Trends and Concepts in the Software Industry I

Summer 2010

HPI Hasso Plattner Institut
IT Systems Engineering | Universität Potsdam

Complexity of enterprise applications

EA requirements are diverse and complex

SAP for Consumer Products:
customers with 1.4 million sales order line items per day

SAP Human Capital Management:
compute the payroll for 500.000 employees within hours

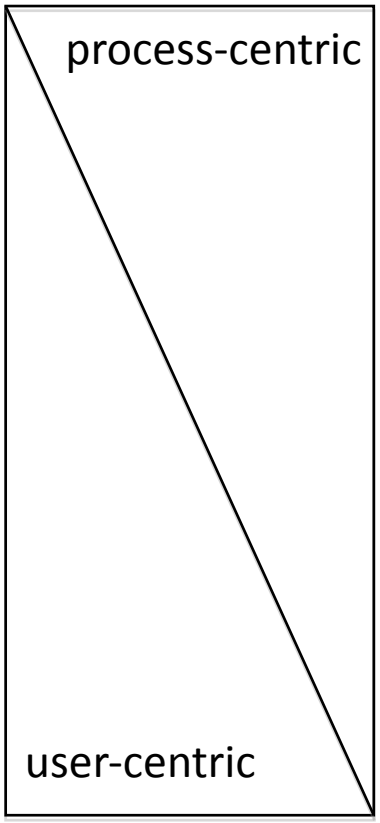
SAP Enterprise Resource Planning:
no maintenance downtime allowed for Apple iTunes Store

SAP Netweaver Portal:
handle 300.000 users in a web portal (40.000 concurrently active)

SAP Supply Chain Management:
customers with over 3 million different product configurations

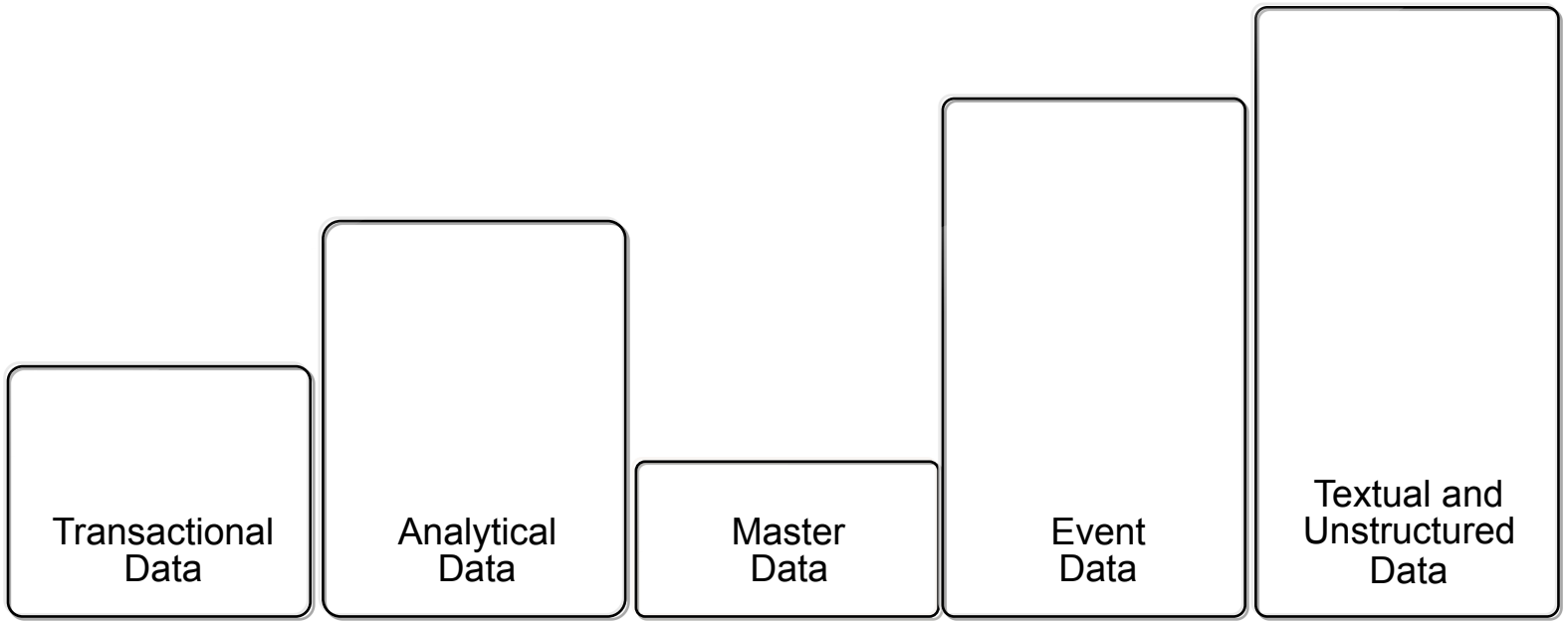
SAP Business Intelligence:
customer with a live database of 40 terabytes in size

EA have different interaction paradigms



- Automated, machine to machine interaction
- Repeated, batch-wise processing of data sets
- Data entry (repetitive, transactional, isolated)
- Transactional work with human workflow
- Collaborative workflow amongst people
- Unstructured, ad-hoc collaboration

EA process data that is not alike



So what ?

Facebook has 100 terabytes of data

So what ?

Facebook has 100 terabytes of new data per day

Reflect on what you've heard so far! Does Facebook …

… have to deal with different customer requirements in various industries in different countries ?

… have to be afraid of being sued when they have a downtime ?

… have the need for consistency ?

… have to provide an infrastructure that supports individual customizing ?

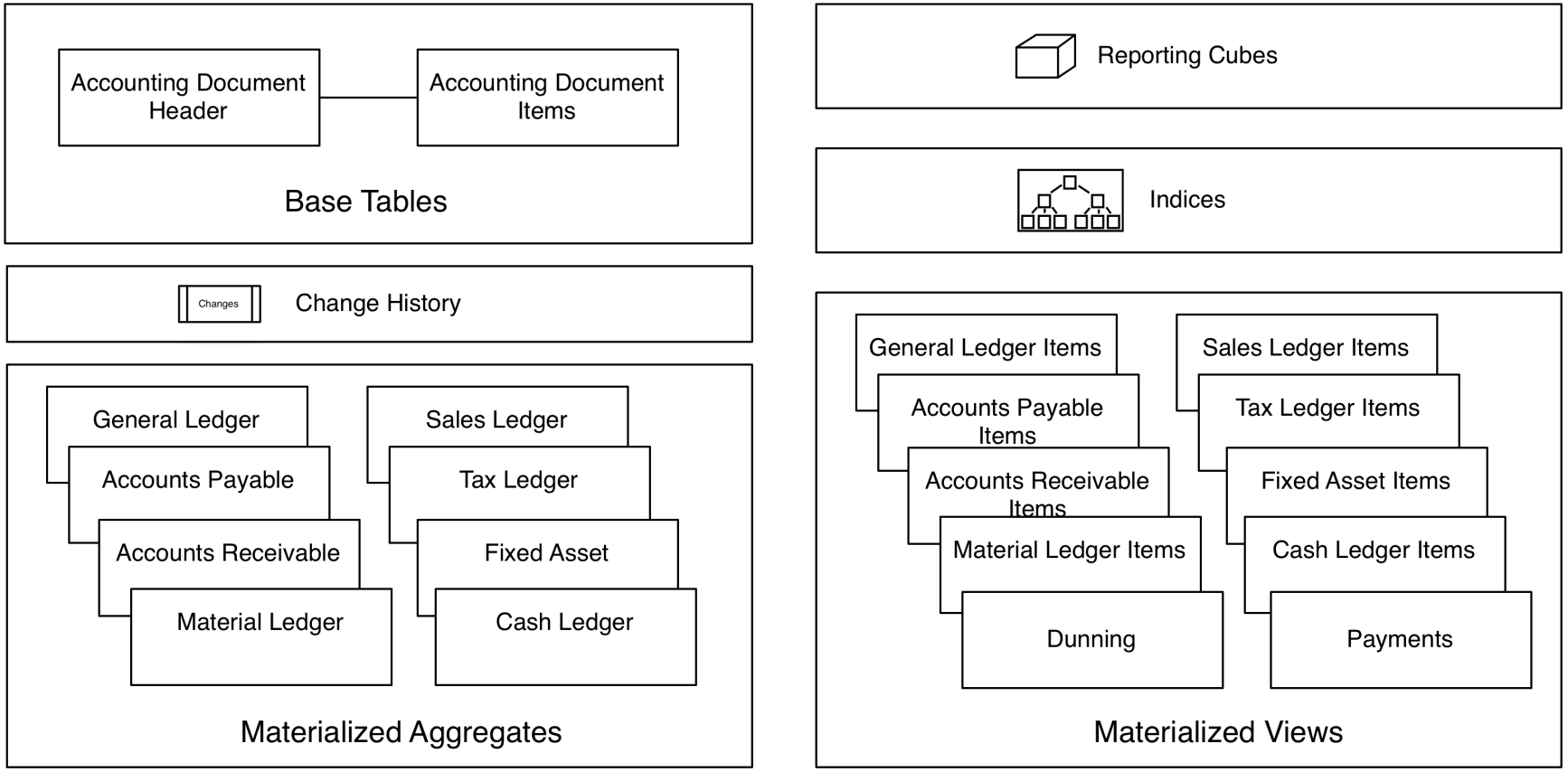
… have to consider different user interaction paradigms ?

… send you an invoice ?

HPI Hasso Plattner Institut
IT Systems Engineering | Universität Potsdam

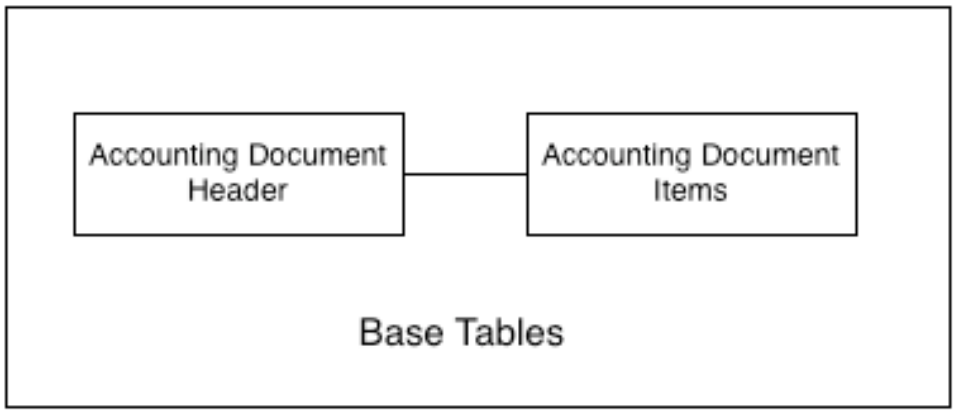
Financial accounting – extended example
OLTP & OLAP

Old financials system



Tomorrow‘s financials system

Only base table and algorithms



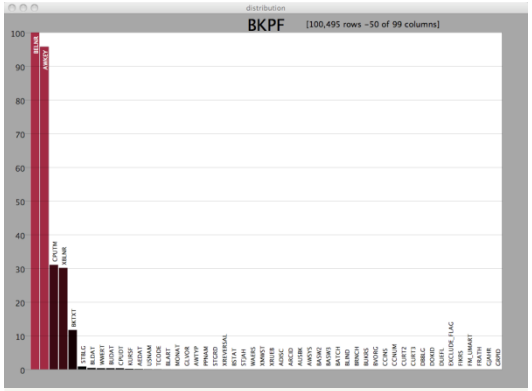
OLTP & OLAP characteristics

OLTP	OLAP/Decision Support Systems
Full row operations	Retrieve small number of columns
Simple Queries	Complex Queries
Detail Row Retrieval	Aggregation and Group By
Inserts/Updates/Selects	Mainly Selects
Short Transactions	Long Transactions
Small Found Sets	Large Found Sets
Pre-determined Queries	Adhoc Queries
Real Time Updates	Batch Updates
„Source of Truth“	Alternative representation

