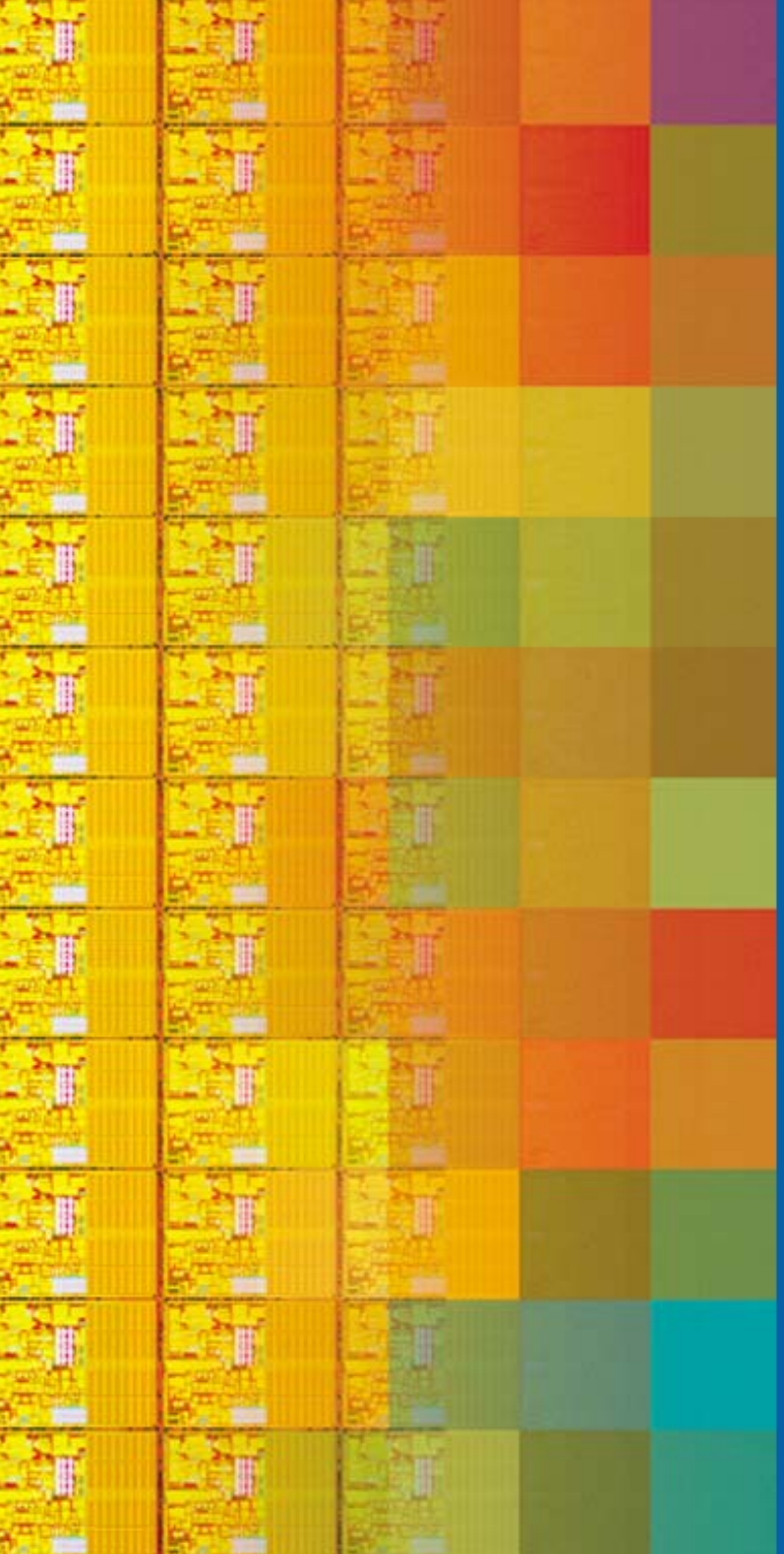


On-Die Interconnects for next generation CMPs

Partha Kundu
Corporate Technology Group (MTL)
Intel Corporation

OCIN Workshop, Stanford University
December 6, 2006

Multi-Core Transition Accelerating

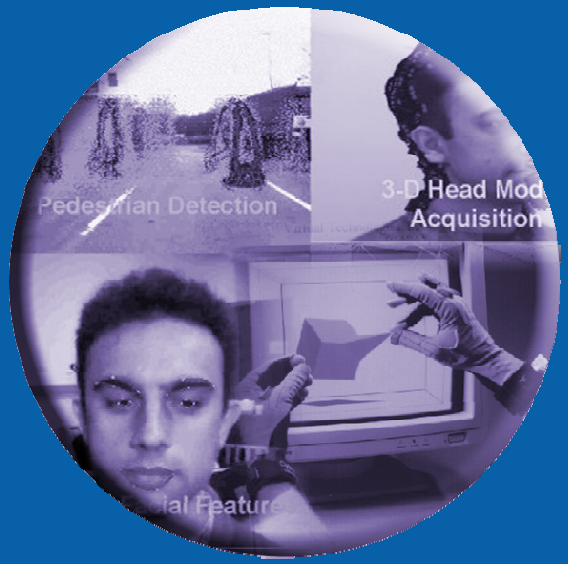
"We notified customers we're pulling in both the desktop and server (launch) of the first quad-core processors into the fourth quarter of this year from the first half of 2007"

"The UltraSPARC T1 processor with CoolThreads technology is the highest-throughput and most eco-responsible processor ever created."

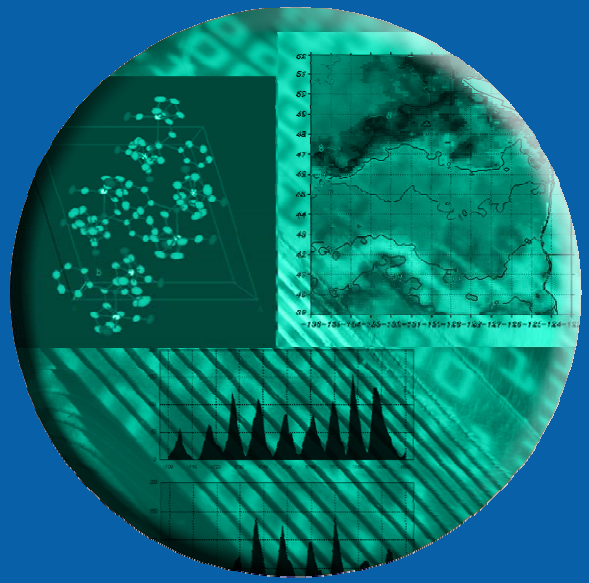
"Azul has been able to pack an industry-leading 24 processor cores on a single-chip, which means that each processor is able to run 24 simultaneous parallel threads"

What will we do with this Compute Power?

