
CASH: Supporting IaaS Customers with a Sub-core Configurable Architecture

Yanqi Zhou¹

Henry Hoffmann²

David Wentzlaff¹



PRINCETON
School of Engineering and Applied Science



Department of
Computer Science



Web Services Have Latency Requirements

2 MINUTE READ

How One Second Could Cost Amazon \$1.6 Billion In Sales



HOW
Loading Time

AFFECTS YOUR
Bottom Line

Econsultancy Subscriber Research & Data Blog Events Training

It's Official – 'Web Stress' is Bad for Business

A Limited Pallet of Choices in EC2

Region:

	Linux/UNIX Usage	Windows Usage
Standard On-Demand Instances		
Small (Default)	\$0.065 per Hour	\$0.125 per Hour
Medium	\$0.125 per Hour	\$0.250 per Hour
Large	\$0.250 per Hour	\$0.500 per Hour
Extra Large	\$0.500 per Hour	\$1.000 per Hour
Micro On-Demand Instances		
Micro	\$0.020 per Hour	\$0.020 per Hour
High-Memory On-Demand Instances		
Extra Large		\$0.570 per Hour
Double Extra Large		\$1.140 per Hour
Quadruple Extra Large		\$2.280 per Hour
High-CPU On-Demand Instances		
Medium		\$0.285 per Hour
Extra Large		\$1.140 per Hour
Cluster Compute Instance		
Eight Extra Large	\$2.400 per Hour	\$2.970 per Hour