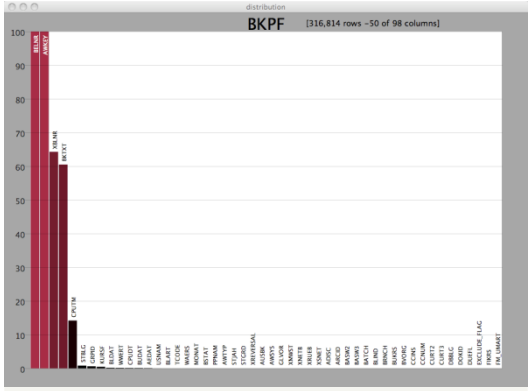
Accounting document headers

~99 attributes

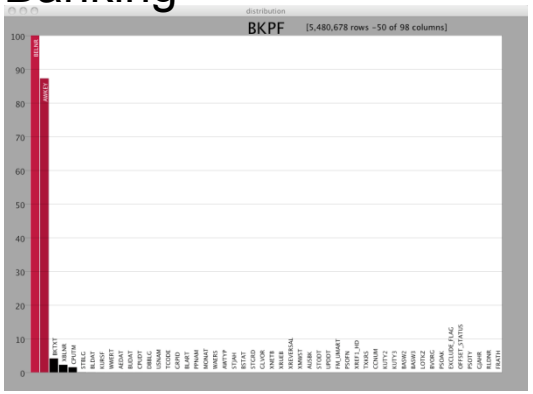
Consumer Goods

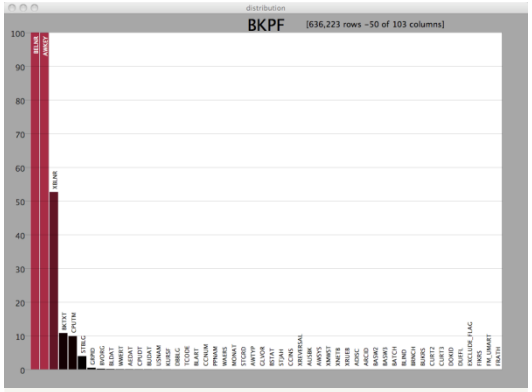


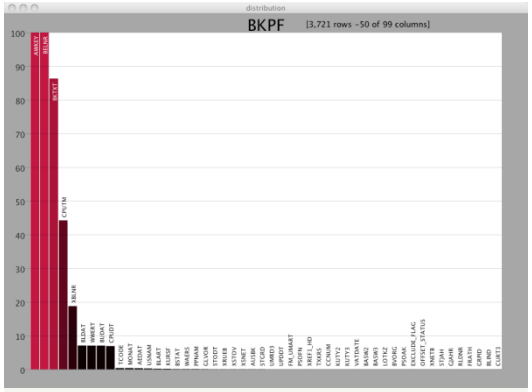
Logistics

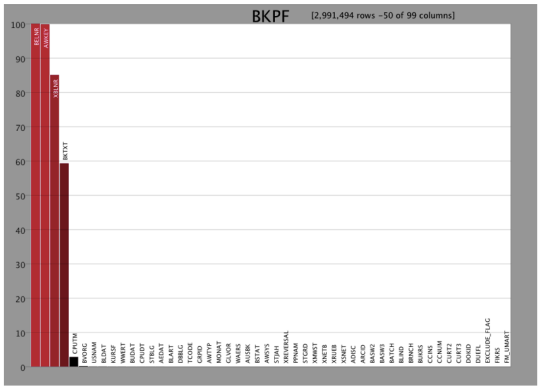


Banking









High Tech

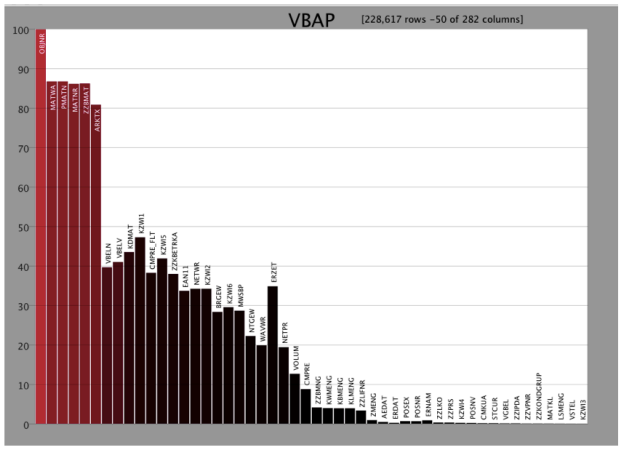
Mechanical/
Industrial Engineering

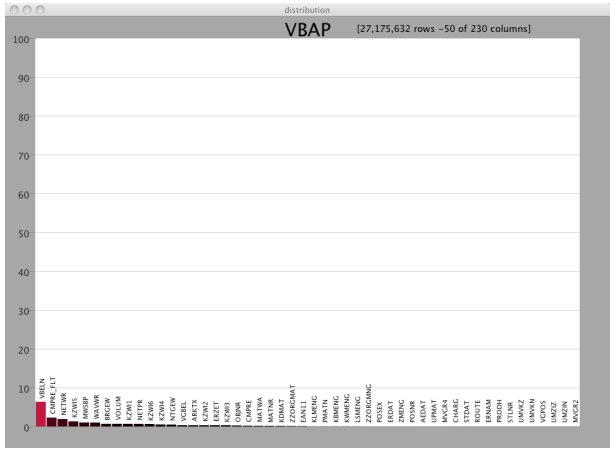
Electrical/
Electronic
Manufact.

Sales order items

~250 attributes

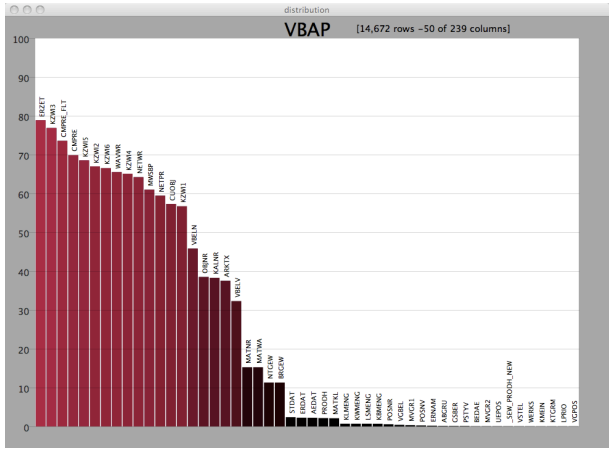
Electrical/
Electronic
Manufact.

