Datenbanksysteme II: Finish

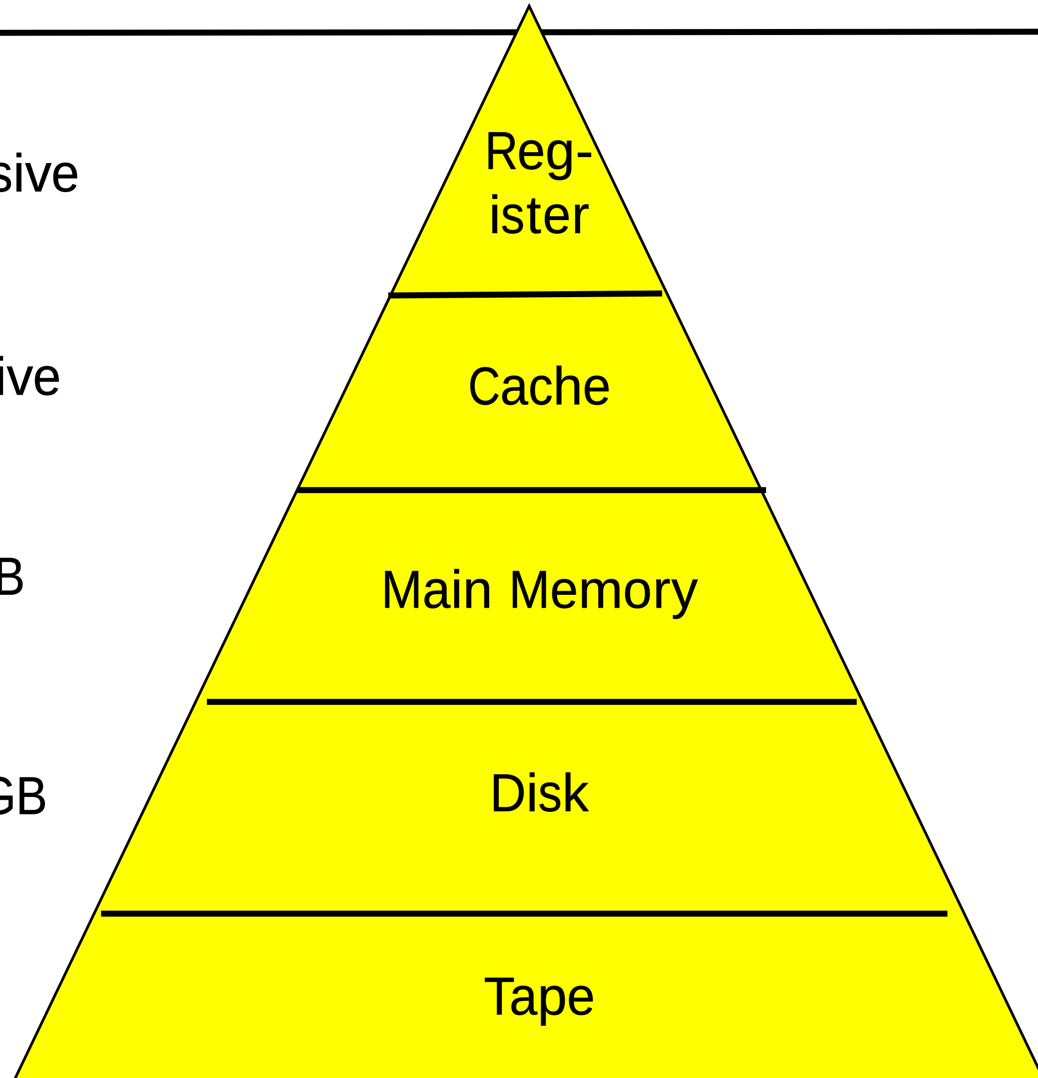
Ulf Leser

- Summary
  - What we learned
  - What we didn‘t learn
- Feedback
- Exams
- Advertisement

