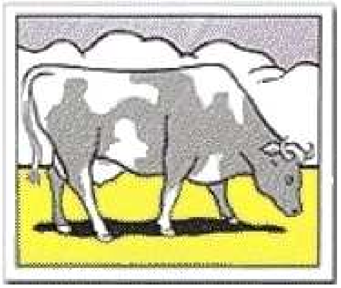
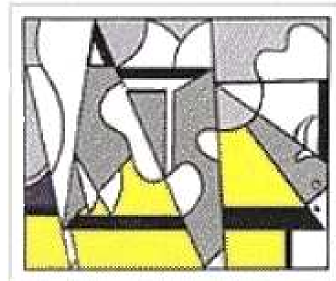
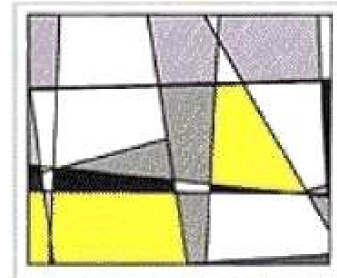




An Evaluation of Server Consolidation Workloads for Multi-Core Designs

Natalie Enright Jerger,
Dana Vantrease, Mikko H. Lipasti

University of Wisconsin - Madison

IEEE International Symposium on Workload Characterization
September 27, 2007

Motivation

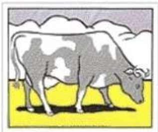
- Multi-core chips are here
- Many-core chips (10-100s of cores) are on the horizon
  - We need applications to evaluate future systems now
  - Programming parallel applications is not getting any easier
- Many-core systems can serve as consolidation platforms for multiple discrete applications

Server Farms

- Consequences of growing server farms
  - Higher management costs/overheads
  - Higher floor space and electricity consumption
  - Less reliable
  - Individual servers may be underutilized
- Virtualization facilitates consolidation of several physical servers onto a single high-end system
  - Reduces management costs/overheads
  - Increases overall utilization

Server Consolidation Overview

- Server workload characteristics well suited to multi-core architectures
  - Multithreaded
    - Take advantage of abundance of on-chip cores
  - Communication intensive
    - Low latency cache-to-cache transfers
    - On-chip sharing
- Potential interference between applications
  - Interconnect bandwidth
  - Memory controllers
  - Cache resources

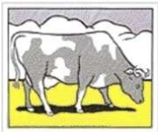


Opportunities with Server Consolidation Workloads

- Explore the relevance of server consolidation workloads to a variety of research communities
- Identify the interactions within server consolidation workloads
  - Open up new avenues of research
  - We focus on cache sharing within and between applications

Outline

- Motivation
- Workloads
- Cache Configurations
- Evaluation Methodology
- Results
- Conclusions



