

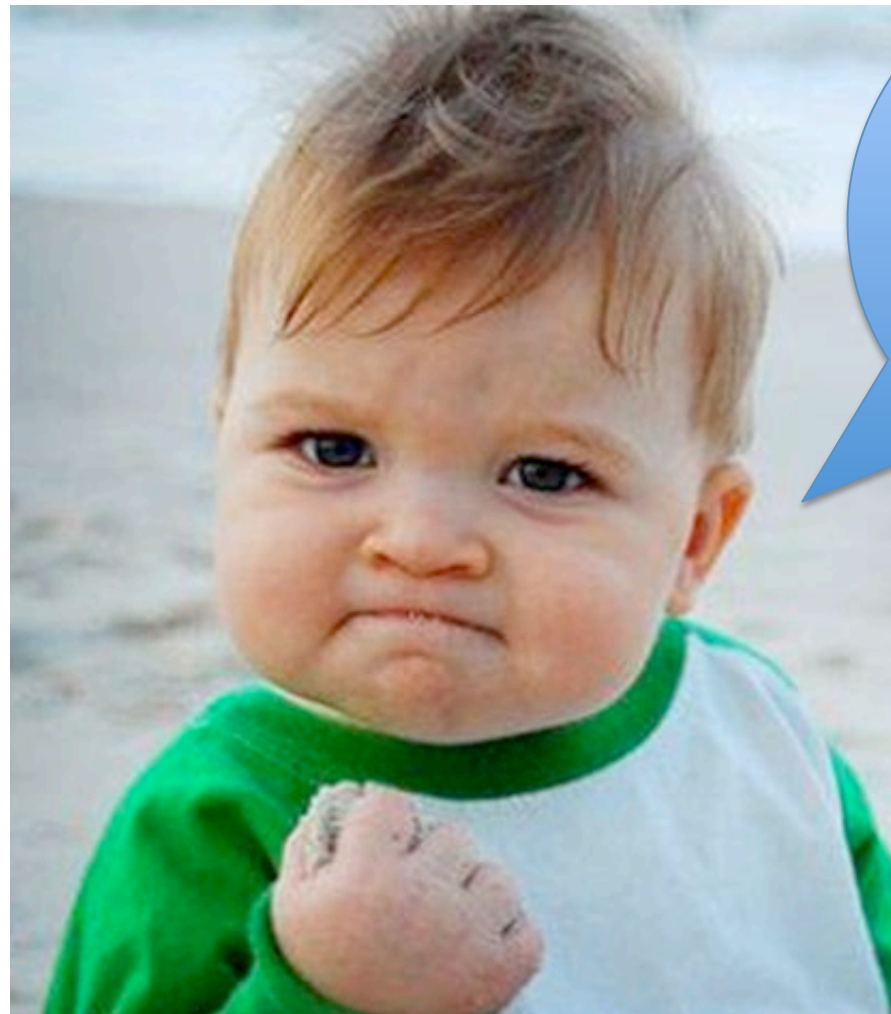
DBs and hardware:
the beginning and sequel
of a beautiful friendship



Anastasia Ailamaki

EPFL and RAW Labs SA

FRIEND OR BULLY?



pipelining
ILP
SMT
multi-core
heterogeneous H/W

recurring headaches

non-uniform communication

underutilization of resources

one size fits none

aspirin or vitamin?

recurring headaches

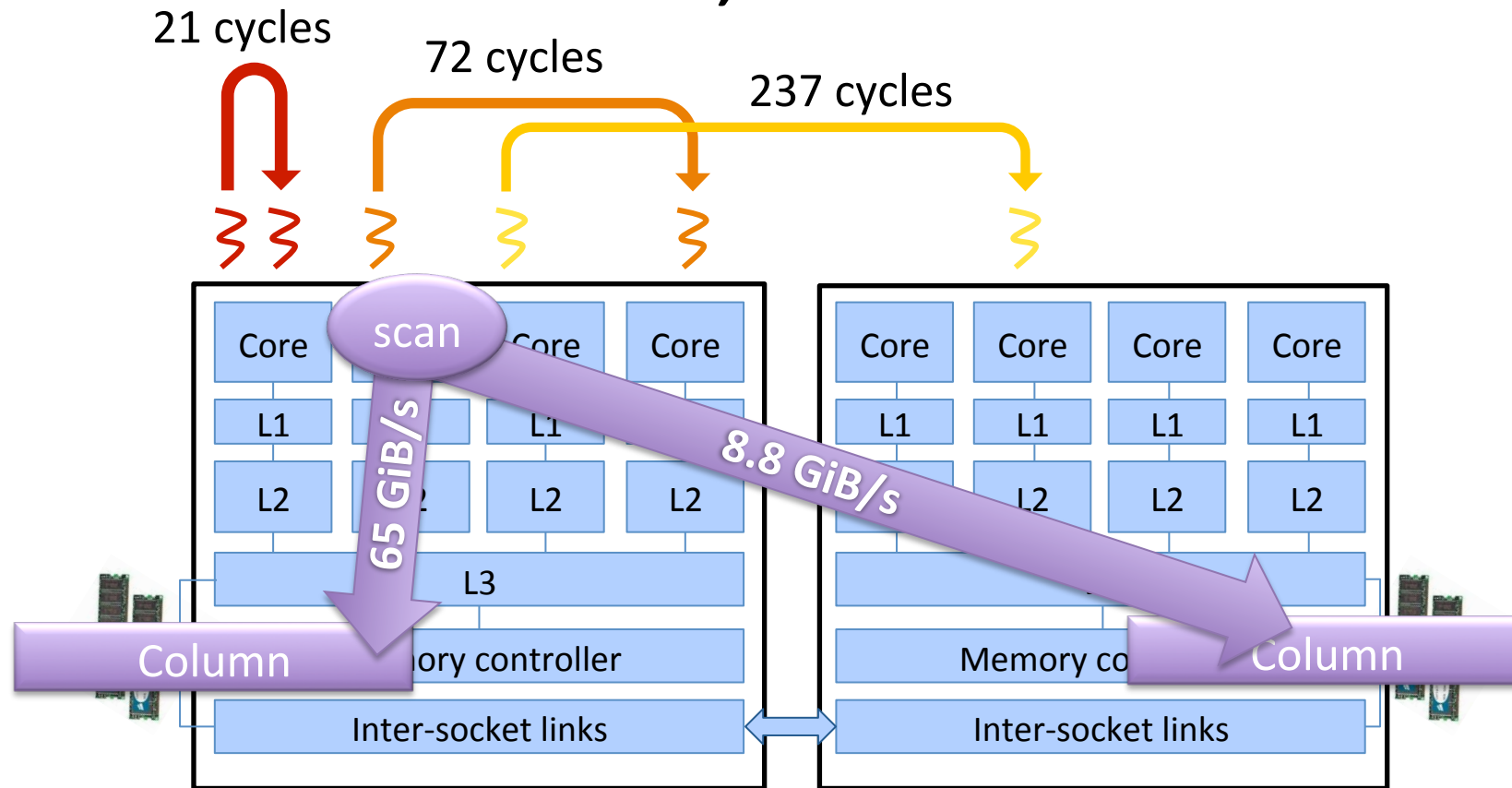
non-uniform communication

underutilization of resources

one size fits none

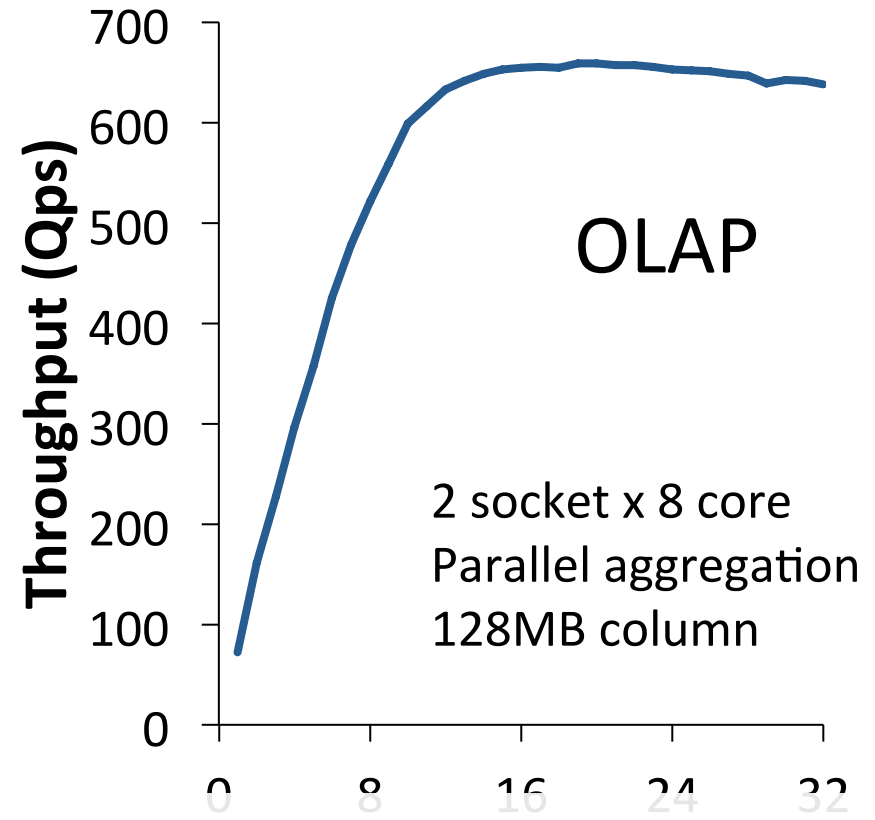
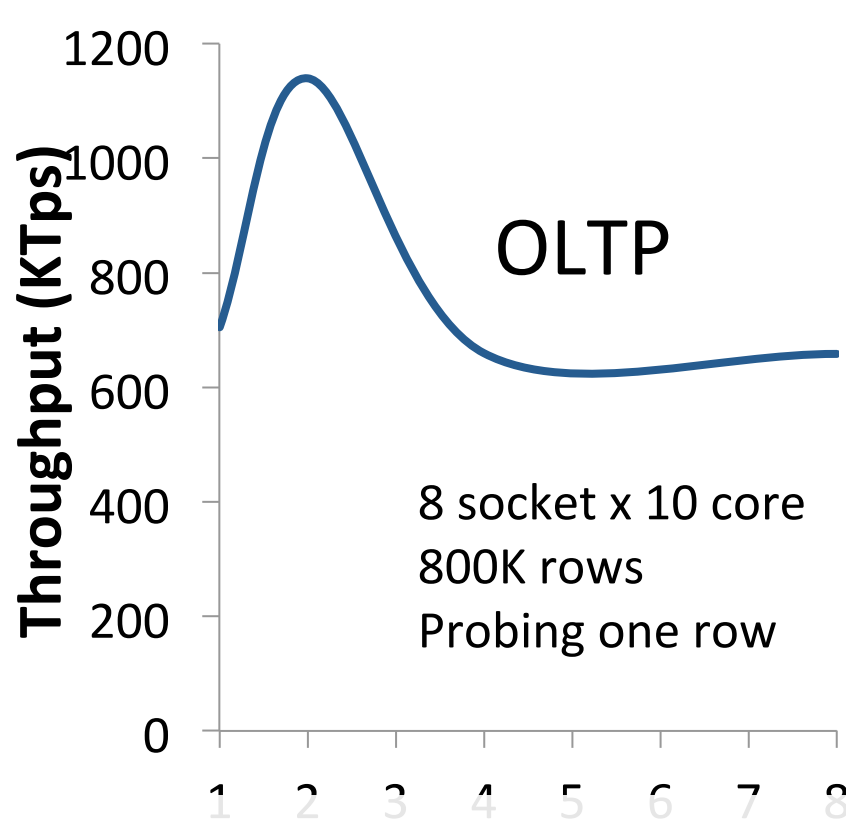
aspirin or vitamin?

