THE GRAINGER COLLEGE OF ENGINEERING
SIEBEL SCHOOL OF COMPUTING AND DATA SCIENCE

CS533: Architecture Interaction with Databases

Josep Torrellas
Presenting: Antonis Psistakis

04/16/2026

ILLINOIS

Memory System Characterization of Commercial Workloads

Barroso et al, ISCA-98

Context & Motivation

DBs and webservers → Largest and fastest-growing segment of the multiprocessor server market

- Disk subsystem innovations + CPU/mem speed gap → memory system design becomes a critical factor for such workloads
- Most current designs: optimized to perform well on scientific and engineering workloads
  - Not ideal for commercial applications
- Key Issues
  - Lack of info on performance requirements of commercial workloads
  - Lack of available applications for widespread study
  - Most representative apps: too large & complex to study
  - Fast moving target

Workloads

OLTP

- Banking system: Transactions update balance in randomly-selected account
- Small computation
- 4 tables updated per transaction → high I/O

DSS

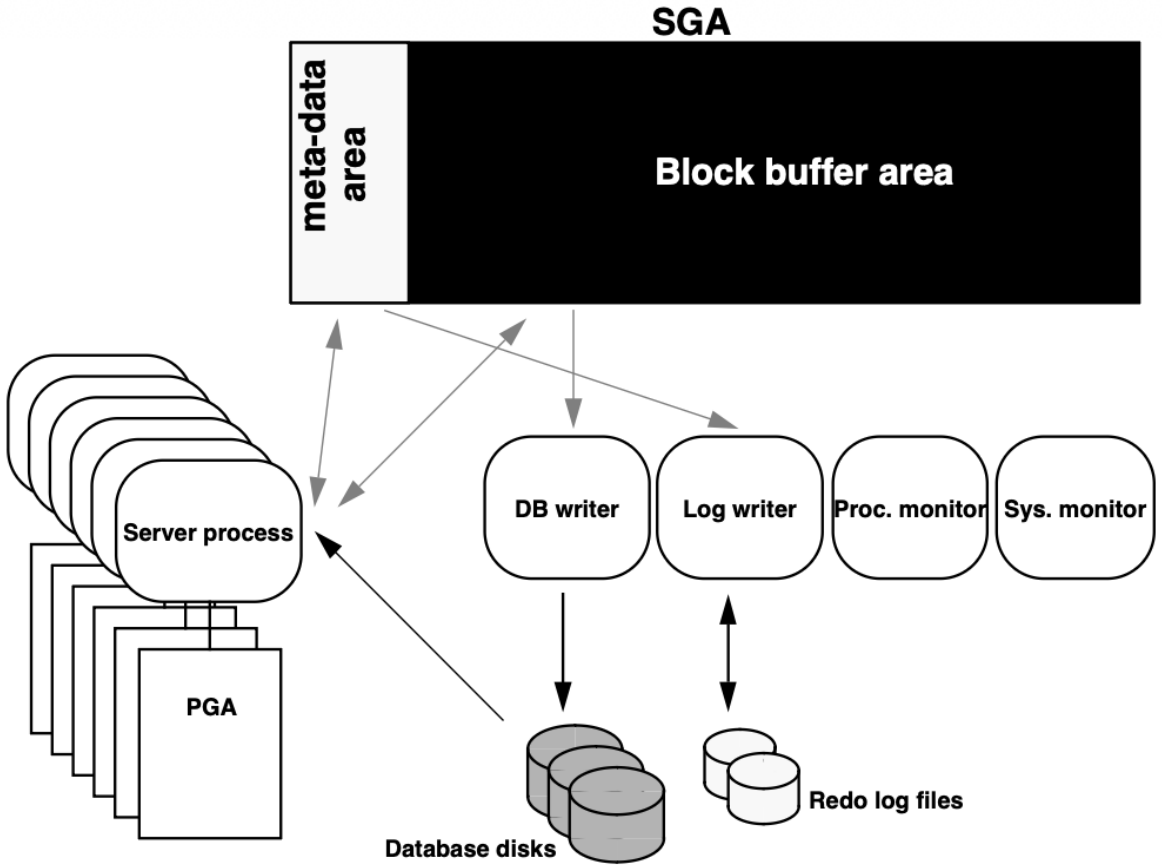
- Read-only queries
- More complicated
- Queries can be parallelized

Web Index Search

- Like DSS
- Read-only

Oracle Database Engine

- Shared data in-memory with multiple server processes
- High context switch frequency (~25K Instructions)
- 7-8 processes per processor



