Georgia Tech

System Abstractions for Resource Scaling on Heterogeneous Platforms

Vishal Gupta, Min Lee
Karsten Schwan

CERCS
Georgia Institute of Technology

http://www.istc-cc.cmu.edu/

Overview – GT CC Research

Elastic and Reliable Cloud Services on Next Generation Server Platforms (+ 'at the edge') – touching elements of all pillars (automation, specialization, big data, ...)

- Kisung Lee -Utility-aware Graph Partitioning for Efficient Graph Query Processing
- Hrishikesh Amur – memory efficient 'groupby-aggregate' and 'KV's (with Michael Kaminsky Dave Anderson, Greg Ganger)
- Liting Hu – 'StreamingData' (large 'windows', coordinated queries, 'Amazon' use cases) and 'multi-DC provisioning' (with Mike Kozuch)
- Chengwei Wang – VScope ,VFocus, Monalytics - online troubleshooting (OS Review)
- Jack Li - VMM comparisons, including 'noisy' VMs
- Qi Zhang - Systems and Techniques for Optimizing Inter-VM Communication Bandwidths

- **Vishal Gupta, Min Lee – Heterogeneous Server Platforms (CPU, memory, ...) (Scott Hahn, ...)**
- Priyanka Tembey – 'Virtual Platforms', 'Islands of Cores' (George Cox, Intel)

- Jian Huang (Intel URO) – 'Nand-Flash VM' (with Microsoft)
- Sudarsun Kannan (Intel URO) – 'Client NVM'

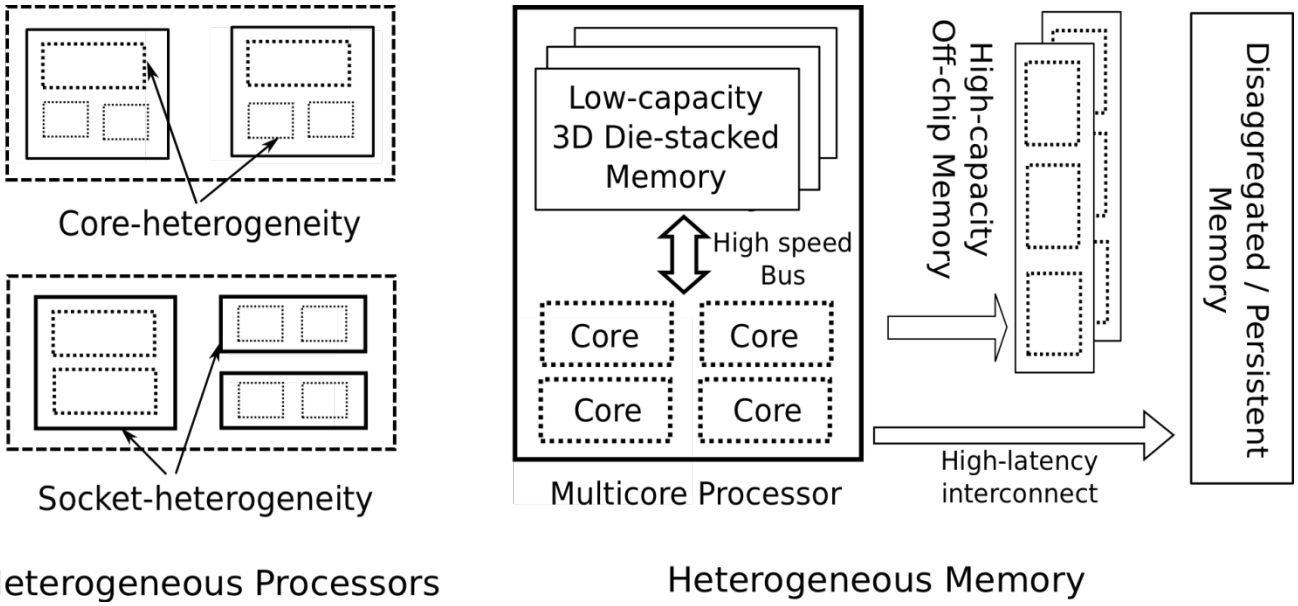
- Jeff Young – PGAS systems (with Alex Merrit) (see his poster)
- Se Hoon Shon – Data Warehousing on GPU
- Ifrah Saeed - working with Se (see her poster)

