Comparative Evaluation of FPGA and ASIC Implementations of Bufferless and Buffered Routing Algorithms for On-Chip Networks

Yu Cai   Ken Mai   Onur Mutlu

Carnegie Mellon University

Motivation

- In many-core chips, on-chip interconnect (NoC) consumes significant power.

  - **Intel Terascale**: ~30%; **MIT RAW**: ~40% of system power

- Must devise low power routing mechanisms

SAFARI

Carnegie Mellon
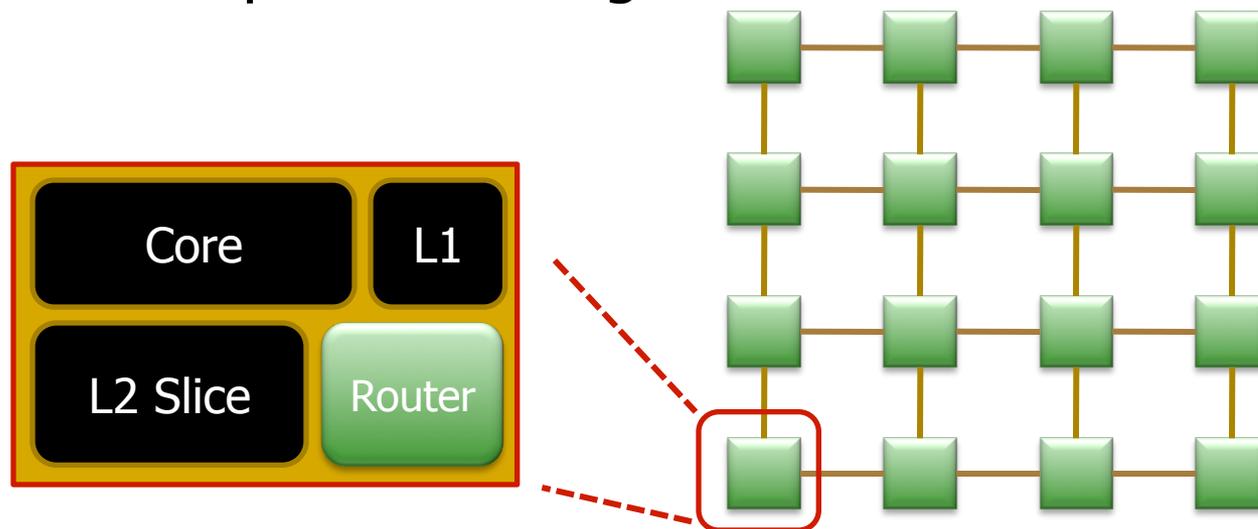
Bufferless Deflection Routing

- **Key idea**: Packets are never buffered in the network. When two packets contend for the same link, one is deflected. (BLESS [Moscibroda, ISCA 2009])

New traffic can be **injected** whenever there is a free output link.
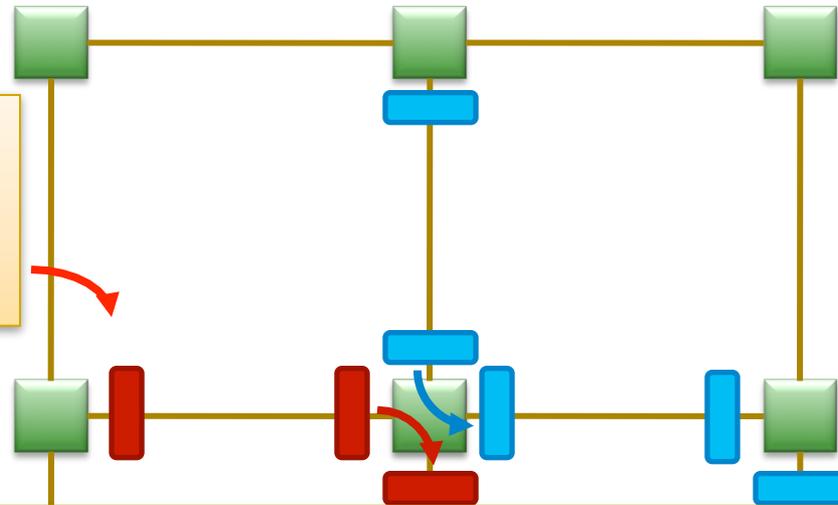
☞ No buffers ➔ lower power, smaller area

☞ Conceptually simple

Carnegie Mellon

Motivation

- Previous work has evaluated bufferless deflection routing in software simulator (BLESS [Moscibroda, ISCA 2009])

  - Energy savings
  - Area reduction
  - Minimal performance loss
  - Unfortunately: No real silicon implementation comparison
    - ➔ Silicon power, area, latency, and performance?

- **Goal**: Implement BLESS in FPGA for apple to apple comparison with **other buffered routing algorithms.**

