ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Ecole Polytechnique Fédérale de Lausanne
EPFL

EPFL History

1853	1946	1968 – Act on the Federal Institutes of Technology	1969	Today

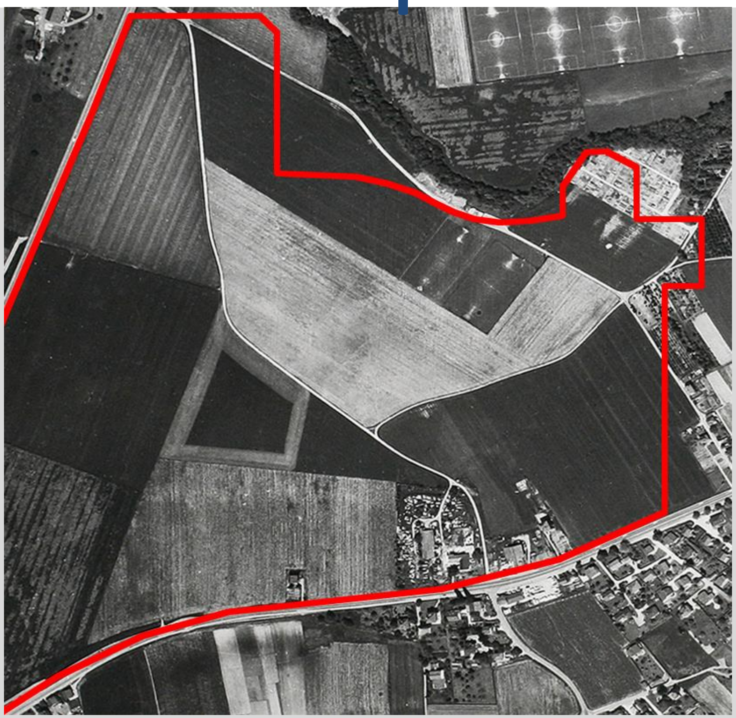
Maison Bischoff, rue Saint-Pierre in Lausanne

Former Hotel Savoy, 29-33 avenue de Cour in Lausanne

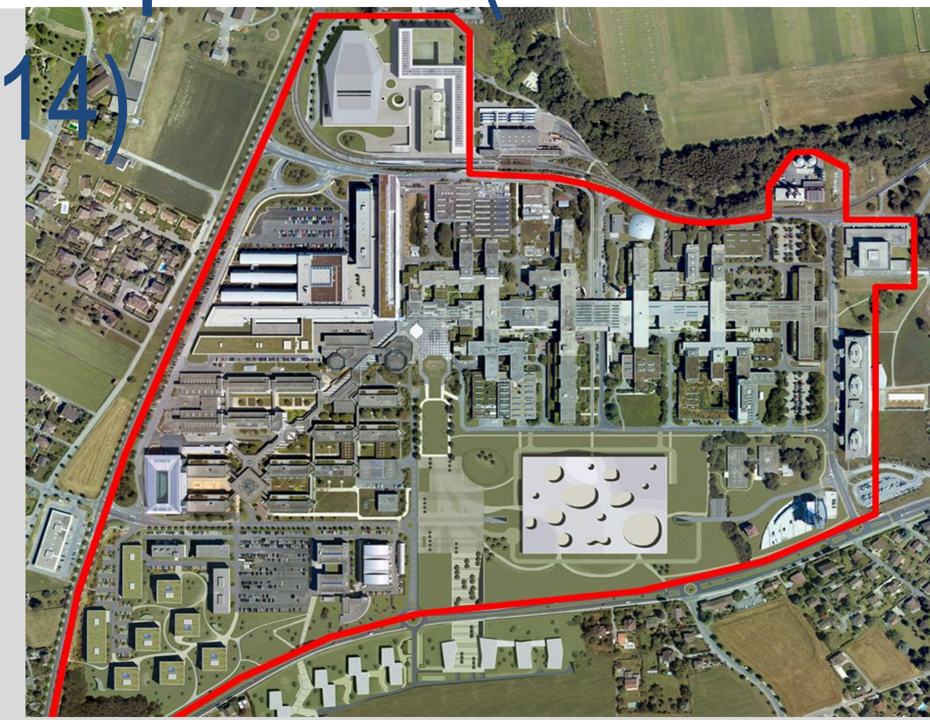
Ecublens-Dorigny

	Ecole spéciale de Lausanne	EPUL	EPFL	
Students	11	399	1,293	9,868 (2013)
Faculty			56	329 (2013)

Fast-paced development (1969-2014)



1969 …



2014 Premises > 54 hectares > approx. 561,800 sq.m

EPFL at the heart of national endeavours

ETH Domain = 2 Federal Institutes of Technology & 4 Research Institutes (2013)

Students in the ETH Domain

Bachelor Students 13,622
ETH Zurich 8,444
EPFL 5,178

Master Students 7,208
ETH Zurich 4,778
EPFL 2,430

PhD Students 5,947
ETH Zurich 3,889
EPFL 2,058

Total 28,055 ETH Zurich 18,187 EPFL 9,868

CHF 2.27 B

EPFL
Lausanne

ETH
Zurich

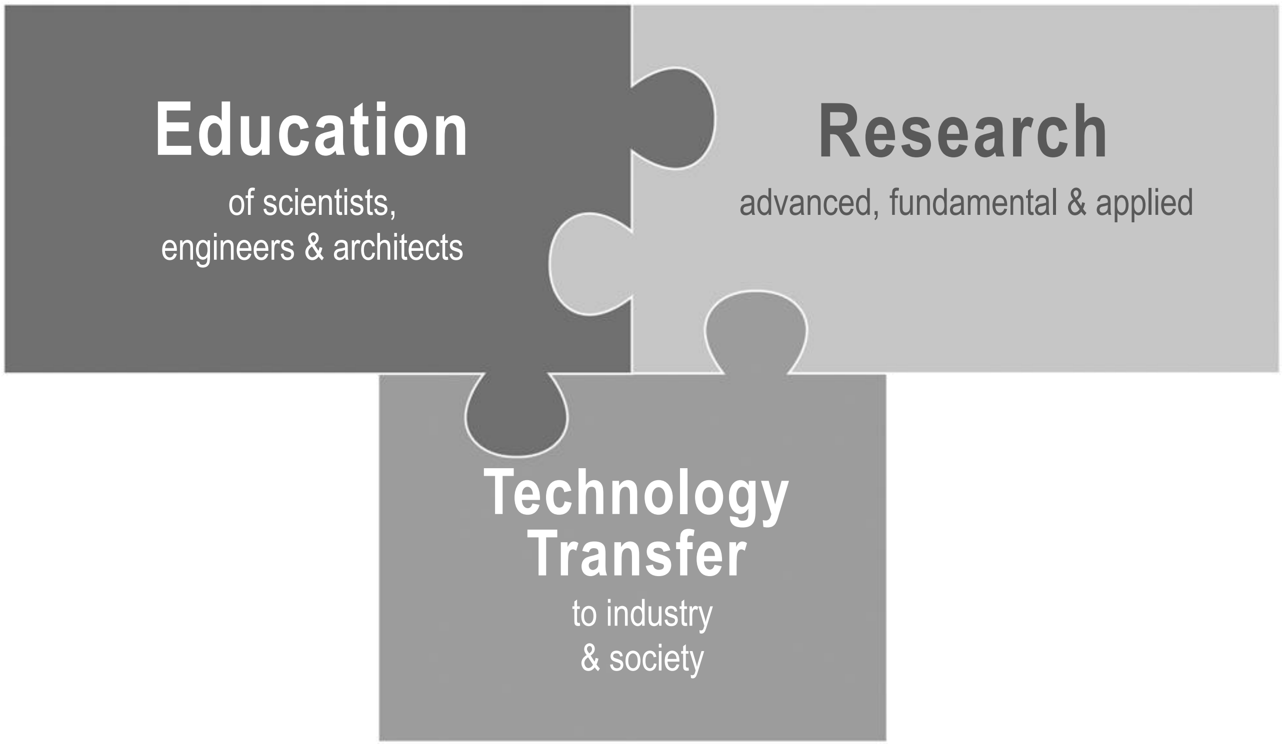
PSI
Paul Scherrer Institute

EMPA
Materials

WSL
Forest, Snow & Landscape

EAWAG
Water management, treatment & protection

EPFL's three missions according to the Federal Act



13 Study Programmes, 316 Research Labs

5 Schools
13 Sections

SB
Mathematics
Chemistry
Physics

IC
Computer Science
Communication Science

ENAC
Architecture, Civil Engineering, Environmental Science & Engineering

STI
Electrical Engineering
Microengineering
Mechanical Engineering
Materials Science

SV
Life Sciences & Technology
Bioiengineering

2 Colleges

CDM
Management of Technology
Financial Engineering

CDH
Human & Social Sciences

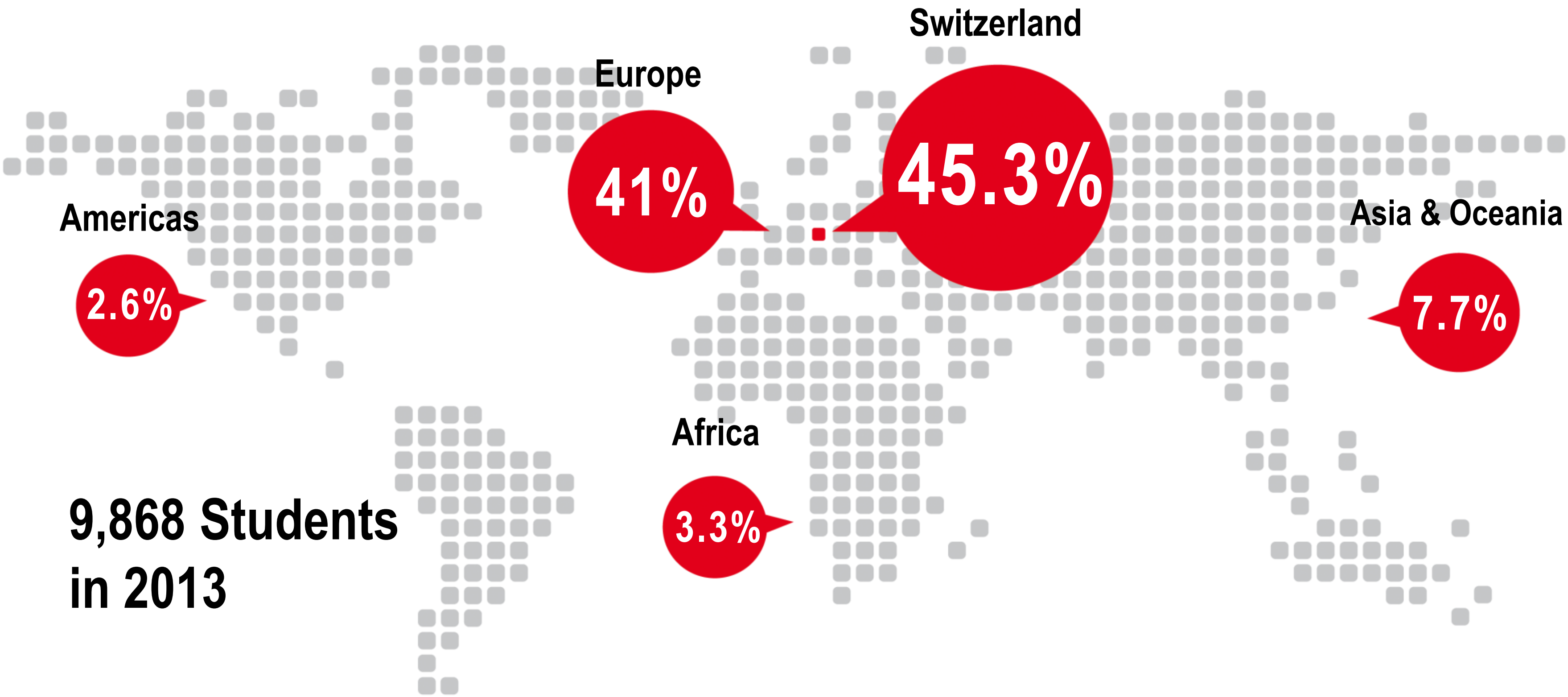
7 Interdisciplinary Centers

Transport, Energy, Neuroprosthetics, Design (EPFL+ECAL Lab) etc...

