

Chip Multi-Threading Keeps the Data Center Cool

David Greenhill

Distinguished Engineer

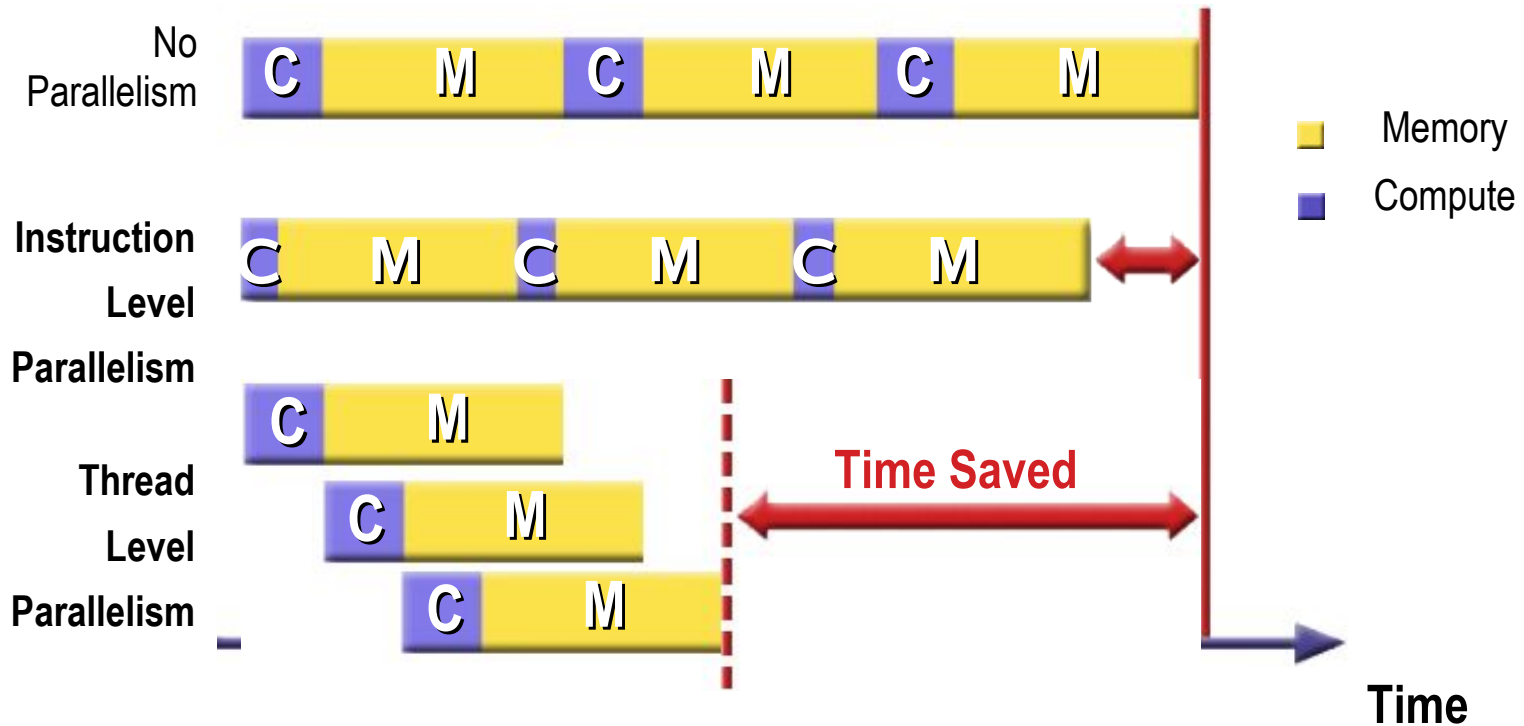
Sun Microsystems Inc

david.greenhill@sun.com

Introduction

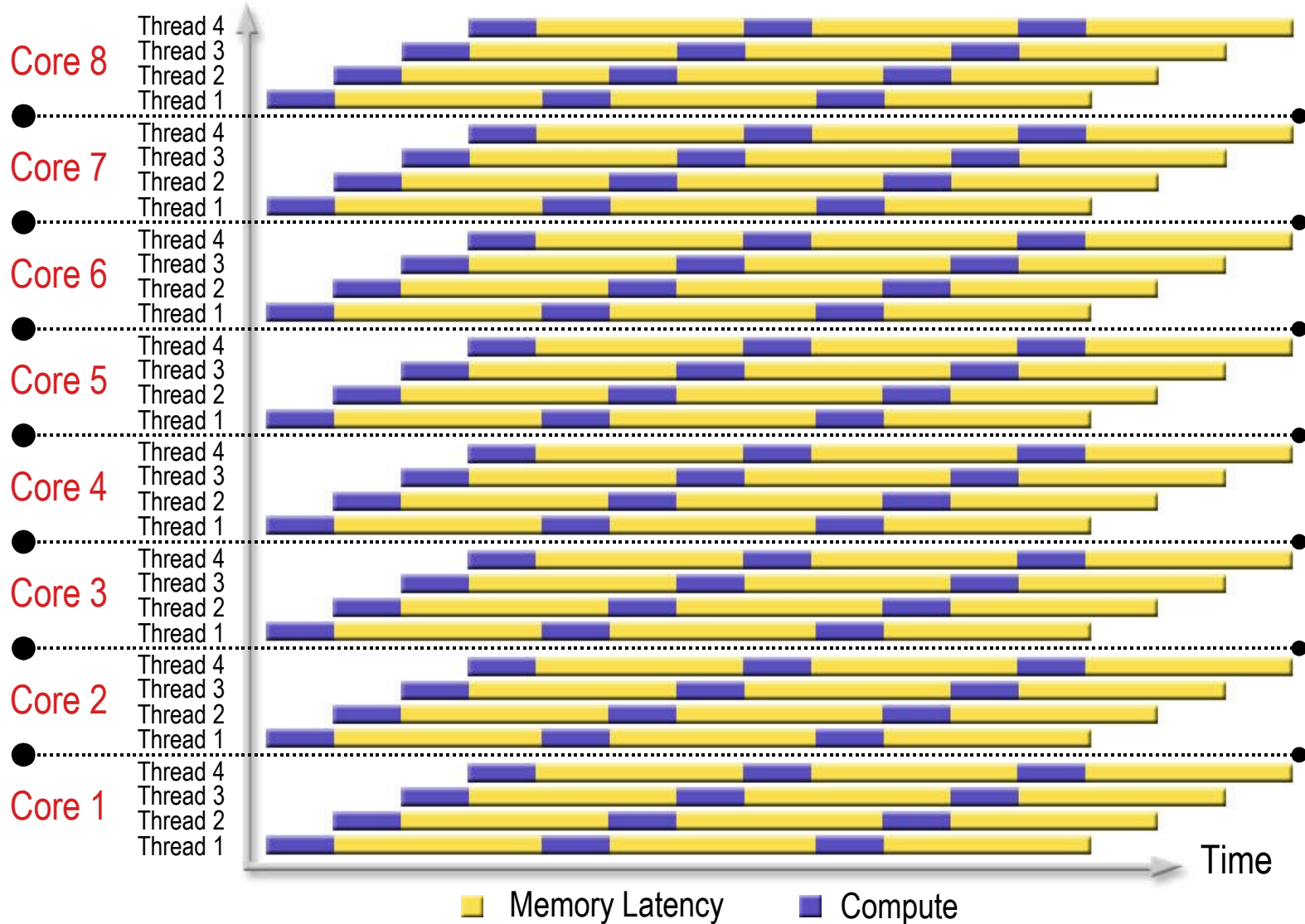
- Market forces are driving the following
 - > More performance
 - > More threaded workloads
 - > Power limited designs
- Technology is driving
 - > Higher power each generation
 - > Less performance from frequency gains
 - > Good scaling of I/O bandwidth through SERDES
- This is exploited by Sun Chip Multi-threaded (CMT) Systems

