EBERHARD KARLS UNIVERSITÄT TÜBINGEN

DBMS´s On A Modern Processor: Where Does Time Go?

Niklas Heinsohn

Seminar: “Turn Head 90 ° - Column Store Database”

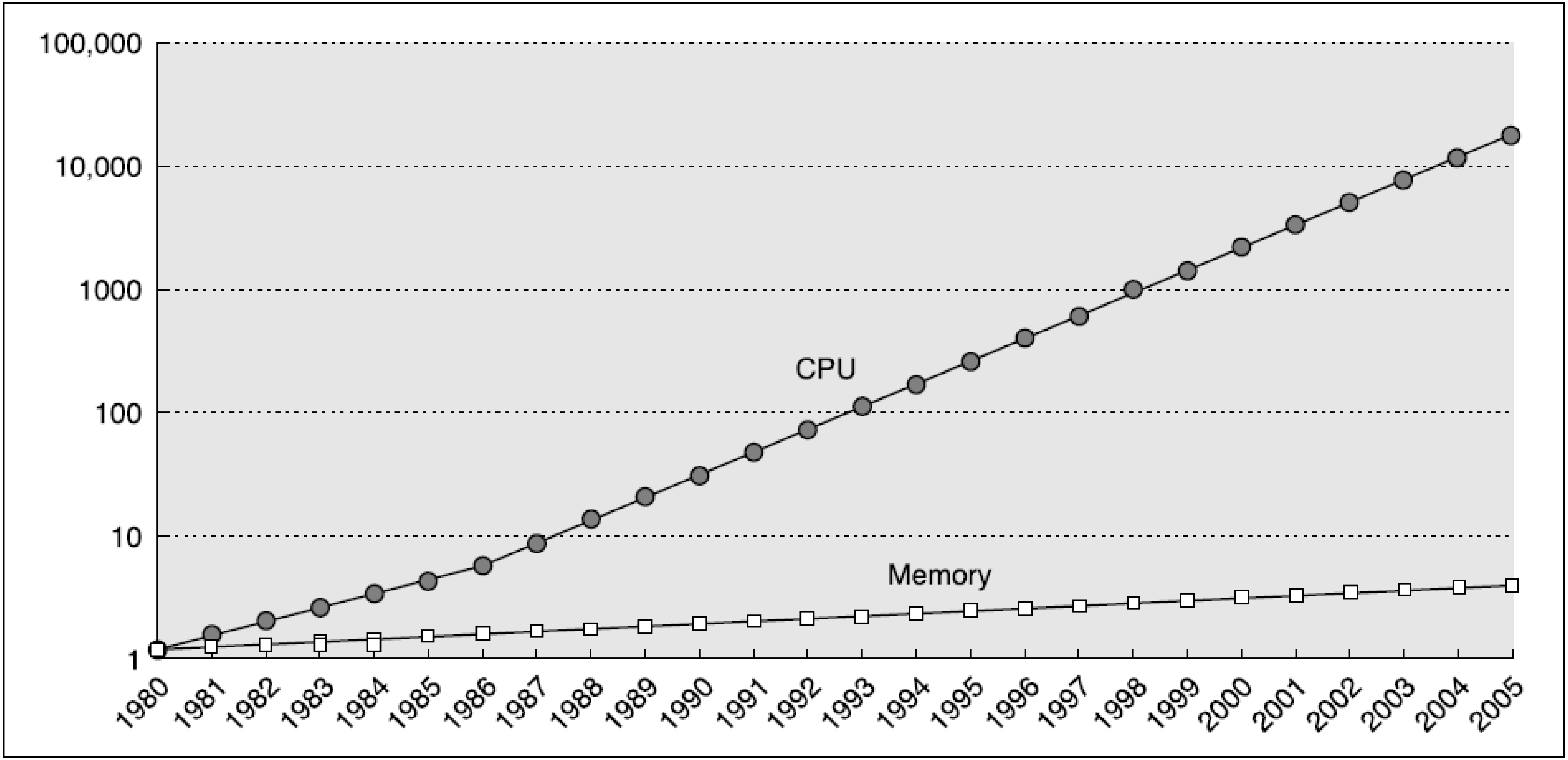
WS 10/11

Eberhard Karls Universität Tübingen

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Half of the execution time is spent on stalls

