Technische Universität München

# Heterogeneity-Conscious Parallel Query Execution: Getting a better mileage while driving faster!

Tobias Mühlbauer, Wolf Rödiger, Robert Seilbeck,
Alfons Kemper, Thomas Neumann

Technische Universität München

Data Management on New Hardware (DaMoN 2014)



http://www.hyper-db.com/

# The Rise of Dark Silicon

Moore's law still valid

- transistor density doubles with each generation

\+ Failure of Dennard Scaling

- proportional scaling of threshold/supply voltages failed in ~2002
- power density is growing

= Dimmed or Dark Silicon: > 50% at 10nm (ITRS roadmap projection)

- processor power budget is constant
- not all transistors can be used simultaneously
- multi-core scaling is just a workaround

⇒ Heterogeneous Processors (CPU including GPGPU, FPGA, ASICs, …)

# Single-ISA Heterogeneous Multi-Cores

Cores implement the same instruction set architecture (ISA)

Cache-coherent access to main memory

Advantages

- Different types of cores for energy efficiency and performance
- Multiple simple cores for high parallel performance, some complex cores for high serial performance (Amdahl's law)
- Single ISA = single implementation
- Avoid over-specialization (GPGPUs, ASICs, FPGAs)

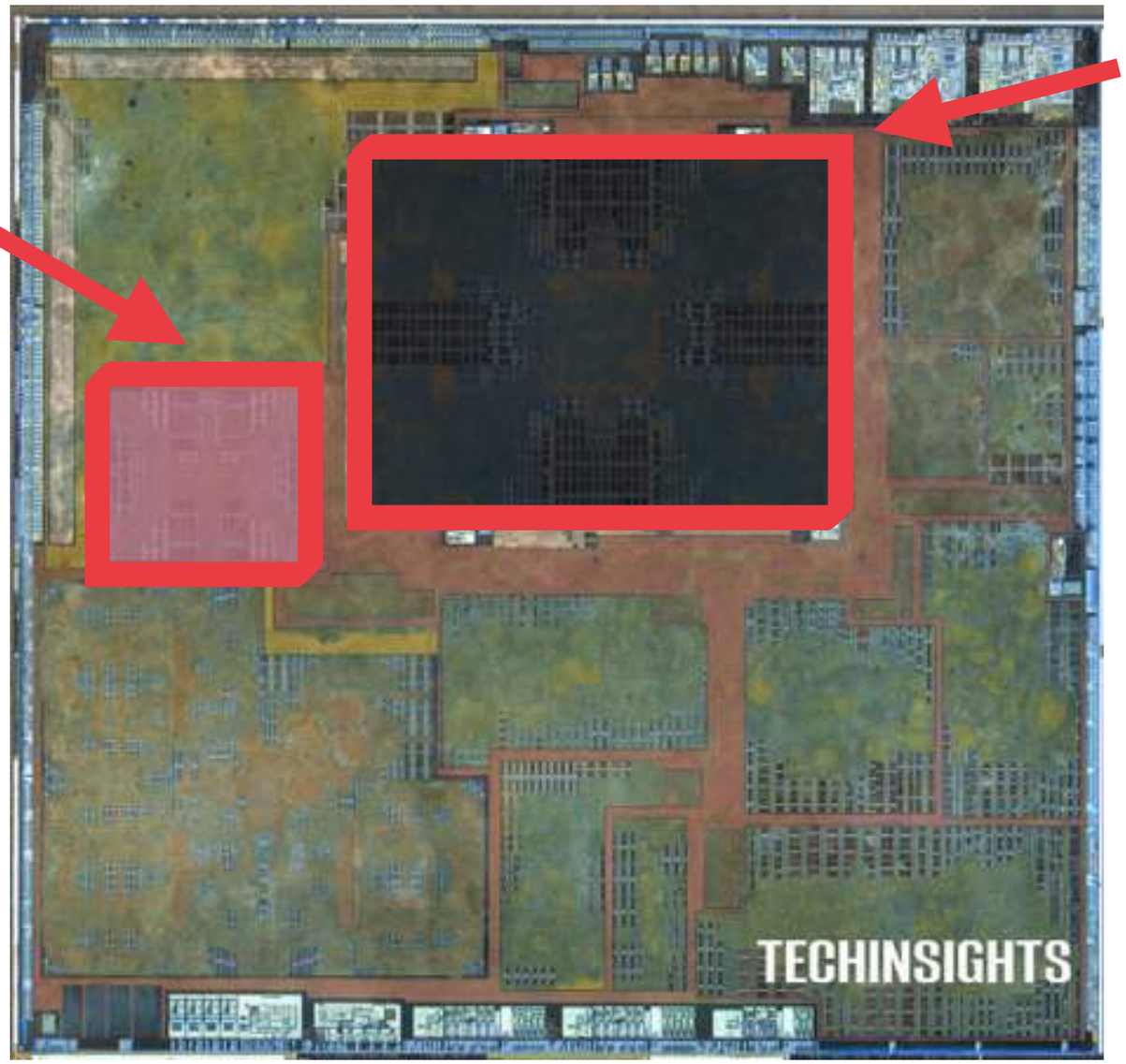
Challenge: mapping jobs to the cores that fit best

- OS has to rely on performance counters, history, …
- DBMS has more valuable knowledge about the work that will be executed

# ARM big.LITTLE (Exynos 5410)

## LITTLE

- 4 cores
- in-order
- max. 2 issues/ cycle
- 8–10 stage pipeline
- 32B cache lines
- 32kB L1 I/D
- 512kB L2 LLC
- 3.8mm² die area




die photo from http://www.techinsights.com/

## big

- 4 cores
- out of order
- max. 3 issues/ cycle
- 15–24 stage pipeline
- 64B cache lines
- 32kB L1 I/D
- 2MB L2 LLC
- 19mm² die area

Cache-Coherent Interconnect (currently only allows one active cluster)

2GB dual-channel LPDDR3 main memory (12.8GB/s peak transfer rate)