Contents

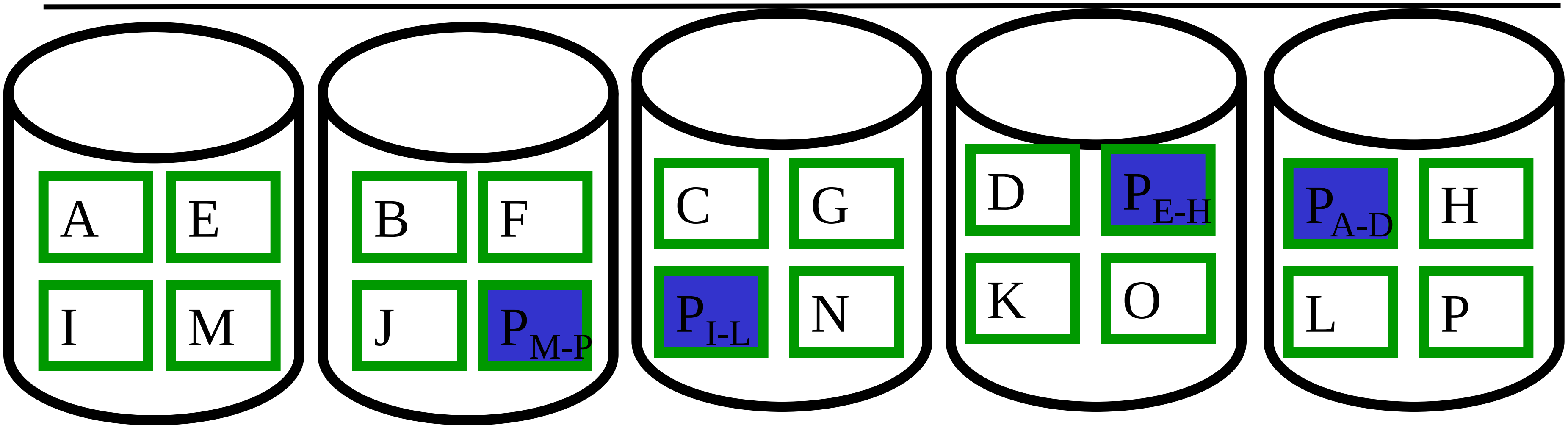
- Introduction
- Overview and architecture
- Storage and access methods
  - B*-Trees, Extensible hashing, index-sequential files ...
  - Multidimensional indexing: Grid-files, kd-Trees, R-Trees ...
- Query processing and optimization
  - Physical relational operators
  - Cost-based optimization
- Recovery
- Transactions and concurrency control

Price versus Access Time



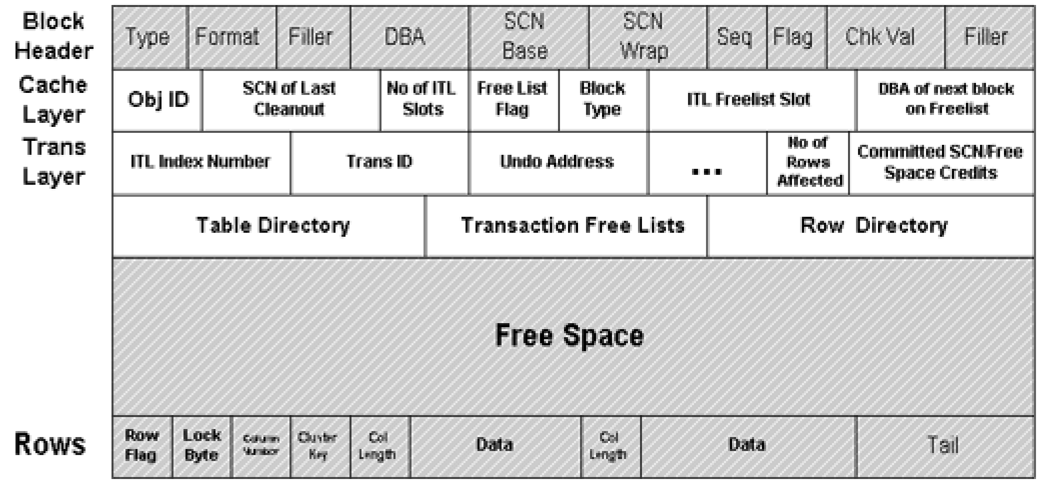
Really expensive	1-10ns / byte
Very expensive	10-60ns / cache line
~ 15 € / GB	100-300ns / block
~ 0,04 € / GB	10-20 ms / block
	sec – min

RAID 5: RAID4, Verteilung der Paritity



- Parity immer auf der Platte, die keinen der Blöcke enthält
- Wesentlich bessere Lastbalancierung als RAID4
  - Paritätsplatte kein Flaschenhals mehr
  - Schreiben erfordert X-1 Platten für die Daten und 1 weitere Platte für Parity
- Langsamer als RAID10, aber platzsparender

Oracle Block Structure



- DBA: Data Block Header: block address (global and relative in tablespace)
- Block type: data, index, redo, …
- Table directory: tables in this block (for clustered data)
- Row directory: offset of tuples in block
- ITL: Interested transaction list – locks on rows in block
  - There is no „lock manager“ in Oracle
  - ITL grows and shrinks – “ITL wait”, INITTRANS, MAXTRANS
  - Locks are not cleaned upon TX end – next TX checks TX-ID