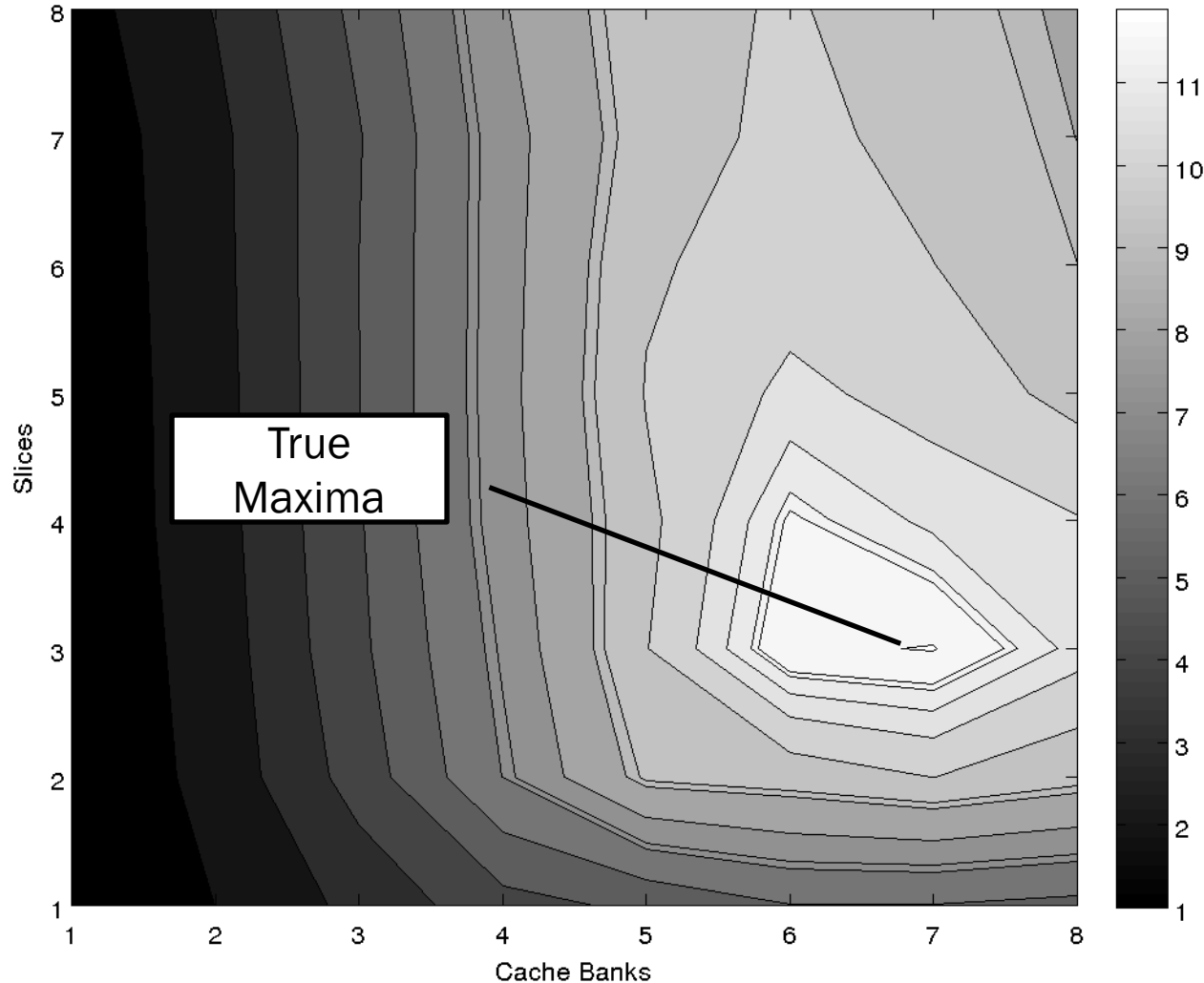
I need more performance

I need to reduce costs

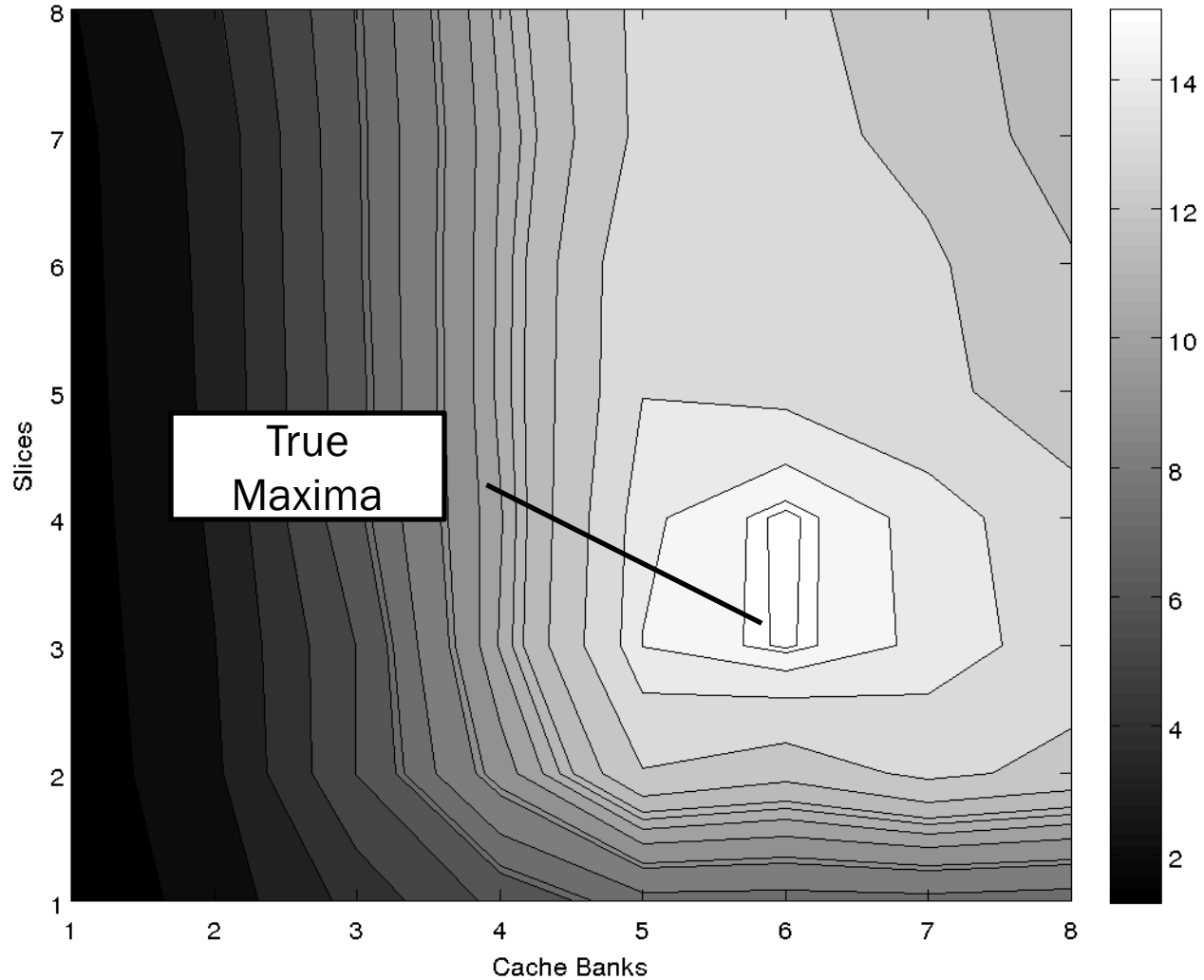


Features	Distributed ILP	TRIPS, CLP	Core Fusion	WiDGET	Conjoined Cores	Clustered	big. LITTLE	Sharing Arch.
Scale up & down	✓	✓	✗	✓	✗	✗	✗	✓
Distributed/switched	✓	✓	✗	✗	✗	✗	✗	✓
Symmetric	✓	✓	✓	✓	✓	✓	✗	✓
Partition L2	✓	✓	✗	✗	✗	✗	✗	✓
Dynamic OoO	✗	✗	✓	✗	✓	✓	✓ ✗	✓
ISA Compatible	✓	✗	✓	✓	✓	✓	✓	✓
Multiple Metrics	✗	✓	✗	✗	✗	✗	✗	✓

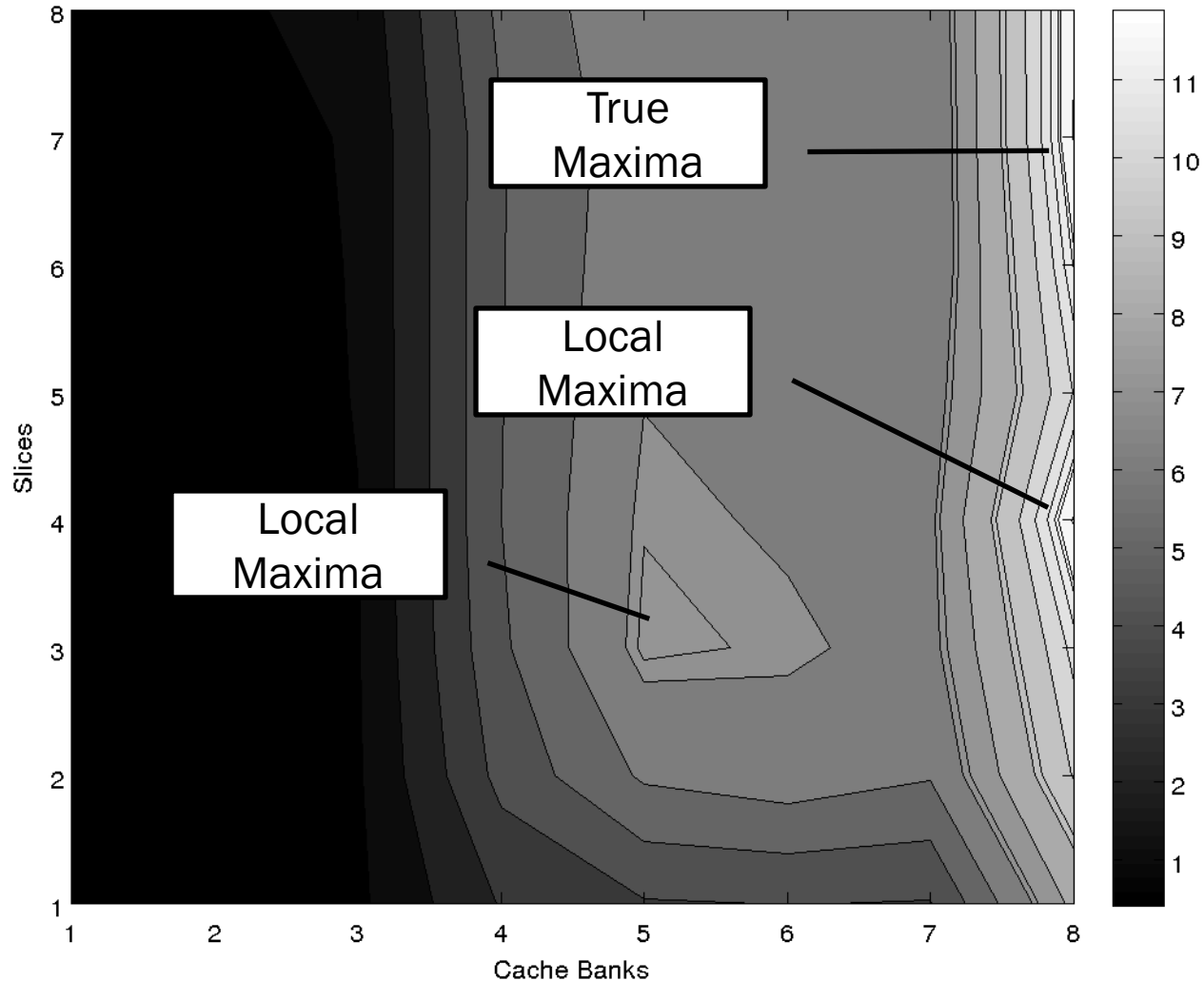
Managing Fine-grain Configurability is Hard



