Lecture 29: Interconnection Networks

Papers:

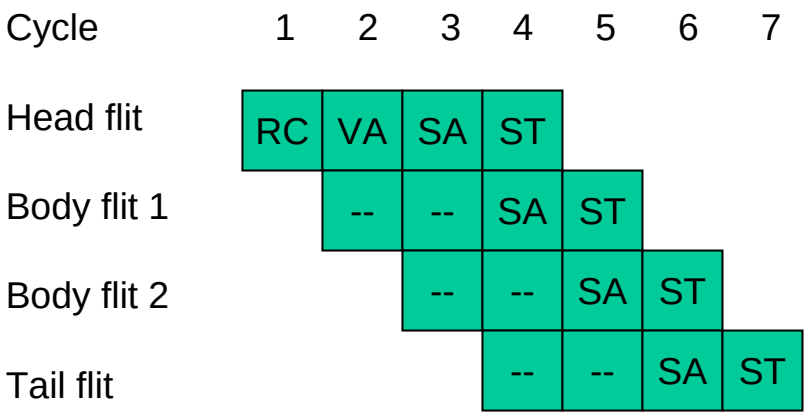
- Express Virtual Channels: Towards the Ideal Interconnection Fabric, ISCA'07, Princeton
- Interconnect Design Considerations for Large NUCA Caches, ISCA'07, Utah

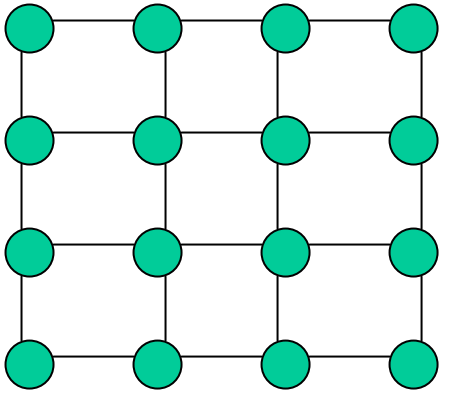
Reminders:

- Take-home final (20%): available this weekend, due May 10th
- Project (50%): due May 10th; peer-review due May 12th
- Project reports: conference paper style, at least 4 pages double-column; abstract, intro, background, proposal, methodology, results, related work, conclusions

Router Pipeline

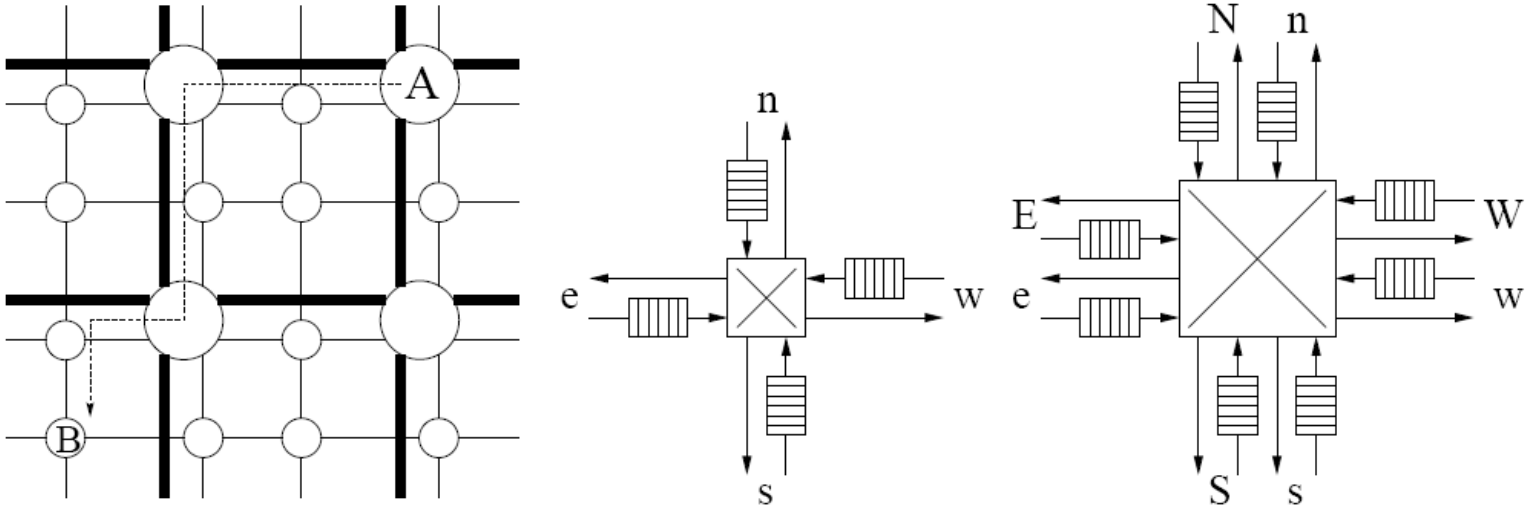
- Four typical stages:
  - RC routing computation: compute the output channel
  - VA virtual-channel allocation: allocate VC for the head flit
  - SA switch allocation: compete for output physical channel
  - ST switch traversal: transfer data on output physical channel





