



# Beispiele und Eigenschaften von Unternehmensanwendungen

Dr. Matthias Uflacker, Stefan Klauck

18. April 2018

- Phase 1
  - **Einführung zu Unternehmensanwendungen (2 Vorlesungen)**
  - Grundlagen von spaltenorientierten Hauptspeicherdatenbanken (4 Vorlesungen)
  - Wöchentliche Übungsblätter
  
- Phase 2
  - Grundlagen des IT-gestützten Rechnungswesens und Planung (3 Vorlesungen)
  - Programmiermodelle für Unternehmensanwendungen (1 Vorlesung)
  - Zwei praktische Programmierübungen
  
- Klausur

- Unternehmensanwendungen
- Beispiele für Unternehmensanwendungen
- Datenbankaneigenschaft von Unternehmensanwendungen
  - OLTP vs. OLAP
  - Trennung in eigenständige Systeme
  - Optimierung eines Systems für beide Anfragearten
- Eigenschaften von Unternehmensdaten
- Zusammenfassung
- Übungsblatt 1

- „Enterprise applications are about the **display, manipulation, and storage of large amounts of often complex data** and the **support or automation of business processes with that data.**“ Martin Fowler „Patterns of Enterprise Application Architecture Patterns“ (2002)
  - Große komplexe Datenmengen, die in der realen Welt existierende Entitäten abbilden
  - Unterstützung/Automatisierung von Geschäftsprozessen auf Basis dieser Daten
  - Nutzen üblicherweise (relationale) Datenbanken

- Klassische Domäne
  - ERP, SCM, CRM (siehe 01\_Unternehmensanwendungen)

- Erweiterte Domäne

Neue Technologien haben Einfluss auf Unternehmensanwendungen.

Informationsintegration aus unterschiedlichen Quellen bietet neue Möglichkeiten zur Anreicherung und dadurch Optimierung der klassischen Unternehmensprozesse.

- Sensordaten in Produktionsüberwachung
- Unstrukturierte Daten wie medizinische Behandlungsberichte und Röntgenbilder
- Posts im Web und auf sozialen Netzwerken zum Einfluss von Produkten

# Beispiele für Unternehmensanwendungen

## Klassische Domäne: Abwicklung eines Kaufauftrags

Frank Körsgen „SAP R/3 Arbeitsbuch“ (2005)

- Prozesse im Vertrieb werden durch Vertriebs-/Verkaufsbelege abgebildet
  - Kundenanfrage
  - Angebot
  - Auftrag
  - Lieferung
  - Rechnung
  - Buchhaltungsbeleg