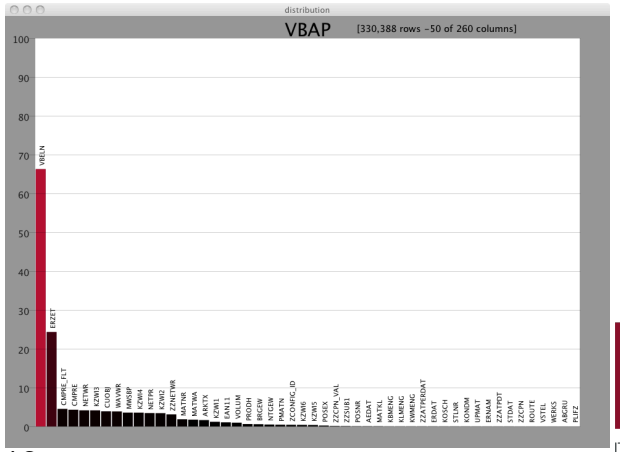




Consumer
Goods

Mechanical/
Industrial
Engineering

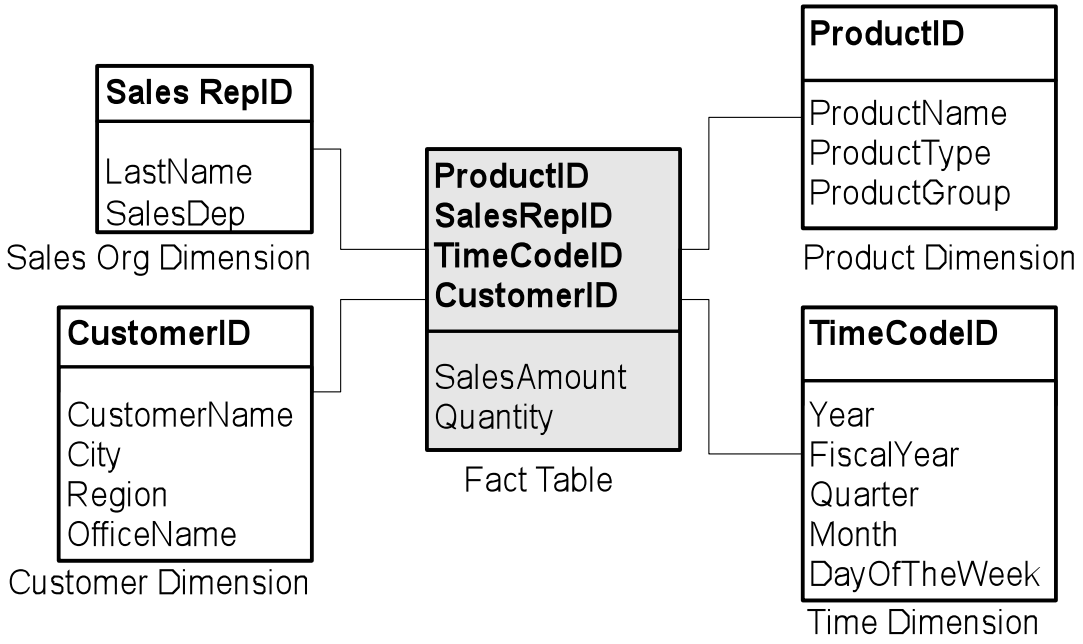


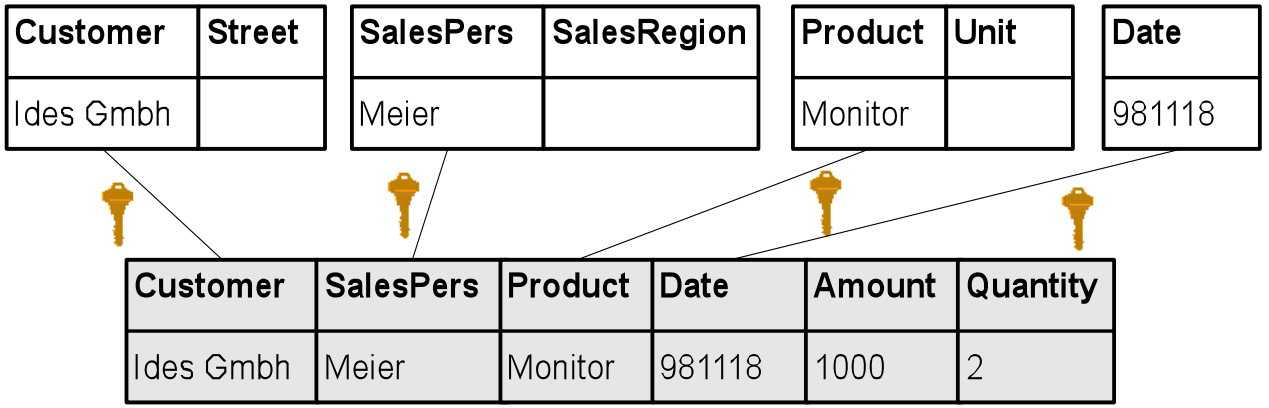


Computer
Hardware/
Software

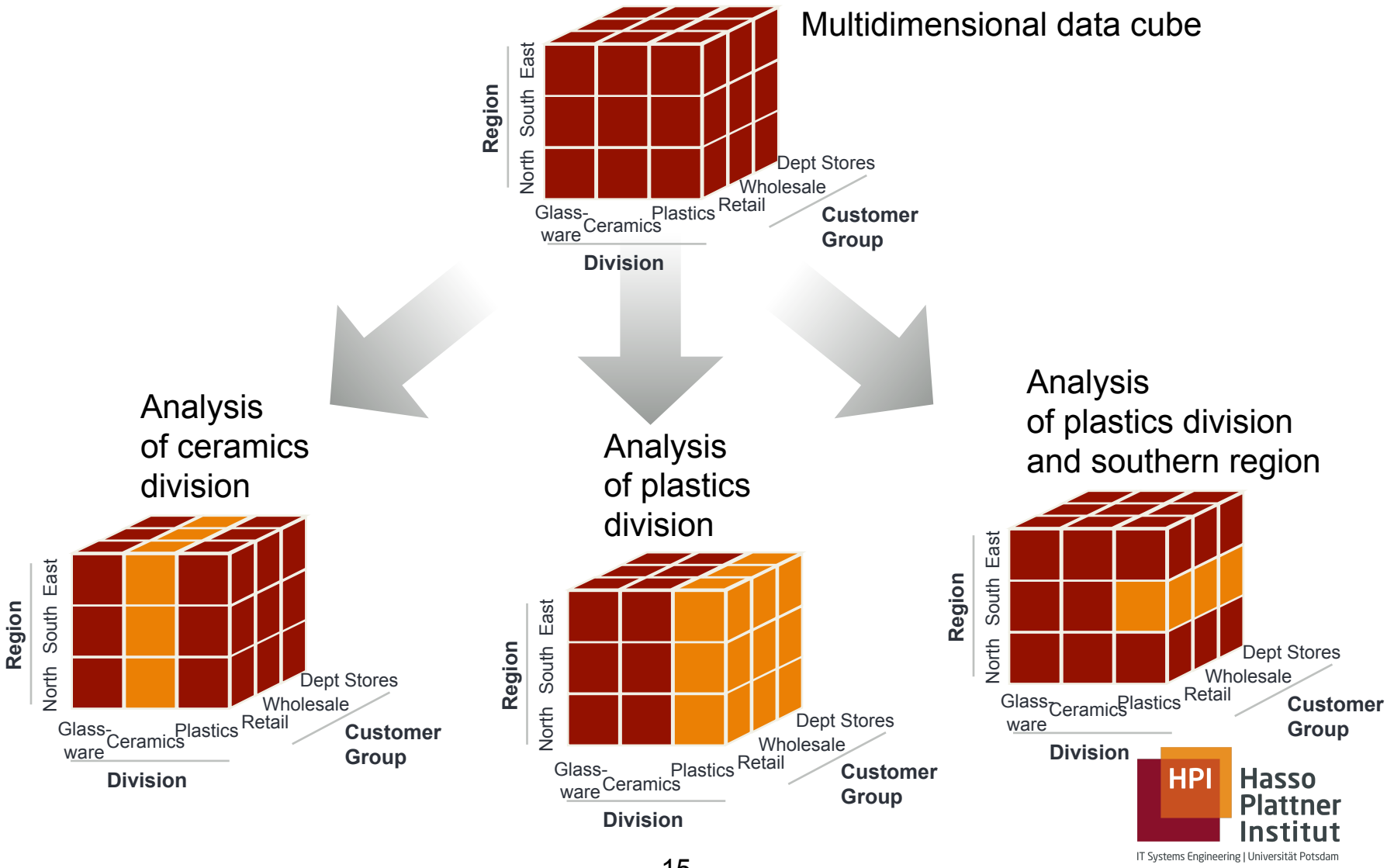
HPI Hasso Plattner Institut
IT Systems Engineering | Universität Potsdam

Star schema example





Multi-dimensional analysis



In-memory databases

Rows vs. columns

Document Number	Document Date	Sold-To Party	Order Value	Status	Sales Organization	...
95769214	2009-10-01	584	10.24	CLOSED	Germany Frankfurt	...
95769215	2009-10-01	1215	124.35	CLOSED	Germany Berlin	...
95779216	2009-10-21	584	47.11	OPEN	Germany Berlin	...
95779217	2009-10-21	454	21.20	OPEN	Germany Frankfurt	...

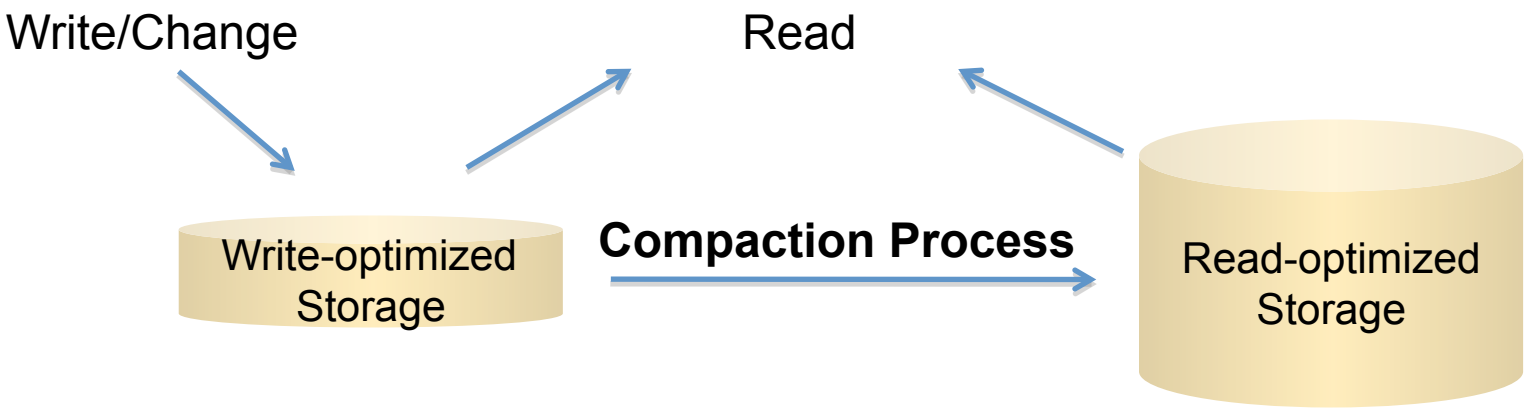
Row Store

95769214 2009-10-01 584 10.24 CLOSED Germany Frankfurt ... 95769215 2009-10-01 1215 124.35 CLOSED Germany Berlin ... 95779216 2009-10-21 584 47.11 OPEN Germany Berlin ... 95779217 2009-10-21 454 21.20 OPEN Germany Frankfurt

Column Store

95769214 95769215 95779216 95779217 ...
2009-10-01 2009-10-21 2009-10-21 2009-10-21 ...
584 1215 584 454 ...
10.24 124.35 47.11 21.20 ...
CLOSED CLOSED OPEN OPEN ...
Germany Frankfurt Germany Berlin Germany Berlin Germany Frankfurt ...

Read- and write-optimized storage



Compression techniques

	Few distinct values	Many distinct values
Ordered	Run-length encoding	Delta representation
Unordered	Bit vector encoding	Dictionary compression

Dictionary compression

Compressed Tables

Document Number	Document Date	Sold-To Party	Order Value
95769214	2009-10-01	584	10.24
95769215	2009-10-01	1215	124.35
95779216	2009-10-21	584	47.11
95779217	2009-10-21	454	21.20

Document Number	Document Date	Sold-To Party	Order Value
0	0	1	0
1	0	2	3
2	1	1	2
3	1	0	1

Dictionaries

Document Number

0	95769214
1	95769215
2	95779216
3	95779217

Order Value

0	10.24
1	21.20
2	47.11
3	124.35

Document Date

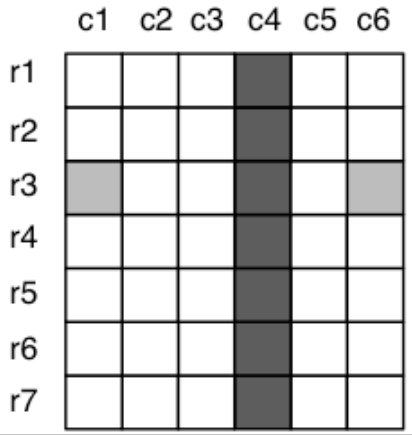
0	2009-10-01
1	2009-10-21

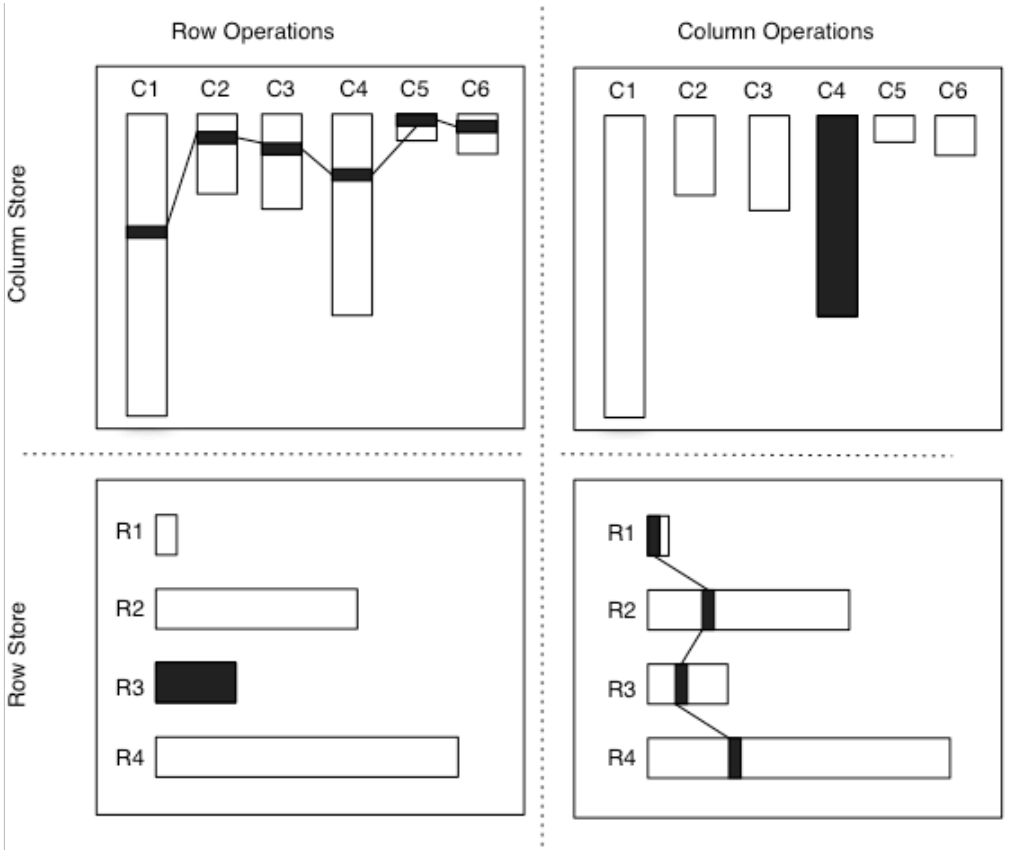
Sold-To Party

0	454
1	584
2	1215

Column store example

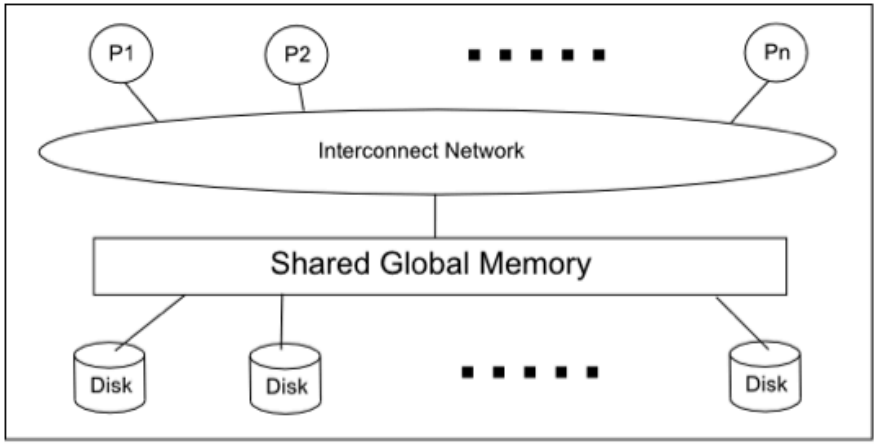
SELECT c1, c4, c6 FROM table WHERE c4 < ?