- Minsung Jang and Ketan Bhardwaj – 'ToTheEdge' – EC ISTC – `Aggregated Edge Platforms'

- Additional students funded from other sources (12+), including HPC 'Big Data' and 'Exascale'

Heterogeneous Platforms: Processor and Memory

Components with different performance/power characteristics



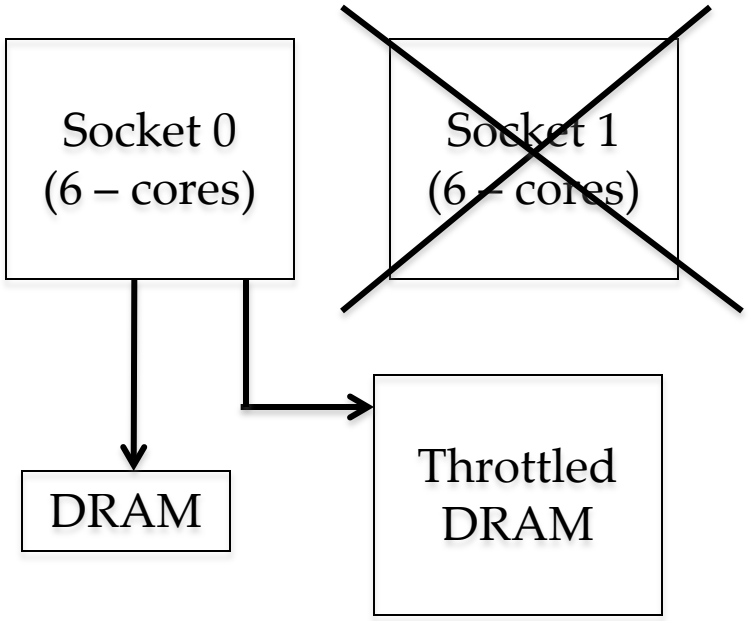
Specific Case for Heterogeneity: Big/Little Cores

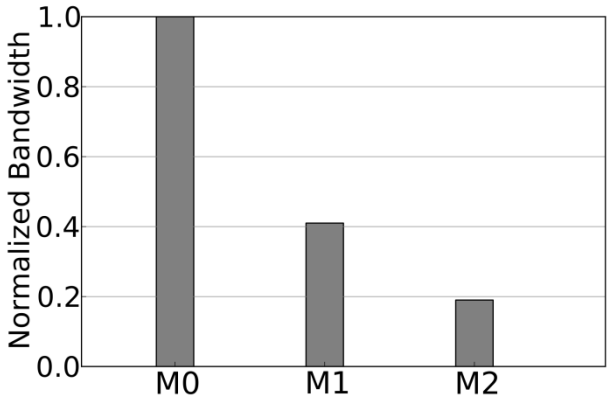
Background: Big/Little Cores

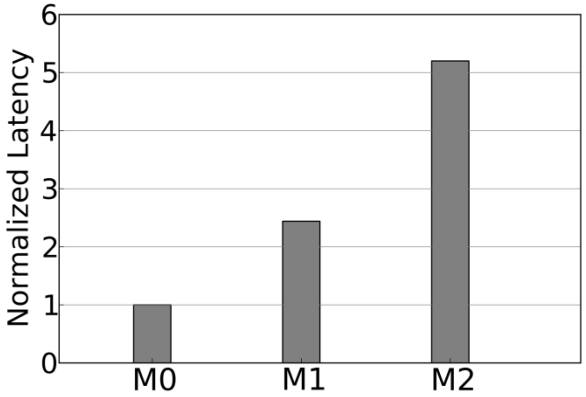
- Brawny vs. Wimpy
  - Evaluation of modern workloads on heterogeneous cores
- Beyond Core (with Ganapati et al. – Intel)
  - Analysis of uncore and heterogeneous memory organizations
- HeteroMates (with Scott Hahn et al. – Intel)
  - High dynamic range on mobile devices using heterogeneous cores
- HeteroVisor (new work presented next)
  - Elastic resource scaling for heterogeneous cloud platforms
  - Key addition: memory scaling

