

Database Systems

WS 08/09

Prof. Dr. Jens Dittrich

Chair of Information Systems Group
<http://infosys.cs.uni-saarland.de>

Wrap-up.

Topics (1/6)

- fundamental system components
- storage media
 - disk
 - flash
 - main memory
- storage management
 - principles
 - page/block mapping and
- data layout - mapping data
 - vertical
 - horizontal
 - column grouping
 - hybrid mappings, PAX, fr
 - compression
 - free memory management

Week 1:

- requirements for a DBMS
- OLTP/OLAP/search/ etc.
- classic architecture
- system design principles
- storage media: tape, hard disks, flash, DRAM, SAN
- random I/O vs. sequential
- DB file system
- raw devices
- pages and blocks
- block assignment
- handling writes

- replacement strategies
- differential files
- DB-buffer

Topics (1/6)

- fundamental system components
- storage media
 - disk
 - flash
 - main memory
- storage management
 - principles
 - page/block mapping and
- data layout - mapping data
 - vertical
 - horizontal
 - column grouping
 - hybrid mappings, PAX, fr
 - compression
 - free memory management

Week 2:

- data item management
- tuple-IDs
- indirect addressing
- NSM
- DSM (column store)
- Fractured Mirrors
- Column Grouping
- PAX
- storage level compression
- impact on a MM DB
- impact on cache
- free memory management
- append only/best fit/hybrid

- storage defragmentation

Topics (2/6)

■ indexing

- one-dimensional
- tree-structured
- partition-based indexing
- bulk-loading
- main-memory indexing
- hash-indexes
- multi-dimensional indexes
- differential indexing
- read-optimized indexing
- write-optimized indexing
- data warehouse indexing
- text indexing: inverted files
- (flash-indexing)

Week 3 (part I):

- index = materialized mapping
- index requirements
- access paths
- primary vs. secondary
- secondary access paths and inversion
- Google example
- one-dimensional indexes
- chain/list/sequence
- tree structures
- B+-trees
- bottom-up index
- ISAM

Topics (2/6)

■ indexing

- one-dimensional
- tree-structured
- partition-based indexing
- bulk-loading
- main-memory indexing
- hash-indexes
- multi-dimensional indexes
- differential indexing
- read-optimized indexing
- write-optimized indexing
- data warehouse indexing
- text indexing: inverted files
- (flash-indexing)

Week 3 (part II)&3b:

- CSS trees
-

- split()
- indirect vs. direct
- clustered vs. non-clustered
- dense vs. sparse
- recursive partitioning
- bulk-loading
- prefix B+-trees
- prefix/suffix compression
- large pages
- cache conscious
- CSB+-tree
- prefetching trees
- fractal trees
- drawbacks of arrays

Topics (2/6)

■ indexing

- one-dimensional
- tree-structured
- partition-based indexing
- bulk-loading
- main-memory indexing
- **hash-indexes**
- multi-dimensional indexes
- differential indexing
- read-optimized indexing
- write-optimized indexing
- data warehouse indexing
- text indexing: inverted files
- (flash-indexing)

Week 4:

- applicability
- array -> hash table
- static hashing
- hash functions
- hash collisions
- open addressing
- probing
- dynamic hashing
- extendible hashing
- ext. hashing vs. kd-tries
- linear hashing
- hashing -> bitmaps

Topics (2/6)

- indexing
 - one-dimensional
 - tree-structured
 - partition-based indexing
 - bulk-loading
 - main-memory indexing
 - hash-indexes
 - multi-dimensional indexes
 - differential indexing
 - read-optimized indexing
 - write-optimized indexing
 - data warehouse indexing
 - text indexing: inverted files
 - (flash-indexing)

Topics (5/6)

- large systems
 - global scale data management
 - map/reduce
 - pig and pig latin
 - search engines
 - data warehousing and OLAP
- write-optimized system concepts
 - OLTP
 - publish/subscribe
 - streaming
 - moving objects
- management of geographical data
 - basic concepts
 - GIS, google maps

Week 5a:

- intersection
- key concatenation
- data types
- query types
- k-nearest-neighbor query
- digital trees
- kd-tree and kd-trie
- quadtree and quadtrie
- kdB-tree
- linearized B⁺-trees
- z- and hilbert-curve
- z-codes and kd-trie
- range query partitioning
- grid-indexes
- grid-directory
- grids vs. hashing
- R-trees
- priority-driven k-NN query
- GIST/XXL

Topics (2/6)

■ indexing

- one-dimensional
- tree-structured
- partition-based indexing
- bulk-loading
- main-memory indexing
- hash-indexes
- **multi-dimensional indexes**
- differential indexing
- read-optimized indexing
- write-optimized indexing
- data warehouse indexing
- text indexing: inverted files
- (flash-indexing)

Week 6b:

- issues in multi-dimensional indexing
- example: image search
- properties of high-dimensional spaces
- node access probabilities
- any index will fail in high-dimensional spaces
- linear search better for high dimensions
- VA-file: modified signature file
- filter index

If you're dealing with high-dimensional data: be careful!