SAFARI

Carnegie Mellon

Bufferless Router
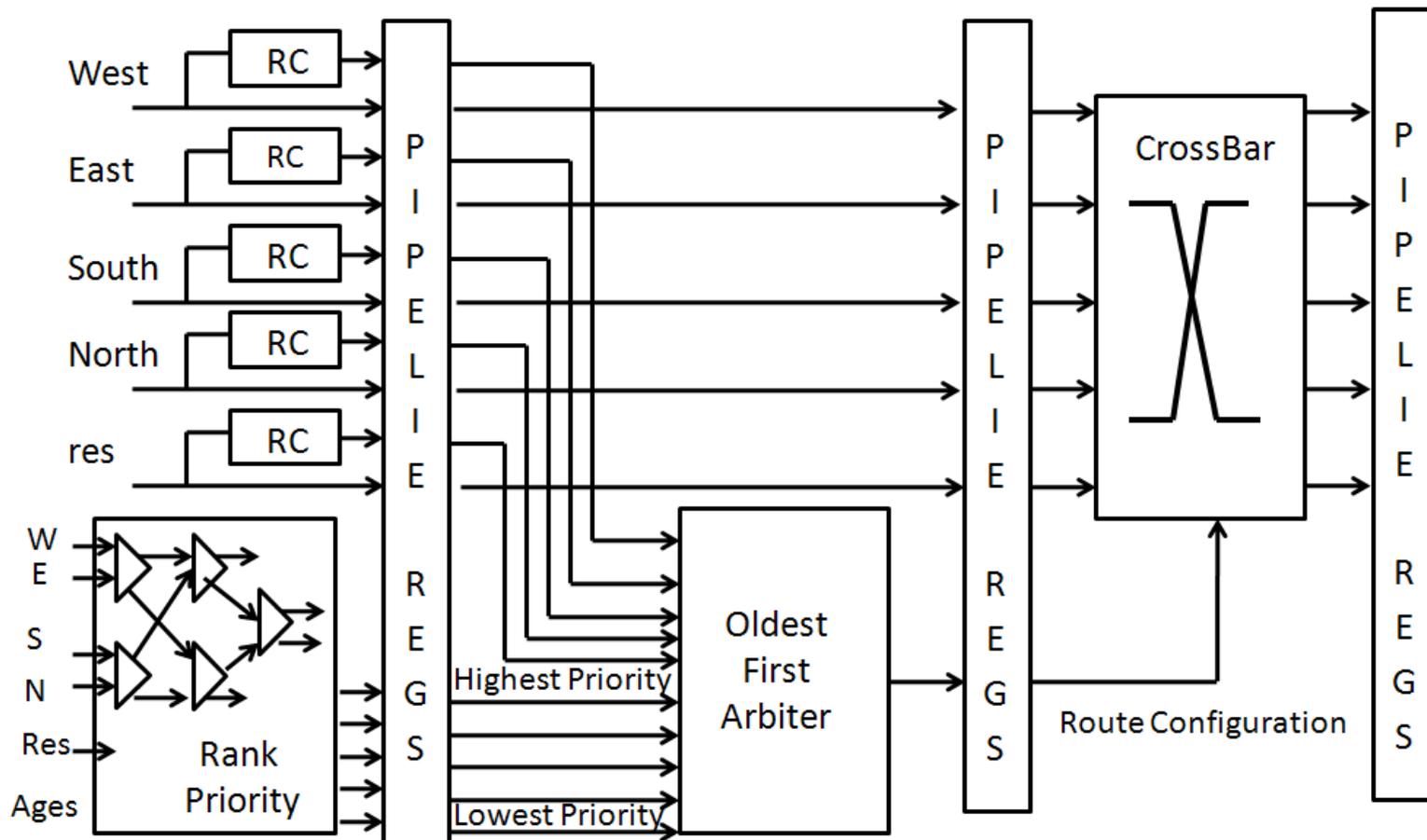
- Pipeline the oldest first algorithm into two stages
  - Rank priority and oldest first arbiter

Carnegie Mellon

Oldest-First Arbiter
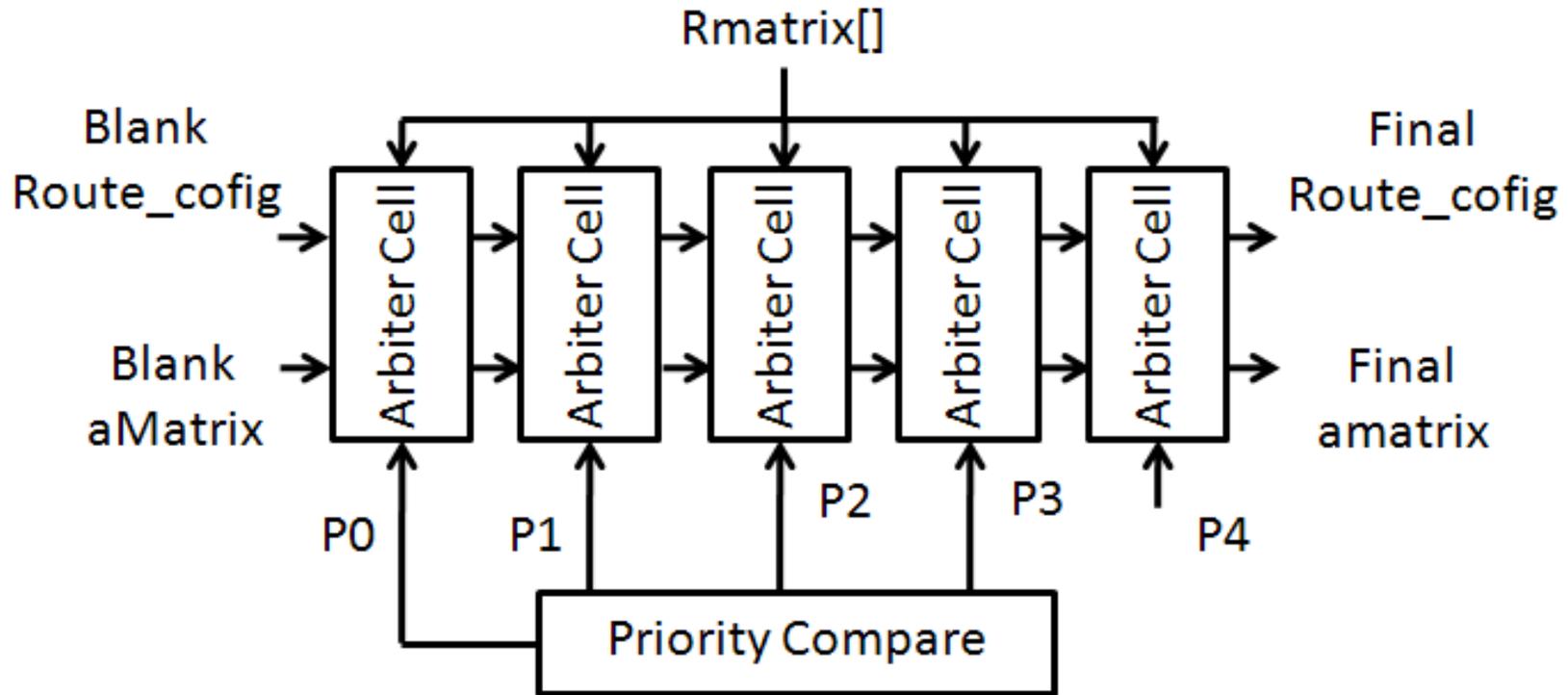
- Latency of arbiter cells linearly correlated with port number
- Latency of each arbiter cells also correlated with port number