Background: Emulation Platform

Memory throttling to emulate heterogeneous memory (for stacked vs. offchip DRAM)



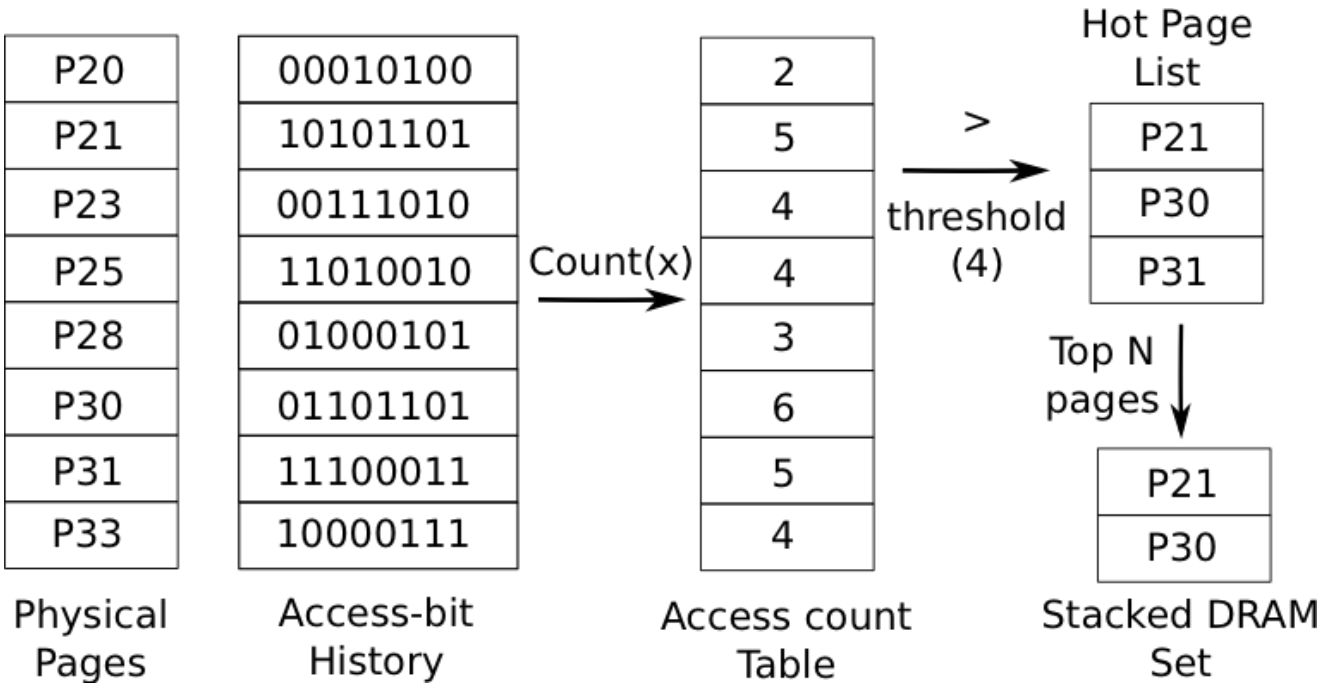




Note: method too inaccurate for NVM

Background: Hot Page Detection

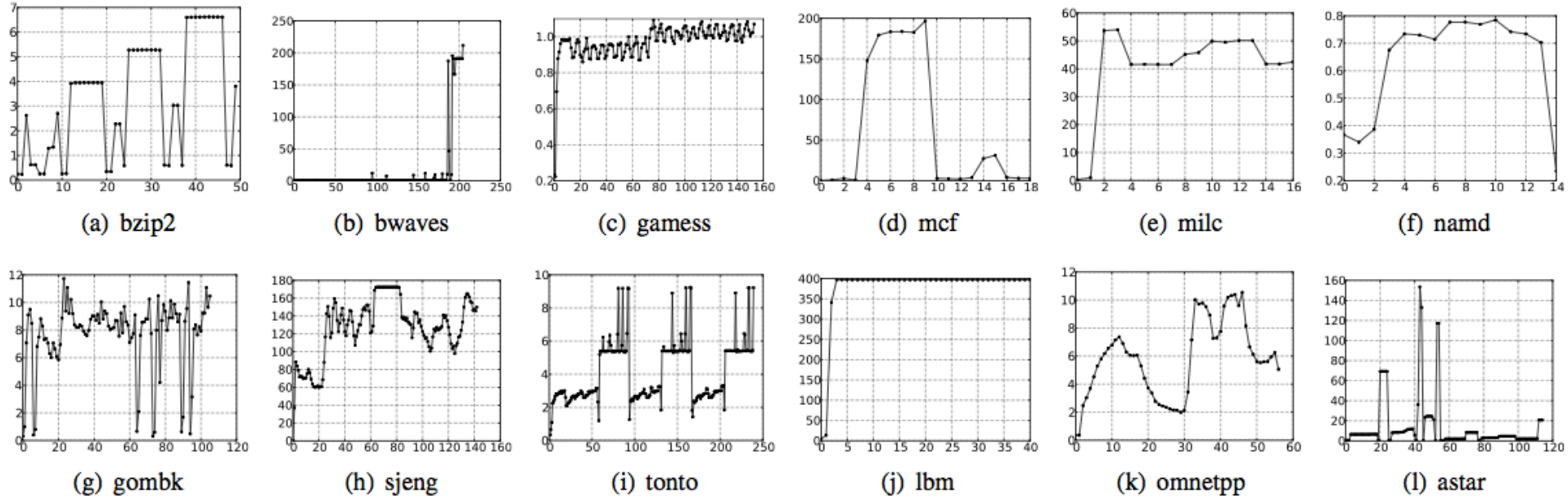