26 Institutes

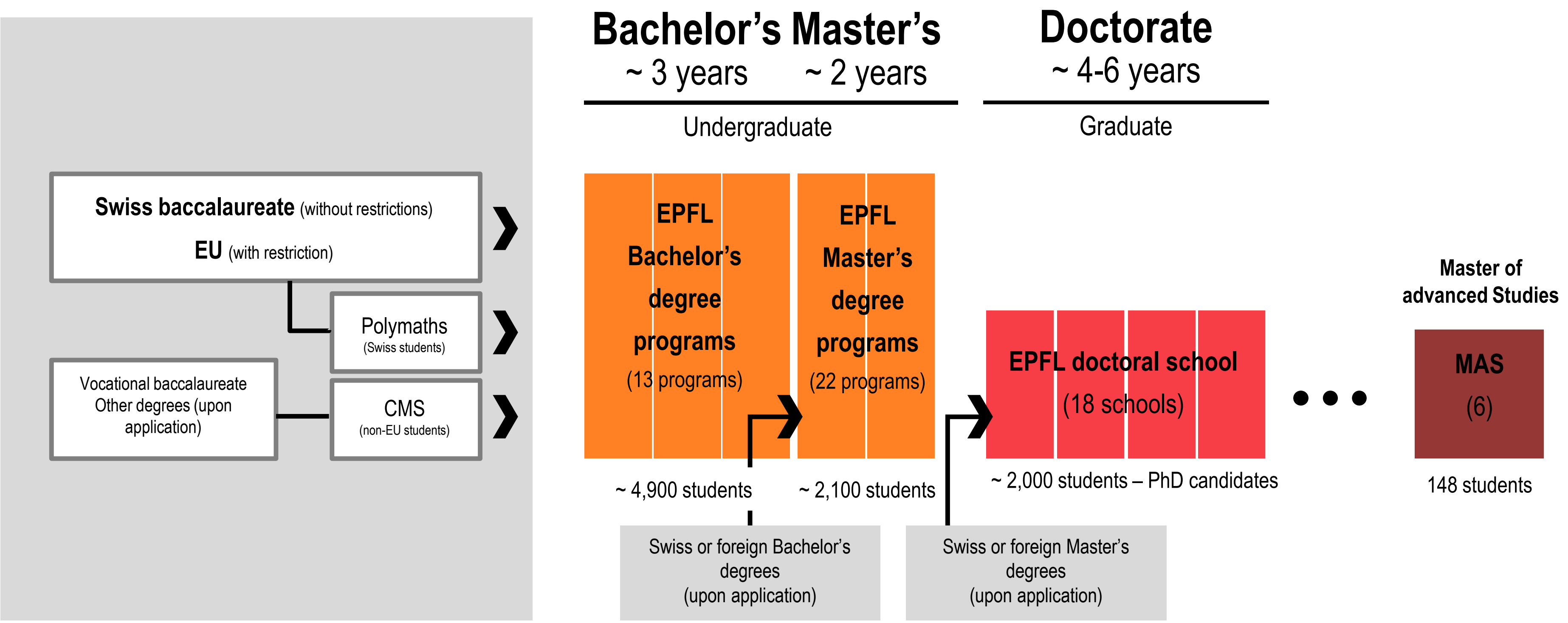
316 Laboratories and Research Groups

Origin of Students (Bachelor+Master+PhD)



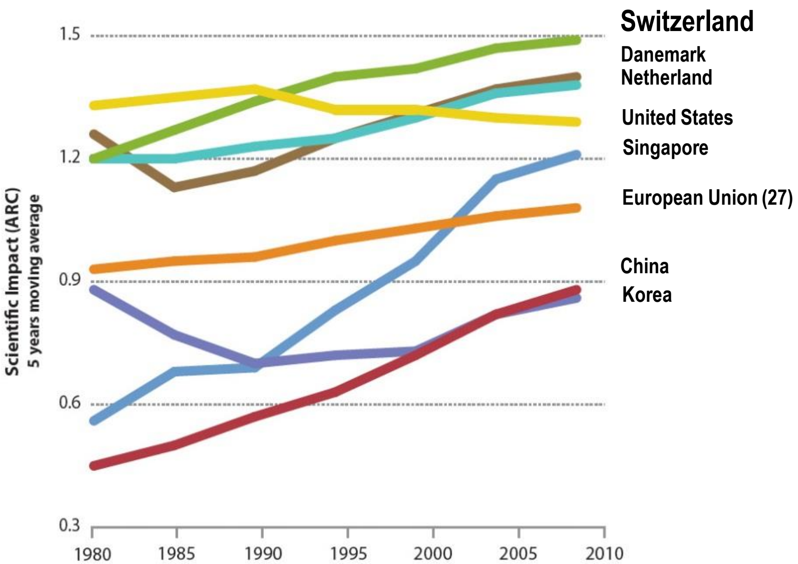
9,868 Students
in 2013

Study programs at EPFL



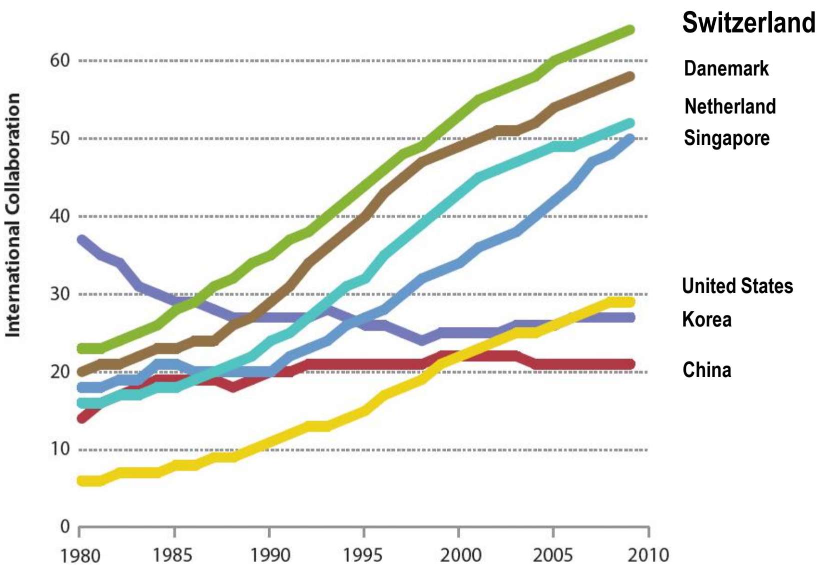
Very productive Swiss science

The world's highest science productivity since 1990 ... (OECD 2010)



Average of relative citations, figures by Thomson Reuters

…and one of the highest collaboration rates



Index of international collaboration, figures by Thomson Reuters

European Ranking 2013 – The EPFL is well-placed

(Eng. & Technology)

QS (Eng. & Technology)	THE (Eng. & Technology)	Shanghaï (Eng./Tech. & Computer Science)	Leiden (PP (top-10%) indicator)
1. University of Cambridge	1. University of Cambridge	1. University of Cambridge	**1. EPFL**
2. ETH Zurich	2. University of Oxford	**2. EPFL**	2. Weizmann Institute of Science
3. Imperial College, London	3. ETH Zurich	3. Imperial College, London	3. University of Cambridge
4. EPFL	4. Imperial College, London	4. University of Manchester	4. ETH Zurich
5. University of Oxford	**5. EPFL**	5. ETH Zurich	5. University of Oxford

Main Criteria

QS	THE	Shanghaï	Leiden
Reputation. Citations/paper. Student/faculty ratio.	Reputation. Citations/paper normalised by publication field, across all fields.	Most cited scientists. Number of papers published. % of papers published in top 20 science journals.	Citations/paper normalised by volume and publication field, across all fields. 300 largest European universities.

Global Ranking 2013 – U.S. on top, EPFL close behind…

(Eng. & Technology)

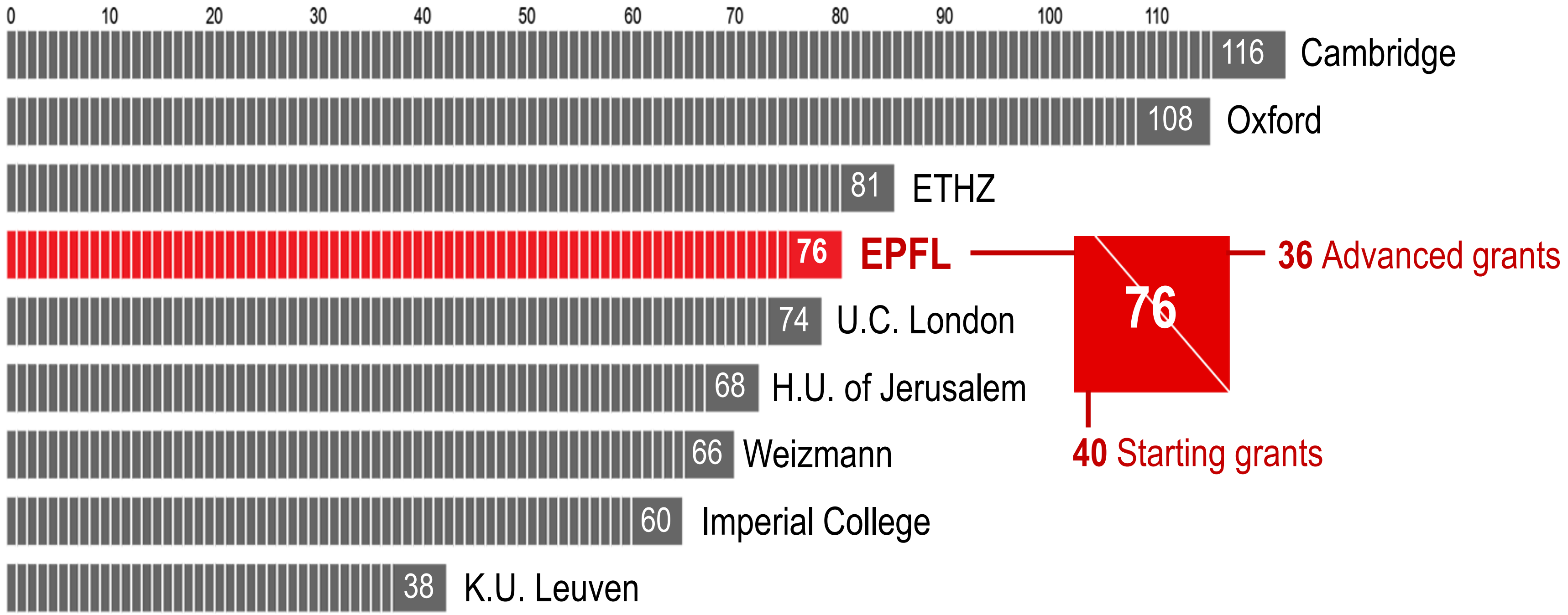
QS (Eng. & Technology)	THE (Eng. & Technology)	Shanghaï (Eng./Tech. & Computer Science)	Leiden (PP (top-10%) indicator – World)
1. MIT	**1. MIT**	**1. MIT**	**1. MIT**
2. Stanford University	2. Stanford University	2. Stanford University	2. Uni. of Calif, Santa Barbara
3. University of Cambridge	3. Uni. of California, Berkeley	3. Uni. of California, Berkeley	3. Stanford University
4. Uni. of California, Berkeley	4. California Institute of Technology	4. University of Texas	4. Princeton University
5. ETH Zurich	5. Princeton University	5. University of Illinois	5. Harvard University
8. EPFL	**15. EPFL**	**15. EPFL**	**13. EPFL**

Main Criteria

QS	THE	Shanghaï	Leiden
Reputation. Citations/paper. Student/faculty ratio.	Reputation. Citations/paper normalised by publication field, across all fields.	Most cited scientists. Number of papers published. % of papers published in top 20 science journals.	Citations/paper normalised by volume and publication field, across all fields. World's 500 largest universities.