# Beispiele für Unternehmensanwendungen

## Klassische Domäne: Abwicklung eines Kaufauftrags

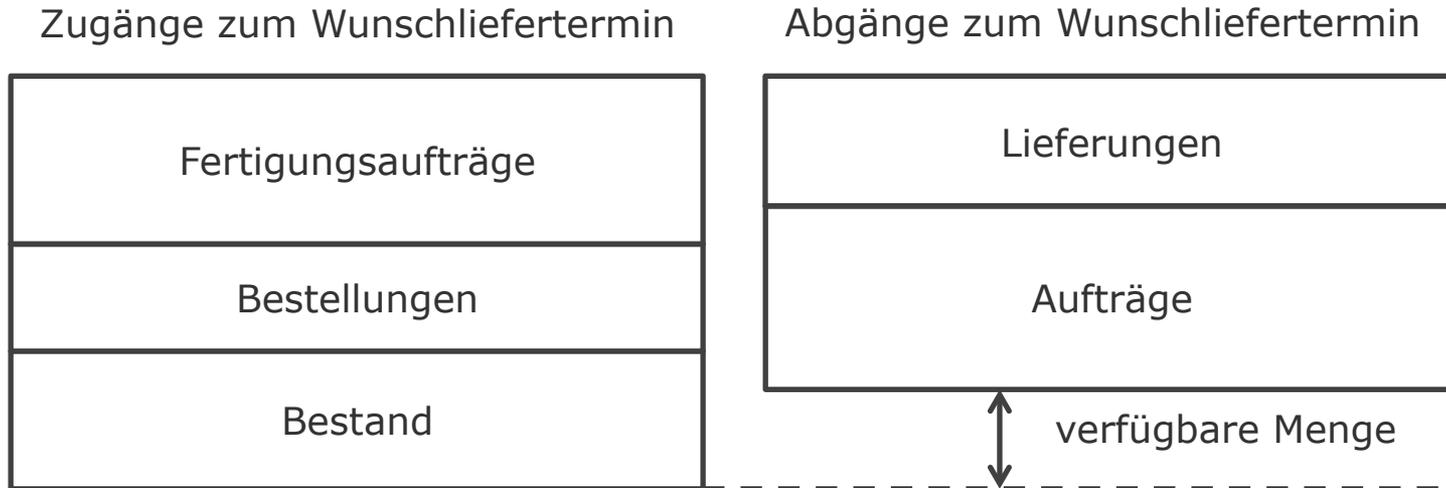
---

- Vorkommende teils aufwändige Prozesse
  - Preisfindung
  - **Verfügbarkeitsprüfung**
  - Bedarfsübergabe (an Einkauf und/oder Fertigung)
  - Versandterminierung
  - Versandstellen- und Routenermittlung
  - Kreditlimitprüfung

# Beispiele für Unternehmensanwendungen

## Klassische Domäne: Abwicklung eines Kaufauftrags

- Verfügbarkeitsprüfung



### ■ Online Transaction Processing (OLTP)

- Beispiele: Auftrag anlegen, Rechnung erstellen, Kontenbuchung, Anlegen und Anzeigen von Stammdaten
- Transaktionsorientierung
- Dateneingabe und -abruf
- Einzelne Tupel

### ■ Online Analytical Processing (OLAP)

- Beispiele: Mahnlauf, Verfügbarkeitsprüfung, Cross-Selling, Operationales Reporting (Auflistung offener Rechnungen)
- „Decision Support“
- Mehrdimensionale Analyseanfragen
- Viele Tupel, aber wenige Attribute

- Unternehmensanwendungen beinhalten sowohl transaktionale als auch analytische Anfragen
- Aber: Anforderungen an OLAP- und OLTP-Datenbanksysteme sind widersprüchlich

(historischer) Ansatz 1:

**Trennung in eigenständige Systeme** (Tuning der Datenspeicherung und des Schemas)

- Data Warehouse: separate optimierte OLAP-Datenbank
  - Effizientere Bearbeitung komplexer analytischer Anfragen
  - Daten vieler Datenbanken können integriert werden

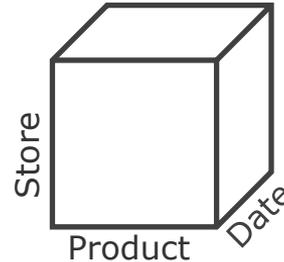
(moderner) Ansatz 2:

**Optimierung eines Systems für beide Anfragearten** (mit Kompromissen) 10

# Unternehmensanwendungen

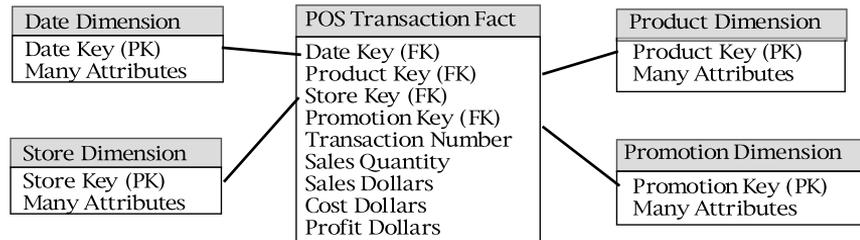
## Systemtrennung: Datensicht im Data-Warehouse

Mehrdimensionale Datensicht:  
Datenwürfel (OLAP-Würfel)



- Relationale Umsetzung im Sternschema mit Denormalisierungen aus Performanzgründen

R(elational)OLAP



Patrick O'Neil „The Star Schema Benchmark and Augmented Fact Table Indexing“ (2009)

- Umsetzung mittels spezialisierter Datenstruktur zum Teil mit materialisierten Aggregaten