NU: MA, and more



unpredictable performance

scalability on multicores

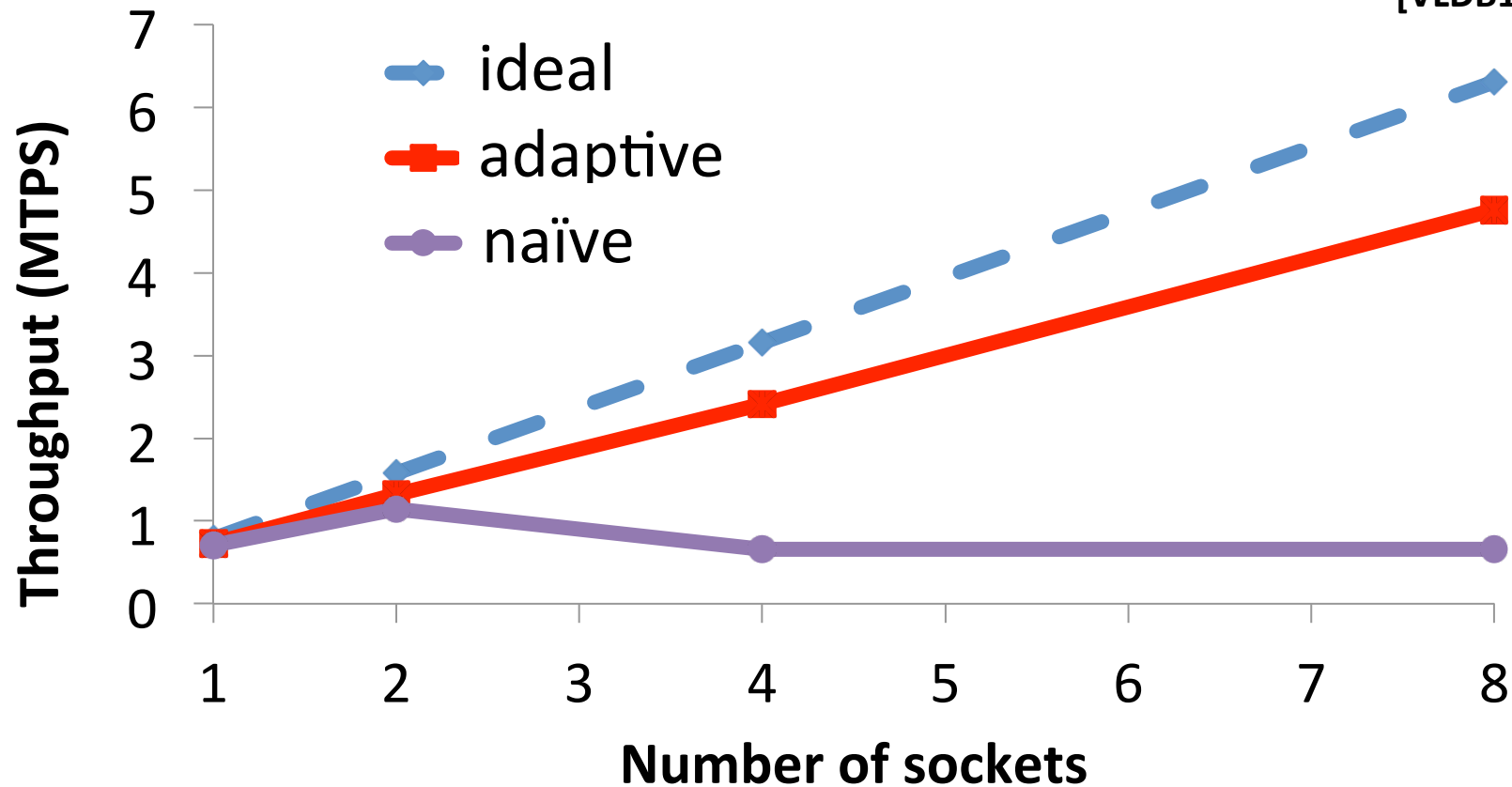


vitamin: adaptive partitioning

access latency memory bandwidth

OLTP: communication islands

[VLDB12, ICDE14]

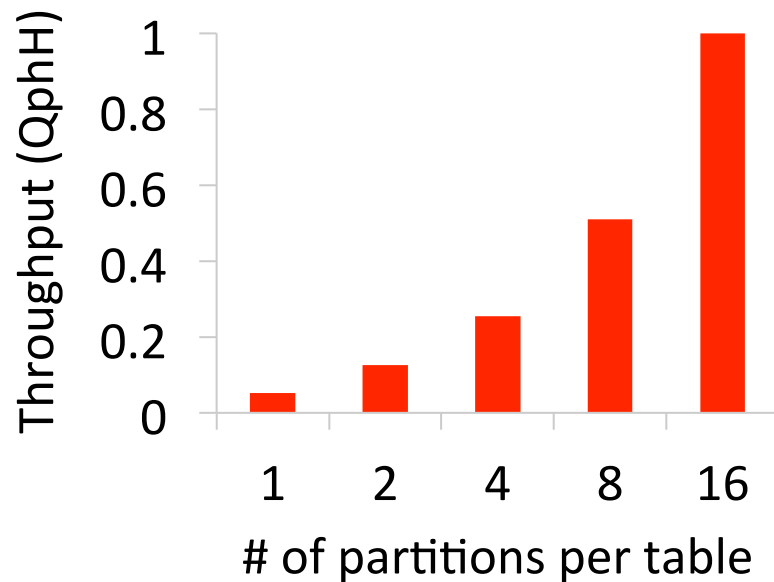


near-linear scalability

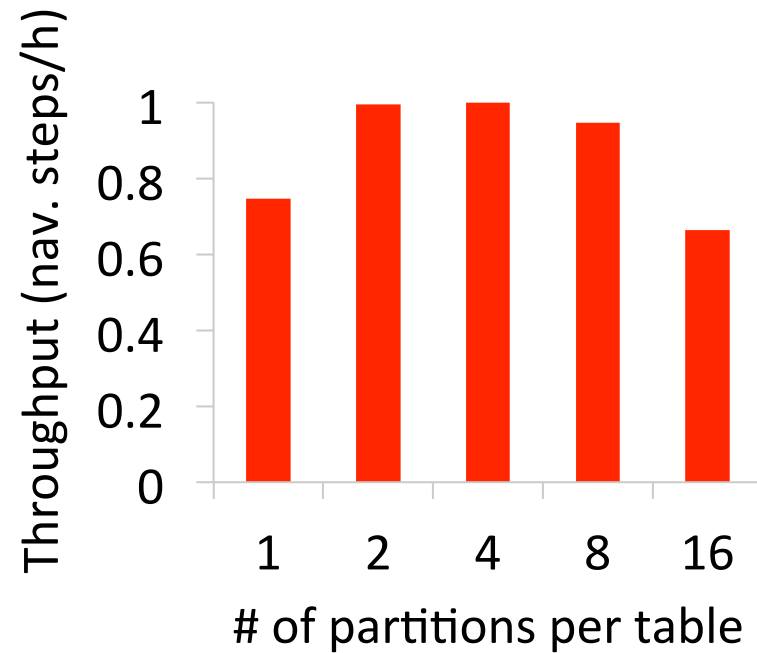
OLAP: partitioning across sockets

[VLDB15]