The CPU – Memory speed gap



A Operation on a Processor

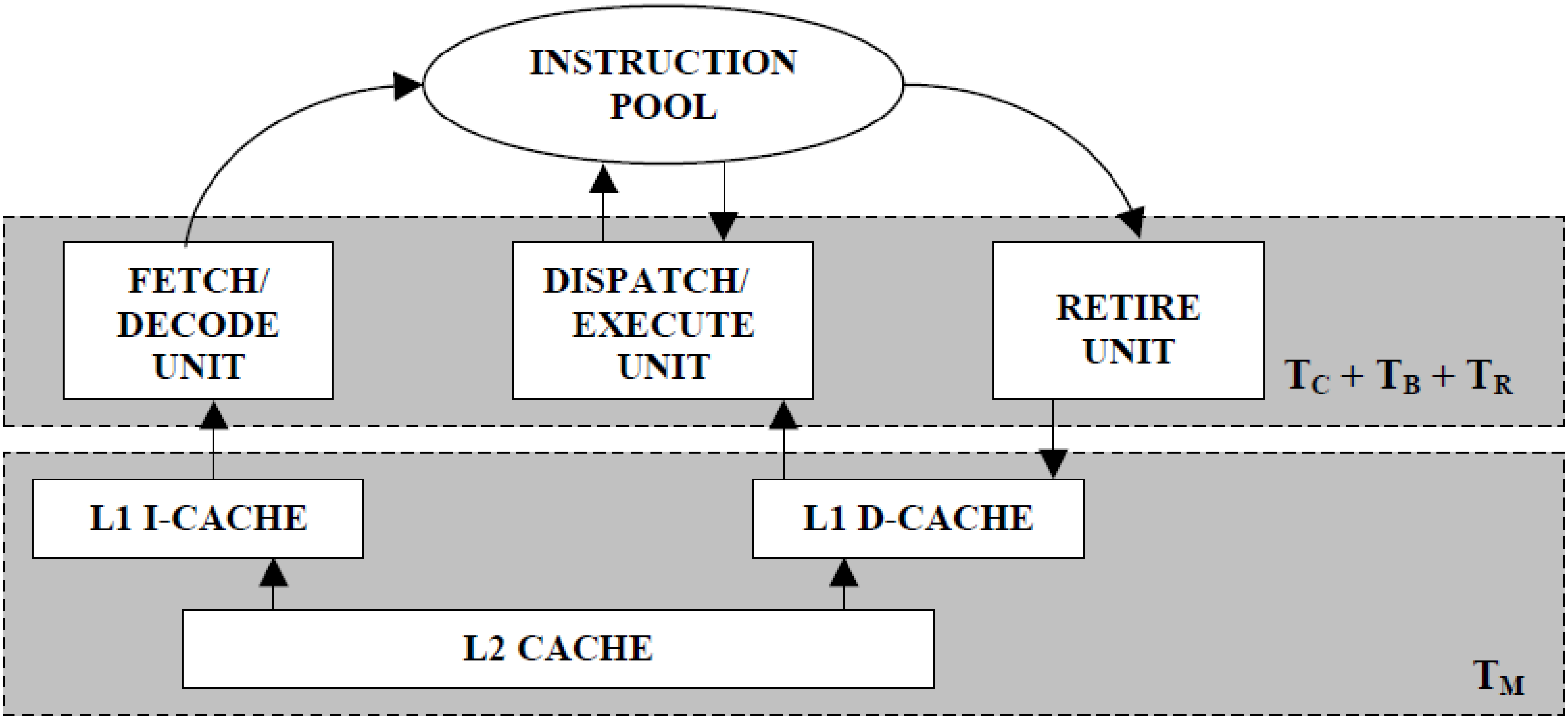


Figure 2.1: *Simplified block diagram of a processor operation*

Measurement

- 4 different DBMS`s
- one operating system
- CPU performance counter

- Sequential range Selection
- Indexed range Selection
- Sequential Join

The Table

```
Create Table R ( a1 integer not null,
                 a2 integer not null,
                 a3 integer not null,
                 <Rest of Fields>)
```

<Rest of Fields> list of integers not used by any Queries

- 1.2 million 100-byte records
- value of a2 uniformly distributed between 1 and 40 000

