Sytem Implications of Aggressive CMT Processors

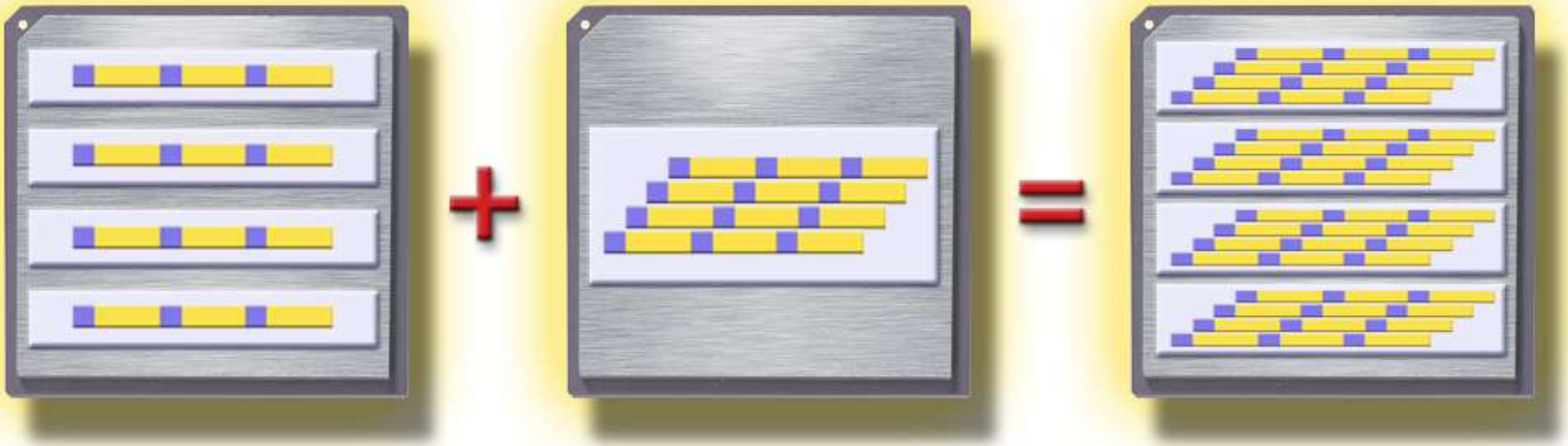
Rick Hetherington
Distinguished Engineer
Chief Architect of Niagara Processors and Systems

Sun Microsystems, Inc., Sunnyvale, CA

Agenda

- Define CMT and Give an example
- A typical CMT system
- System Implications of CMT
  - Performance
  - Power
  - Reliability
- Conclusion

CMT Definition and Example



CMP
(Chip MultiProcessing)

