

# Data Warehouse

## Inhaltsverzeichnis

Einsatzgebiete	2
Data-Warehouse Implementierungen	2
Big Data Anwendung	4
Literaturverzeichnis	4

## Einsatzgebiete

Data-Warehouse-Systeme werden in verschiedensten Gebieten aus analytischen Zwecken verwendet. Vor allem aber in den Bereichen Marketing und Verkauf. Hier werden oft Daten zu Kauf- und Suchverhalten im Web gesammelt. Mit Hilfe dieser Daten können Marketing- und/oder Verkaufsstrategien entwickelt werden, welche den Unternehmen mehr Profit bringen könnten.

Es gibt noch andere Anwendungsbereiche die für die Nutzung von Data Warehouses in Frage kommen. So könnte man diese Informationssysteme auch für die Vorhersage, des Bevölkerungswachstums eines bestimmten Landes über einen gewissen Zeitraum, verwenden und ebenfalls den proportionalen Bedarf an Nahrung, Ausbildung und Jobs vorhersagen. [1] [2]

*Mehr hab ich da auch noch nicht*

## Data-Warehouse Implementierungen

### Central Data Warehouse

Als Central Data Warehouse wird das klassische Data Warehouse bezeichnet. Dieses kümmert sich um die Aufnahme der Daten aus den Betriebsdatenbanken und die Ausgabe in den Data Marts. Alle Datensets werden hierbei von einer zentralen Datenbank gemanagt. Das kann man sich so vorstellen wie in der Abbildung 1. [1][2]

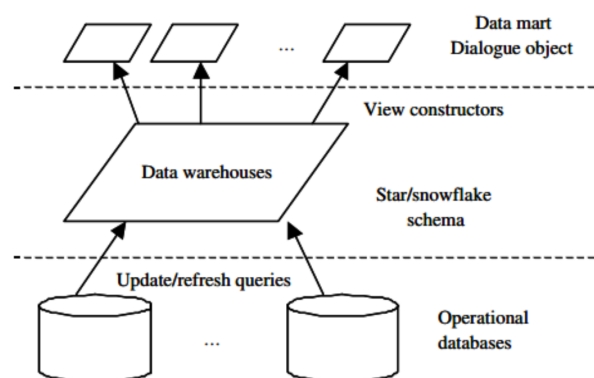


Abbildung 1: Generelle Data Warehouse Architektur [1]

### **1. Die Betriebsdatenbank-Ebene (Operational database)**

Hier werden die einzelnen Daten extrahiert, welche von Interesse für das Data Warehouse sind und dort dann auch gespeichert werden. Diese Daten-Inputs erhält man mittels Update- oder Aktualisierungsanweisungen. Die Befehle wiederum können als einfache Abfragen, auch Queries genannt, durchgeführt werden. So kann man davon ausgehen, dass diese Abfragen in SQL ausformuliert werden.

### **2. Die Data Warehouse-Ebene**

Die Ebene des Data Warehouse kann man sich ähnlich wie ein Datenbanksystem vorstellen. Der Unterschied liegt darin, dass es sich um ein wesentlich einfacheres Schema handelt und keine komplexen Transaktionen betrachtet werden müssen. Schreiboperationen finden ausschließlich in Form von Aktualisierungsbefehlen statt, welche das Data Warehouse und die Betriebsdatenbanken verbinden. Alle anderen Operationen lesen nur die Daten vom Data Warehouse, um so Views für die Data Marts oder Dialogue Objects zu generieren, welche wiederum als OLAP-Schnittstelle fungieren.

### **3. Die OLAP-Ebene**

Die oberste Ebene, die OLAP-Ebene, besteht aus vielen Dialogue Objects, welche von OLAP Anwendungen bearbeitet werden. Jeder User hat eine Sammlung von Dialogue Objects, also Data Marts, geöffnet. Es können immer wieder neue User dazukommen, welche neue Dialogue Objects erstellen können, ohne dass die bereits existierenden geschlossen werden müssen. Allerdings werden diese Dialoge nicht ewig gespeichert, denn sobald der User das System verlässt, werden alle seine Data Marts entfernt. Die Hauptaufgabe dieser Ebene ist die Verwaltung der Data Marts von den Usern und die Erstellung von Neuen. [1]

## **Distributed Data Warehouse**

Wie der Name schon sagt, geht es beim Distributed Data Warehouse, darum mit verteilten Systemen zu arbeiten. Hierbei steht vor allem der Vorteil im Fokus, dass man einzelne Fragmente gewissen Knoten in einem System zuweist. Diese wiederum können die Fragmente replizieren, also werden die Daten an mehreren Knoten gespeichert. Die Idee dahinter ist die Verbesserung der Gesamtleistung eines solchen Systems.

Für die Aktualisierung der Daten, die man erhalten möchte, werden schon vorgefertigte Queries offline ausgeführt, sodass die OLAP-Ebene nur noch die Daten aus dem Data Warehouse auslesen muss. Aus diesem Grund werden replizierende Architekturen von solchen Systemen bevorzugt, weil alle Views für die Data Marts schon lokal generiert werden können. Abbildung 2 zeigt ein Beispiel für ein Distributed Data Warehouse Architektur. [1]

[2]

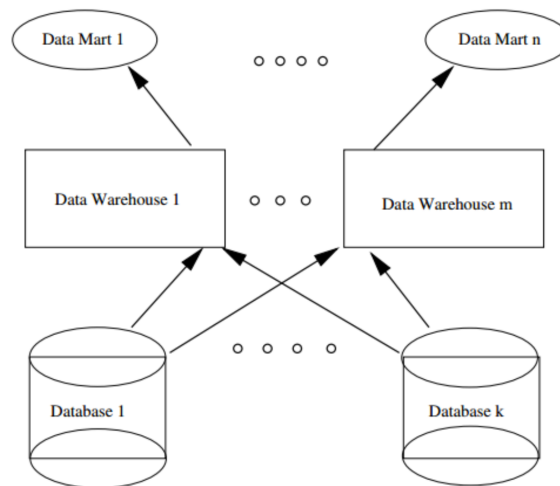


Abbildung 2: Beispiel für eine Distributed Data Warehouse Architektur [1]

## Virtual Data Warehouse

Bin noch nicht dazugekommen. [3]

## Big Data Anwendung

Da habe ich noch nichts passendes gefunden.

## Literaturverzeichnis

- [1] Designing Distributed Data Warehouses and OLAP Systems, Jane Zhao, Massey University, Information Science Research Centre & Department of Information Systems, 2012
- [2] Virtual Data Warehouse Modeling Using Petri Nets for Distributed Decision Making, Nabendu Chaki and Bidyut Biman Sarkar, Journal of Information Technology, 2010
- [3] Trade and Tricks: Traditional vs. Virtual Data Warehouse, Dr. Paresch V. Virparia, Dr. Sanjay H. Buch and Roohana F. Parabia, International Journal of Advanced Engineering & Application, Jan. 2010