The Queries

Sequential Selection

select Avg(a3)
from R
where a2<Hi and a2>Lo

Range Selection

non-clustered index on R.a2

Sequential Join

S defined same way as R
select Avg(R.a3)
from R,S
where R.a2 = S.a1

Execution Time

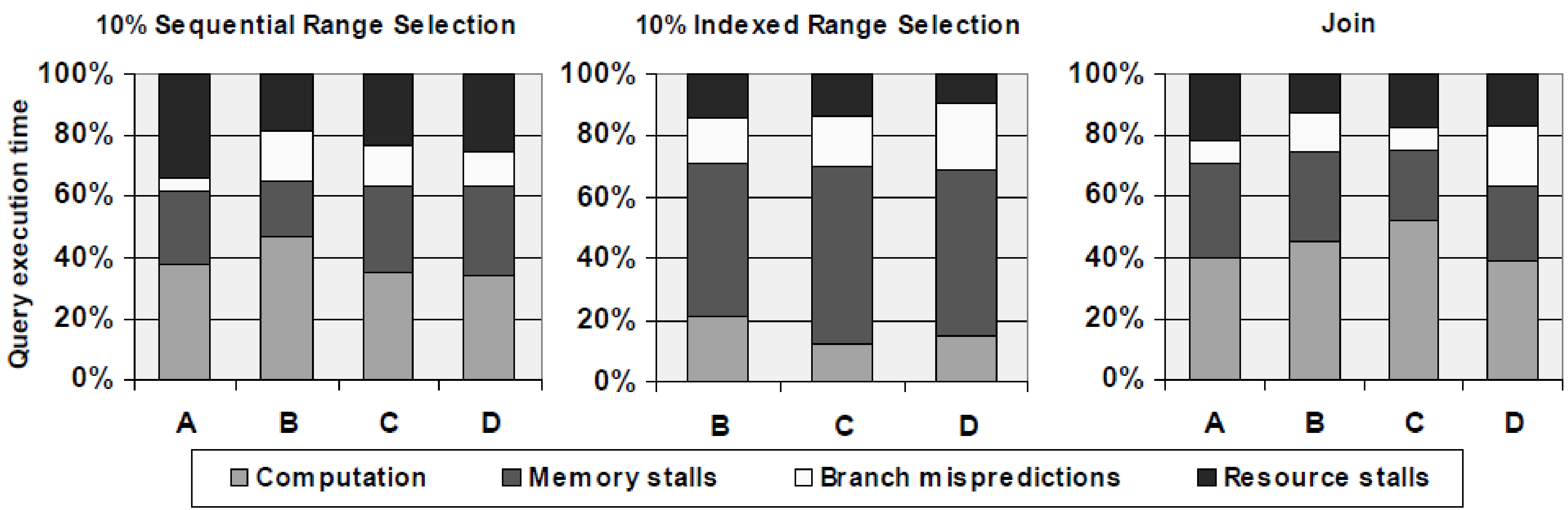


Figure 5.1: *Query execution time breakdown into the four time components.*

Methods to cover stall time

- Non-blocking Cache
- Out-of-Order execution
- Speculative execution

Resource related stalls

- Functional unit stalls
- Dependency stalls
- Instruction length decoder stalls

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Branch predictions

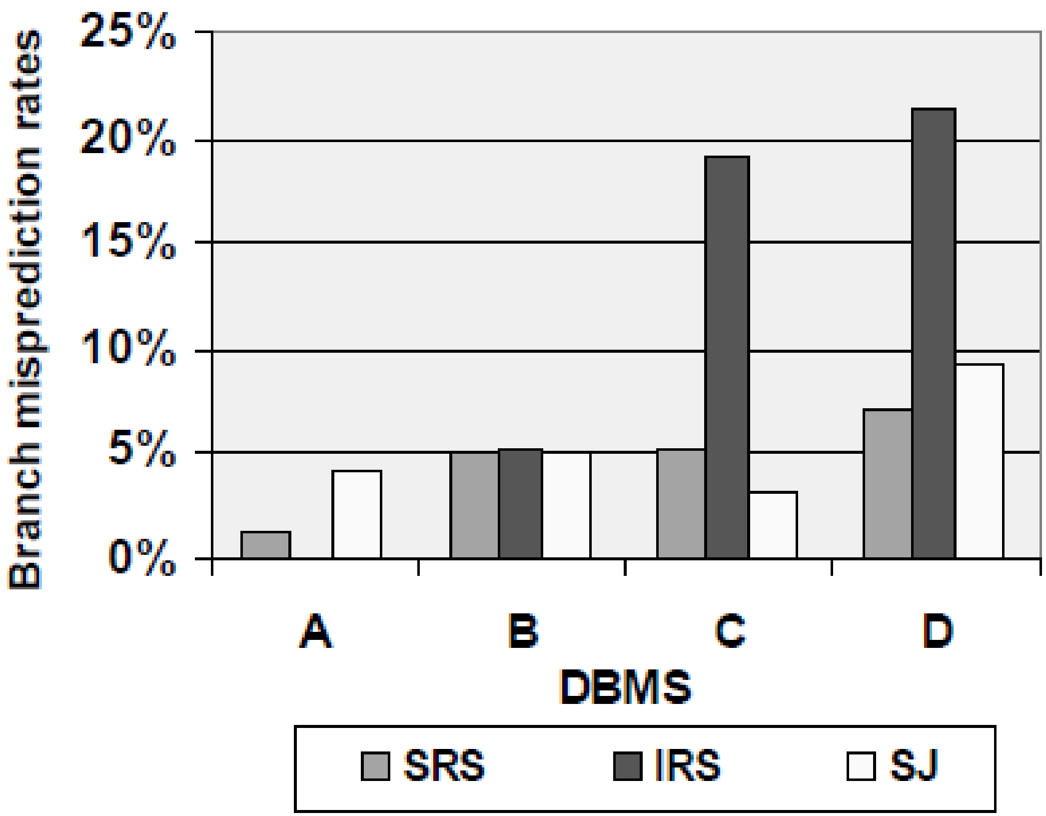


Figure 5.4: *Branch misprediction rates. SRS: sequential selection, IRS: indexed selection, SJ:* join.