Workloads

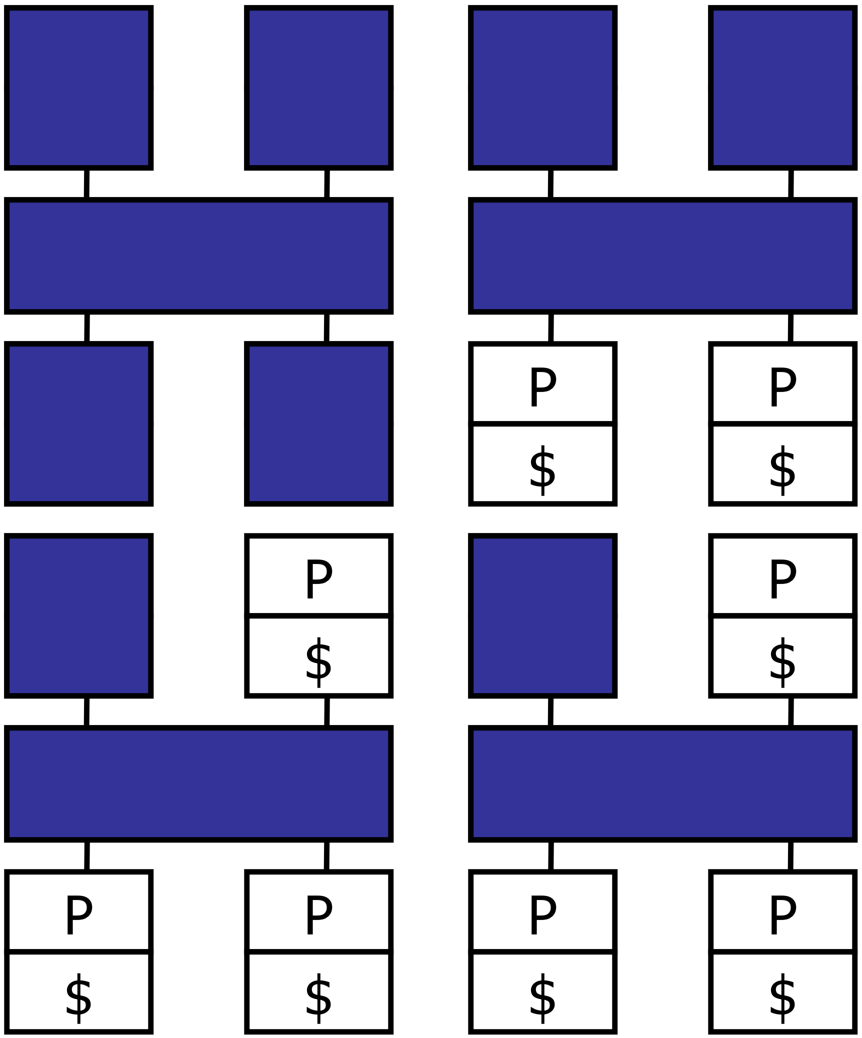
Workloads	Description	Setup	Execution
SPECjbb	Order processing application for wholesaler. Emphasizes the middle-tier business logic and performance of Java-based middleware.	3-tier client-server with 6 warehouses	6400 requests with 15 seconds of warm-up time
SPECweb	World-wide web server	3-tier, Zeus Web Server 3.3.7	300 HTTP requests
TPC-H	Decision Support System	IBM DB2 v6.1	Query #12 (shipping modes and order priority) on 512 MB database with 1GB of memory
TPC-W	Web commerce modeling online bookstore	IBM DB2 v6.1	Browsing mix for 25 web transactions

Workload Characteristics

	Percent of accesses resulting in $-to-$ transfer			Memory Footprint
	Clean	Dirty	Clean+Dirty	
SPECjbb	13%	2%	15%	Large
SPECweb	49%	3%	52%	Medium
TPC-H	30%	39%	69%	Small
TPC-W	34%	3%	37%	Large

- Many different combinations possible

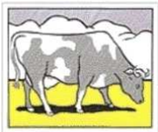
Target CMP



- Affinity
  - Maximizes sharing
- Round Robin/Interleaved
  - Maximizes available cache
- Round Robin – Affinity Hybrid
- Random
  - Result of load balanced scheduling

Cache Configurations

- Vary degree of sharing
  - Private
  - Partially shared
  - Shared
- Vary assignments of threads to cores and caches
  - Affinity
  - Round Robin
  - RR-Affinity Hybrid
  - Random



Simulation Methodology (1)

- PHARMsim
  - Full-system simulator
  - Runs AIX OS
- Methodology
  - Each application isolated in its own Virtual Machine
    - Runs its own copy of OS
    - Statically assign a portion of global memory – private address space

Simulation Methodology (2)

- Each VM:
  - Loads 4-processor checkpoint
    - Snapshot of the workload
    - System already booted and warmed
    - Alleviates overhead of booting the OS
  - Runs both user and operating system code
    - Fixed number of transactions
    - Ensures that the same set of transactions are run

Simulation Methodology (3)

- At startup the system:
  - Loads four VMs
    - Each VM runs a pre-made 4-processor checkpoint
    - Allows a mix and match of VMs for a variety of workload combinations
  - Assigns a VM to a subset of cores
    - All of a VM's threads run within their assigned domain
    - Assignment is maintained throughout simulation