European Recognition - ERC Grants (European Research Council)

EPFL 4th in Europe, cumulative data for 2007-2013



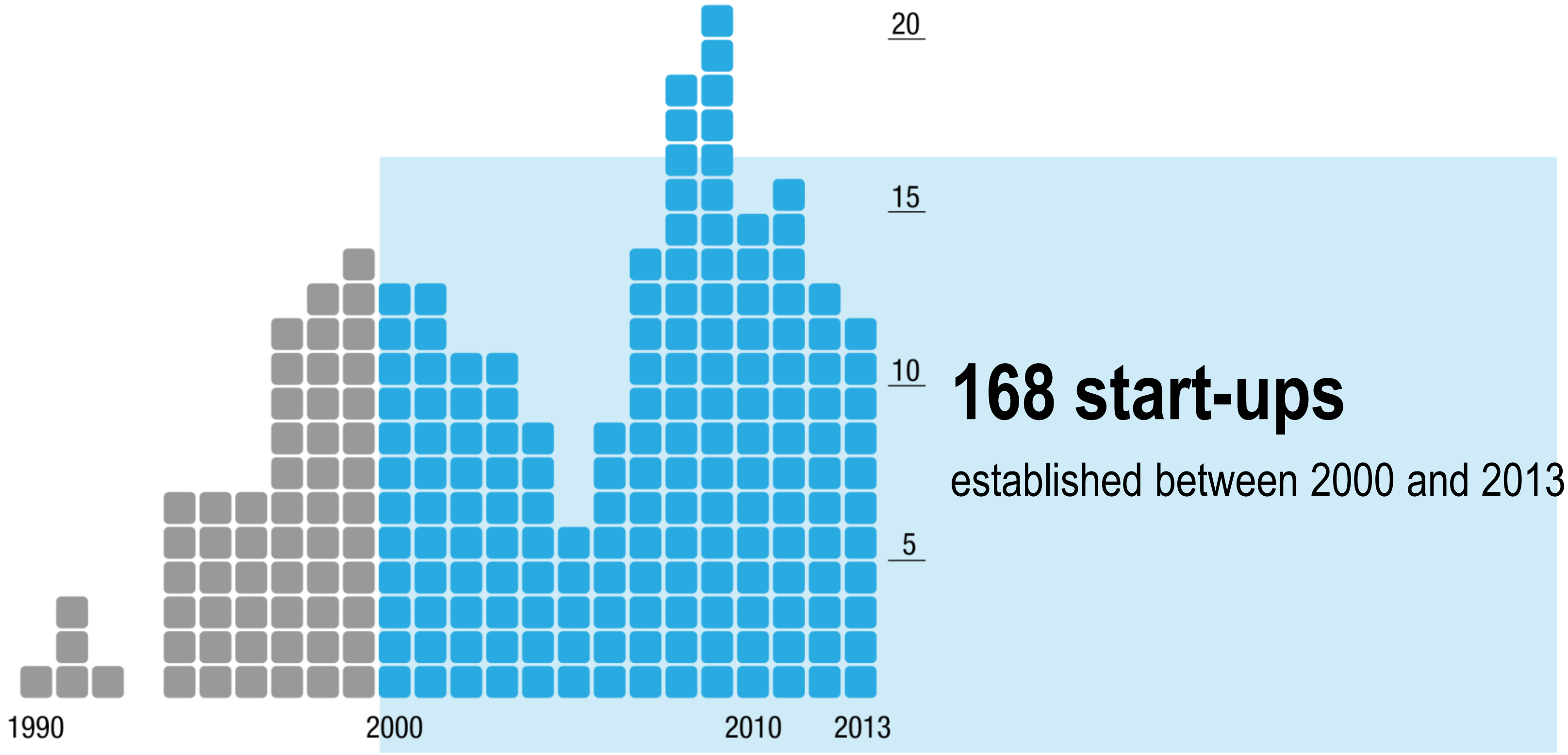
Great Technology Adventures

Solar Impulse

Alinghi (twice winner: 2003, 2007)

HYDROS: Hydroptère.ch

Start-Up development

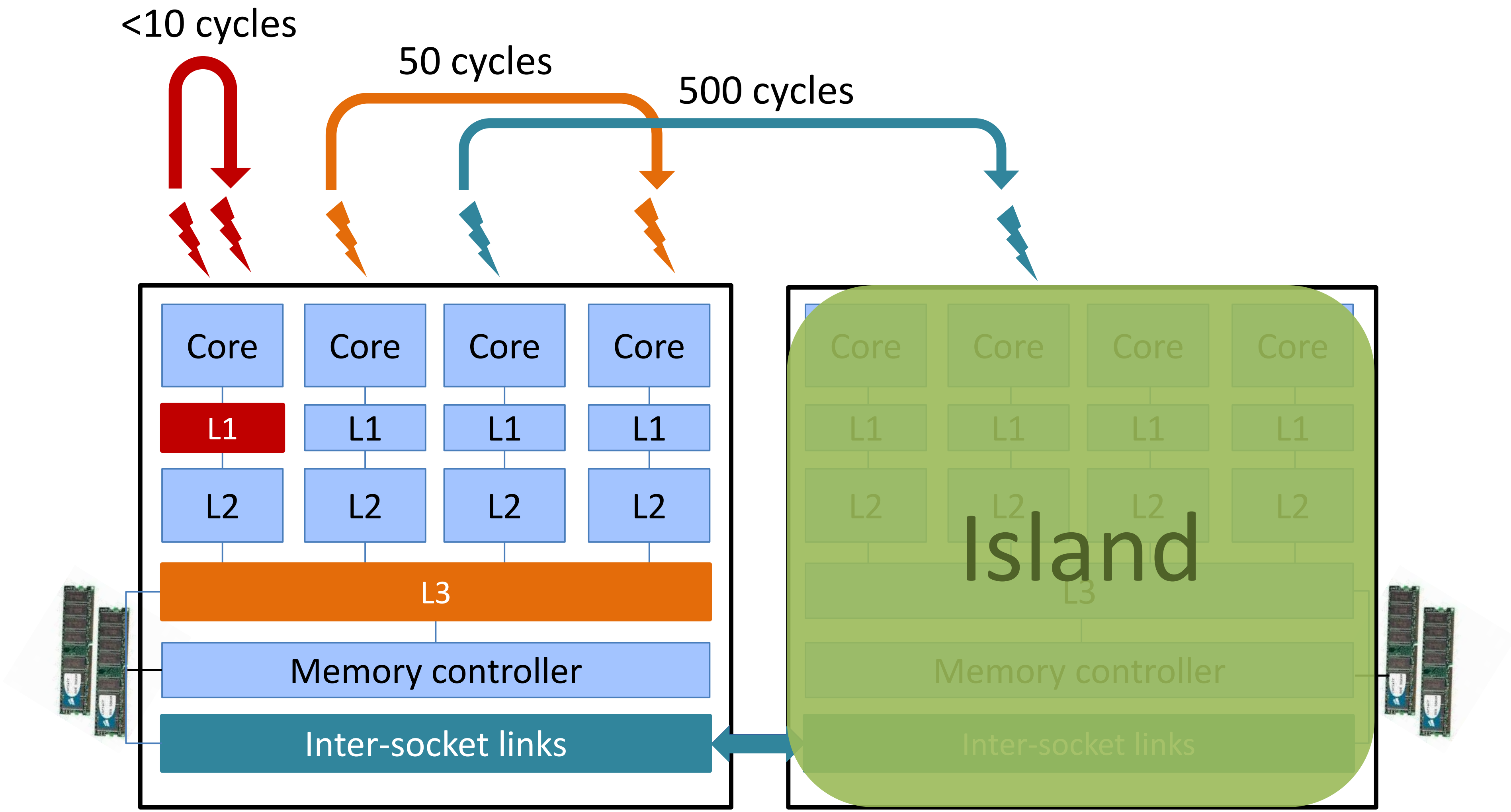


EPFL Campus

High-Performance Transaction Processing on Non-uniform Hardware Topologies

Danica Porobic
danica.porobic@epfl.ch
DIAS, EPFL

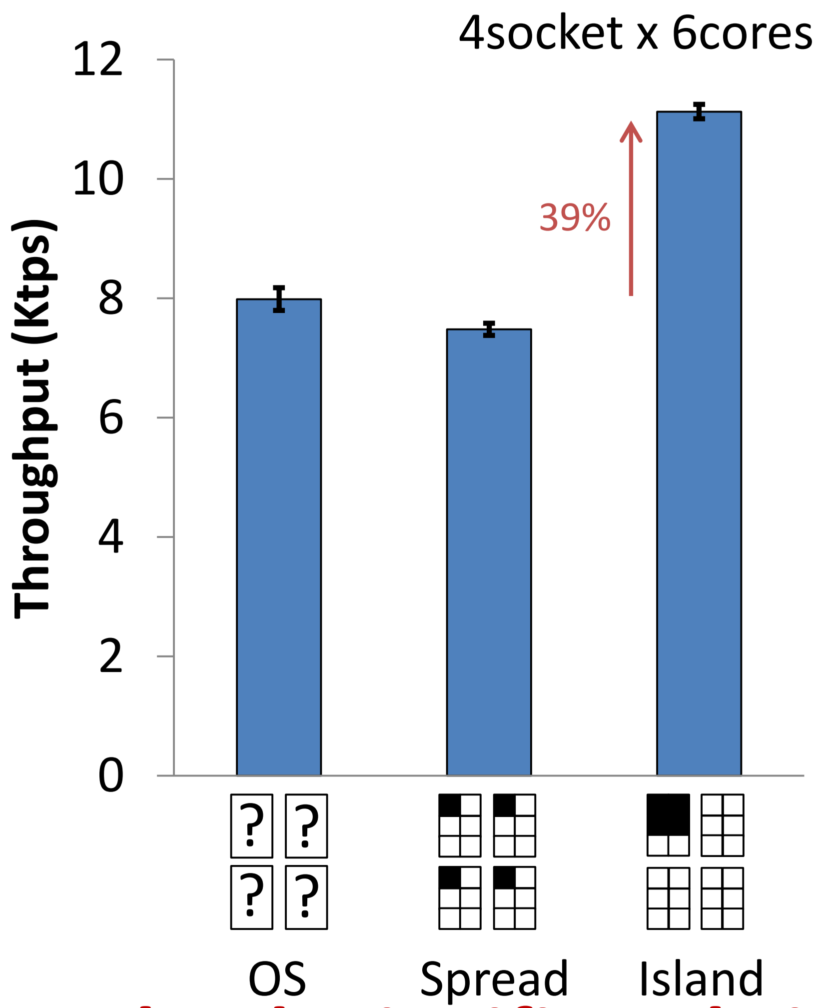
Multisocket multicores



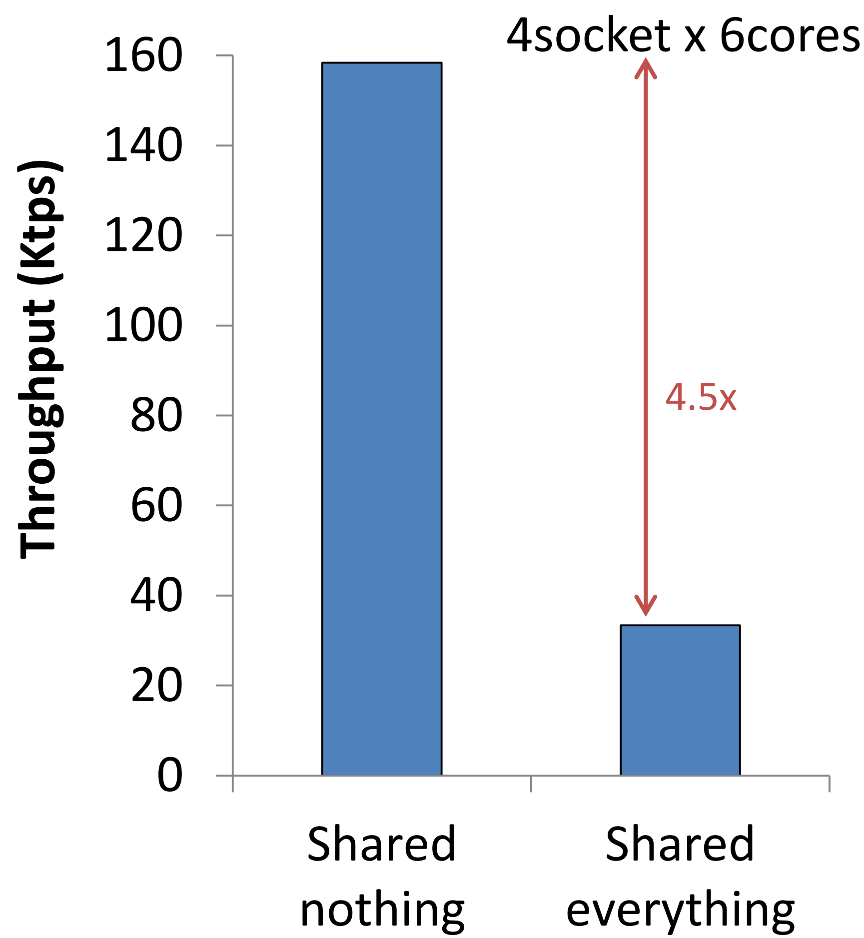
Communication latencies vary by an order-of-magnitude