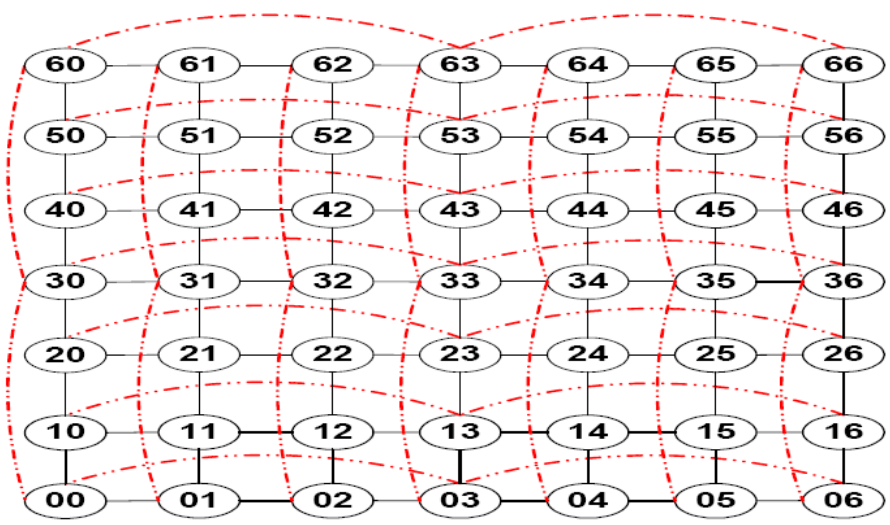
Express Physical Channels

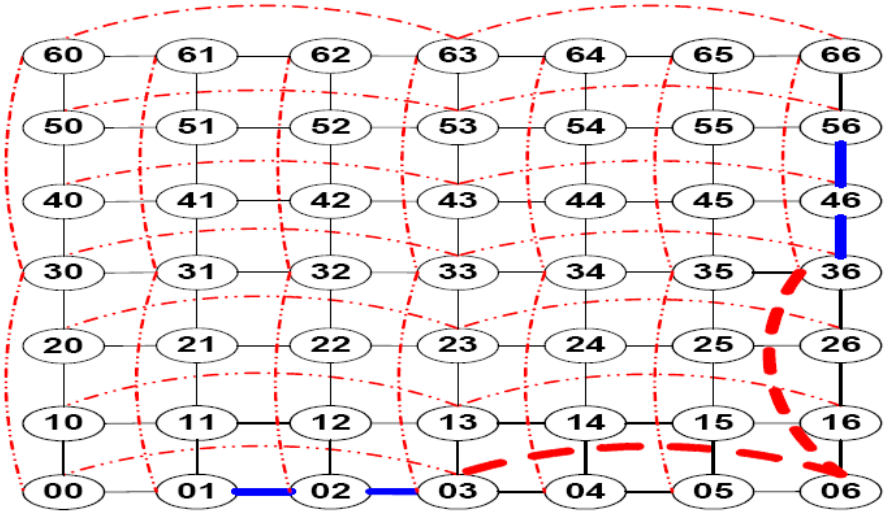
- Express channels connect non-adjacent nodes – flits traveling a long distance can use express channels for most of the way and navigate on local channels near the source/destination (like taking the freeway)
- Helps reduce the number of hops
- The router in each express node is much bigger now



Express Virtual Channels

- To a large extent, maintain the same physical structure as a conventional network (changes to be explained shortly)
- Some virtual channels are treated differently: they go through a different router pipeline and can effectively avoid most router overheads