1 hot table



3 hot tables



fewer partitions → higher throughput

recurring headaches

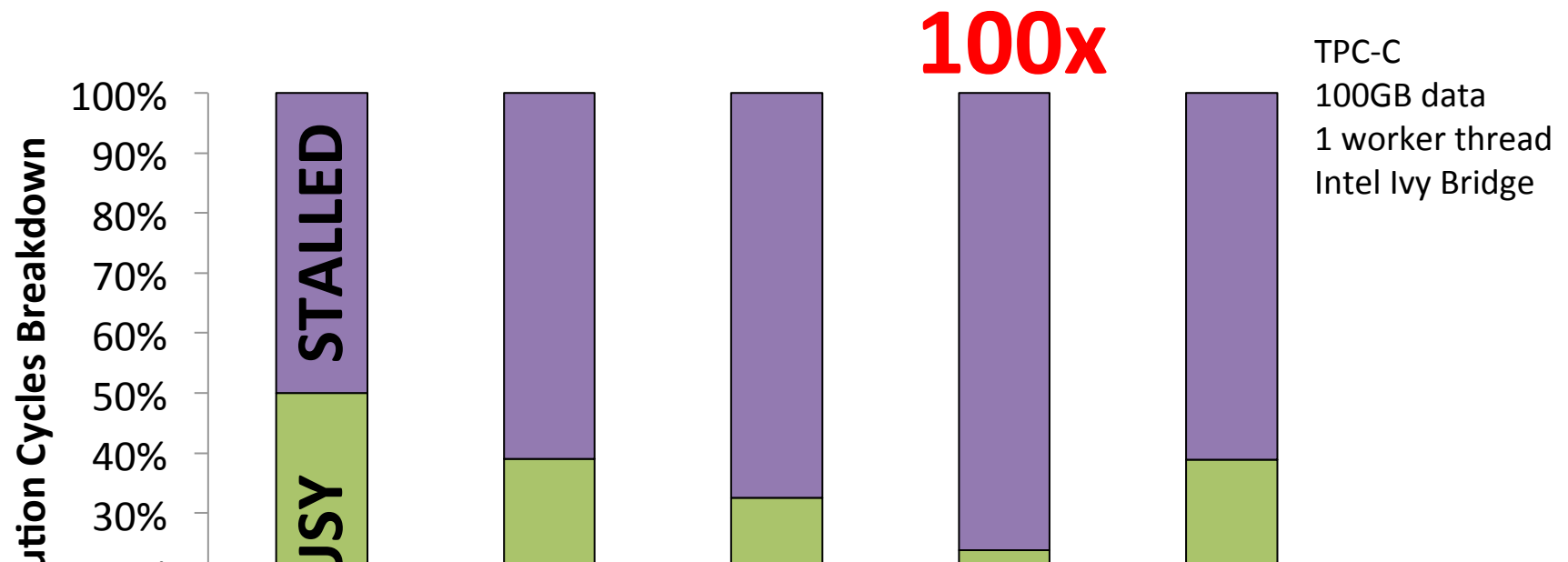
non-uniform communication

underutilization of resources

one size fits none

aspirin or vitamin?

processor is idle most of time



use cores

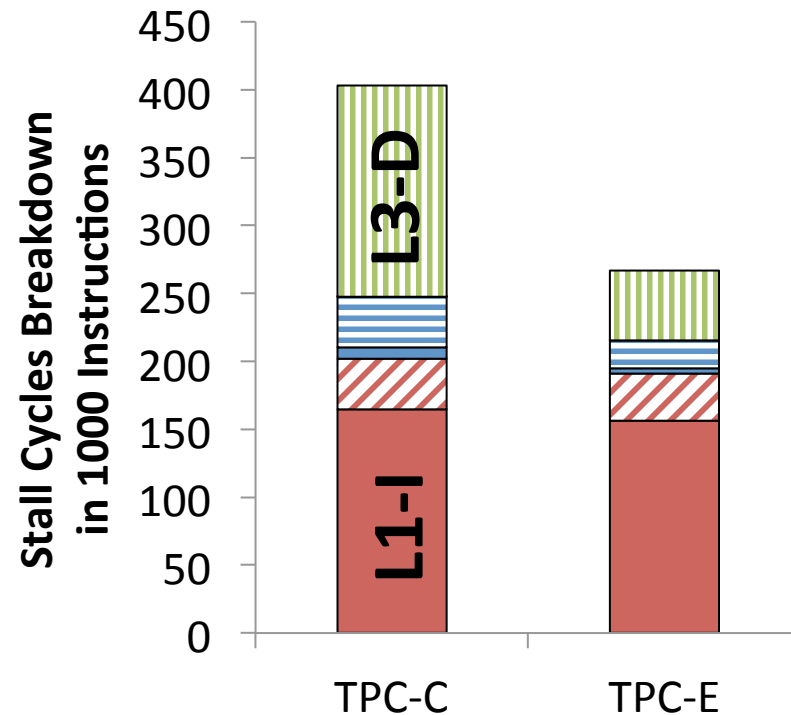
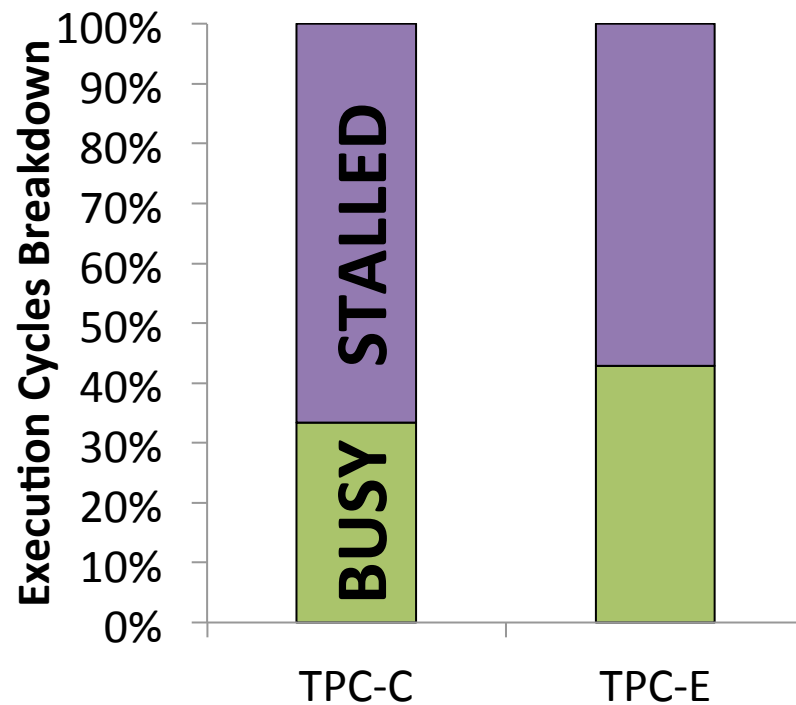
improve performance

save energy

instruction stalls (still) rule

100GB data on Shore-MT
Intel Xeon E5-2660

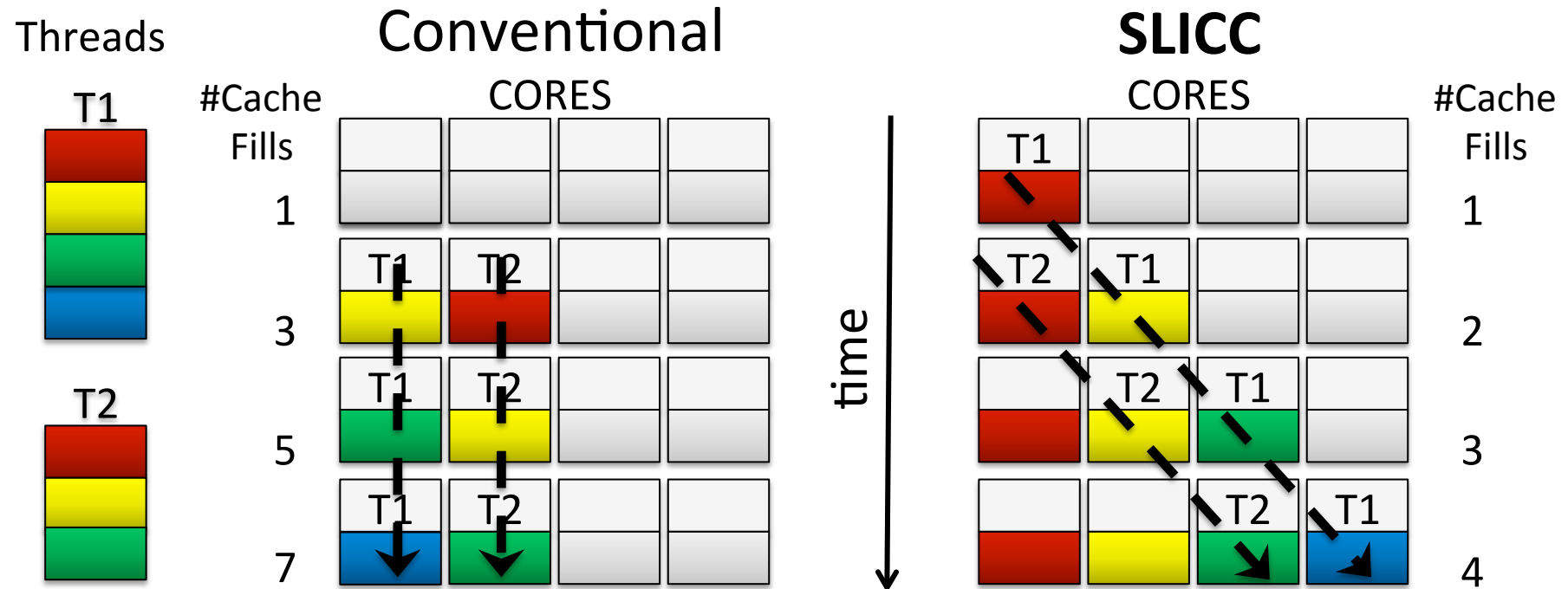
[EDBT13,DAMON13]