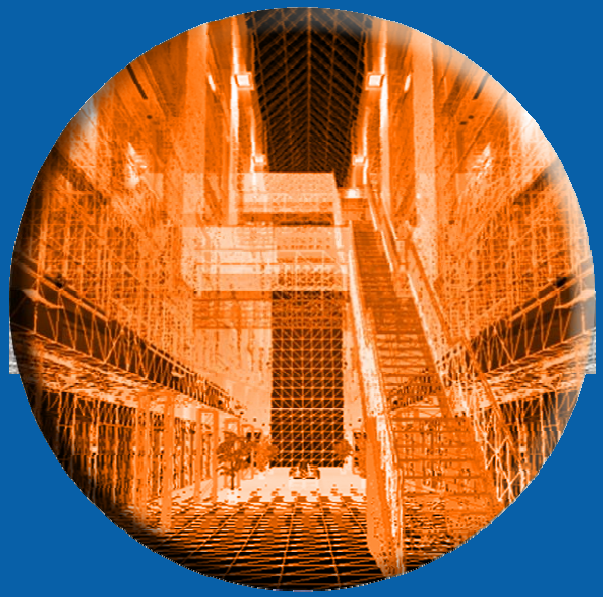
Recognition



Mining



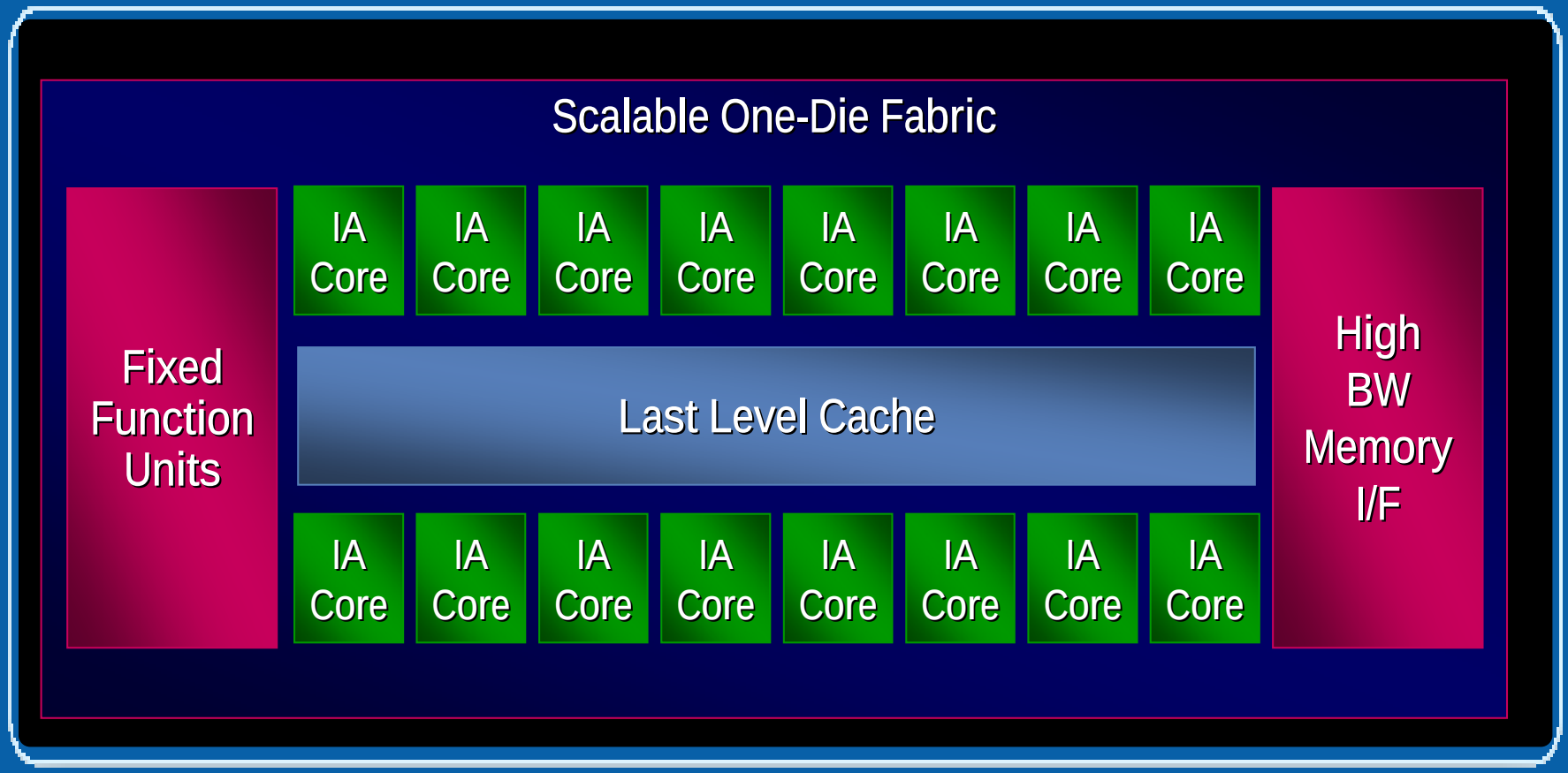
Synthesis



Emerging 'Killer' Applications
The RMS Suite

Source : "Cool Codes for Hot Chips" Keynote by Justin Rattner, CTO, Intel, Aug. 2006