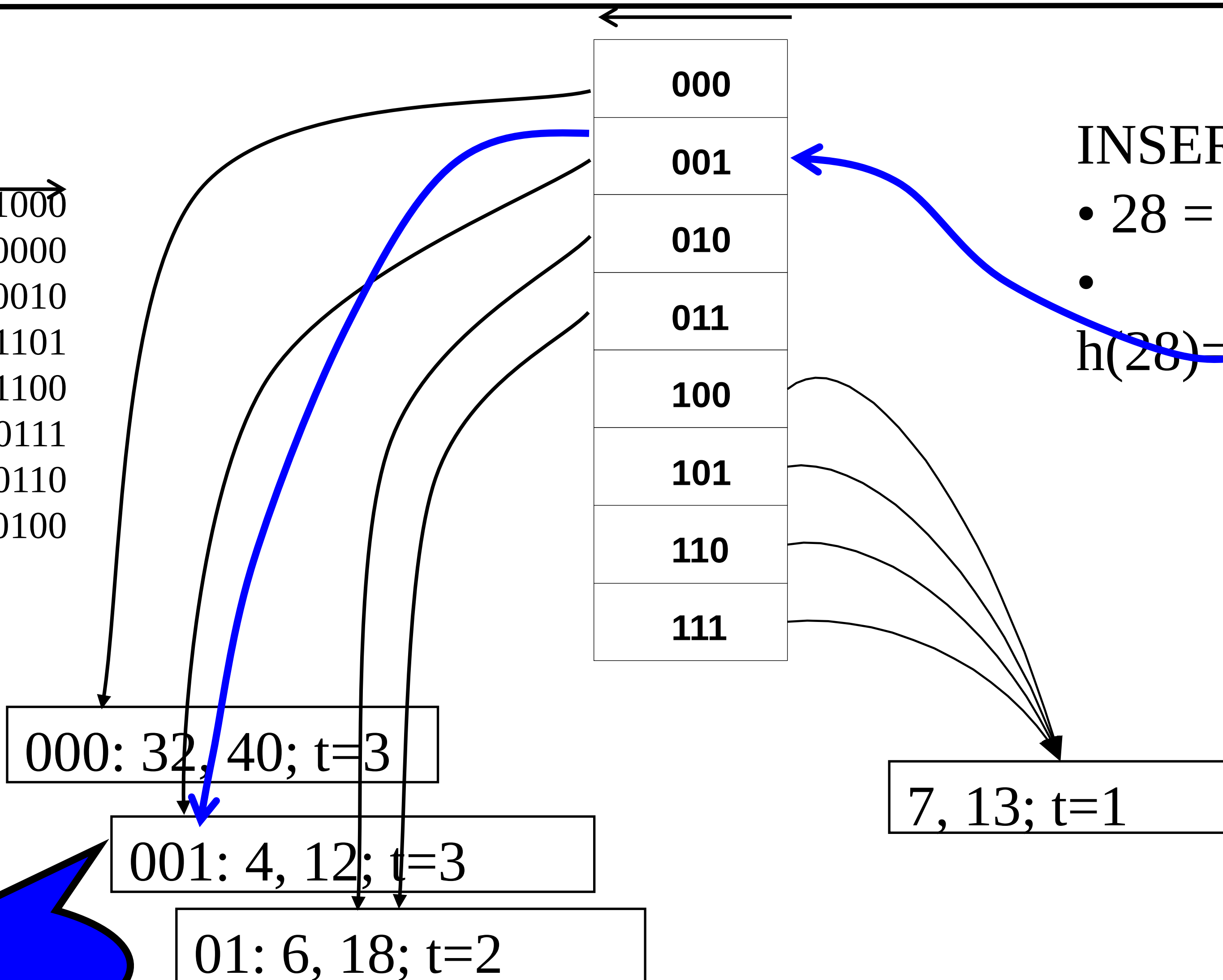
Inserting Values

Current content

40 = 101000
32 = 100000
18 = 010010
13 = 001101
12 = 001100
7 = 000111
6 = 000110
4 = 000100

INSERT(28)