# Contributions

1) Analysis of parallel query execution on a single-ISA heterogeneous system (ARM big.LITTLE system)

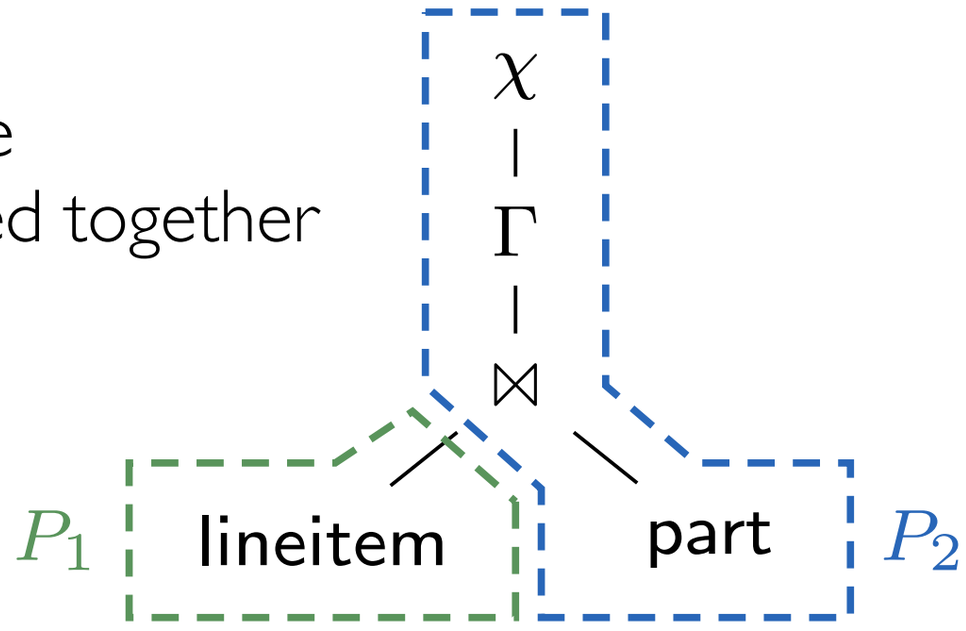
2) Analysis of parallel database operators on the LITTLE and big cluster

3) A heterogeneity-conscious DMBS-controlled job-to-core mapping approach for parallel query execution

# Parallel Query Execution in HyPer

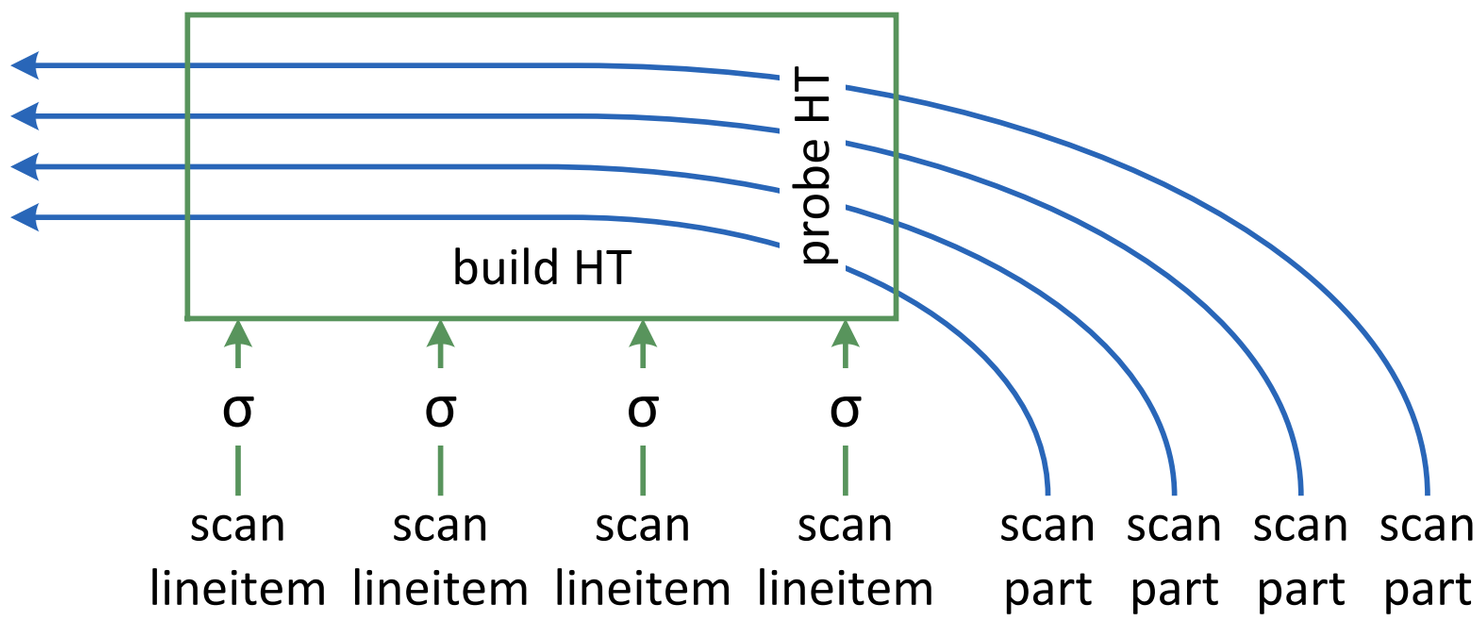
## Data-centric code generation

- operators that do not require intermediate materialization are interleaved and compiled together into pipelines
- tight work loops: keep data in registers

