

# Einführung in Hauptspeicherdatenbanken

Harald Zankl

Probevorlesung

13. 01., 13:15–14:00, HS C



- Organisation
- Überblick
- Konklusion

## Organisation

- Einführung in Hauptspeicherdatenbanken (703650, VO1, 2 ECTS)
- Modul: Neuere Datenbankmodelle (VO2+VO1+PS2, 10 ECTS)
- Wahlmodul Master Informatik (SS 2015)
- Zeit & Ort: Mittwoch, 13:15–14:00, HS C

## Organisation

- Einführung in Hauptspeicherdatenbanken (703650, VO1, 2 ECTS)
- Modul: Neuere Datenbankmodelle (VO2+VO1+PS2, 10 ECTS)
- Wahlmodul Master Informatik (SS 2015)
- Zeit & Ort: Mittwoch, 13:15–14:00, HS C

## Literatur

- H. Plattner, A. Zeier,  
In-Memory Data Management: Technology and Applications,  
2. Auflage, Springer, 2012  
 
- H. Plattner,  
A Course in In-Memory Data Management: The inner Mechanics of  
In-Memory Databases, Springer, 2013  
 

## Zeitplan

VO	Datum	Thema
1	4. März	Organisation, Überblick
2	11. März	
3	18. März	
4	25. März	
5	22. April	
6	29. April	
7	6. Mai	
8	13. Mai	
9	20. Mai	
10	27. Mai	
11	3. Juni	
12	10. Juni	
13	17. Juni	
14	24. Juni	Klausur

## Zeitplan

VO	Datum	Thema
1	4. März	Organisation, Überblick
2	11. März	Verwendung
3	18. März	
4	25. März	
5	22. April	
6	29. April	
7	6. Mai	
8	13. Mai	
9	20. Mai	
10	27. Mai	
11	3. Juni	
12	10. Juni	Verwendung
13	17. Juni	Verwendung
14	24. Juni	Klausur

## Zeitplan

VO	Datum	Thema
1	4. März	Organisation, Überblick
2	11. März	Verwendung
3	18. März	Realisierung
4	25. März	Realisierung
5	22. April	
6	29. April	
7	6. Mai	Realisierung
8	13. Mai	Realisierung
9	20. Mai	Realisierung
10	27. Mai	
11	3. Juni	
12	10. Juni	Verwendung
13	17. Juni	Verwendung
14	24. Juni	Klausur

## Zeitplan

VO	Datum	Thema
1	4. März	Organisation, Überblick
2	11. März	Verwendung
3	18. März	Realisierung
4	25. März	Realisierung
5	22. April	technische Grundlagen
6	29. April	technische Grundlagen
7	6. Mai	Realisierung
8	13. Mai	Realisierung
9	20. Mai	Realisierung
10	27. Mai	
11	3. Juni	
12	10. Juni	Verwendung
13	17. Juni	Verwendung
14	24. Juni	Klausur



## Zeitplan

VO	Datum	Thema
1	4. März	Organisation, Überblick
2	11. März	Verwendung
3	18. März	Realisierung
4	25. März	Realisierung
5	22. April	technische Grundlagen
6	29. April	technische Grundlagen
7	6. Mai	Realisierung
8	13. Mai	Realisierung
9	20. Mai	Realisierung
10	27. Mai	fortgeschrittene Konzepte
11	3. Juni	fortgeschrittene Konzepte
12	10. Juni	Verwendung
13	17. Juni	Verwendung
14	24. Juni	Klausur

## Zeitplan

VO	Datum	Thema
1	4. März	Organisation, Überblick
2	11. März	Enterprise Applications (OLTP, OLAP)
3	18. März	Realisierung
4	25. März	Realisierung
5	22. April	technische Grundlagen
6	29. April	technische Grundlagen
7	6. Mai	Realisierung
8	13. Mai	Realisierung
9	20. Mai	Realisierung
10	27. Mai	fortgeschrittene Konzepte
11	3. Juni	fortgeschrittene Konzepte
12	10. Juni	Verwendung
13	17. Juni	Verwendung
14	24. Juni	Klausur