vitamin: follow L1-I contents

SLICC: chasing locality in L1-I

[MICRO12,ISCA13,VLDB15]



improved performance by 70%

recurring headaches

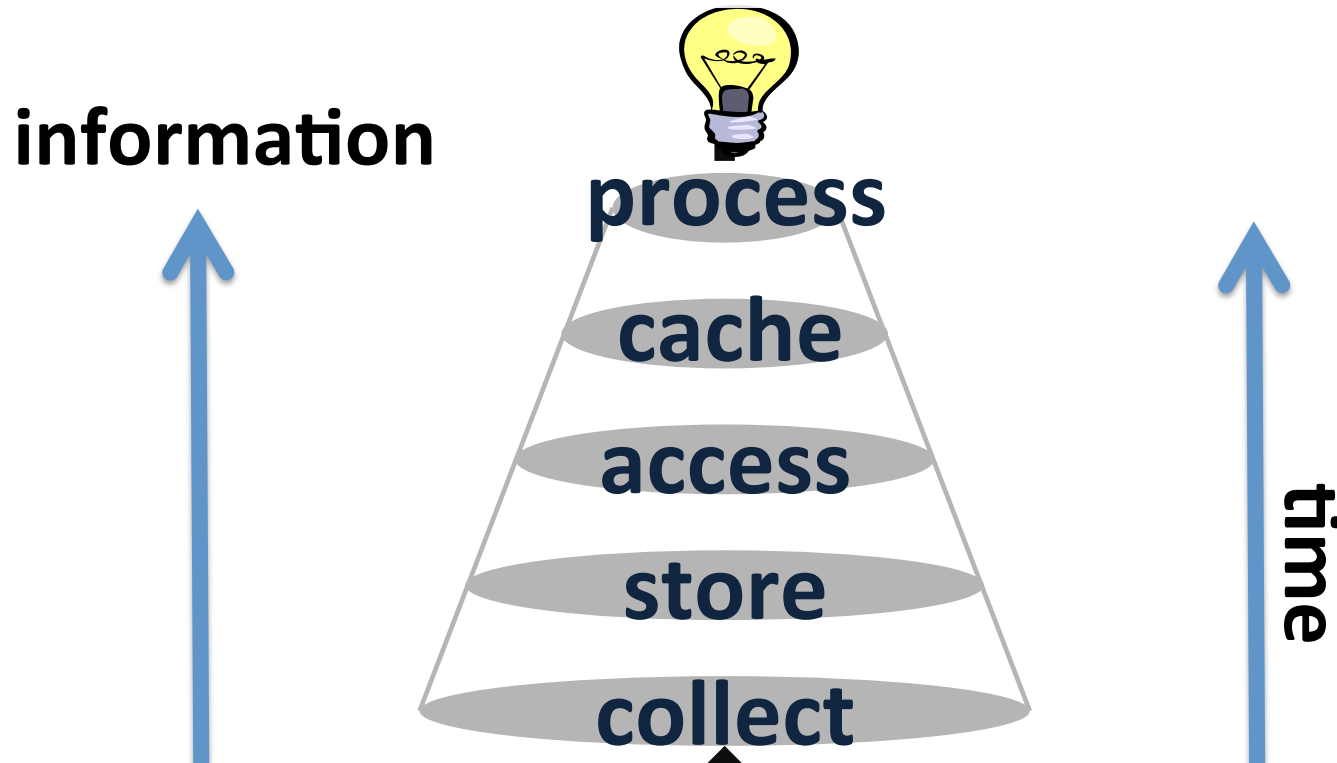
non-uniform communication

underutilization of resources

one size fits none

aspirin or vitamin?

one DBMS fits all

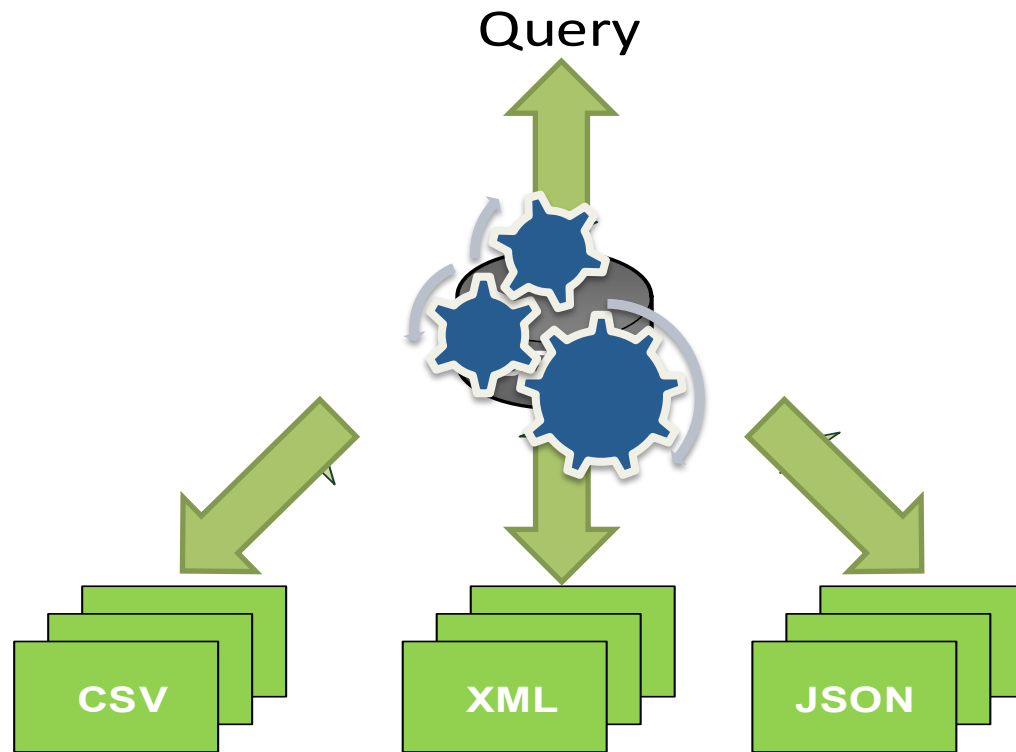


but: data and apps vary, so...
→ one DBMS fits one app/format/...

one DBMS fits NONE



vitamin: just-in-time databases



access only useful raw data
(not clear *which* data)

no copying or altering data

no vendor lock-in

custom (per-query) engine

RAWdb: an evolving query engine

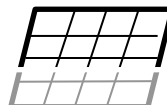
how to build a just-in-time database

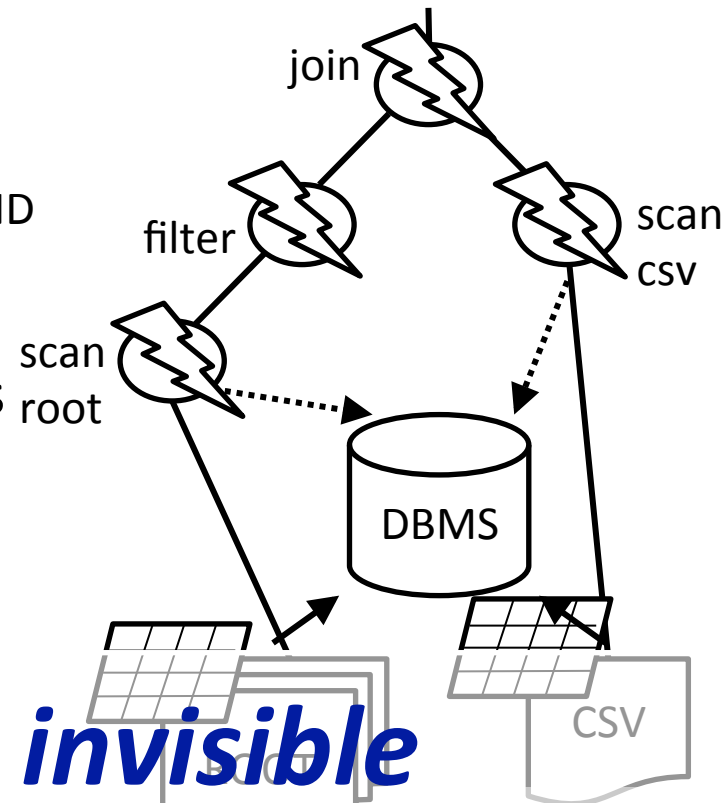
[[CIDR11,SIGMOD12,VLDB14]

```
SELECT event.jet...  
FROM csv, root  
WHERE csv.RunNumber == root.RunNumber AND  
      root.EF_2mu13 == TRUE AND ...
```