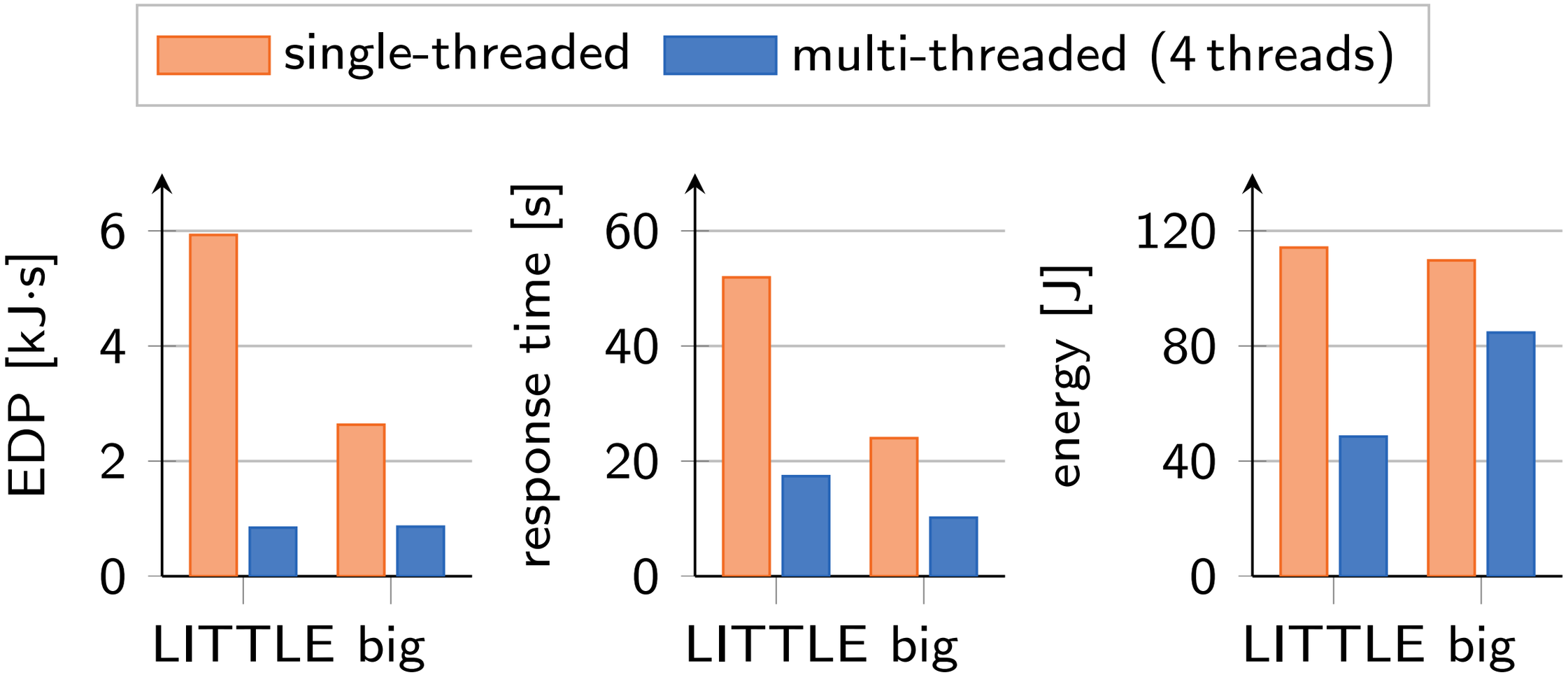



## Morsel-driven parallelism

- execute a pipeline in parallel
- morsels are fragments of input tuples
- all operators are parallelized (lock-free)
- elastic
- NUMA-aware



