Seminar
Recent Trends in Database Research

Summer Term 2013

Lehrgebiet Informationssysteme

Weiping Qu
qu@cs.uni-kl.de

DBIS
Datenbanken und
Informationssysteme

**AG Datenbanken und
Informationssysteme**

HiS

**AG Heterogene
Informationssysteme**

Goals

a) Familiarize yourself with a scientific topic independently

b) Find scientific literature on web or in the library, cite correctly
   - ❑ ACM digital library
   - ❑ DBLP
   - ❑ CiteseerX
   - ❑ Google

c) Prepare a written composition, presentation, and discussion

d) Time Management

e) Don't copy and paste! (not from papers or web articles directly)

Written composition

Task 1: Searching for Literature (2 weeks later, May 2nd)

- ❑ Some basic literature is provided by your supervisor
- ❑ You have to search for further literature
- ❑ Collect a list of related literature and send it to your supervisor

Task 2: Annotated Table of Contents (TOC) (2 weeks later, May 15th)

Prepare an annotated TOC for your supervisor
(approx. 1.5 pages)

Task 3: Written Composition (again 7 weeks later, July 3rd)

- ❑ LNCS Layout
- ❑ PDF Format required
- ❑ Length: 6000–8000 Words (net.) ≅ 15–20 Pages
- ❑ Correct and complete bibliography

Task 4: Presentation (July 12th,15th,19th)

Links

Homepage

➢ http://wwwlgis.informatik.uni-kl.de/cms/courses/seminar/

LNCS

➢ http://www.springer.com/computer/lncs?SGWID=0-164-6-793341-0

ACM Digital Library

➢ http://dl.acm.org/

DBLP

➢ http://www.informatik.uni-trier.de/~ley/db/

CiteSeerX

➢ http://citeseerx.ist.psu.edu/index

Presentation

a) Length: 60 Minutes

❑ 45 minutes for the presentation

❑ 15 minutes for the discussion

b) Presentation:

❖ Projector (private or one of our notebooks)

c) You must submit your presentation electronically, one week after you gave your talk

d) First talks are scheduled for

❑ Friday, July 12th, 9:30-12:30 & 13:30-16:30

❑ Room 36/336

Topic I

Topic 1:   Caching for flash-based databases

Topic 2:   Flash-based caching for databases - energy efficiency and performance

Topic 3:   Use of phase-change memory (PCM) in modern DBMS architecture

Topic 4:   SystemML: declarative machine learning on MapReduce

Topic 5:   Online Reorganization in memory-based databases

Topic 6:   Data sharing across concurrent queries

Topic 7:   Query processing on raw files

Topic 8:   Hybrid storage databases

Topic 9:   Analysis algorithms in spatial database

Topic 10:   Optimization in MapReduce workflow

Topic 11:   Data security in cloud storage

Topic 12:   New hardware for database management system (Multi-cores, GPU, FPGA)

Topic 13:   Cost model for MapReduce

Topic 14:   Resource estimation for cost-based query optimization (statistical, machine learning)

Topic 15:   Data flow optimization for multiple execution engines

How you successfully pass the seminar

➢ **"Nongraded Certificate":**

❑ Meet the deadlines!

❑ Don't copy and paste!

❑ Decent presentation

❑ Be present when others give their talks

➢ **"Graded Certificate"**

• See above

Criteria for grading:
❑ Quality of your written composition
❑ Quality of your presentation (including your slides)
❑ Discussion
❑ Meeting the deadlines
❑ Overall impression of your supervisor
❑ ...

Schedule

- ➢ Wednesday, April 17th: Kick-off meeting
- ➢ Thursday, May 2nd: Deadline for literature list
- ➢ Wednesday, May 15th: Deadline for annotated TOC
- ➢ Wednesday, July 3rd: Deadline written composition
- ➢ Final presentations:
  - ❑ Friday, July 12th: 9:30-12:30 & 13:30-16:30
  - ❑ Monday, July 15th: 10:00-13:00 & 14:00-16:00
  - ❑ Friday, July 19th: 10:00-13:00 & 14:00-16:00

All deadlines are strict!

Questions

Select you topic now!

Topic 1: Caching for flash-based databases

Topic 2: Flash-based caching for databases - energy efficiency and performance

Topic 3: Use of phase-change memory (PCM) in modern DBMS architecture

Topic 4: SystemML: declarative machine learning on MapReduce

Topic 5: Online Reorganization in memory-based databases

Topic 6: Data sharing across concurrent queries

Topic 7: Query processing on raw files

Topic 8: Hybrid storage databases

Topic 9: Analysis algorithms in spatial database

Topic 10: Optimization in MapReduce workflow

Select you topic now!

Topic 11:   Data security in cloud storage

Topic 12:   New hardware for database management system (Multi-cores, GPU, FPGA)

Topic 13:   Cost model for MapReduce

Topic 14:   Resource estimation for cost-based query optimization (statistical, machine learning)

Topic 15:   Data flow optimization for multiple execution engines