n cores/processor

Fine Grain Interleaved Threading

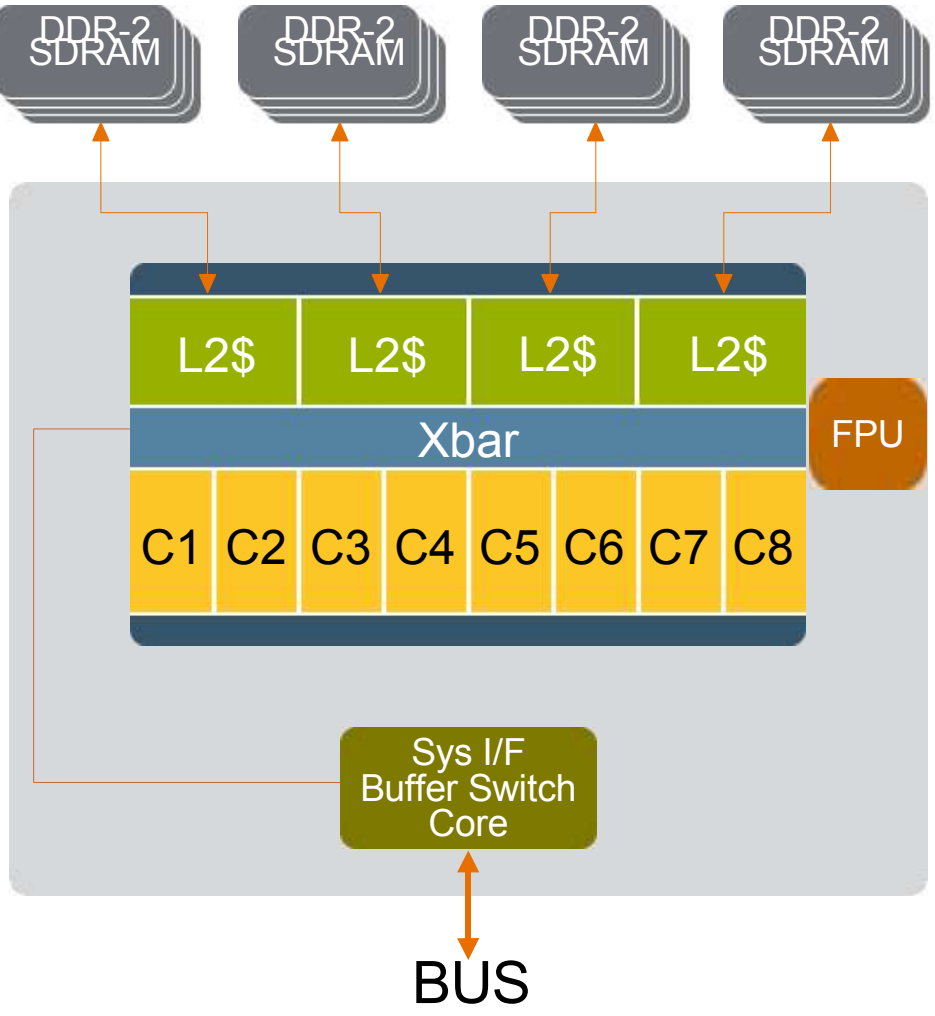
m threads/core

CMT
(Chip MultiThreading)