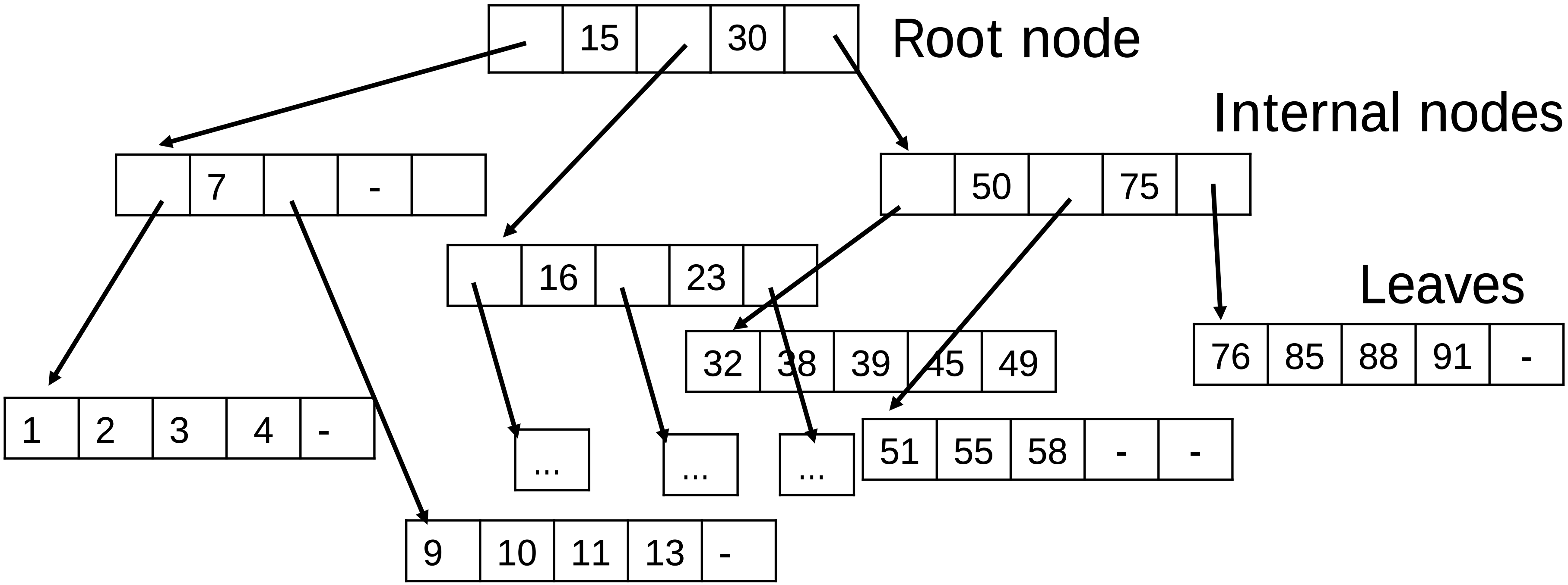
- 28 = 011100
- h(28)=001110



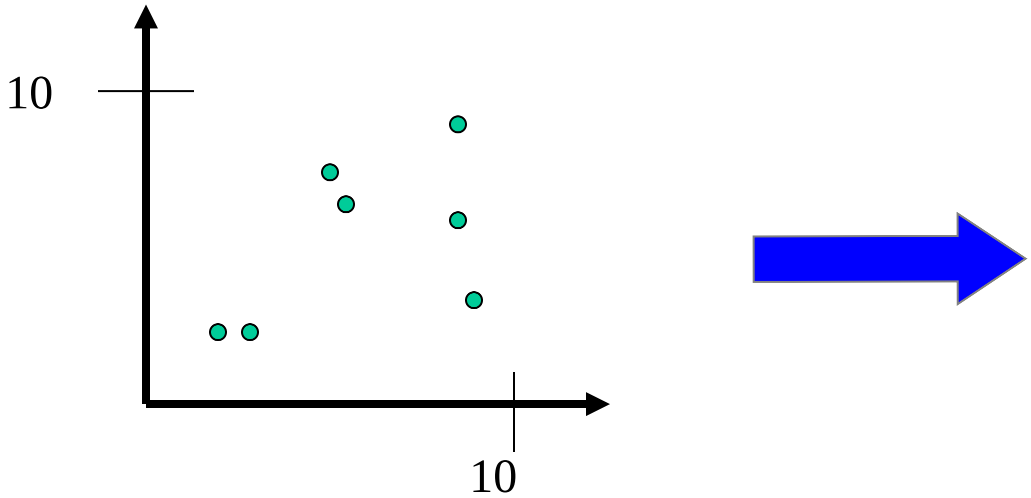
d=t; Overflow

B-Trees

- B-Tree is a multi-level index with variable number of levels
- Height adapts to table growth / shrinkage