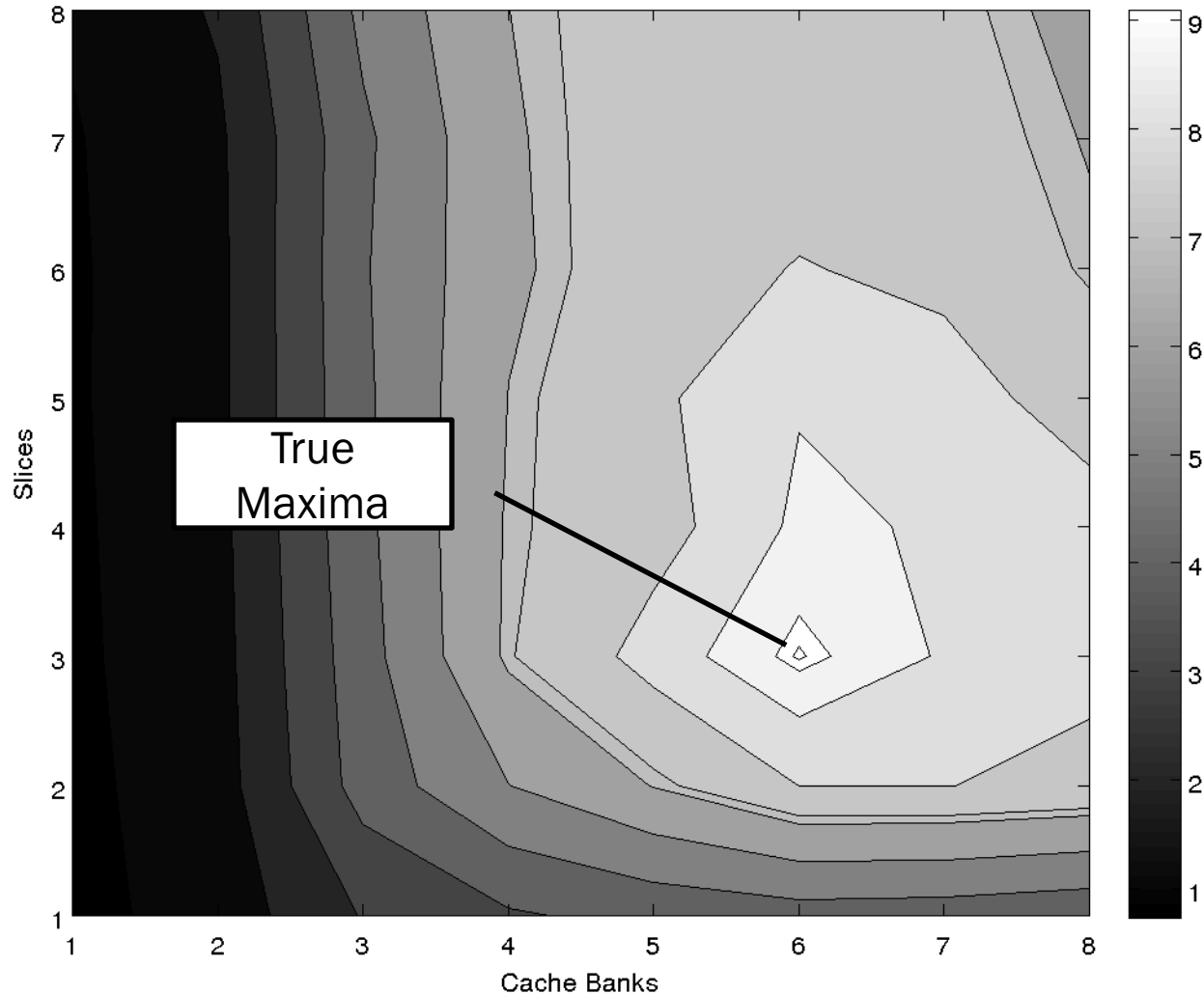
Managing Fine-grain Configurability is Hard



Managing Fine-grain Configurability is Hard

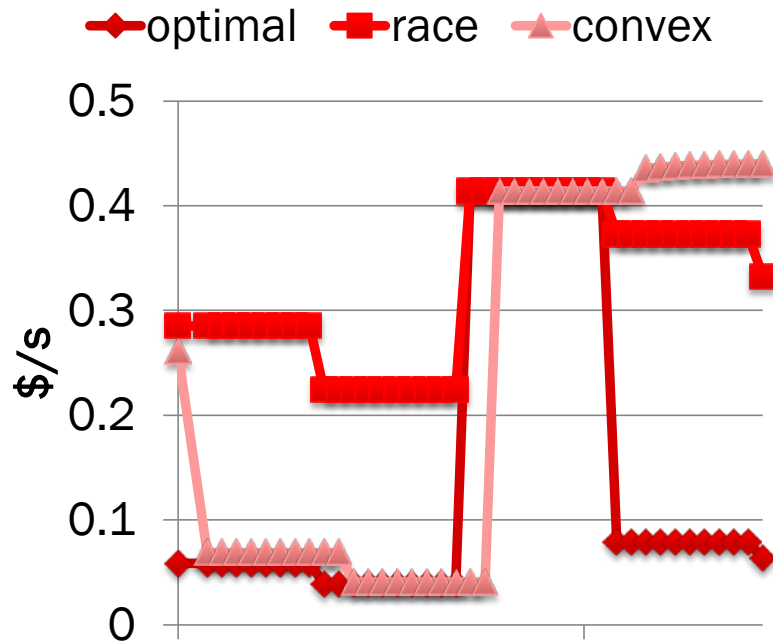


Managing Fine-grain Configurability is Hard

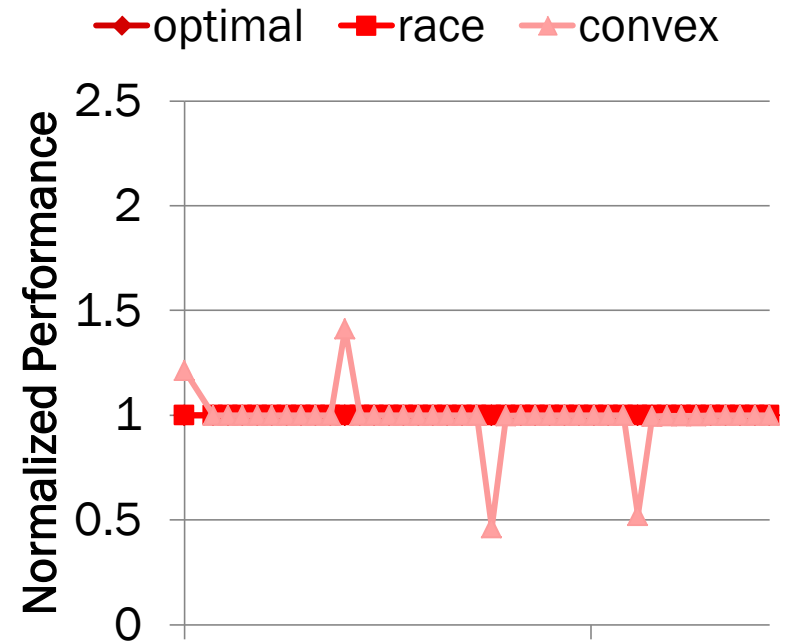


Benefit of Managing Fine-grain Configurability

Cost



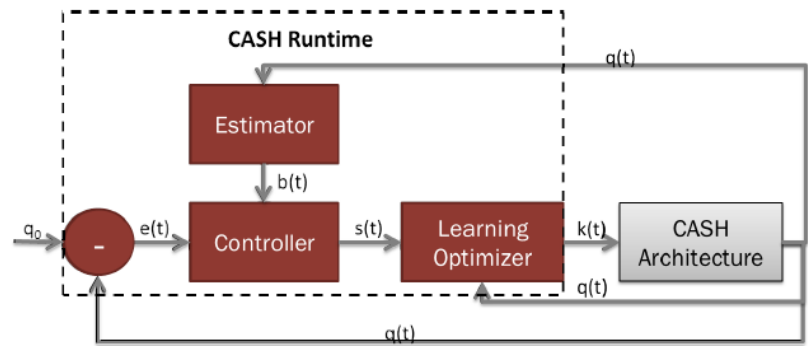
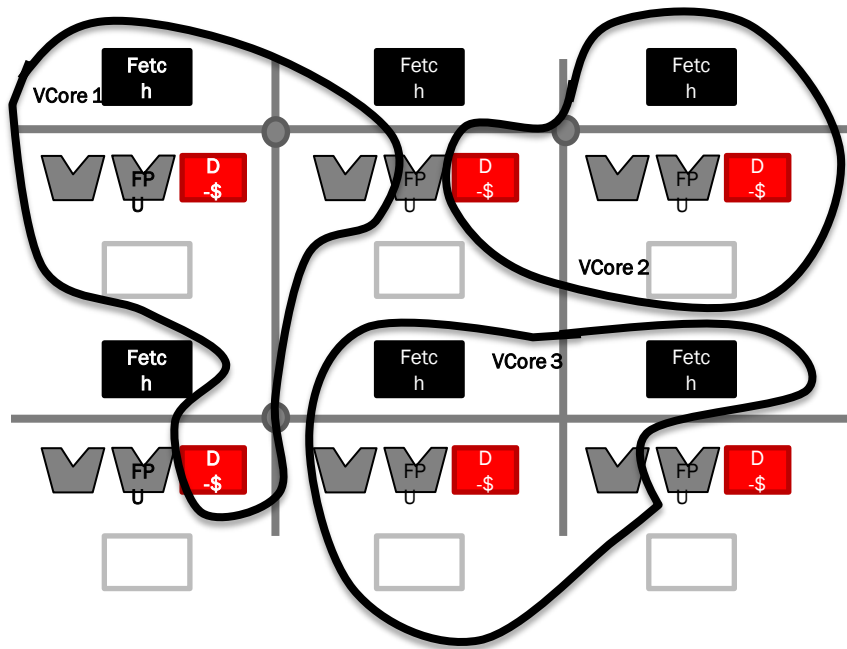
Performance