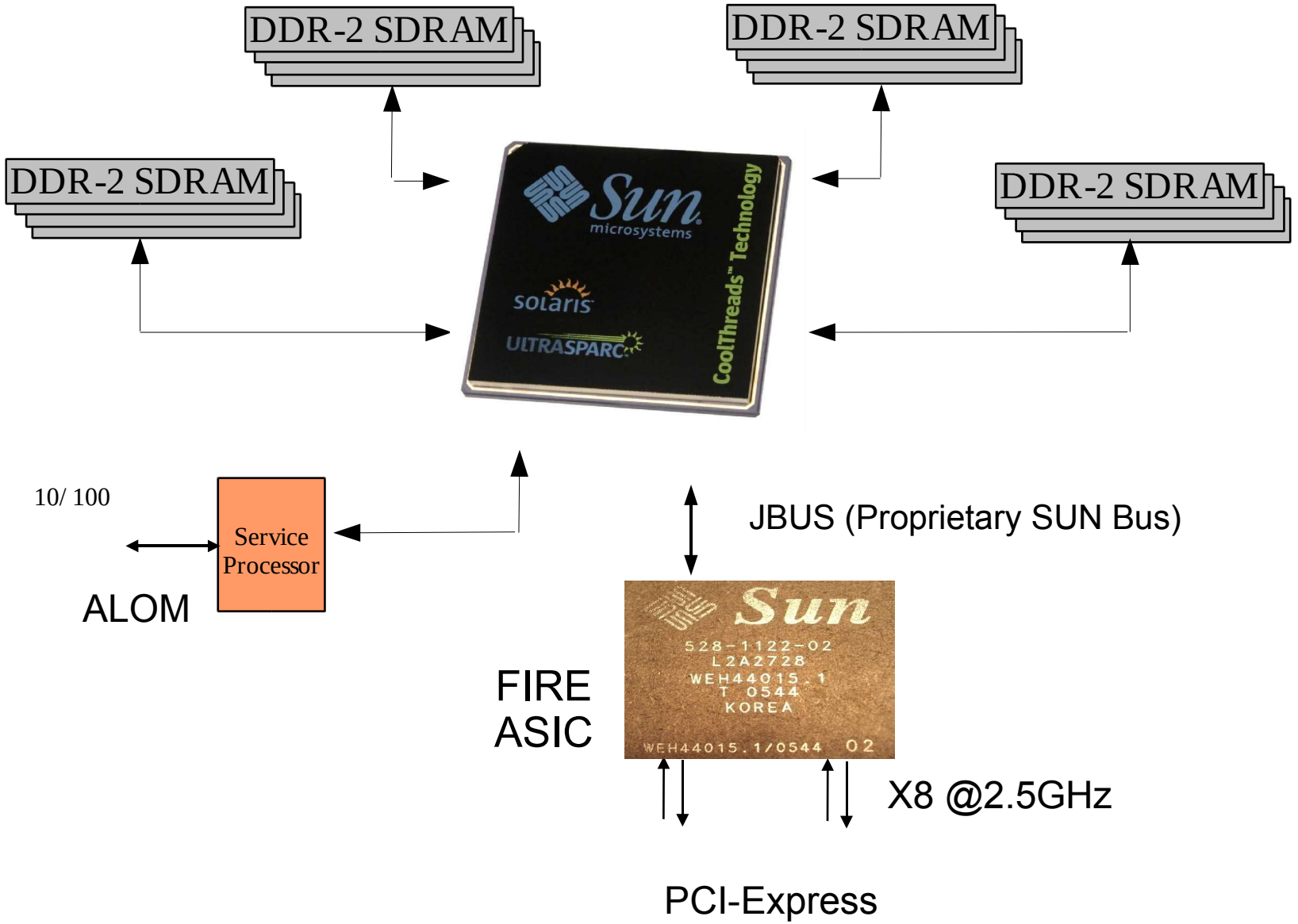
n x m threads/ processor

CMT Example..... US T1

- 8 Cores and 4 threads/core
- 32 Total Threads
- Full Crossbar to Memory/IO
- 4 way Banked L2 Cache
- Memory controller/bank
- One FPU
- I/O through proprietary bus