(b) VCs acquired from nodes 01 to 56

Router Pipelines

- If Normal VC (NVC):
  - at every router, must compete for the next VC and for the switch
  - will get buffered in case there is a conflict for VA/SA
- If EVC (at intermediate bypass router):
  - need not compete for VC (an EVC is a VC reserved across multiple routers)
  - similarly, the EVC is also guaranteed the switch (only 1 EVC can compete for an output physical channel)
  - since VA/SA are guaranteed to succeed, no need for buffering
  - simple router pipeline: incoming flit directly moves to ST stage
- If EVC (at EVC source/sink router):
  - must compete for VC/SA as in a conventional pipeline
  - before moving on, must confirm free buffer at next EVC router

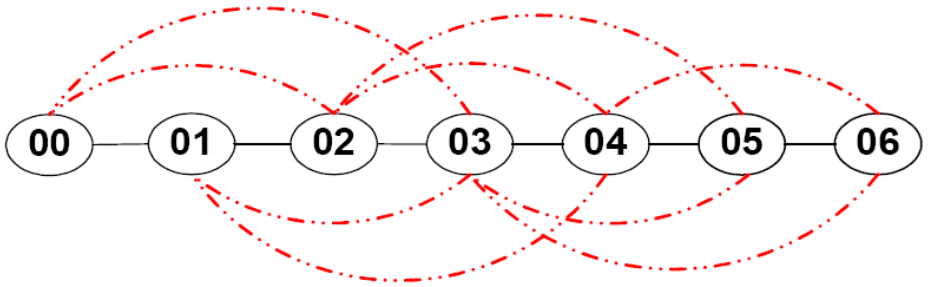
Bypass Router Pipelines

- Non aggressive pipeline in a bypass node: an express flit simply goes through the crossbar and then on the link; the prior SA stage must know that an express flit is arriving so that the switch control signals can be appropriately set up; this requires the flit to be preceded by a single-bit control signal (similar to cct-switching, but much cheaper)

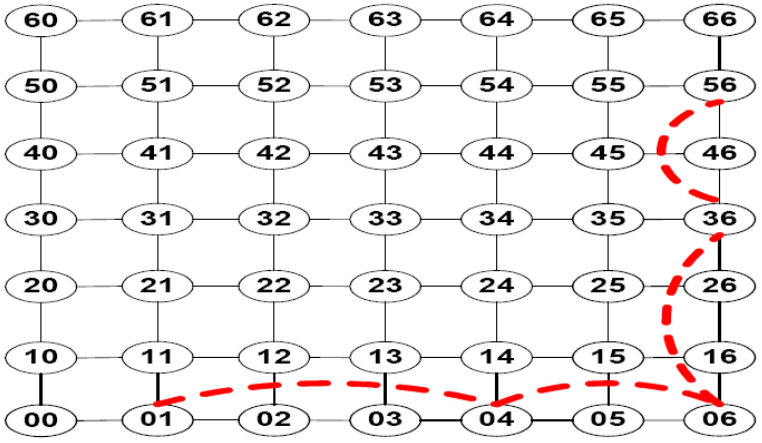
- Aggressive pipeline: the express flit avoids the switch and heads straight to the output channel (dedicated hardware)… will still need a mechanism to control ST for other flits

Dynamic EVCs

- Any node can be an EVC source/sink
- The EVC can have length 2 to l_{max}



(a) Dynamic EVCs along the X dimension (assuming $l_{max} = 3$)