Using page-table access-bits for detecting hot pages



For NVM, now adding functionality to distinguish read/write accesses

Background: Working-set Size Tracking

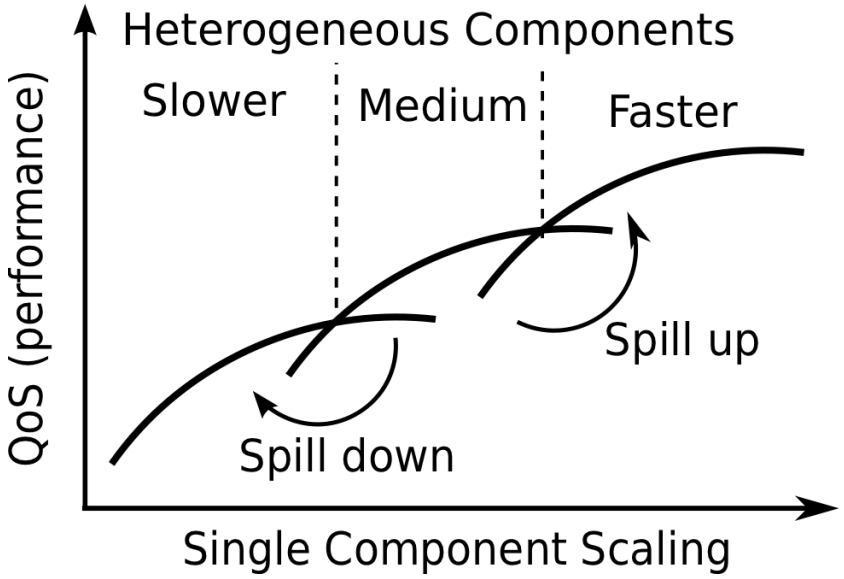
Diversity across applications requiring dynamic management