Logical UST1 System



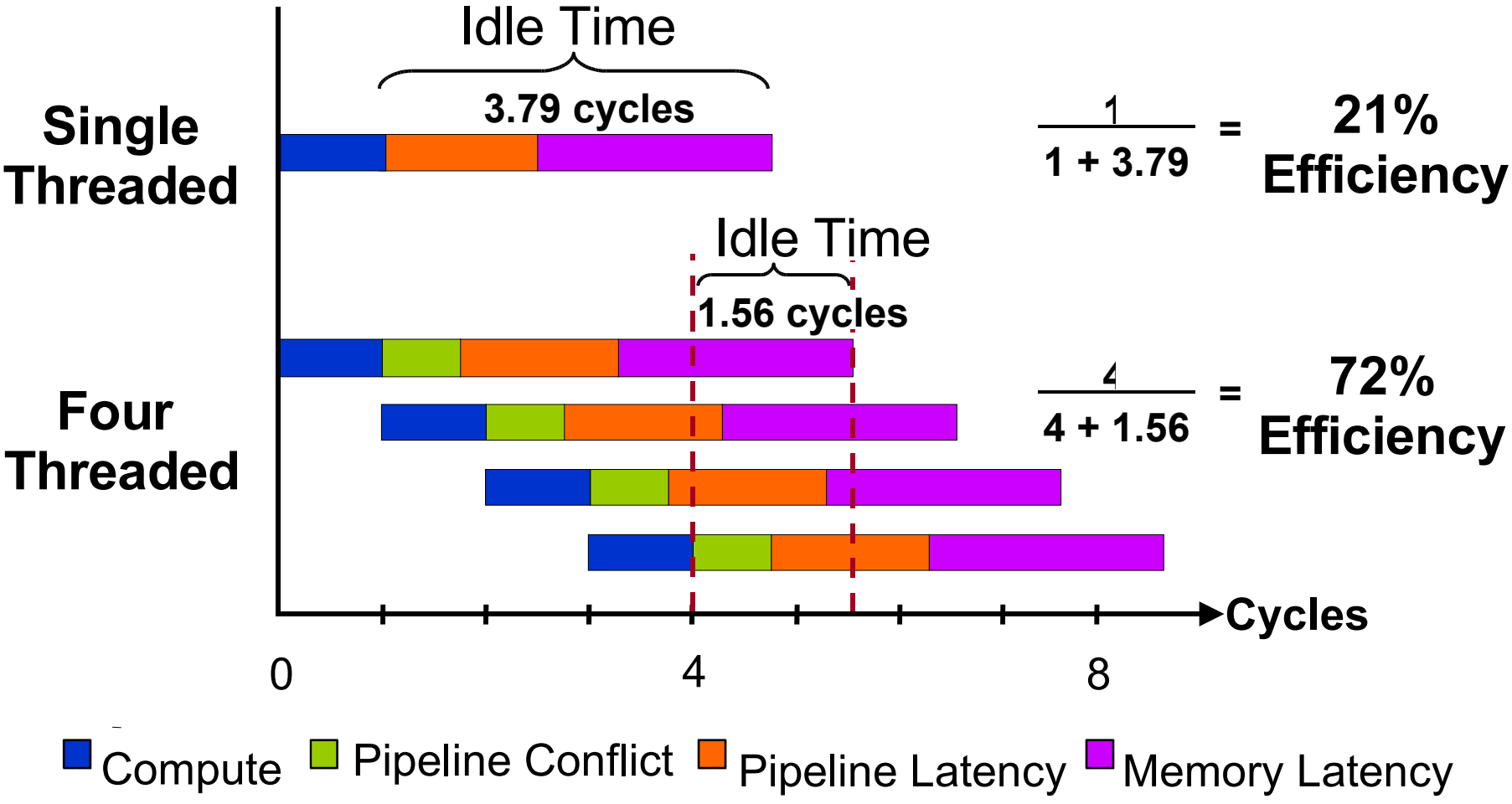
UltraSparc T1000 System

Sun Fire T1000

Performance Benefits of CMT

- **Fine Grain Threaded Core**
  - High aggregate IPC across 4 threads
  - Efficient use of core die area with high utilization rates
- **On Chip Coherence Fabric**
  - Fabic is 100 times shorter (mm as opposed to Ms)
  - No cache to cache transfers- can be up to 800 nsec
  - Spin-Lock Wait loops are radically reduced if not eliminated
  - Linear scale across 32 threads for most applications
- **On Chip Memory Controllers**
  - No North Bridge
  - Lowest possible Memory Latency

SpecJBB Execution Efficiency



CMT Performance

N1 Benchmark Performance 3/6/2006

Benchmark	OLTP	SpecJBB2005	Specweb2005		JappServer2004 +
	T2000	T2000	T2000-Ecomm	T2000-Bank	T2000
IPC/Core	0.41	0.67	0.58	0.65	0.56
I$ Miss	8.23%	2.09%	4.93%	4.58%	7.71%
D$ Miss	7.19%	3.54%	4.28%	3.40%	5.06%
L2 LD Miss	1.08%	0.84%	0.61%	0.36%	0.59%
L2 I Miss	0.15%	0.01%	0.09%	0.07%	0.34%
Path-len	1.41 Mil.	122 k			
Thruput	171 ktpm	63 kbops	19500 conns.	16000 conns.	616 Jops

Notes:

+ With Bea

CMT System Performance

Benchmark	Performance
Lotus iNotes(2)	19000 users
SPECjbb2005	63089 BOPS
SPECweb2005B	16000 sessions
SPECweb2005E	19500sessions
SPECjAp-d	615 JOPS
SPECjAp-m	3329 JOPS