Comparing Modern CPU Design Techniques

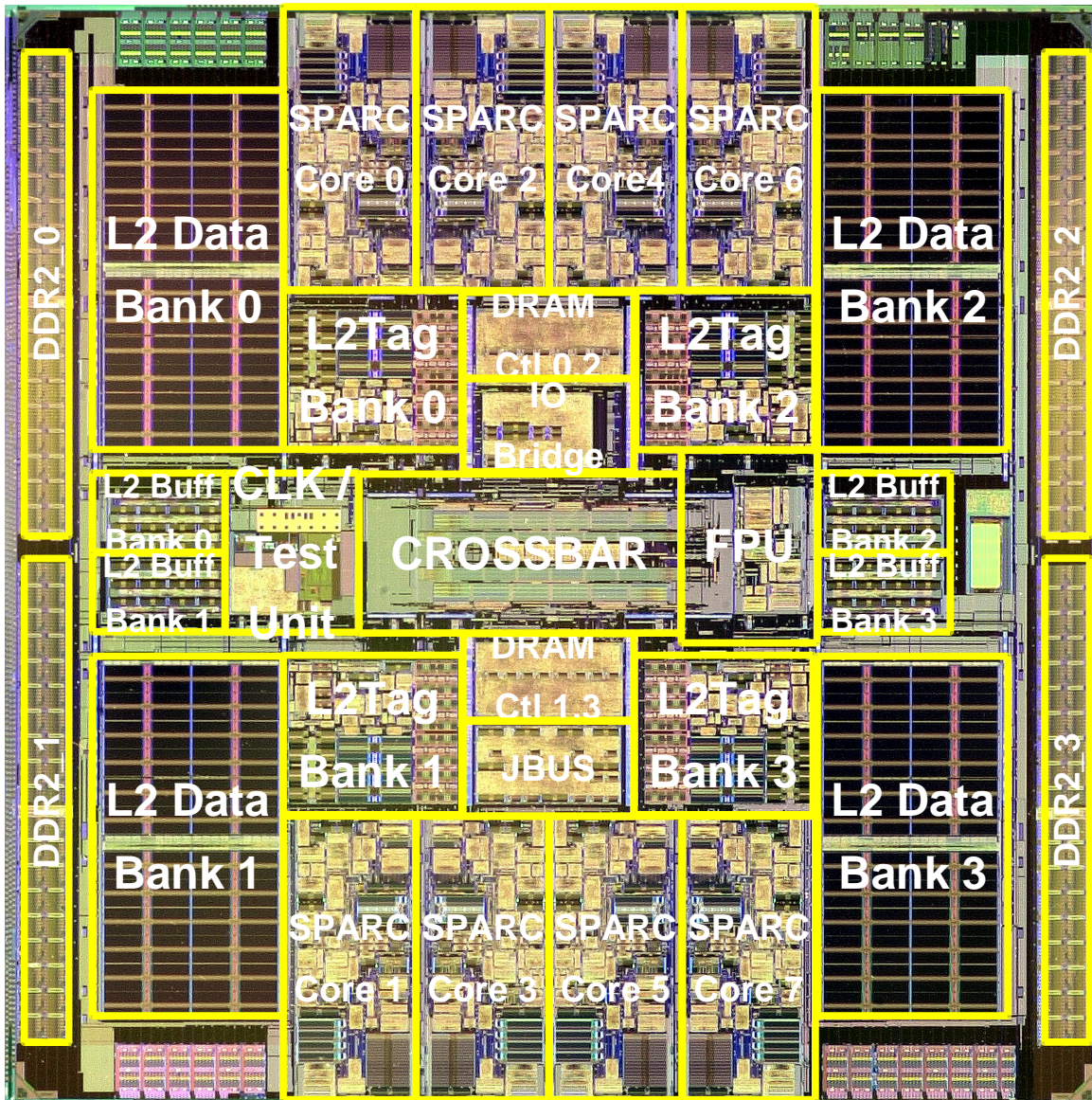


- ILP Offers Limited Headroom
- TLP Provides Greater Performance Efficiency

CMT – Multithreaded Cores



Niagara Micrograph and Overview



Features:

- 8 64-bit Multithreaded SPARC Cores
- Shared 3MB L2 Cache
- 16KB I-Cache per Core
- 8KB D-Cache per Core
- 4 144-bit DDR2 channels
- 3.2 GB/sec JBUS I/O

Technology:

- TI's 90nm CMOS Process
- 9LM Cu Interconnect
- 63 Watts @ 1.2GHz/1.2V
- Die Size: 378mm²
- 279M Transistors
- Flip-chip ceramic LGA

CMOS Power

(Ignoring 2nd order terms)

