

Computer architecture is all about interconnect

(it is now and it will be more so in 2010)

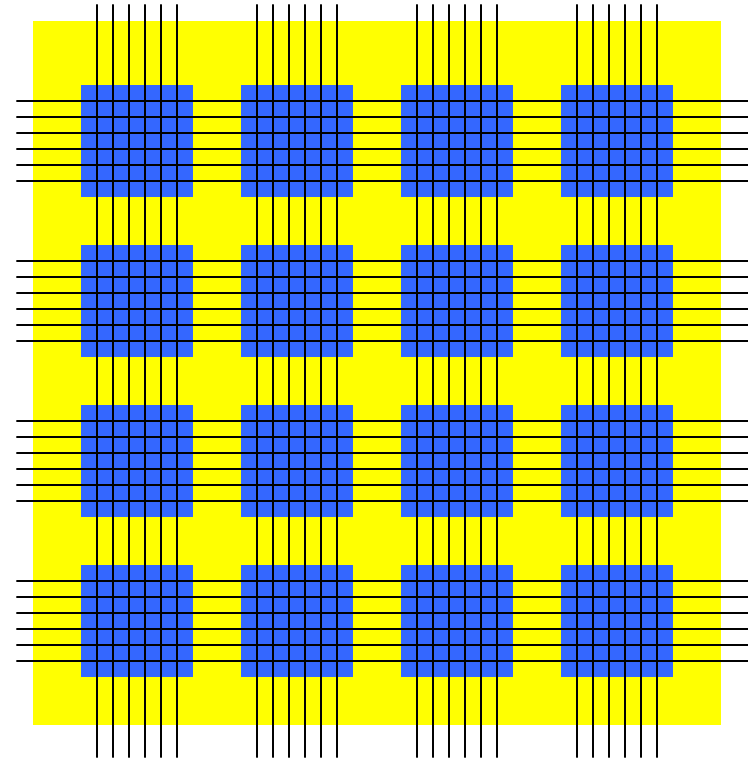
HPCA Panel
February 4, 2002

William J. Dally
Computer Systems Laboratory
Stanford University
billd@csl.stanford.edu

Computer architecture is primarily about interconnect

Two subproblems:

1. Schedule a computation in time and space to minimize communication
 - Placement, scheduling, staging
 - Compile-time and run-time
2. Architect efficient interconnection networks to move data through a computation
 - Topology, routing, flow-control
 - Bandwidth, latency, power
 - Depends on technology



Interconnect is the limiting factor because...

- Interconnect dominates delay
- Interconnect dominates power
- Interconnect is becoming even more dominant over time as technology 'improves'
 - Memory (cells and arrays) and processing getting better with technology, interconnect gets worse
- The 'memory' problem is really an interconnect problem
 - Memory (cells and arrays) are plenty fast, its getting to them that's slow
- The 'processor' problem is really an interconnect problem
 - Easy to do lots of operations – its hard to move the data between them

Interconnect dominates delay

Operation	Delay	
	(0.13um)	(0.05um)
32b ALU Operation	650ps	250ps
32b Register Read	325ps	125ps
Read 32b from 8KB RAM	780ps	300ps
Transfer 32b across chip (10mm)	1400ps	2300ps
Transfer 32b across chip (20mm)	2800ps	4600ps

2: 1 global on-chip comm to operation delay
9: 1 in 2010

Power is a matter of distance (interconnect)

Operation	Power	
	(0.13um)	(0.05um)
32b ALU Operation	5pJ	0.3pJ
32b Register Read	10pJ	0.6pJ
Read 32b from 8KB RAM	50pJ	3pJ
Transfer 32b across chip (10mm)	100pJ	17pJ
Execute a uP instruction (SB-1)	1.1nJ	130pJ
Transfer 32b off chip (2.5G CML)	1.3nJ	400pJ
Transfer 32b off chip (200M HSTL)	1.9nJ	1.9nJ

