
6. Zusammenfassung

Architektur



- Überblick über Prozesse beim Laden der Files

AGENDA

1. Architektur
2. Vorbereiten der Files
3. Laden der Files
4. Ergebniskontrolle
5. Demo
6. Zusammenfassung

Konfiguration Vorbereitung Files

■ Source Files

- Gepackte oder einzelne Files
- Periodische Überprüfung falls File fehlt (Scheduler Job)



■ Load Files

- Entpackte oder kodierte Source Files
- Abbruch falls File fehlt
- Registrierung falls File vorhanden



■ Zuordnung zu Filegruppen

- Themengebiete (bspw. „Kunden“)
- Vollständigkeit wird geprüft – erst wenn alle Files zu einer Gruppe da sind wird geladen

■ Zuordnung zu Lieferanten (optional)

- Ein File kommt mehrfach aus verschiedenen Quellen

Vorbereitung der Files

- Keine Verarbeitung per Skript außerhalb DB nötig
- Verarbeitung der Files mit der DB
 - PL/SQL (utl_file Package)
 - Lesen, Schreiben, Kopieren, Löschen
 - Überprüfung von Existenz und Größe
 - Java in der DB
 - Liste der Files in Verzeichnis (Name, Eigenschaften)
 - Packen und Entpacken von Winzip und Gzip Archiven
 - Host Call, z.B. Start von SQLLoader
 - Java in DB importiert
 - Open Source Library JTAR: Packen und Entpacken von TAR Archiven

AGENDA

1. Architektur
2. Vorbereiten der Files
3. Laden der Files
4. Ergebniskontrolle
5. Demo
6. Zusammenfassung

Konfiguration Laden der Files

- Schnittstelle
 - Load Files mit gleicher Struktur
 - Zuordnung zu einer Staging Table

- Versionierung
 - Ein File kann in mehreren Versionen existieren (mehr/weniger/breitere Spalten)
 - Die Definition der Versionen wird in Tabellen abgespeichert
 - Kann vor Ladeprozess festgelegt werden
 - Oder wird im Ladeprozess automatisch erstellt

Dynamisches Laden der Files

- Strukturinformation in SQLLoader Control File (ctl – File)
 - Name Staging Table
 - Spaltenname / Datentypen (optional)
 - INTEGER_EXTERNAL(length) -> NUMBER(length)
 - DECIMAL_EXTERNAL(length) -> NUMBER
 - CHAR(length) -> VARCHAR2(length)
 - DATE(length) -> DATE
 - Sonstiges -> VARCHAR2(4000)
- Alternativ: zusätzlich CREATE TABLE Skript für Staging Table (sql – File)
 - Unterstützte Datentypen: NUMBER(prec, scale), CHAR(length), VARCHAR2(length) und DATE
- Abgleich mit Version der Schnittstelle in Datenbank

Dynamisches Laden der Files

- Bei neuer Version:

- Anpassen / Anlegen Staging Table
- Erzeugen External Table Skript mit SQLLoader:

```
sqlldr control = ... .ctl external_table=GENERATE_ONLY log= ... .log
```

- Anpassen und Ausführen von Skript in DB

- Laden der Files

- Befehl generieren + ausführen: INSERT /*+ APPEND */ INTO ..
- Kontrolle von Badfile und Errorlogtable

Statisches Laden der Files

- Kein ctl – File vorhanden
 - Verwende die aktuellste Version der Schnittstelle in DB
 - Staging Table und External Table müssen in dieser Version angelegt sein
- Laden der Files
 - Wie beim dynamischen Laden