# Initial Results: HyPer on big.LITTLE



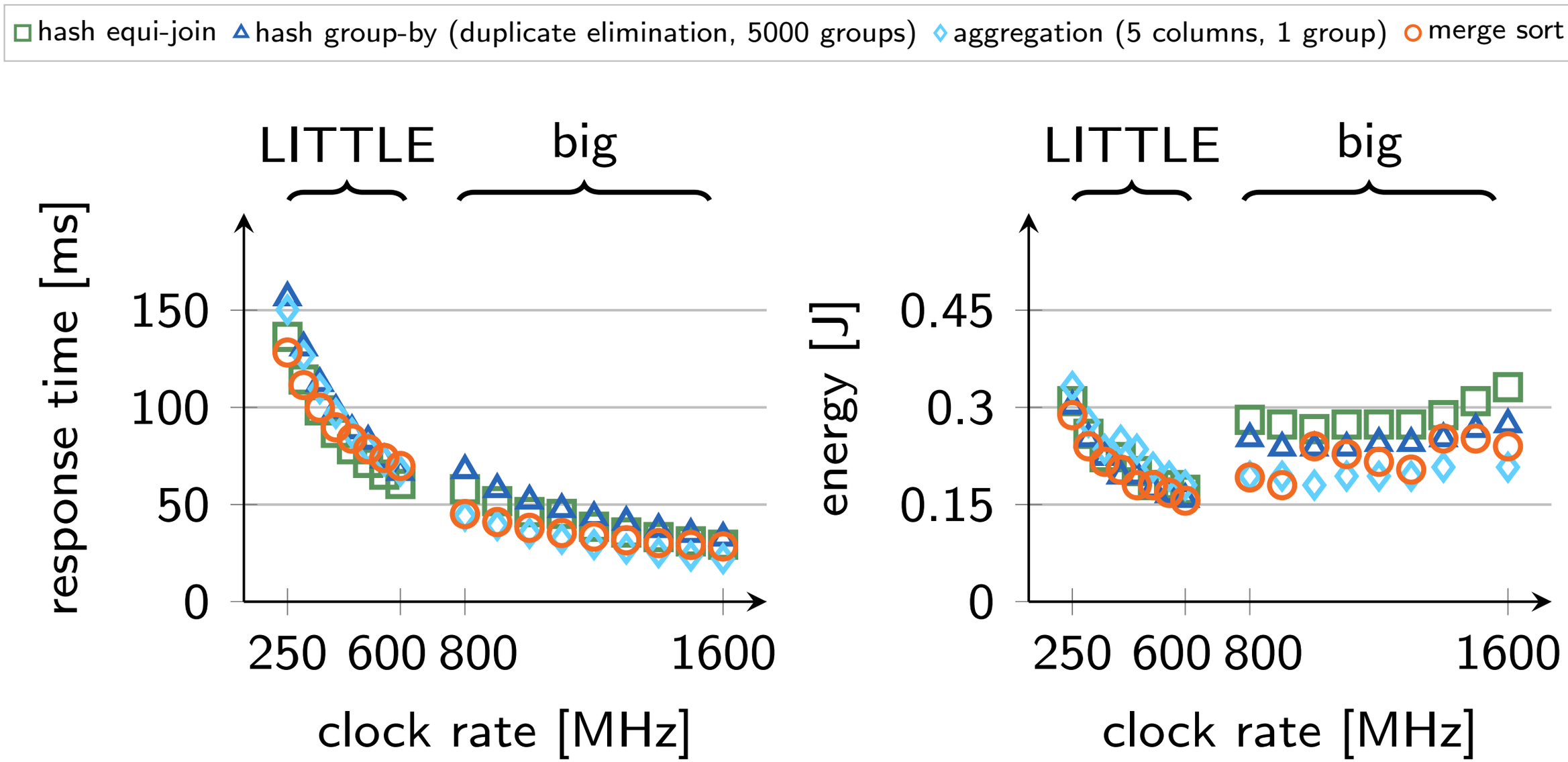

TPC-H scale factor 2, running all 22 queries, **OS ondemand**

**Energy Delay Product EDP = energy consumed × response time**

Single-threaded execution: LITTLE core has worse EDP than big core

Multi-threaded execution: LITTLE cluster and big cluster have equal EDP

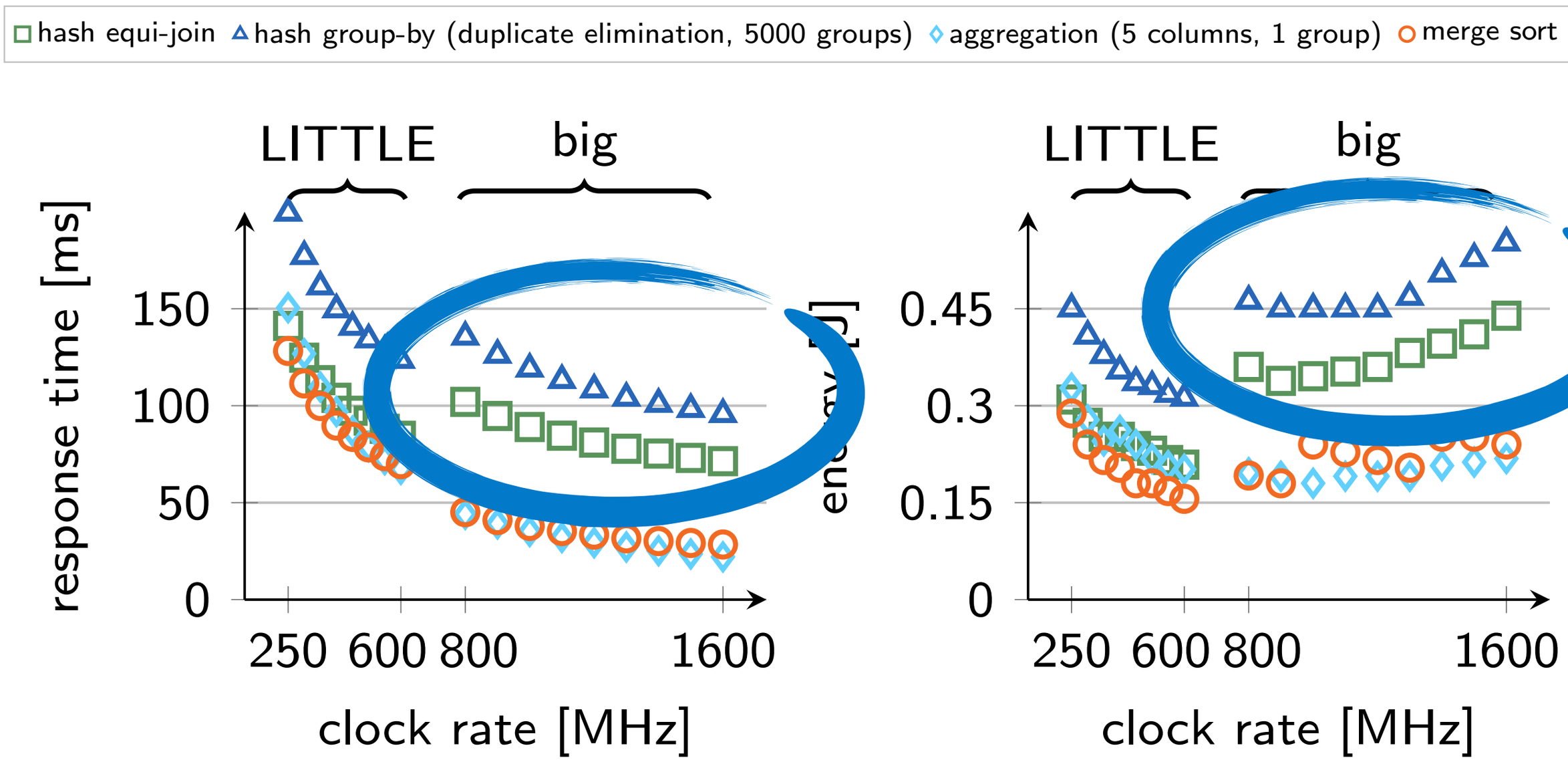
# Core Database Operator Analysis (i)




