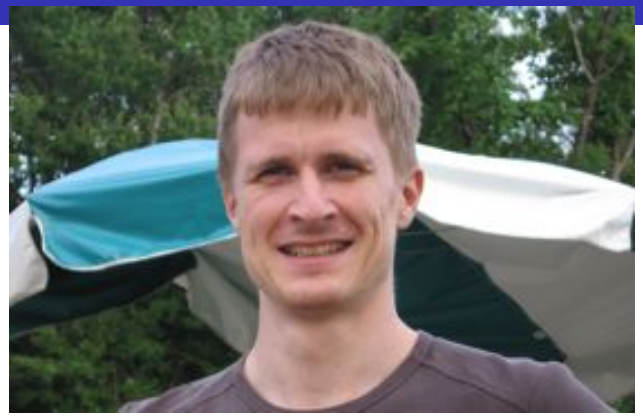
Data Processing on Modern Hardware

Jens Teubner, ETH Zurich, Systems Group
jens.teubner@inf.ethz.ch

Fall 2012

A Few Words About Me



Jens Teubner
CAB E 77.1
+41 44 632 4407
jens.teubner@inf.ethz.ch

1996–2001	Diploma in Physics, U Konstanz
2001–2005	Research assistant, DBIS Group, U Konstanz
2005–2007	Research assistant, Database Group, TU München
Oct 2006	PhD in Computer Science (XML query processing)
2007–2008	Postdoc, IBM T. J. Watson Research Center, NY, USA
since 8/2008	Senior Researcher, Systems Group, ETH Zurich

Topic: Database systems on modern computing hardware,
mainly using FPGAs and hardware-accelerated networks

A Motivating Example (Memory Access)

Task: sum up all entries in a two-dimensional array.

Alternative 1:

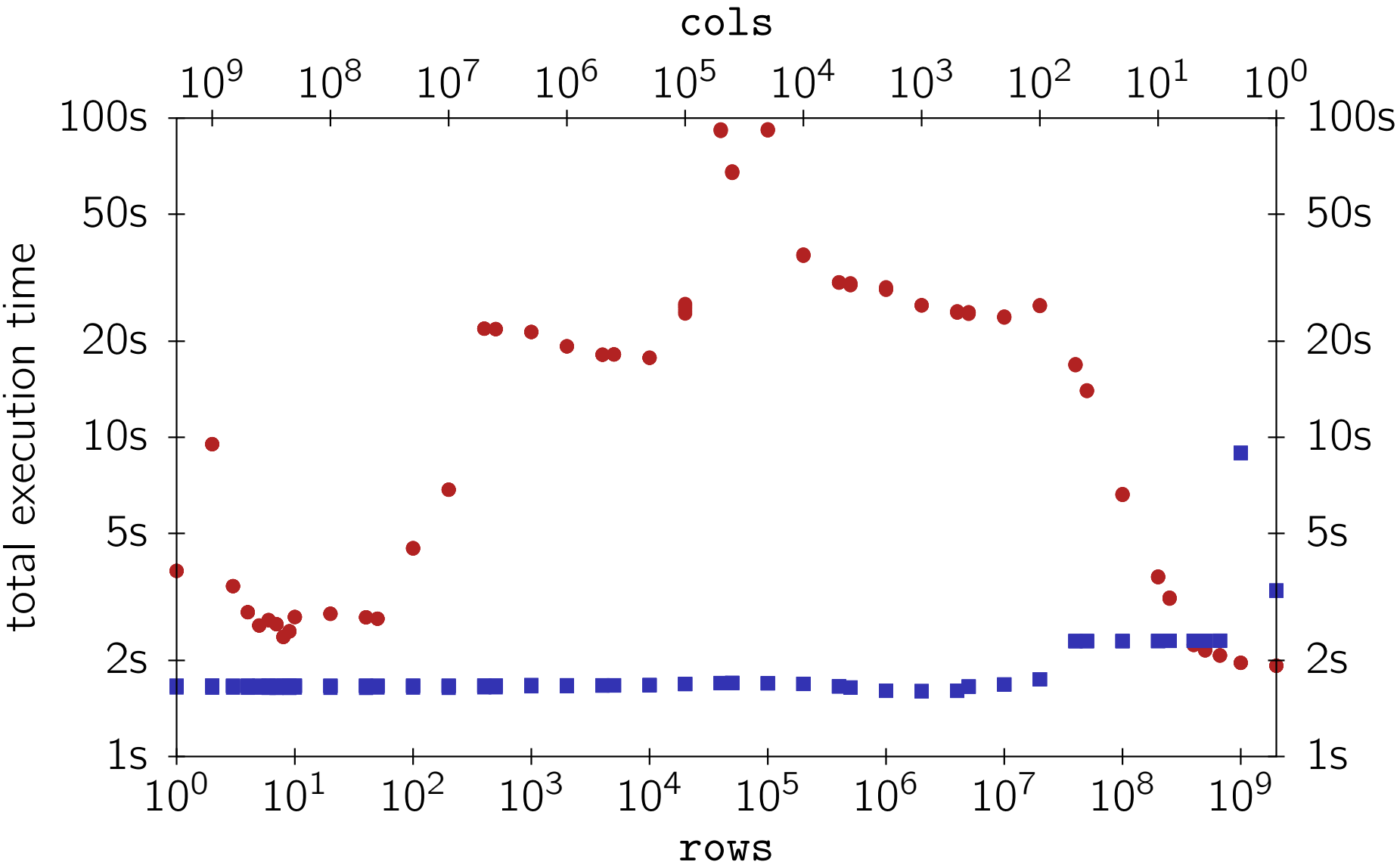
```
for (r = 0; r < rows; r++)
    for (c = 0; c < cols; c++)
        sum += src[r * cols + c];
```

