

# Performance by Design

Wie werden performante ETL-Prozesse entwickelt?

**Reinhard Mense**

Geschäftsführer ARETO Consulting

**ARETO**  
*consulting*

# Die Herausforderung

Große Datenmengen

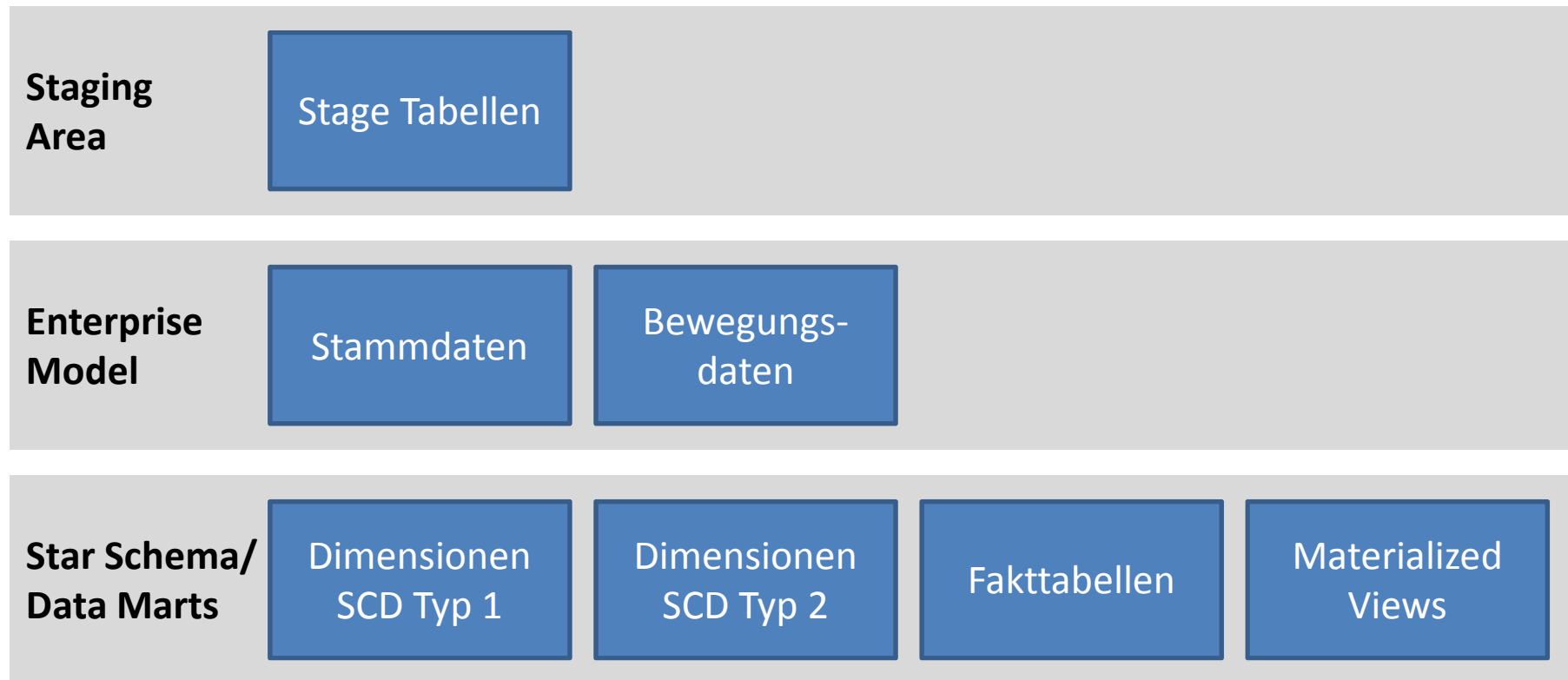


Zunehmendes historisches Datenvolumen

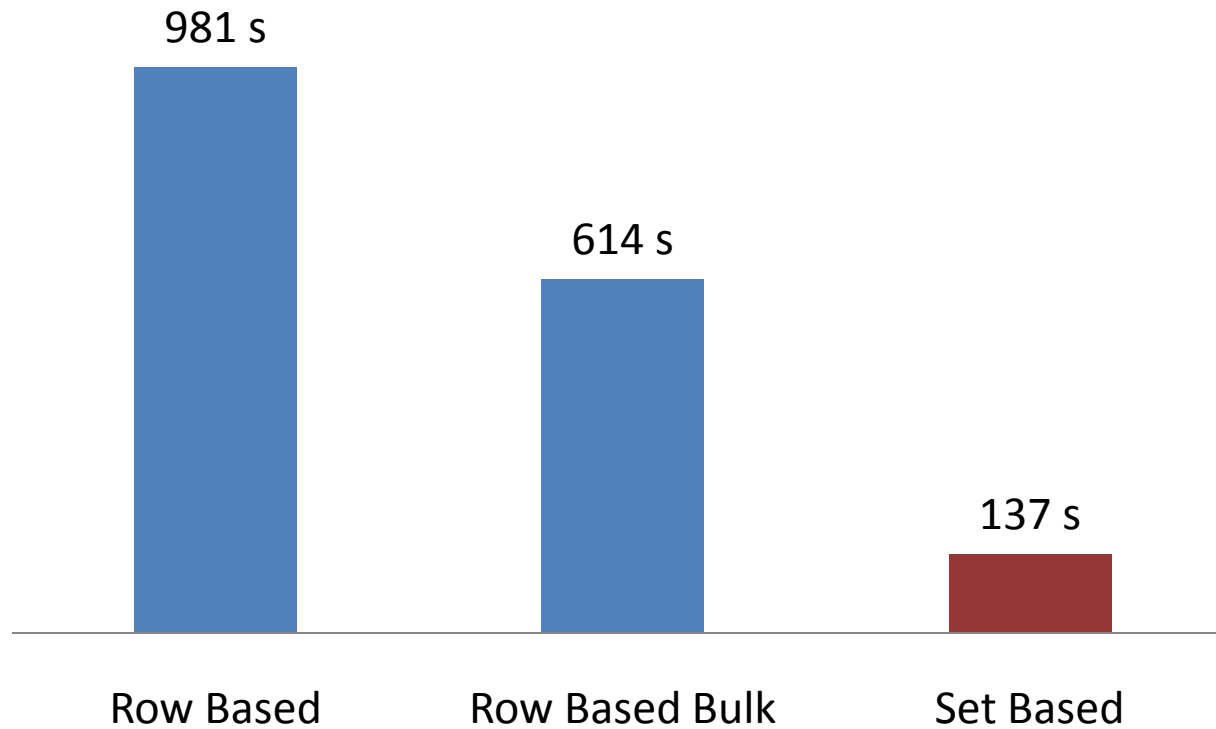
**Was kann man machen?**

Weniger!  
&  
Parallel!