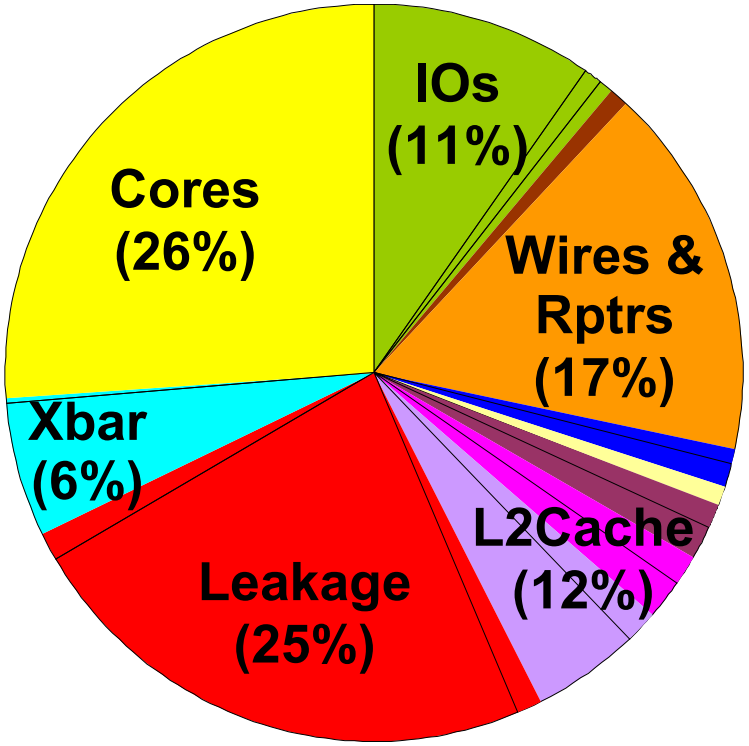
CMT Power Advantage

63W @ 1.2GHz / 1.2V
< 2 Watts / Thread

- Simple Single Issue In-Order 6 stage Pipe
- No Speculation, No Branch Prediction
- Fully static design
- Fine granularity clock gating for datapaths (30% flops disabled)
- No North Bridge – Memory Controllers On Chip
- No External Coherence Fabric with Power Consuming Snoops and Cache-to Cache xfers

Chip Power

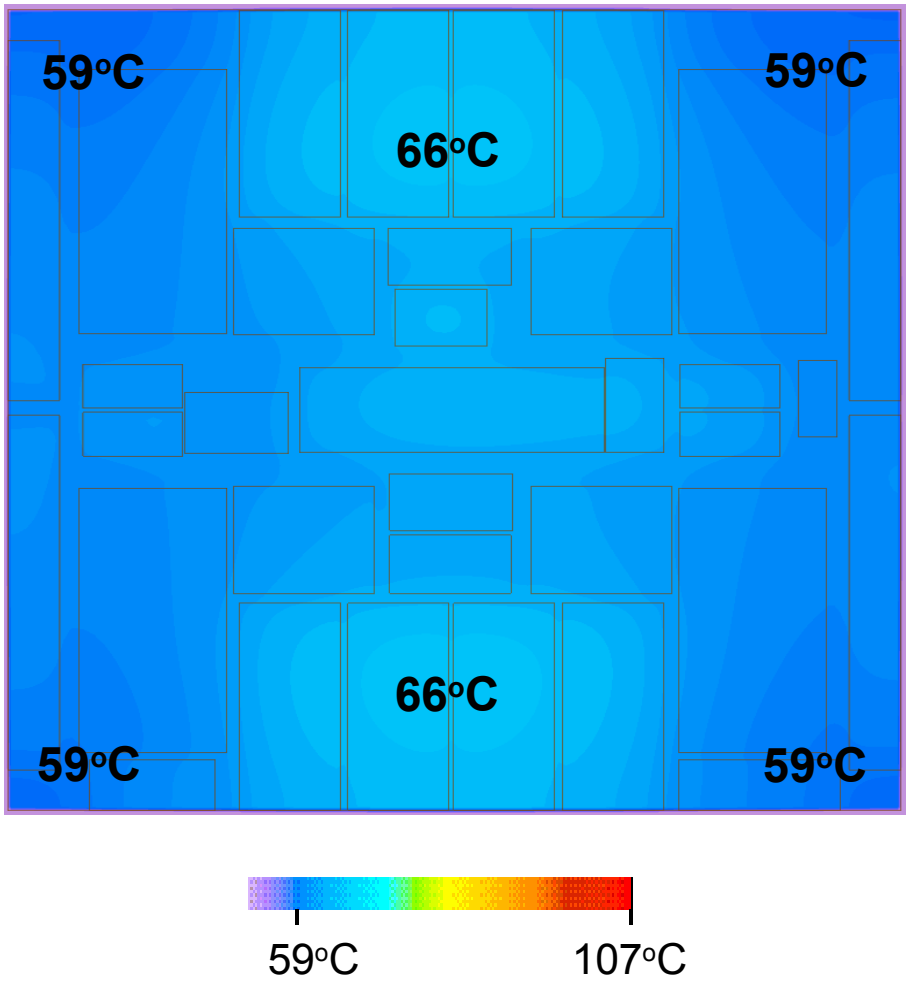
63W @ 1.2GHz / 1.2V
< 2 Watts / Thread



- Fully static design
- Fine granularity clock gating for datapaths (30% flops disabled)
- Lower 1.5 P/N width ratio for library cells
- Interconnect wire classes optimized for power x delay
- SRAM activation control

