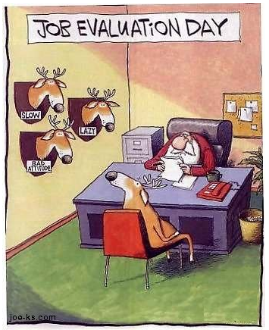
Advanced Computer Architecture

Evaluating Systems

Fall 2016



Pejman Lotfi-Kamran

Adapted from slides originally developed by Profs. Hill, Hoe, Falsafi and Wenisch of CMU, EPFL, Michigan, Wisconsin

Where Are We?

Fr	Sa	Su	Mo	Tu
	27-Shahrivar		29-Shahrivar	
	3-Mehr		5-Mehr	
	10-Mehr		12-Mehr	
	17-Mehr		19-Mehr	
	24-Mehr		26-Mehr	
	1-Aban		3-Aban	
	8-Aban		10-Aban	
	15-Aban		17-Aban	
	22-Aban		24-Aban	
	29-Aban		1-Azar	
	6-Azar		8-Azar	
	13-Azar		15-Azar	
	20-Azar		22-Azar	
	27-Azar		29-Azar	
	4-Dey		6-Dey	

- This Lecture
  - Evaluation
- Next Lecture:
  - Coherence

Commercial Server Software

Commercial Server Software

- Primary market for multiprocessor systems
- Examples:
  - Database systems: Oracle, DB2, SQLServer, PostGres, MySQL
  - Business apps: SAP, BAAN, PeopleSoft
  - Data analysis: MapReduce, large scale graph processing
  - Web-servers
    - Static content
    - Dynamic content: database integration + business logic
    - Web 2.0: user-supplied content
  - Infrastructure apps: memcached, J2EE

Why study database apps?

- They are economically important
- They share characteristics of many other apps (filesystems, web search, etc.)
- The vendors have spent a lot of time optimizing (generally, they won't have silly bottlenecks)

Key characteristics

- Large, complex, monolithic software systems
- Designed for MP systems
  - Clusters (distributed databases)
  - Shared Memory
- Subsumes many OS functions
  - File system
  - Scheduling and multi-threading
  - Memory management
- Designed for high reliability (ACID properties)
  - Atomicity: a transaction happens or doesn't
  - Consistency: the state of the DB remains consistent
  - Isolation: transactions are independent
  - Durability: once performed, transactions are permanent

How are they different from Sci Apps?

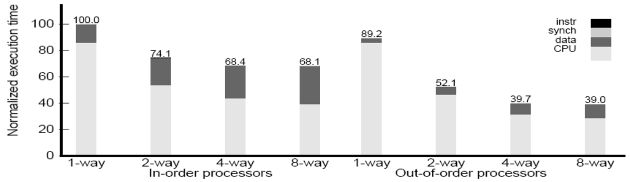
- Requires tuning: knowledge-intensive, difficult
- Competitive market:
  - deliberate obfuscation/ benchmark gaming
- Large instruction footprints (I$ matters)
- Huge data footprints (TLBs matter)
- Weird access types (non-cacheable, etc.)
- Latency, not bandwidth bound
- Dynamic memory allocation, garbage collection
- More pointer-chasing, fewer arrays
- No single obvious "working set"
  - multiple working sets with varying temporal locality
- Unpredictable sharing patterns
- Data & lock contention

Standardized Benchmarks

Transaction Processing Council (TPC)
- Strict scaling, disclosure, auditing rules
- Running these for real is hard: big hardware, 20-50 engineers, months of effort
- Running them in simulations is also hard: scaling, non-determinism

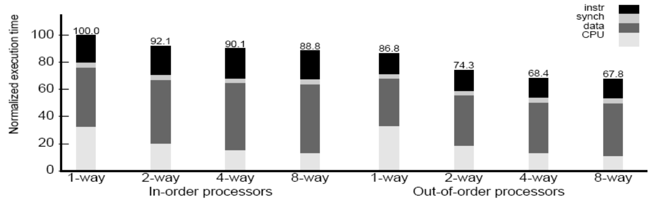
Two flavors of benchmark
- Online transaction processing (OLTP): TPC-C
  - Lots of small transactions
  - Lots of locking, concurrency, I/O; memory-latency bound
- Decision support system (DSS): TPC-H
  - Large, complex read-only queries
  - Often compute bound (given enough disks)
  - Highly parallel, data partitioning, parallel operators

DSS – Impact of ILP



- Lots of ILP
- Multiple issue helps, but OoO helps more
- L2 hits account for most data stalls

OLTP – Impact of ILP



- Instruction cache misses & synch are now issues
- Less ILP, but 2-way still helps a lot
- Coherence (dirty) & DTLB misses most read stalls

Emerging Scale-Out Workloads
(Ferdman et al., ASPLOS 2012)

Scale-Out Datacenters

Massive Scale

- $100+ M investment
- 5-20 MW power budget

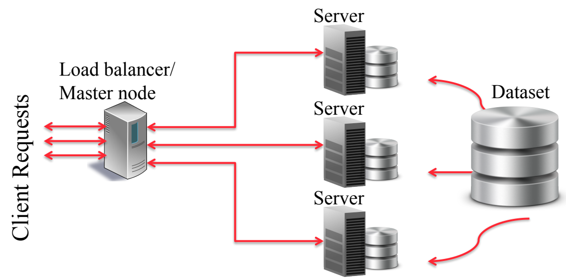
Why?