(b) VCs acquired from nodes 01 to 56

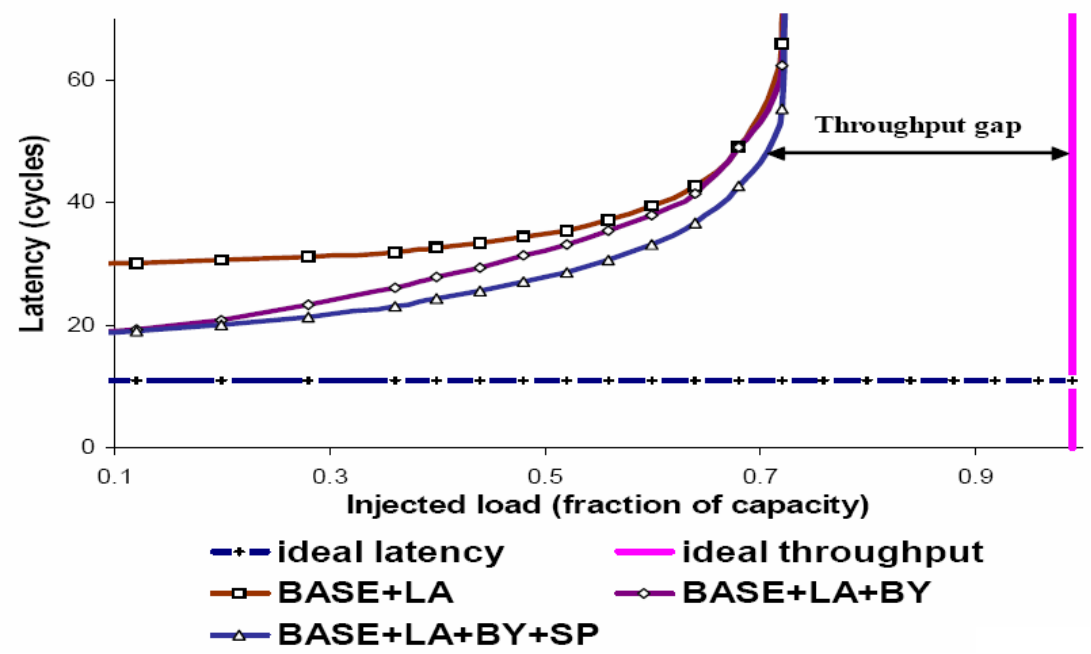
VC Allocation

- All the VCs at a router are now partitioned into l_{max} bins
- More buffers for short-hop EVCs
- Flow control credits have to propagate l_{max} nodes upstream
- Can also dynamically allocate buffers to EVCs (although one buffer must be reserved per EVC to avoid deadlock)
- EVCs can potentially starve NVCs at bypass nodes: if a bypass node is starved for *n* cycles, it sends a token upstream to prevent EVC transmission for the next *p* cycles

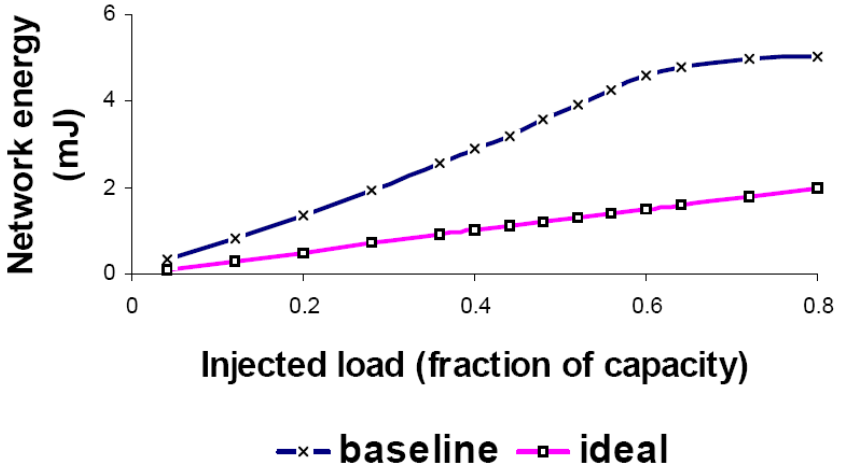
Ideal Network

- Fully-connected: every node has a dedicated link to every other node
- Bisection bandwidth: $L_{edge} = 2 \cdot \frac{N}{2} \cdot \frac{N}{2} \cdot W_{pitch} \cdot c_{width}$
- For a 7x7 network, L_{edge} will be 69mm and chip area will be 4760mm^2 (for a single metal layer)
- An ideal network will provide the least latency, least power, and highest throughput, but will have an inordinate overhead, as specified above

Approaching the Ideal Network



- The interconnection network models employ different forms of speculative router pipelines