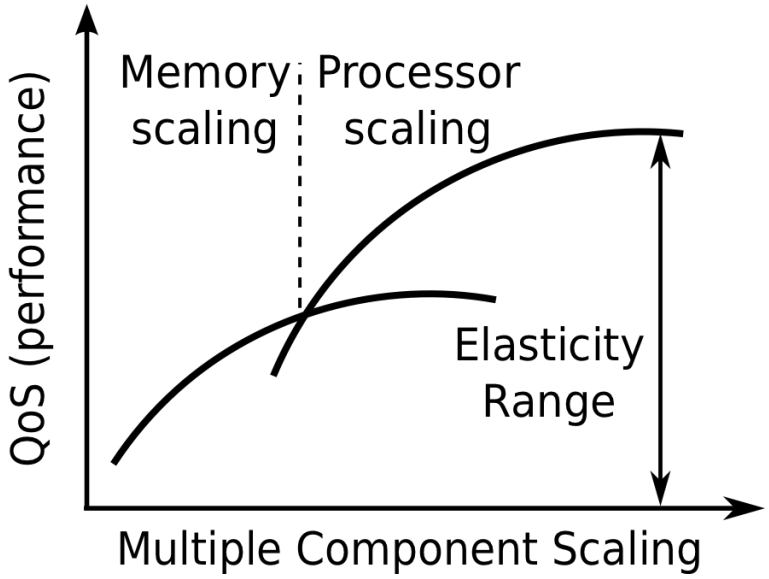


SPEC CPU2006

HeteroVisor: Scaling using Heterogeneity

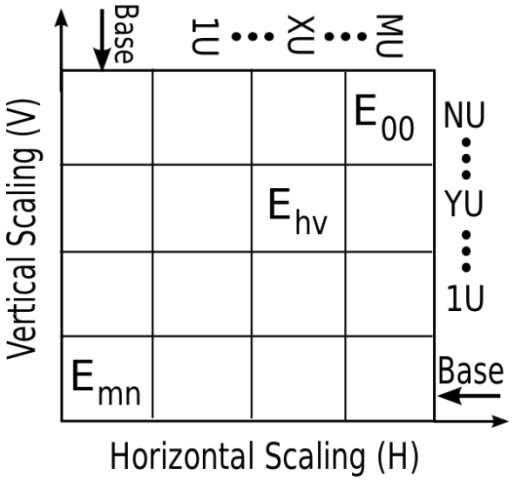
Key Idea: use both processor and memory heterogeneity to extend scaling range

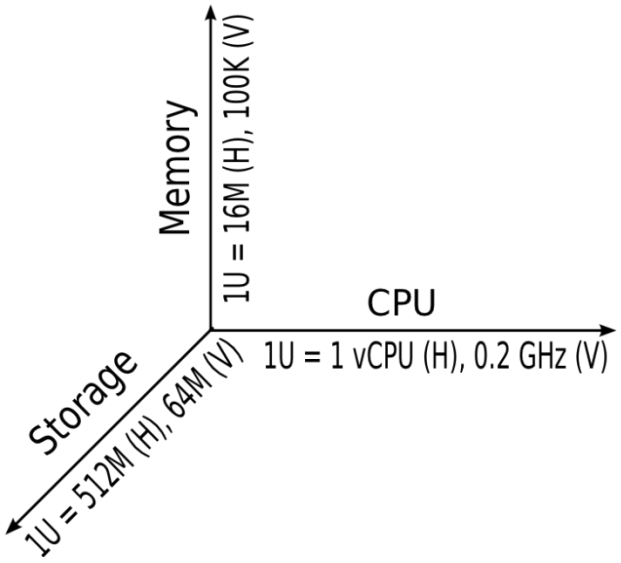




Concepts: Elasticity (E) States

System-wide abstraction to dynamically request resources
Incorporates horizontal/vertical scaling and different platform resources