Memory stalls

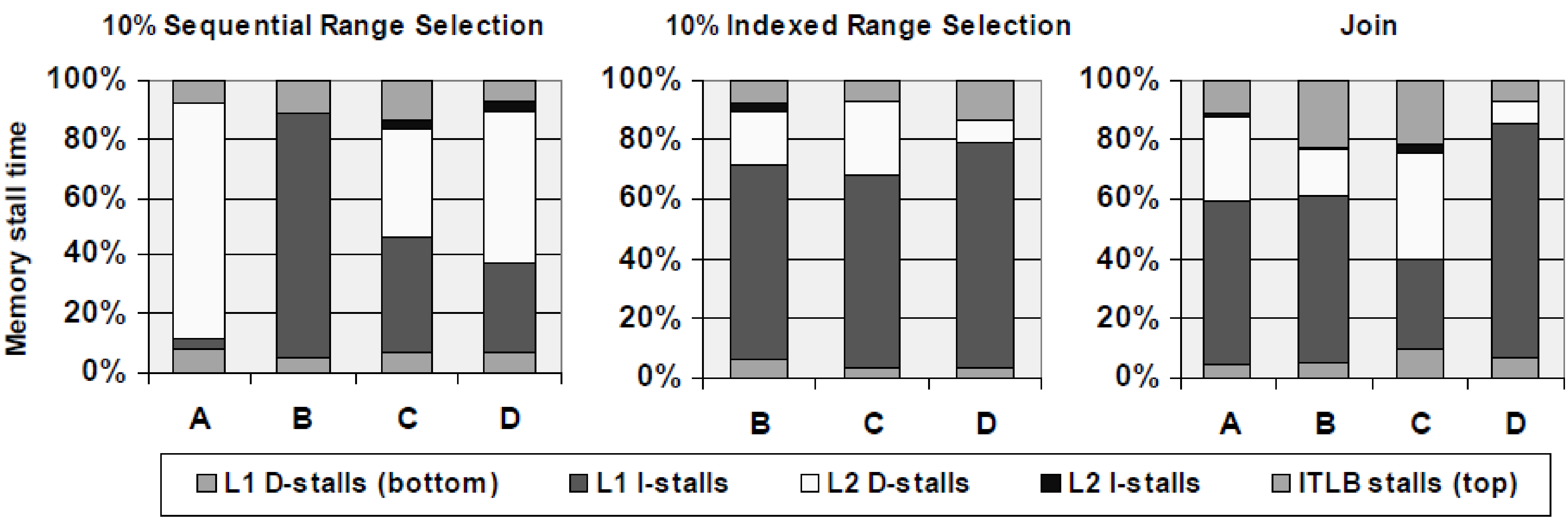


Figure 5.2: *Contributions of the five memory components to the memory stall time (T_M)*

Second-level cache data stalls

- L2 cache miss rate between 40% and 90%
- Larger records increase stall time
- Multiple L2 data misses can overlap with each other
- Increasing cache size leads to longer latencys

First-level cache instruction stalls

- L1 I-misses are a bottleneck
- Difficult to overlap
- Between 4% and 40% of execution time
- ? Larger records cause more L1 I-misses ?
  - Inclusion (L1 cache may contain blocks present in L2-Cache)
  - Frequent boundary crossings in L2 cache

Translation lookaside buffer (TLB)

- Virtual to physical address translation
- improve virtual address translation speed
- Average miss rate ~1 %
- Hit takes ~ 0.5 to 1 CPU clock cycles
- Miss takes ~ 10 to 100 CPU clock cycles

Results

- 50% of execution time the CPU stays stalled
- 90% of memory stalls are due to
  - Second Level data misses
  - First Level instruction misses
- There is always a bottleneck

Sources

- "DBMSs On A Modern Processor: Where Does Time Go?" By Anastassia Ailamaki, David J. DeWitt, Mark D. Hill, David A. Wood
- Slide "Database Systems and Modern CPU Architecture" "Memory Close to the CPU: Caches", 5. Juli 2009 by Torsten Grust
- http://www.pro-datenrettung.net/uploads/pics/festplatte-schnittstellenpin-abgebrochen.jpg
- http://www.n24.de/media/import/dpainfoline/dpainfoline_20100102_13/jpeg-14784400EF32095A-20100102-img_23434826originallarge-4-3-800-207-0-2871-1996.jpg
- http://www.gnurf.net/v3/wp-content/uploads/2008/03/025-magician-01.png