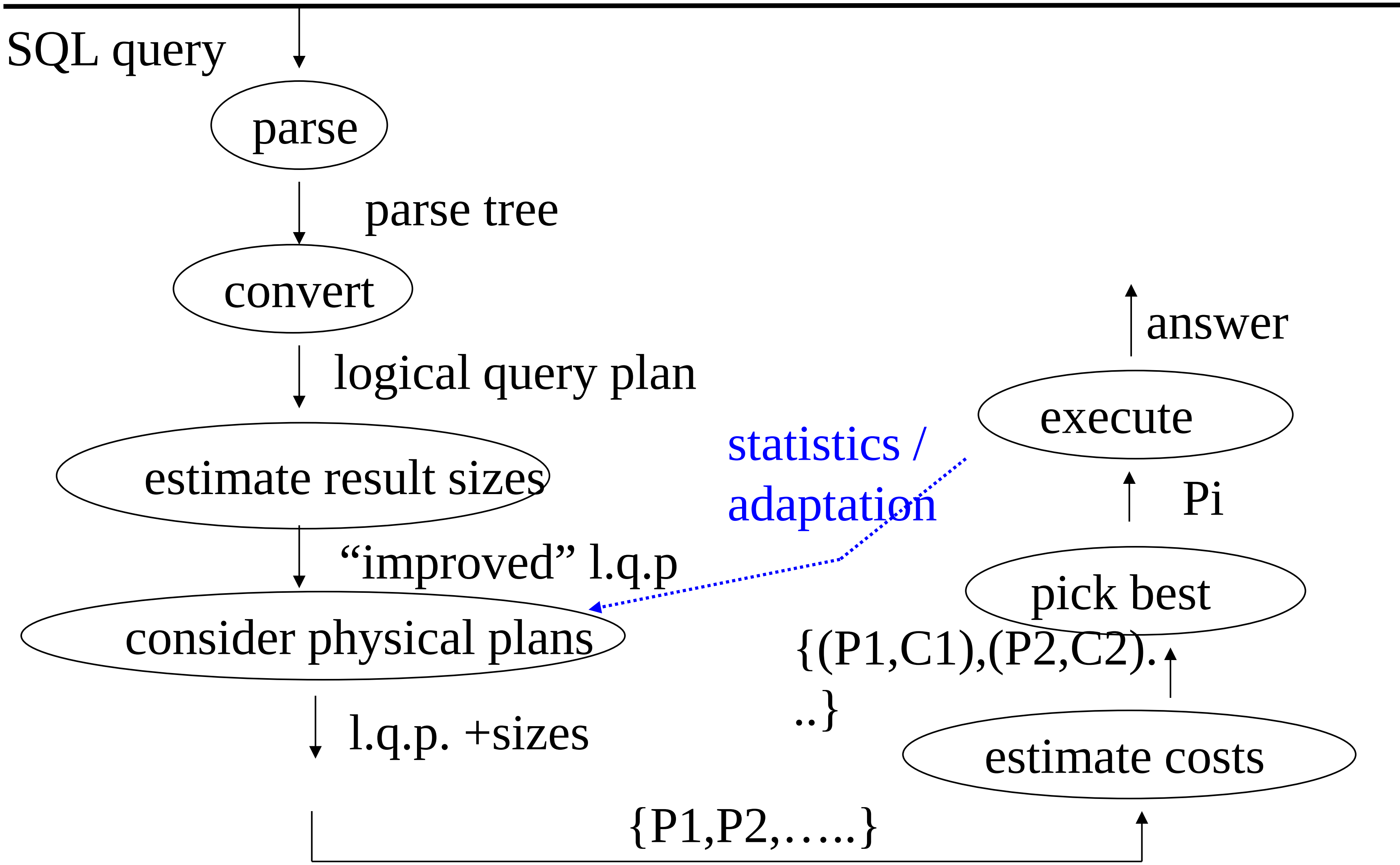
Example: 2D objects



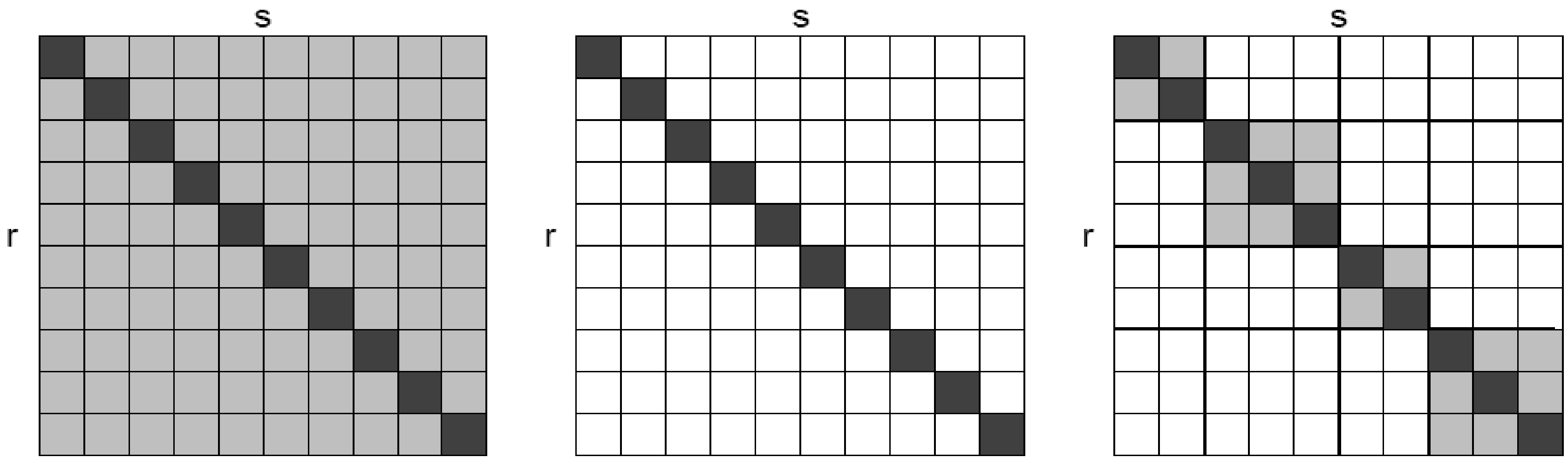
Point	X	Y
P1	2	2
P2	2,5	2
P3	4,5	7
P4	4,7	6,5
P5	8	6
P6	8	9
P7	8,3	3

- Objects are points in a 2D space
- Queries
  - Exact: All objects with coordinates (X1, Y1)
  - Box: Find all points in a given rectangle
  - Partial: All points/rectangles with X (Y) coordinate between ...