Tera-Scale Prototype

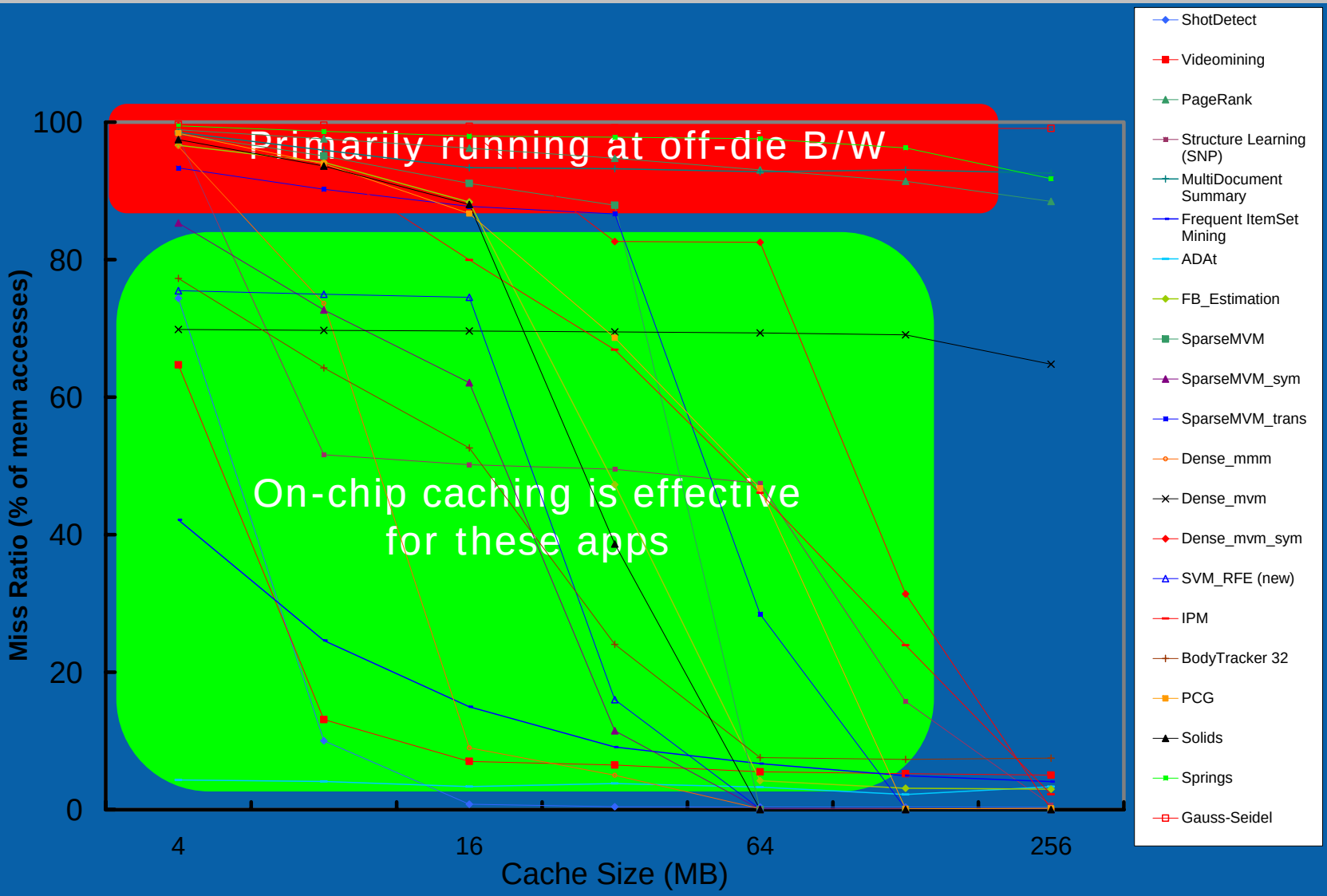


Source : "Cool Codes for Hot Chips" Keynote by Justin Rattner, CTO, Intel, Aug. 2006

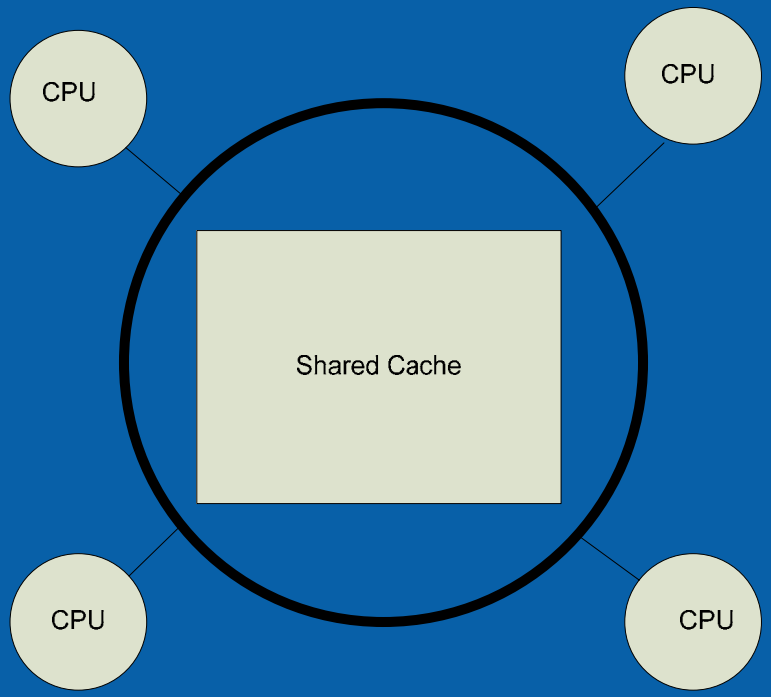
Overview of Talk

- Establish Importance of On-die Interconnects
- Walk through Case Study of a router design
- Evaluate against Goals
- Conclusions