M(ultidimensional)OLAP

# Unternehmensanwendungen

## Nachteile der Systemtrennung

---

- Datenredundanz
- Synchronisierung der Systeme durch Extract, Transform, Load (ETL)
  - Kostenintensiver Prozess
  - ETL verursacht eine Zeitverzögerung
- Unterschiedliche Daten Schemas erzeugen ein höhere Komplexität für Anwendungen, die beide Systeme nutzen
- (OLAP-Systeme halten eine vordefinierte Untermenge aller Daten)

# Unternehmensanwendungen

## Kombination von OLTP und OLAP in einem System

---

- Eine Datenbasis als „Wahrheitsquelle“ („Single source of truth“)
- Ziel sind ad-hoc-Analysen auf dem transaktionalen Schema ohne materialisierte Sichten (z.B. vorberechnete Aggregate) – „Mixed Workloads“ auf einem System
- -> Vereinfachte Anwendungen (nur eine DB mit allen Datenpunkten anfragen) und Datenbankstrukturen (weniger Indexe, keine materialisierte Sichten)
- Ermöglicht durch modernere Hardware
- Dennoch: für spezielle Charakteristiken optimierte Datenbanksysteme sind „One size fits all“ Systemen überlegen

Michael Stonebraker „One Size Fits All: An Idea Whose Time Has Come and Gone“ (2005)

# Unternehmensanwendungen

## Vielseitige Anfrageeigenschaften („Mixed Workloads“)

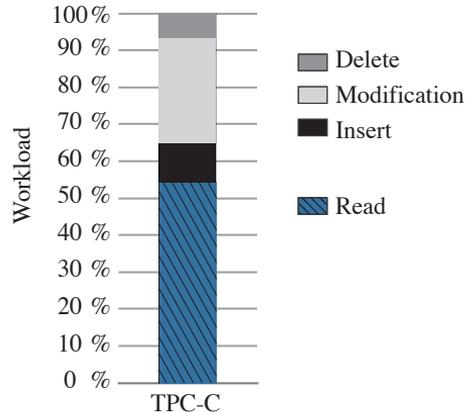
---

- Vereinen Eigenschaften von OLTP **und** OLAP
  - Komplette Zeilenoperationen **und** Materialisierung weniger Spalten
  - Einfache **und** komplexe Abfragen
  - Tuple Wiederherstellung **und** Aggregationen, Gruppierungen, Joins
  - INSERTS, UPDATES **und** viele SELECTS
  - Kurze **und** lange Transaktionen
  - Vordefinierte **und** ad-hoc Abfragen

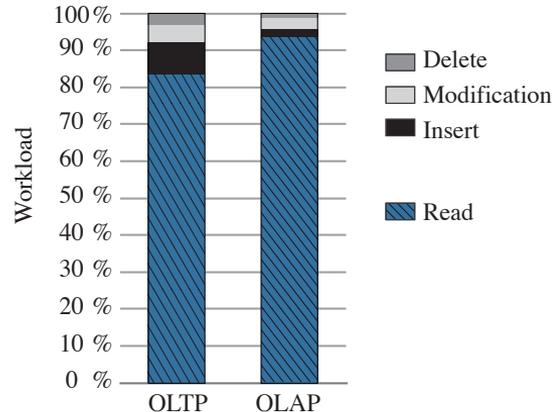
# Unternehmensanwendungen

## Anfrageeigenschaften: OLTP Datenzugriffsmythos

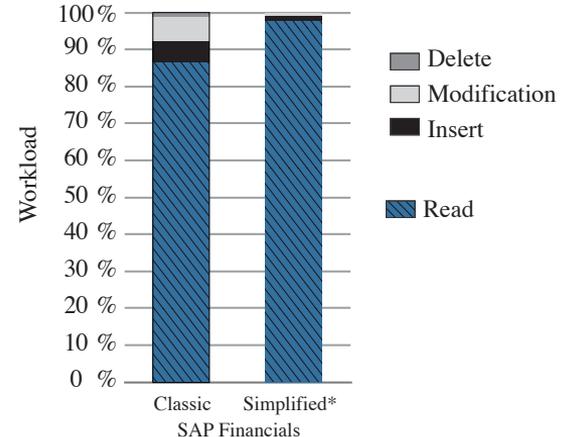
- Workload Analysen von Unternehmensanwendungen zeigen: OLTP und OLAP Workloads unterscheiden sich NICHT wesentlich bezüglich ihres Schreibanteil