## Zeitplan

VO	Datum	Thema
1	4. März	Organisation, Überblick
2	11. März	Enterprise Applications (OLTP, OLAP)
3	18. März	Design (am Beispiel SanssouciDB)
4	25. März	Realisierung
5	22. April	technische Grundlagen
6	29. April	technische Grundlagen
7	6. Mai	Realisierung
8	13. Mai	Realisierung
9	20. Mai	Realisierung
10	27. Mai	fortgeschrittene Konzepte
11	3. Juni	fortgeschrittene Konzepte
12	10. Juni	Verwendung
13	17. Juni	Verwendung
14	24. Juni	Klausur

## Zeitplan

VO	Datum	Thema
1	4. März	Organisation, Überblick
2	11. März	Enterprise Applications (OLTP, OLAP)
3	18. März	Design (am Beispiel SanssouciDB)
4	25. März	Datenlayout (zeilen-/spaltenorientiert)
5	22. April	technische Grundlagen
6	29. April	technische Grundlagen
7	6. Mai	Realisierung
8	13. Mai	Realisierung
9	20. Mai	Realisierung
10	27. Mai	fortgeschrittene Konzepte
11	3. Juni	fortgeschrittene Konzepte
12	10. Juni	Verwendung
13	17. Juni	Verwendung
14	24. Juni	Klausur

## Zeitplan

VO	Datum	Thema
1	4. März	Organisation, Überblick
2	11. März	Enterprise Applications (OLTP, OLAP)
3	18. März	Design (am Beispiel SanssouciDB)
4	25. März	Datenlayout (zeilen-/spaltenorientiert)
5	22. April	Speicherhierarchie (Festplatte, Hauptspeicher, Cache)
6	29. April	technische Grundlagen
7	6. Mai	Realisierung
8	13. Mai	Realisierung
9	20. Mai	Realisierung
10	27. Mai	fortgeschrittene Konzepte
11	3. Juni	fortgeschrittene Konzepte
12	10. Juni	Verwendung
13	17. Juni	Verwendung
14	24. Juni	Klausur

## Zeitplan

VO	Datum	Thema
1	4. März	Organisation, Überblick
2	11. März	Enterprise Applications (OLTP, OLAP)
3	18. März	Design (am Beispiel SanssouciDB)
4	25. März	Datenlayout (zeilen-/spaltenorientiert)
5	22. April	Speicherhierarchie (Festplatte, Hauptspeicher, Cache)
6	29. April	<b>Komprimierung</b>
7	6. Mai	Realisierung
8	13. Mai	Realisierung
9	20. Mai	Realisierung
10	27. Mai	fortgeschrittene Konzepte
11	3. Juni	fortgeschrittene Konzepte
12	10. Juni	Verwendung
13	17. Juni	Verwendung
14	24. Juni	Klausur

## Zeitplan

VO	Datum	Thema
1	4. März	Organisation, Überblick
2	11. März	Enterprise Applications (OLTP, OLAP)
3	18. März	Design (am Beispiel SanssouciDB)
4	25. März	Datenlayout (zeilen-/spaltenorientiert)
5	22. April	Speicherhierarchie (Festplatte, Hauptspeicher, Cache)
6	29. April	Komprimierung
7	6. Mai	Speicherorganisation (Differential Buffer)
8	13. Mai	Realisierung
9	20. Mai	Realisierung
10	27. Mai	fortgeschrittene Konzepte
11	3. Juni	fortgeschrittene Konzepte
12	10. Juni	Verwendung
13	17. Juni	Verwendung
14	24. Juni	Klausur

## Zeitplan

VO	Datum	Thema
1	4. März	Organisation, Überblick
2	11. März	Enterprise Applications (OLTP, OLAP)
3	18. März	Design (am Beispiel SanssouciDB)
4	25. März	Datenlayout (zeilen-/spaltenorientiert)
5	22. April	Speicherhierarchie (Festplatte, Hauptspeicher, Cache)
6	29. April	Komprimierung
7	6. Mai	Speicherorganisation (Differential Buffer)
8	13. Mai	Operationen (Schreiben: insert, update, delete)
9	20. Mai	Realisierung
10	27. Mai	fortgeschrittene Konzepte
11	3. Juni	fortgeschrittene Konzepte
12	10. Juni	Verwendung
13	17. Juni	Verwendung
14	24. Juni	Klausur