iRMS Data Size estimates

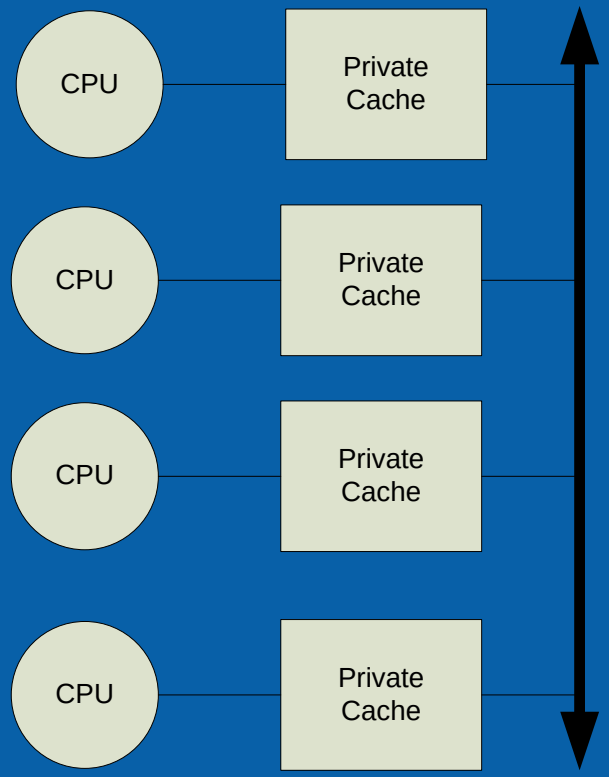


* Data collected on complete application run on a hardware cache emulator



No data replication

All data goes over on-die interconnect



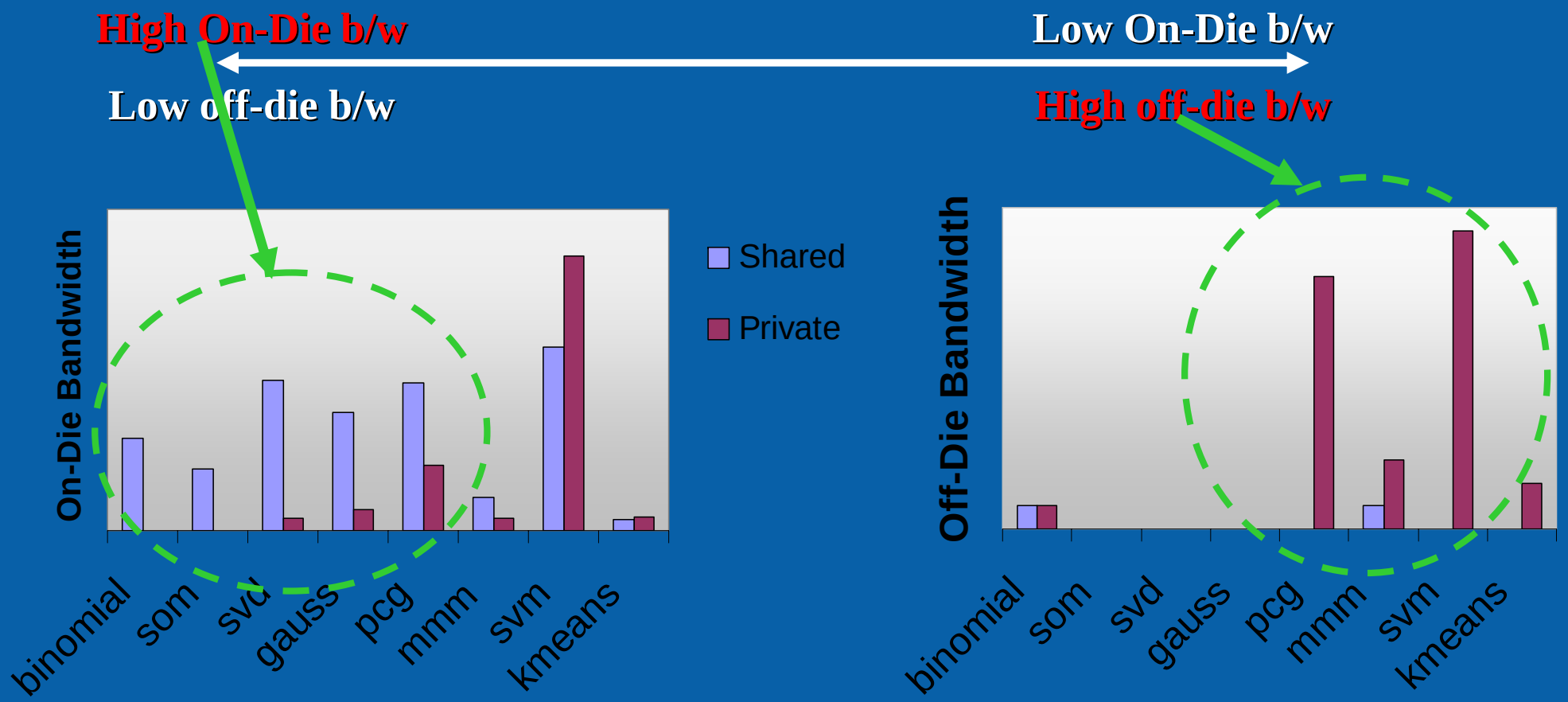
Possible data replication

primarily dirty blocks go over on-die interconnect

High On-Die B/W ⟷ **Low On-Die B/W**

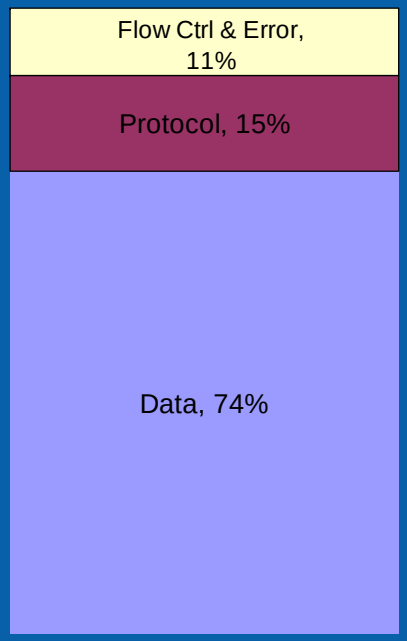
Low off-die B/W ⟷ **High off-die B/W**

sharing exists in some of the RMS kernels



Manage Off-Die bandwidth via better On-Die Network

Need for Scalability

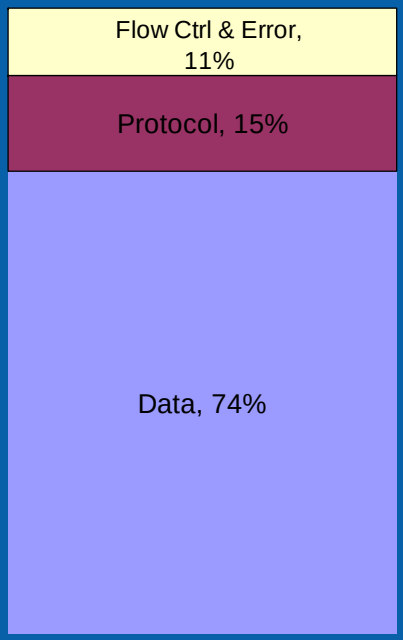


Bandwidth Components

Bandwidth Growth over time

- Data grows with cores
- Protocol grows faster than cores
- Error growing due to process

Need for Scalability

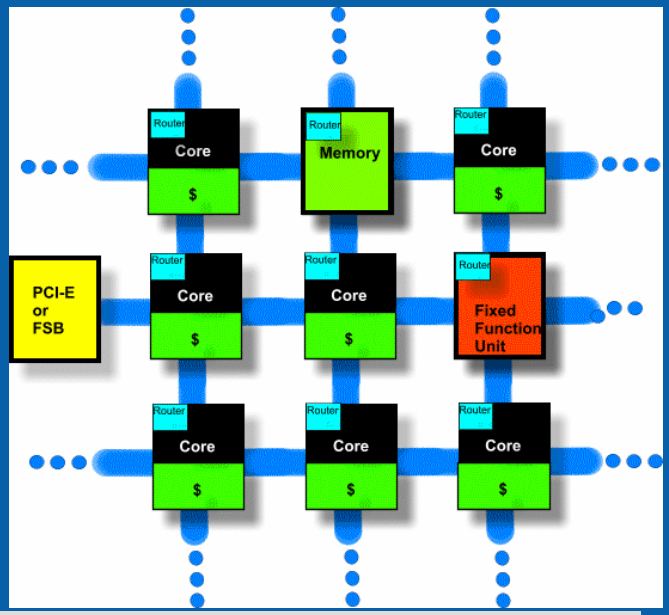


Bandwidth Components

Bandwidth Growth over time

- Data grows with cores
- Protocol grows faster than cores
- Error growing due to process