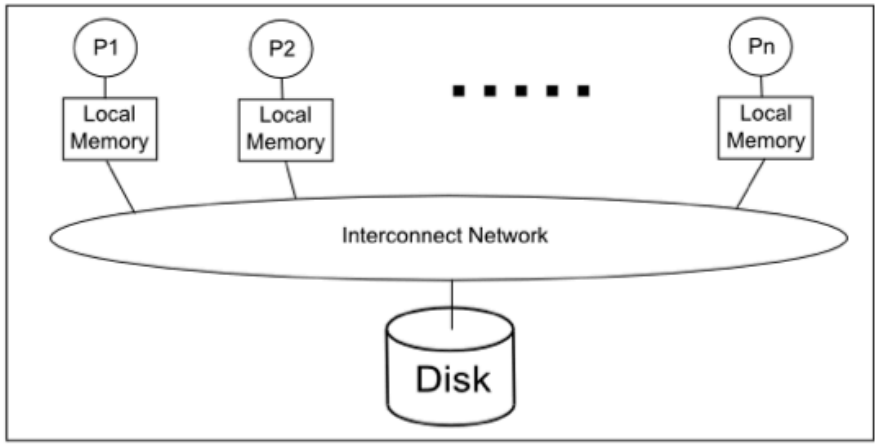


Parallelization

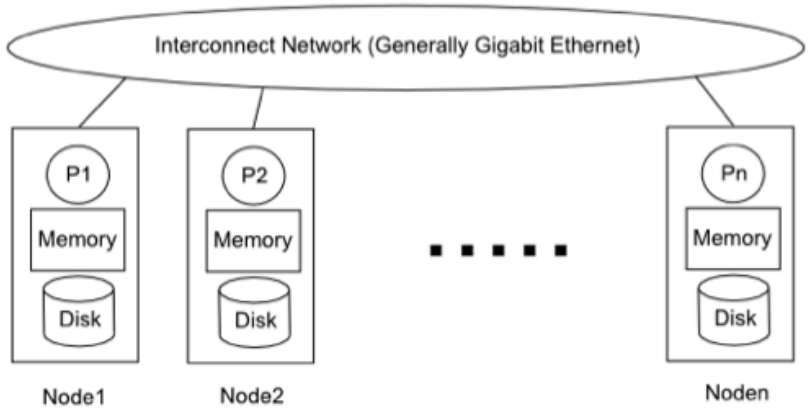
Parallel system architectures



(a) Shared Memory Architecture - Single Machine



(b) Shared Disk Architecture - Single Machine



(c) Shared Nothing Architecture - Multiple Machines

Operator parallelism

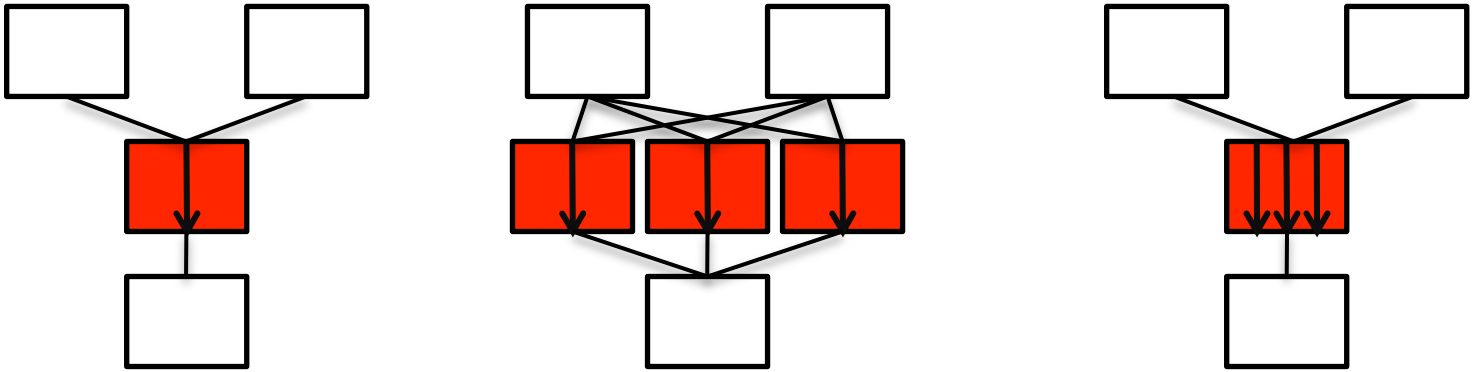
Alternative 1

Alternative 2

Serial

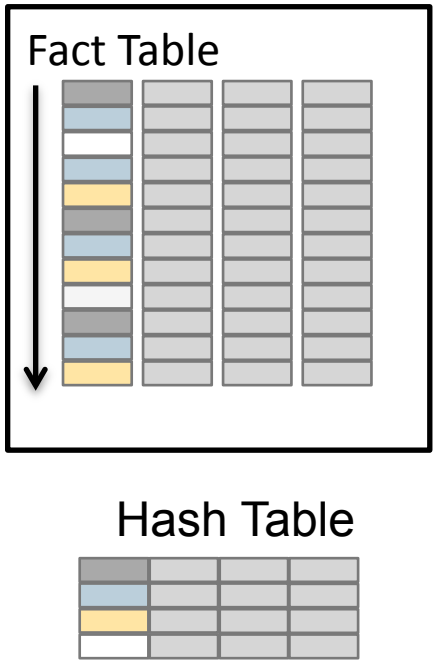
Inter-Operator Parallelism

Intra-Operator Parallelism

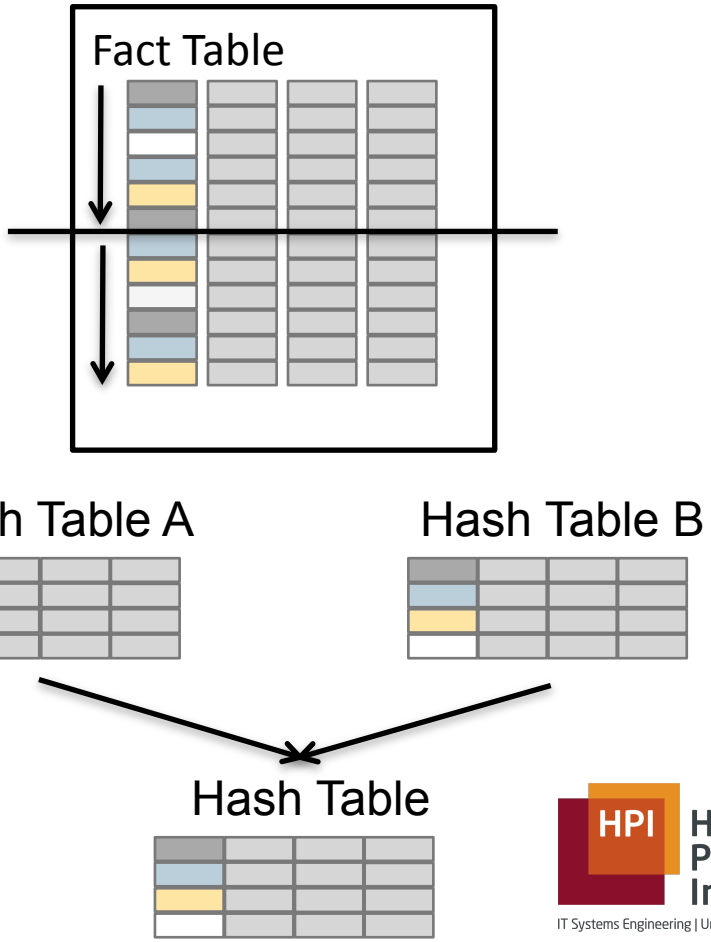


Aggregation

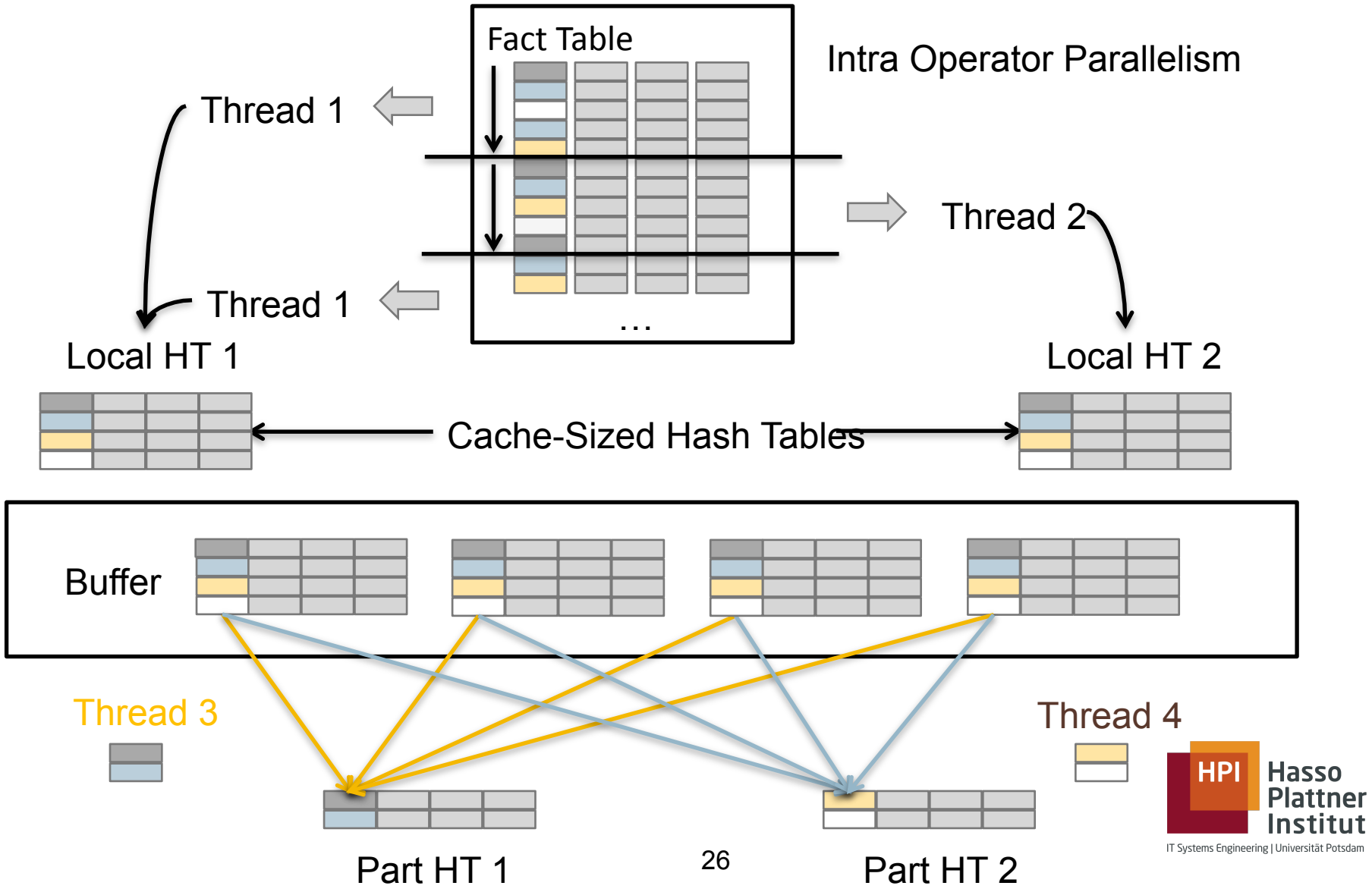
Serial



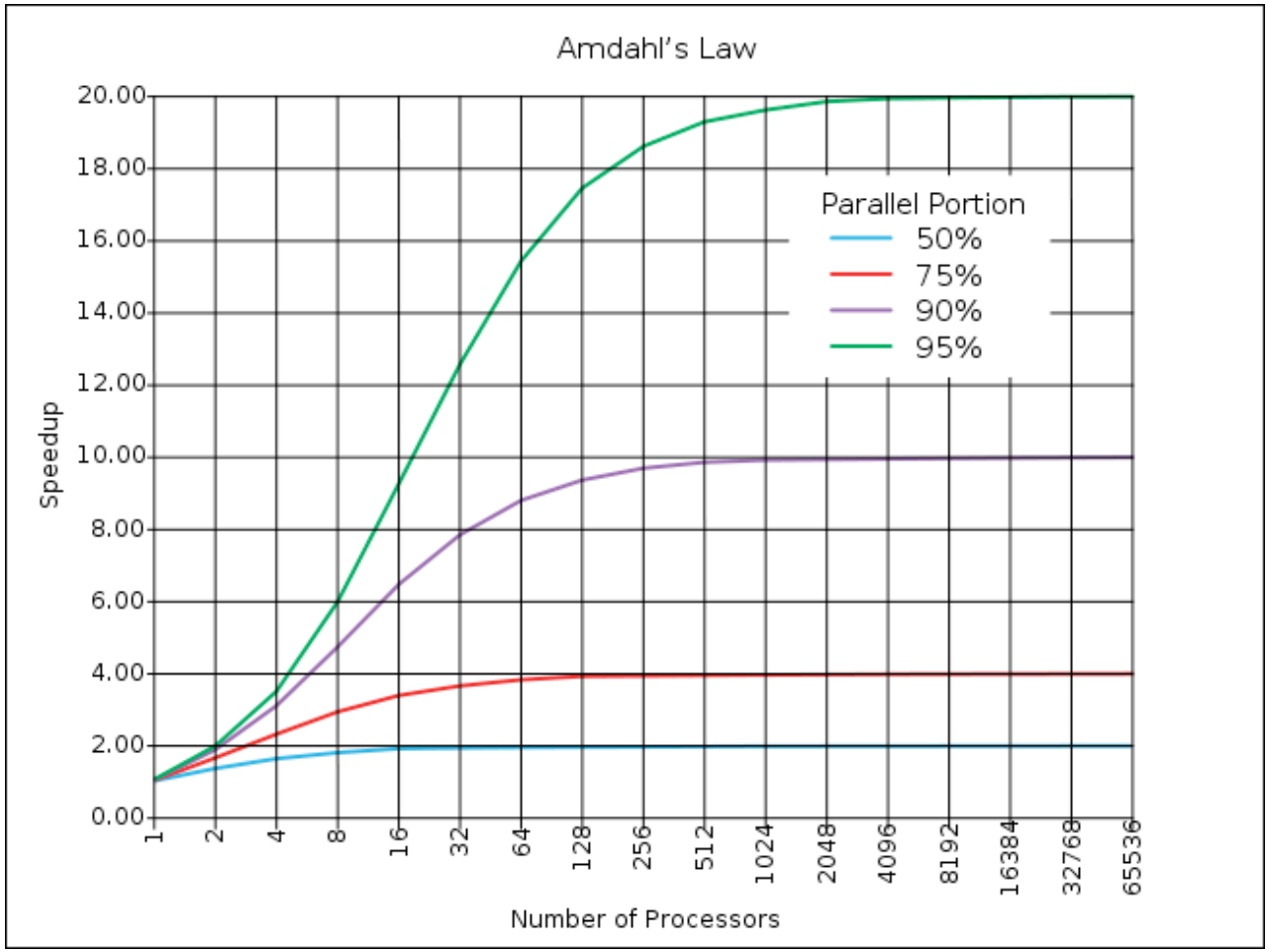
Inter Operator Parallelism



Aggregation w. multiple threads

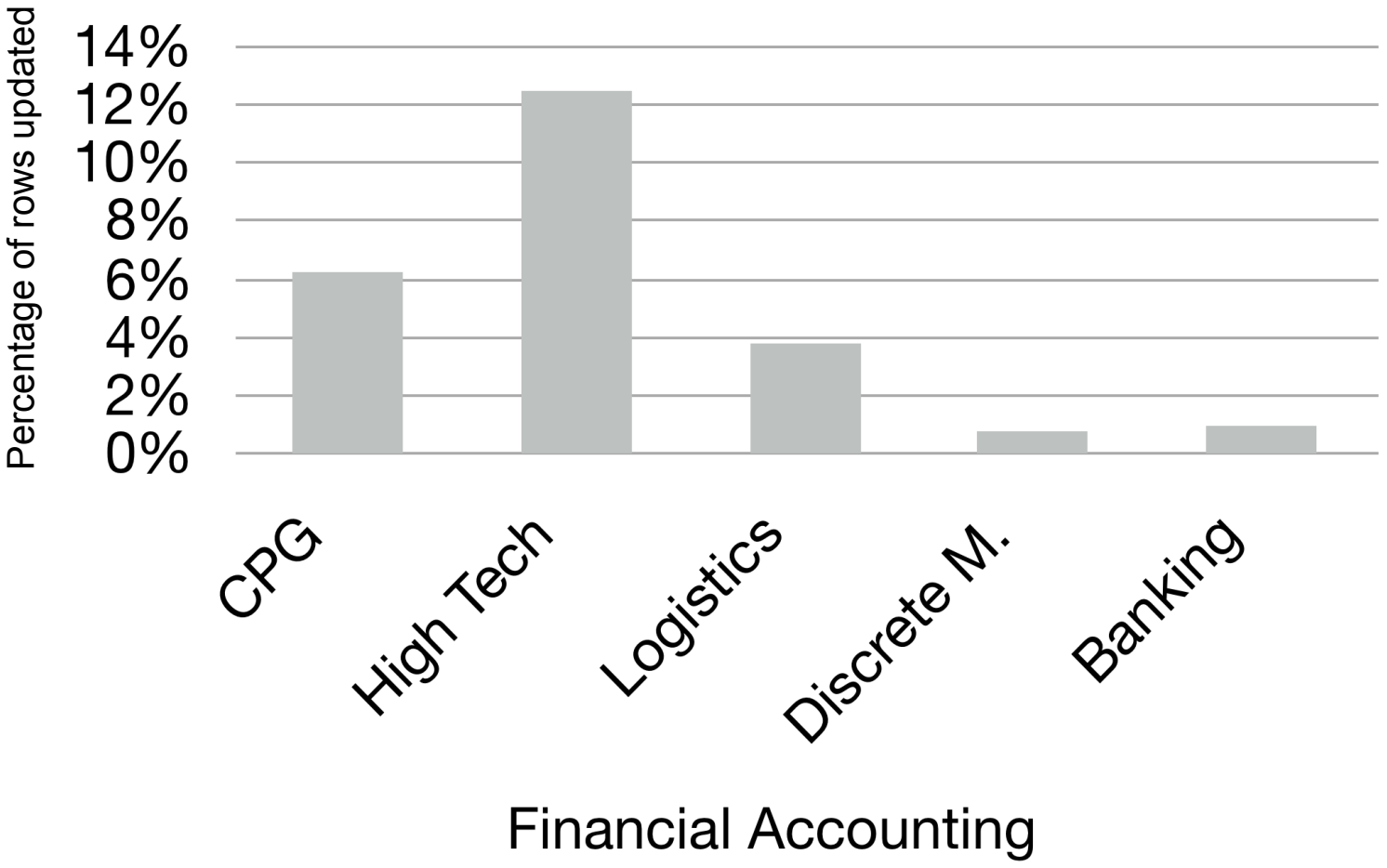


Limits of parallelism

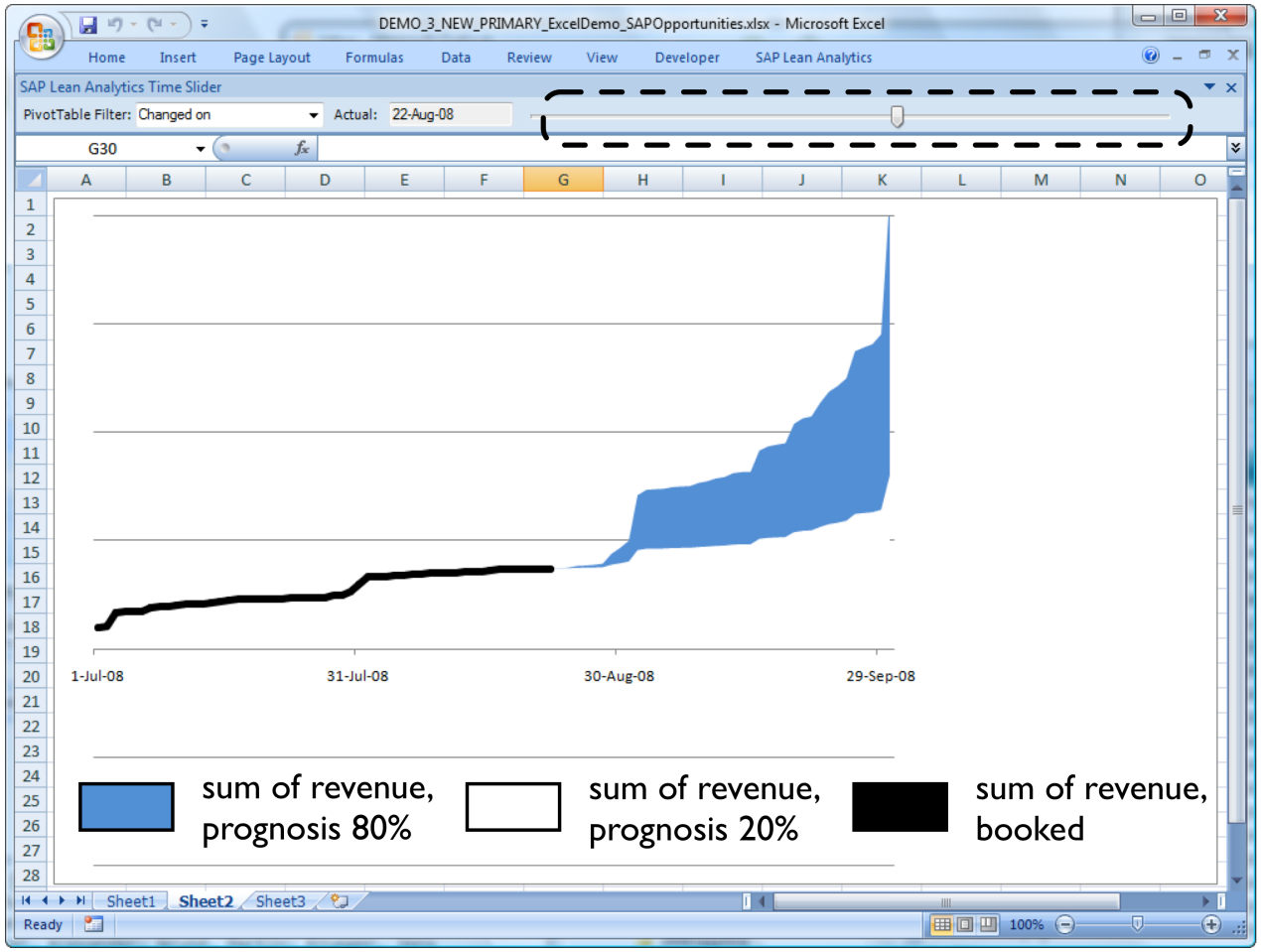


Insert-only and historical data

Value updates



Sales pipeline forecast



HPI Hasso Plattner Institut