 Code Generate the Access Paths

 Code Generate the Query

 Build Position and Data Caches



RAWdb is *invisible*
adapts to *data* and *queries*

... containing "group" run numbers
physics events

aspirin → lag behind



catch me



if you can



adaptive data partitioning
chasing instruction locality
just-in-time databases

BACKUP

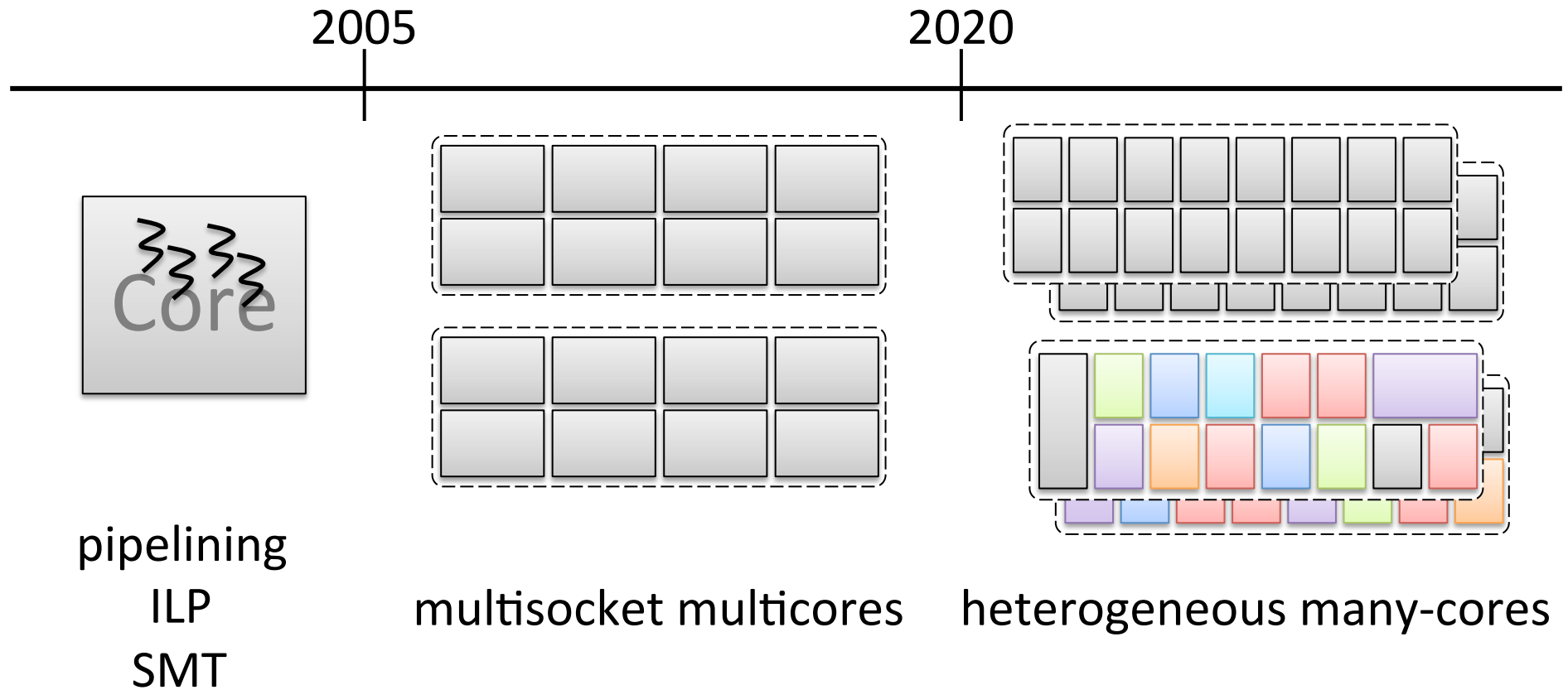


Databases and Hardware



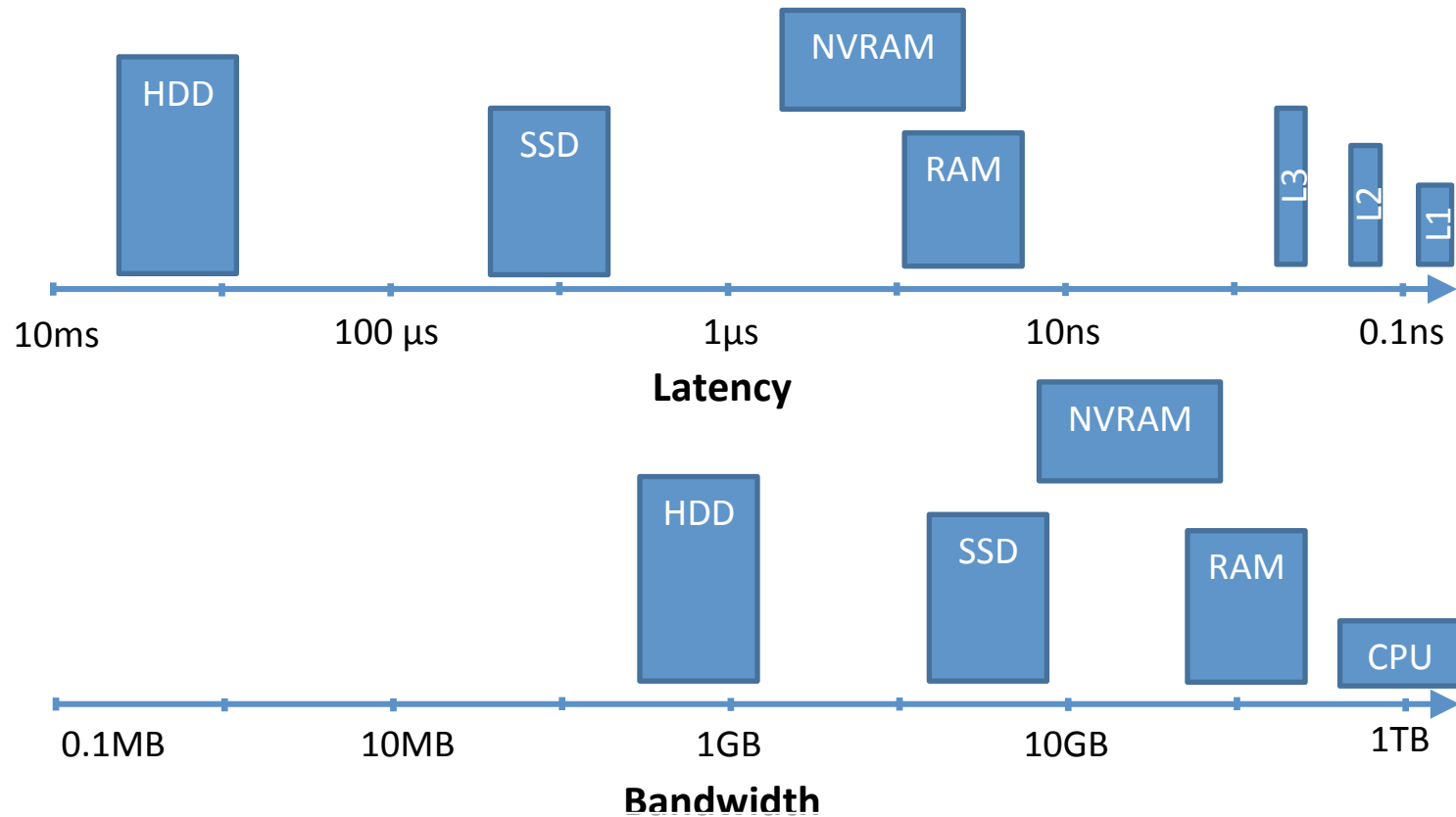
- abundant compute power
 - increasing heterogeneous parallelism
 - deep storage hierarchies
- to use it
 - adapt to hardware and application
 - Just-in-Time Databases