Impact of Islands: TPC-C Payment

Thread Placement



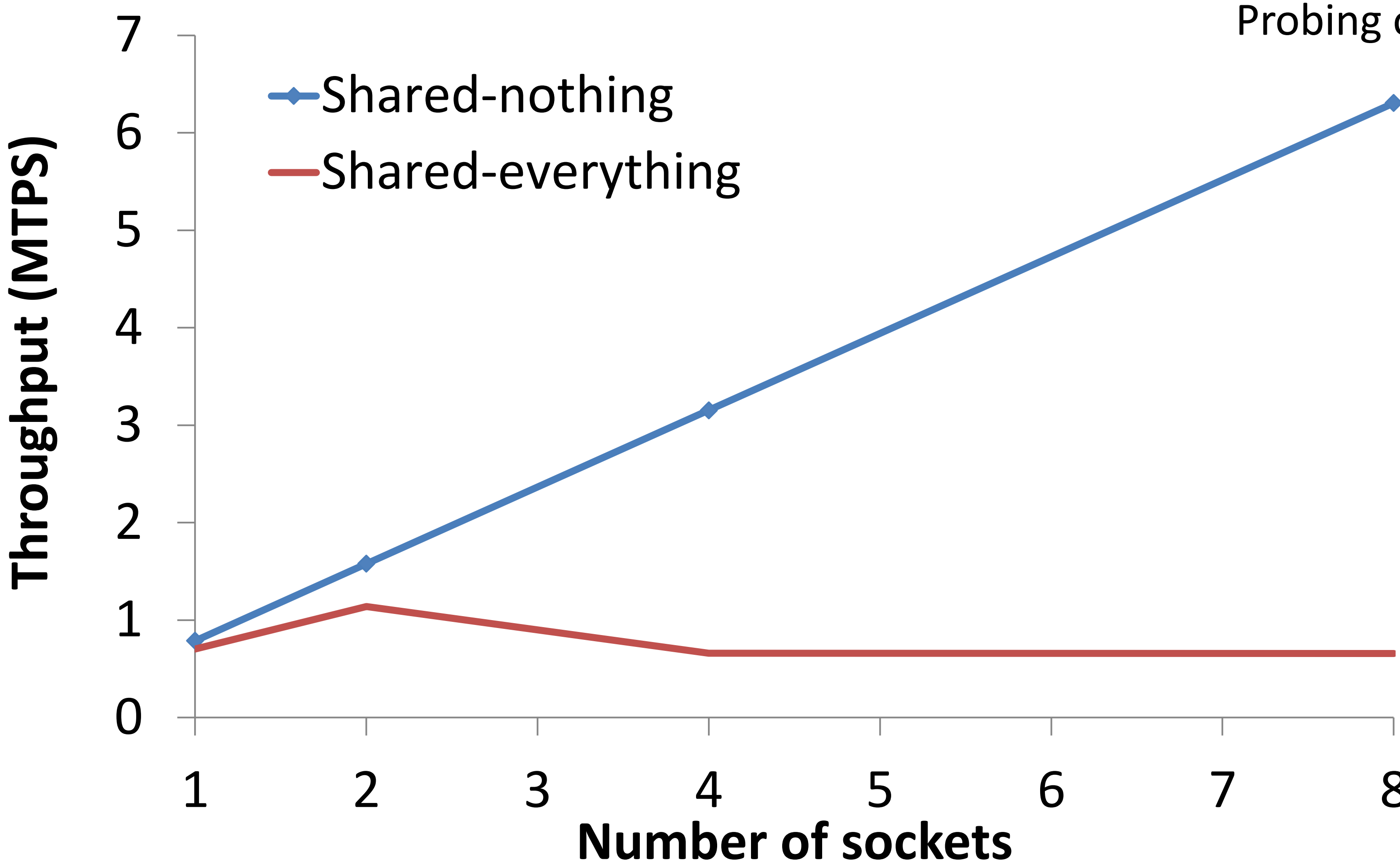
Communication granularity



Islands significantly impact OLTP applications

Scaling-up on a 8-socket machine

8socket x 10cores
Probing one row

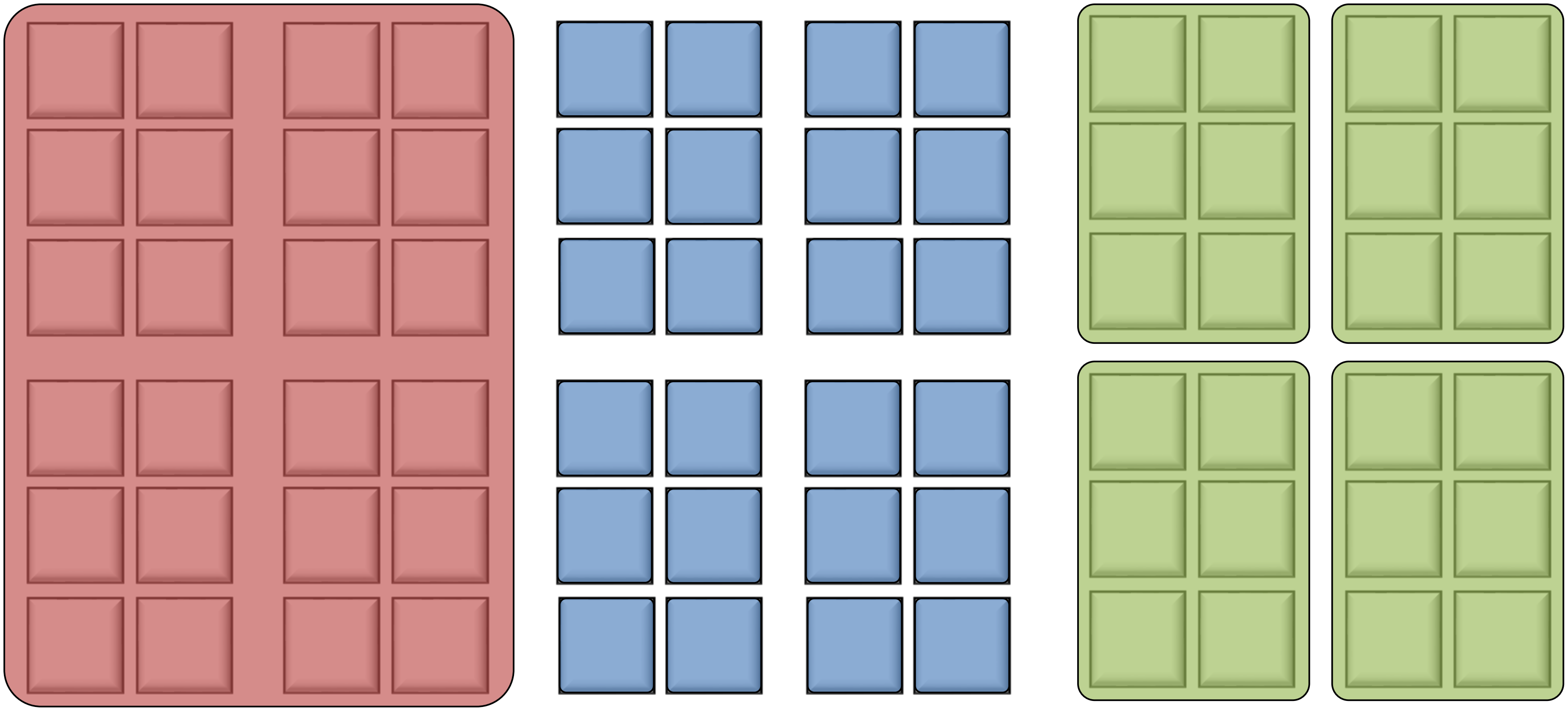


Islands significantly challenge scalability

Outline

- Introduction
- Impact of Hardware Islands on OLTP
- Adaptive Transaction Processing for Islands
- Conclusions

OLTP System Configurations



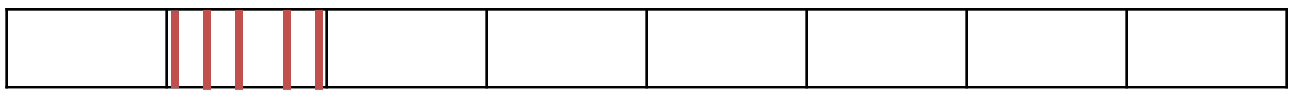
Shared-everything

Shared-nothing

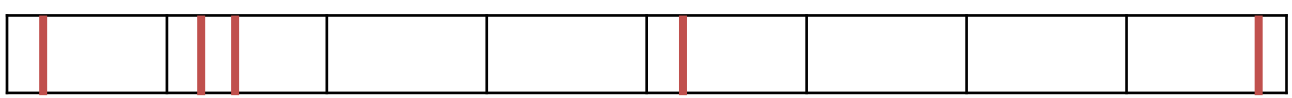
Island
shared-nothing

Partition sensitive microbenchmark

- Single-site version
  - Probe/update N rows from the local partition

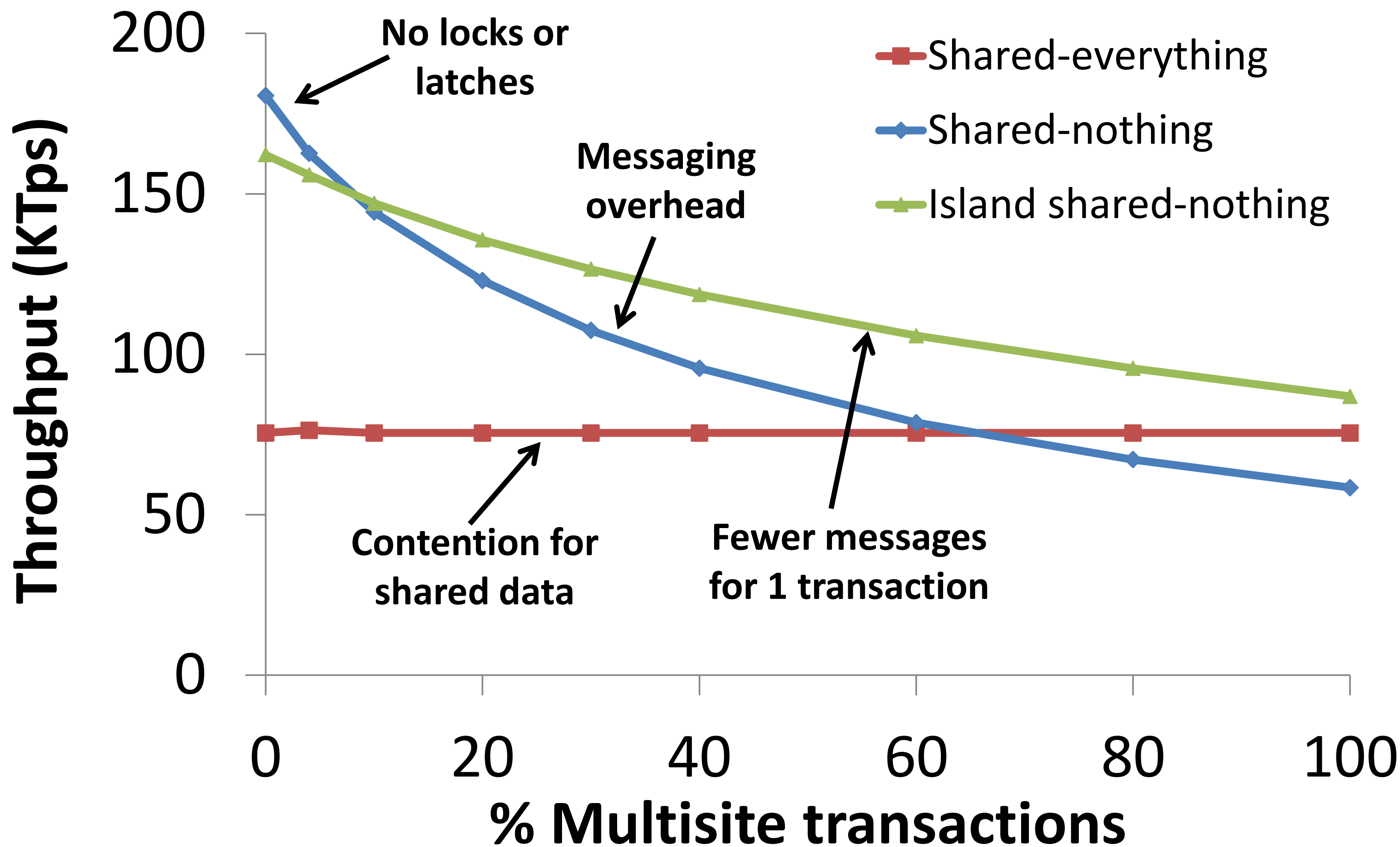


- Multi-site version
  - Probe/update 1 row from the local partition
  - Probe/update N-1 rows uniformly from any partition
  - Partitions may reside on the same instance