(a) TPC-C Benchmark



(b) Workload Analysis  
Krueger et al. VLDB'11



(c) SAP Financials  
Workload Analysis

# Unternehmensanwendungen

## Eigenschaften von Unternehmensdaten

---

- Viele Tabellen haben hunderte von Attributen
- Viele Attribute werden in der Regel NICHT benutzt
- Für viele Attribute dominieren NULL oder DEFAULT-Werte
- Viele Attribute haben eine geringe Kardinalität (Anzahl verschiedener Attributwerte)
  
- Tabellen sind breit und spärlich gefüllt -> erlaubt effiziente Komprimierung

- Detaillierte Informationen in

Jens Krüger „Fast Updates on Read-Optimized Databases Using Multi-Core CPUs“ (2011)  
[http://www.vldb.org/pvldb/vol5/p061\\_jenskrueger\\_vldb2012.pdf](http://www.vldb.org/pvldb/vol5/p061_jenskrueger_vldb2012.pdf)

- Unternehmensanwendungen beinhalten sowohl transaktionale als auch analytische Anfragen; Insbesondere analytische Anforderungen sind mit der Zeit gewachsen
- Tabellen von Unternehmensanwendungen sind breit und spärlich gefüllt
- Zwei Ansätze: Separat optimierte Systeme vs. **Vereinigung in einem System**
  - Vereinigung überwindet viele Nachteile der Systemtrennung
  - Entsprechende Datenbankarchitekturen beinhalten viele Optimierungen, die den Anfrage- und Datencharakteristiken gerecht werden (kommende Vorlesungen)

# Unternehmensanwendungen

## Übungsblatt 1

Unternehmensanwendungen

Sommersemester 2018: Übung 1

Enterprise Platform and  
Integration Concepts

Fachgebiet | Hasso Plattner Institut  
Universität Potsdam



HPI

### Aufgabe 1

Lesen Sie Hasso Plattners Paper ["The Impact of Columnar In-Memory Databases on Enterprise Systems"](#) und beantworten Sie folgende Fragen:

- Was sind „transaction-maintained“ Aggregate?
- Welche Nachteile haben diese?
- Beschreiben Sie die grundlegende Idee des „Aggregate Caches“?
- Welche Anwendungen werden im Paper genannt, die sowohl analytische als auch transaktionale Anfragemuster aufweisen? Kennen Sie weitere Anwendungen?
- Warum werden Inserts in Unternehmensanwendungen durch den Umstieg auf spaltenbasierte Hauptspeicherdatenbanken nicht unbedingt langsamer als in einer Zeilendatenbank?
- In welchem Fall ist eine zeilenbasierte Datenbank mit materialisierten Aggregaten einer Spaltendatenbank für Mengenoperationen überlegen?
- Erklären Sie warum spaltenbasierte Datenbanken die Entwicklung neuer Unternehmensanwendungen vereinfachen können.
- Was ist die Idee hinter der Datenpartitionierung von Unternehmensdaten in „heiße“ und „kalte“ Daten?

### Aufgabe 2

Installieren Sie die [SAP HANA Tools](#).

Fügen Sie in der Oberfläche die folgende HANA-Instanz zu Ihren (Datenbank-)Systemen hinzu:

(Der Zugriff auf die Datenbank ist nur aus dem HPI-Netz möglich.)

Host Name: syene  
Instance Nummer: 02  
User Name: STUDENTS  
Password: Students2

Die Nutzung der SQL-Konsole des SAP HANA Tools ist für die zweite Übung empfohlen.

### Abgabeanweisung

Die Aufgaben dürfen in Zweierteams bearbeitet werden.

Senden Sie eine pdf-Datei mit ihren Lösungen mit den Betreff „[Unternehmensanwendungen Übung1] NACHNAME1 NACHNAME2“ an [stefan.klauck@hpi.de](mailto:stefan.klauck@hpi.de).

**Abgabefrist: 30. April Anywhere on Earth (AoE).**

Die pünktliche Abgabe und das Bestehen der Übungsblätter sind Voraussetzung für die Zulassung zur Klausur.