## Zeitplan

VO	Datum	Thema
1	4. März	Organisation, Überblick
2	11. März	Enterprise Applications (OLTP, OLAP)
3	18. März	Design (am Beispiel SanssouciDB)
4	25. März	Datenlayout (zeilen-/spaltenorientiert)
5	22. April	Speicherhierarchie (Festplatte, Hauptspeicher, Cache)
6	29. April	Komprimierung
7	6. Mai	Speicherorganisation (Differential Buffer)
8	13. Mai	Operationen (Schreiben: insert, update, delete)
9	20. Mai	Operationen (Lesen: scan, search, join)
10	27. Mai	fortgeschrittene Konzepte
11	3. Juni	fortgeschrittene Konzepte
12	10. Juni	Verwendung
13	17. Juni	Verwendung
14	24. Juni	Klausur

## Zeitplan

VO	Datum	Thema
1	4. März	Organisation, Überblick
2	11. März	Enterprise Applications (OLTP, OLAP)
3	18. März	Design (am Beispiel SanssouciDB)
4	25. März	Datenlayout (zeilen-/spaltenorientiert)
5	22. April	Speicherhierarchie (Festplatte, Hauptspeicher, Cache)
6	29. April	Komprimierung
7	6. Mai	Speicherorganisation (Differential Buffer)
8	13. Mai	Operationen (Schreiben: insert, update, delete)
9	20. Mai	Operationen (Lesen: scan, search, join)
10	27. Mai	<b>Performanz</b>
11	3. Juni	fortgeschrittene Konzepte
12	10. Juni	Verwendung
13	17. Juni	Verwendung
14	24. Juni	Klausur

## Zeitplan

VO	Datum	Thema
1	4. März	Organisation, Überblick
2	11. März	Enterprise Applications (OLTP, OLAP)
3	18. März	Design (am Beispiel SanssouciDB)
4	25. März	Datenlayout (zeilen-/spaltenorientiert)
5	22. April	Speicherhierarchie (Festplatte, Hauptspeicher, Cache)
6	29. April	Komprimierung
7	6. Mai	Speicherorganisation (Differential Buffer)
8	13. Mai	Operationen (Schreiben: insert, update, delete)
9	20. Mai	Operationen (Lesen: scan, search, join)
10	27. Mai	Performanz
11	3. Juni	Wiederherstellung
12	10. Juni	Verwendung
13	17. Juni	Verwendung
14	24. Juni	Klausur

## Zeitplan

VO	Datum	Thema
1	4. März	Organisation, Überblick
2	11. März	Enterprise Applications (OLTP, OLAP)
3	18. März	Design (am Beispiel SanssouciDB)
4	25. März	Datenlayout (zeilen-/spaltenorientiert)
5	22. April	Speicherhierarchie (Festplatte, Hauptspeicher, Cache)
6	29. April	Komprimierung
7	6. Mai	Speicherorganisation (Differential Buffer)
8	13. Mai	Operationen (Schreiben: insert, update, delete)
9	20. Mai	Operationen (Lesen: scan, search, join)
10	27. Mai	Performanz
11	3. Juni	Wiederherstellung
12	10. Juni	Auswirkungen (Praxis)
13	17. Juni	Verwendung
14	24. Juni	Klausur

## Zeitplan

VO	Datum	Thema
1	4. März	Organisation, Überblick
2	11. März	Enterprise Applications (OLTP, OLAP)
3	18. März	Design (am Beispiel SanssouciDB)
4	25. März	Datenlayout (zeilen-/spaltenorientiert)
5	22. April	Speicherhierarchie (Festplatte, Hauptspeicher, Cache)
6	29. April	Komprimierung
7	6. Mai	Speicherorganisation (Differential Buffer)
8	13. Mai	Operationen (Schreiben: insert, update, delete)
9	20. Mai	Operationen (Lesen: scan, search, join)
10	27. Mai	Performanz
11	3. Juni	Wiederherstellung
12	10. Juni	Auswirkungen (Praxis)
13	17. Juni	Ausblick
14	24. Juni	Klausur

## Enterprise Applications

- Bsp: Telekommunikation, Patientendaten, Soziale Medien

## Enterprise Applications

- Bsp: Telekommunikation, Patientendaten, Soziale Medien
  - viele Daten
  - komplexe Abfragen

## Enterprise Applications

- Bsp: Telekommunikation, Patientendaten, Soziale Medien
  - viele Daten (**OLTP**)
  - komplexe Abfragen
- aktuell: Zweiteilung