Exemplary Workflow



Comparing Join Methods



Nested-Loops-Join

Merge-Join

Hash-Join

Correlated Subquery without Aggregation

```
SELECT o.o_id
FROM   order o
WHERE  o.o_id IN (SELECT d.o_id
                  FROM   delivery d
                  WHERE  d.o_id = o.o_id AND
                         d.date-o.date<5)
```

- Subquery materialization not possible
- Naïve computation requires one execution of subquery for each tuple of outer query
- Solution: Rewrite into join
  - Again: Caution with duplicates (if o:d is 1:n, DISTINCT required)

```
SELECT DISTINCT o.o_id
FROM   order o, delivery d
WHERE  o.o_id = d.o_id AND
       d.date-o.date<5
```

Joins and Projection/Selection

- Rule 6. Exchange of selection and join

  If *Cond* contains only attributes of E_1, then:

  $$\sigma_{Cond}(E_1 \bowtie_{Cond1} E_2) \equiv \sigma_{Cond}(E_1) \bowtie_{Cond1} E_2$$

- Rule 7. Exchange of selection and union/difference

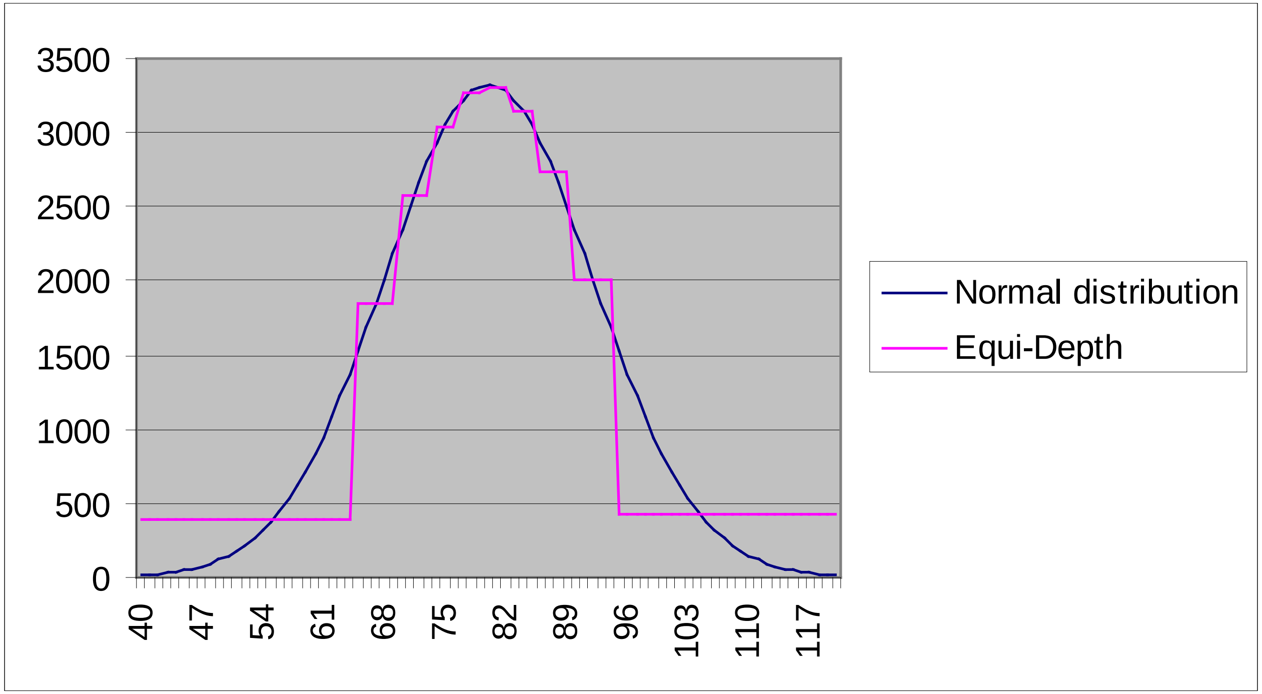
  $$\sigma_{Cond}(E_1 \cup E_2) \equiv \sigma_{Cond}(E_1) \cup \sigma_{Cond}(E_2)$$