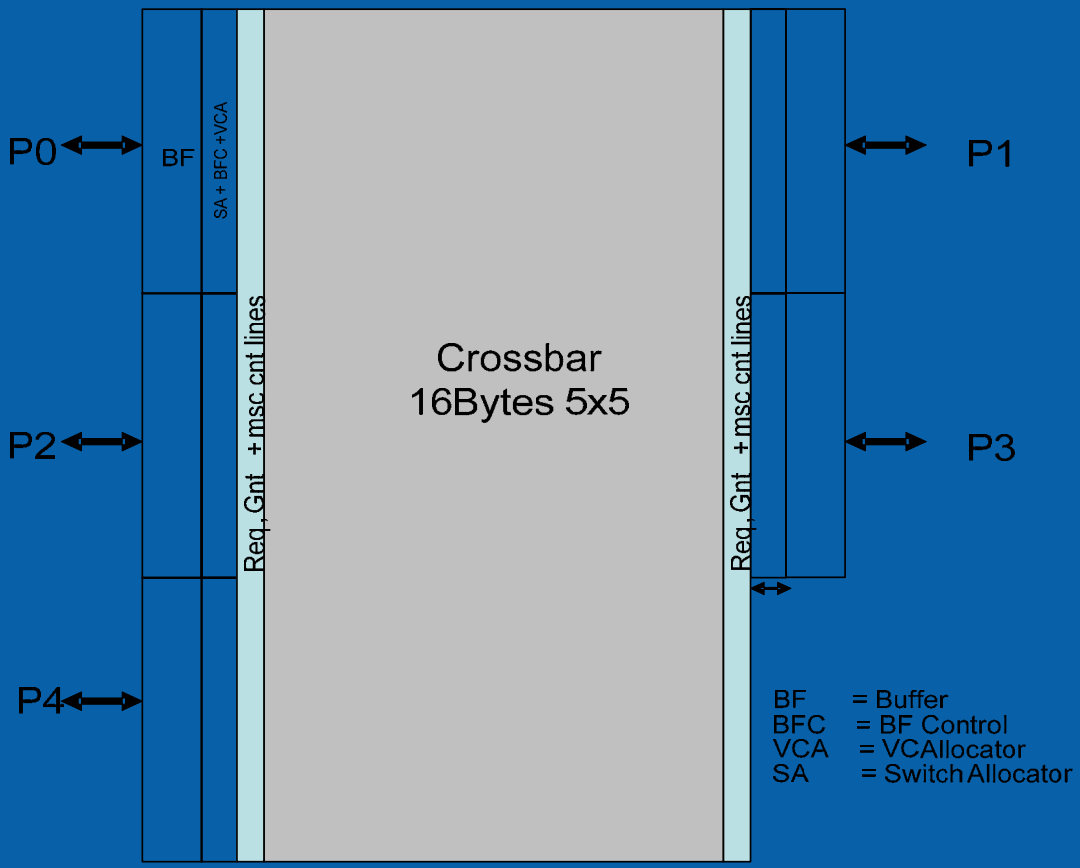
Need scalable network

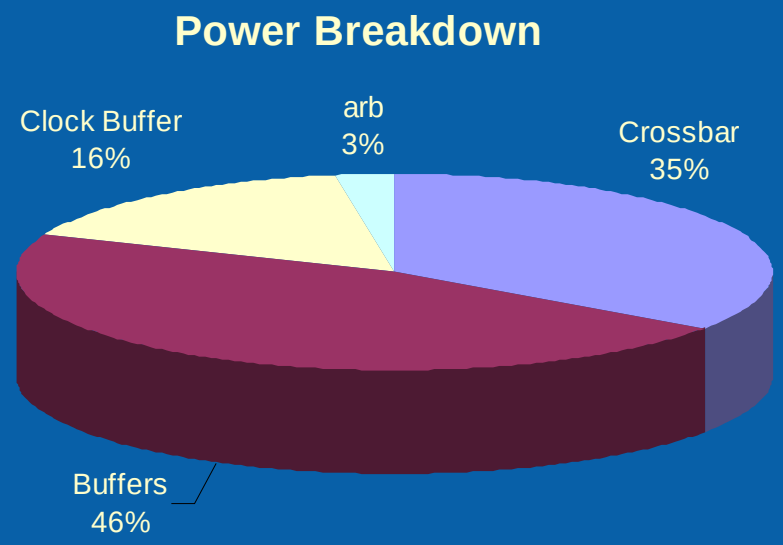


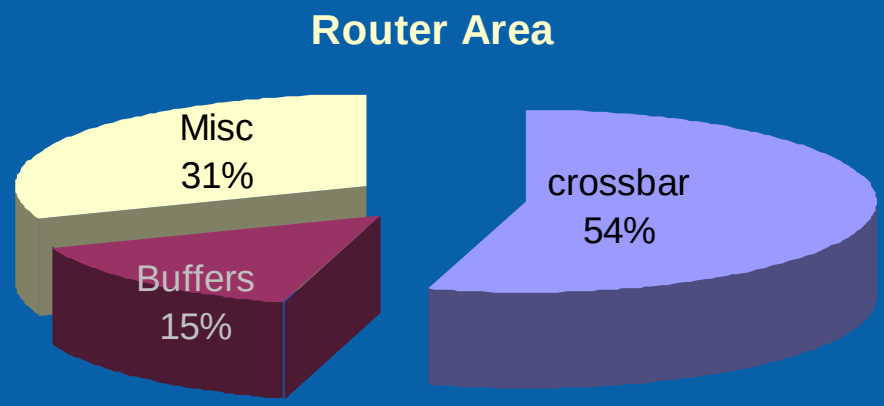
Network Parameters	
Size	6x6 mesh
Link Sizing	16B, >3Ghz
Traffic Classes	Request, response, data
Data Block Size	64 Bytes
Switching & Flow Control	Wormhole w/VC flow control
Error Control	end-to-end

Case Study of a Router

5-port Switch (overview)



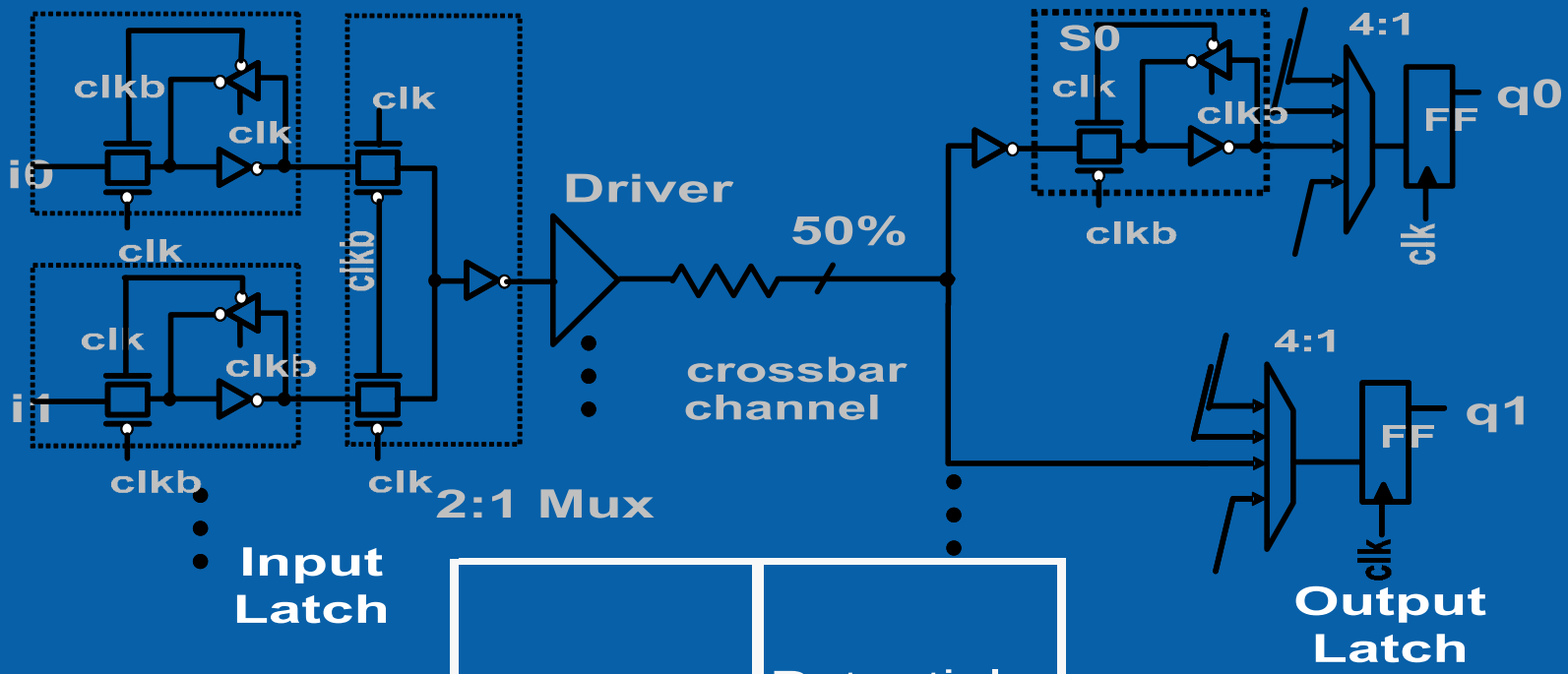




Design/uArchitecture Goals:

- **Reduce Crossbar area (and power)**
- **Reduce Buffer power**
- **Maximize throughput of network**