$$\text{Power} = aCV^2F + \text{Leakage (P,V,T)}$$

a = activity factor

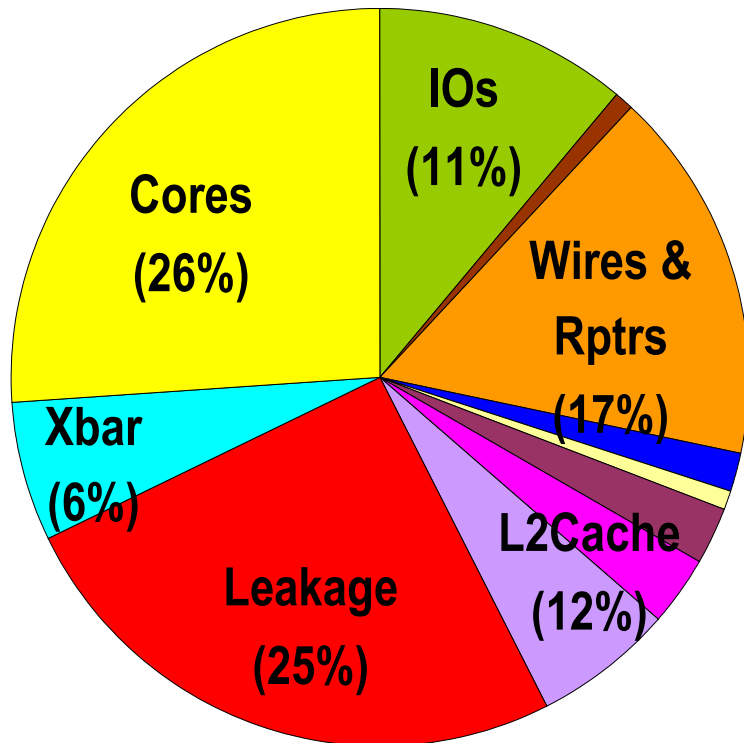
C = capacitance of nodes

V = voltage

F = chip frequency

“Niagara” T1 Chip Power

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



 **SPARC Cores**

 **Leakage**

 **Wires & Rptrs**

 **L2Data**

 **L2Tag Unit**

 **L2 Buffer Unit**

 **Crossbar**

 **Global Clock**

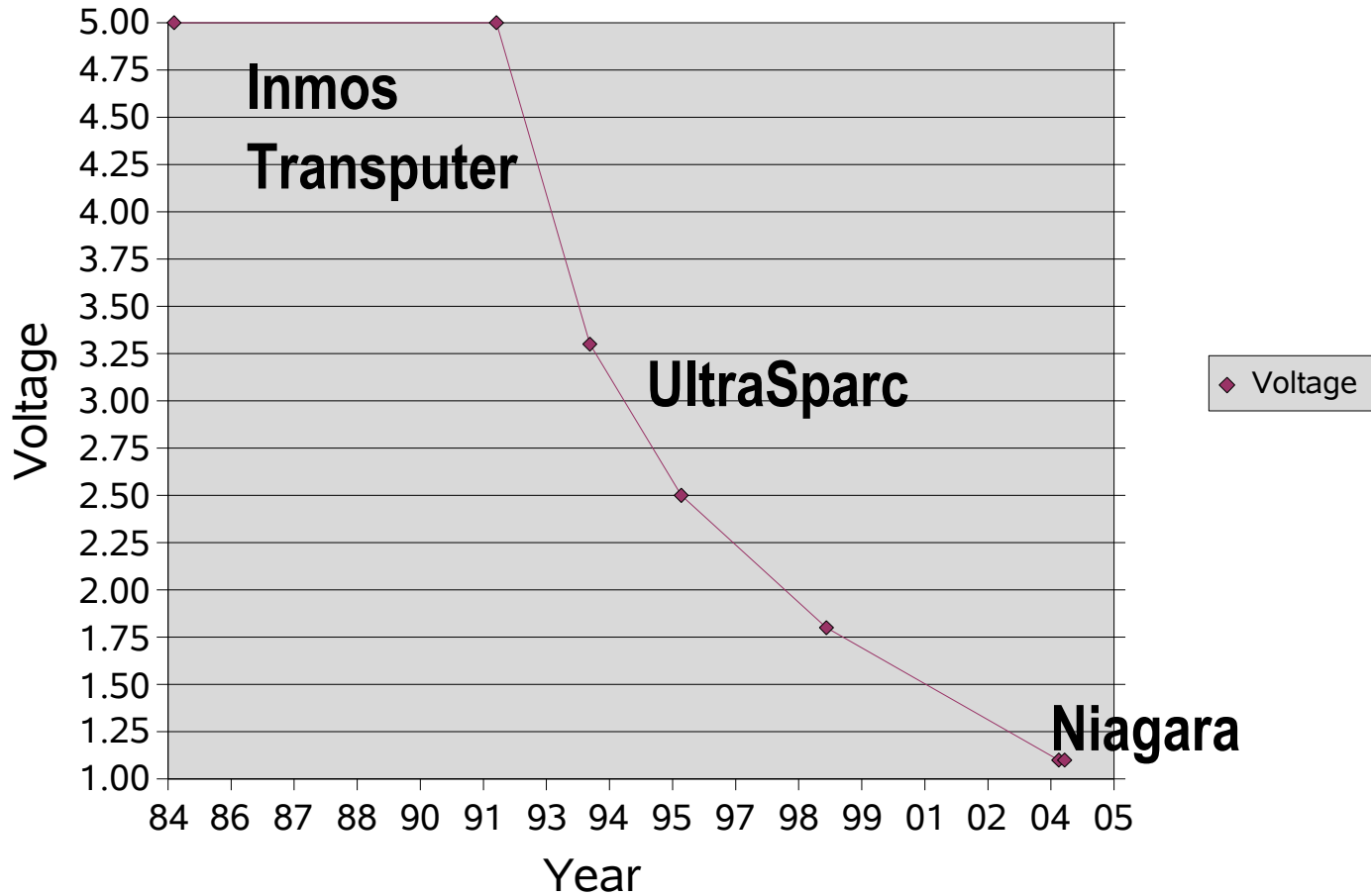