# Was muss der ETL-Prozess leisten?

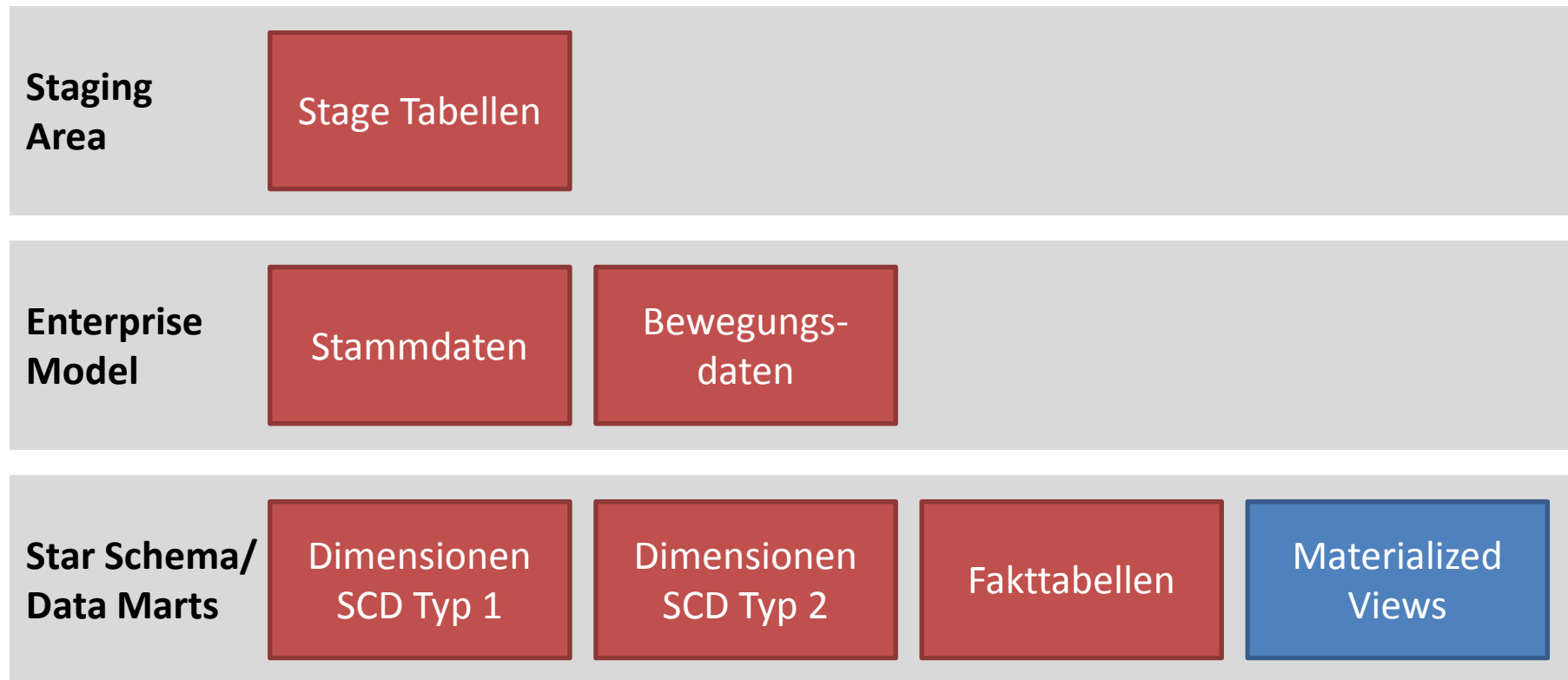


# Set Based statt Row Based

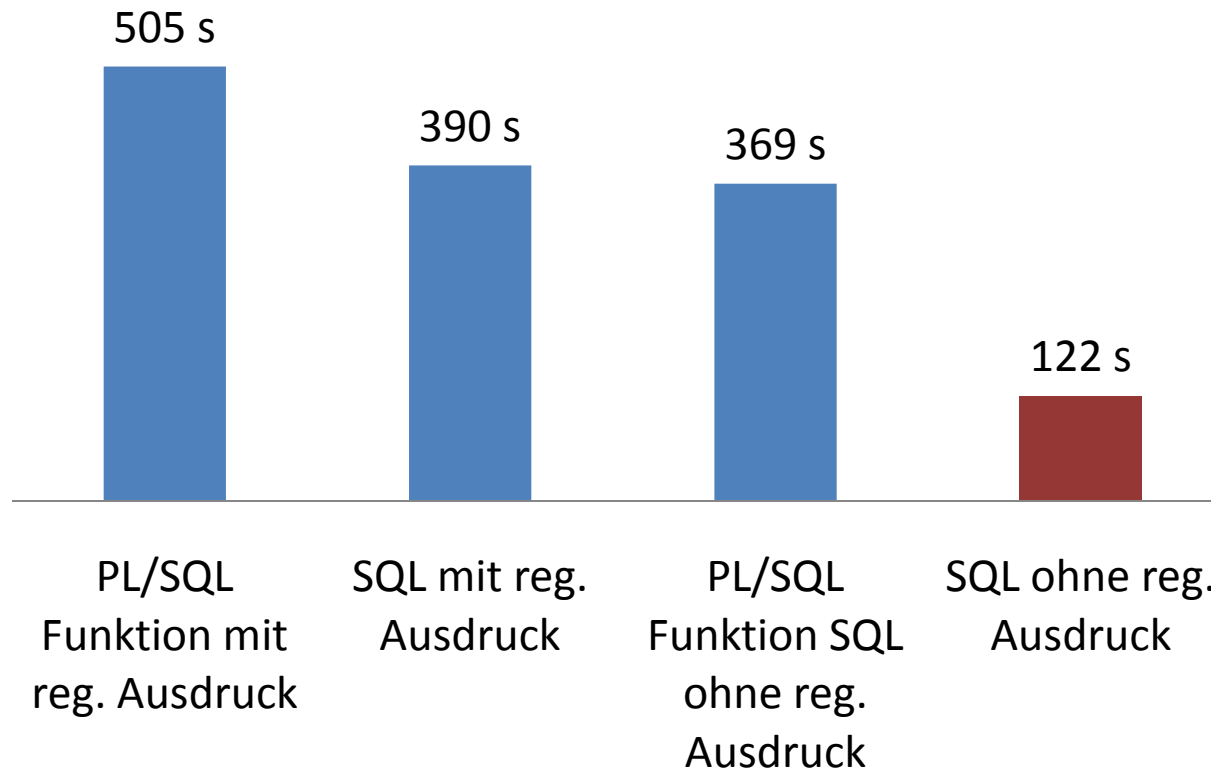


Beispiel: 10 Millionen Datensätze einfügen

# Set Based statt Row Based



# SQL statt