Observations:

- Fine-grain configurability is extremely helpful given current constraints
- Such configurability *should* produce non-convex optimization spaces
- Much research needs to be done on managing these non-convex spaces

CASH: Architecture and Runtime for Managing Fine-grain Configurability

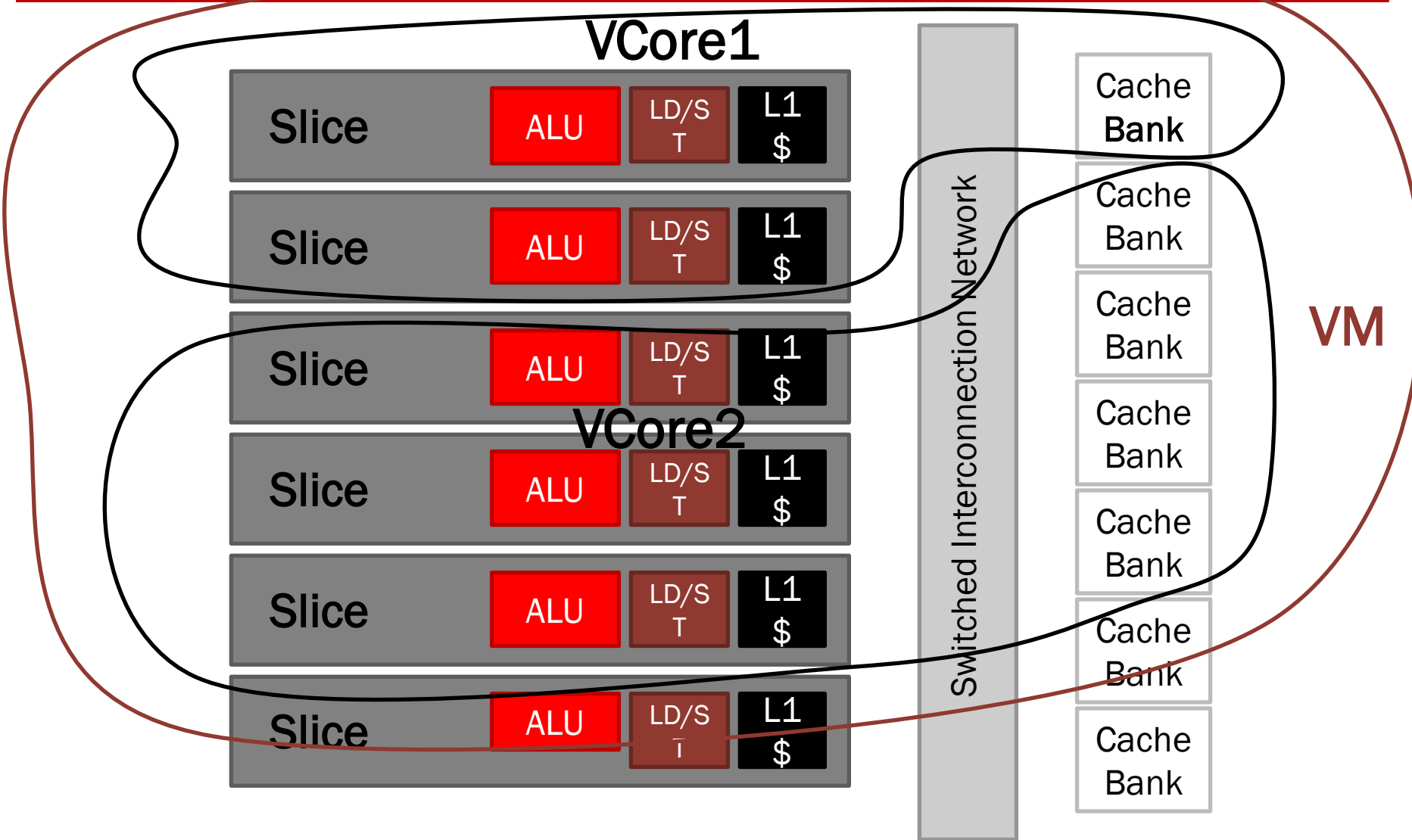


Solution:

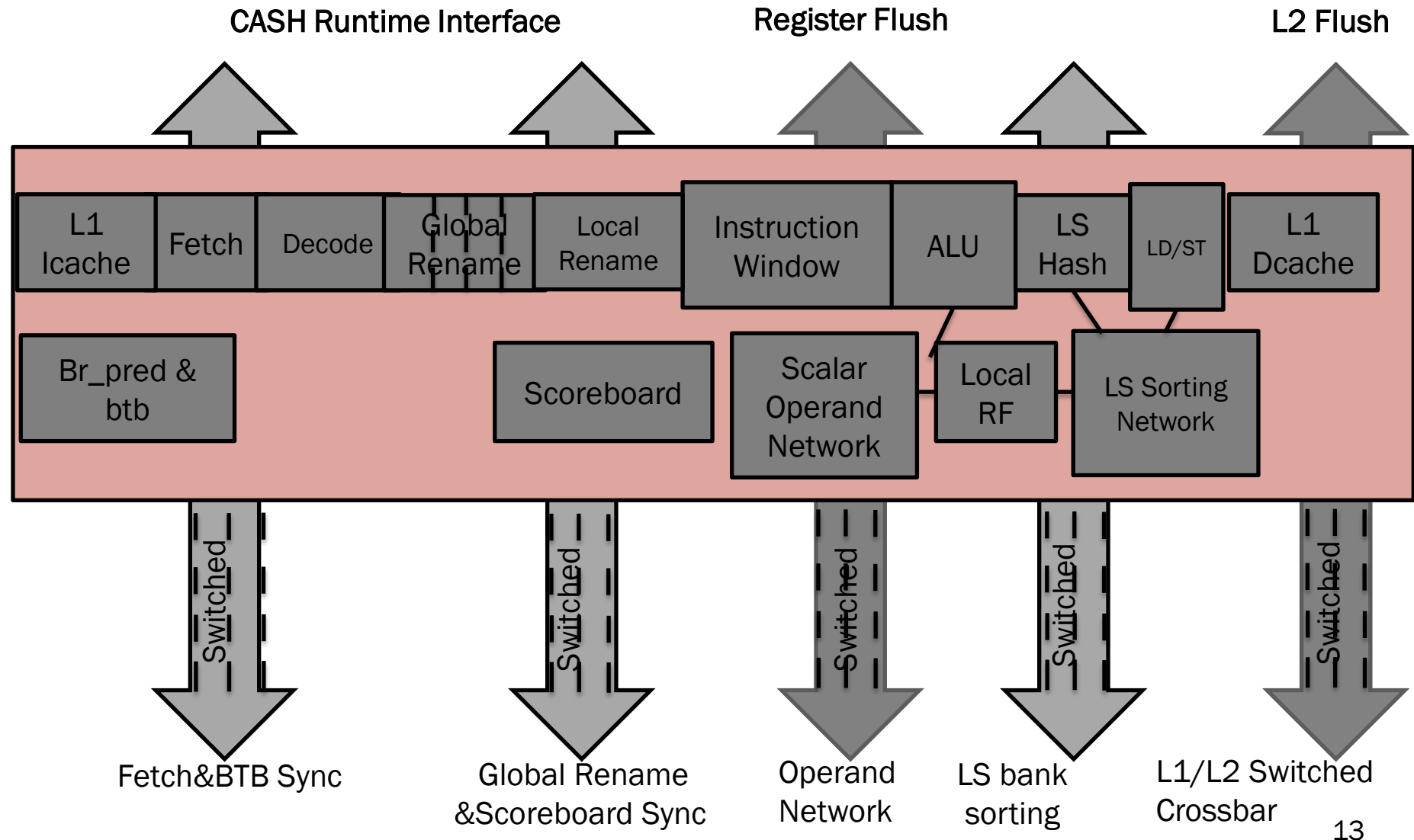
- Statically homogeneous, dynamically heterogeneous architecture
- Lightweight runtime system manages configurability online