SAFARI

Carnegie Mellon

Bufferless Router with Buffers
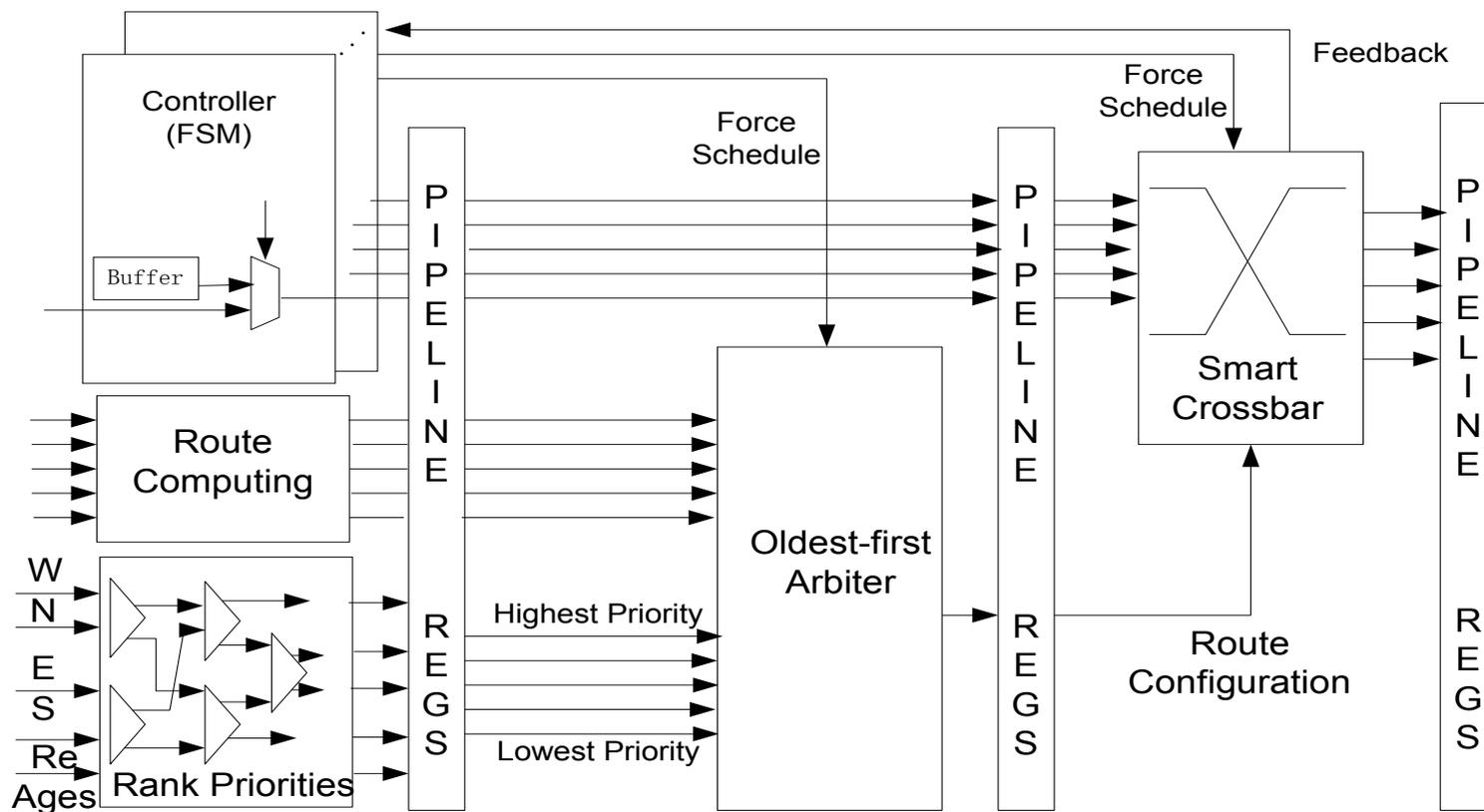
- Force schedule can force the flits to be assigned an output port, although it is not the desired port