PL/SQL



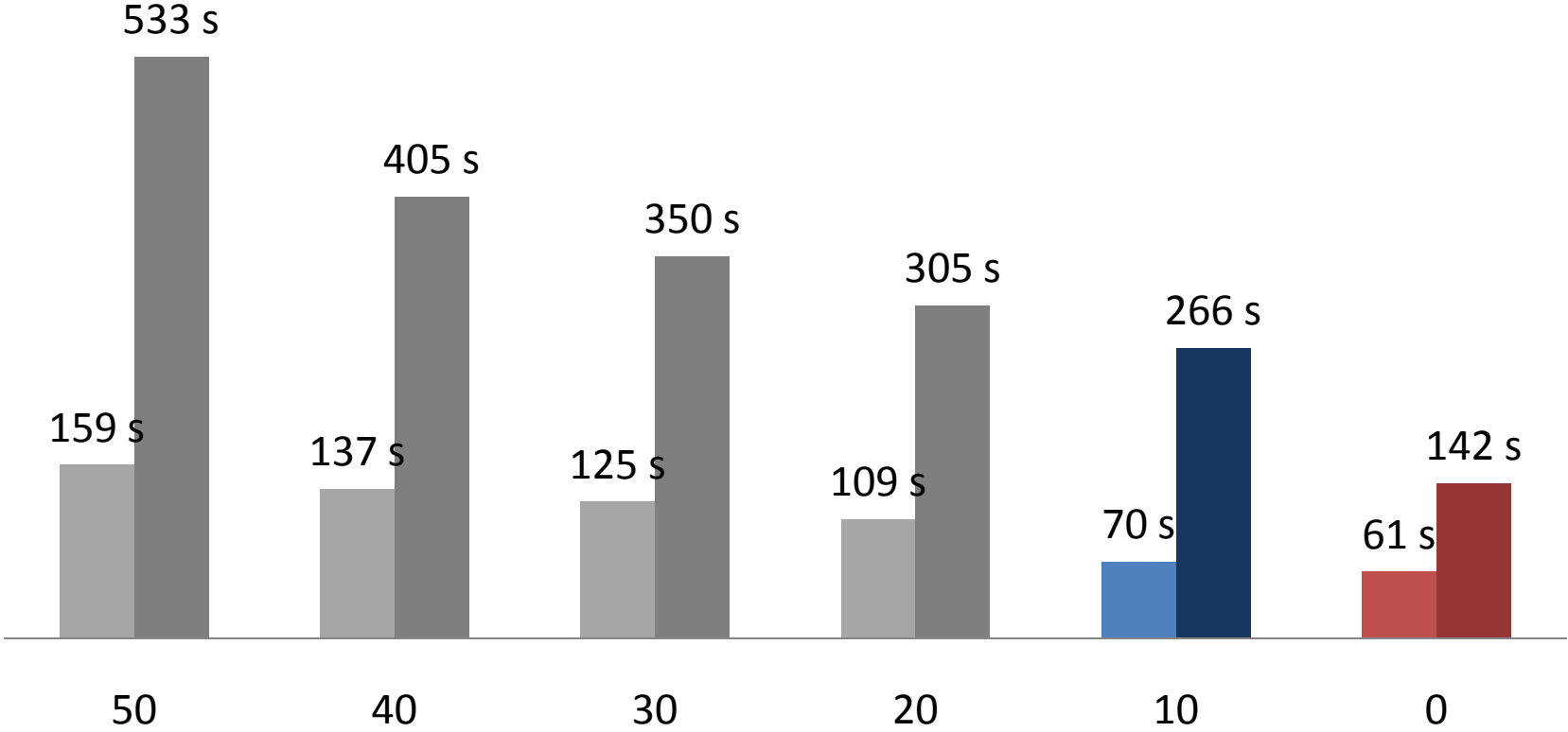
Beispiel: IP-Adresse in 10 Millionen Zeichenketten suchen

# SQL statt PL/SQL





# PCTFREE 0



Beispiel: 10 Millionen Datensätze einfügen und auslesen für Materialized View Refresh