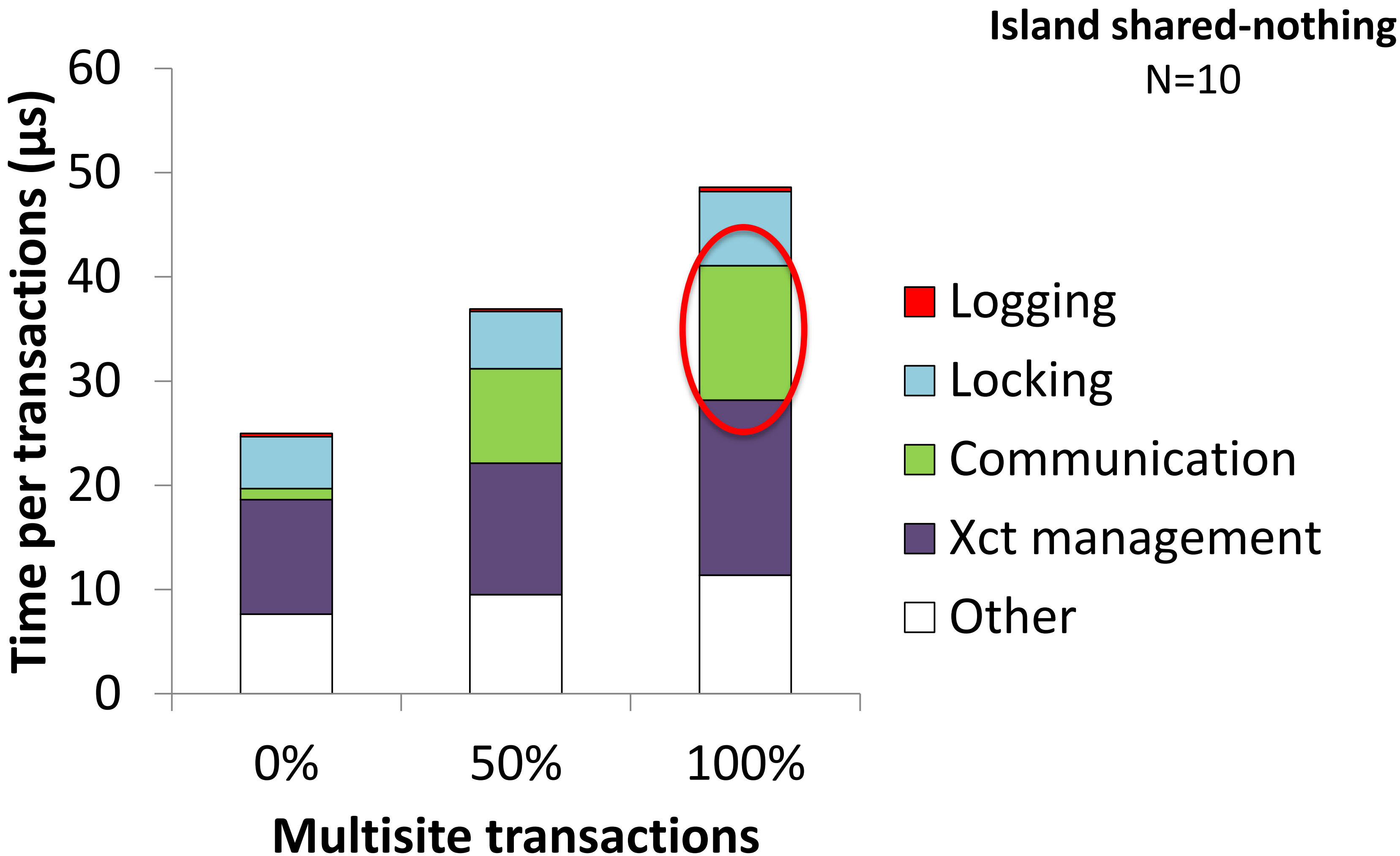
Multi-site transactions: reads

4socket x 6cores
N=10



More nodes -> faster performance drop

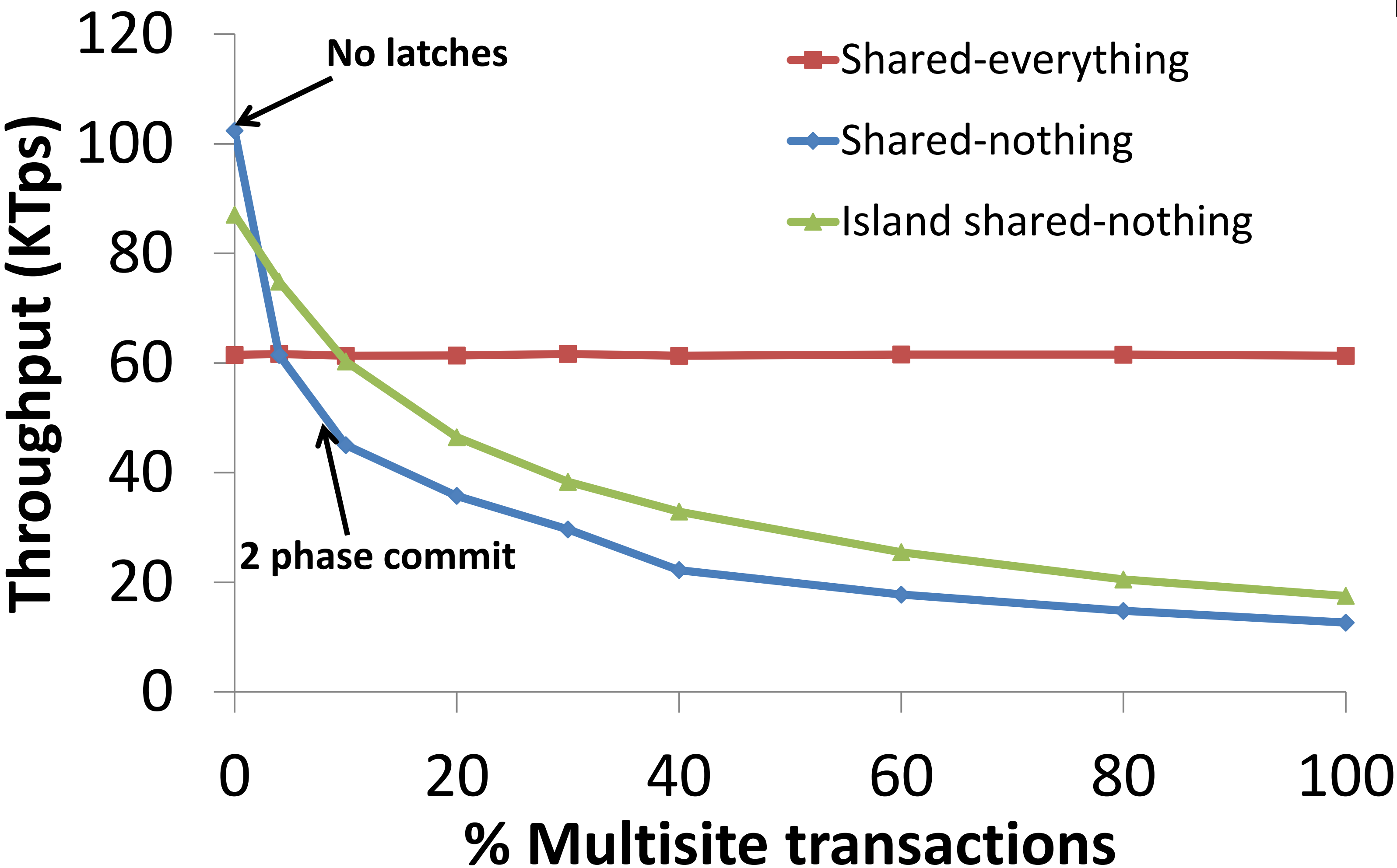
Where are the bottlenecks? Read case



Communication overhead dominates

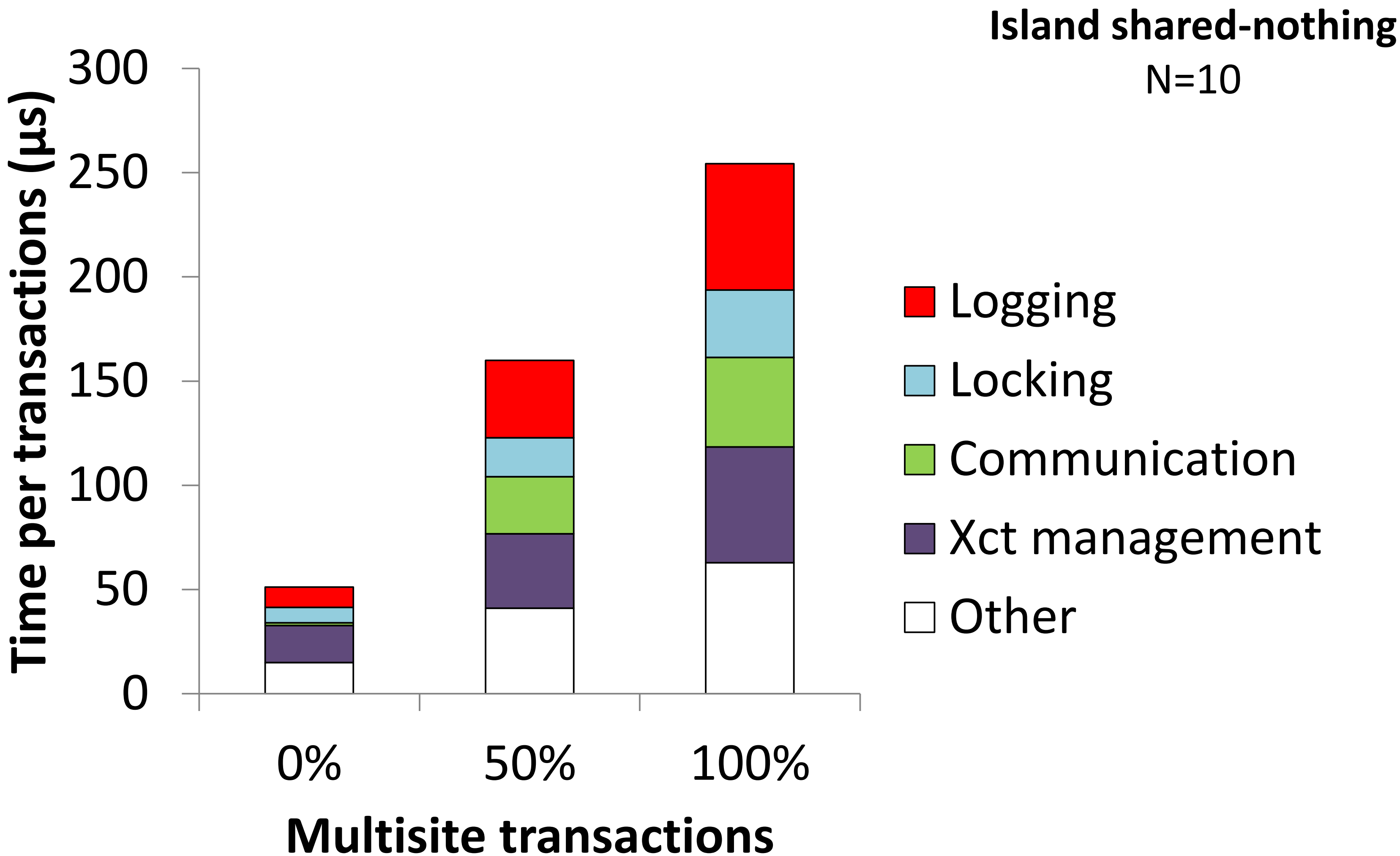
Multi-site transactions: updates

4socket x 6cores
N=10



Update distributed transactions are more expensive

Where are the bottlenecks? Update case



Several overheads contribute to the cost

OLTP on Hardware Islands

- Shared-everything: stable, but non-optimal
- Shared-nothing: fast, but sensitive to workload
- Island shared-nothing: a robust middle-ground

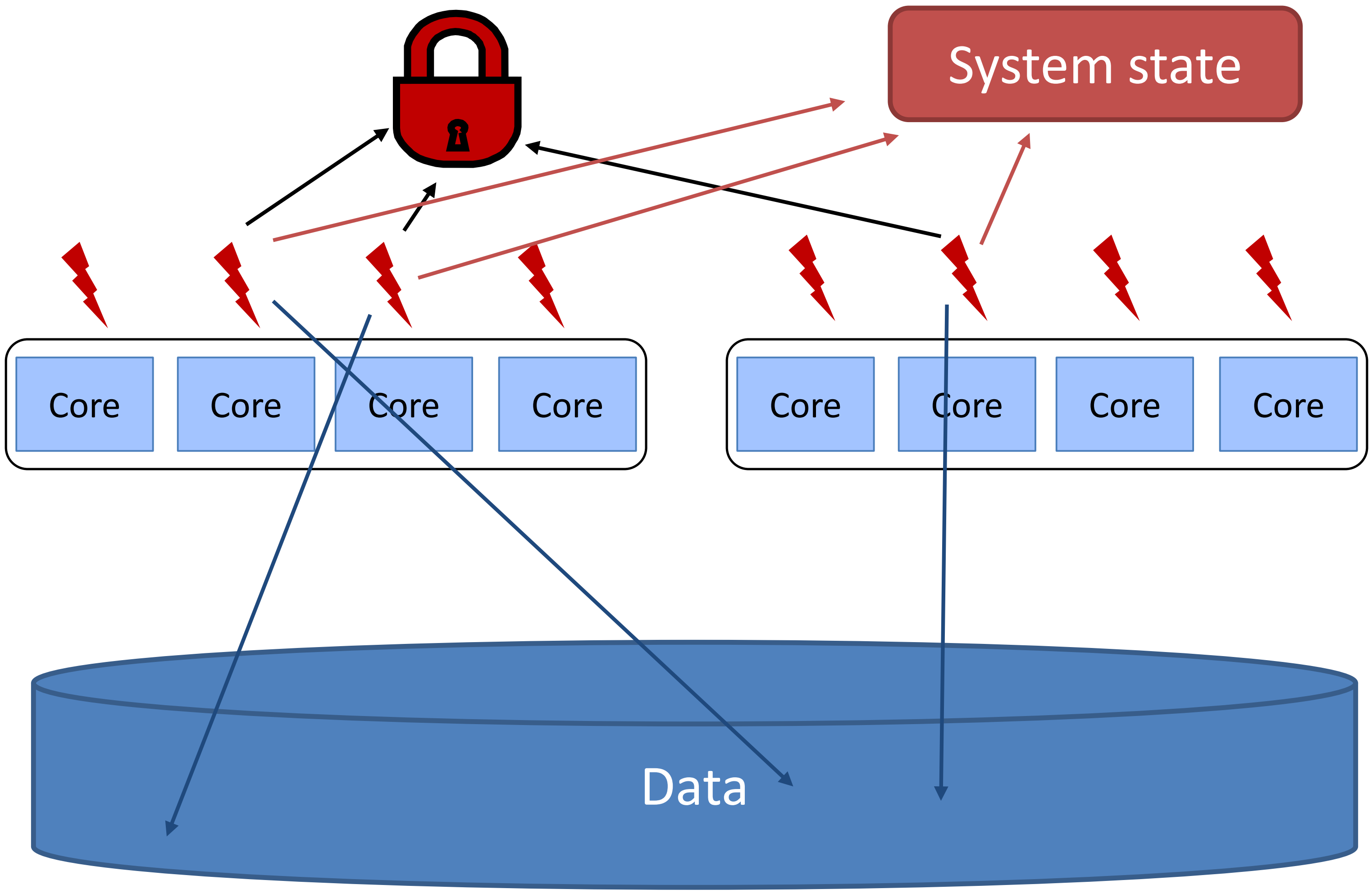
- Challenges:
  - Optimal configuration depends on workload and hardware
  - Expensive repartitioning due to physical data movement