Working set in LLC

4-way parallel processing of operators

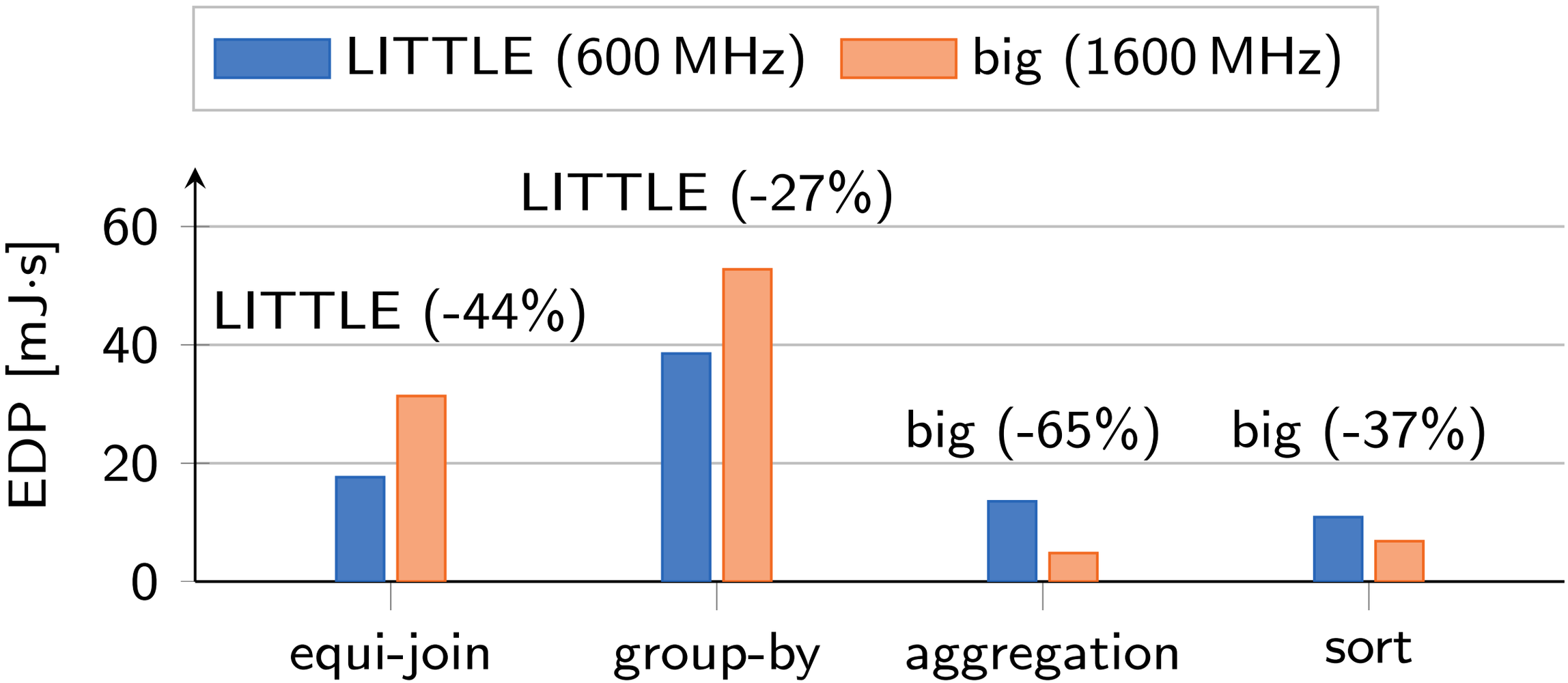
# Core Database Operator Analysis (ii)




Working set exceeding LLC

4-way parallel processing of operators

# Core Database Operator Analysis (iii)




Working set exceeding LLC

4-way parallel processing of operators

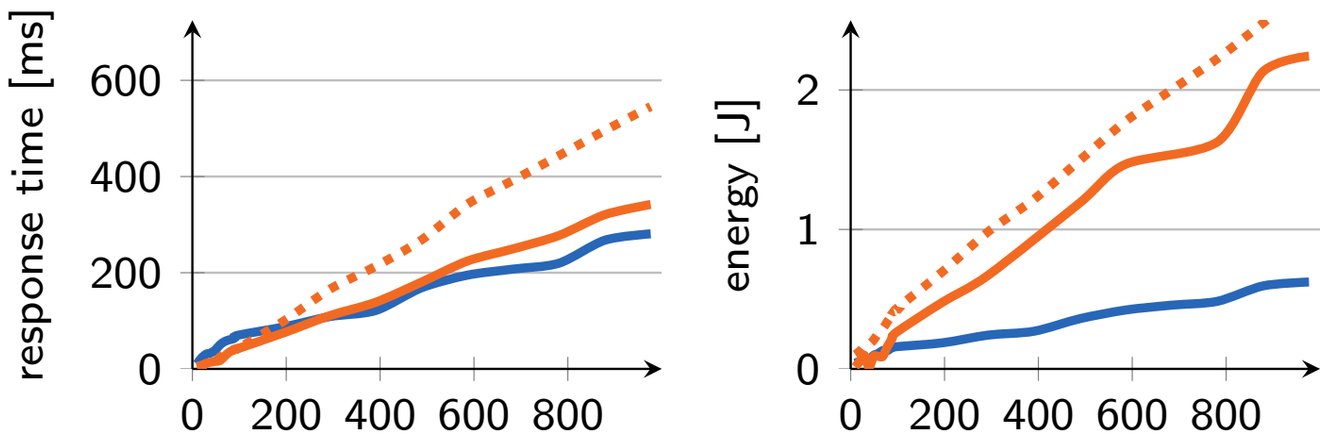
**Even with varying implementations, trend stays the same**