- Massive data sets, complex queries
- Global demand

Applications

- Web search, media streaming, social connectivity
- Scale-out by design

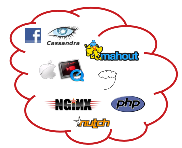
Characteristics of Scale-Out Apps



- Many independent requests/tasks
- Huge dataset split into shards
- Minimal communication among servers

How Efficient are Today's Servers?

- Created benchmark suite
  - Diverse set of cloud workloads
  - Quantified high-level behavior
- Studied off-the-shelf hardware
  - Used performance counters
  - Identified needs of cloud apps



Modern CPUs don't match needs of cloud apps

Cloud Suite 1.0
(released @ parsa.epfl.ch/cloudsuite, tutorial @ ISCA 2012)

Hardware

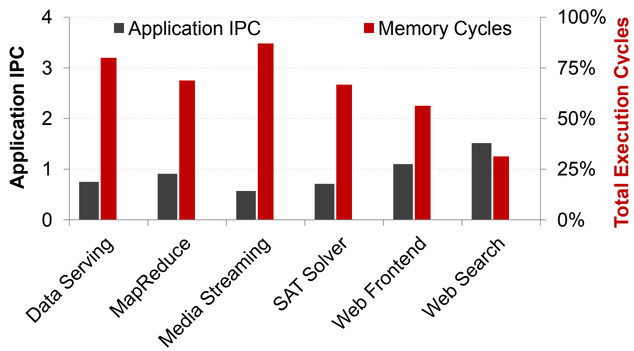
Dell PowerEdge M1000e

Dell Blades M610

Two Intel x5670 2.9GHz 6-core, 12MB LLC

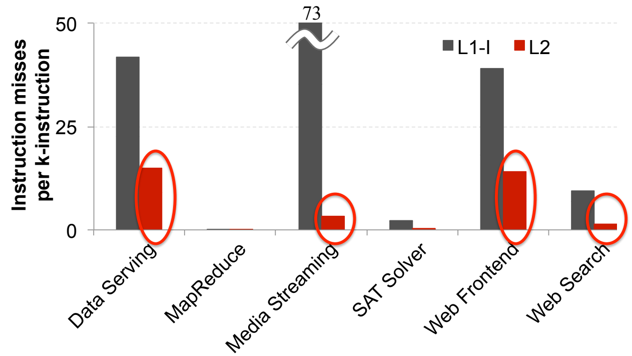
24GB RAM

Core Inefficiencies



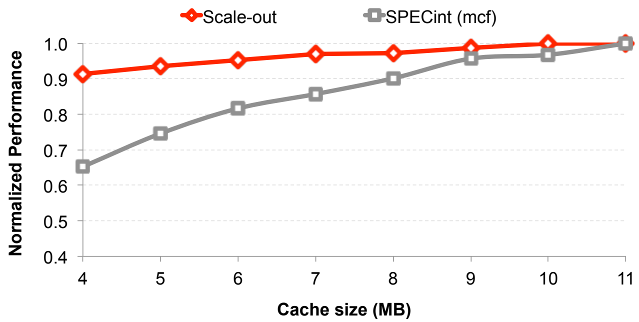
Execute ~1 instruction per cycle

Instruction-Fetch Misses



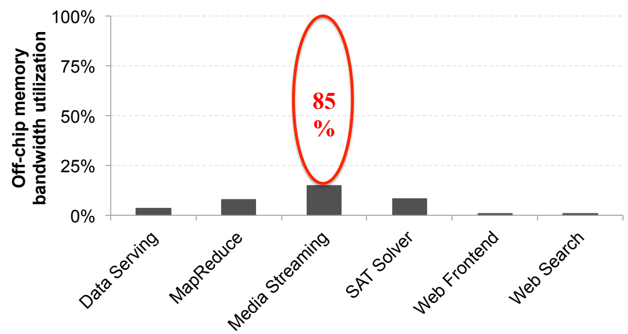
Suffer severe i-cache miss penalties

LLC Sensitivity



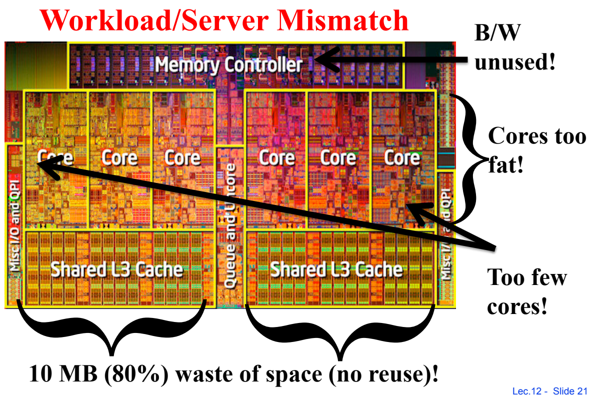
Minimal performance from large LLC

Off-chip Memory Bandwidth



Off-chip BW severely underutilized

Clearing the Clouds in a nutshell [ASPLOS 2012]



Summary

- Accurate metrics to evaluate designs
- Don’t forget the laws!
- Tools
- Wide spectrum of parallel workloads