Baseline buffered Router

Carnegie Mellon
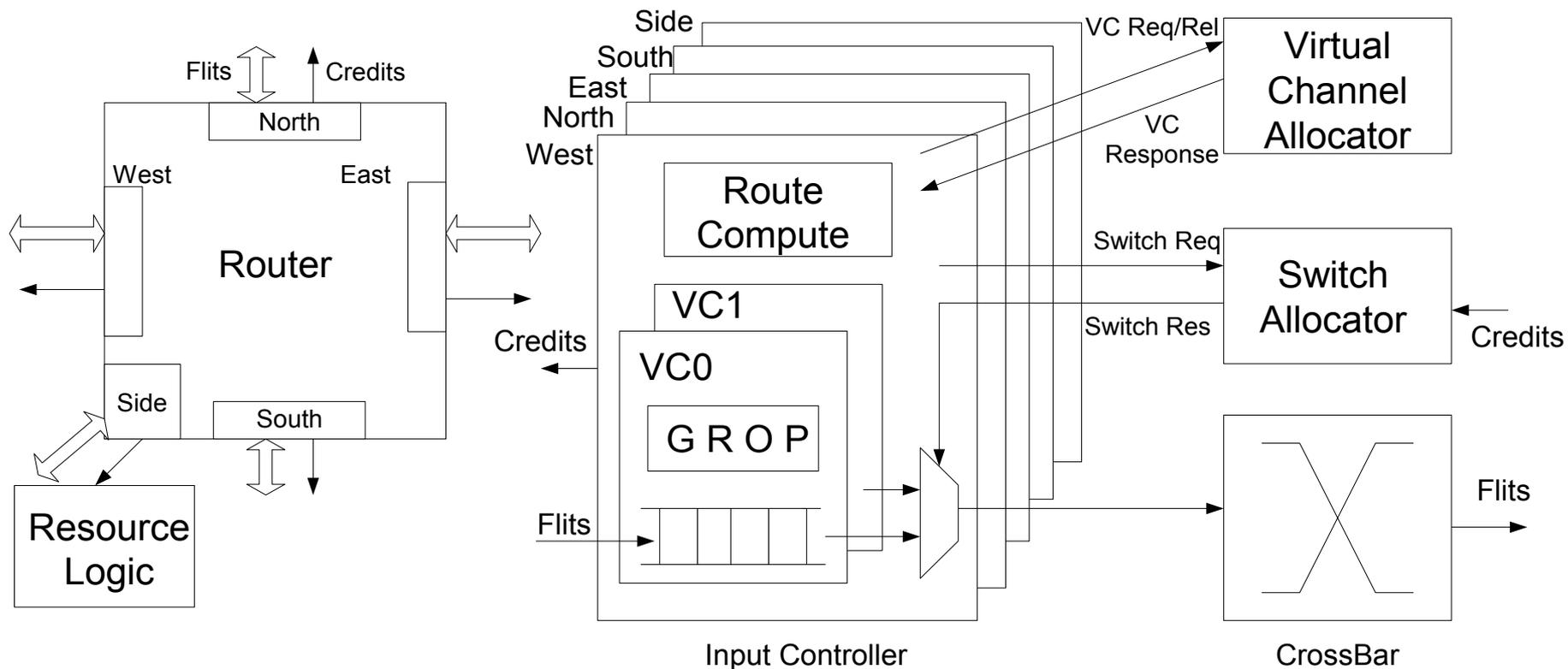
Emulation Platform Setup

- **FPGA platform (Berkeley Emulation Enginee II)**
  - Virtext2Pro FPGA
- **Network topology**
  - 3x3 and 4x4 Mesh network
  - 3x3 and 4x4 Torus network
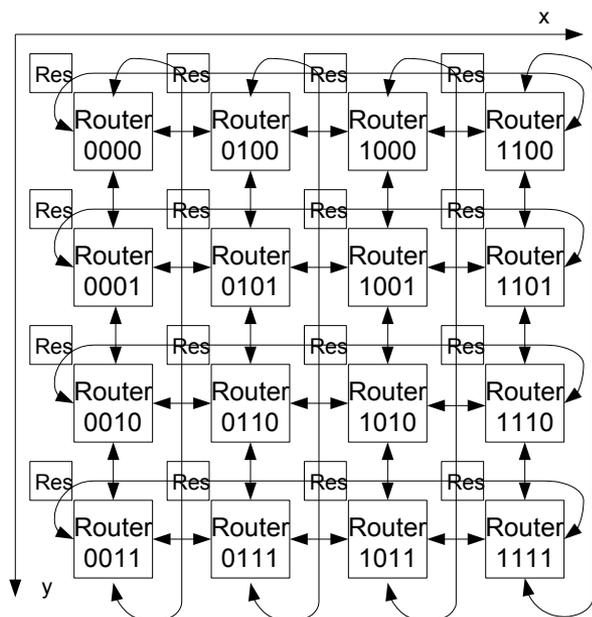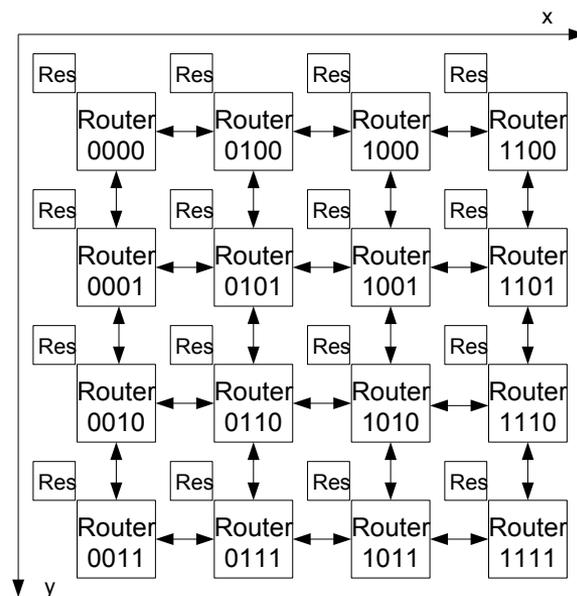
(a) 4x4 Torus Network   (b) 4x4 Mesh Network