# Parallel Hash Equi-Join




## Join build

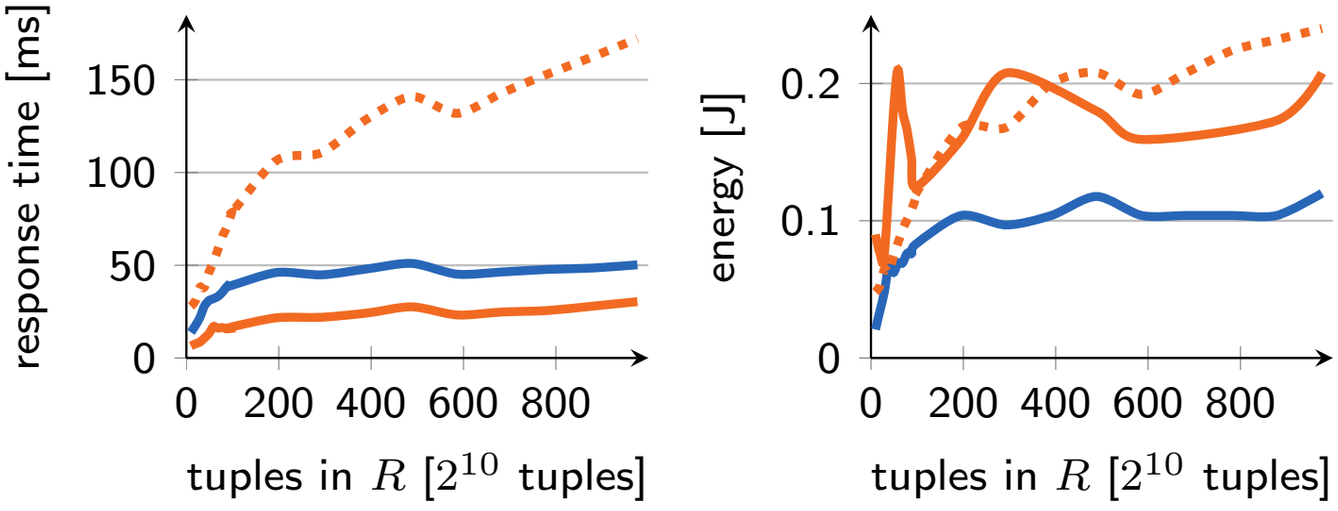
- LITTLE cluster has better EDP
- atomic CAS to build hash table has worse performance on big cluster




join build

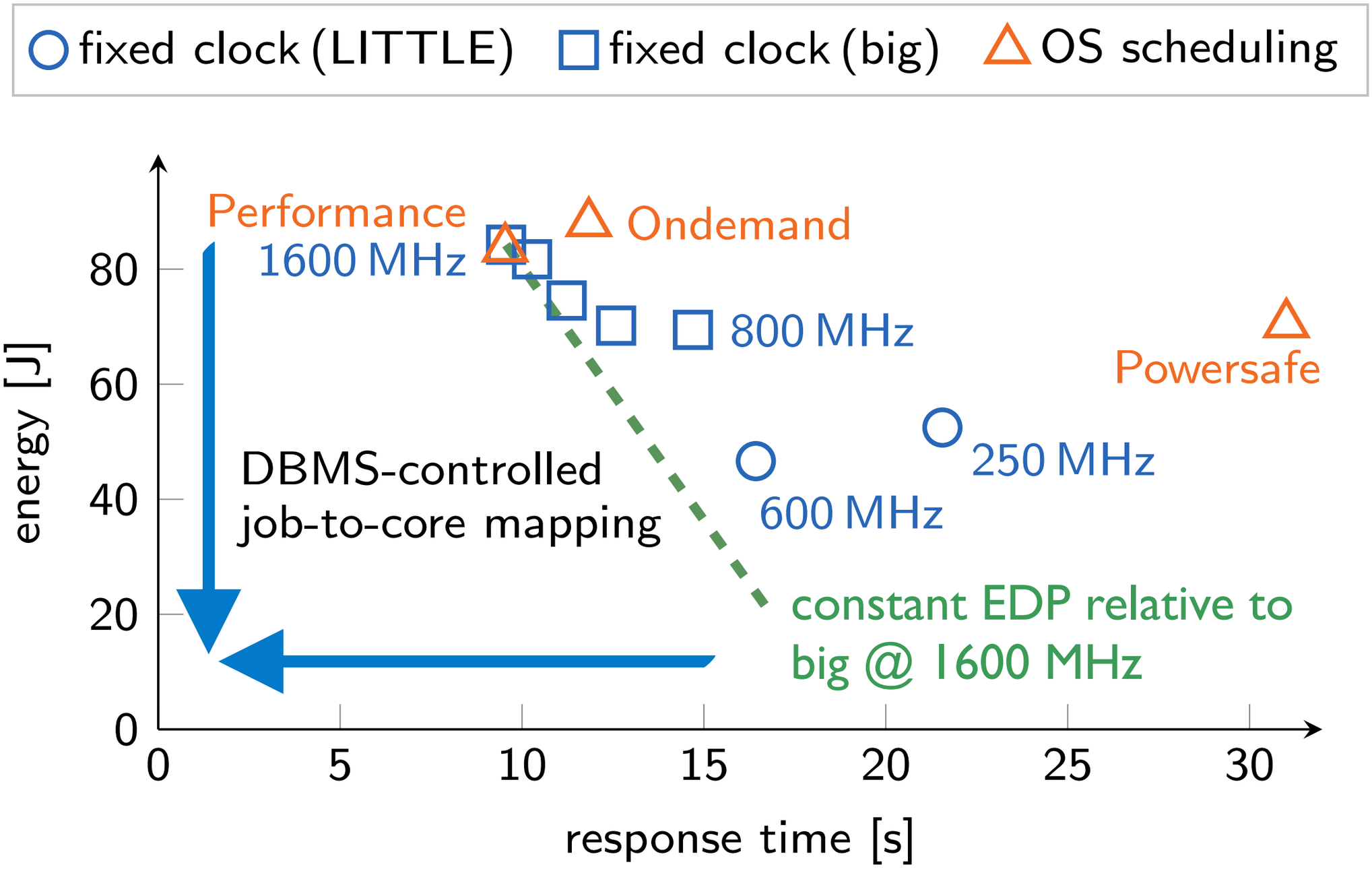
## Join probe

- big and LITTLE cluster have almost equal EDP
- pointer chasing vs hash table in LLC