 **Floating Point**

 **IOs**

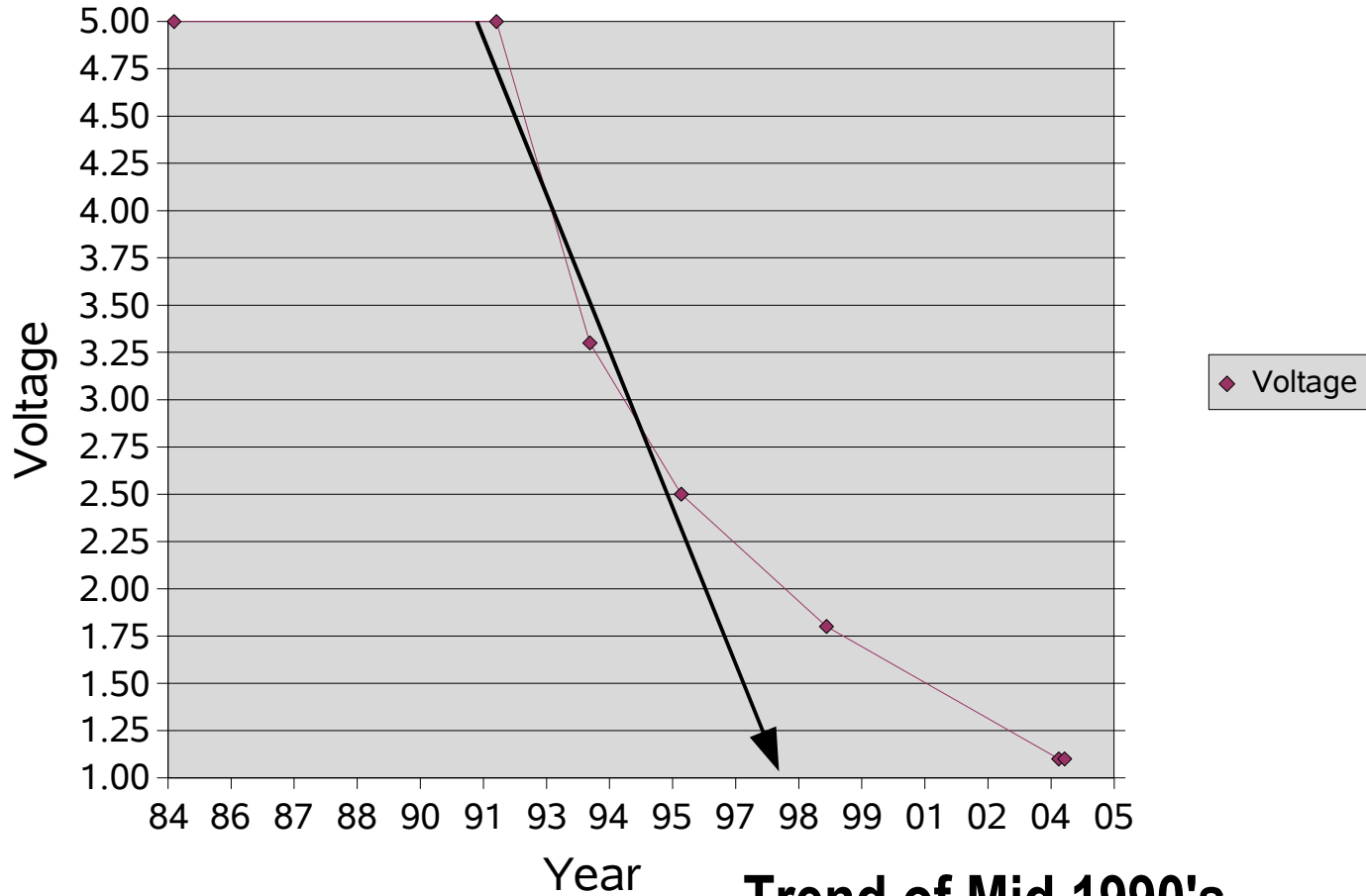
 **Misc Units**

- 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

Personal View of Processor Designs

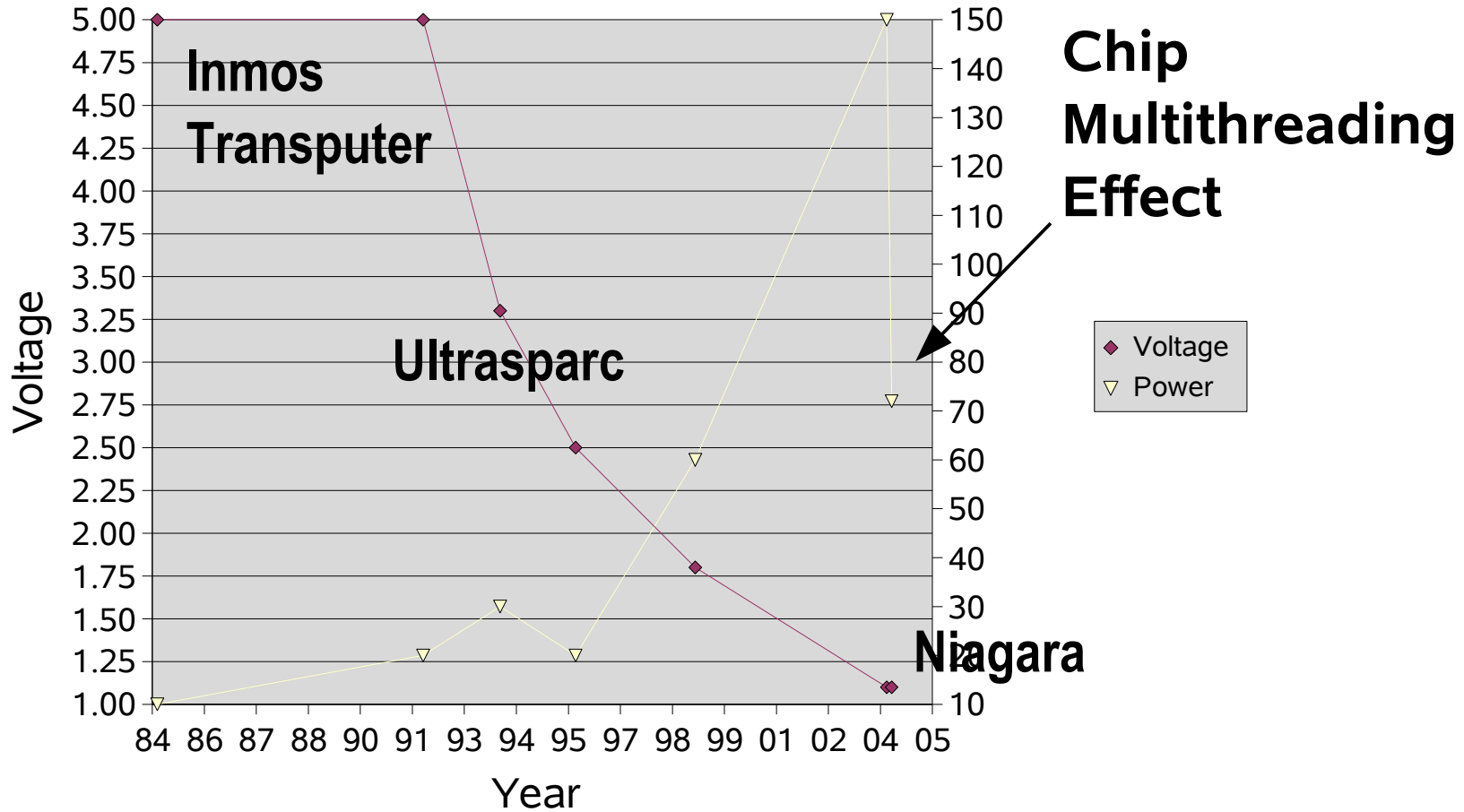


Personal View of Processor Designs



**Trend of Mid 1990's
not sustainable**

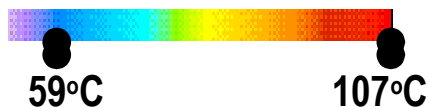
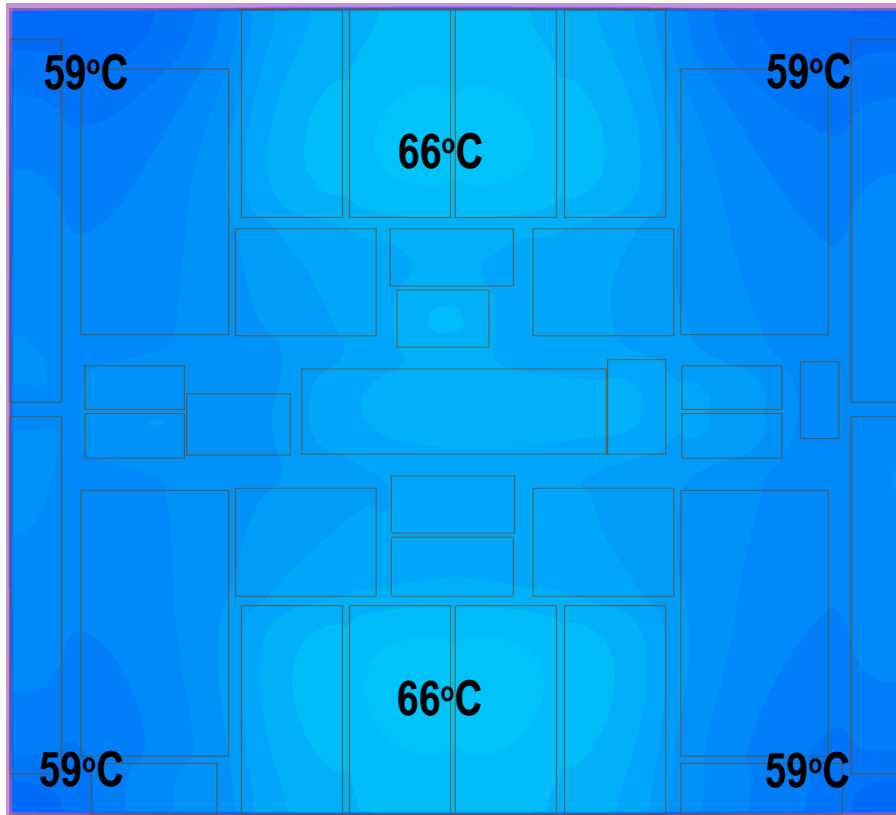
Personal View of Processor Designs



Niagara low power

- Niagara low power style
 - > No speculation
 - > No out of order
 - > No complex branch prediction
 - > No predication
 - > Short pipeline
 - > Moderate clock frequency
 - > Static CMOS design
 - > Threading to cover memory latency
- Typical competitor
 - > Lots of speculation
 - Out of order etc
 - > Wide issue
 - > Deep pipelines
 - > High frequency
 - > Lots of dynamic circuits
 - > Long stall when memory is accessed