# PCTFREE 0

**Staging Area**

Stage Tabellen

**Enterprise Model**

Stammdaten

Bewegungsdaten

**Star Schema/  
Data Marts**

Dimensionen  
SCD Typ 1

Dimensionen  
SCD Typ 2

Fakttabellen

Materialized Views

# Direct Path Load einsetzen

**Staging Area**

Stage Tabellen

**Enterprise Model**

Stammdaten

Bewegungsdaten

**Star Schema/  
Data Marts**

Dimensionen  
SCD Typ 1

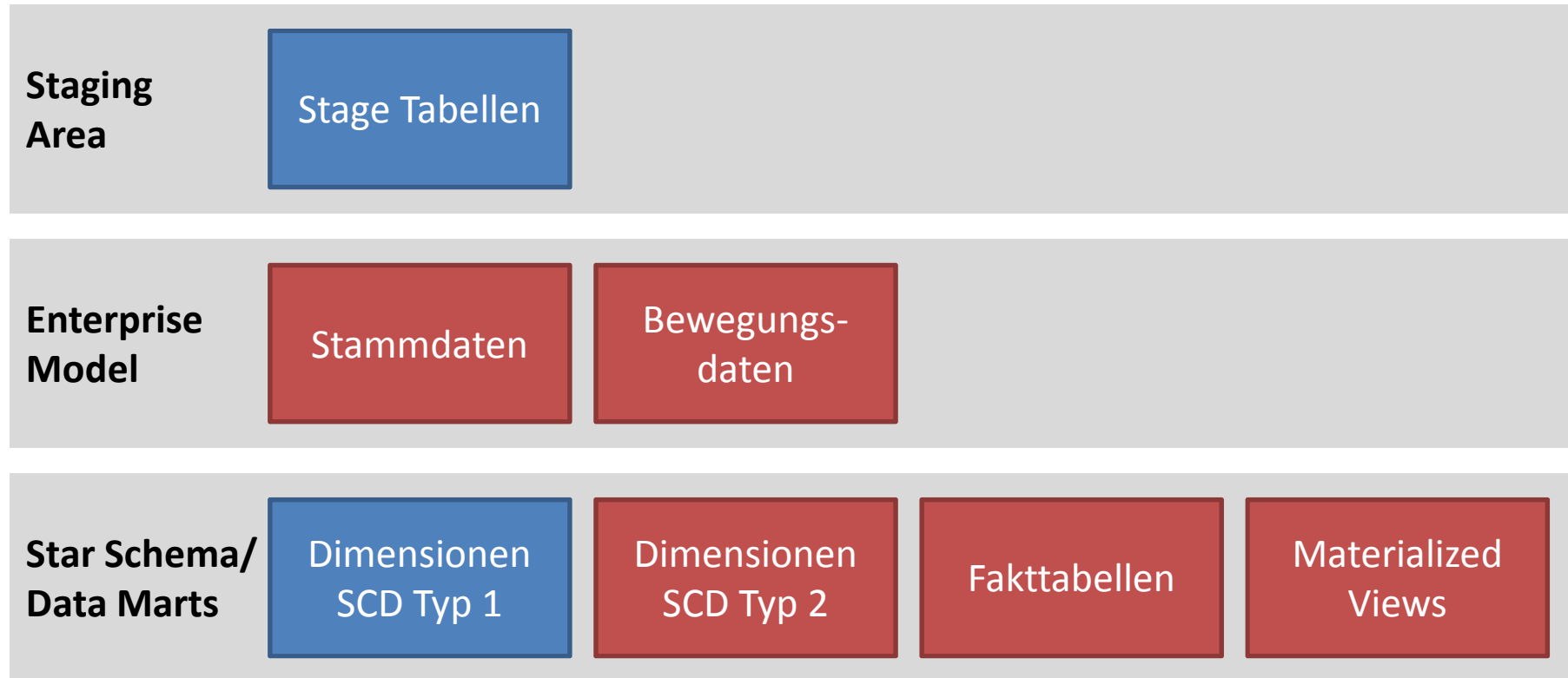
Dimensionen  
SCD Typ 2

Fakttabellen

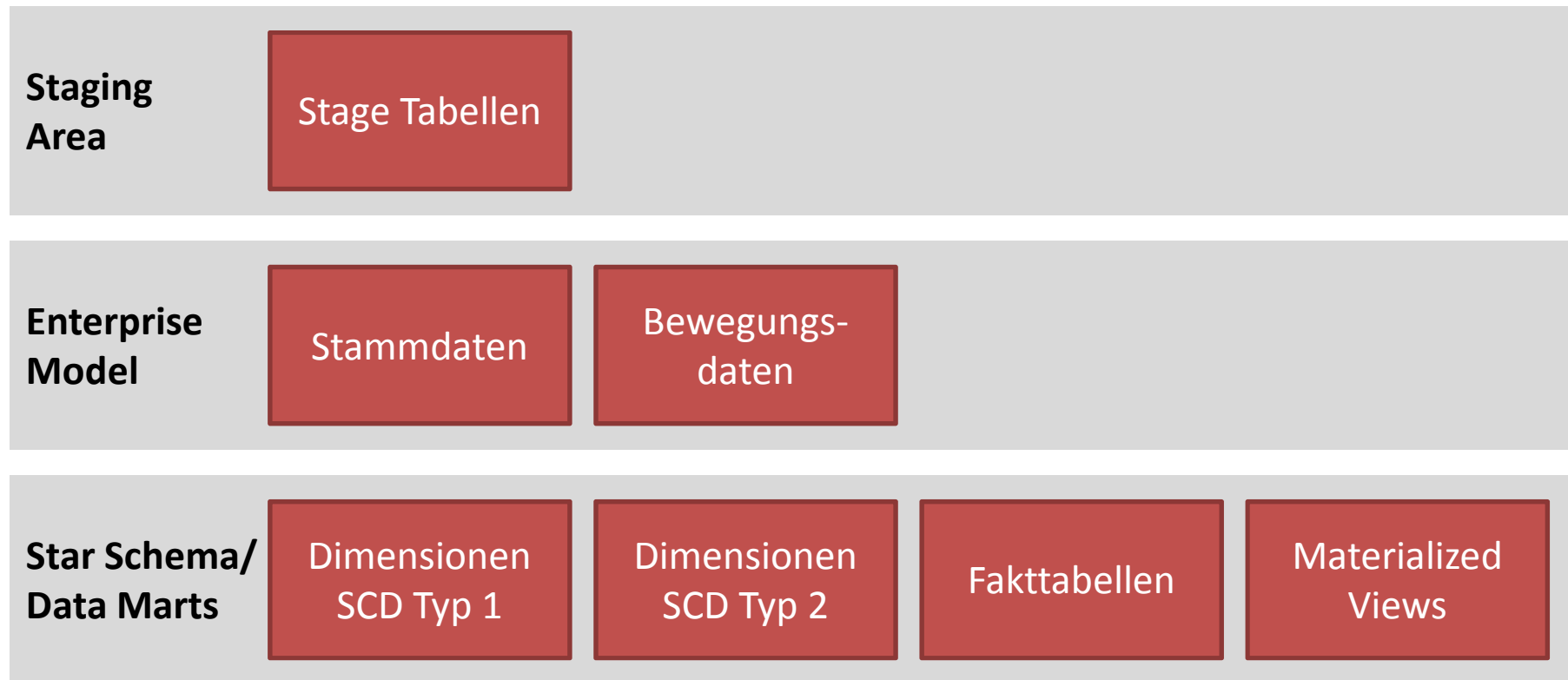
Materialized Views

# Partitionierung

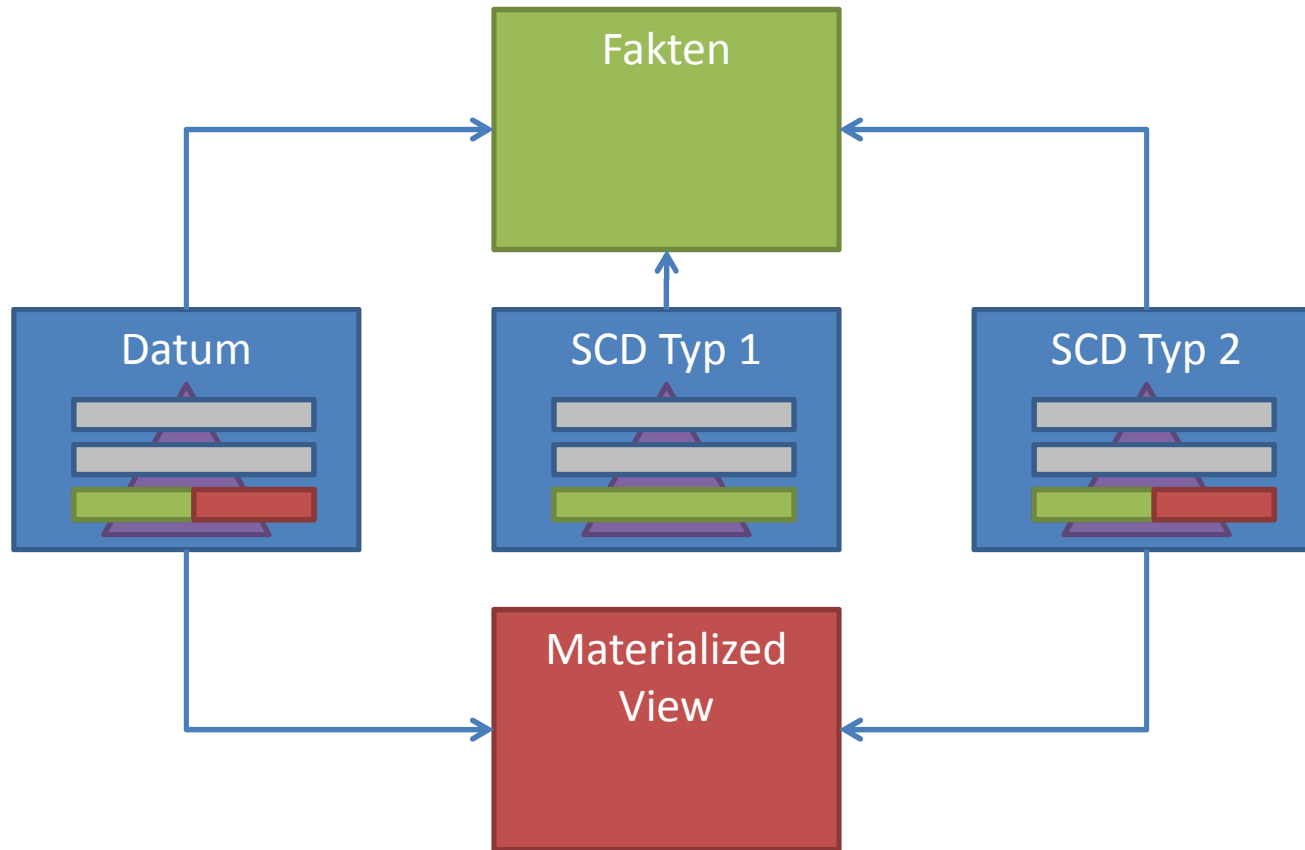