Can we have one system that adapts to all cases?

Outline

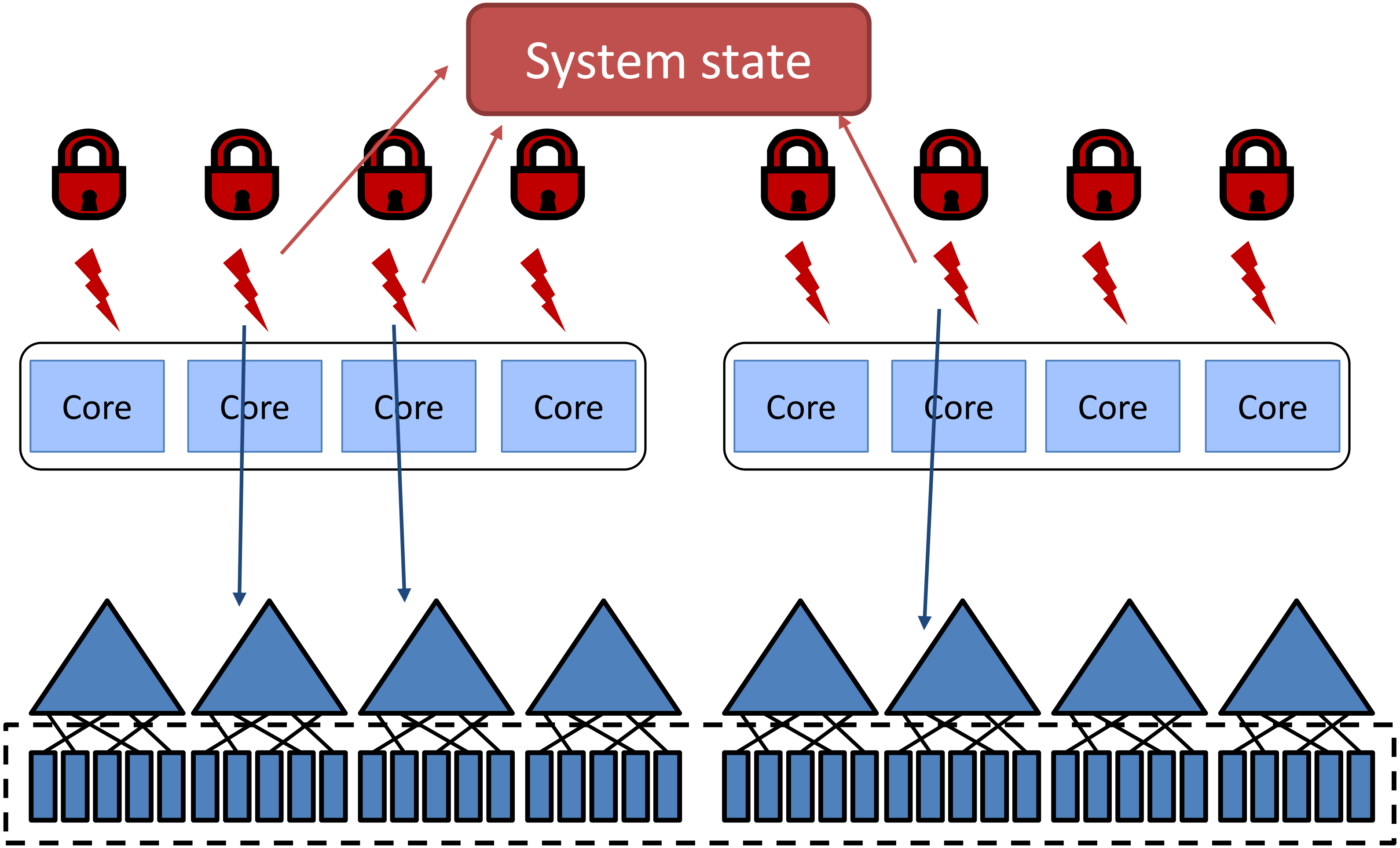
- Introduction
- Impact of Hardware Islands on OLTP
- Adaptive Transaction Processing for Islands
- Conclusion