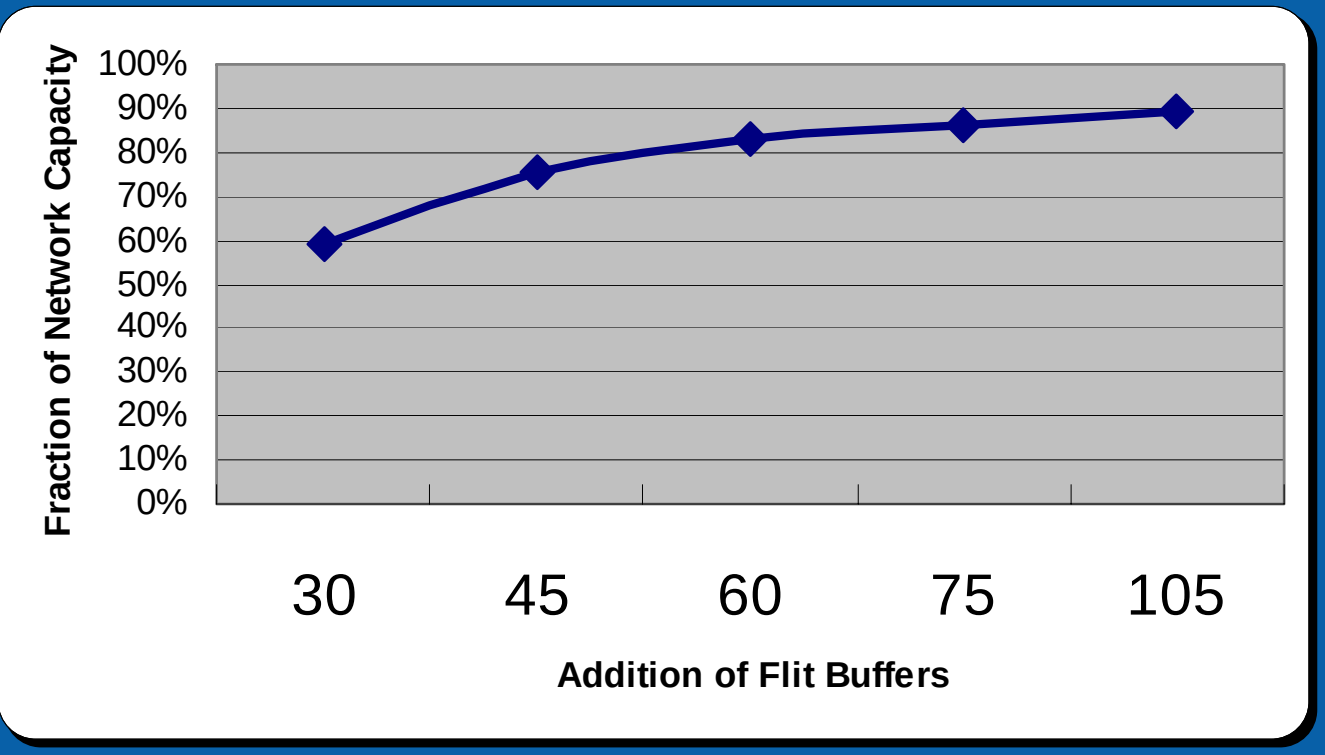
Double-pumped Crossbar



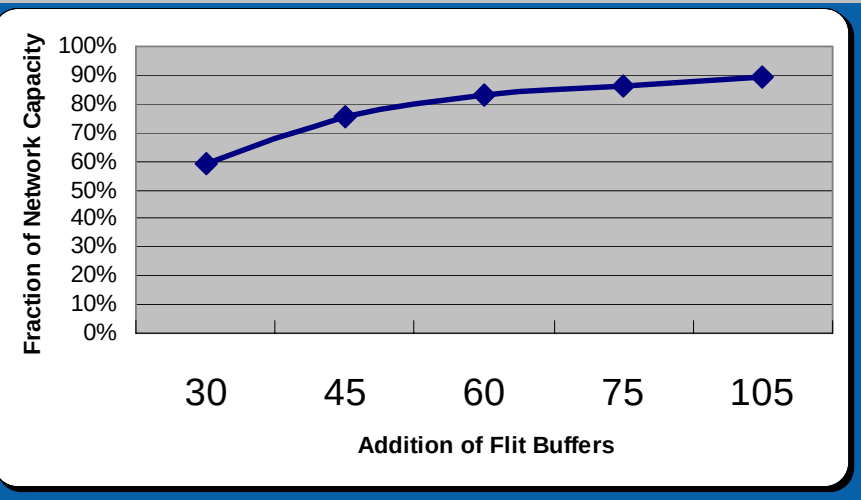
	Potential Reduction
Channel Width	50%
Channel Area	25%
Channel Power	17%
Channel Delay	17%

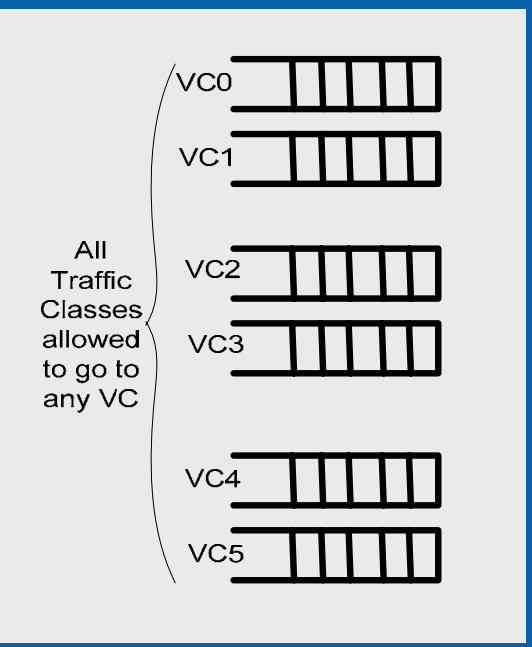
Source : Vangal et al "A six-port 57GB/s double pumped non-blocking router core" Sym. On VLSI Circuits, June 2005