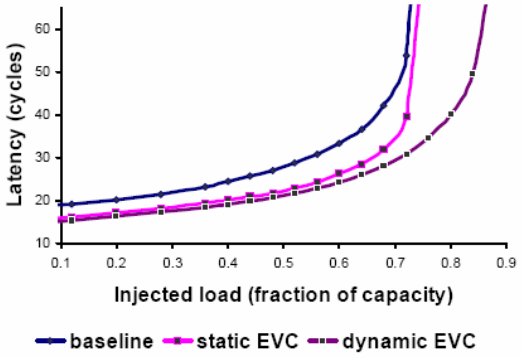
Results

Table 1: Baseline process and network parameters

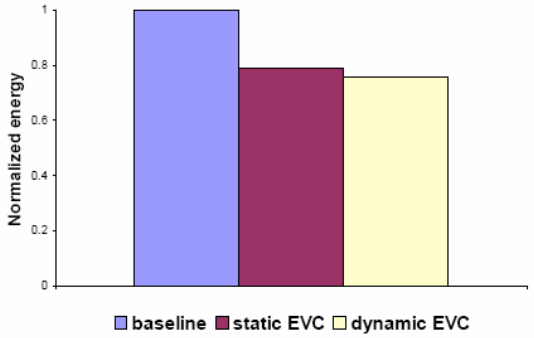
Technology	65 nm
V_{dd}	1.1 V
$V_{threshold}$	0.17 V
Frequency	3 GHz
Topology	7-ary 2-mesh
Routing	Dimension-ordered (DOR)
Traffic	Uniform random
Number of router ports	5
VCs per port	8
Buffers per port	24
Flit size/channel width (c_{width})	128 bits
Link length	1 mm
Wire pitch (W_{pitch})	0.45μm

Table 2: EVC-specific parameters

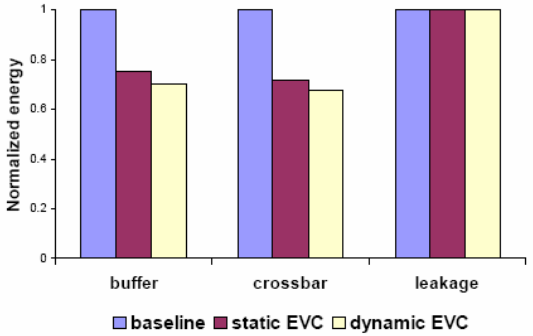
EVC pipeline	Aggressive express pipeline
Buffer management	dynamic
Buffers per port	24
Static EVC-specific parameters	
EVC length	2 hops
NVCs per port	4
EVCs per port	4
Dynamic EVC-specific parameters	
l_{max}	2
NVCs per port	2
EVCs per bin	6
Starvation-avoidance parameters	
n	20
p	3