Architecture: Sharing Architecture

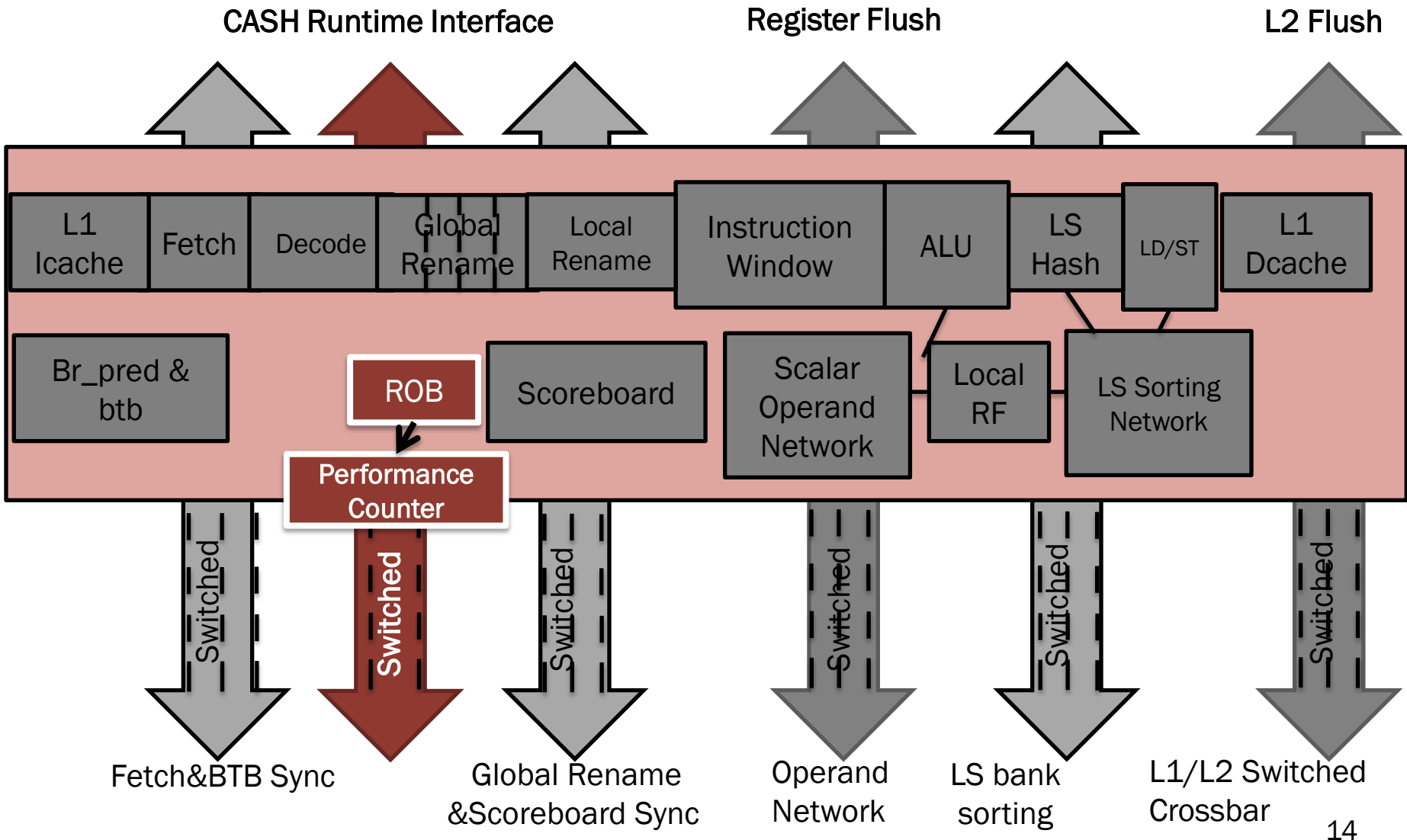
[Zhou & Wentzlaff ASPLOS 2014]



Slice Details



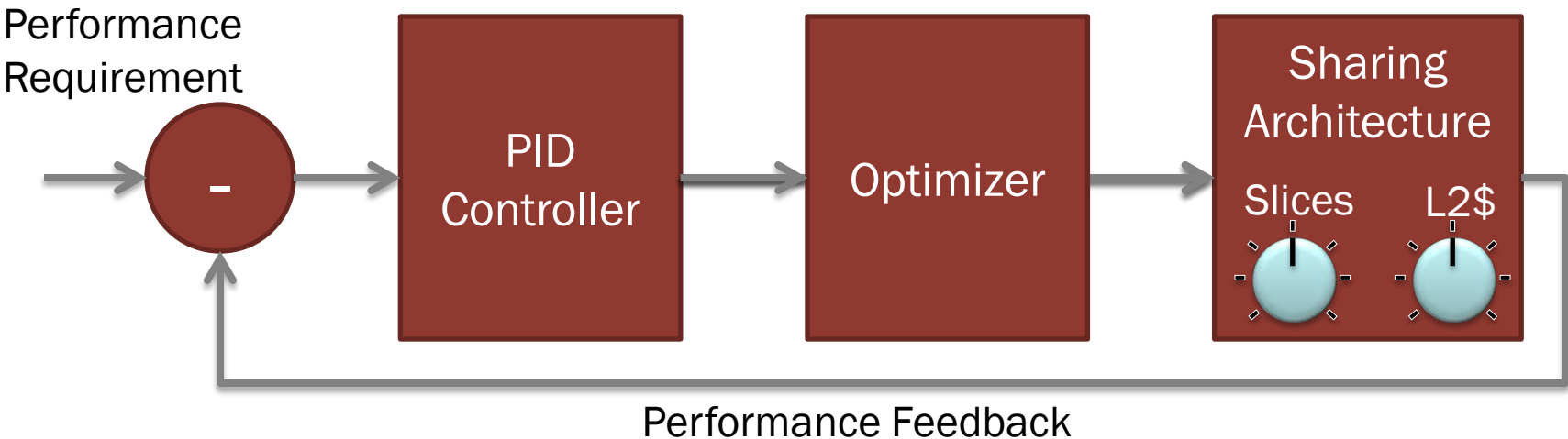
Aggregating Performance Data Across Slices