adding compute elements



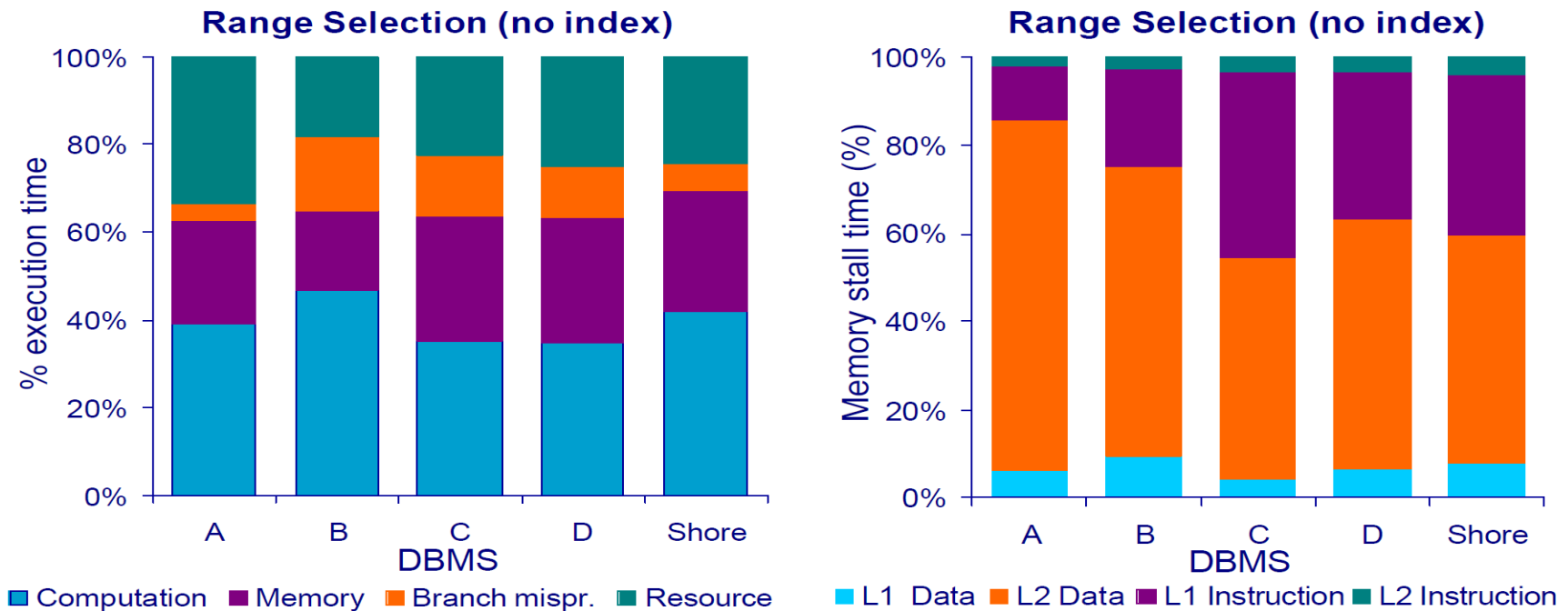
abundant heterogeneous parallelism

storage depths



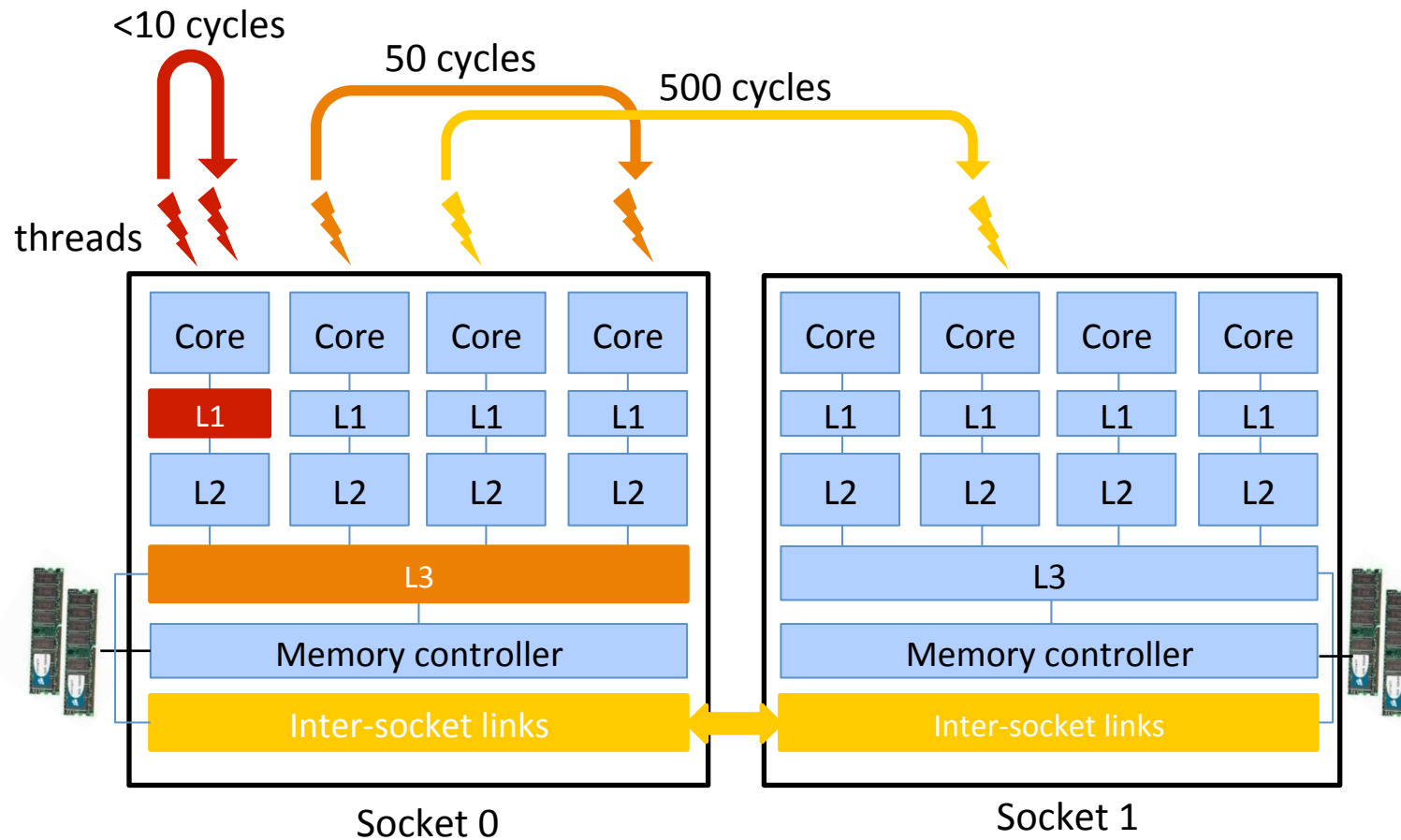
deep memory hierarchy

once upon a time ...



processor stalled 50% of time

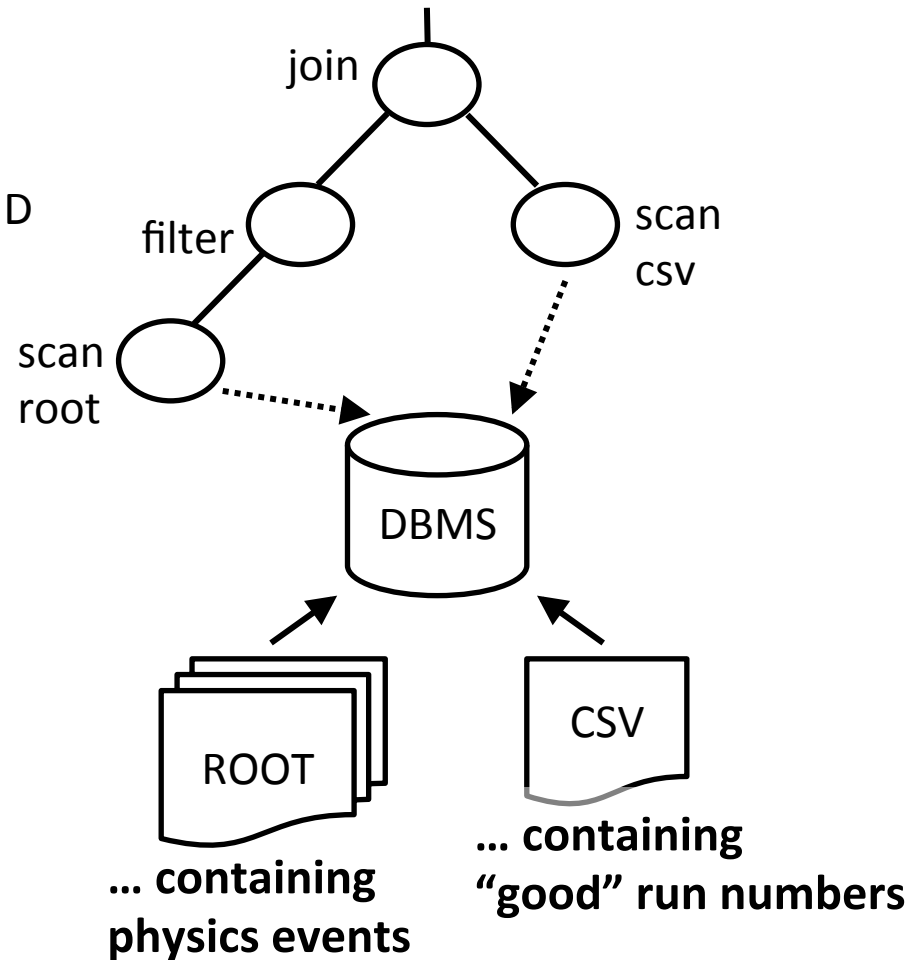
non-uniform communication latencies



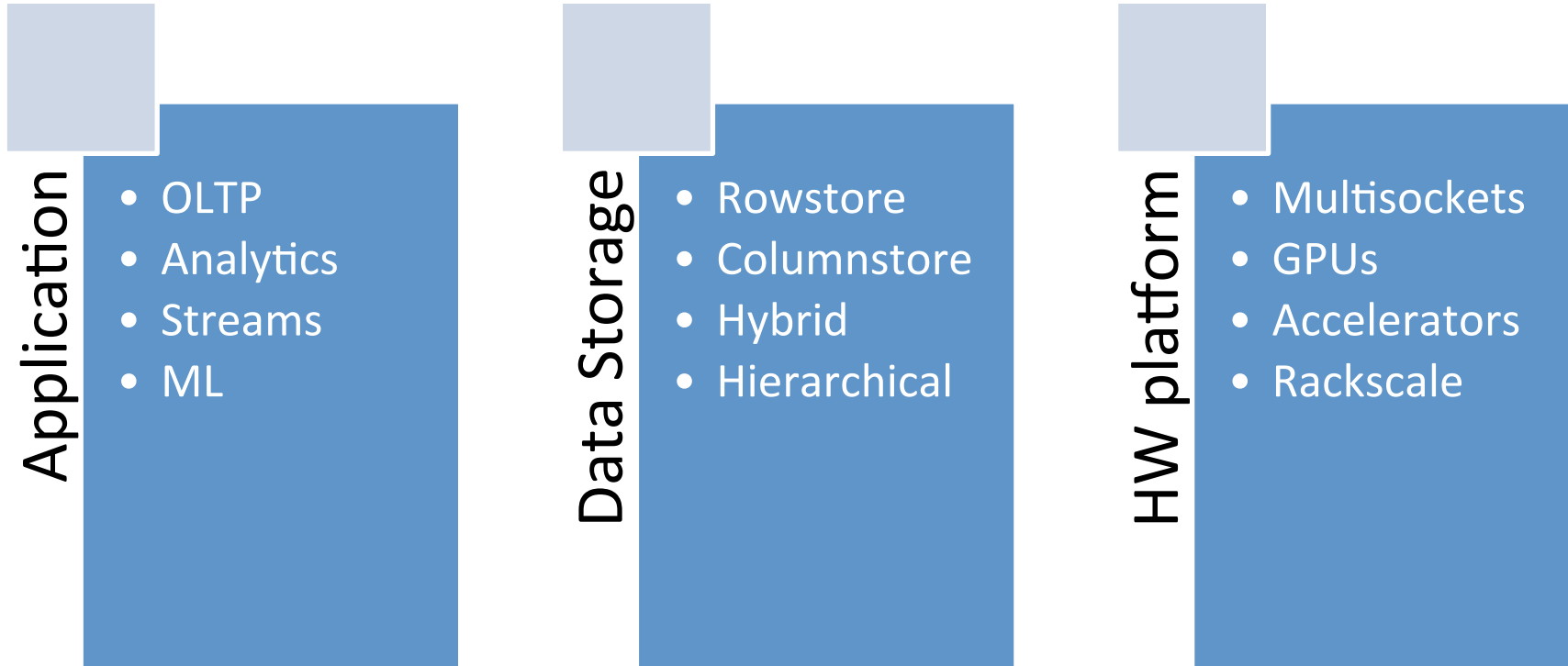
non-uniform

How to build a just-in-time data base

```
SELECT event.jet...  
FROM csv, root  
WHERE csv.RunNumber == root.RunNumber AND  
      root.EF_2mu13 == TRUE AND ...
```



