



# Realtime Data Warehouse

Auslagern und Wiederherstellen operativer Daten

Steffen Kantorek

Hamburg, 21.02.2012



# Agenda

---

1. Ausgangssituation
2. Herausforderung
3. Prozess
4. Fazit



# Ausgangssituation

## Herausforderung und Motivation



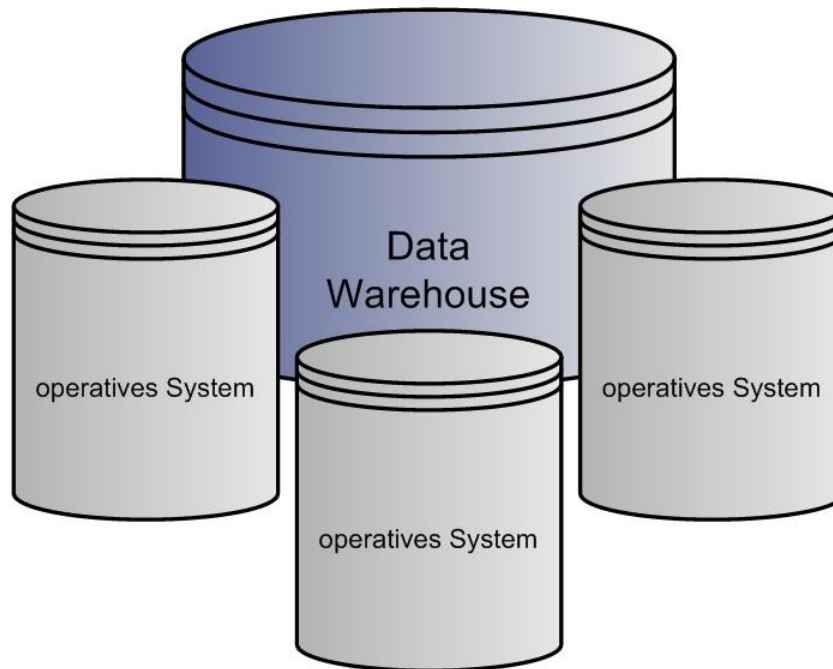
# Ausgangssituation die Systeme

## operatives System

- ausführen der Geschäftsprozesse
- hohe Performance

## Data Warehouse

- Speicherort für historische Daten
- Ausgelegt für die Verwaltung großer Datenmengen
- dient der Archivierung, dem Reporting ...



## Eckdaten

- Wertpapierabwicklungssystem mit einem durchschnittlichen Datenaufkommen von 1 Mio. Datensätzen am Tag
- ca. 1% manuelle Änderungen pro Tag
- 95% der Änderungen innerhalb der ersten sieben Tage
- keine technische Früherkennung möglich, welche Daten geändert werden müssen
- Daten werden auf das Data Warehouse ausgelagert



### **Idee: Daten aus dem Data Warehouse wiederherstellen!**

- Vorteile beider Systeme nutzen
  - Geschwindigkeit des operativen Systems
  - Handling von großen Datenmengen im Data Warehouse
- Herausforderung: Wiederherstellung der Daten zu einem beliebigen Zeitpunkt



Auslagern

Reorganisieren

Wiederherstellen

#### realtime Auslagerung in das Data Warehouse

Beim Auslagern der Daten in das Data Warehouse werden die Daten mittels Oracle Streams Realtime an das Data Warehouse übertragen.

- Capture von Insert und Update Statements
- Übertragung als Logical Change Record (LCR) an das Data Warehouse

Im Data Warehouse werden die Daten beim Apply an die Struktur des Systems angepasst.

- Trigger auf den Zieltabellen, welche auch bei DML Operationen zünden („DBMS\_DDL“-Package Firing Once = False)
- Auflösen von Foreign Key Beziehungen zum Reduzieren von Zugriffszeiten

#### Hinweis

Zu diesem Zeitpunkt sind auf dem operativen System und dem Data Warehouse die Daten redundant vorhanden.



## Der Prozess

### Reorganisieren der Daten im operativen System

Auslagern

Reorganisieren

Wiederherstellen

#### Reorganisieren der Daten im operativen System

Um eine stabile Performance im operativen System zu gewährleisten, müssen die Daten aus dem operativen System reorganisiert werden.

- Reorganisieren von Daten, die älter sind als sieben Tage
- Kann aufgrund der Datenredundanz jedoch zu einem frei wählbaren Zeitpunkt erfolgen

#### Hinweis

Löschungen werden im Capture Prozess nicht berücksichtigt und somit nicht an das Data Warehouse übertragen.



Auslagern

Reorganisieren

Wiederherstellen

#### Wiederherstellen der Daten für das operative System

Werden Daten zur nachträglichen Bearbeitung im operativen System benötigt, müssen diese manuell wiederhergestellt werden.

- kein Capture Prozess im Data Warehouse
- Aufruf einer Prozedur zum Erzeugen eines Logical Change Record(LCR)
- Aufbau des LCR in der Struktur des Zielsystems
- Propagieren des LCR in der korrekten Reihenfolge um Constraint Verletzungen zu verhindern

Es besteht darüber hinaus die Möglichkeit im operativen System einen DML-Error Handler zu definieren, welcher Fehler beim Wiederherstellen im operativen System verwaltet.