DSS Workload

Decision Support Systems

- Business analysis
- Long running SQL queries
- Spanning a large fraction of the DB
- Read-only, amenable to intra-query parallelism

	Tables Used	Selects	Joins	Sorts/ Aggr.	User	Kernel	Idle
Q1	1	1 FS	-	1	94%	2.5%	3.5%
Q4	2	1 FS, 1 I	-	1	86%	4.0%	10%
Q5	6	6 FS	5 HJ	2	84%	4.0%	12%
Q6	1	1 FS	-	1	89%	2.5%	8.5%
Q8	7	8 FS	7 HJ	1	82%	5.0%	13%
Q13	2	1 FS, 1 I	1 NL	1	87%	4.0%	9.0%

SQL operations

- Select, Join, Sort, Aggregate

Leverage multiple processes per processor

- Parallelize each query across each process
- Hide I/O latency
- Low kernel utilization

Alta Vista: Web Index Search

- Searching the internet before it was cool! (Google)
- Large search index, 200GB
- The entire DB is memory mapped
- Addresses page faults – how?
  - Leverage multiple threads: During search a thread often gives up the processor, before its quantum due to high page fault rate

Experiments

Hardware platform

- Server with 4 300-MHz processors
- Hardware event counters
- Latency to memory: 260ns

Simulator platform

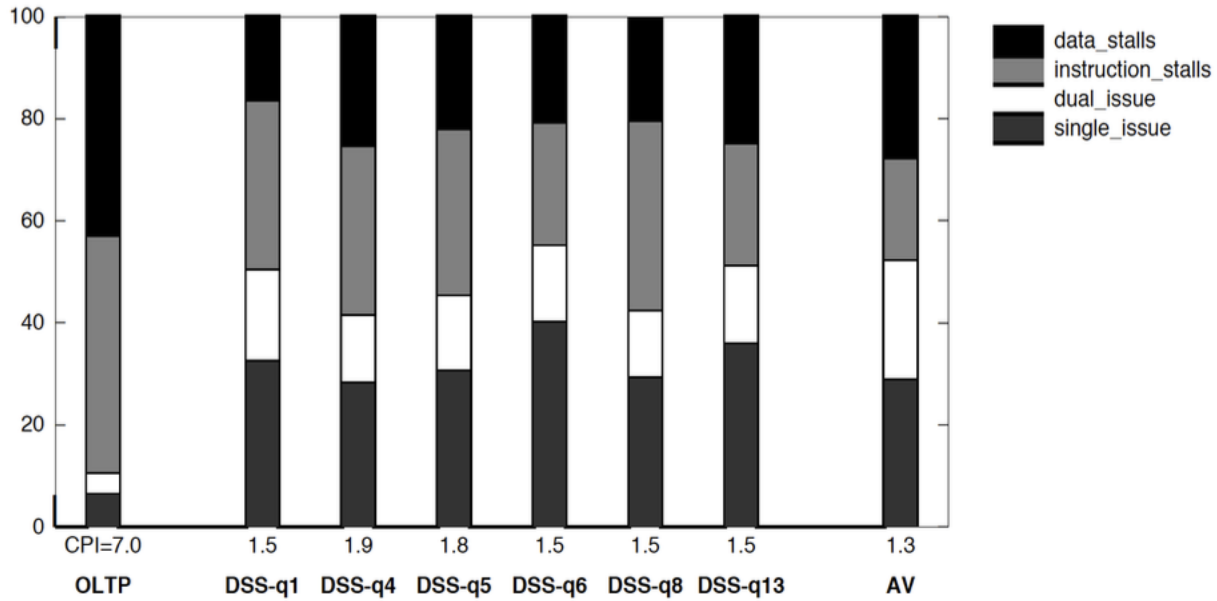
- SimOS: Simulates App + OS

Cache Hierarchy

- L1/First level – 8KB
- L2/Secondary Cache (Scache) – 96KB
- L3/Board Level Cache (Bcache) – 2MB
- Line size
  - First level: 32B
  - Lower levels: 64B
- Cache-to-cache transfer: 125~