  $$\sigma_{Cond}(E_1 - E_2) \equiv \sigma_{Cond}(E_1) - \sigma_{Cond}(E_2)$$

- Rule 8. Exchange of selection and natural join

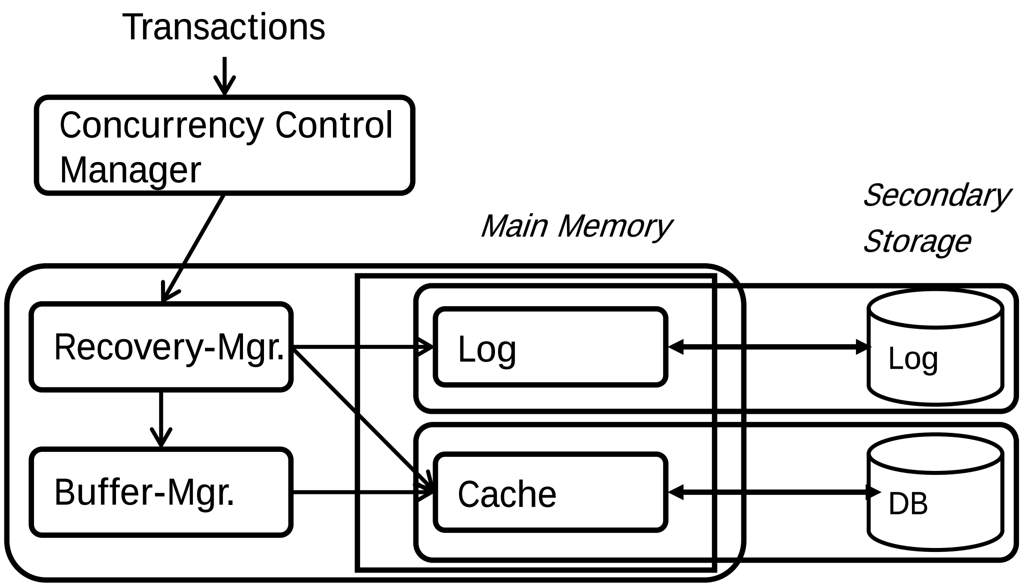
  $$\sigma_{Cond}(E_1 \bowtie E_2) \equiv \sigma_{Cond}(E_1) \bowtie \sigma_{Cond}(E_2)$$

Equi-Depth



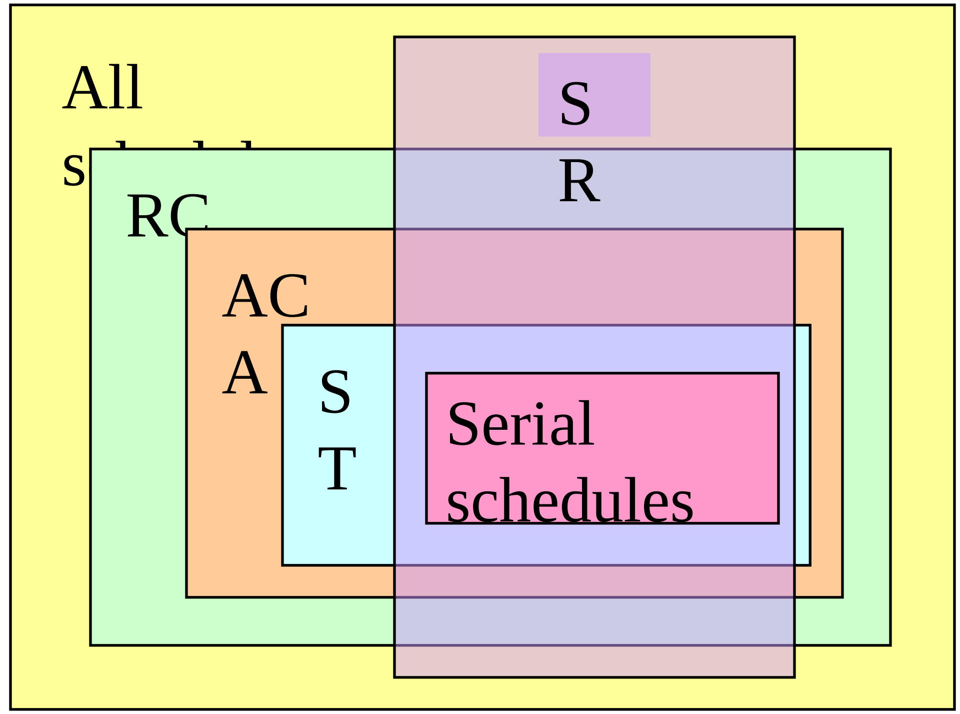
- Chose borders such that total frequency in each bucket is app. equal
  - Problem if one value is more frequent than |R|/|buckets| - use / combine with other types of histograms (later)

Architecture of a Recovery Manager



- In the following, we talk of "objects"
  - Usually means tuple
  - Could also be block or attribute value (more later)

Relationships



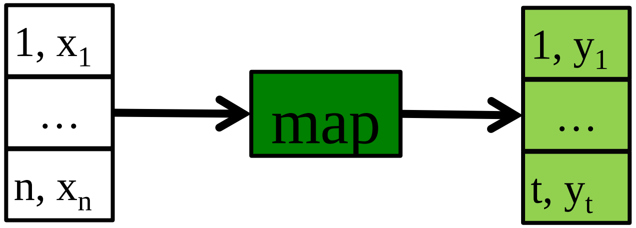
- RC: Recoverable schedules
- ACA: Schedules avoiding cascading aborts
- ST: Strict schedules
  - Usually, we want strict schedules in databases
- SR: Serializable schedules