## OLTP

- Online Transactional Processing
- viele/einfache Transaktionen
- konventionelle DB



## Enterprise Applications

- Bsp: Telekommunikation, Patientendaten, Soziale Medien
  - viele Daten (OLTP)
  - komplexe Abfragen (OLAP)
- aktuell: Zweiteilung

### OLTP

- Online Transactional Processing
- viele/einfache Transaktionen
- konventionelle DB

### OLAP

- Online Analytical Processing
- wenige/komplexe Analysen
- Datawarehouse

## Enterprise Applications

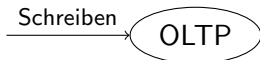
- Bsp: Telekommunikation, Patientendaten, Soziale Medien
  - viele Daten (OLTP)
  - komplexe Abfragen (OLAP)
- aktuell: Zweiteilung

### OLTP

- Online Transactional Processing
- viele/einfache Transaktionen
- konventionelle DB

### OLAP

- Online Analytical Processing
- wenige/komplexe Analysen
- Datawarehouse



## Enterprise Applications

- Bsp: Telekommunikation, Patientendaten, Soziale Medien
  - viele Daten (OLTP)
  - komplexe Abfragen (OLAP)
- aktuell: Zweiteilung

### OLTP

- Online Transactional Processing
- viele/einfache Transaktionen
- konventionelle DB

### OLAP

- Online Analytical Processing
- wenige/komplexe Analysen
- Datawarehouse



## Nachteile der Trennung

- getrennte Architekturen erfordern ETL (extract-transform-load)
- Analyse von alten Daten oder Redundanz (views, cubes)

## Hauptspeicherdatenbank (SanssouciDB)

- Multi-core (Blades @ 64 Cores @ 2TB Hauptspeicher)
- Spaltenorientiert
- Insert-only
- Wörterbuch Kodierung (Komprimierung)

## Hauptspeicherdatenbank (SanssouciDB)

- Multi-core (Blades @ 64 Cores @ 2TB Hauptspeicher)
- Spaltenorientiert
- Insert-only
- Wörterbuch Kodierung (Komprimierung)

## Vorteile

- OLTP und OLAP in einem System
- Effizienz (Analysen in Echtzeit)
- horizontale (scale out) vs. vertikale (~~scale up~~) Skalierung
- Datenverarbeitung nahe an Daten

## Hauptspeicherdatenbank (SanssouciDB)

- Multi-core (Blades @ 64 Cores @ 2TB Hauptspeicher)
- Spaltenorientiert
- Insert-only
- Wörterbuch Kodierung (Komprimierung)

## Vorteile

- OLTP und OLAP in einem System
- Effizienz (Analysen in Echtzeit)
- horizontale (scale out) vs. vertikale (~~scale up~~) Skalierung
- Datenverarbeitung nahe an Daten

## Nachteile

## Hauptspeicherdatenbank (SanssouciDB)

- Multi-core (Blades @ 64 Cores @ 2TB Hauptspeicher)
- Spaltenorientiert
- Insert-only
- Wörterbuch Kodierung (Komprimierung)

## Vorteile

- OLTP und OLAP in einem System
- Effizienz (Analysen in Echtzeit)
- horizontale (scale out) vs. vertikale (scale up) Skalierung
- Datenverarbeitung nahe an Daten

## Herausforderungen

## Hauptspeicherdatenbank (SanssouciDB)

- Multi-core (Blades @ 64 Cores @ 2TB Hauptspeicher)
- Spaltenorientiert
- Insert-only
- Wörterbuch Kodierung (Komprimierung)

## Vorteile

- OLTP und OLAP in einem System
- Effizienz (Analysen in Echtzeit)
- horizontale (scale out) vs. vertikale (scale up) Skalierung
- Datenverarbeitung nahe an Daten

## Herausforderungen

- verteiltes System, Speicherorganisation



# Layouts

id	name	zip	country
1	Kunde1	6020	Österreich
2	Kunde2	9900	Österreich
3	Kunde3	80331	Deutschland

## Layouts

id	name	zip	country
1	Kunde1	6020	Österreich
2	Kunde2	9900	Österreich
3	Kunde3	80331	Deutschland

zeilenorientiert:

1	Kunde1	6020	Österreich	2	Kunde2	9900	Österreich	3	...
---	--------	------	------------	---	--------	------	------------	---	-----

## Layouts

id	name	zip	country
1	Kunde1	6020	Österreich
2	Kunde2	9900	Österreich
3	Kunde3	80331	Deutschland

zeilenorientiert:

1	Kunde1	6020	Österreich	2	Kunde2	9900	Österreich	3	...
---	--------	------	------------	---	--------	------	------------	---	-----

spaltenorientiert:

1	2	3	Kunde1	Kunde2	Kunde3	6020	9900	...
---	---	---	--------	--------	--------	------	------	-----

## Layouts

id	name	zip	country
1	Kunde1	6020	Österreich
2	Kunde2	9900	Österreich
3	Kunde3	80331	Deutschland

zeilenorientiert:

1	Kunde1	6020	Österreich	2	Kunde2	9900	Österreich	3	...
---	--------	------	------------	---	--------	------	------------	---	-----

spaltenorientiert:

1	2	3	Kunde1	Kunde2	Kunde3	6020	9900	...
---	---	---	--------	--------	--------	------	------	-----

## Eigenschaften

Operation	zeilenorientiert	spaltenorientiert
Schreiben (insert/update/delete)	+	-
Lesen (scan/search/join)	-	+
Komprimierung	o	+
Attribute hinzufügen/entfernen	-	+
Index	nötig	nicht nötig
	OLTP	OLAP

## Festplatten

- sind langsam
- werden immer billiger (2TB  $\approx$  200 Euro)
- Zugriffszeit nicht exakt prognostizierbar (Position Lesekopf)

## Festplatten

- sind langsam
- werden immer billiger (2TB  $\approx$  200 Euro)
- Zugriffszeit nicht exakt prognostizierbar (Position Lesekopf)

## Hauptspeicher

- ist schnell
- wird immer billiger (2TB  $\approx$  20.000 Euro)
- Zugriffszeit exakt prognostizierbar

## Festplatten

- sind langsam
- werden immer billiger (2TB  $\approx$  200 Euro)
- Zugriffszeit nicht exakt prognostizierbar (Position Lesekopf)

## Hauptspeicher

- ist schnell
- wird immer billiger (2TB  $\approx$  20.000 Euro)
- Zugriffszeit exakt prognostizierbar

## Cache

- Level 1, Level 2, Level 3
- Faktor 10 schneller als RAM
- Spaltenorientierung reduziert Page Misses
- Prefetching

## Warum Komprimierung

- Flaschenhals: Speicher
- leichtgewichtige vs. ~~schwergewichtige~~ Komprimierung
- Anwendung: Enterprise Applications (redundante Daten,  $\approx 10\%$ )



## Warum Komprimierung

- Flaschenhals: Speicher
- leichtgewichtige vs. ~~schwergewichtige~~ Komprimierung
- Anwendung: Enterprise Applications (redundante Daten,  $\approx 10\%$ )

## Verwendete Methoden

Wörterbuch Kodierung, Präfix Kodierung, Lauflängen Kodierung, Cluster Kodierung, Indirekte Kodierung, Delta Kodierung

## Warum Komprimierung

- Flaschenhals: Speicher
- leichtgewichtige vs. schwergewichtige Komprimierung
- Anwendung: Enterprise Applications (redundante Daten,  $\approx 10\%$ )

## Verwendete Methoden

Wörterbuch Kodierung, Präfix Kodierung, Lauflängen Kodierung, Cluster Kodierung, Indirekte Kodierung, Delta Kodierung

spaltenorientiert:

...	Kunde2	Kunde3	...	Österreich	Österreich	Deutschland
-----	--------	--------	-----	------------	------------	-------------

## Warum Komprimierung

- Flaschenhals: Speicher
- leichtgewichtige vs. schwergewichtige Komprimierung
- Anwendung: Enterprise Applications (redundante Daten,  $\approx 10\%$ )

## Verwendete Methoden

Wörterbuch Kodierung, Präfix Kodierung, Lauflängen Kodierung, Cluster Kodierung, Indirekte Kodierung, Delta Kodierung

spaltenorientiert:

...	Kunde2	Kunde3	...	Österreich	Österreich	Deutschland		
<table border="1"> <tr> <td>Deutschland</td> </tr> <tr> <td>Österreich</td> </tr> </table>							Deutschland	Österreich
Deutschland								
Österreich								

## Warum Komprimierung

- Flaschenhals: Speicher
- leichtgewichtige vs. schwergewichtige Komprimierung
- Anwendung: Enterprise Applications (redundante Daten,  $\approx 10\%$ )

## Verwendete Methoden

Wörterbuch Kodierung, Präfix Kodierung, Lauflängen Kodierung, Cluster Kodierung, Indirekte Kodierung, Delta Kodierung

spaltenorientiert:

...	Kunde2	Kunde3	...	2	2	1
-----	--------	--------	-----	---	---	---

1  $\mapsto$  Deutschland

2  $\mapsto$  Österreich

## Warum Komprimierung

- Flaschenhals: Speicher
- leichtgewichtige vs. schwergewichtige Komprimierung
- Anwendung: Enterprise Applications (redundante Daten,  $\approx 10\%$ )

## Verwendete Methoden

Wörterbuch Kodierung, **Präfix Kodierung**, Lauflängen Kodierung, Cluster Kodierung, Indirekte Kodierung, Delta Kodierung

spaltenorientiert:

...	Kunde2	Kunde3	...	2	2	1
-----	--------	--------	-----	---	---	---

1  $\mapsto$  Deutschland

2  $\mapsto$  Österreich

## Warum Komprimierung

- Flaschenhals: Speicher
- leichtgewichtige vs. schwergewichtige Komprimierung
- Anwendung: Enterprise Applications (redundante Daten,  $\approx 10\%$ )

## Verwendete Methoden

Wörterbuch Kodierung, Präfix Kodierung, **Laufängen Kodierung**, Cluster Kodierung, Indirekte Kodierung, Delta Kodierung

spaltenorientiert:

...	Kunde2	Kunde3	...	2	2	1
-----	--------	--------	-----	---	---	---

1  $\mapsto$  Deutschland

2  $\mapsto$  Österreich

## Speicherorganisation: Problem

Spaltenorientiertes Layout langsam für Schreiben (insert/update/delete)

## Speicherorganisation: Problem

Spaltenorientiertes Layout langsam für Schreiben (insert/update/delete)

## Lösung

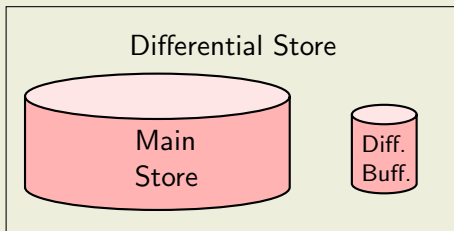
Differential Store



## Speicherorganisation: Problem

Spaltenorientiertes Layout langsam für Schreiben (insert/update/delete)

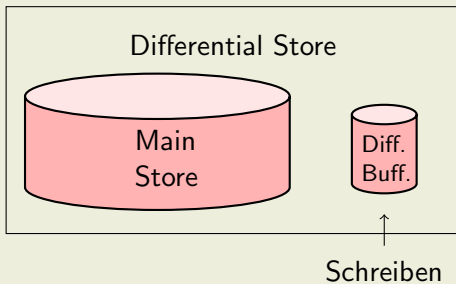
## Lösung



## Speicherorganisation: Problem

Spaltenorientiertes Layout langsam für Schreiben (insert/update/delete)

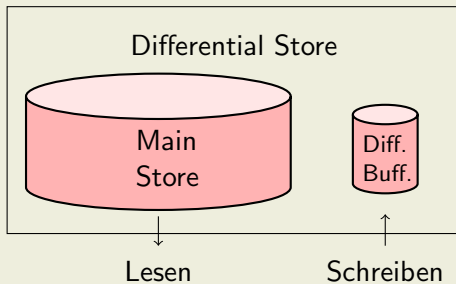
## Lösung



## Speicherorganisation: Problem

Spaltenorientiertes Layout langsam für Schreiben (insert/update/delete)