## historischer Daten



# Parallele Verarbeitung

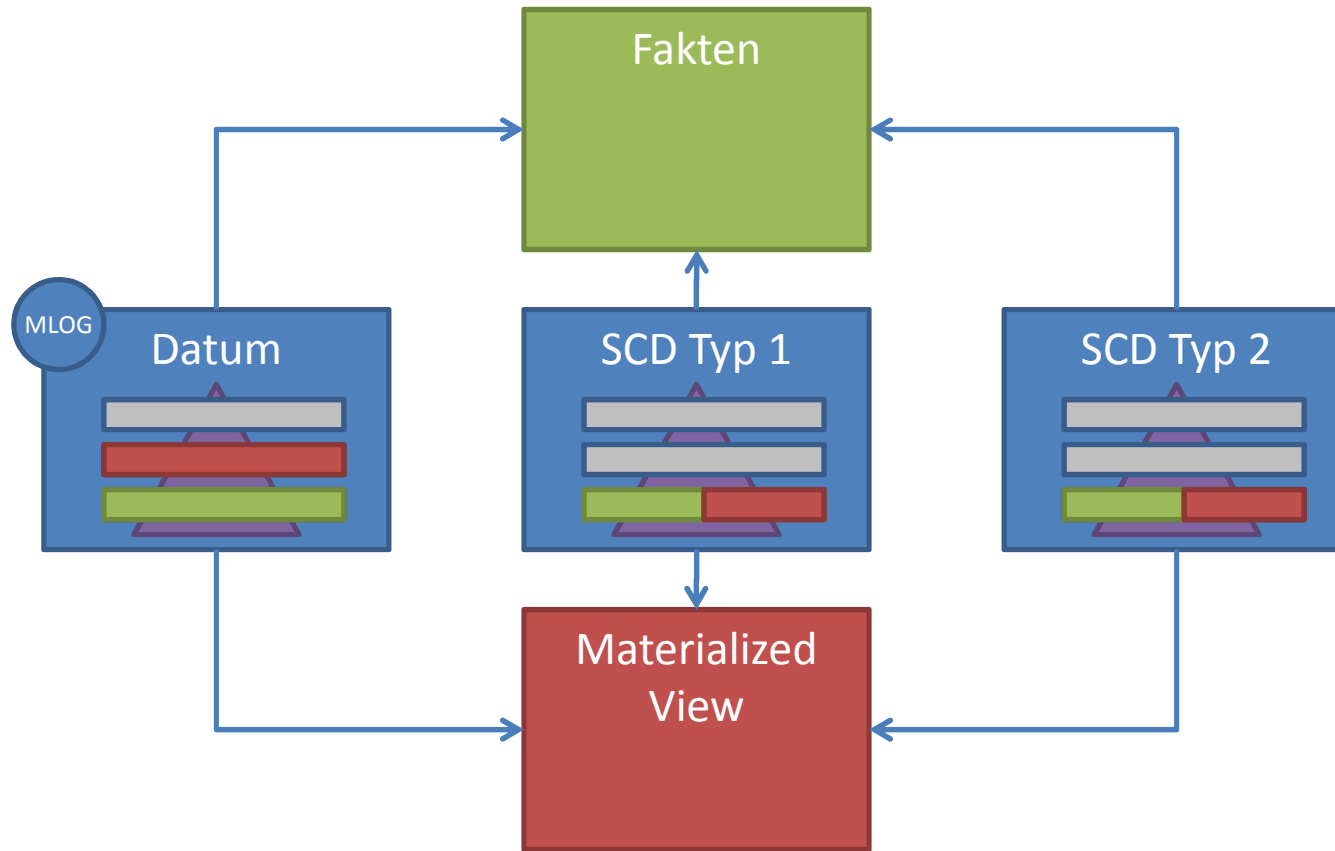


# Materialized Views Refresh



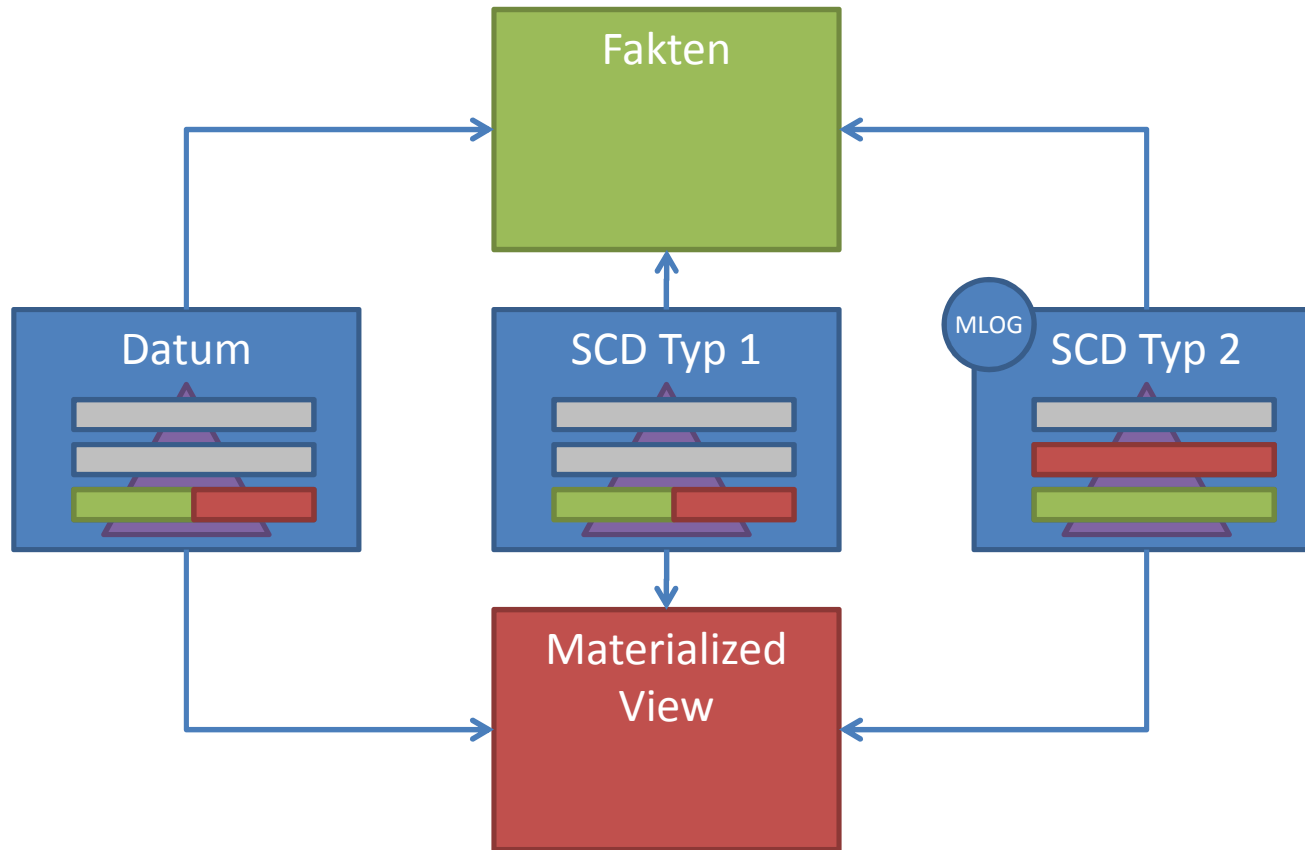
→ inkrementell & konstant (Fast PCT)