Emulation Configuration

- **Router**
  - Buttered router with 1, 2 and 4 virtual channels
  - Bufferless router
 - Bufferless router with one flit buffer in the input port
- **Traffic patterns**

TABLE I. Evaluated traffic patterns

UR	Router chooses a random destination among other routers with equal probability and sends a packet to that destination. The probability is equal among the other routers
NN	Each node sends a packet to one of its immediate neighbors with equal probability
TOR	Node $\{X, Y\}$ sends packets to node $\{X+k/2-1, y\}$ mod k for the k-ary network (k=4)
TR	Node $\{X, Y\}$ sends packets to node $\{Y, X\}$
BC	Node with address $\{b0, b1, b2, b3\}$ in bits sends packets to the node with address NOT$\{b0, b1, b2, b3\}$ in bits
HS	All the nodes send the packet to a certain single node. The hot spot can act as receiver only or can be transmitter and receiver.

Carnegie Mellon

FPGA Area Comparison

TABLE II. FPGA area usage of various routers

	BLESS	BLESS(Buf)	1-VC	2-VC	4-VC
Slices	634	930 (2%)	989 (2%)	2098(6%)	4638(14%)
FF	335	695 (1%)	947 (1%)	1764(2%)	3480(5%)
LUTs	1090	1671 (2%)	1392(2%)	3133(4%)	7362(11%)

Carnegie Mellon
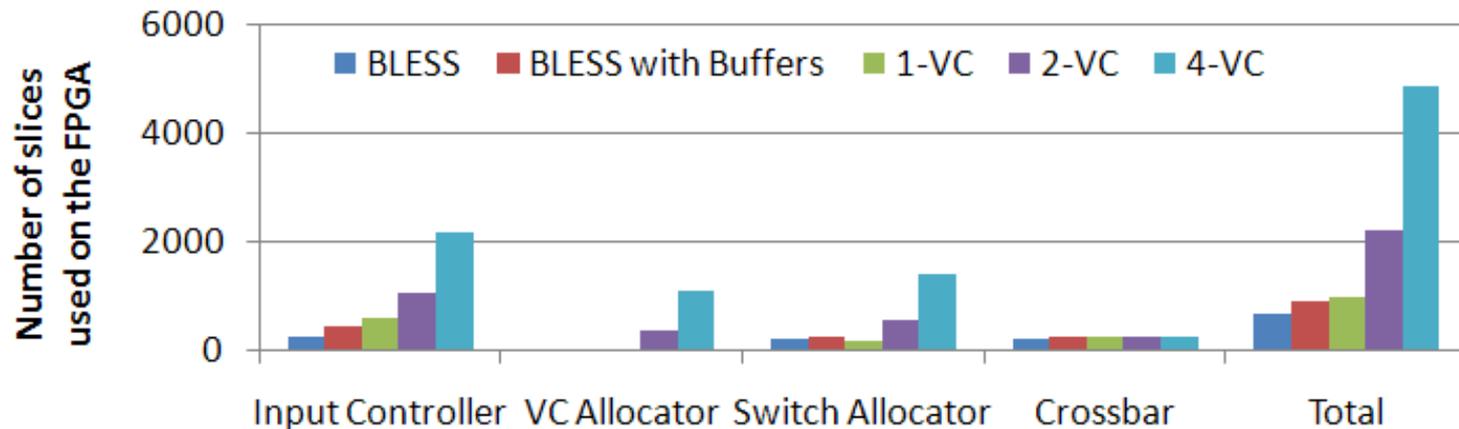
FPGA Frequency Analysis

TABLE III. Critical path length of various routers

	BLESS	BLESS(Buf)	1-VC	2-VC	4-VC
Delay (ns)	13.285	17.335	15.948	34.081	71.788
Freq. (MHz)	75.27	57.69	62.704	29.342	13.93