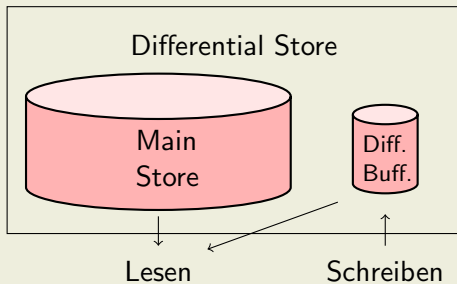
## Lösung



## Speicherorganisation: Problem

Spaltenorientiertes Layout langsam für Schreiben (insert/update/delete)

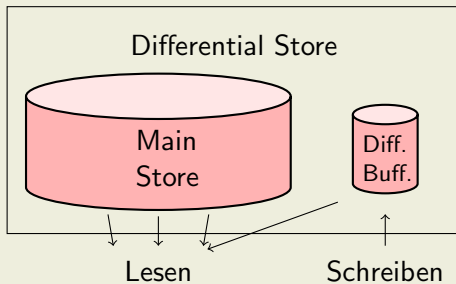
## Lösung



## Speicherorganisation: Problem

Spaltenorientiertes Layout langsam für Schreiben (insert/update/delete)

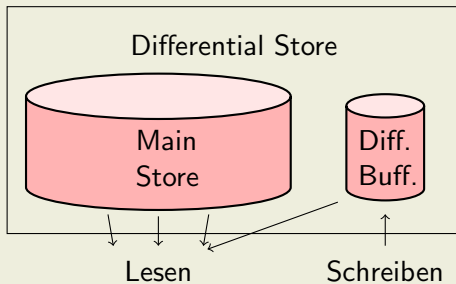
## Lösung



## Speicherorganisation: Problem

Spaltenorientiertes Layout langsam für Schreiben (insert/update/delete)

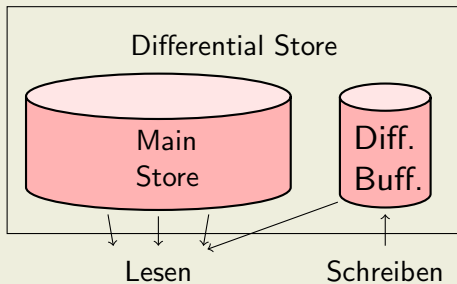
## Lösung



## Speicherorganisation: Problem

Spaltenorientiertes Layout langsam für Schreiben (insert/update/delete)

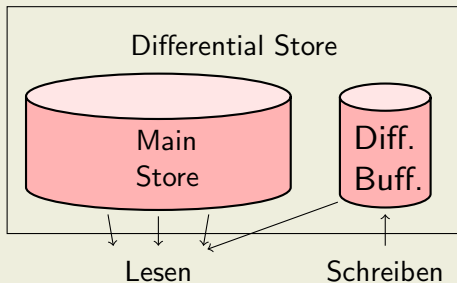
## Lösung



## Speicherorganisation: Problem

Spaltenorientiertes Layout langsam für Schreiben (insert/update/delete)

## Lösung



## Herausforderung: Merge

- integriere Differential Buffer in Main Store
- wann: Nacht/Wochenende bzw. extern



## Operationen (parallel/verteilt)

- SQL
- insert, delete, update, join, scan

## Operationen (parallel/verteilt)

- SQL
- insert, delete, update, join, scan

## Partitionierung

id	name	zip	country
1	Kunde1	6020	Österreich
2	Kunde2	9900	Österreich
3	Kunde3	80331	Deutschland

- horizontal vs. vertikal

## Operationen (parallel/verteilt)

- SQL
- insert, delete, update, join, scan

## Partitionierung

id	name	zip	country
1	Kunde1	6020	Österreich
2	Kunde2	9900	Österreich
3	Kunde3	80331	Deutschland

- **horizontal** vs. vertikal
- zeilenorientiert spaltenorientiert

## Operationen (parallel/verteilt)

- SQL
- insert, delete, update, join, scan

## Partitionierung

id	name	zip	country
1	Kunde1	6020	Österreich
2	Kunde2	9900	Österreich
3	Kunde3	80331	Deutschland

- horizontal vs. vertikal
- zeilenorientiert + horizontal vs. spaltenorientiert

## Operationen (parallel/verteilt)

- SQL
- insert, delete, update, join, scan

## Partitionierung

id	name	id	zip	country
1	Kunde1	1	6020	Österreich
2	Kunde2	2	9900	Österreich
3	Kunde3	3	80331	Deutschland

- horizontal vs. **vertikal**
- zeilenorientiert + horizontal vs. spaltenorientiert + vertikal