Optimize just-in-time for any



Application

- OLTP
- Analytics
- Streams
- ML

Data Storage

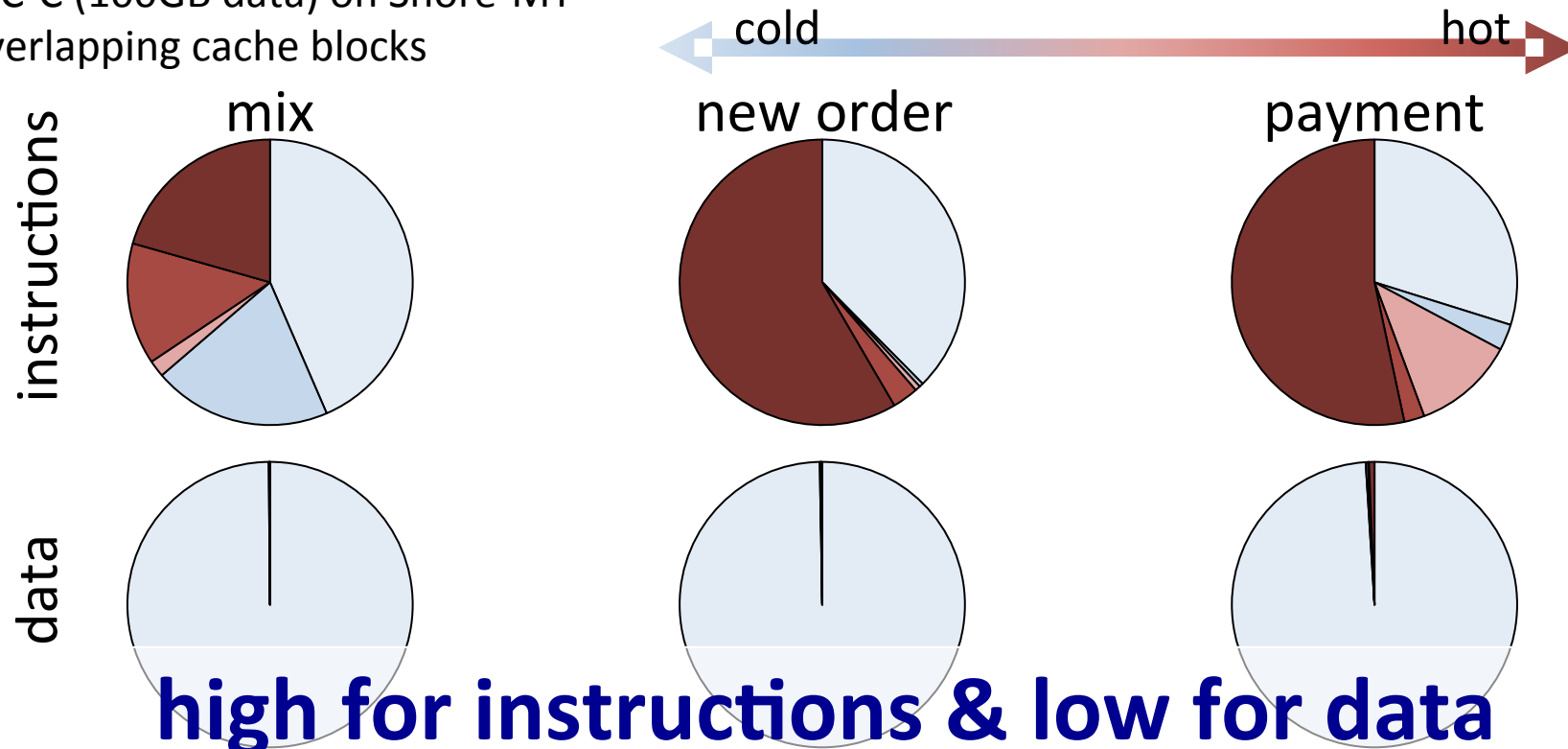
- Rowstore
- Columnstore
- Hybrid
- Hierarchical

HW platform

- Multisockets
- GPUs
- Accelerators
- Rackscale

obs: overlap across transactions

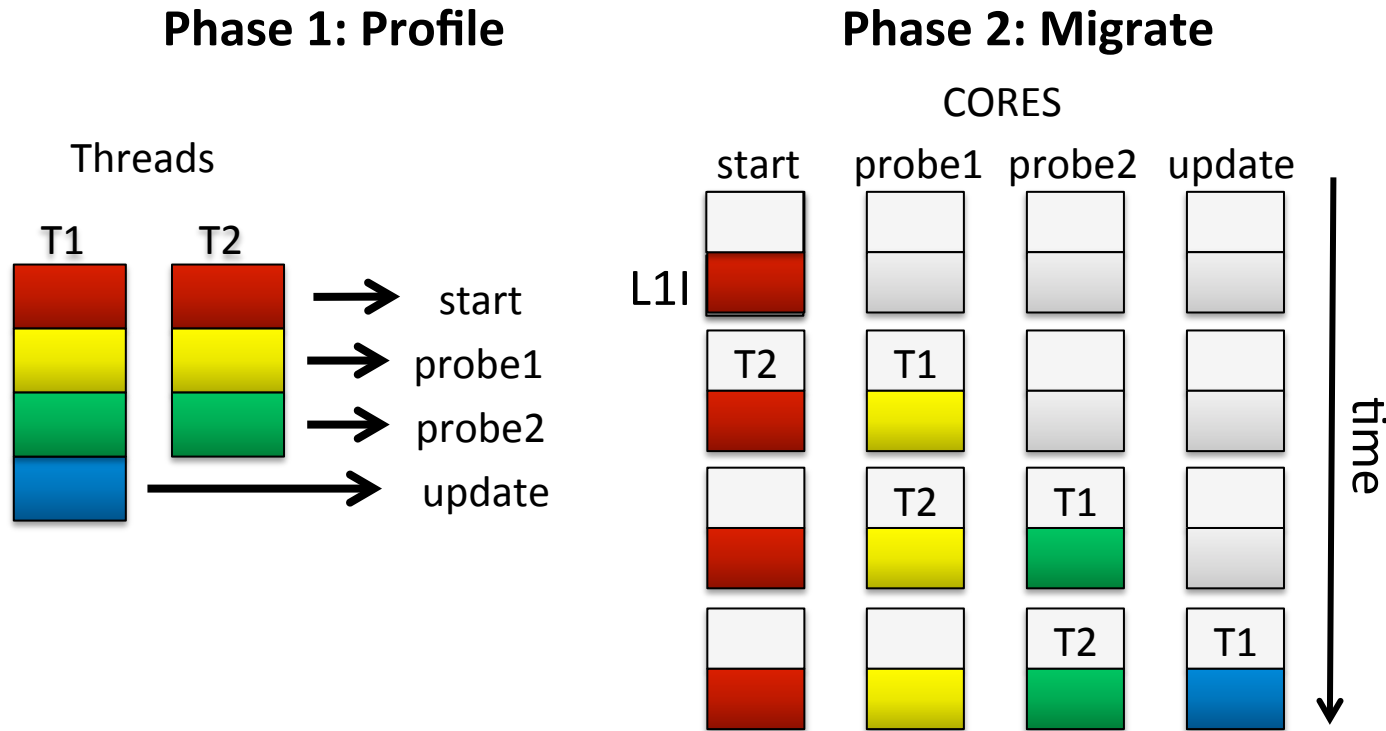
TPC-C (100GB data) on Shore-MT
Overlapping cache blocks