IT Systems Engineering | Universität Potsdam

Use case: inventory management

Characteristics on material movements

(Table *MSEG from HANA2 customer ERP instance)*

- 80M material movements in MSEG
  (SAP ERP 6.0 used due to huge amounts of realistic data)
- 117.869 different materials,
  100 different storage locations
- Amount of movements/material
  scale from 1 to 1.407.401
- Data schema:
  - 2 **base** tables
  - 2 tables for materializing **aggregates**
    (plant/material and storage location/material)

Compression on material movements

(Table *MSEG from HANA2 customer ERP instance)*

	TREX	mySQL
MSEG wo/ indices	10.1 GB	51GB
MSEG w/ indices (mandt, mblnr, matnr, lgort, werks,)	10.7GB (+5.9%)	51GB + 9.4GB (+18.4%)

Used MSEG: 80M records

Characteristics on material movements

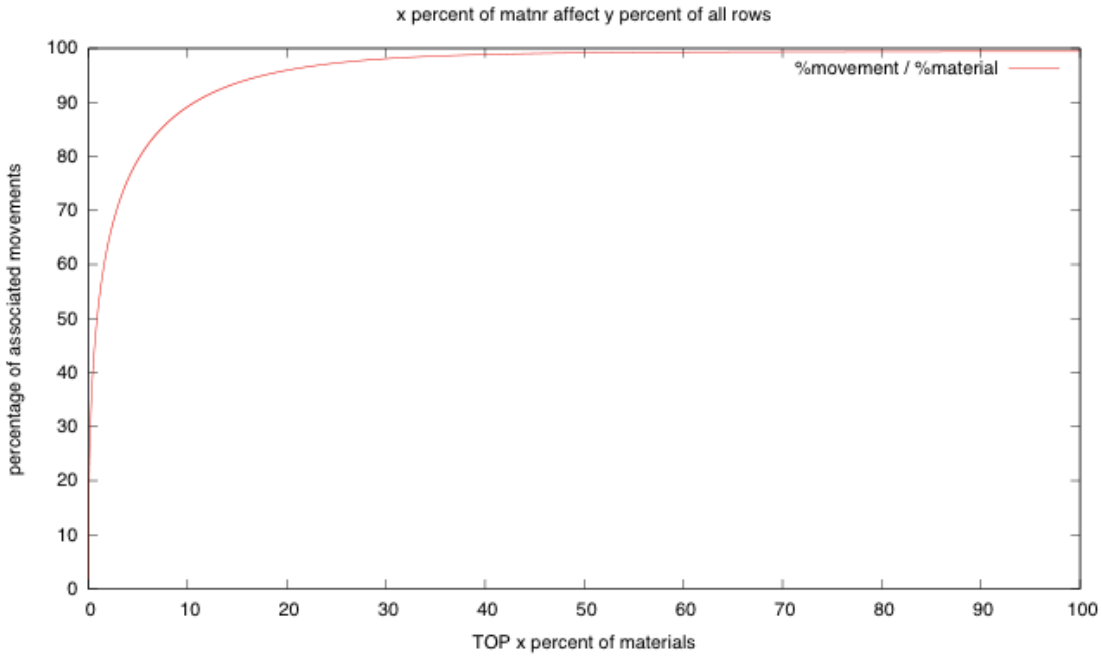
(Table *MARD from HANA2 customer ERP instance)*

- 628.413 records, 21MB on TREX, 144M mySQL (materialized **aggregates**)