## Operationen (parallel/verteilt)

- SQL
- insert, delete, update, join, scan

## Partitionierung

id	name	id	zip	country
1	Kunde1	1	6020	Österreich
2	Kunde2	2	9900	Österreich
3	Kunde3	3	80331	Deutschland

- horizontal vs. vertikal
- zeilenorientiert + horizontal vs. spaltenorientiert + vertikal

## Insert-only

- erlaubt historische Abfragen
- effizienter

## Scheduling

- Motivation: OLTP (schnell) und OLAP (langsam) zur selben Zeit
- Realisierung: getrennte Thread Pools (für OLTP und OLAP)
- Scheduling Algorithmen

## Scheduling

- Motivation: OLTP (schnell) und OLAP (langsam) zur selben Zeit
- Realisierung: getrennte Thread Pools (für OLTP und OLAP)
- Scheduling Algorithmen

## Data Aging

- aktive Daten                      passive Daten



## Scheduling

- Motivation: OLTP (schnell) und OLAP (langsam) zur selben Zeit
- Realisierung: getrennte Thread Pools (für OLTP und OLAP)
- Scheduling Algorithmen

## Data Aging

- aktive Daten (20%) vs. passive Daten (80%)

## Scheduling

- Motivation: OLTP (schnell) und OLAP (langsam) zur selben Zeit
- Realisierung: getrennte Thread Pools (für OLTP und OLAP)
- Scheduling Algorithmen

## Data Aging

- aktive Daten (20%) vs. passive Daten (80%)

## Benchmarking

- Full-table Scan
- Schritt Scan
- TPC-C

## Problem

- Stromausfall: dynamischer RAM verliert Daten (volatil)

## Problem

- Stromausfall: dynamischer RAM verliert Daten (volatil)

## Lösung

- persistenter Speicher
  - Logs
  - Checkpoints

## SAP HANA

- High Performane Analytic Appliance
- 2010 vorgestellt
- Hauptspeicherdatenbank
- integriert viele Konzepte von SanssouciDB

## SAP HANA

- High Performane Analytic Appliance
- 2010 vorgestellt
- Hauptspeicherdatenbank
- integriert viele Konzepte von SanssouciDB

## Anwendungsentwicklung

- engere Kooperation zwischen Business Logik und Datenbank
- Stored procedures

## SAP HANA

- High Performane Analytic Appliance
- 2010 vorgestellt
- Hauptspeicherdatenbank
- integriert viele Konzepte von SanssouciDB

## Anwendungsentwicklung

- engere Kooperation zwischen Business Logik und Datenbank
- Stored procedures

## Fallstudien

- Charité (Berlin)
- Hilti (Liechtenstein)

## Einsatz von Hauptspeicherdatenbanken

- Firmen skeptisch (wenig Erfahrungswerte)
- Vorschlag: hybride Systeme (zwischenzeitlich)



## Einsatz von Hauptspeicherdatenbanken

- Firmen skeptisch (wenig Erfahrungswerte)
- Vorschlag: hybride Systeme (zwischenzeitlich)

## neue Hardware (Cloud Computing)

- Skalierbarkeit
- Verfügbarkeit
- Risikominimierung

## Einsatz von Hauptspeicherdatenbanken

- Firmen skeptisch (wenig Erfahrungswerte)
- Vorschlag: hybride Systeme (zwischenzeitlich)

## neue Hardware (Cloud Computing)

- Skalierbarkeit
- Verfügbarkeit
- Risikominimierung

## neue Software (neue Möglichkeiten)

- Echtzeitanalyse möglich
- Smart Meter, Patientendaten

## Zusammenfassung

- Hauptspeicherdatenbanken
- Merkmale
  - Multi-core, Spaltenorientiert, Insert-only, Komprimierung
- gelöste Herausforderungen
  - verteiltes System, Speicherorganisation
- Einsatzgebiet
  - Enterprise Applications

## Zusammenfassung

- Hauptspeicherdatenbanken
- Merkmale
  - Multi-core, Spaltenorientiert, Insert-only, Komprimierung
- gelöste Herausforderungen
  - verteiltes System, Speicherorganisation
- Einsatzgebiet
  - Enterprise Applications

Vielen Dank für die Aufmerksamkeit