# Materialized Views Refresh



→ inkrementell & konstant (Fast PCT)

# Materialized Views Refresh

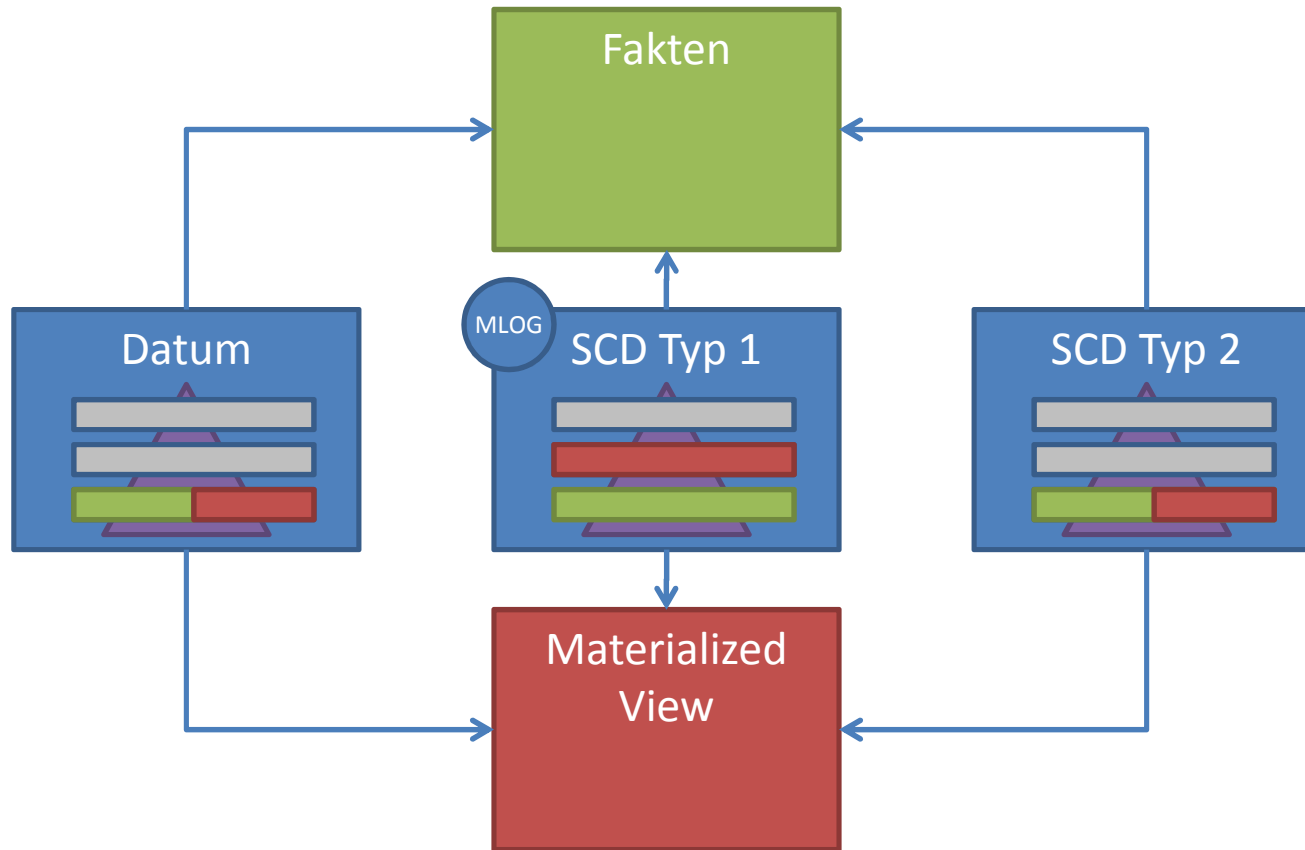


→ inkrementell (Fast)

→ inkrementell & konstant (manuell)



# Materialized Views Refresh



- inkrementell (Fast)
- inkrementell (manuell)

I/O reduzieren

Zugriff auf historische  
Daten vermeiden

Materialized Views  
Refresh optimieren

Features von Oracle 11g



**ARETO Consulting GmbH**

51429 Bergisch Gladbach  
[www.areto-consulting.de](http://www.areto-consulting.de)

**Reinhard Mense**

Geschäftsführer ARETO Consulting  
[reinhard.mense@areto-consulting.de](mailto:reinhard.mense@areto-consulting.de)