Monitoring Results

Cycle Breakdown



OLTP

- CPI very high → Poor perf

DSS

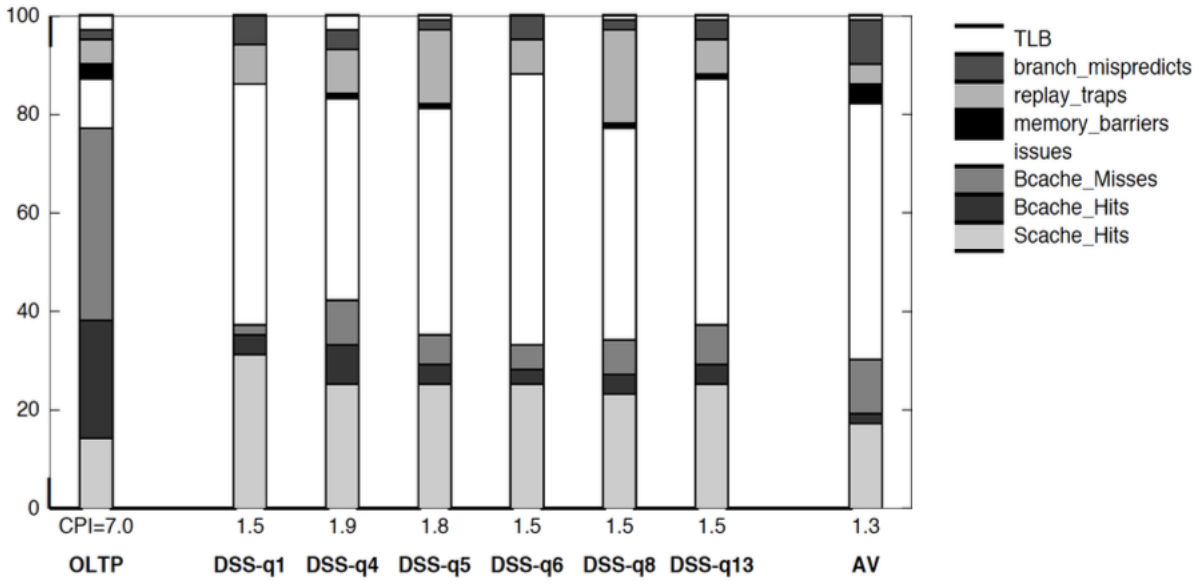
- Lower CPI (more efficient)

AltaVista

- Best performance

Monitoring Results

Detailed Cycle Breakdown



OLTP

- L3 misses → workload overwhelms caches
- Also important: L1 and L2 misses (L3/Bcache hits)

DSS & AltaVista

- Better locality

Cache Characterization

	OLTP	DSS-Q1	DSS-Q4	DSS-Q5	DSS-Q6	DSS-Q8	DSS-Q13	AltaVista
Icache (global)	19.9%	9.7%	8.5%	4.6%	5.9%	3.7%	6.7%	1.8%
Dcache (global)	42.5%	6.9%	22.9%	11.9%	11.3%	11.0%	12.4%	7.6%
Scache (global)	13.9%	0.8%	2.3%	1.0%	0.6%	1.0%	1.0%	0.7%
Bcache (global)	2.7%	0.1%	0.5%	0.2%	0.2%	0.3%	0.3%	0.3%
Scache (local)	40.8%	3.6%	10.7%	5.7%	3.9%	6.0%	6.1%	7.6%
Bcache (local)	19.1%	13.0%	21.3%	23.9%	30.7%	27.9%	31.3%	41.2%
Dirty miss fraction	15.5%	2.3%	2.2%	10.6%	1.7%	8.4%	3.3%	15.8%

OLTP

- High miss rate; multiple levels
- Dirty Misses (cache-to-cache): ~15%

DSS

- L1 (global) misses -- unsuccessful at keeping working set
- L2 can hold the state
- Low/No Dirty Misses

AltaVista

- Effective on-chip caches (global)