Alternative 2:

```
for (c = 0; c < cols; c++)
    for (r = 0; r < rows; r++)
        sum += src[r * cols + c];
```

Both alternatives touch the same data, but in different order.

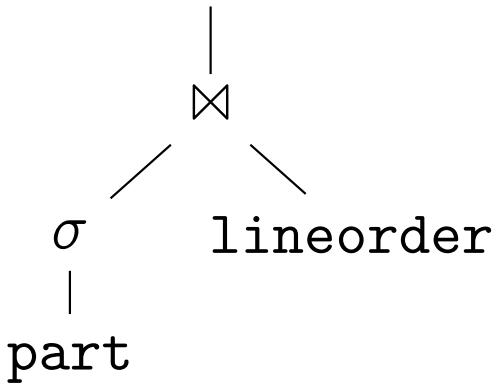
A Motivating Example (Memory Access)



A Motivating Example (Multi-Core)

Task: run parallel instances of the query

```
SELECT SUM(lo_revenue)
  FROM part, lineorder
 WHERE p_partkey = lo_partkey
   AND p_category <= 5
```



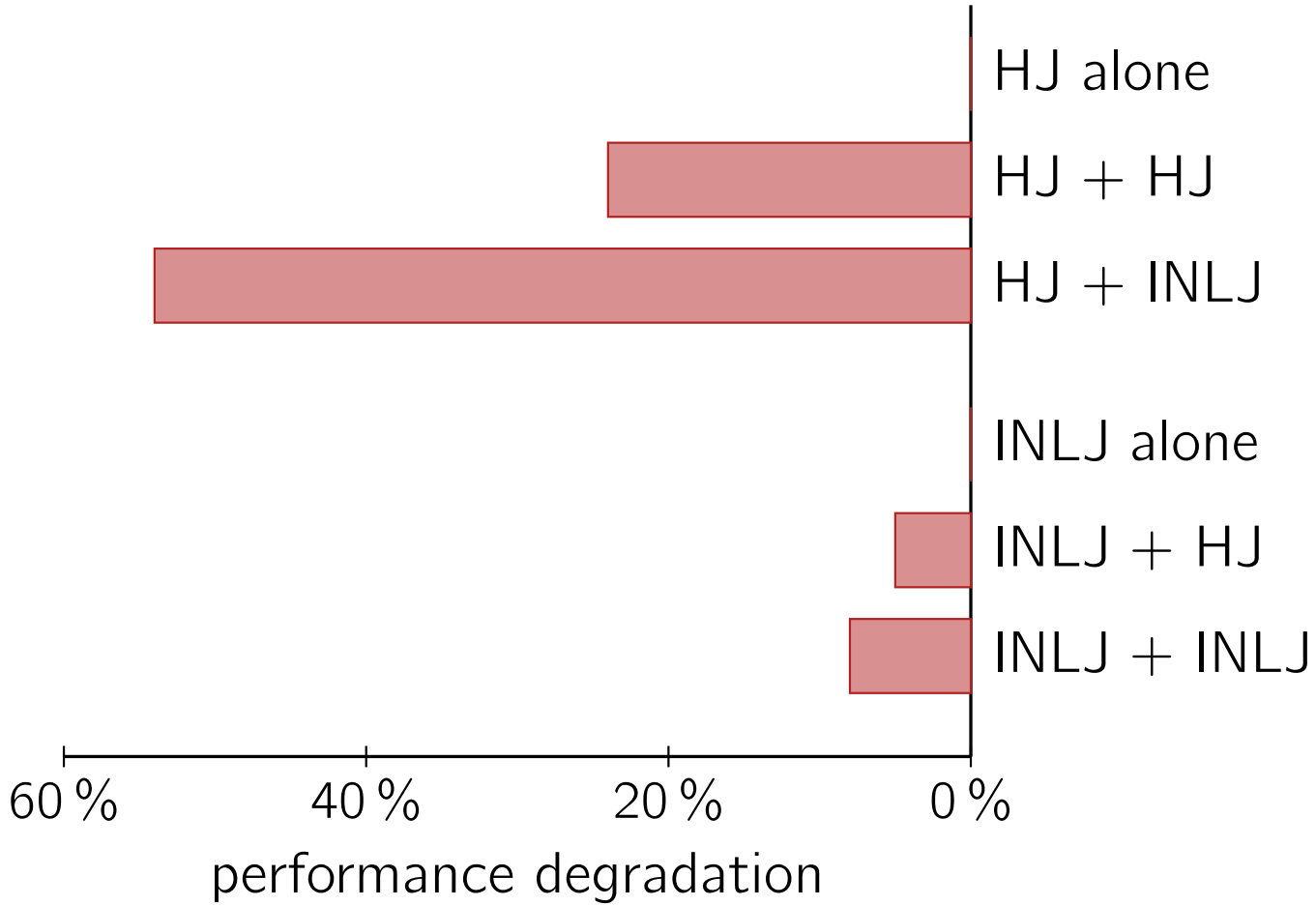