Workload Mixes

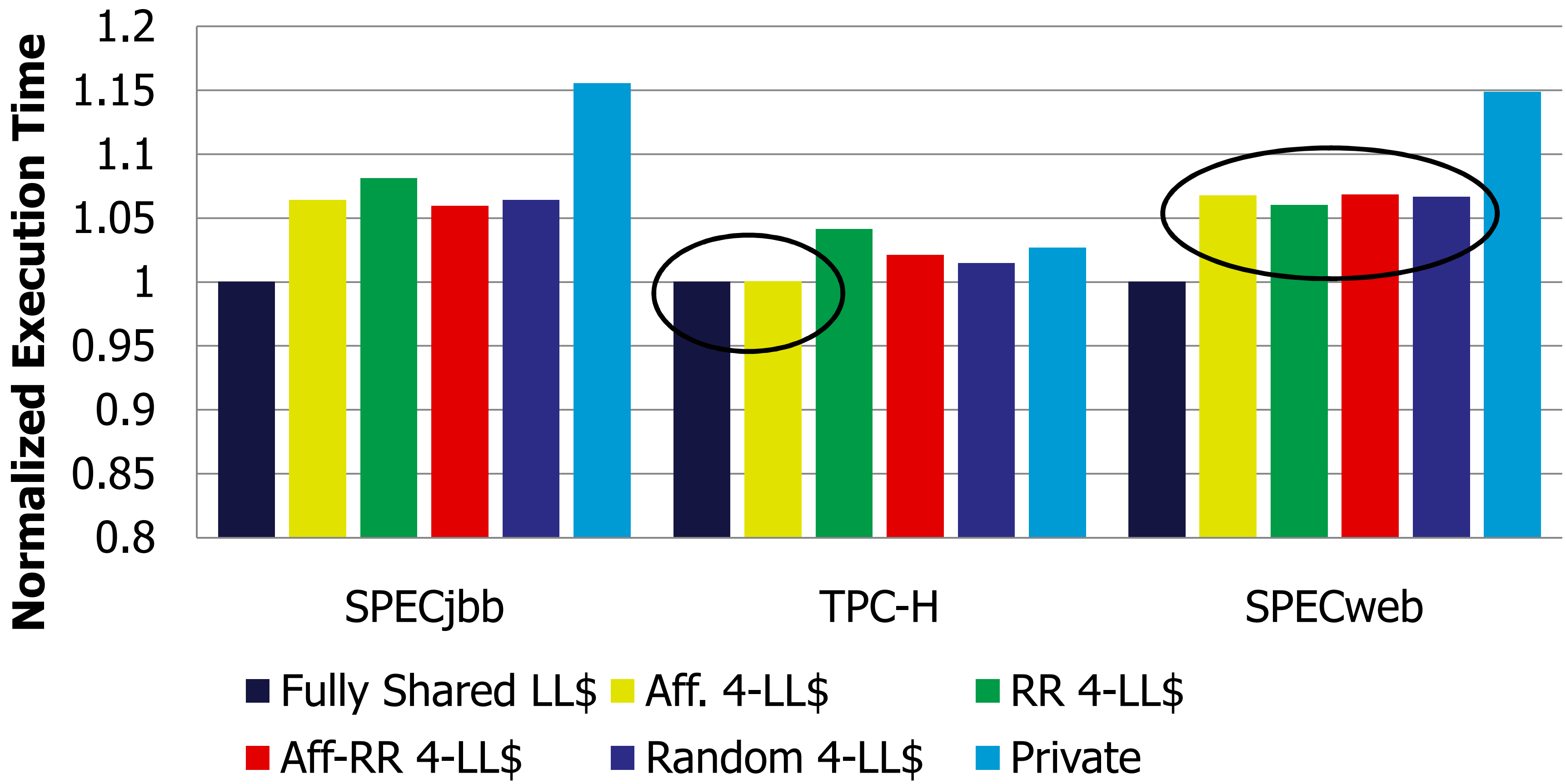
Heterogeneous Mixes	
Mix 1	TPC-W (3) + TPC-H (1)
Mix 2	TPC-W (2) + TPC-H (2)
Mix 3	TPC-W (1) + TPC-H (3)
Mix 4	SPECjbb (3) + TPC-H (1)
Mix 5	SPECjbb (2) + TPC-H (2)
Mix 6	SPECjbb (1) + TPC-H (3)
Mix 7	SPECjbb (3) + TPC-W (1)
Mix 8	SPECjbb (2) + TPC-W (2)
Mix 9	SPECjbb (1) + TPC-W (3)