Carnegie Mellon
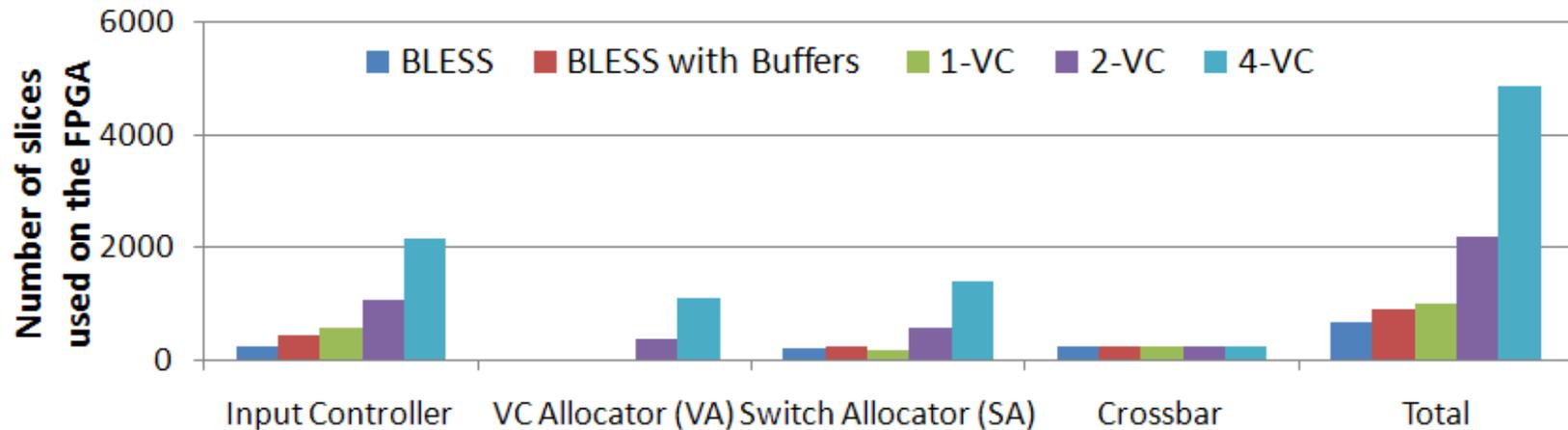
FPGA Power Analysis
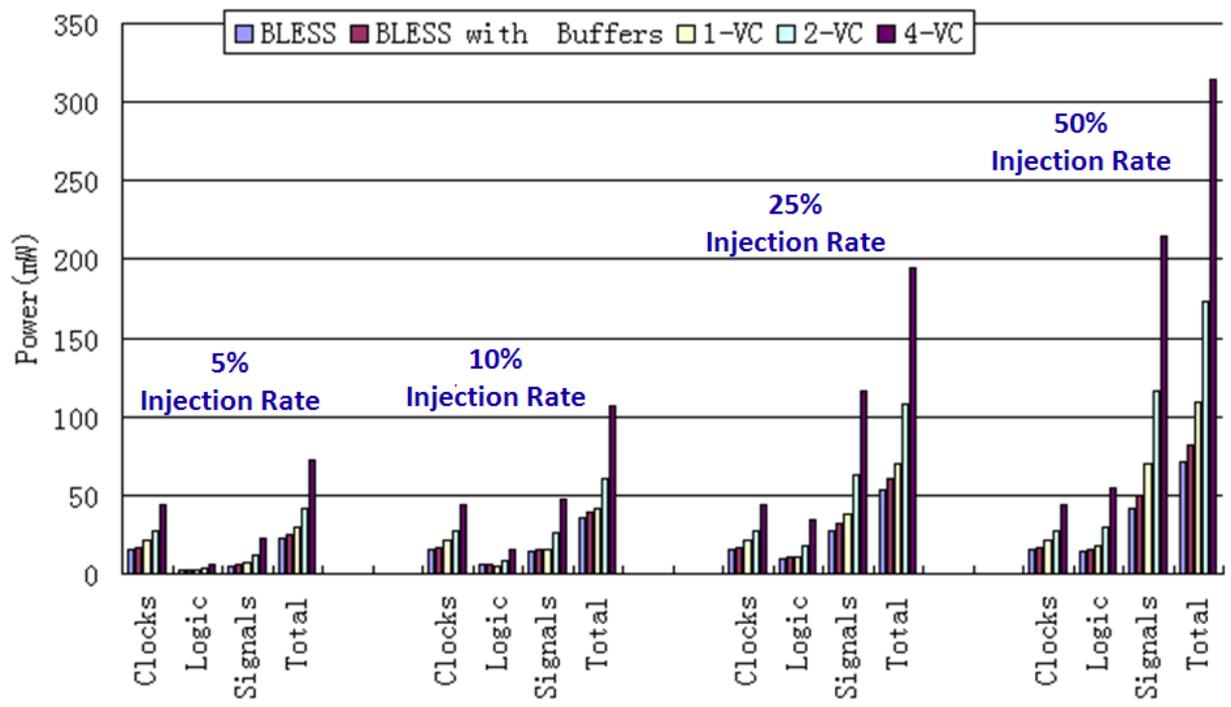
Network Power Analysis
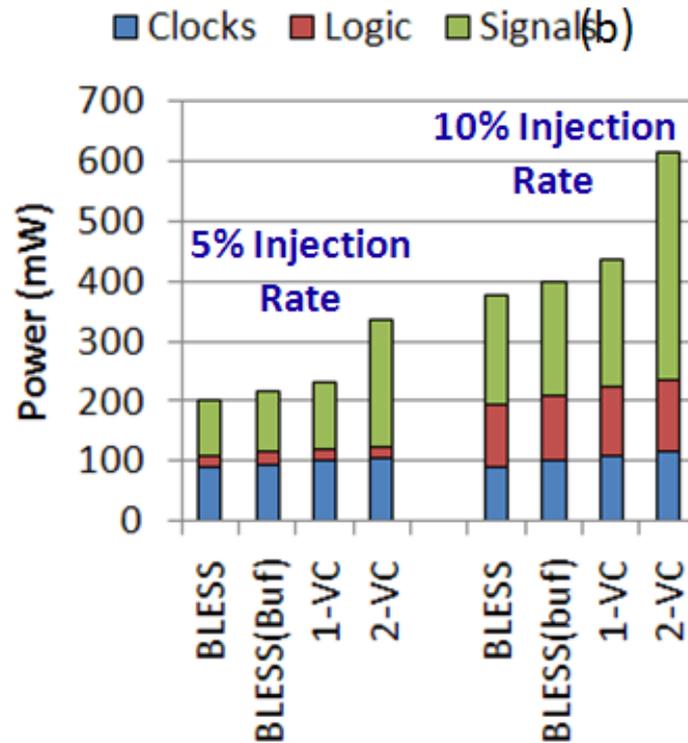