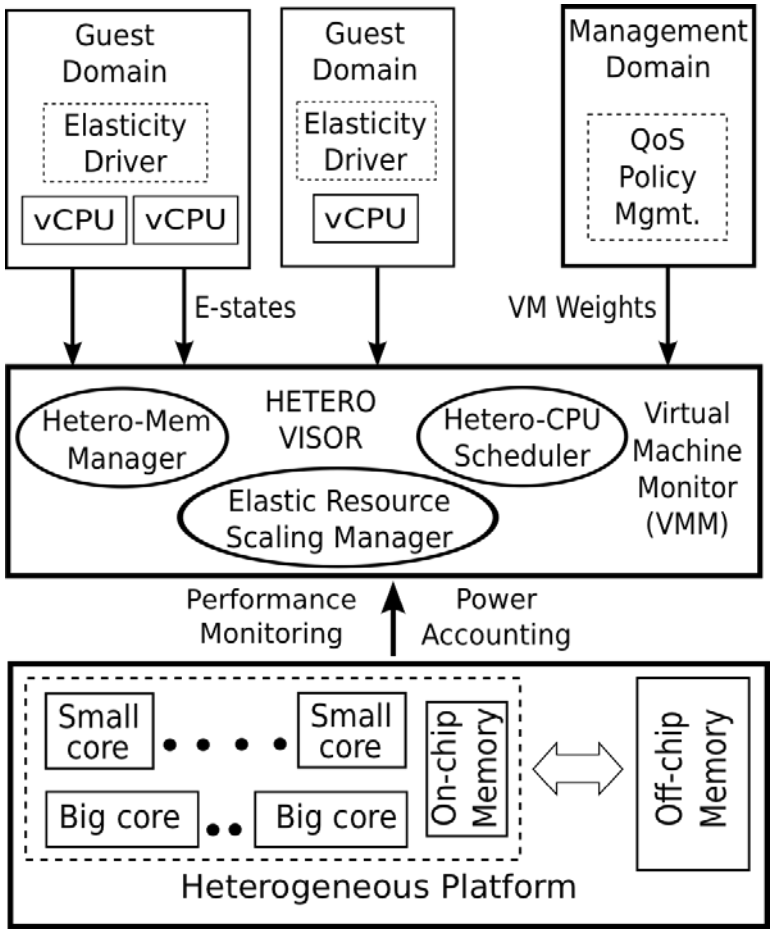




HeteroVisor Design

Concepts:

- Elasticity Drivers and E-States
- Scaling Manager
  based on E-States
  implemented in VMM
- co-exists with and leverages
  existing homogeneous managers



Elasticity Driver

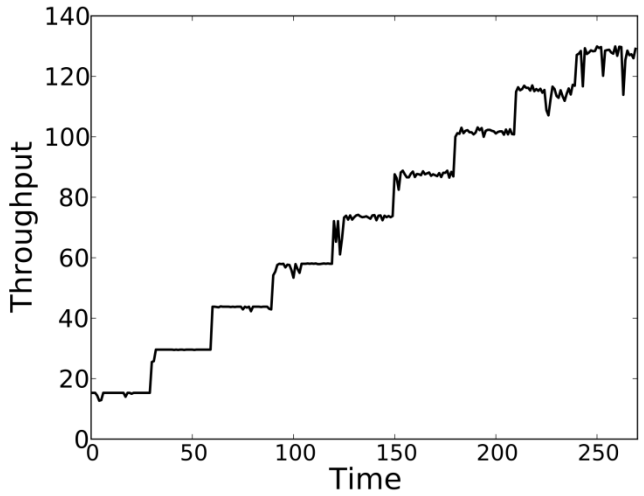
Scaling heuristic using activity factor and QoS metric