- Homogeneous Mixes of TPC-W, TPC-H, SPECjbb, and SPECweb

Simulation Parameters

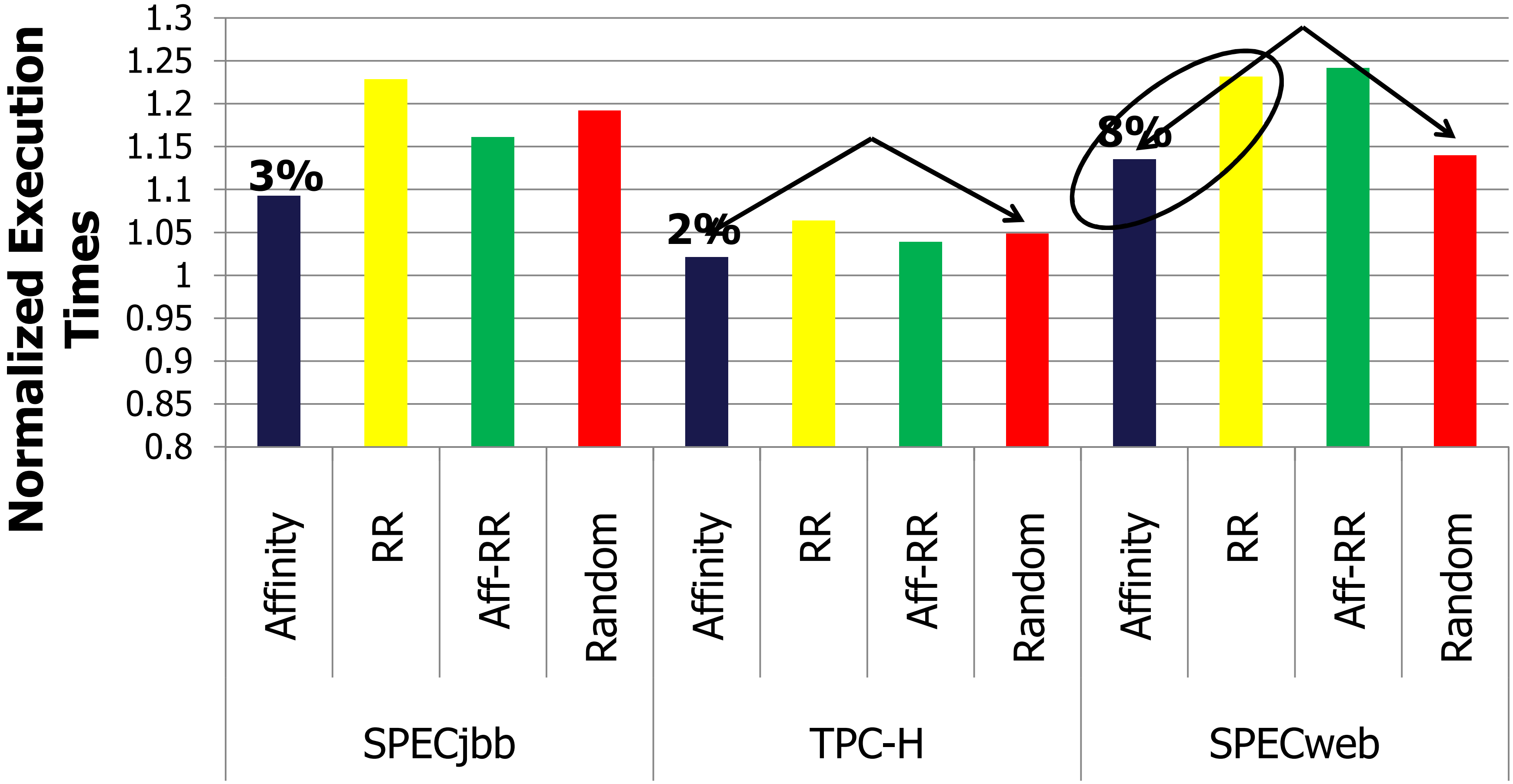
Machine Configuration	
Core	16 single-threaded, in-order
Interconnect	2-D Packet Switched Mesh
L0 Cache	Split I/D, 8 KB each
L1 Cache	Private 64 KB
L2 Cache	Shared, 16MB
Memory Latency	150 cycles
Cache Coherence	SGI Origin Directory Protocol
Thread to Core Assignment	Round Robin, Affinity, Round Robin-Affinity Hybrid, Random