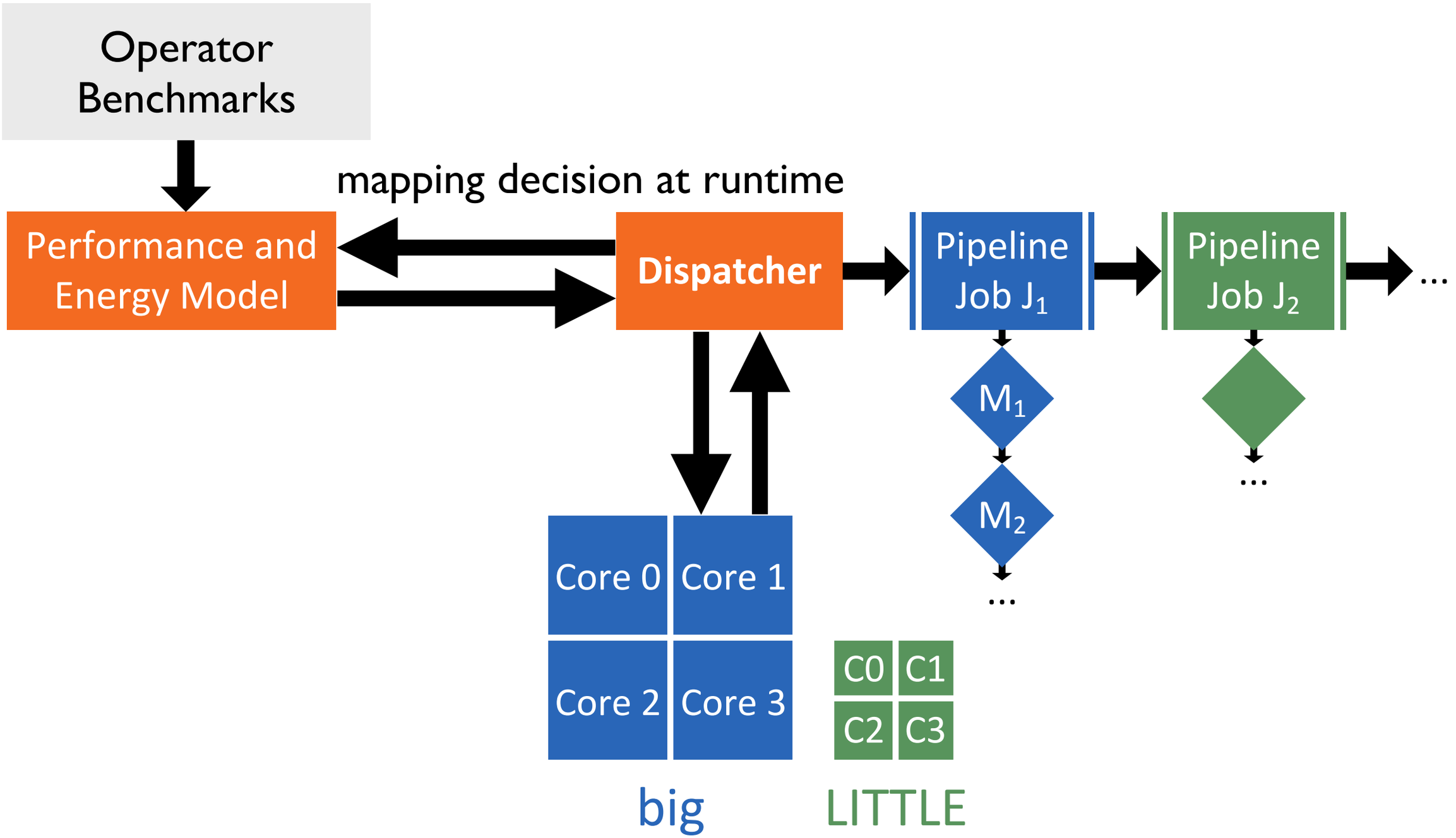



join probe

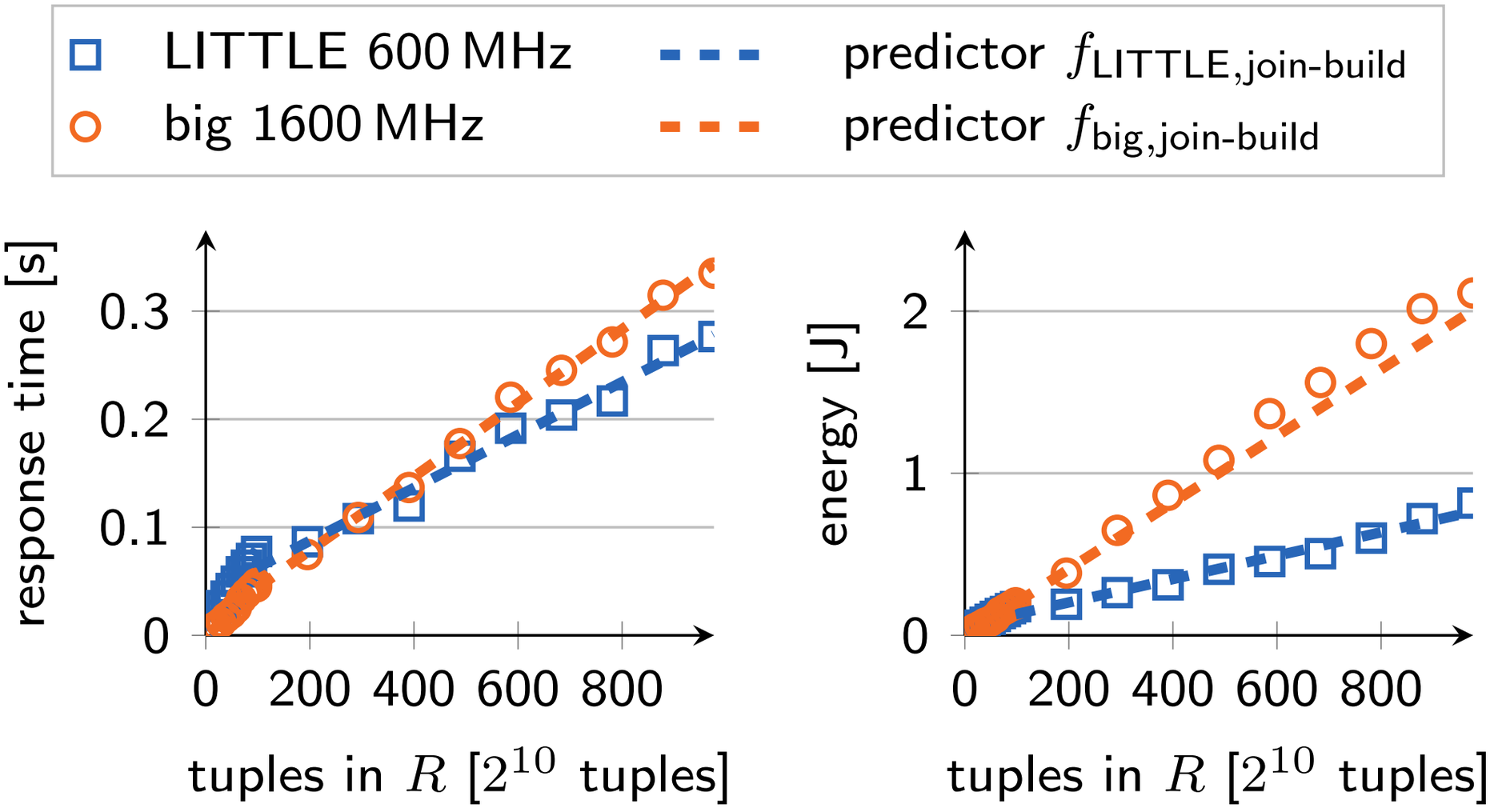
# Getting a better mileage …
# … while driving faster!