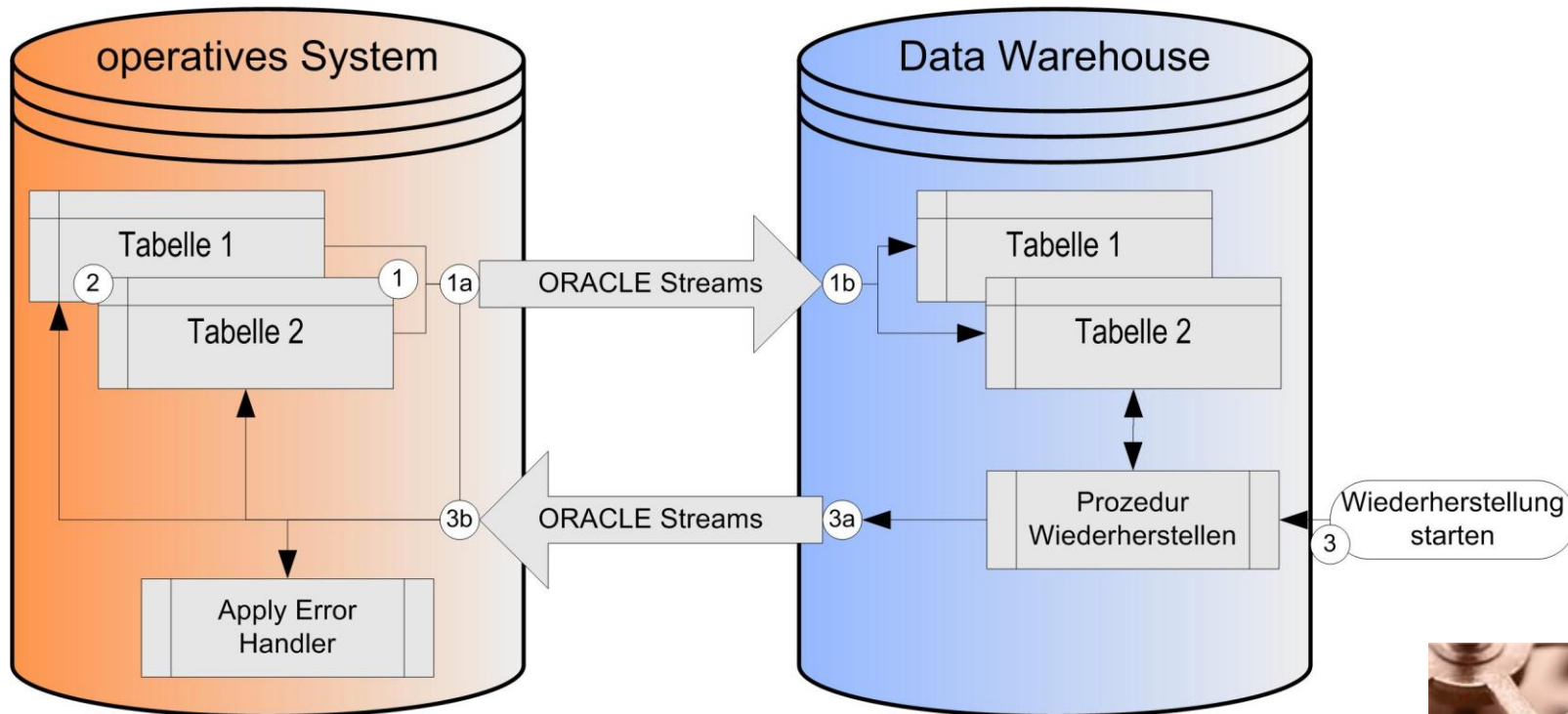
#### Hinweis

Aufgrund definierter Regeln findet beim Wiederherstellen keine erneute Übertragung vom operativen System in das Data Warehouse statt.





## Realtime Data Warehouse – Auslagern und Wiederherstellen operativer Daten



- 1) Realtime Auslagerung in das Data Warehouse
  - 1a) Daten Capturen und an das Data Warehouse übertragen
  - 1b) Anpassen der Daten auf die Struktur im Data Warehouse
- 2) Reorganisation der Daten im operativen System
- 3) Wiederherstellen der Daten für das operative System
  - 3a) Übertragung der Daten in das operative System
  - 3b) Persistieren der Daten im operativen System und ggf. vorhandene Fehler verwalten



## Fazit

- komfortable Möglichkeit, Daten nachträglich zu bearbeiten
- Datenmenge im operativen System wird gering gehalten
- unterschiedliche Datenstrukturen sind kein Problem
- Streams Kenntnisse werden benötigt

Achtung ist geboten bei große Datenmengen innerhalb einer Transaktion. Diese können dafür Sorgen, dass die Replizierung sehr lange dauert. Daher sollte man sich genau überlegen welche Daten benötigt werden.

### Hinweis:

Weiterführende Informationen mit Code Beispielen gibt es in der aktuellen DOAG News (Nr.1, Februar 2012).



Vielen Dank für Ihre Aufmerksamkeit

Steffen Kantorek  
Softwareentwickler  
Berenberg Bank