- keeps track of amount of stock at certain location
- Select single
  - **TREX**: simple lookup wo/ index: 2ms, w/ index: 0.5ms
  - **mySQL**: key lookup w/ index: 10-20ms uncached, 1-4ms cached

Question:

How many different materials generate xx% of all material movements?

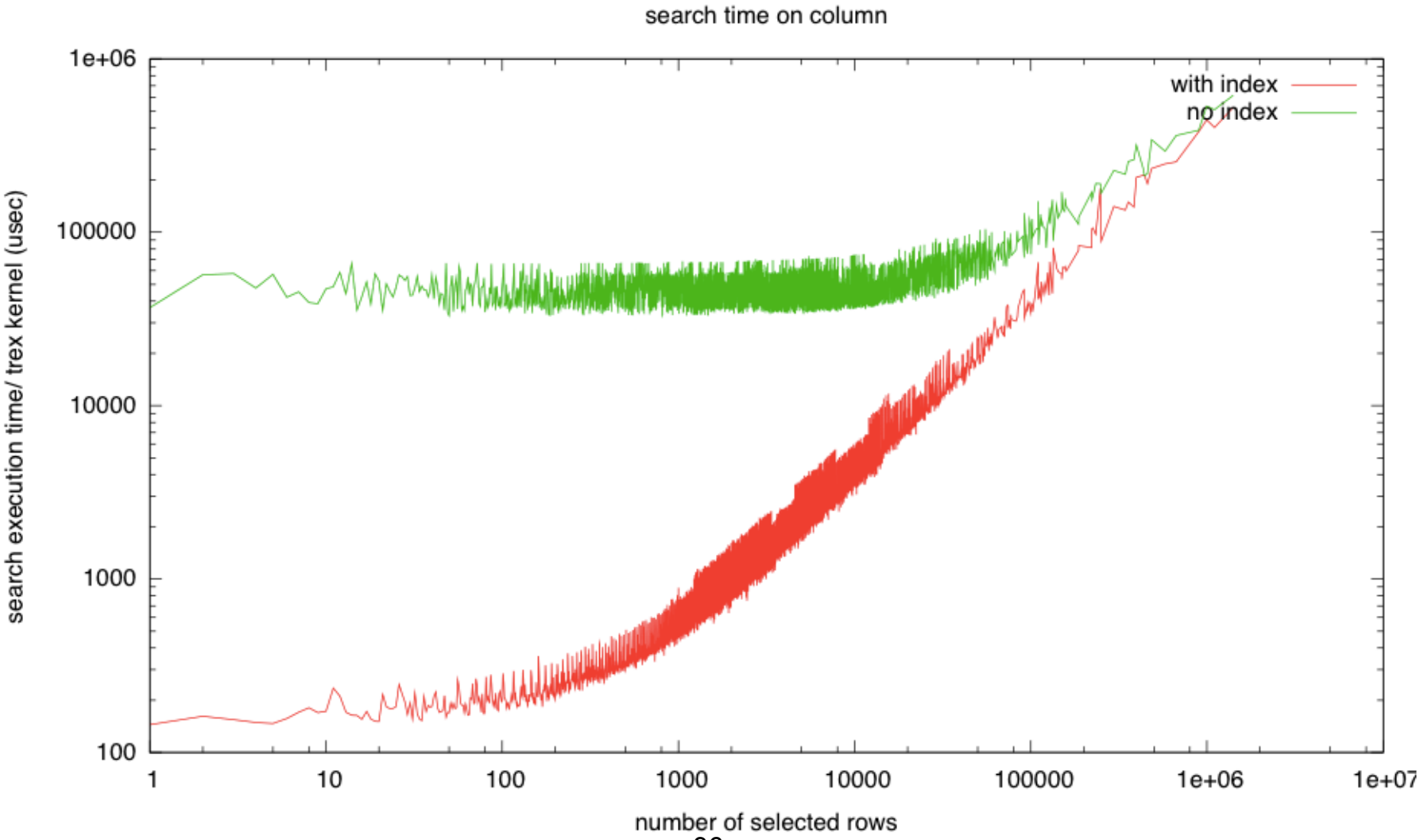
# Materials	# Materials in %	Material Movements	Movements in %
1	0.00	1407401	1.75
4	0.00	4404921	5.47
12	0.01	8253393	10.25
30	0.03	12092634	15.02
68	0.06	16127953	20.03
231	0.20	24182013	30.03
531	0.45	32216729	40.01
1037	0.88	40276183	50.01
1874	1.59	48319632	60
3347	2.84	56374387	70
6142	5.21	64425848	80
8592	7.29	68451732	58
12612	10.70	72477980	90
20819	**17.66**	**76504573**	**95**



Answer: 17% of all different materials generate **95%** of the material movements.

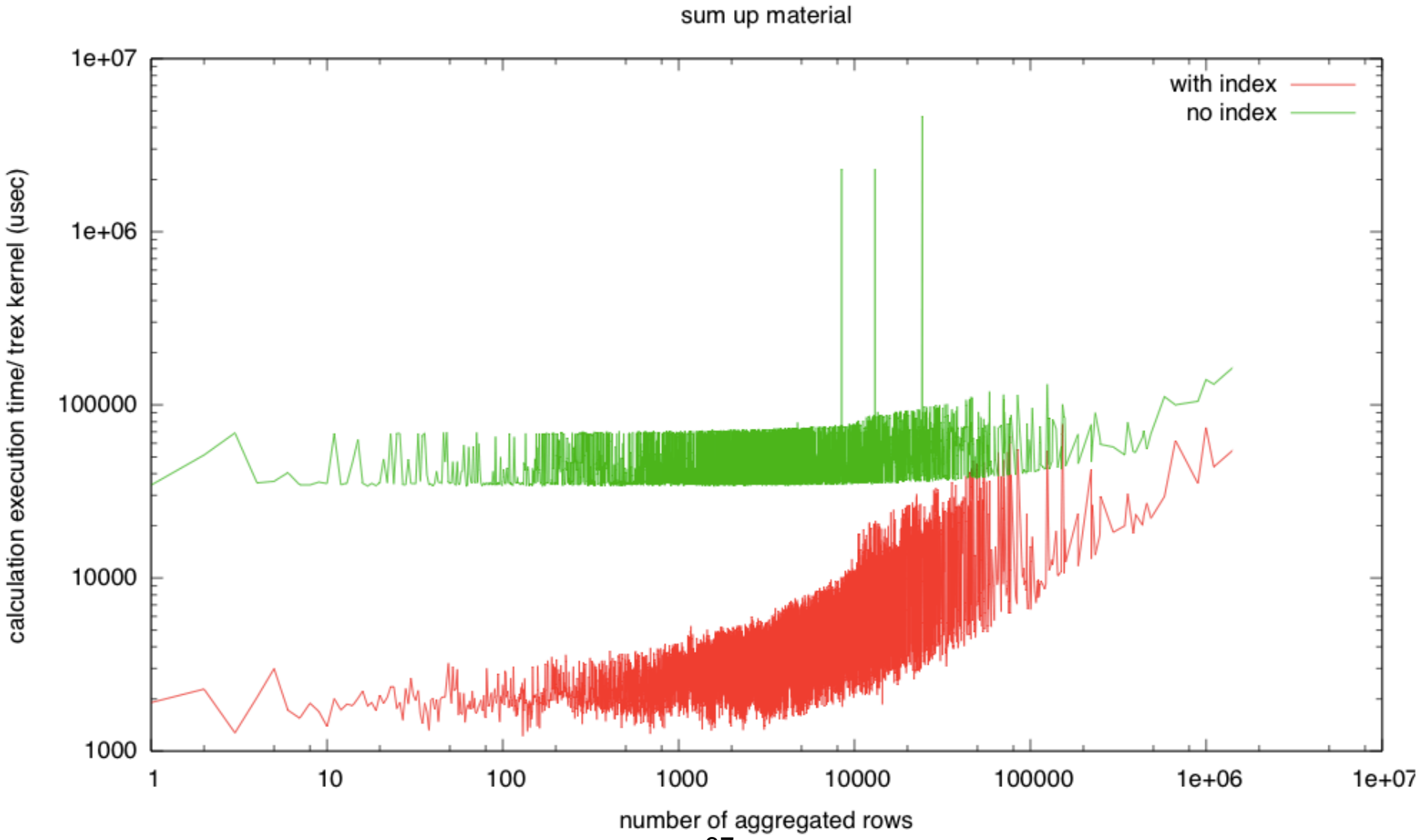
OLTP-style query: single select operation