Controlling Application Performance

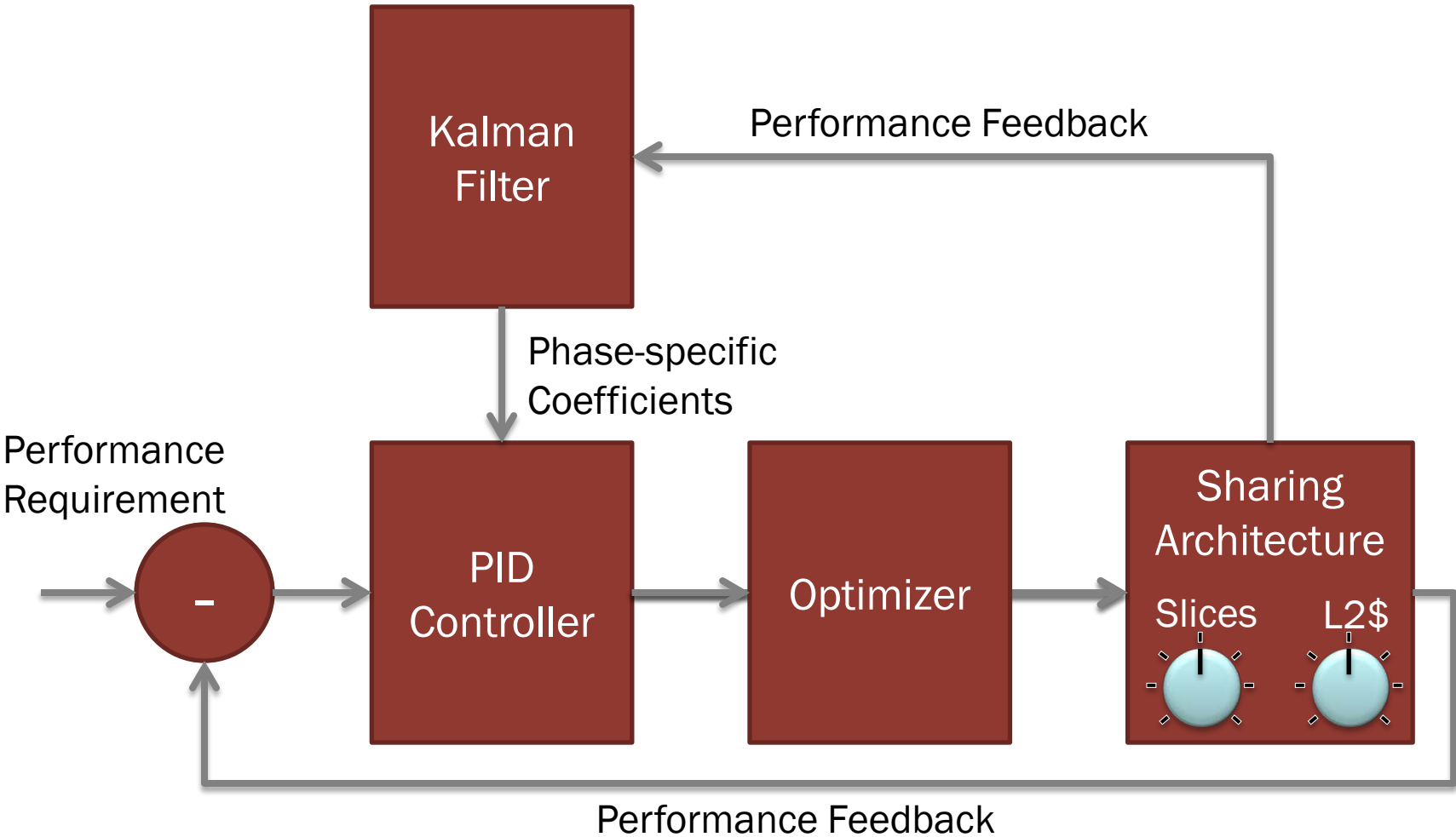
All problems in computer science can be solved by another level of indirection...
 -- David Wheeler

Corollary: All optimization problems can be solved by another level of feedback...