(a) Network performance



(b) Router energy

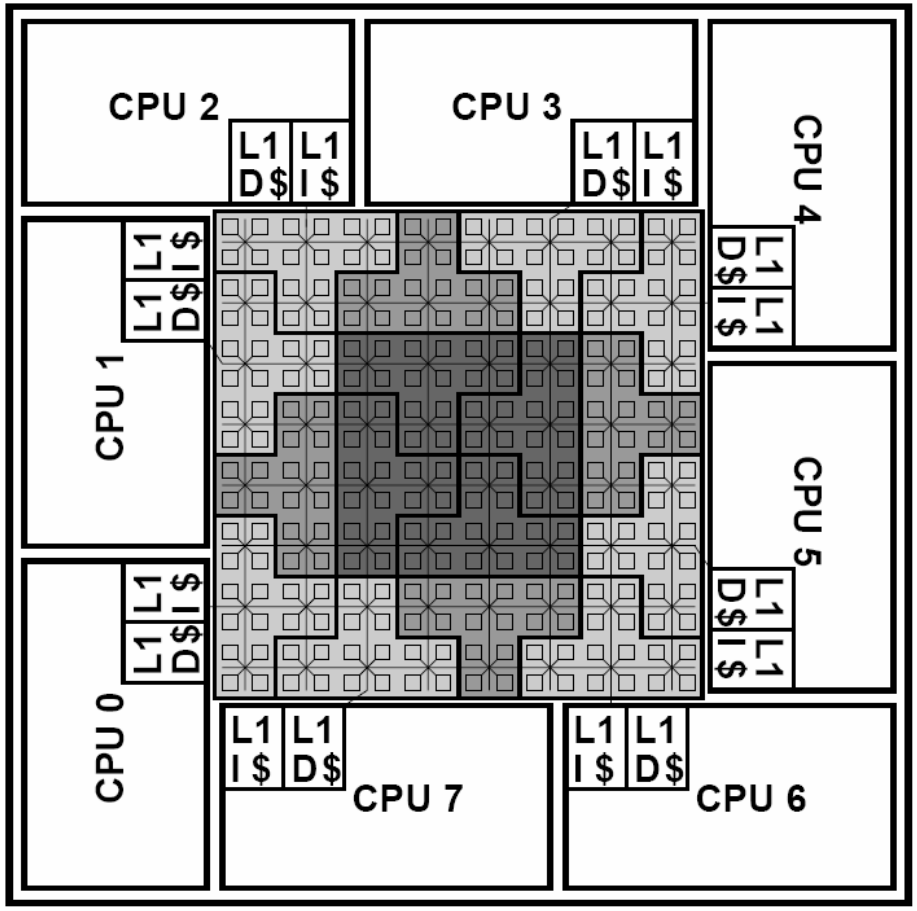


(c) Energy saving from different router components

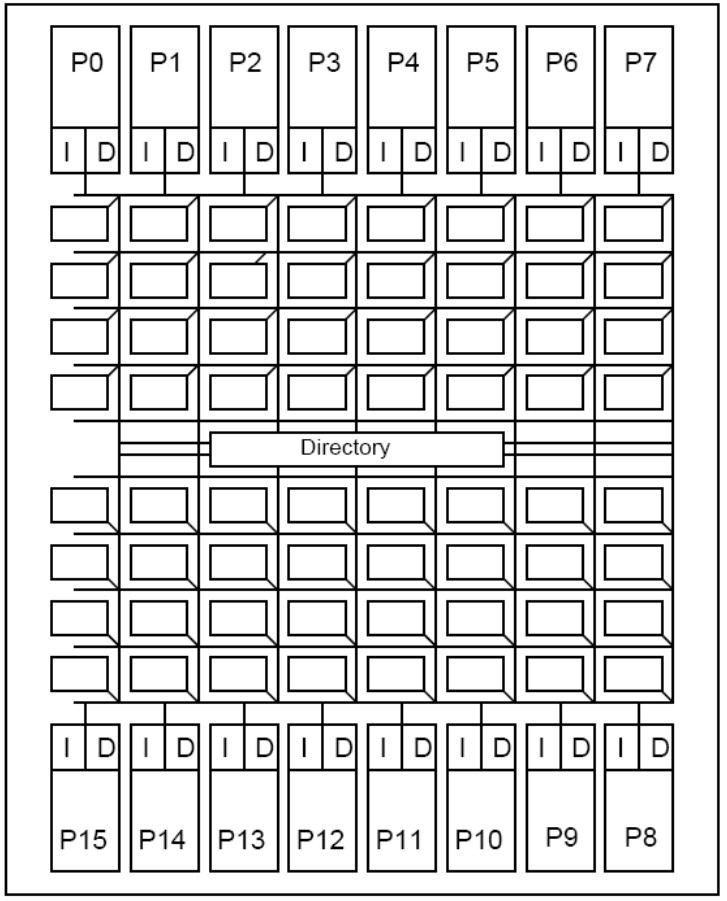
Figure 11: Uniform random traffic results

- Roughly 40% of all nodes are bypassed

Non-Uniform Cache Access



(a) CMP Substrate: 16 CPUs 8x8 Banks



From Beckmann et al. (MICRO'04) and Huh et al. (ICS'05)

Improving NUCA Methodologies

- Design space exploration: iterate over bank counts, organizations, and destinations to compute the optimal cache structure