Buffer Management



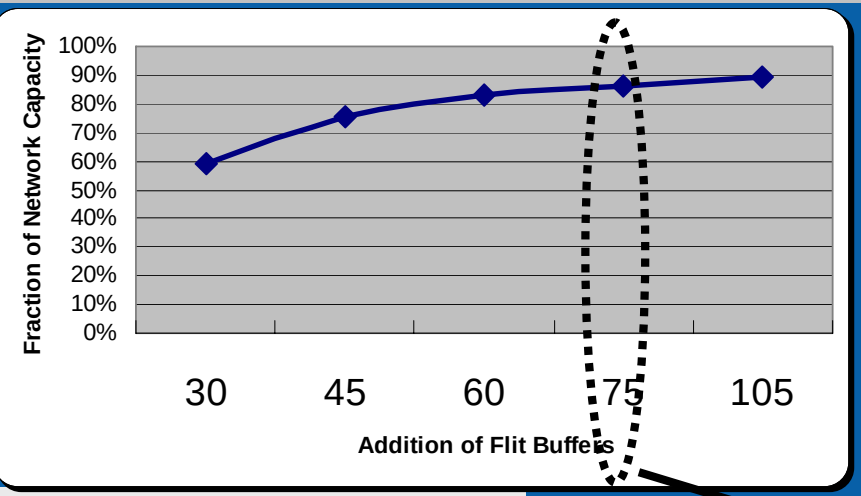
Buffer Management

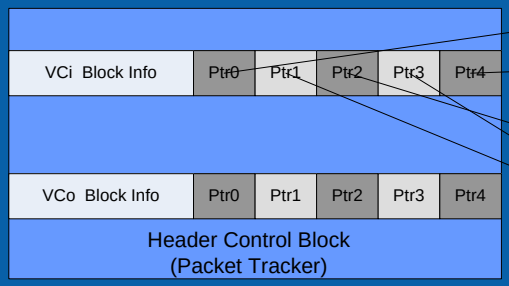


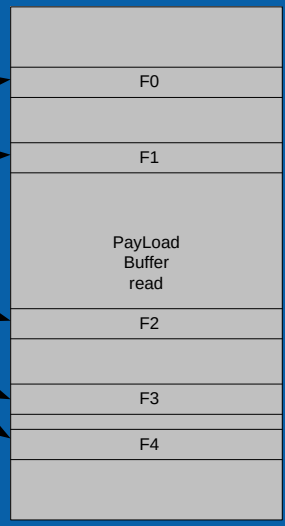


Statically Assigned Buffers
SAMQ with simple (VCT) flow control

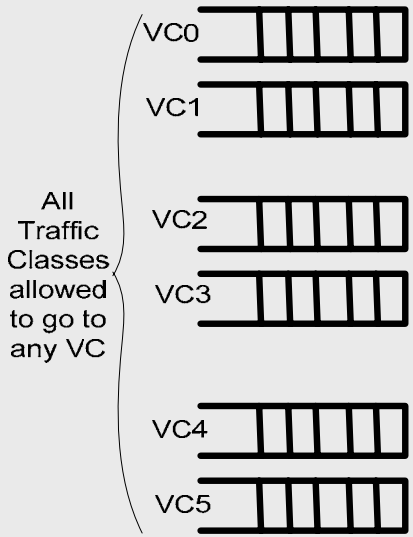
Buffer Management



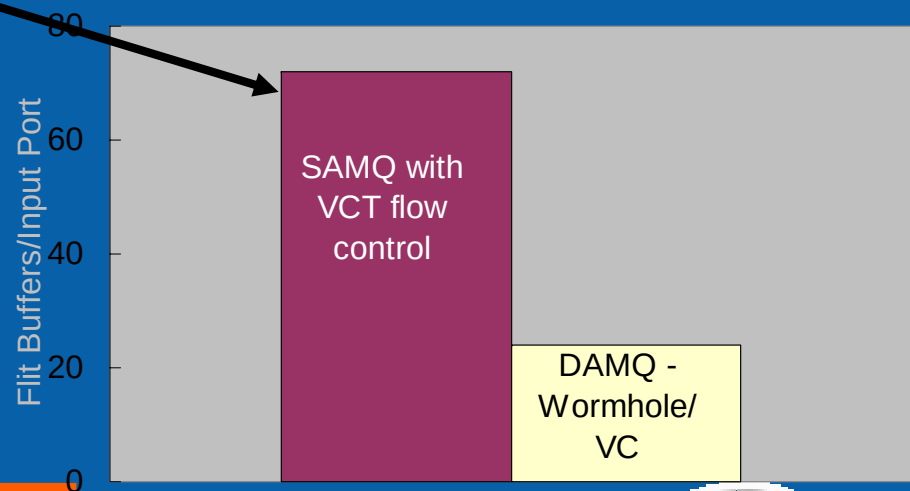




Dynamically Assigned Buffers
DAMQ-WormHole with
Virtual Channel Flow Control

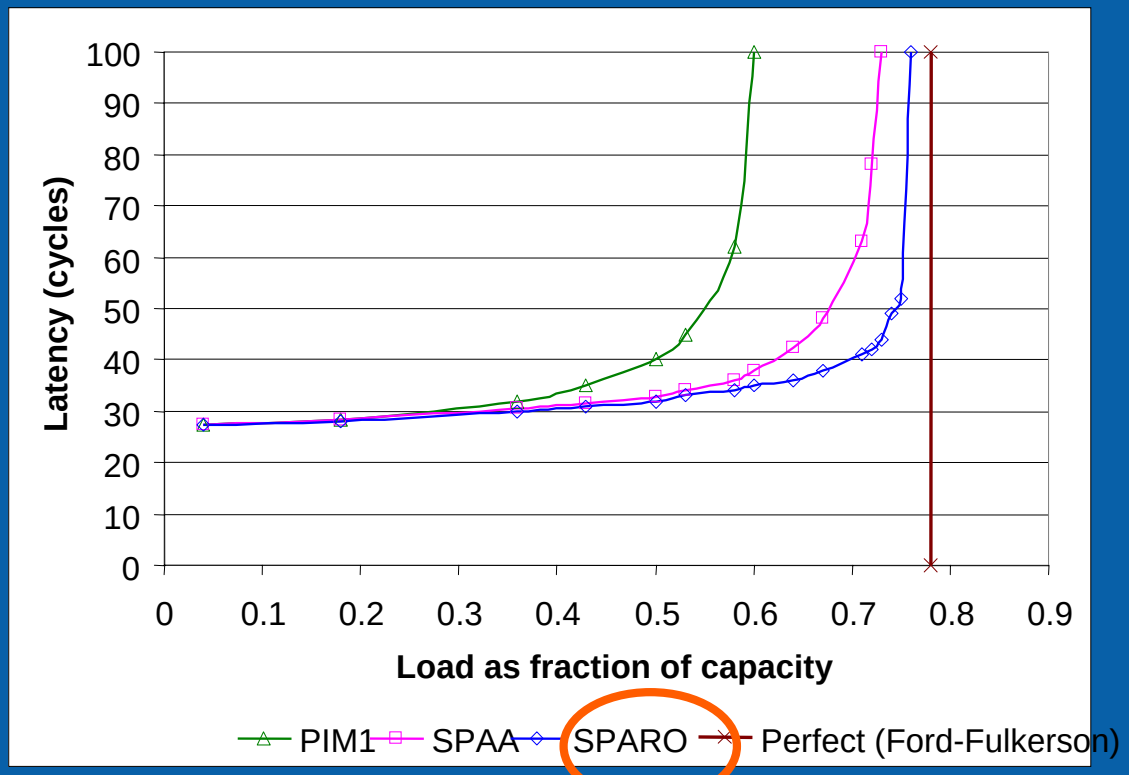


Statically Assigned Buffers (SAMQ with VCT flow control)