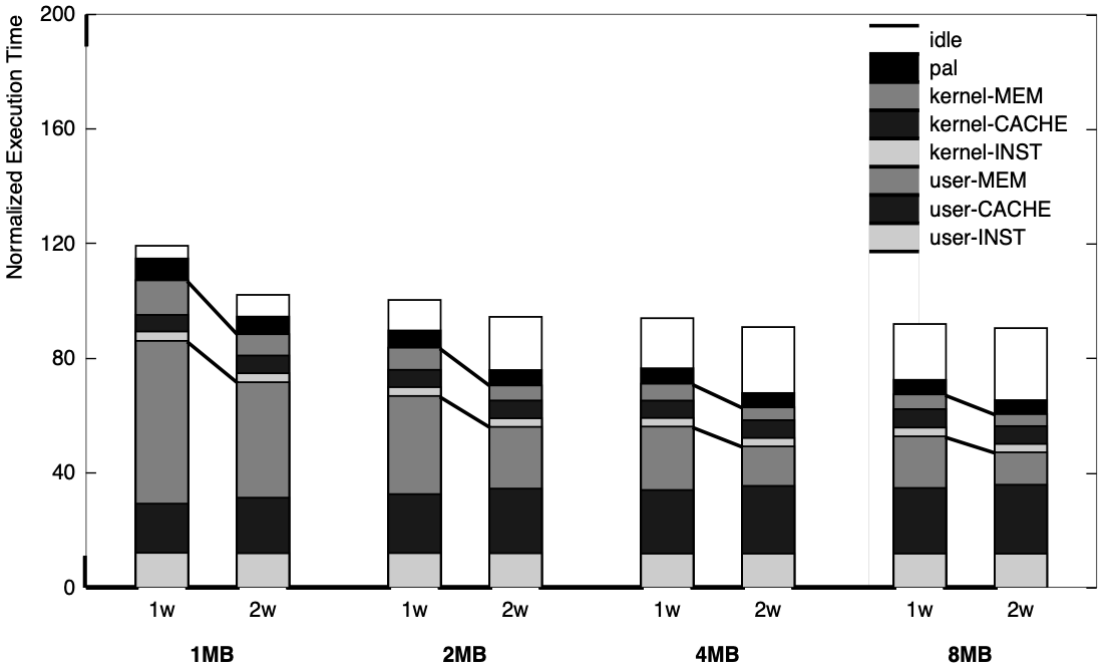
Simulation Methodology

- L1-D and L1-I: 32 KB
- L2: 2MB

	OLTP			DSS Q6		
Event	SimOS	HW	error	SimOS	HW	error
Instructions	99.5	100.0	0.5%	101.1	100.0	1.1%
Bcache misses	2.25	2.39	5.7%	0.15	0.14	9.1%
User cycles	66.5%	71%		88.5%	89%	
Kernel cycles	16.9%	18%		2.2%	2.5%	
PAL cycles	6.0%			1.0%		
Idle cycles	10.7%	11%		8.4%	8.5%	

Sharing Patterns

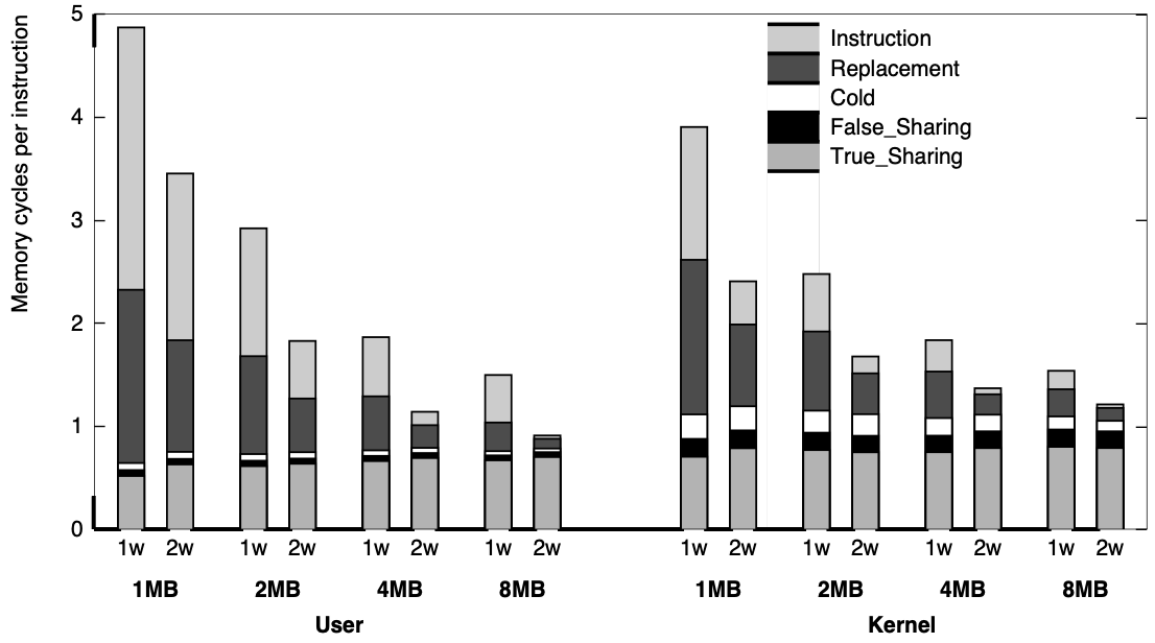
OLTP Execution Time



- Kernel much smaller than user
- Both user + kernel benefit from higher assoc. caches
- Cache and memory stall still important
- Idle time increases