SPARC Cores
L2Data
Crossbar
IOs
Leakage
L2Tag Unit
Global Clock
Misc Units
Wires & Rptrs
L2 Buffer Unit
Floating Point

CoolThreads™ Advantages



- Improved reliability with lower and more uniform junction temperatures
  - Increased lifetime
  - Total failure rate reduced by ~8X (vs 105°C)
- Optimized performance/ reliability trade-off
  - Frequency guardbands due to CHC, NBTI, etc. reduced by > 55%
  - Reduced design margins (EM/NBTI)
  - Less variation across die

CMT System Power/Performance

T2000

BM	*Performance*	*Avg Power(W)*
Lotus iNotes	19000 users	312
SPECjbb2005	63089 BOPS	298
SPECweb2005B	16000 sessions	329
SPECweb2005E	19500 sessions	330
SPECjAp-d	615 JOPS	320
SPECjAp-m	3329 JOPS	321

T1000

Web Consol	4900 Users	182
SPECjbb2005	51400 BOPS	166

Legal Disclosures

(1)SPECweb2005: Sun Fire T2000 (8 cores, 1 chip) 14001.SPEC, SPECweb reg tm of Standard Performance Evaluation Corporation. Sun Fire T2000 results submitted to SPEC. Other results from www.spec.org as of December 6, 2005.Sun Fire T2000 server power consumption taken from measurements made during the benchmark run.

(2)SPECjappSERVER 2004: Sun Fire T2000 (8 cores, 1 chip) 615.6 JOPS @Standard. Sun FireT2000 (8 cores, 1 chip) 436.71 JOPS @Standard Sun Fire T2000 (8 cores, 1chip) 3328.80 JOPS @Standard. SPEC, SPECjappSERVER reg tm of Standard Performance Evaluation Corporation. Sun Fire T2000 results submitted to SPEC. Sun Fire T2000 server power consumption taken from measurements made during the benchmark run..

(3) SPECjbb2005: Sun Fire T1000 Server (1 chip, 8 cores, 1-way) 51,540 bops, 12,885/JVM -Sun Fire T2000 (1 chip, 8 cores, 1-way) 63,378 bops, 15845 bops/JVM SPEC, SPECjjj reg tm of Standard Performance Evaluation Corporation. Sun Fore T1000 results submitted to SPEC. Other results as of 12/6/2005 on www.spec.org. Sun Fore T1000 and T2000 server power consumption taken from measurements made during the benchmark run.

(4) NotesBench R6iNotes: Sun Fire T2000 (1x1200 Mhz UltraSPARC T1, 32GB), 4 paritions, Solaris[TM] 10, Lotus[R] Domino 7.0, 19,000 users, $4.36 per user, 16,601 NoteMark tpm, 400 ms avg rt. Sun Fire T2000 server power consumption taken from measurements made during the benchmark run..

CMT System Reliability And Availability

Availability and Mean Time Between System Interrupt are summary metrics

They are driven by

- Rate
  - How often do faults occur?
- Robustness
  - Do faults cause system outages?
  - Can the system be repaired online?
- Recovery
  - How quickly can we return to nominal operation?