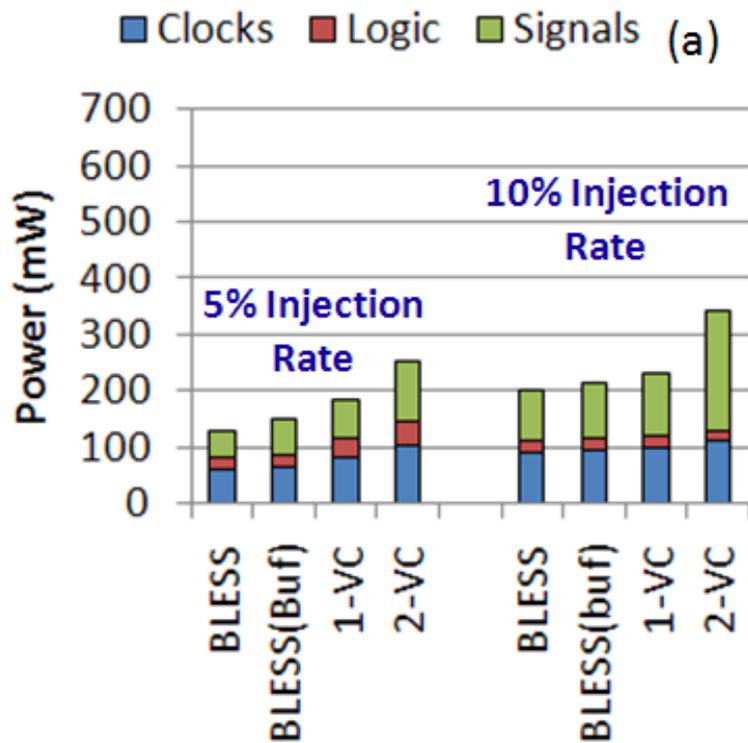
Packet Latency (4x4 Torus)
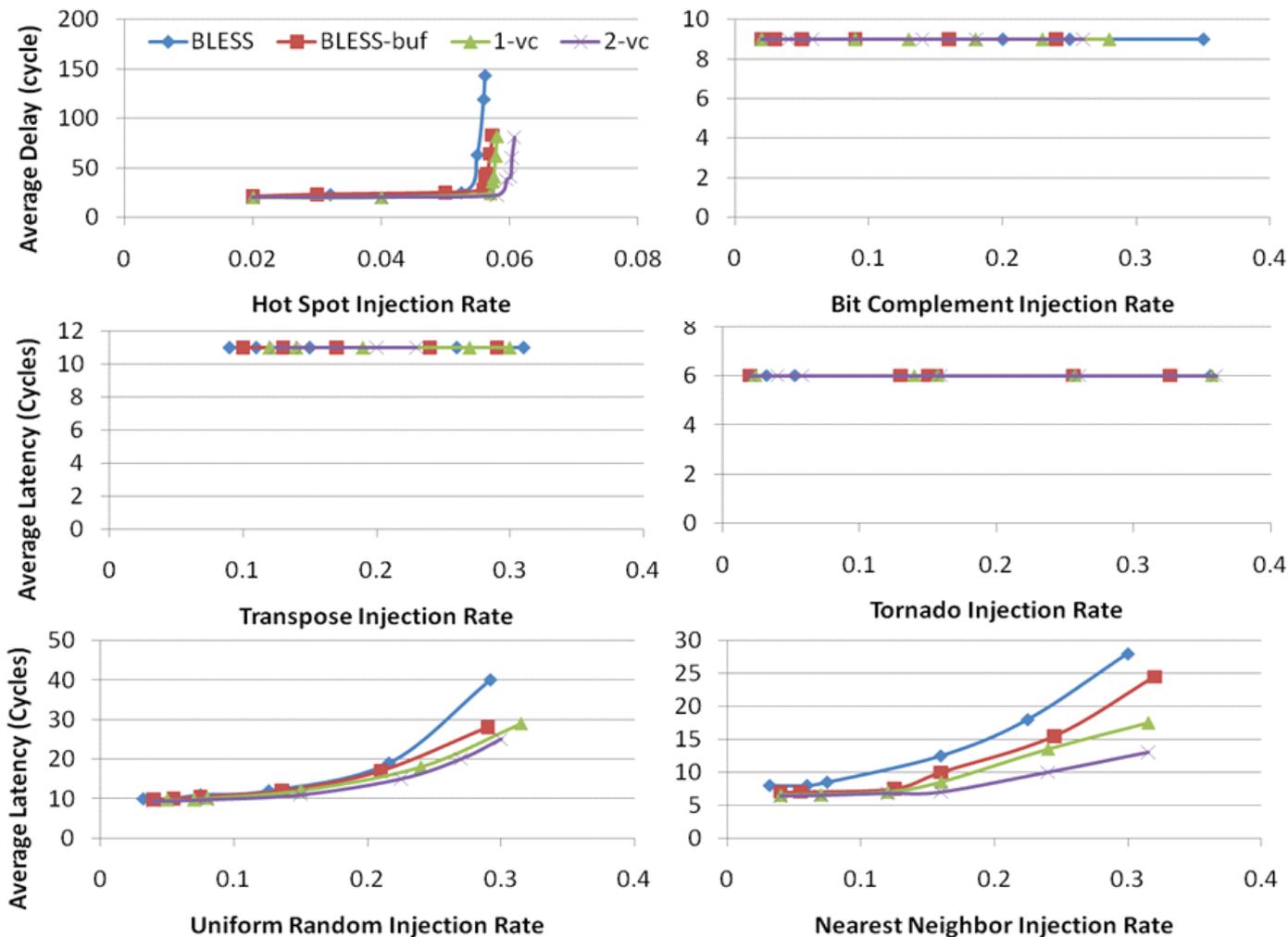
Packet Latency (4x4 Mesh)
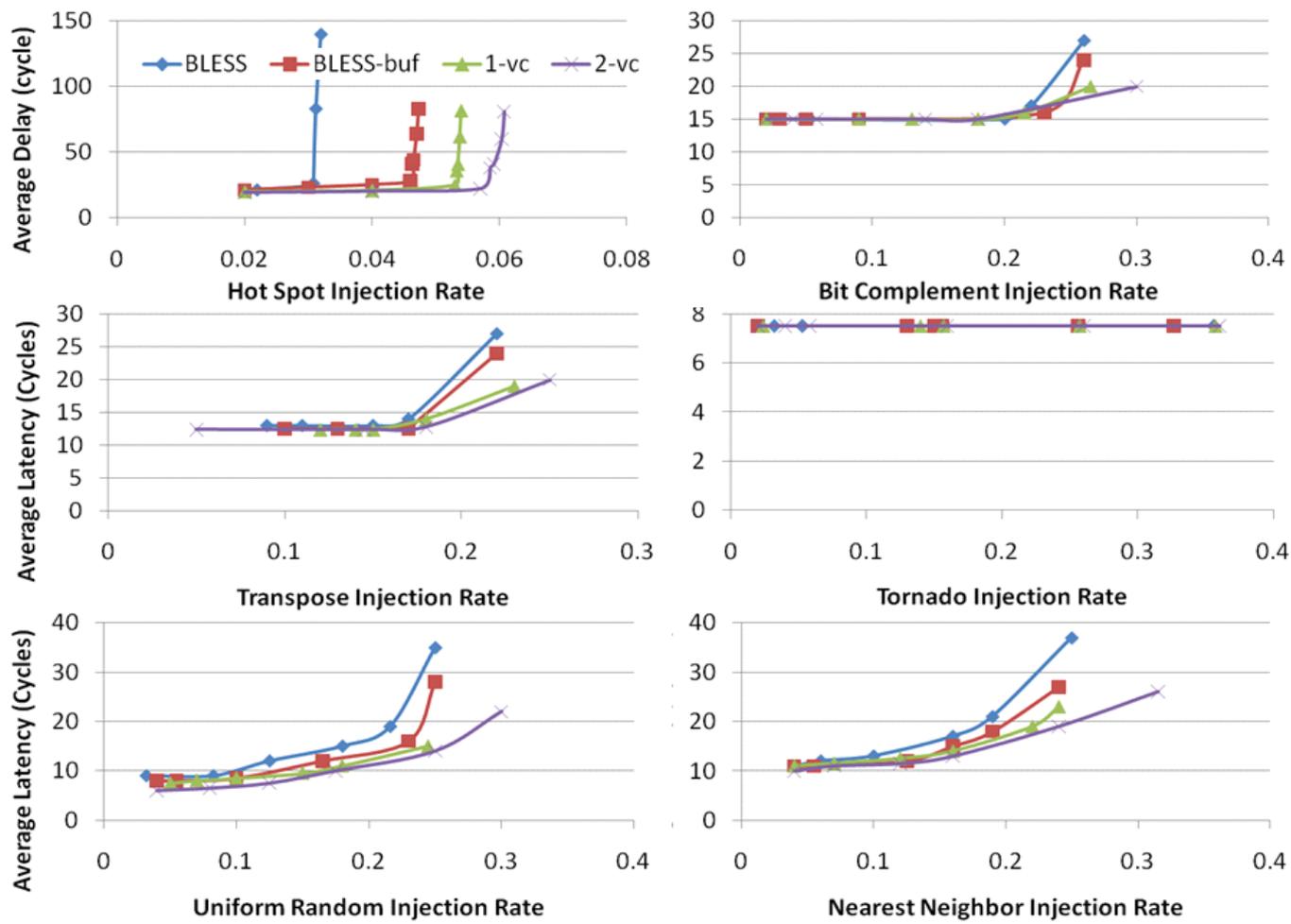
ASIC Implementation results