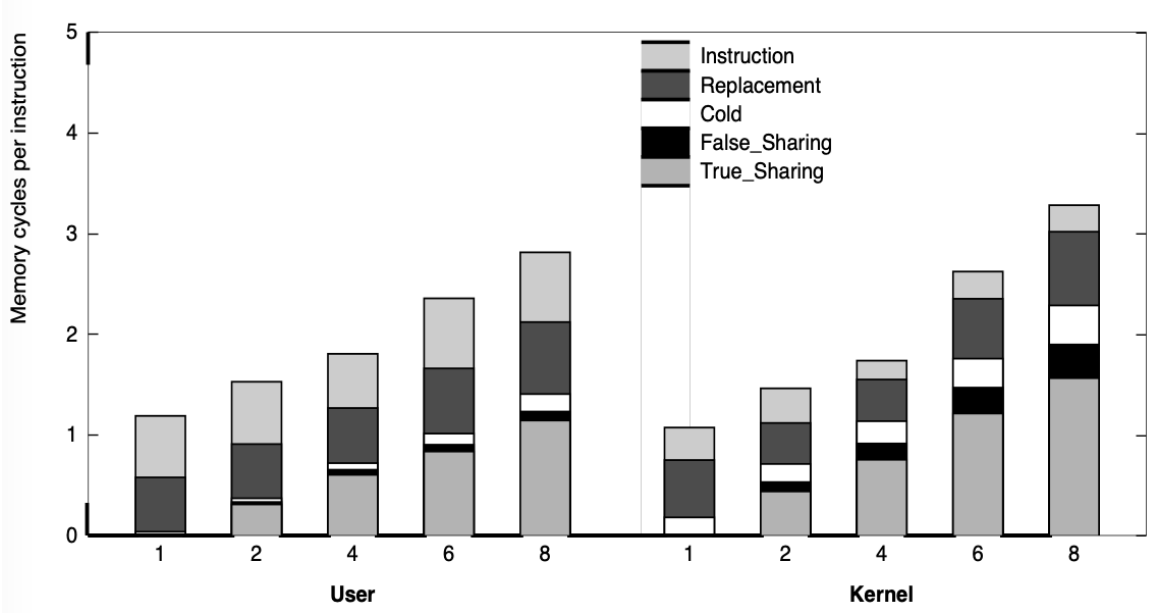
Sharing Patterns

Contribution of Bcache Misses (OLTP)



- Benefits from larger & higher assoc. caches
- Communication misses dominate
  - Most due to true (not F) sharing