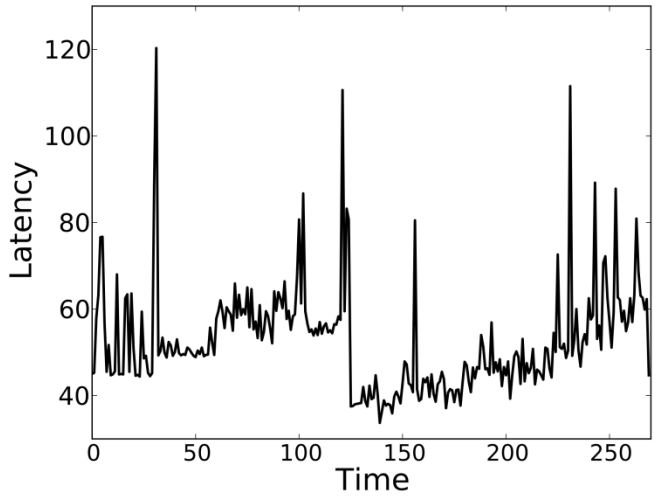
Use of history counters to avoid oscillations

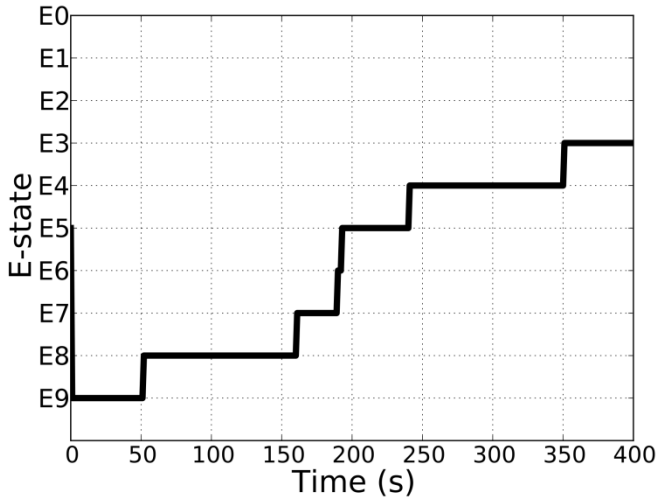
1. IF activity > actv_hi OR QoS < qos_lo
2. Scale up
3. ELSE IF activity < actv_lo AND QoS > qos_hi
4. Scale down
5. ELSE
6. No change

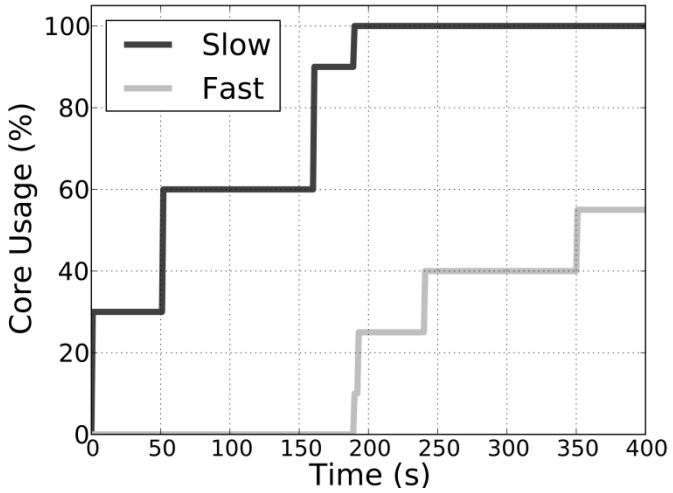
Elastic Scaling in Action

CPU results, only!