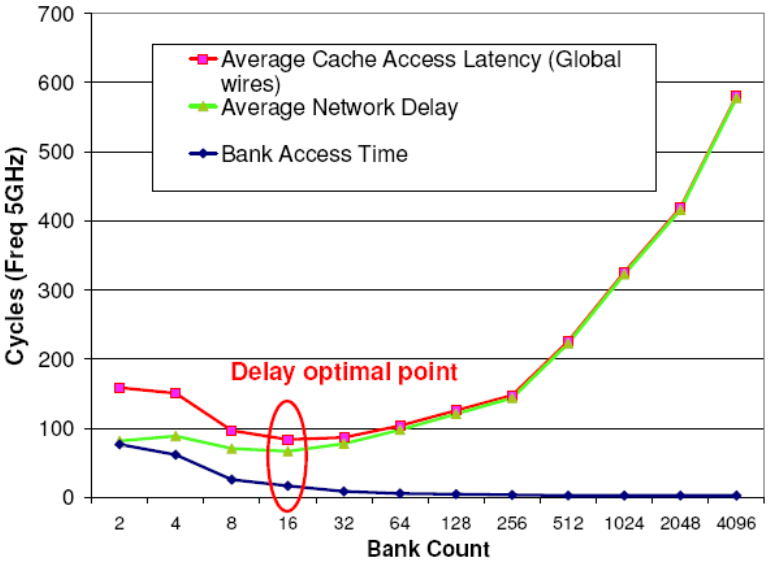
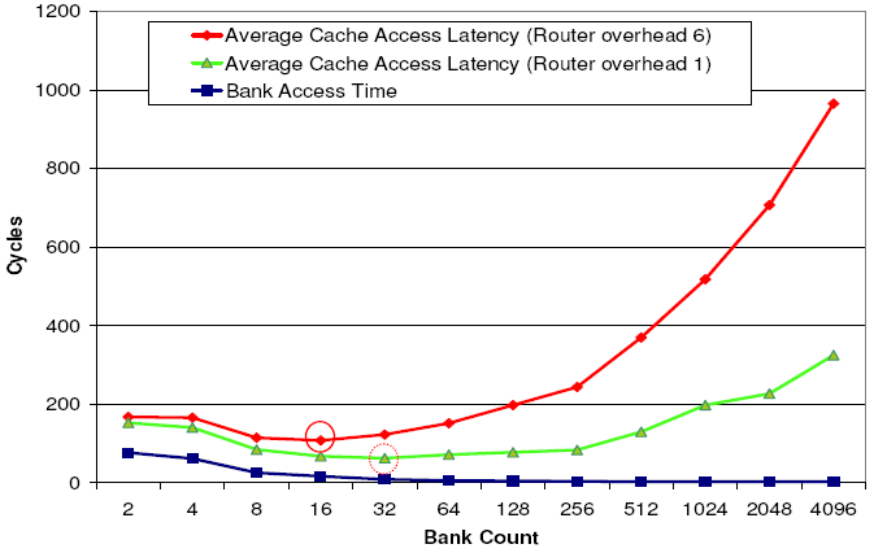
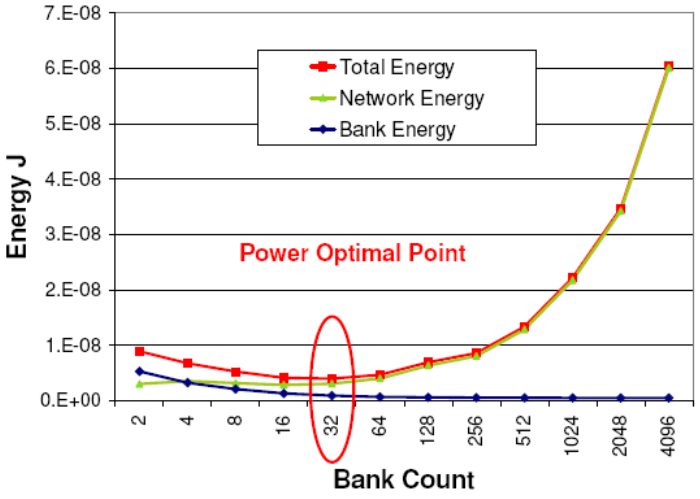


Figure 5: Access latency for a 32 MB cache.





Forms of Heterogeneity

- Different networks for data and address – the latter has lower bandwidth demands and can employ faster wires on higher metal layers
- Parts of the address are more critical: the index bits are transmitted on low-latency links so that cache access can begin early – the rest of the address arrives in time for tag comparison