Achieve High Throughput
@ low(er) power/area

Switch Allocator



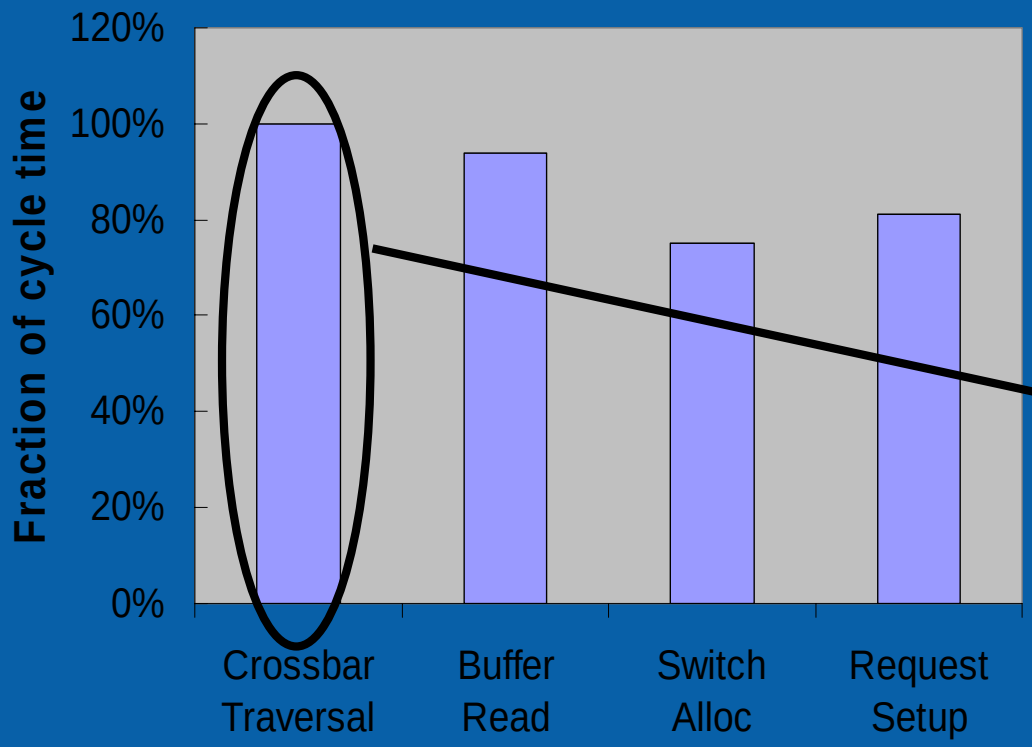
- Need to generate 4 requests per cycle
- Adapts to load conditions using heuristic

Achieve High Throughput @ manageable latency

Proprietary Switch Allocator achieves high matching efficiency

Pipeline Design

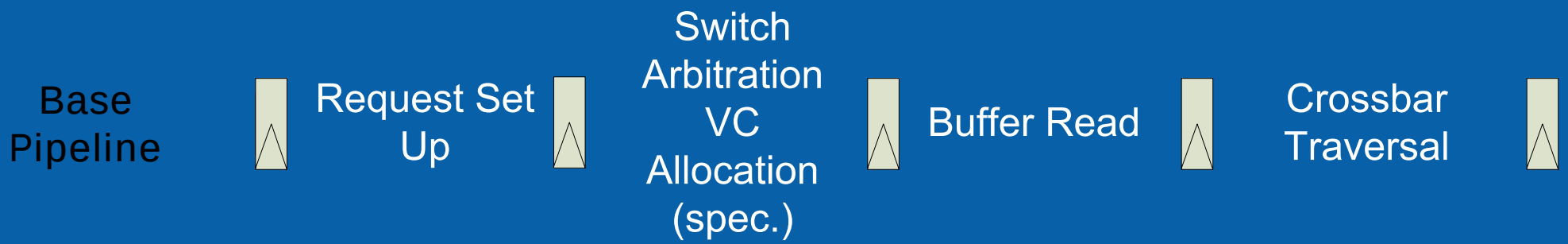
Request Set Up | Switch Arbitration VC Allocation (spec.) | Buffer Read | Crossbar Traversal



- 4-stage pipeline
- Buffer Read not in parallel with Switch Arbitration

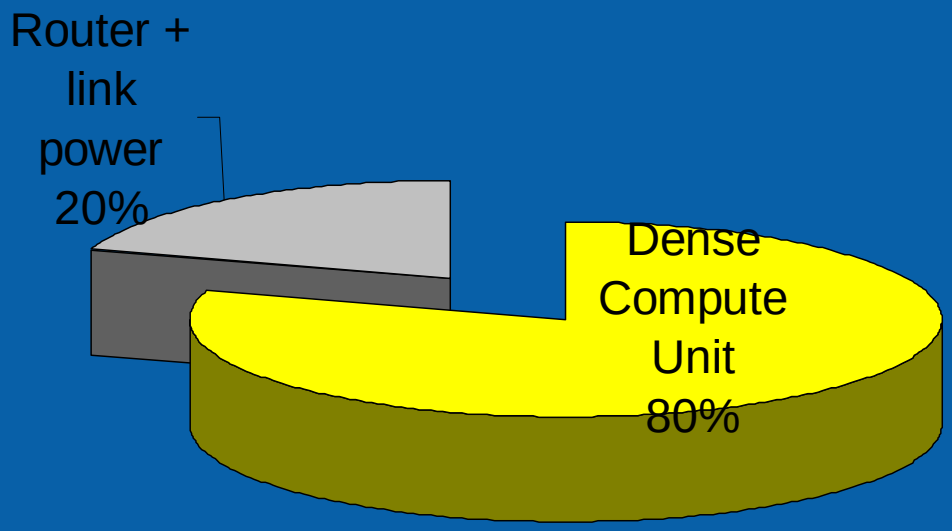
Crossbar traversal sets the cycle time

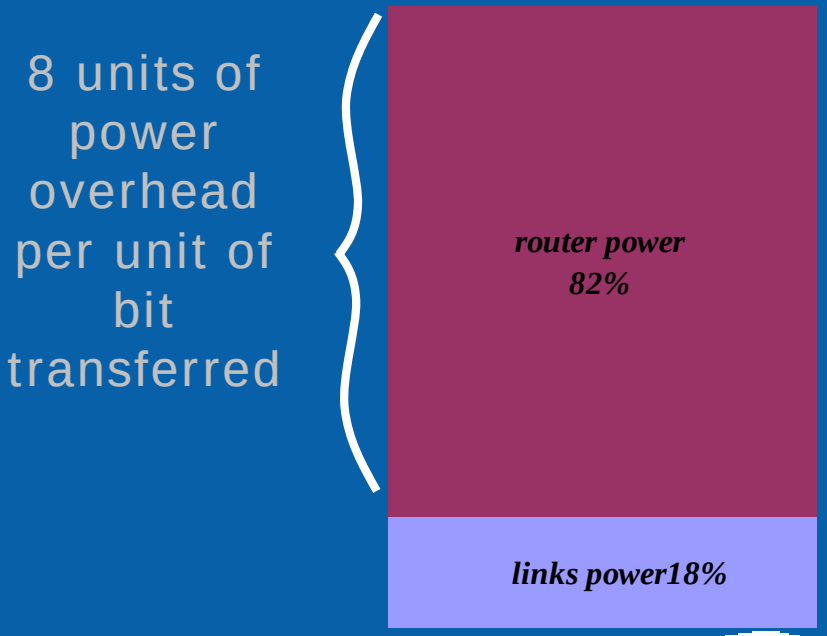
Pipeline Design



- Choose Pipeline frequency to Maximize Switching rate
- Optimize for load conditions

Power Challenges for ODI





Interconnect Power Currently Exceeding budget!

Miscellaneous Issues

- Increased Soft Error and Process Variability impacts design
  - design to detect and/or correct errors (latency, bandwidth impact)
  - routing for fault tolerance

- Clocking power is high (16%)
  With wide links cost of GALS approaches may be higher

Conclusions

- Scalable High Performance on-die interconnect would be required in future CMPs
- We do achieve high network throughput
  Many of the techniques are borrowed from previous research
- But significant challenge is to fit within power and area

Acknowledgments

Co-Leads : Jay Jayasimha, Yatin Hoskote

Aniruddha Vaidya, Sriram Vangal, Arvind P. Singh, Chris Hughes, Y-K Chen, Ioannis Schoinas, Akhilesh Kumar, Sailesh Kottapalli, Jeffrey Chamberlain, Li-Shiuan Peh, Amit Kumar, Niraj Jha