SELECT matnr FROM mseg WHERE matnr=?



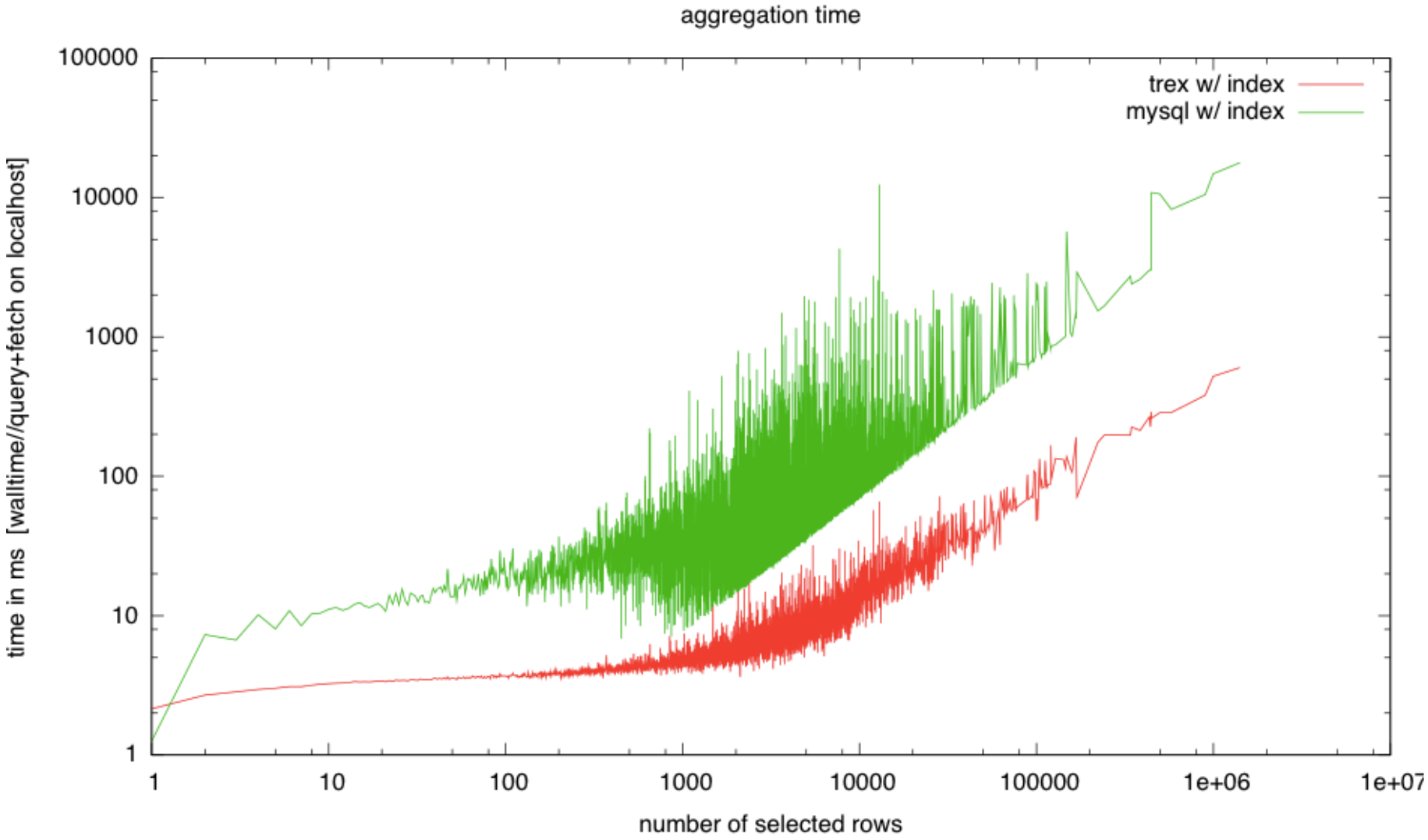
Compute stock at specific storage location on-the-fly

Select sum(menge) from mseg where lgort=0100 and matnr=?



Compute stock at specific storage location on-the-fly

Select sum(menge) from mseg where lgort=0100 and matnr=?



Use case: dunning

Quantity structure (I)

- European division of consumer packaged goods (CPG) company
- **5.5 TB** database size
- ERP 2005 Release ECC 6.0 EhP3
- Accounting documents
  - **23 million** headers / 8 GB in main memory
  - **252 million** items / 50 GB in main memory

Quantity structure (II)

- **Financials**
  - 250mio accounting document line items,
  - 380k open items accounts receivable,
  - 200k overdue open items
- **Master data**
  - 200k customers
  - 10k customers with open items

Dunning (I)

- Dunning is the process of scanning through unpaid invoices to identify ones that are overdue, generating reminder notices for those orders, and tracking which notices have been sent.
- Today: background process

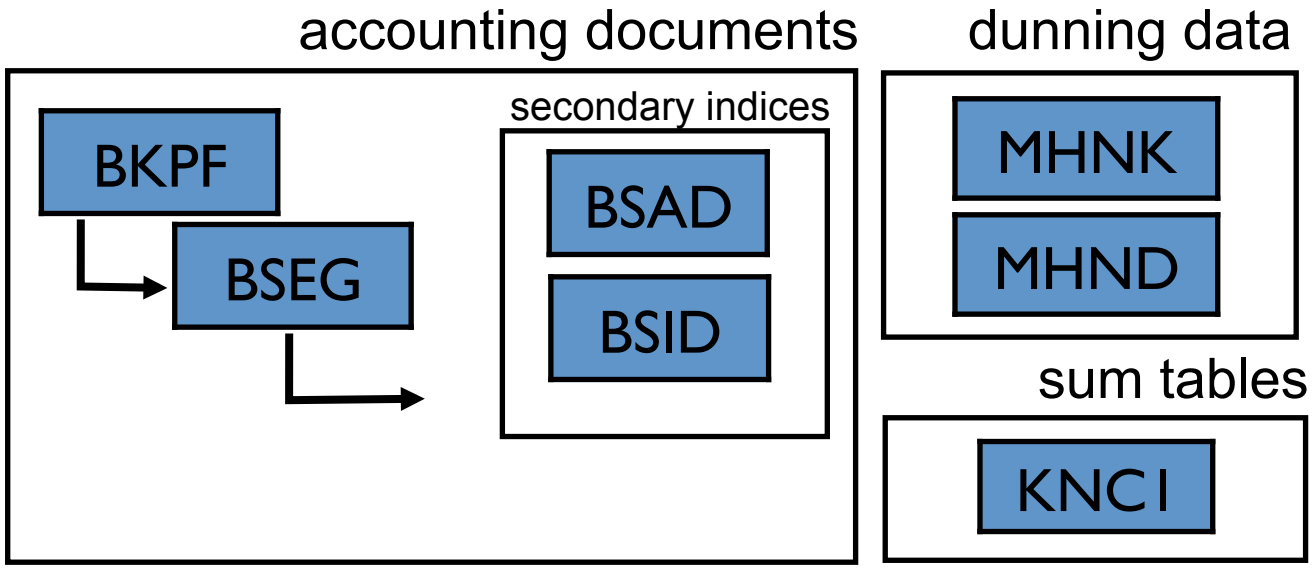
Dunning schema (I)

- Tables **grouped** into:
  - Transactional tables (BSEG, BKPF)
  - Sum tables and secondary indices (BSID, BSAD, KNC1
  - Pre-calculated and materialized tables

➔We use **transactional tables** only

➔Change documents removed

Dunning schema (II)



Dunning (II)

- Select accounts to be dunned, **for each**:
  - Select open account items from BSID, **for each**:
    - Calculate due date
  - Select dunning procedure, level and area
  - Create MHNK entries
- Create and write dunning item tables

Dunning (II)

1 SELECT

- Select accounts to be dunned, **for each**:

10000 SELECT's

  - Select open account items from BSID, **for each**:
    - Calculate due date

10000 SELECT's

  - Select dunning procedure, level and area
  - Create MHNK entries

31000 Entries

- Create and write dunning item tables

Dunning (II)

- Select accounts to be dunned, **for each**:
  - Select op
    - Calcula
  - Select du
  - Create MHNK entries
- Create and

One single stored procedure executed within TREX

Calculated on-the-fly

Dunning (III)

- TREX capabilities **enable** new dunning possibilities
- **Customer Segmentation** (Identification of customers which are very reliable and should not receive a standard dunning letter)
  - Calculates the average of **all BSEG items**
  - Searches all customers which never had an item with a dunning level higher one

Acceleration: 800x

(Quantity: 250 mio items, 380k open, 200k due)

#	Operation	Original	Version 1	Variant 2	Variant 3
1	Select open items		0.63s	1.01s (incl. T047 & ⊞	0.6s (incl. T047 & ⊞
2	Due date, dunning level		27s	Deferred to aggregation	0.5s
3	Filter 1 (verify dunning levels)		~19s	1.1s	0.5s
4	Filter 2 (check last dunning)		~15s	0.8s	0.4s
5	Generate MHNK (aggregate)		done in #1	1.2s	Done in #1
6	Generate MHND (execute filters)		done in #1	140ms	Done in #1
	Total	**~20 minutes**	**~1 minute**	**~3.0s** **(#3, #4 exec. in parallel)**	**~1.5s** **(#2, #3, #4 exec. in parallel)**