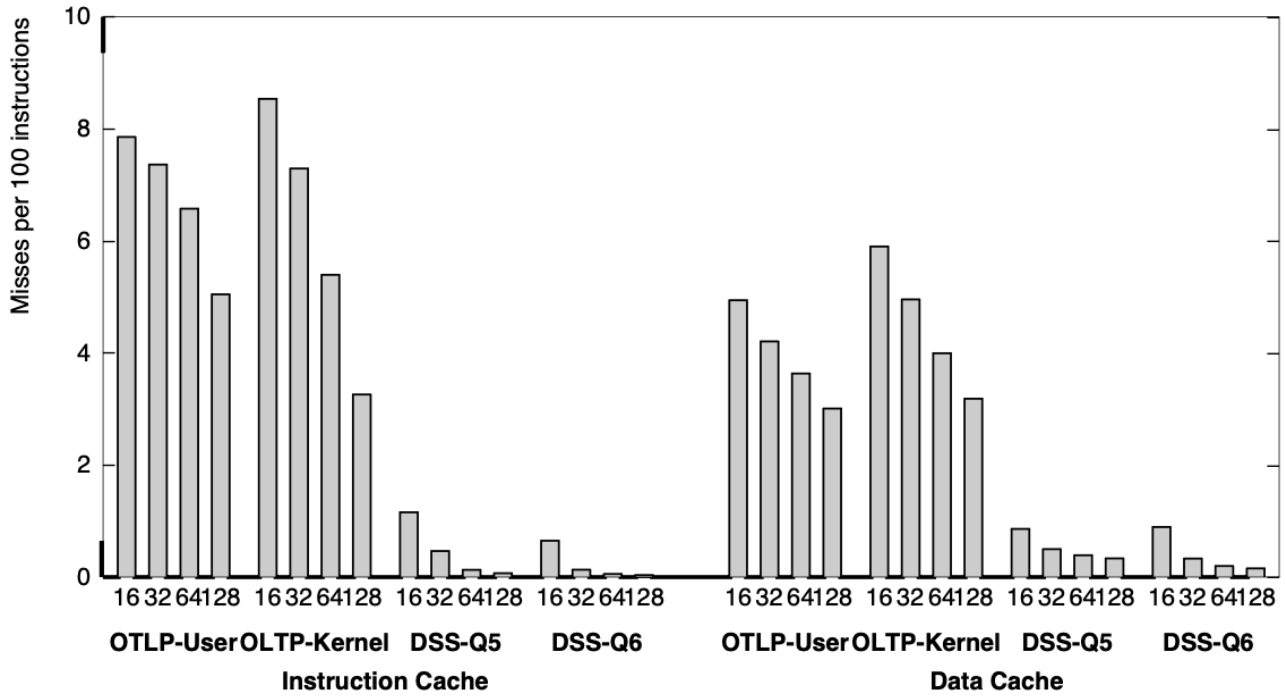
Number of Processors



If P goes up:

- Ratio of true to false sharing does not change
- Communication misses are primarily affected

Cache Sizes



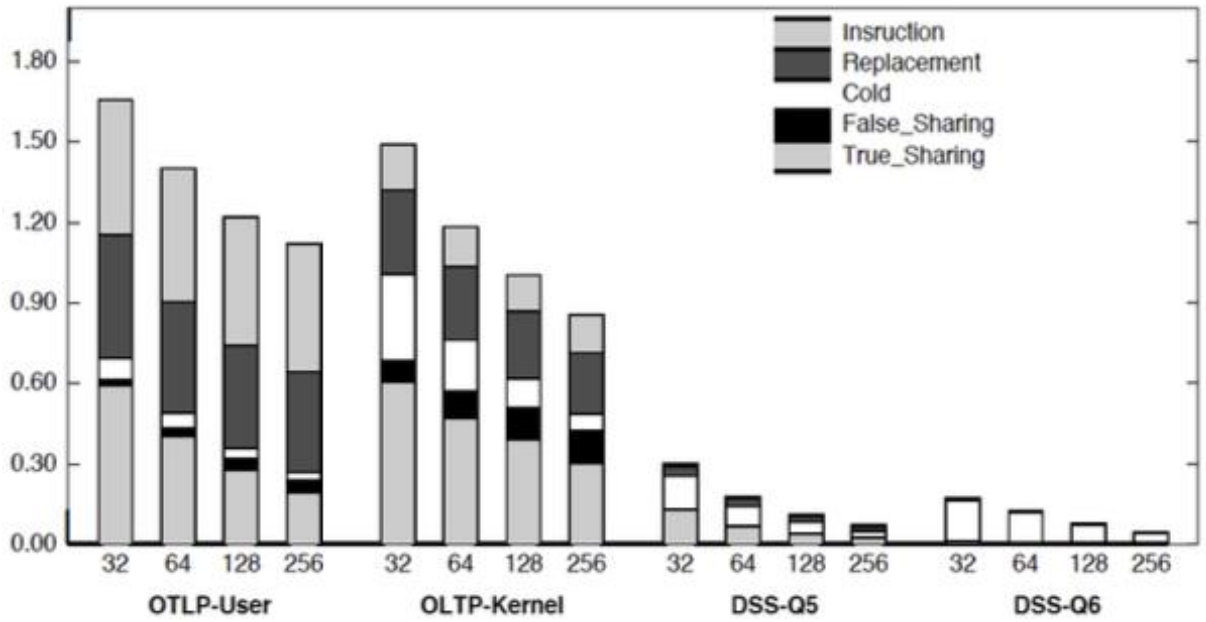
On-chip cache miss-rates for OLTP and DSS for split I/D caches (16KB-128KB).

- Larger on-chip caches good for DSS and OLTP
- Larger on-chip caches cache most misses in DSS effectively

Sensitivity to Cache Line Size

Changing line size of the L3

- Data communicated among processors (true sharing) has good spatial locality
- Cold misses also spatial locality
- No impact on replacement misses



Summary

- OLTP
  - I and D locality only captured with large L3 caches
  - High communication miss rate (dirty misses)
- DSS and AltaVista
  - Sensitive to the size and latency of L1
  - Less sensitive to off-chip caches sizes/latencies
- Workloads require different server designs

Thank You

Questions?