Topics (3/6)

- operator models
 - push-model
 - pull-model
- operator implementation
 - general idea
 - join algorithms for relational
 - other operators
- query processing
 - scanning & “naive plans”
 - canonical plan computation

Week 7 (part I):

- logical and physical operators
- push vs. pull
- iterators
- iterator implementation
- granularity of iteration
- external sorting
- simple two-way merge sort
- m-ary merge sort
- blocked I/O
- double buffering
- replacement selection
- online merge

Topics (3/6)

- operator models
 - push-model
 - pull-model
- operator implementation
 - general idea
 - join algorithms for relational
 - other operators
- query processing
 - scanning & “naive plans”
 - canonical plan computation

Week 7 (part II):

z-code join

- join algorithms
- nested-loops
- index nested-loops
- sort-merge
- duplicates
- sweep areas
- simple hash-join
- grace hash-join
- interval partitioning
- hash partitioning
- double-pipelined hash join
- aggregation and grouping
- early aggregation
- non-relational joins
- plane sweep

Topics (3/6)

- operator models
 - push-model
 - pull-model
- operator implementations
 - general idea
 - join algorithms for relational and multidimensional data
 - other operators
- query processing
 - scanning & “naive plans“
 - canonical plan computation

WS 08/09

Prof. Dr. Jens Dittrich / Information Systems Group / infosys.cs.uni-saarland.de

Topics (4/6)

- query optimization
 - query rewrite
 - cost-based
- data recovery
 - single versus multiple instance
 - ARIES
- parallelization of data and queries
 - horizontal partitioning
 - vertical partitioning
 - replication
 - map-reduce
 - multi-cores

WS 08/09

Prof. Dr. Jens Dittrich / Information Systems Group / infosys.cs.uni-saarland.de

Week 9:

- query rewrite
- canonical form
- rewrite rules
- equivalences
- query unnesting
- cost-based optimization
- cost models
- selectivities
- join order
- histograms
- I/O-cost model
- plan enumeration
- explain plan

Topics (4/6)

- query optimization
 - query rewrite
 - cost-based
- data recovery
 - quick recap of transaction management
 - single instance recovery
- transaction handling
 - scheduling of transactions
 - concurrency control
 - implementing isolation levels
- parallelization of data and queries
 - horizontal partitioning, vertical partitioning
 - distributed query processing
 - multi-cores
 - map-reduce

Week 10:

- recap of transaction management
- ACID
- nested TAs
- crash recovery
- possible errors
- write strategies
- logging
- stable storage
- WAL
- log records
- compensation log record

- logical/physical/physiological
- ARIES
- transaction table
- dirty page table
- fuzzy checkpoints
- three phases
- analysis
- redo
- undo
- fault tolerance

Topics (4/6)

- query optimization
 - query rewrite
 - cost-based
- data recovery
 - quick recap of transaction recovery
 - single instance recovery
- **transaction handling**
 - scheduling of transactions
 - concurrency control
 - implementing isolation levels
- parallelization of data and queries
 - horizontal partitioning, vertical partitioning
 - distributed query processing
 - multi-cores
 - map-reduce

Week 10b:

- concurrency vs. batched-processing
- lost updates
- dirty data
- phantoms
- serializability
- cascading aborts
- database scheduler
- locking
- two-phase locking
- strict 2PL
- deadlocks
- waits-for-graph
- timestamping
- multi-granularity locking
- multi-threaded indexes
- optimistic synchronization
- implementing isolation levels

Topics (4/6)

- query optimization
 - query rewrite
 - cost-based
- data recovery
 - quick recap of transaction recovery
 - single instance recovery
- transaction handling
 - scheduling of transactions
 - concurrency control
 - implementing isolation levels
- parallelization of data and query processing
 - horizontal partitioning, vertical partitioning
 - distributed query processing
 - multi-cores
 - map-reduce

Week 11:

- parallelization
- multi-core
- shared-nothing
- GPU
- horizontal partitioning
- vertical partitioning
- replication
- mix
- split vs. merge paradigm
- parallel joins
- table shipping
- semijoin

- bloomjoin
- repartitioning joins=grace hash join
- distributed query optimization
- left-deep vs. right-deep plans
- distributed TAs
- two-phase commit
- coordinator
- processing nodes
- distributed deadlocks

Topics (5/6)