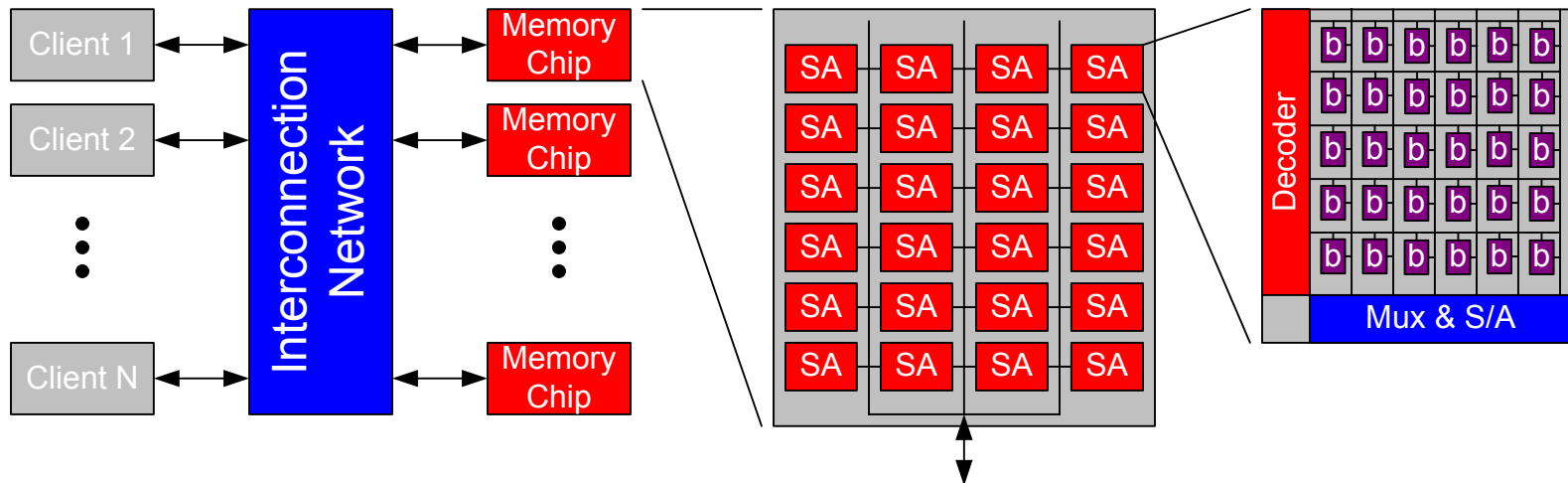
300: 20: 1 off-chip to global to local ratio in 2002
1300: 56: 1 in 2010

Interconnect is becoming more significant over time

- It doesn't scale with technology
 - Memory (cells and arrays) and processing (ALU)s
 - Delay $\propto L$
 - Energy $\propto L^3$
 - System-wide interconnect
 - Delay – constant
 - Energy – constant*
 - Global on-chip interconnect
 - Delay – increasing
 - Energy $\propto L^2$
- Interconnect dominates delay and energy now and its getting worse
- As chips get denser, need on-chip interconnection networks to connect subsystems
 - 32-bit CPU today is $< 0.25\text{mm}^2$ (4,000 on a 10mm chip)
 - In 2010 it will be 0.037mm^2

The 'memory problem' is really an interconnect problem

- Speed, density, and power of memory cell and subarray is scaling with Moore's law
- Latency and bandwidth are limited not by memory performance but by the *interconnect* between the client and the memory
- Viewing this as an interconnect problem rather than a memory problem makes the solution easier and brings a well-developed technology to bear.



The 'processor problem' is really an interconnect problem

- Lots of interconnection networks in a processor
 - From instruction fetch to reservation stations
 - Bypass network for results
 - From 'global' register file to ALUs
 - Can't assume these are free – or even a single cycle
- These are becoming increasingly critical
 - Clustered ALUs
- OOO scheduling hardware is aimed at managing interconnect latency
 - In the form of cache misses
- But all of this communication is 'hidden'

Architectures need to make interconnect a first-class citizen

- Interconnect dominates power and delay of modern computer systems
 - Much in the guise of 'memory latency' or 'superscalar overhead'
- Need to make this communication visible so it can be optimized
 - Schedule where an instruction takes place, not just when
 - Manage communication across a storage hierarchy explicitly
- Legacy ISAs don't do this
- Also need to do a better job of building these interconnection networks – especially the on-chip ones
 - Topology, routing, and flow control
- Need to consider physical location and interconnect when evaluating new processor 'mechanisms'
 - Need to build it (or at least lay it out) to stay honest.

Some rebuttals

- Horst
 - “Even Microsoft can’t find a way to waste the cycles”
- Basil
 - “Memory hopeless” – there is hope with good interconnect
 - “Interconnect fine” – on what planet do people not want more bandwidth?
 - “Processor sandbox ... doesn’t matter” – agreed.
- Jose
 - “Academia follows industry trends...” – Industry follows Stanford.
- Anant
 - “Dally will say...” I was going to say that, but I was thirsty
 - RAW == Warp – why do you need 13-15 cycles per hop?
 - Never trust someone who doesn’t put numbers on the axes of their graphs
- Yale
 - “Processor’s are sexy” – in California we are open minded about orientation
 - “Its all about funding” – but funding sources focus on near term

It's not a question of interconnect vs. processor vs. memory

- Processor and memory research needs to become 'interconnect aware'
 - Interconnect accounts for most of the delay and power in processors and memory systems today (and more so in 2010)
- Interconnect research needs to consider a broad range of applications
 - System-area networks
 - Cabinet-area processor/memory networks
 - On-chip processor/SOC networks
 - On-chip memory interconnect
 - ALU-ALU bypass networks
 - Instruction distribution/dispatch networks
- Interconnect research is harder but has greater potential returns
- Need to take a long-term (5-10 year) view.