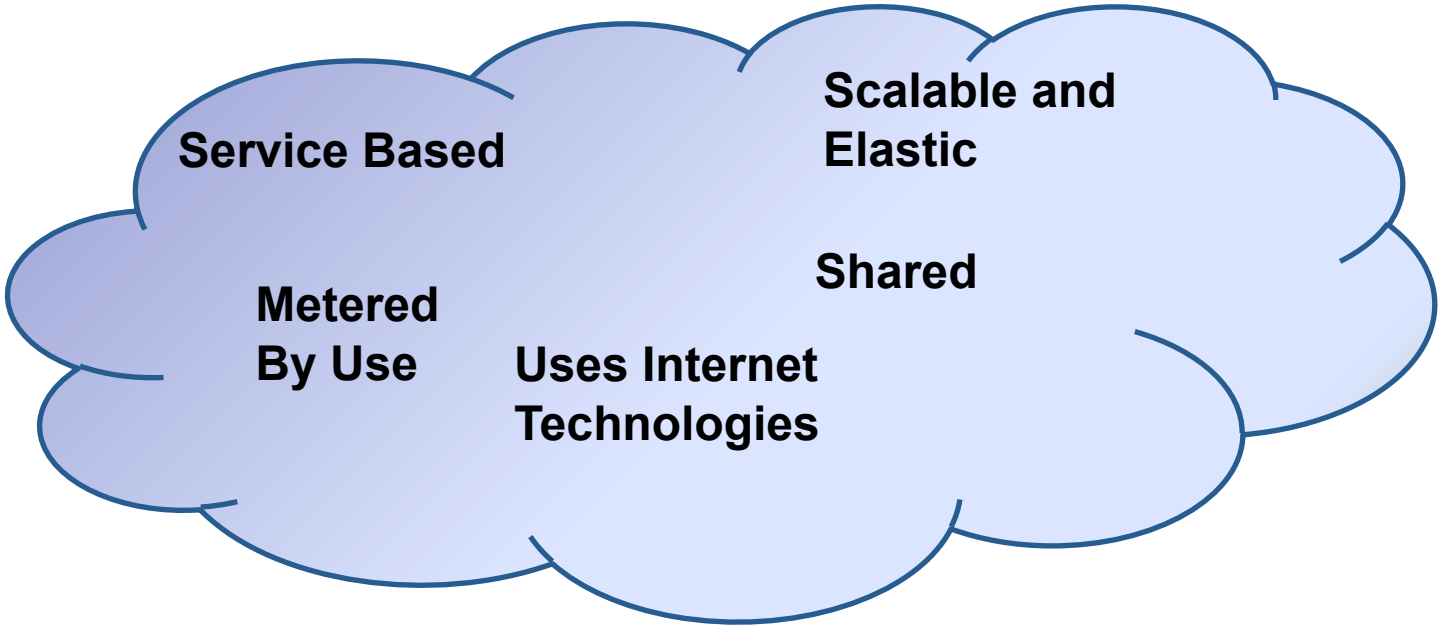
Hardware: 4 CPUs x 6 cores, 256 GB RAM

Cloud computing

Cloud Computing

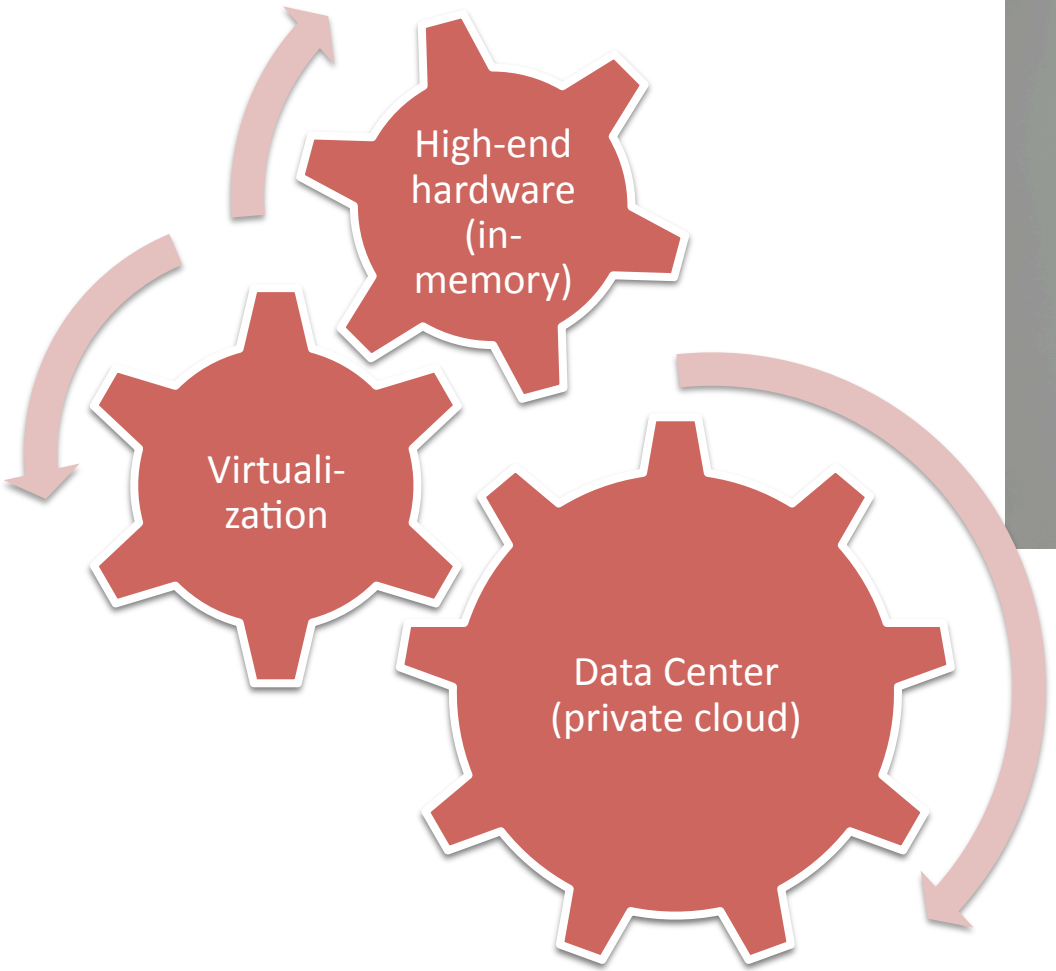
As Defined by Business Analysts:

Gartner‘s Five Attributes of Cloud Computing

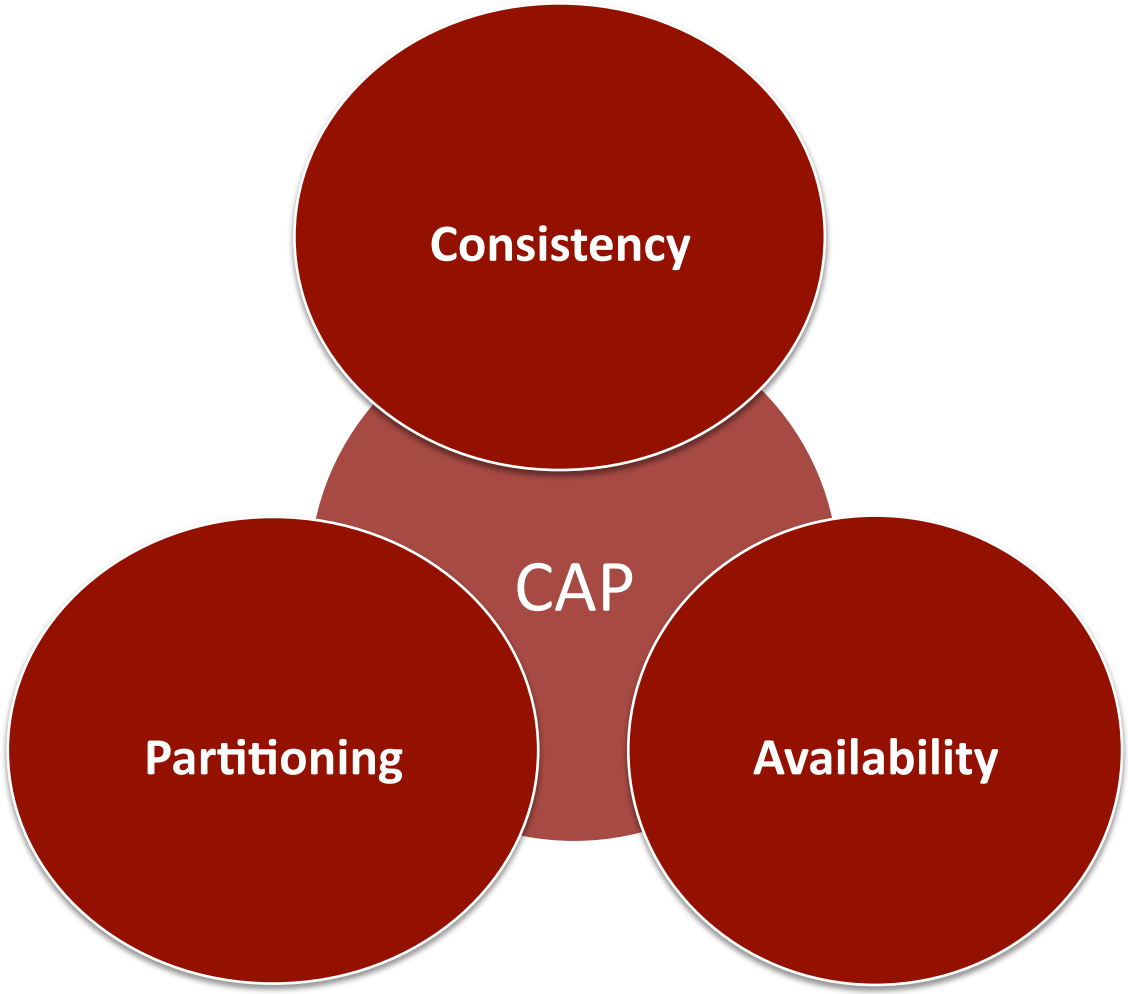


But isn‘t this only another marketing buzzword for existing technologies?

Cloud Computing Technologies



In distributed systems …



… cannot be achieved at the same time!

ACID vs. eventual consistency

ACID

- **Atomicity**
- **Consistency**
- Isolation
- Durability

Eventual Consistency

- **Achieve Partioning and Availability**
- **Weaken consistency**

Commodity vs. high end server clusters

Commodity Clusters	High End Clusters
• Fewer bytes scanned per second per server • Cheap, but fail easily	• More bytes scanned per second per server • Expensive but more reliable • Less communication costs

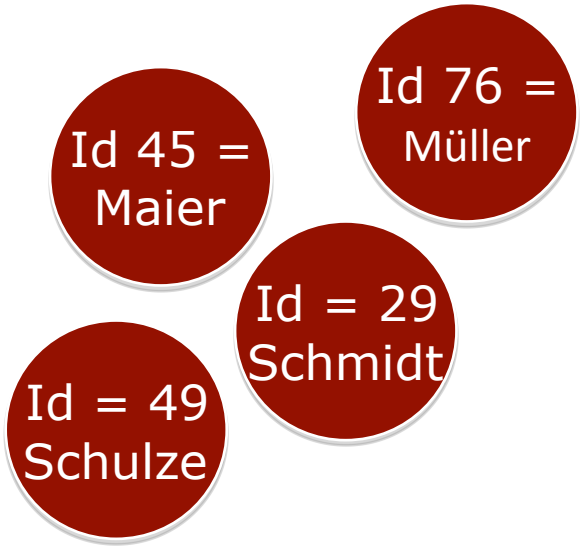
Google™

SAP

HPI Hasso Plattner Institut
IT Systems Engineering | Universität Potsdam

Key value stores vs. DBMS

No SQL Approach
Distributed Key Value Stores



SQL / Relational Approach

Row Based

Id	Cust	Adr.
29	Schmidt	…
45	Maier	…
49	Schulze	…
76	Müller	…

Column Based

Id	Cust	Adr.
29	Schmidt	…
45	Maier	…
49	Schulze	…
76	Müller	…