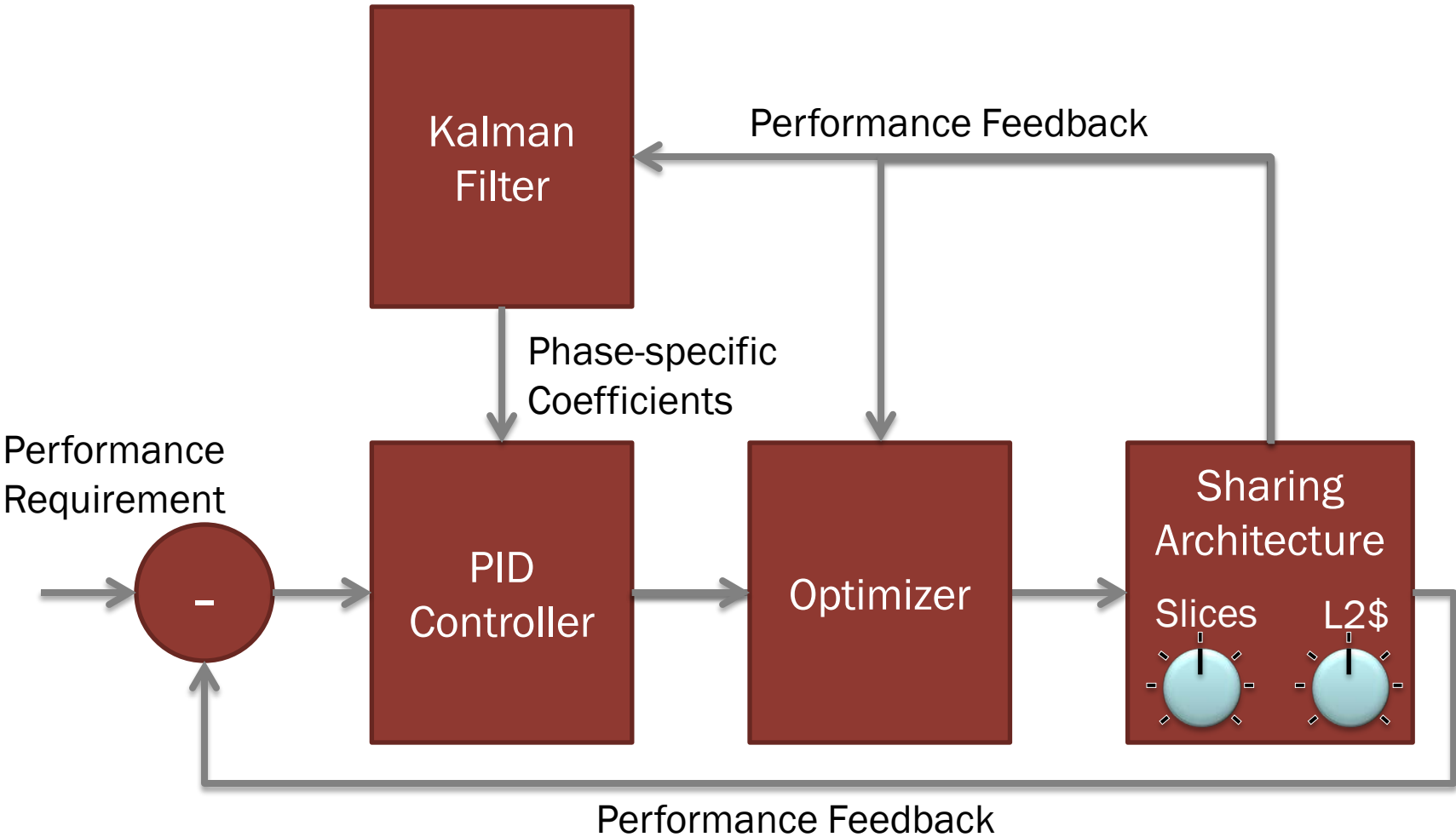
Controlling Application Performance

Adapting to Phases



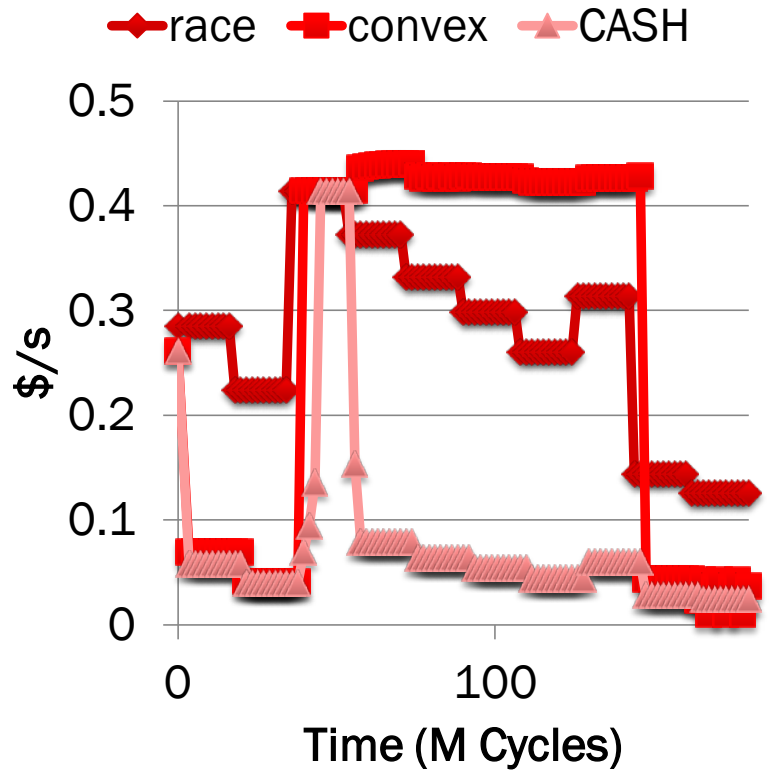
Controlling Application Performance

Learn Online

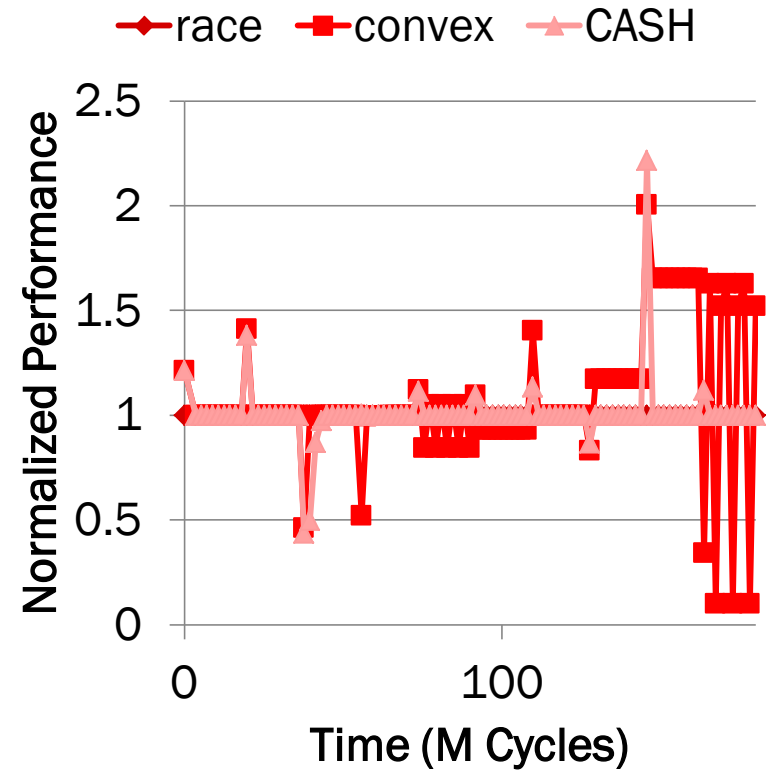


Benefit of Managing Fine-grain Configurability

Cost

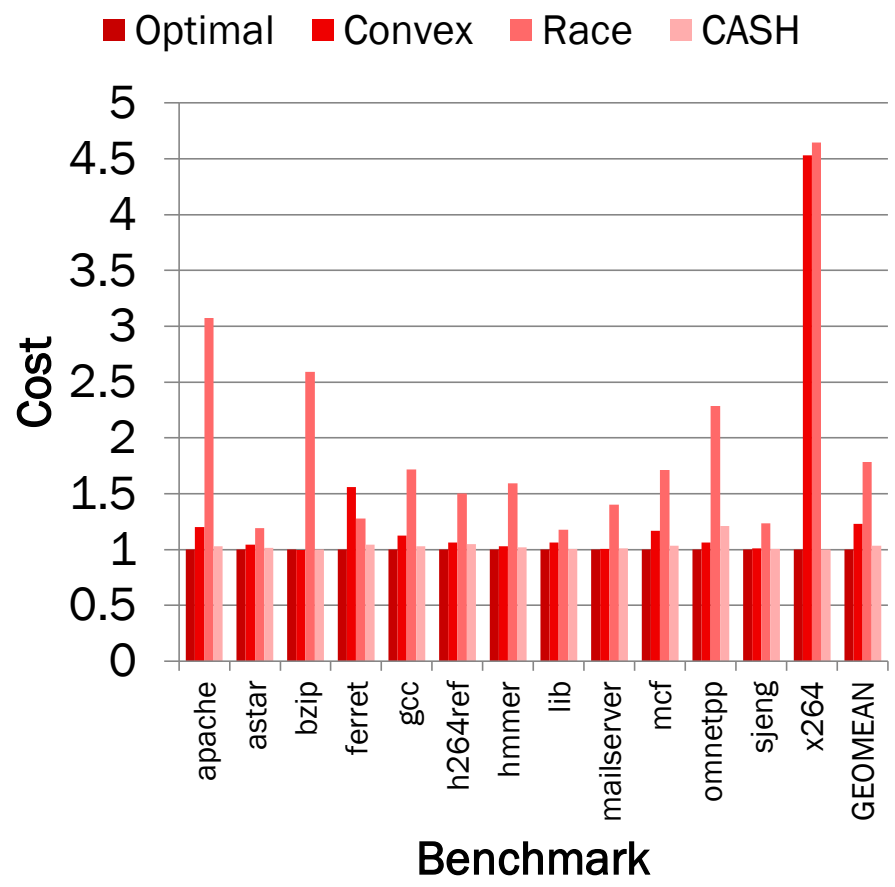


Performance

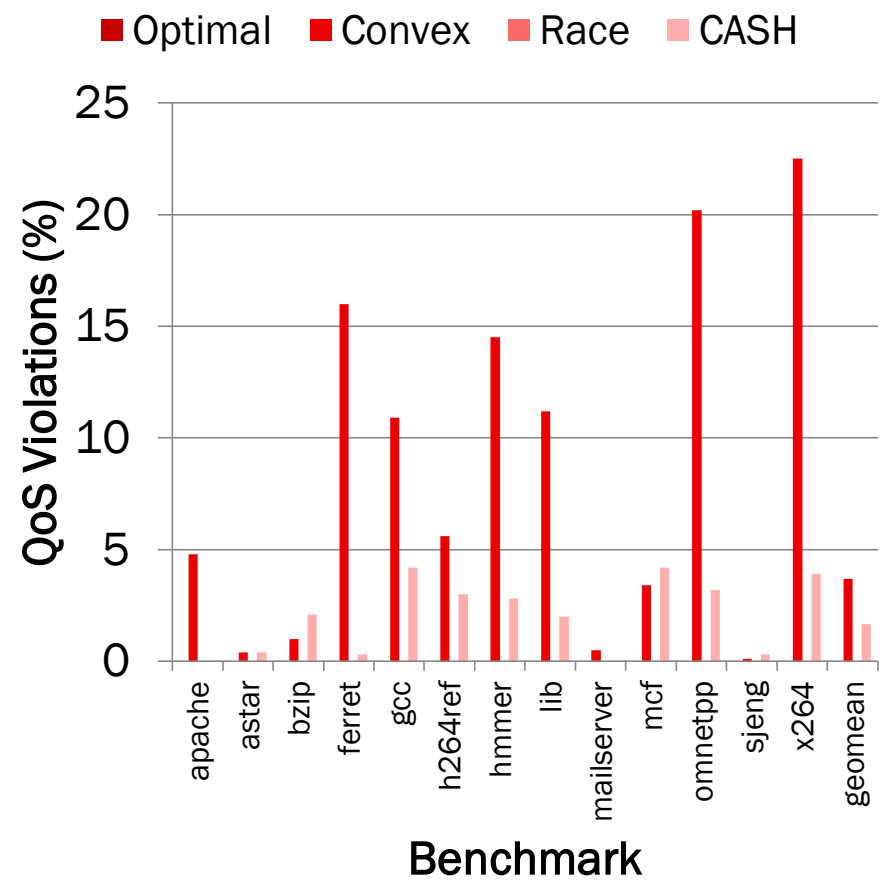


Comparison of Fine-grain Management Techniques