To implement ⋈ use either

- a **hash join** or
- an **index nested loops join**.

Results taken from 'Lee, Ding, Chen, Lu, and Zhang. MCC-DB: Minimizing Cache Conflicts in Multi-core Processors for Databases. *VLDB 2009*.'

A Motivating Example (Multi-Core)

Co-run independent instances on different CPU cores.



Concurrent queries may seriously affect each other's performance.

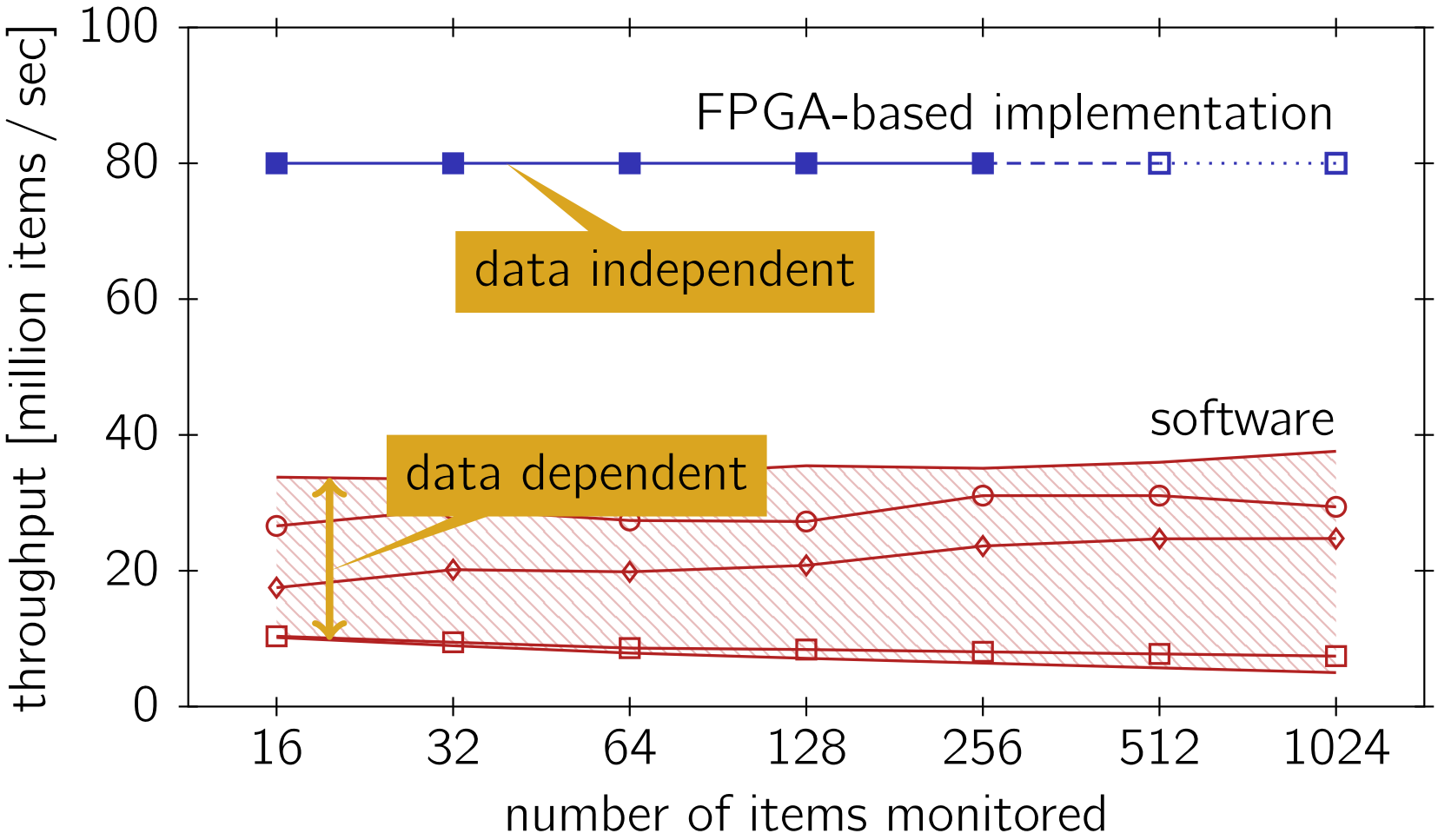
A Motivating Example (Non-Commodity Hardware)