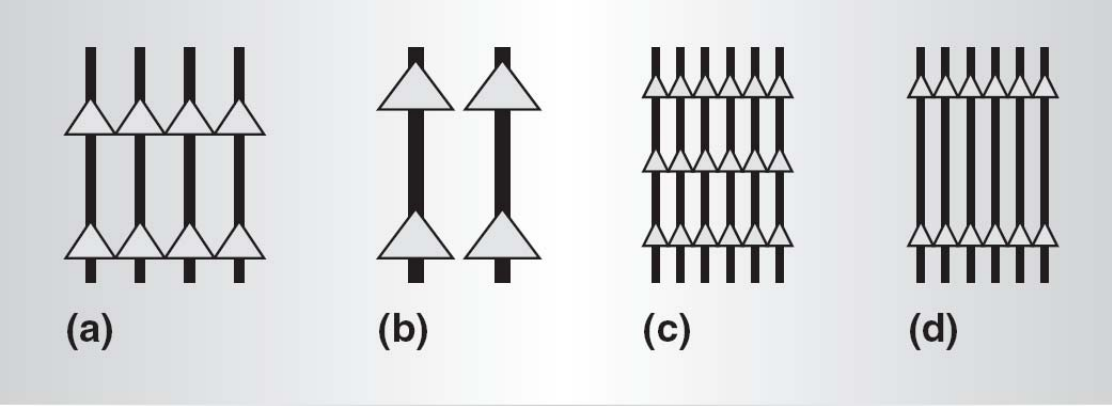
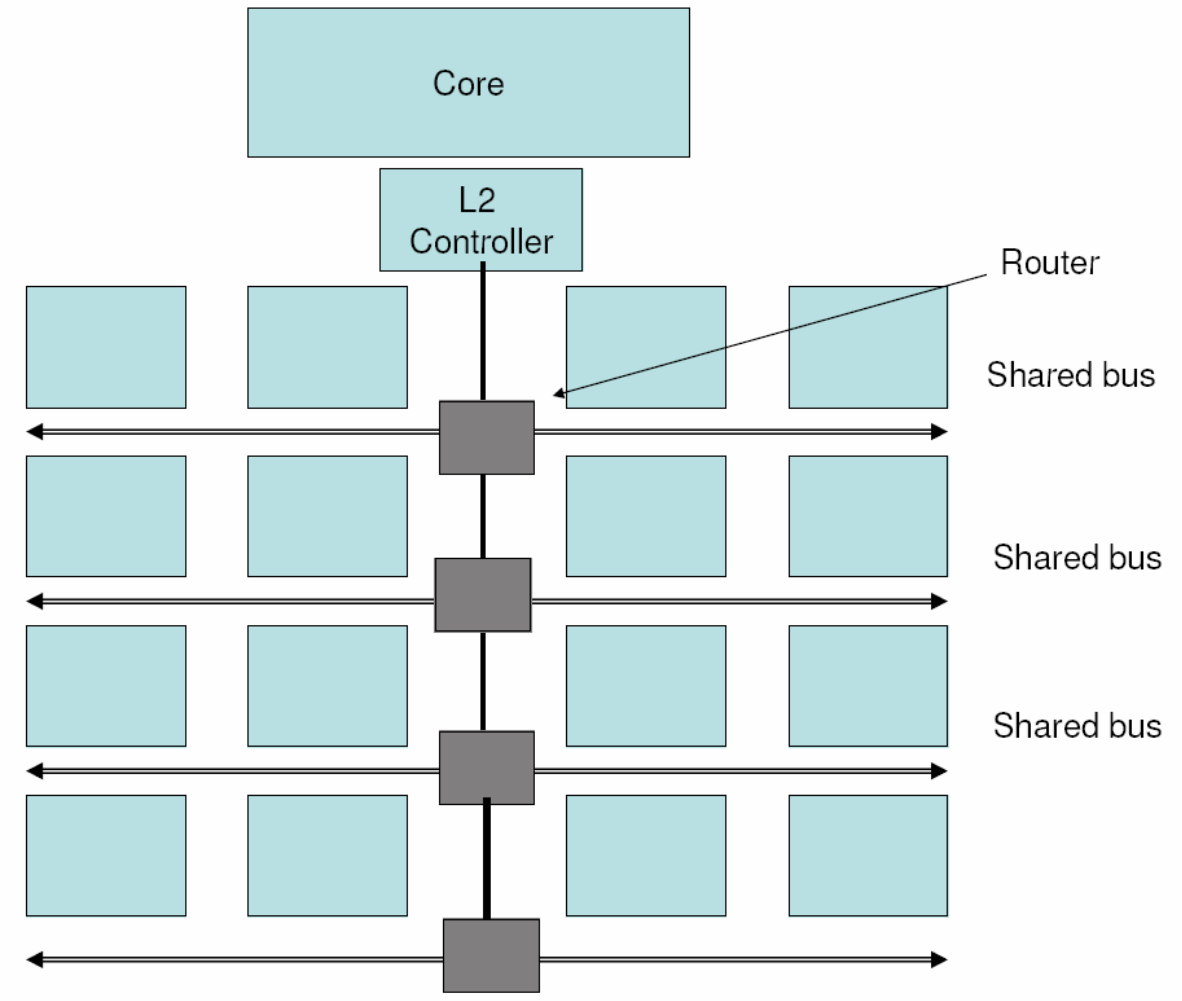


Figure 1. Wire types: B wires (a), L wires (b), W wires (c), and PW wires (d). B, L, and W wires differ in width and spacing. PW wires are W wires with a reduced size and number of repeaters.

Hybrid Network Topology



Title

- Bullet