- large systems
 - global scale data management
 - map/reduce
 - pig and pig latin
 - search engines
 - data warehousing and OLAP
- write-optimized system classes
 - OLTP
 - publish/subscribe
 - streaming
 - moving objects
- management of geographic information
 - basic concepts
 - GIS, google maps

Week 12a:

- one size does not fit all
- OLTP != OLAP
- OLAP ≈ IR
- OLTP -> main memory
- single threading - (almost) no concurrency control
- grids
- high availability without ARIES

Topics (5/6)

- large systems

- global scale data management
 - map/reduce
 - pig and pig latin
 - search engines
 - data warehousing and OLAP

- write-optimized systems

- OLTP
- publish/subscribe
- streaming
- moving objects

- management of geographic information

- basic concepts
- GIS, google maps

Week 12b:

- Petabyte databases
- other reasons
- Yahoo! PNUTS
- relaxed consistency
- eventual cs.
- timeline cs.
- API
- system architecture
- tablets and storage units
- tablet controllers and routers

- message brokers
- regions
- range queries
- indirect addressing
- per-record mastership
- recovery
- applications

Topics (5/6)

- large systems
 - global scale data management
 - map/reduce
 - pig and pig latin
 - search engines
 - data warehousing and OLAP
- write-optimized systems
 - OLTP
 - publish/subscribe
 - streaming
 - moving objects
- management of geographic information
 - basic concepts
 - GIS, google maps

Week 13a:

- map() = splitting
- implicit grouping
- reduce() = aggregation
- use-cases
- inverted index construction
- implementation
- master and workers
- local files vs. GFS
- fault tolerance
- backup tasks
- early aggregation

- hadoop
- multi-core
- GPUs
- map-reduce merge

Topics (5/6)

- large systems
 - global scale data management
 - map/reduce
 - pig and pig latin
 - search engines
 - data warehousing and OLAP
- write-optimized systems
 - OLTP
 - publish/subscribe
 - streaming
 - moving objects
- management of geographic information
 - basic concepts
 - GIS, google maps

Week 13b:

optimizing

- pig = front-end to map/reduce
- pig latin: declarative flow language
- python-like data model
- sequence of steps
- CoGroup
- Join
- Group
- map/reduce in pig latin
- plan generation

Topics (5/6)

- large systems
 - global scale data management
 - map/reduce
 - pig and pig latin
 - search engines
 - data warehousing and OLAP
- write-optimized systems
 - OLTP
 - publish/subscribe
 - streaming
 - moving objects
- management of geographic information
 - basic concepts
 - GIS, google maps

Week 14a:

- query mediation vs. data warehousing
- trade-offs
- Extract, transform, load (ETL)
- star schema
- dimension tables
- snowflake schema
- galaxy schema
- naive plans
- cross product plan
- semijoin plan
- indexing constraints
- materialized views
- view lattice
- join index
- Types 1, 2, and 3

Topics (5/6)

■ large systems

- global scale data management
- map/reduce
- pig and pig latin
- search engines

■ data warehousing and OLAP

■ write-optimized systems

- OLTP
- publish/subscribe
- streaming
- moving objects

■ management of geographic data

- basic concepts
- GIS, google maps

Week 14b:

- join index implementation
- RIDS vs. bitmaps
- value bitmaps
- bitmap operations
- SIMD
- bitmap compression
- bitmap join index
- star join using bitmap join indexes
- join indexes in parallel databases

- limitations of replication
- join index replication
- further improvements
- projection index
- bit-sliced index
- queries on a bit-sliced index
- compression
- MOLAP

Planned Topics not covered

- dataspace systems
 - big vision
 - pay-as-you-go information integration
 - system examples (iMeMex)
- moving objects
- publish/subscribe
- streaming
- information integration
- differential indexing, LSM, stepped merge, etc.
- additional cache-aware stuff
- flash indexing
- advanced concurrency

Project

- latest hand-in: March 15, midnight
- oral exams: March 23 until March 27
- I will offer “consulting“ for your project:
 - office hours: wednesdays at 2pm, E1 1, Room 220.1
 - or arrange meeting by email
 - until March 15

Bachelor and Master Theses

- if you are interested in any of the topics handled in this lecture, contact me directly:
 - jens.dittrich @ cs.
 - office hours Wednesday from 2pm to 3pm

Tutoren gesucht

VL Informationssysteme

SoSe 2009

Kontakt: Prof. Jens Dittrich

infosys.cs.uni-saarland.de

Datenbanken

Informationssysteme

SQL

Oracle 11g

Suchmaschinen

Datenmodellierung

relationales Modell

RDF, OO, XML, Graphen

Datenbankentwurf

Anfragesprachen

Applikationen

Datenbanken und Web

Datenbankkonnektivität

Next: Final exam.

Thursday

Feb 12

12:15 pm

HS I & II

closed book.

Q&A