Task: in a long stream of items, find those items that occur most often.

Algorithm *Space-Saving* (Metwally *et al.*, TODS, vol. 31(3), 2006):

```
1 foreach stream item x ∈ S do
2     find bin b_x with b_x.item = x ;
3     if such a bin was found then
4         b_x.count ← b_x.count + 1 ;
5     else
6         b_min ← bin with minimum count value ;
7         b_min.count ← b_min.count + 1 ;
8         b_min.item ← x ;
```

A Motivating Example (Non-Commodity Hardware)



Jens Teubner, René Müller, and Gustavo Alonso. FPGA Acceleration for the Frequent Item Problem. *ICDE 2010*.

Course Content

- Cache Awareness
  - How can we place data in memory and access it in a way that makes optimal use of memory caches?
- Query Execution
  - How can we tune our algorithms to fit modern processor architectures?
- Multi-Core Architectures
  - How can we exploit the parallelism provided by multi-core architectures?
- Specialized Hardware
  - How can we (mis-)use specialized hardware for data processing (*e.g.*, GPUs, FPGAs, modern NICs)?

Course Organization

Lecture:

- Thursdays, 13–15h, Room CAB G 59
- Course website:
  `http://www.systems.ethz.ch/courses/fall2012/DPMH`
  Please visit this website **regularly**. We will frequently post new information during the semester.

Exercises:

- Mondays, 15–16h, Room CAB G 59 (right after lecture)
- Held by **Cagri Balkesen** (`cagri.balkesen@inf.ethz.ch`)
- First exercise: today
- Exercise material is part of the course content!

Exam

There will be a **written exam** in the **exam session**.

- Duration: 90 min
- More details when the semester end approaches.

Course Setup

- I'd like to make this course highly **interactive**.
  - Please speak up, discuss, ask questions!
- The material we discuss is relevant in **practice**.
  - We'll provide practical examples and exercises.
  - You achieve maximum fun factor if you verify techniques on **your** machine.

Course Material

This is **not** a standard course (often even "bleeding edge").

- There is **no real textbook** for this course.
  Computer architecture basics are covered in "Computer Architecture: A Quantitative Approach" by Hennessy and Patterson, though.
- I'll make **lecture slides** available on the web.
- Most material is taken out of **research papers**.
  - I'll give references to those papers.
  - These are all good and easy-to-read papers.
- If you attended "Architecture & Implementation of DBMS," you'll recognize some ideas again.

MonetDB

MonetDB implements many of the techniques we'll talk about.

- MonetDB is open-source: `http://monetdb.cwi.nl/`
- Support for SQL and XQuery[1], multi-platform
- Numerous tools to look "under the hood" of MonetDB.
- Primary development: CWI Amsterdam
- Try it out yourself!

[1] `http://www.pathfinder-xquery.org/`

We want **You**

Avalanche is a related research project within the Systems Group.

- The project is highly active and internationally very successful.
- **You** could help us:
  - Labs, semester projects, Master projects, etc.
- **Approach me** if you'd like to contribute to a vibrant project at the forefront of research.

More information about Avalanche:

`http://www.systems.ethz.ch/research/projects/avalanche/`