higher overlap among same-type transactions

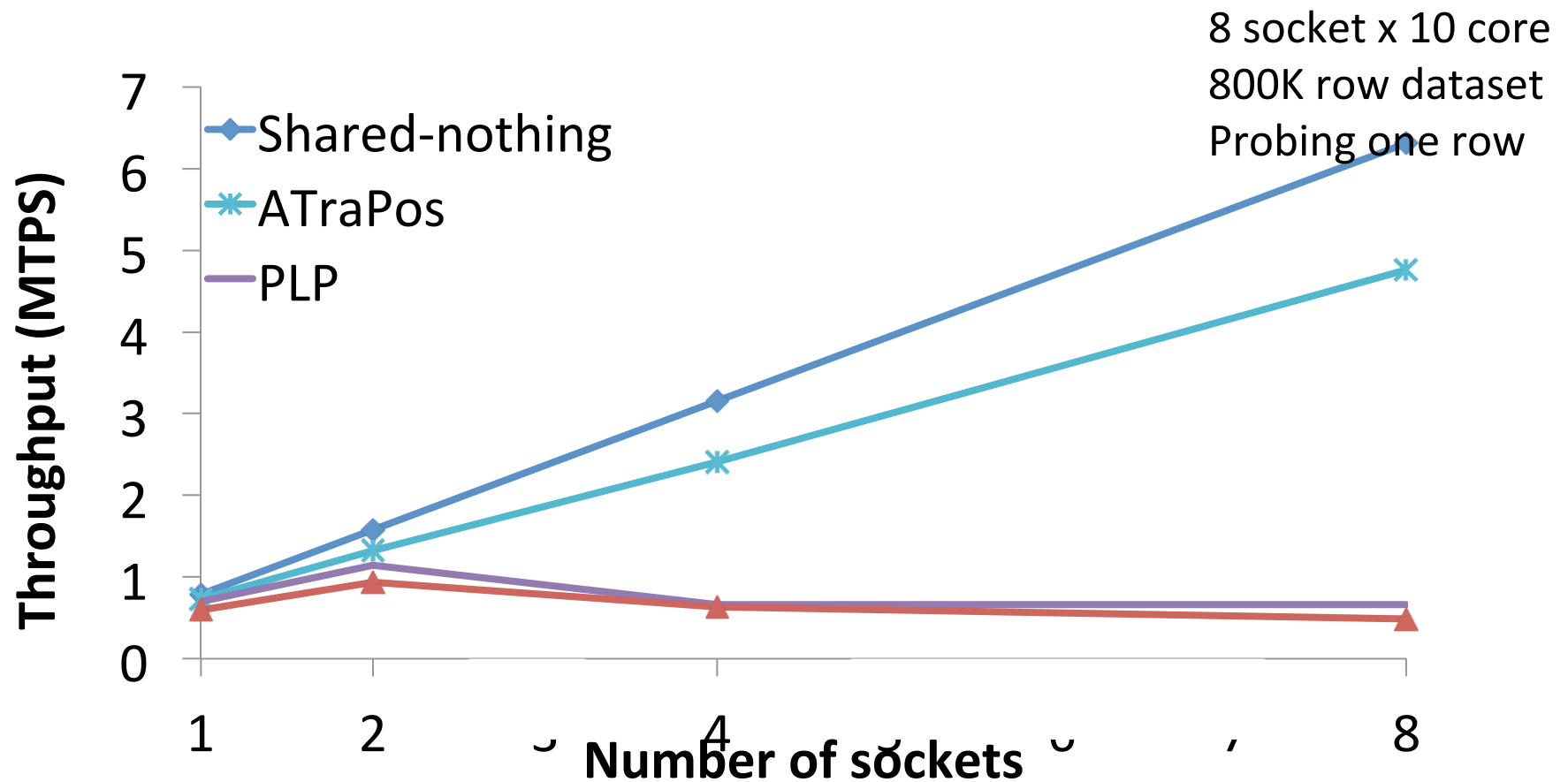
advanced instruction chasing

ADDICT



**compared to SLICC: minimal changes to hw & 2x
higher performance improvement**

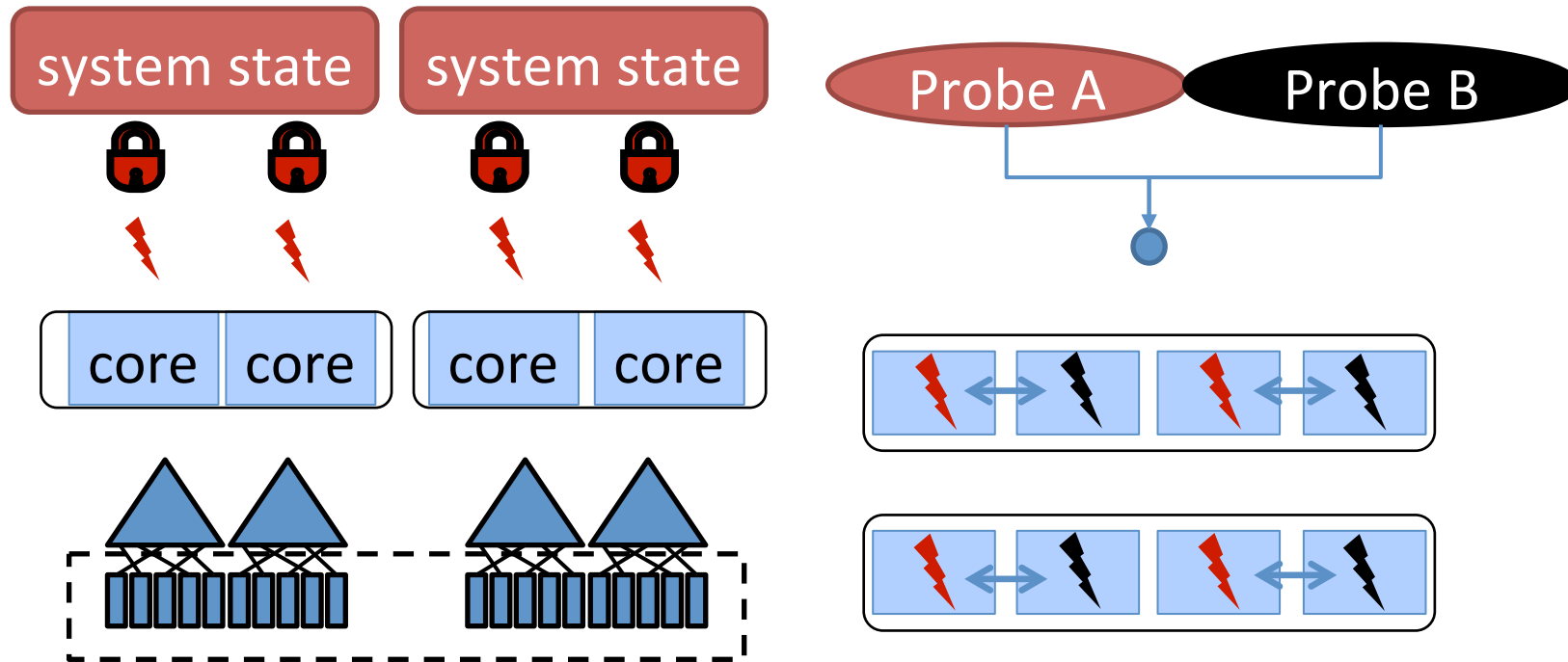
Perfectly partitionable workload



Island awareness brings scalability

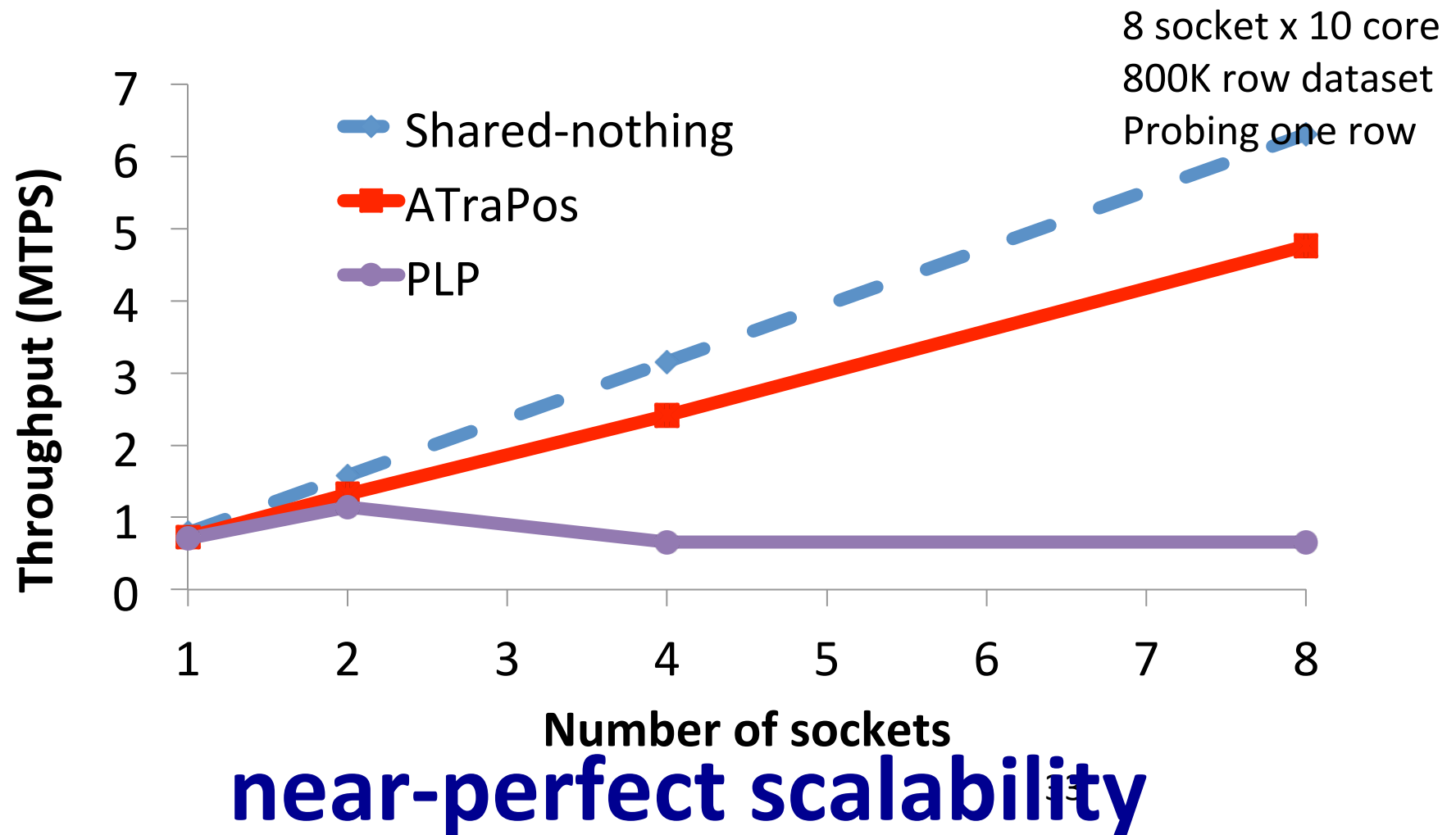
- Try harder to maximize use of hardware
- make databases invisible

ATraPos: adaptive transaction processing



adaptive partitioning and placement

adapting to islands



one DBMS fits one app/data/...