# Heterogeneity-Conscious Dispatching



# Performance Energy Model (PEM)

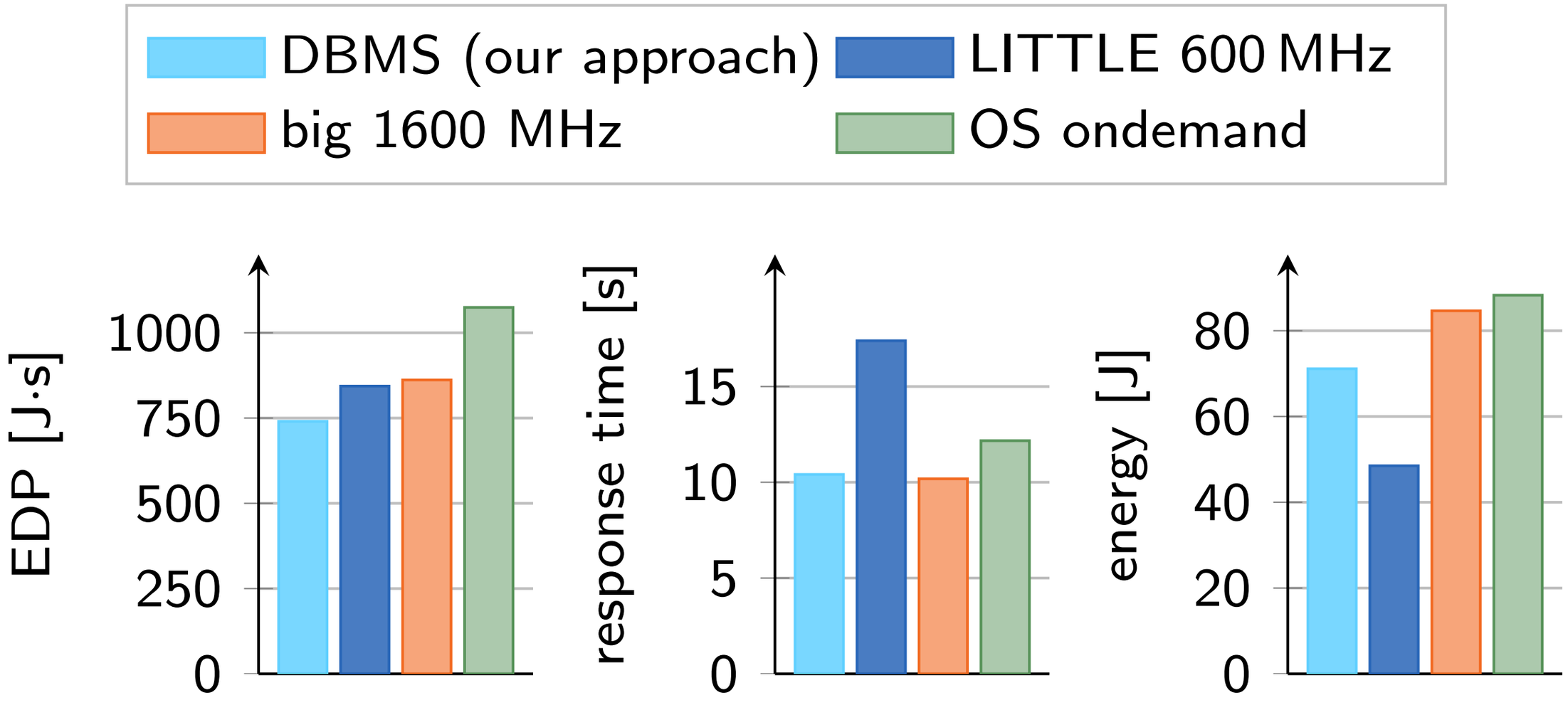



join build PEM

2 segments: working set in LLC and exceeding LLC

linear regression models based on benchmarks, data from query processing

# Evaluation: TPC-H Scale Factor 2




31%/12%/14% improved EDP over OS ondemand/LITTLE/big

**Some queries faster and more energy efficient (e.g., query 14)**

Query 14: 63%/45% improved EDP over OS ondemand/big

# Conclusion

Heterogeneous single-ISA multi-core processors are an interesting design space in light of dark silicon and Amdahl's law:

In our experiments ***4 LITTLE cores constantly outperform 1 big core*** (but occupy the same die area)

These processors are ***no free lunch for parallel database systems***: job-to-core mapping is challenging and is better performed by the DBMS rather than the OS (more knowledge)

A heterogeneity-conscious ***DBMS-controlled job-to-core mapping*** approach for parallel query execution achieves a higher performance while using less energy ⇒ ***We get a better mileage while driving faster!***