AGENDA

1. Architektur
2. Vorbereiten der Files
3. Laden der Files
4. Ergebniskontrolle
5. Demo
6. Zusammenfassung

Ergebniskontrolle

- Protokollierung der Ladeläufe in Logtabellen
 - Tabelle für Ergebnis Filegruppe
 - Status Ladeprozess
 - Evtl. Name von fehlendem Source / Load File
 - Tabelle für Ergebnis Einzelfiles
 - Status Ladeprozess
 - Schnittstelle und Version mit der geladen wurde
 - Anzahl geladener Datensätze
 - Evtl. Name von Badfile und Errorlogtable
 - Tabelle für Einzelschritte beim Laden

| MSG_PROC | MSG_TEXT | ADDO... | ADDON_MSG |
|--------------------|--|---------|--|
| f_execCmd | Start: Ausführung des Kommandos: C:\Oracle\product\... | 0 | |
| extTableDefinition | OTHERS at point: execmd | -20101 | ORA-20101: Fehler bei Aufruf von SQLLoader 1 |

Ergebniskontrolle

- Protokollierung in Staging Table
 - Referenz auf Logtabellen
 - Schnittstelle + Version mit der geladen wurde
 - Filedatum
- Verteilung von Nachrichten per Email
 - Konfigurierbar

AGENDA

1. Architektur
2. Vorbereiten der Files
3. Laden der Files
4. Ergebniskontrolle
5. Demo
6. Zusammenfassung

Demos

- Fall 1: Laden von File mit ctl-File
 - Source File st_products_20120109.zip
 - Load Files st_products_20120109.txt; st_products_20120109.ctl
- Fall 2: Laden von File mit CTL-File mit geänderter Struktur (neue Version)
 - Source File st_products_20120111.zip
 - Load Files st_products_20120111.txt; st_products_20120111.ctl
- Fall 3: Laden von File ohne CTL-File: das File wird mit der neuesten Version (höchste VersionsNr) eingespielt
 - Source File st_products_20120112.zip
 - Load File st_products_20120112.txt

Demos

- Fall 4: Laden von Files mit 2 Lieferanten
 - Source Files: cust_mun_20120627.zip; cust_stut_20120627.zip
 - Load Files: cust_mun_20120627.txt; cust_stut_20120627.ctf;
cust_mun_20120627.txt; cust_stut_20120627.ctf

- Fall 5: Laden von Files mit 2 Lieferanten und Fehlern in den Daten
 - File 1 (mun): Eintrag in Errorlogtable (ungültige Zahl)
 - File 2 (stut): erzeugt Badfile (ungültiges Datum)
 - Source Files: cust_mun_20120628.zip; cust_stut_20120628.zip
 - Load Files: cust_mun_20120628.txt; cust_stut_20120628.ctf;
cust_mun_20120628.txt; cust_stut_20120628.ctf

AGENDA

1. Architektur
2. Vorbereiten der Files
3. Laden der Files
4. Ergebniskontrolle
5. Demo
6. Zusammenfassung

Zusammenfassung

- Flexible Konfiguration
 - Durch Einträge in Tabellen
- Funktionalität komplett in der Datenbank abgelegt
 - Dadurch gute Einbindung in ETL-Prozess
- Dynamische Anpassung an Strukturänderungen der Files
 - Strukturinformation in ctl – File (und optional sql – File)
- Vollständige Kontrolle über Ladeprozess
 - Alle Files angeliefert
 - Alle Files vollständig geladen
- Ergebnisprotokollierung in Datenbank
- Benachrichtigung per Email (konfigurierbar)
- Einbindung möglich in **biGenius™**



DEN TRIVADIS-
~~STAND GÄNDEN~~
SIE AUF

**EBENE 3,
STAND NR. 304**

Trivadis GmbH

Thomas Mauch

Industriestraße 4
70565 Stuttgart

Tel. +49-711-90 36 32 30

Fax +49-711-90 36 32 59

info@trivadis.com

www.trivadis.com

BASEL BERN LAUSANNE ZÜRICH DÜSSELDORF FRANKFURT A.M. FREIBURG I.BR. HAMBURG MÜNCHEN STUTTGART WIEN

22

2012 © Trivadis

Flexible Schnittstelle für Flat Files in das DWH
20.11.2012

trivadis
makes IT easier. ■ ■ ■