Isolated Workload Performance



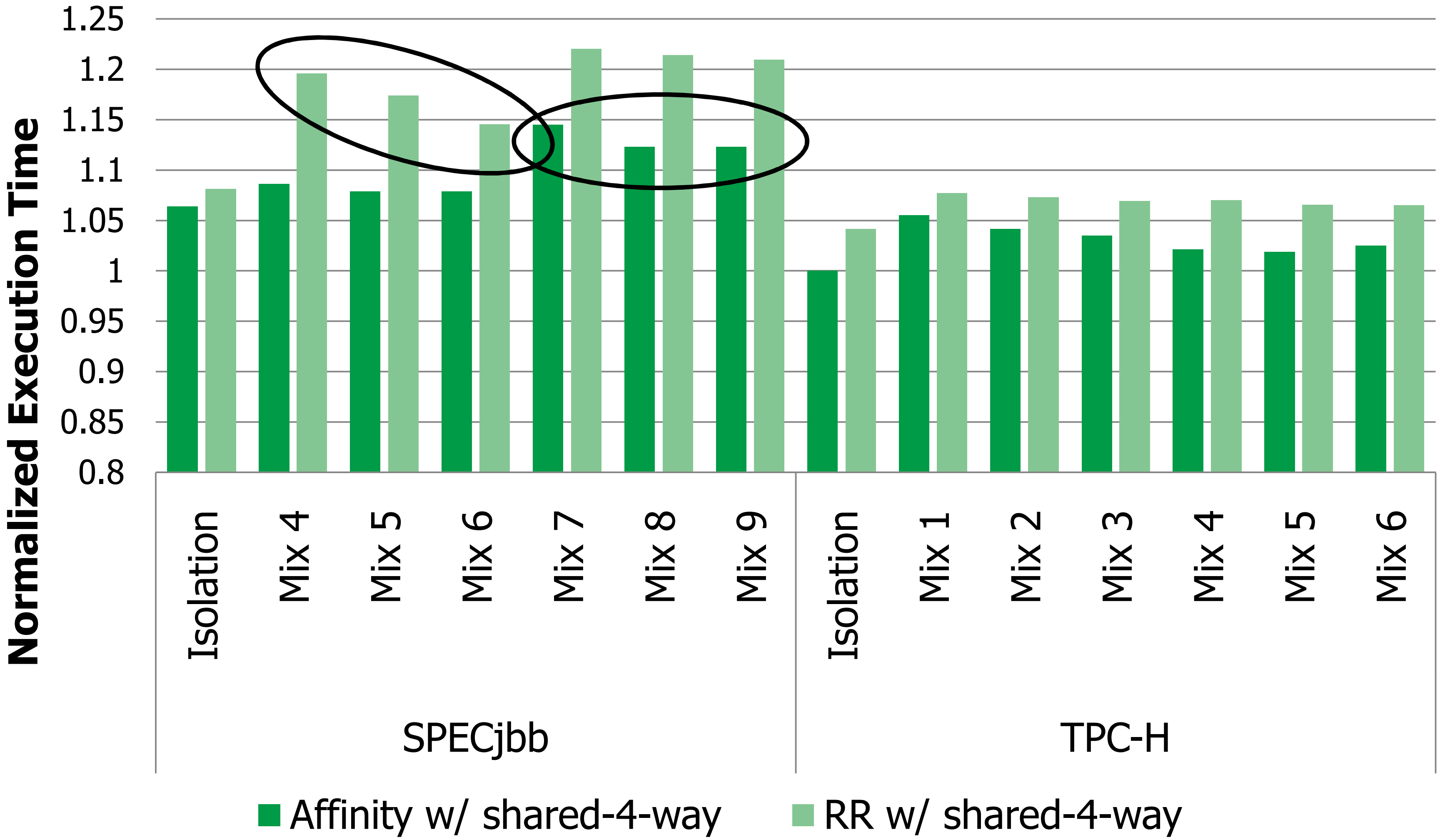
- Affinity mapping *almost* always the best choice

Homogeneous Mix Performance



- Performance differences between placement policies become more pronounced in consolidated environment

Heterogeneous Mix Performance



Implications and Related Work

- Quality of Service
  - [Iyer, Sigmetrics 07], [Nesbit, MICRO-39]
- Coherence Protocols
  - Virtual Hierarchies [Marty, ISCA 07]
- Methodology
  - SMT Methodology

Conclusions

- Server consolidation workloads provide a framework on which to evaluate many-core architectures now
  - Illustrate this by looking at degrees of cache sharing
- Open up new and interesting avenues of research
  - Architecture, Operating Systems, Virtual Machines

Questions?