Current Status: Extension to 'Memory'

- Implementation finished, somewhat stable
- Evaluation with memcached using Twitter dataset (Initial results available for DRAM/Stacked DRAM)
- Now working to add NVM
  - Update hot page tracking
  - Update page migration methods
  - Obtain accurate NVM delays – diff. write/read
- Early submission version of paper available
- Questions and open issues: multi-resource elasticity drivers, dealing with VMs competing for resources

Conclusions and Next Steps

Demonstrated so far:

- Utility of elasticity driver idea
- Importance of joint CPU/Memory Mgt

Many open questions:

- Going beyond concurrent mgt of CPU and Memory – exponential explosion of possible configurations

  => Automation for elasticity drivers
- Useful global policies – QoS/Throughput/…

- Running multiple VMs – VM interference
- Dealing with NVM persistence property vs. NVM for capacity

Backup Slides

Managing Heterogeneity

SW-exposed

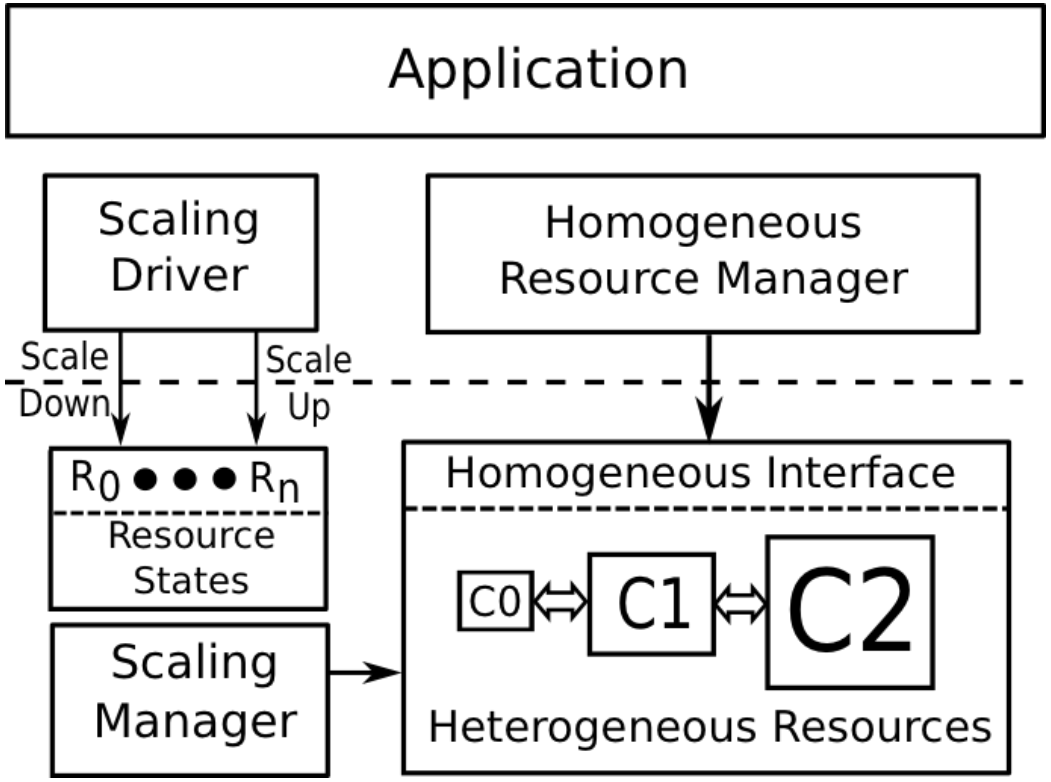
More flexibility
Too disruptive

SW-managed/hidden

HW-managed

Easier adoption
Too restrictive

Solution Approach



Advantages

- Decouples heterogeneity from resource management operations; hide heterogeneity for wider adoption
- Provides a way for applications to guide resource allocation to suit their needs
- Generic interface applicable across different resources/levels of heterogeneity

Background: Heterogeneous Memory Organization