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

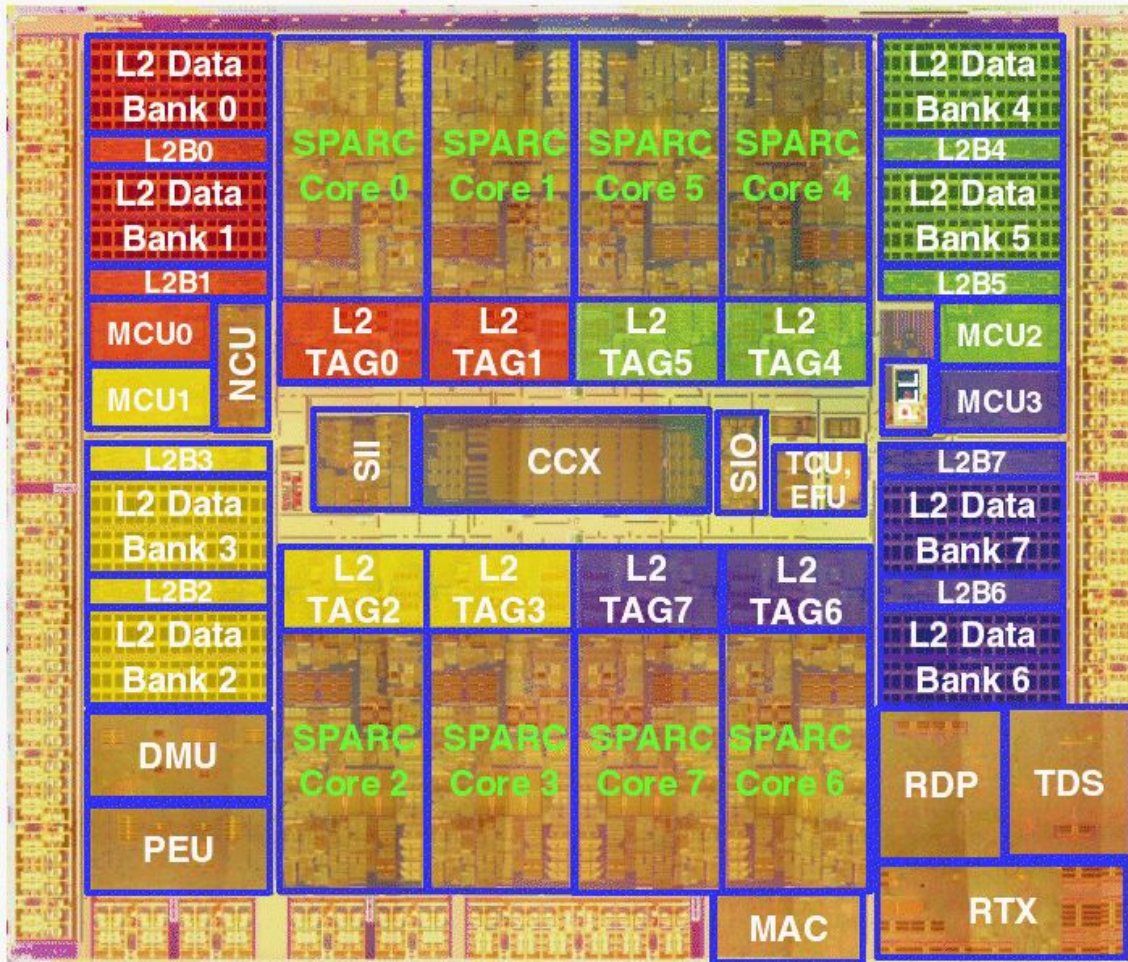
Data Center Constraints

- Many data centers are maxed out
 - > Some constrained by cooling limits
 - > Others by electrical substations
- Getting new buildings & equipment is expensive
- In some locations e.g. financial centers its impossible
- Performance of the data center is constrained by performance/watt of the servers

Niagara-2

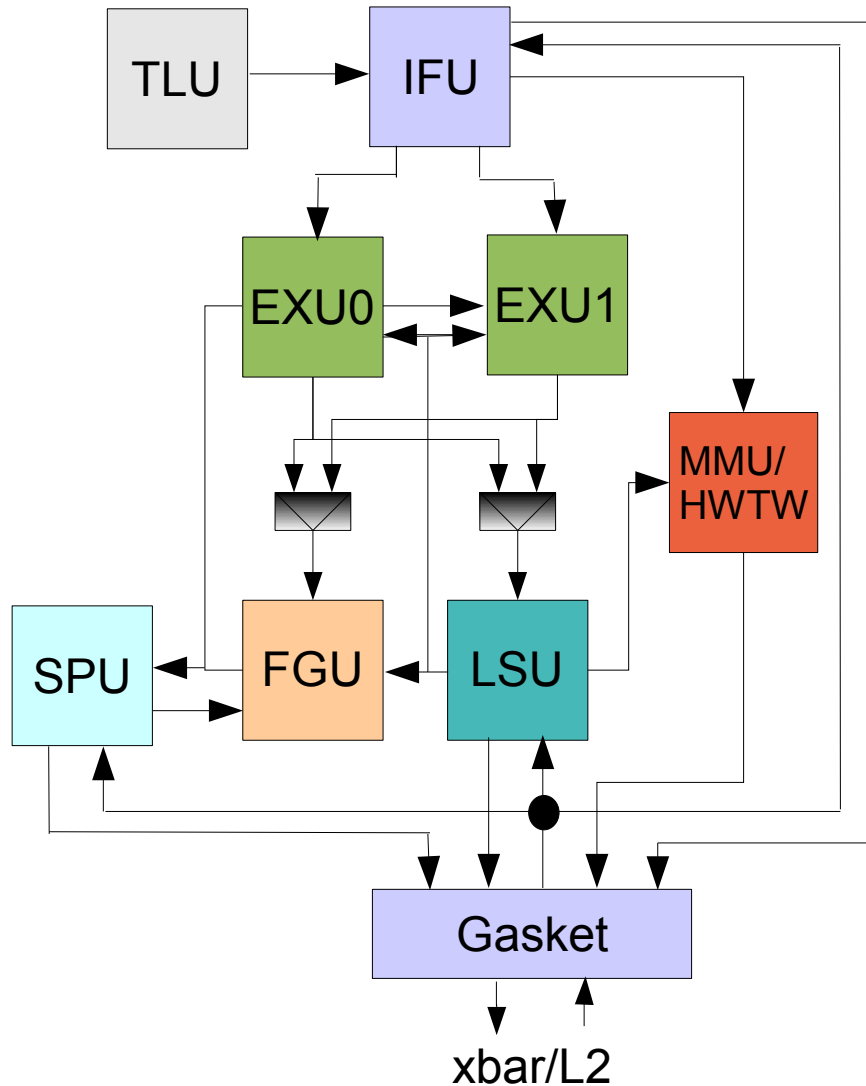
- Double throughput versus UltraSparc T1
 - > Maintain Sparc binary compatibility
 - > <http://opensparc.sunsource.net/nonav/index.html>
- Improve throughput / watt
- Improve single-thread performance
- Integrate important SOC components
 - > Networking
 - > Cryptography