Critical path of transaction execution



Many accesses to shared data structures

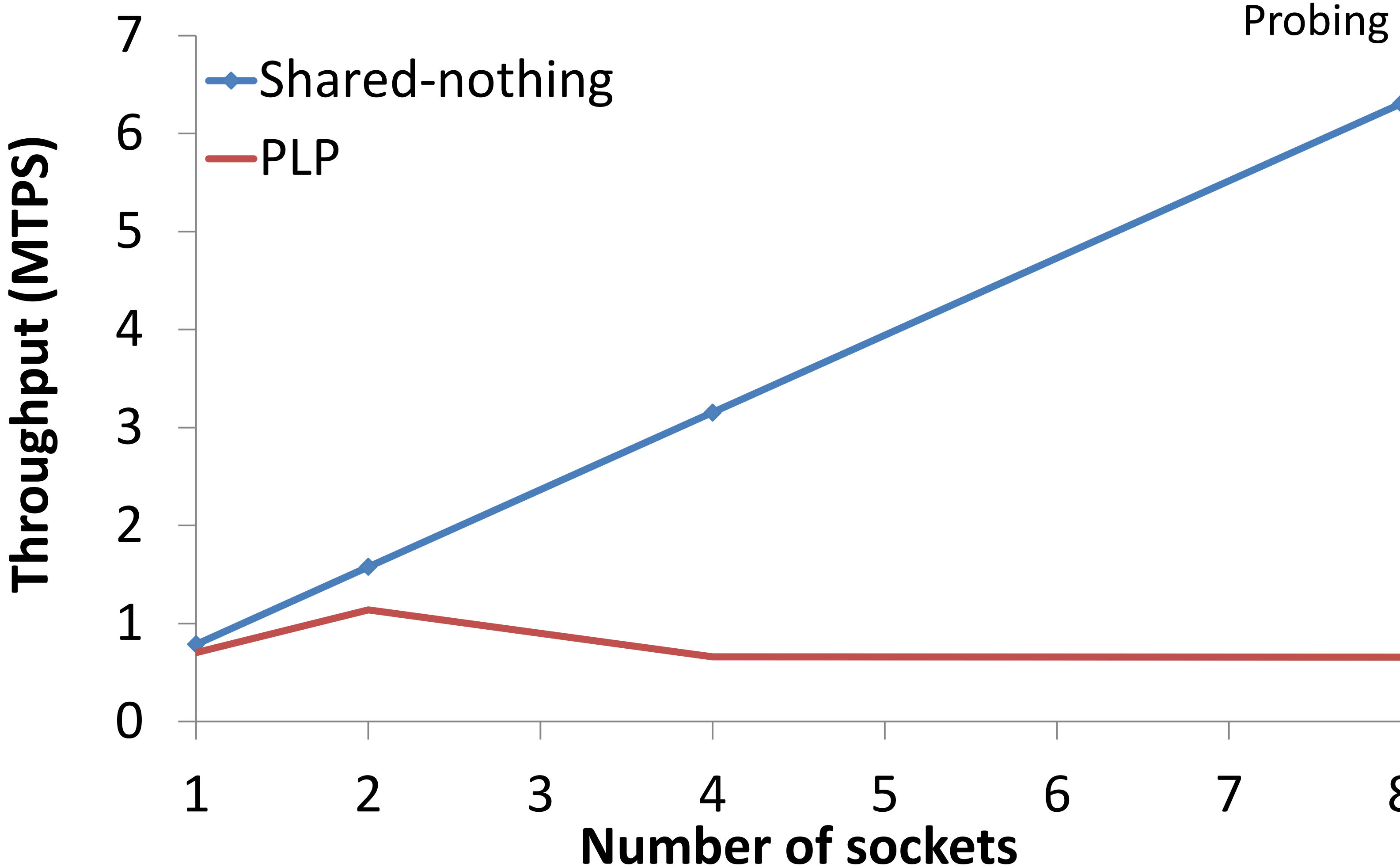
PLP: Physiologically partitioned SE



System state is still shared

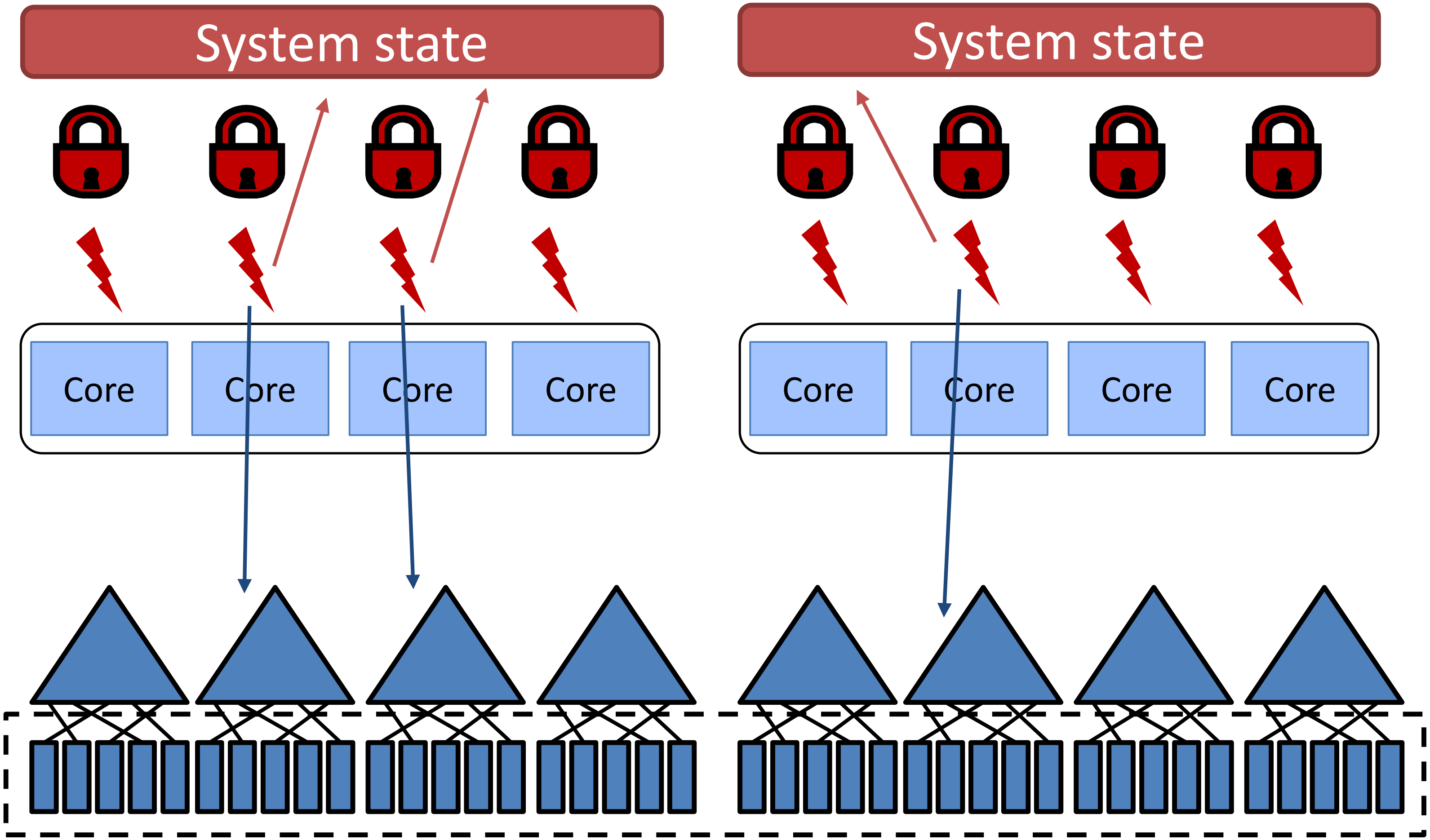
Perfectly partitionable workload

8socket x 10cores
Probing one row



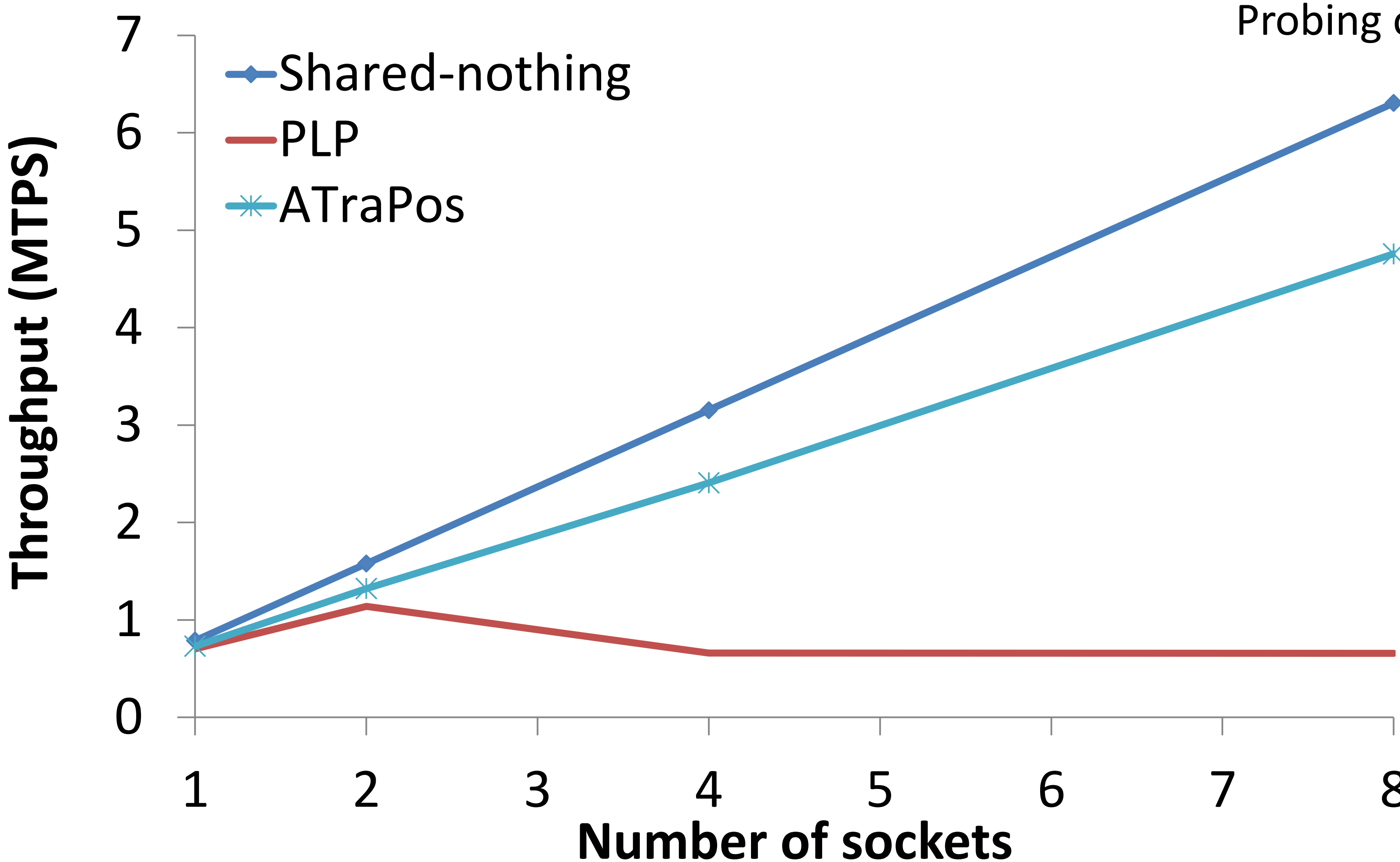
Inter-socket accesses to system state are a bottleneck