Programming Model

- Inspired by functional programming concepts map and reduce
- Operates on key-value pairs

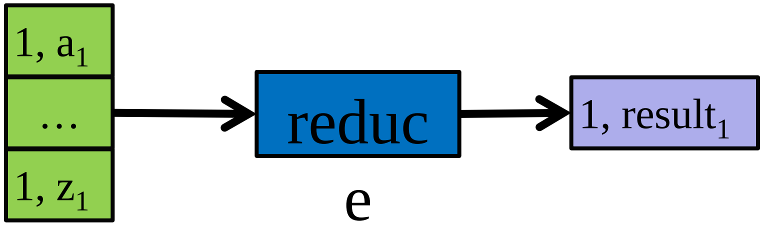
- Map
  - Process key-value pairs individually
    → with UDF
  - Generates key/value pairs
  - Example (LISP):

    (mapcar ’1+ ’(1 2 3 4)) ⇒ (2 3 4 5)



- Reduce
  - Merges intermediate key-value pairs with same key
  - Example (LISP):

    (reduce ’+ ’(1 2 3 4)) ⇒ 10



Further Topics

- Distributed Databases, 2-phase-commit
- Parallel databases
- Special-purpose extensions (multimedia, spatial, text)
- Techniques for very large databases (DWH)
- Techniques for integrating heterogeneous data (Info Int)
- More complex data analysis (OLAP, data mining, analytics)
- Database programming and tuning
- XML, Xpath
- Object-oriented databases, OQL
- ...

Bücher

- Vor allem
  - Saake, Heuer, Sattler: "Datenbanken: Implementierungstechniken", MTP Verlag, 2. Auflage, 2005
  - Garcia-Molina, Ullman, Widom: “Database System Implementation”, Prentice Hall, 2000
- Eher weniger
  - Härder, Rahm: “Datenbanksysteme. Konzepte und Techniken der Implementierung”, Springer, 2. Auflage 2001

Werbung

- Studien- und Diplomarbeiten
- Unsere Forschungsthemen
  - Text Mining (in biomedizinischen Texten)
  - Verteilte Analyse sehr großer Datenmengen (BigData)
  - Statistische Analyse biomedizinischer Daten
  - Ihr Lieblingsthema (wenn es halbwegs zum Profil passt)
- Oft interdisziplinär
  - Charite, MPI's, FUB, Linguisten, Firmen, …
- Immer: Intensive Betreuung

Diplom- / Magister/ Masterarbeiten

Anne Isberner: Similarity Search on Tabular Data (Expose)
Diplomarbeit Informatik
März 2016 - Oktober 2016
Betreuung: Ulf Leser, Felix Naumann (HPI Potsdam)

David Salomon: Bestimmung der Reccurence Matrix für die Reccurence Quantification Analyse mittels Approximate Nearest Neighbour Search (Expose)
Diplomarbeit Informatik
März 2016 - Oktober 2016
Betreuung: Ulf Leser, Norbert Marwan (GFZ Potsdam), Tobias Rawald

Tim Repke: Extraction of Citation Data from Websites based on Visual Cues (Expose)
Masterarbeit Informatik
März 2016 - September 2016
Betreuung: Ulf Leser, Niels Pinkwart

Maurice Bleuel: Implementation and Evaluation of the TPC-DI Benchmark for Data Integration Systems (Expose)
Masterarbeit Informatik
Januar 2016 - Juli 2016
Betreuung: Ulf Leser

Sascha Baese: Skalierbare Indexierung humaner Mutationsprofile durch Inverted Files (Expose)
Diplomarbeit Informatik
November 2015 - August 2017
Betreuung: Ulf Leser, Stefan Sprenger, NN

Peter Moor: Development of a Mutation Panel for Neuroendocrine Tumor Research (Expose)
Masterarbeit Bioinformatik (FU)
Dezember 2014 - Februar 2016
Betreuung: Ulf Leser, Heike Siebert (FUB)

Studien-/ Bachelorarbeiten

Monika Leung: Parallelization of a Bioinformatics application with a Workflow Language: A critical analysis of a parallel grid search optimization of the LIMMA algorithm based on the Cuneiform Workflow Language (Expose)
Bachelorarbeit Informatik
May 2016 - September 2016
Betreuung: Ulf Leser, Raik Otto, Joachim Fischer

Jan-Niklas Rössler: Eine kritische, komparative Analyse von Methoden zur Untersuchung differentieller Genexpression (Expose)
Bachelorarbeit Bioinformatik
May 2016 - August 2016
Betreuung: Ulf Leser, Raik Otto, Rosario Michael Piro (FUB)

Ausblick

- Sommer 2017
  - BA: Algorithmen und Datenstrukturen
  - BA: Einführung in die Bioinformatik (Starlinger)
  - BA: 2*Proseminar: Wissenschaftliches Arbeitens
  - MA: Seminar Text Classification