Niagara-2 Chip Overview



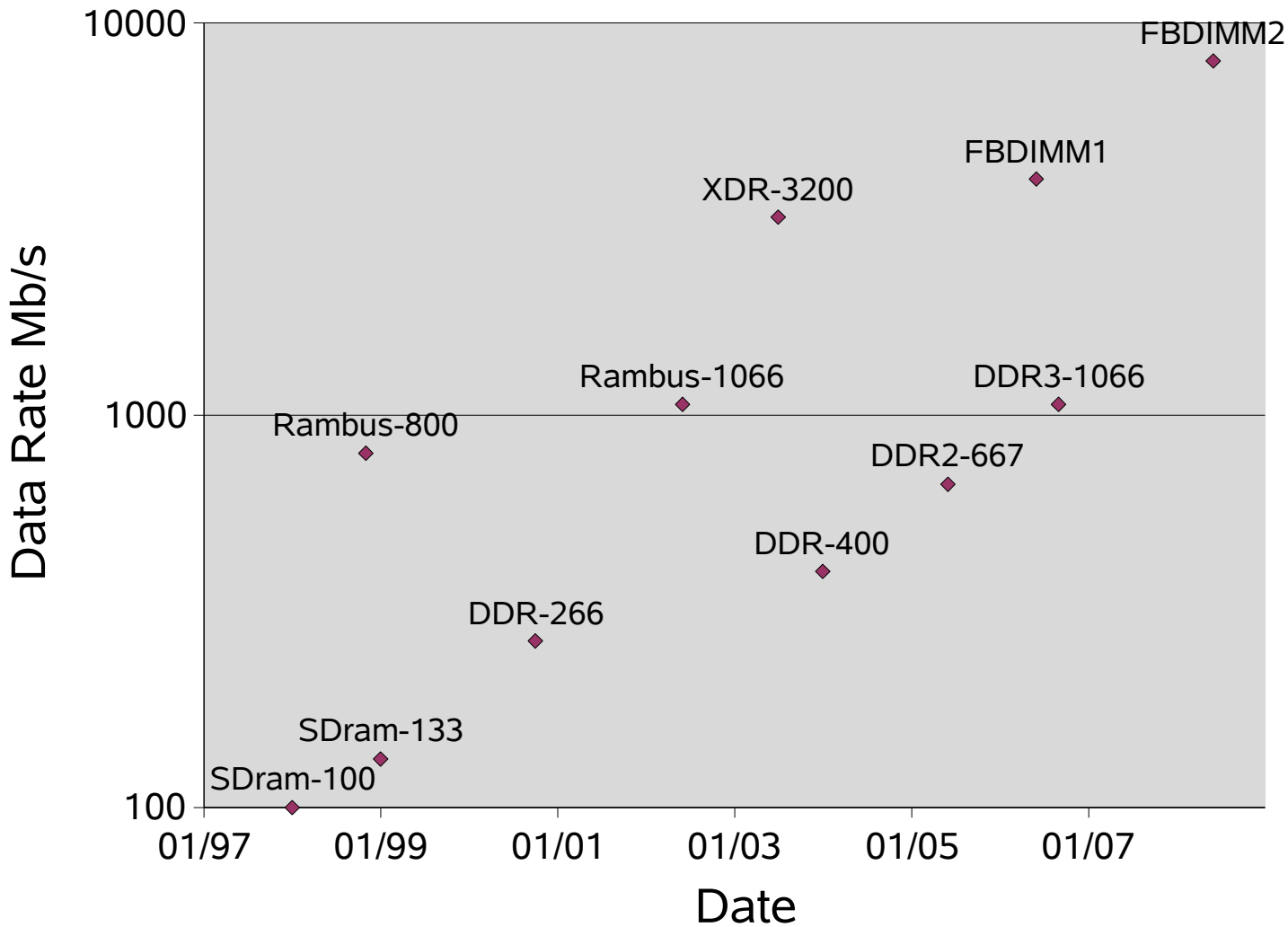
- 8 Sparc cores, 8 threads each
- Shared 4MB L2, 8-banks, 16-way associative
- Four dual-channel FBDIMM memory controllers
- Two 10/1 Gb Enet ports w/onboard packet classification and filtering
- One PCI-E x8 1.0 port
- 711 signal I/O, 1831 total

Sparc Core Block Diagram



- IFU – Instruction Fetch Unit
 - > 16 KB I\$, 32B lines, 8-way SA
 - > 64-entry fully-associative ITLB
- EXU0/1 – Integer Execution Units
 - > 4 threads share each unit
 - > Executes one integer instruction/cycle
- LSU – Load/Store Unit
 - > 8KB D\$, 16B lines, 4-way SA
 - > 128-entry fully-associative DTLB
- FGU – Floating/Graphics Unit
- SPU – Stream Processing Unit
 - > Cryptographic acceleration
- TLU – Trap Logic Unit
 - > Updates machine state, handles exceptions and interrupts
- MMU – Memory Management Unit
 - > Hardware tablewalk (HWTW)
 - > 8KB, 64KB, 4MB, 256MB pages

Dram Data Rates Versus Time



Dram Issues

- Niagara I uses DDRII – 400
- Niagara II uses FBDIMM 4Gb/s
 - > Higher data rate is good
 - > AMB power & cost is a problem
- Need to amortize the serialization cost across more memories
 - > Stacking technologies
 - > More Dram/DIMM
 - > Other configurations of buffers to fanout to DDR DIMMs

Niagara & Niagara II Packages



Future packaging requirements

- Challenges for future packages
- Similar pin counts but:
- Data rates keep increasing to match higher processor performance
- Costs getting squeezed – particularly in entry level servers

At the system level

Yesterdays Vertical System

E10000

1997

32 x US2

77.4 ft³

2000 lbs

13,456 W

52,000 BTUs/hr



Today's Horizontal System

T2000

2005

1 x US T1

0.85 ft³

37 lbs

~300 W

1,364 BTUs/hr



91x smaller
54x lighter
44x cooler



T2000

Conclusion

- Chip multithreading does keep the data center cool
- 2 generations of the Niagara processor
- Future challenges
 - > Very high data rate interfaces
 - > Managing power is an on going challenge at both CPU, Memory and System Levels

Thankyou !

David Greenhill

Distinguished Engineer

david.greenhill@sun.com