ATraPos: Island-aware SE



Perfectly partitionable workload

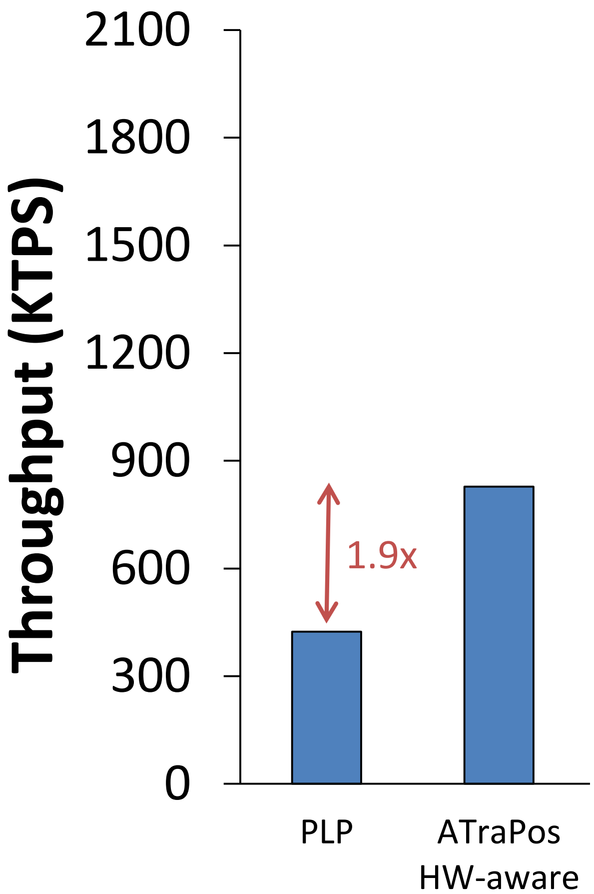
8socket x 10cores
Probing one row

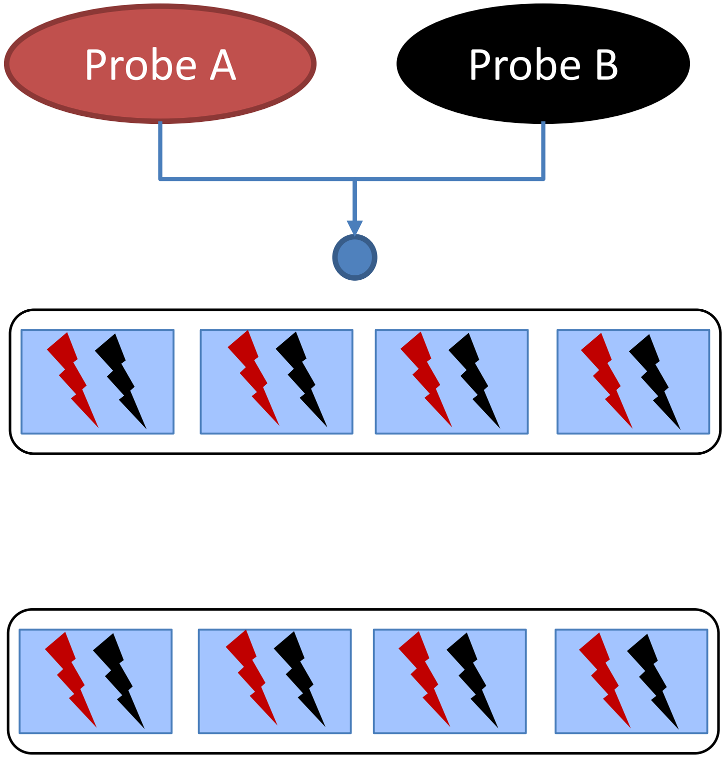


Island awareness brings scalability

Naive partitioning and placement

8 socket x 10 core
800K rows per table
Probing 1 row each from **A** and **B**

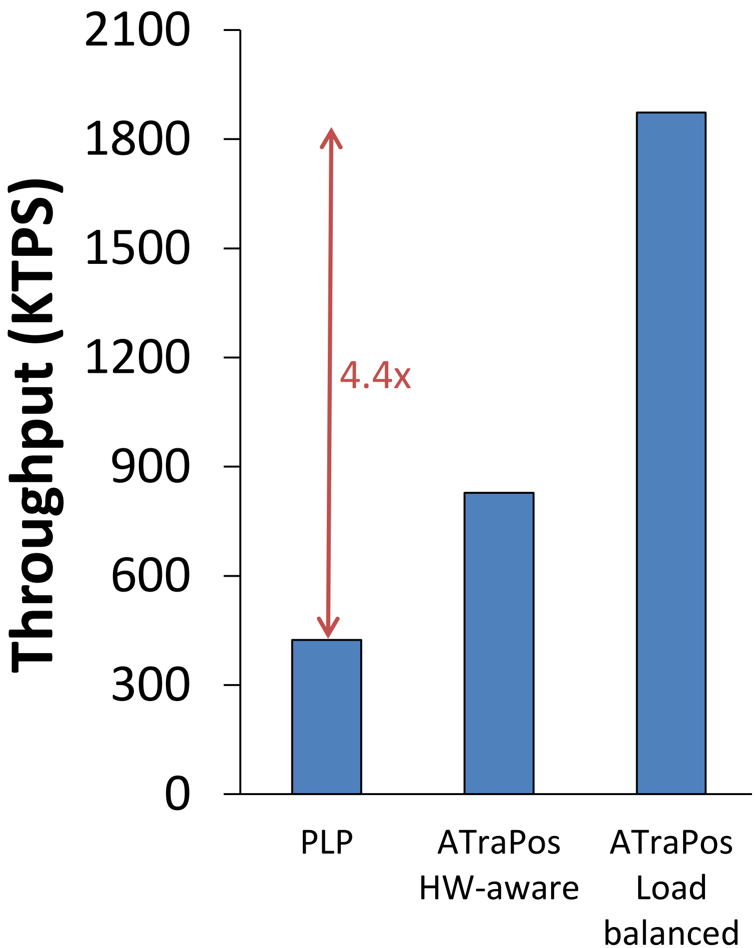


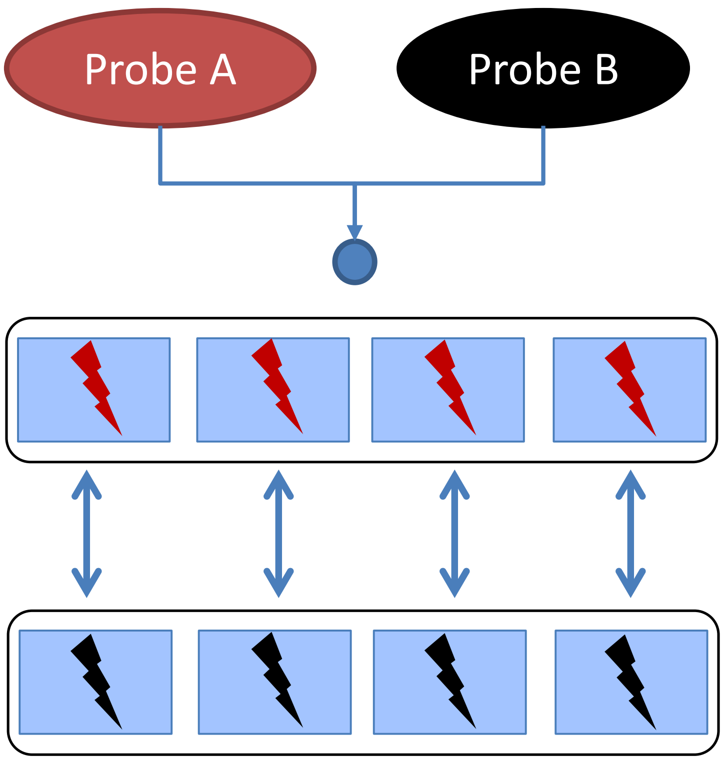


Cores are overloaded with contending threads

ATraPos partitioning and placement

8 socket x 10 core
800K rows per table
Probing 1 row each from **A** and **B**





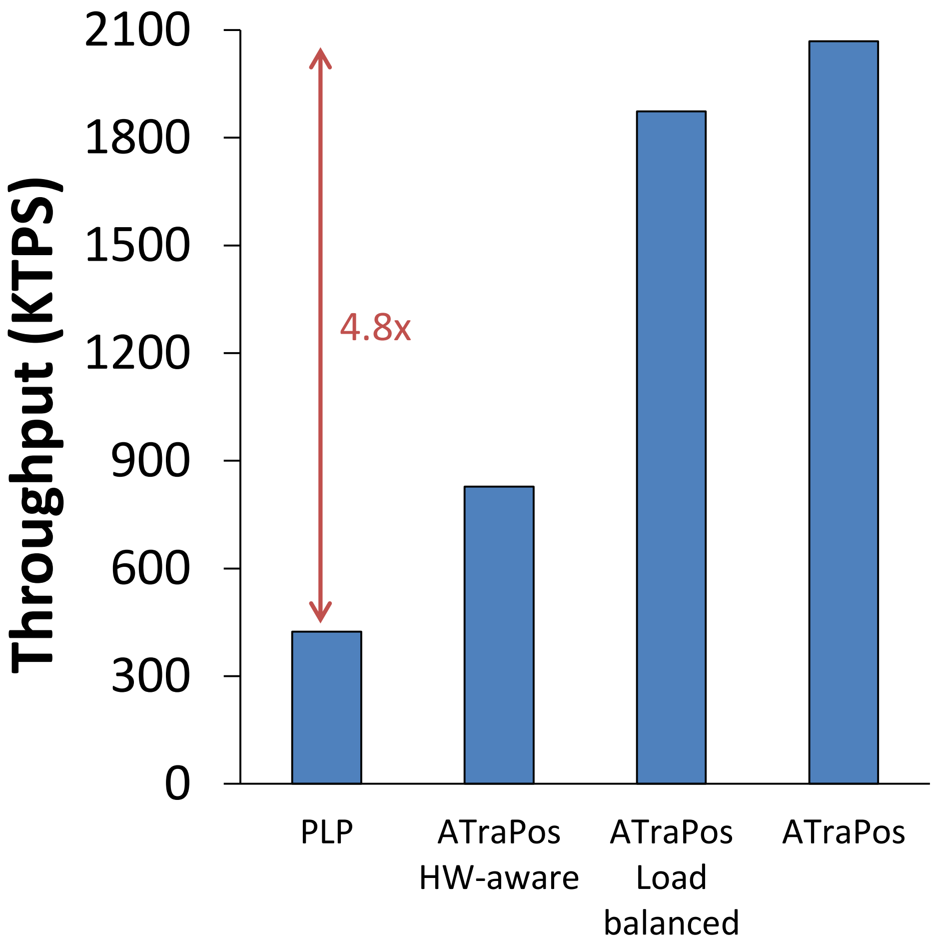
Ignoring Islands -> synchronization overhead

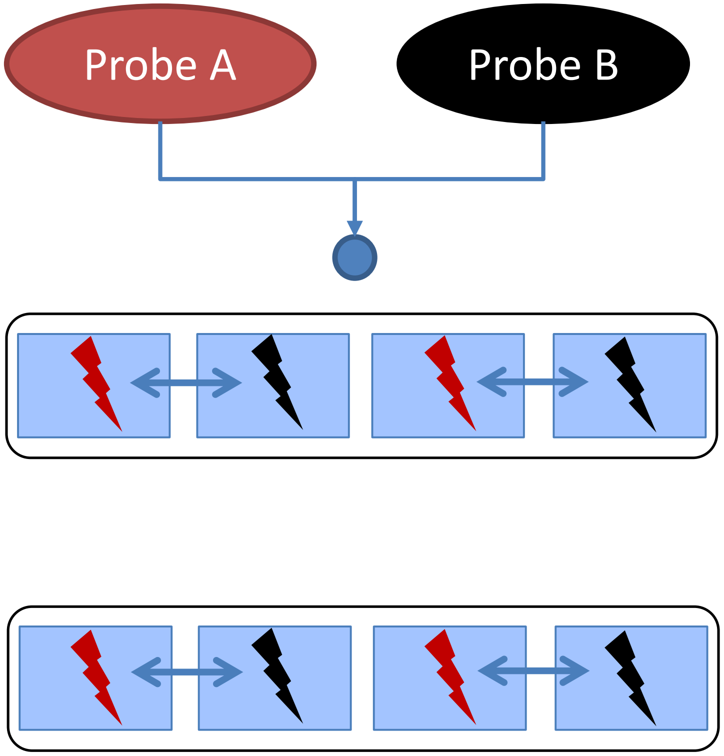
ATraPos partitioning and placement

8 socket x 10 core

800K rows per table

Probing 1 row each from **A** and **B**

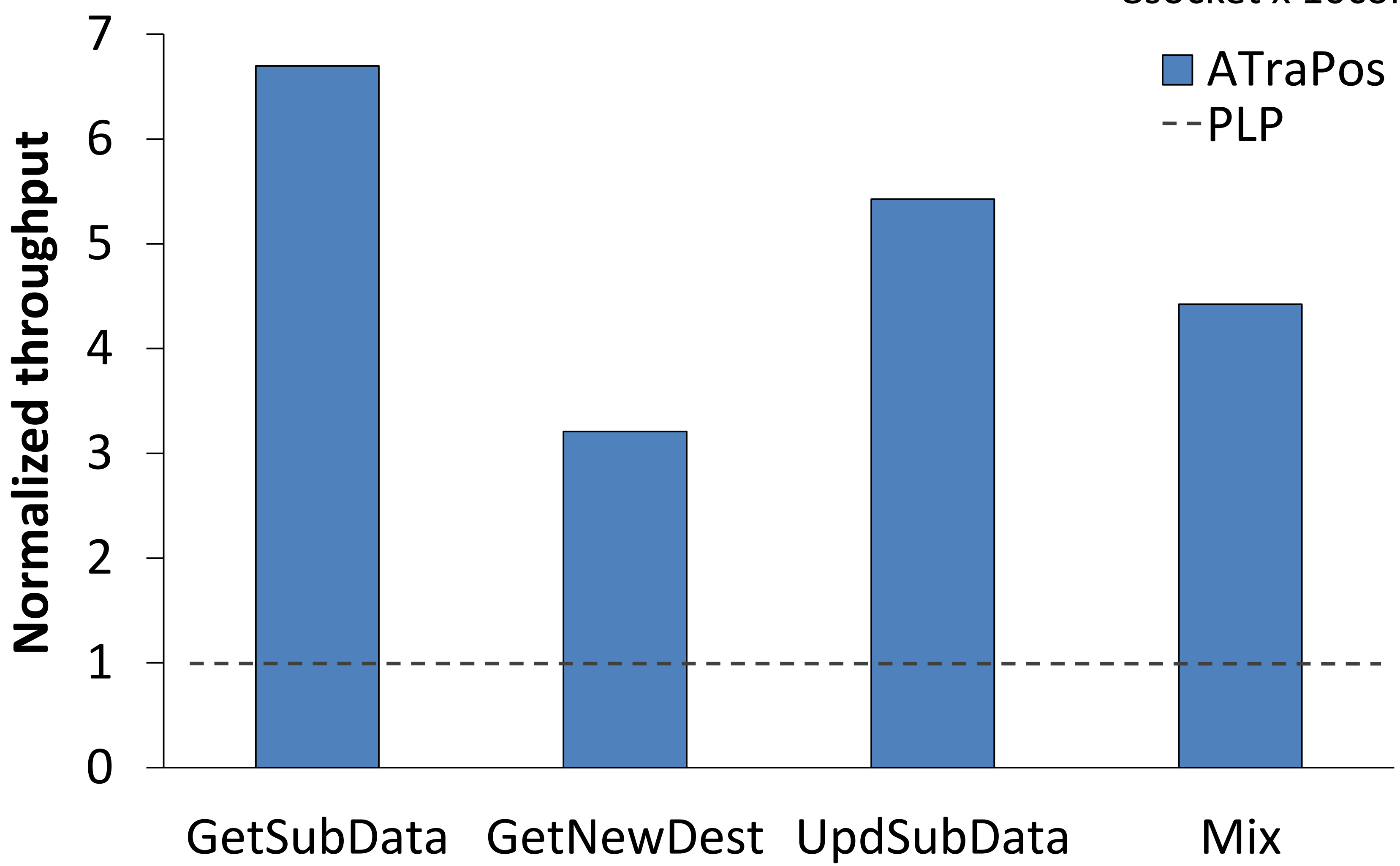




ATraPos: balanced load + reduced synchronization

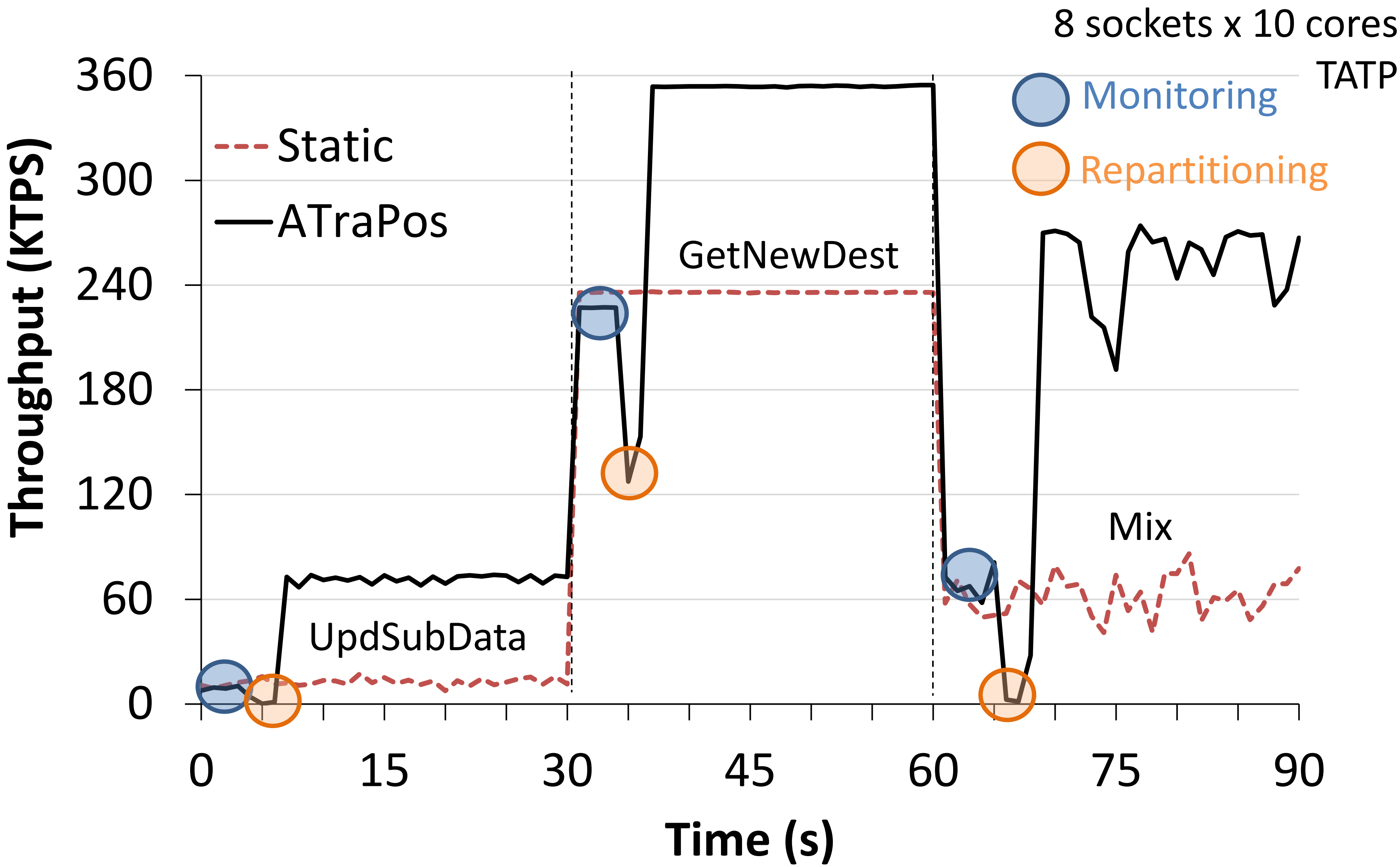
TATP - speedup over PLP

8socket x 10cores



ATraPos improves performance of TATP by 3.1-6.7x

Dynamic workloads



ATraPos gracefully adapts to any change

Summary

- OLTP on Hardware Islands
  - Shared-everything: stable, non-optimal performance
  - Shared-nothing: fast, sensitive to workload
  - Island shared-nothing: a robust middle ground
  - Optimal configuration depends on workload and hardware

- Adaptive transaction processing for Islands
  - Minimal inter-socket accesses in the critical path
  - Workload and HW-aware partitioning and placement
  - Lightweight monitoring and repartitioning

Thank you!