Cost



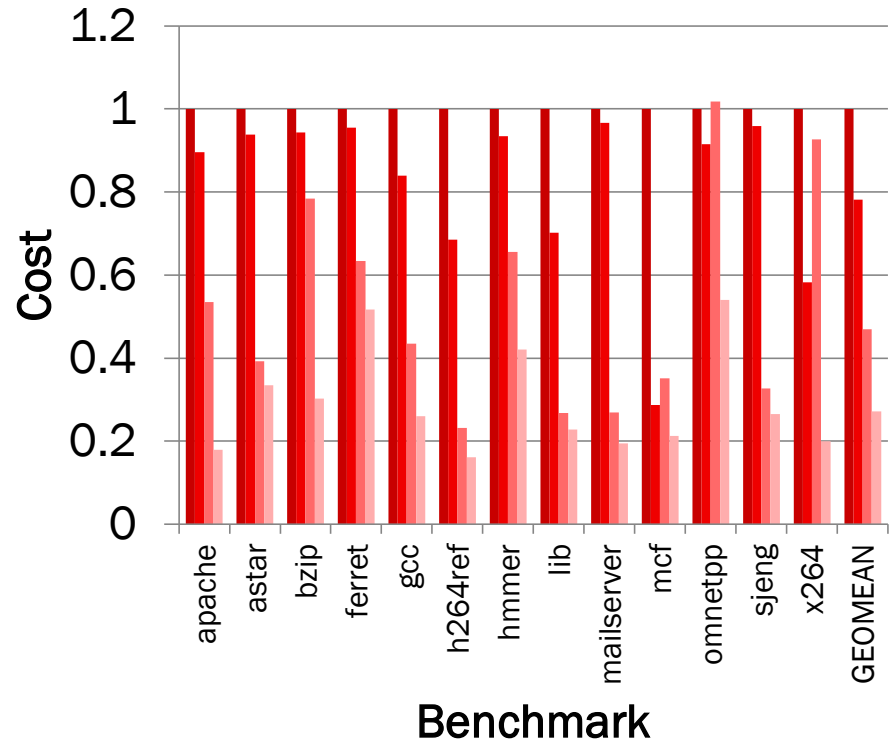
QoS Violations



Comparisons to Coarse-grain Reconfigurable

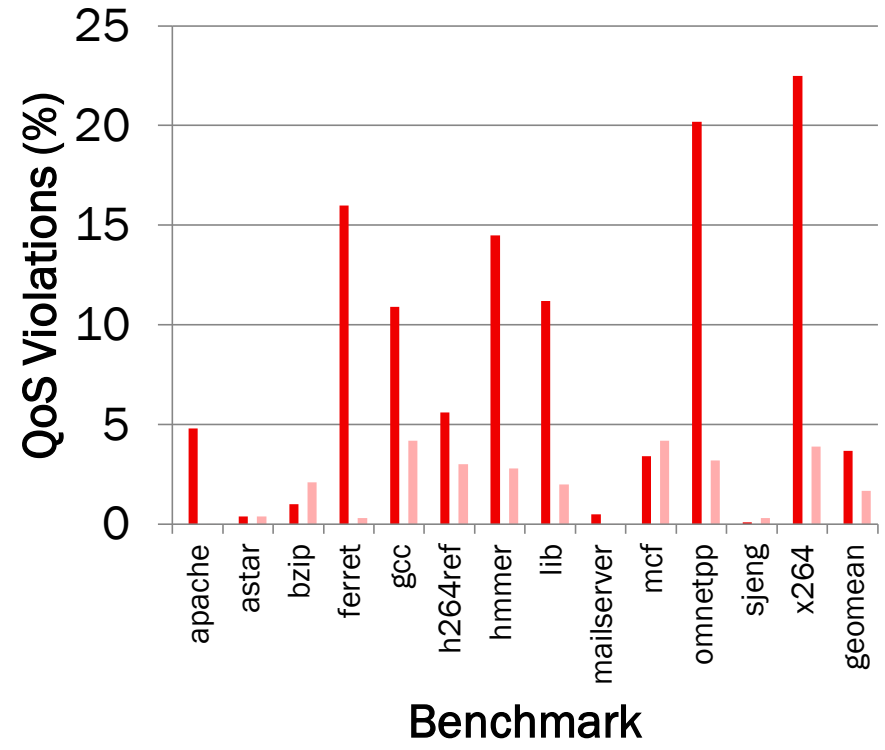
Cost

- Coarse Race
- Coarse Adapt
- Fine Race
- CASH Adapt



QoS Violations

- Coarse Race
- Coarse Adapt
- Fine Race
- CASH Adapt



CASH: Supporting IaaS Customers with a Sub-core Configurable Architecture

Yanqi Zhou Henry Hoffmann David Wentzlaff