Rate

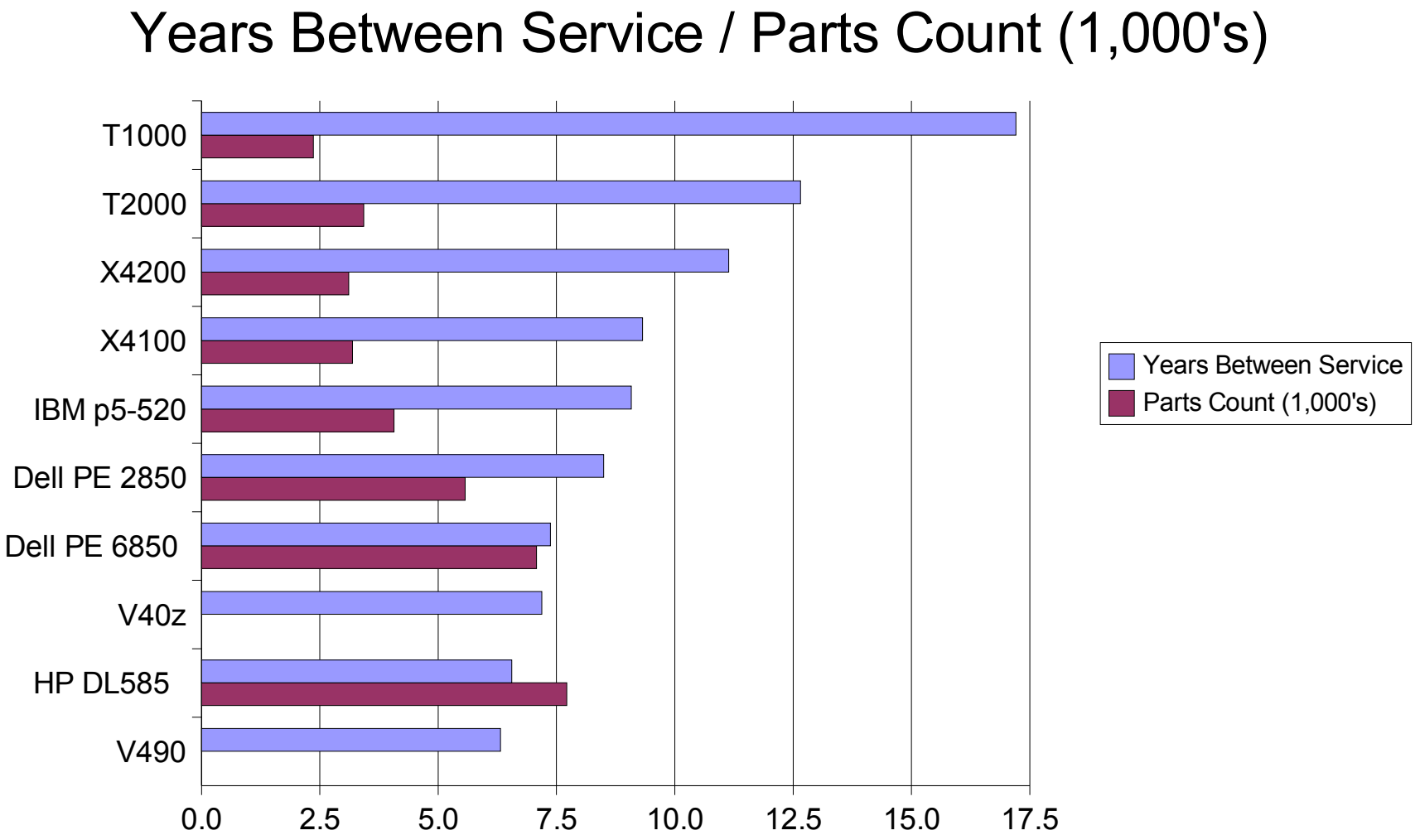
- Rate is driven by
  - How many parts are used
    - Redundancy increases rate
    - High levels of integration reduce rate
- The lower the rate, the more reliable the component
- CMT in UST1 enable systems with lowest part count

CMT System Reliability And Availability

Robustness

- Robustness increases with redundancy
  - CMT Cores are redundant and can be offlined
- Error detection and correction
  - Parity with retry, CRC, SEC-DED
  - Not Unique to CMT
- Failure prediction based on correctable error counts
  - Solaris 10 Fault Management Architecture
  - Diagnose Engines
    - offline threads with repeated reconverable errors

CMT System Reliability And Availability



Conclusion

- Parallelism is now and will be for some time the main focus of processor architects.
- Chip-Multi Threading is the most effective means to achieve high levels of parallelism
- CMT processors yield throughtput 4-8 times OOO single thread designs....
- CMT processors deployed in systems have a significant power advantage over conventional non-Threaeded designs- generally 2-4 times less comsumed power.
- CMT systems have greatly reduced system parts count as well as high level of internal redundancy for robust crash avoidance.

Acknowledgements

- Performance
  - Cong Fu, Denis Sheahan, Pallab Bhattacharya
- Technology
  - Sonia Leon, David Greenhill
- Benchmarking
  - Chris Boire, Brad Carlile
- RAS Engineering
  - Tarek Durbas