Table IV: 65nm ASIC results for all the routers

Router	Area (μm^2)	Power (mW)	Max Frequency
BLESS	12353	2.14	214.6MHz (4.66ns)
BLESS (Buf)	16854	2.74 (2.61)	204.5MHz (4.89ns)
1-VC	20139	3.03 (2.79)	197.2MHz (5.07ns)
2-VC	36152	5.12 (2.40)	100.4MHz (9.93ns)
4-VC	79900	10.11 (2.37)	50.3MHz (19.9ns)

SAFARI

Carnegie Mellon

Conclusion

- First realistic and detailed implementation of deflection-based bufferless routing in on-chip networks

- Bufferless routing is efficiently implementable in mesh and torus based on-chip networks, leading to significant area (38%), power consumption (30%) reductions over the best buffered router implementation on a 65nm ASIC design

- We identify that oldest-first arbitration, which is used to guarantee livelock freedom of bufferless routing algorithms, as a key design challenge of bufferless router design
  - With careful design, oldest-first arbitration can be pipelined and can outperform virtual channel arbitration

Carnegie Mellon

Thank You.

Comparative Evaluation of FPGA and ASIC Implementations of Bufferless and Buffered Routing Algorithms for On-Chip Networks

Yu Cai   Ken Mai   Onur Mutlu

Carnegie Mellon University