ECE 8823 A / CS 8803 - ICN

Interconnection Networks

Spring 2017

http://tusharkrishna.ece.gatech.edu/teaching/icn_s17/

Lecture 9: Flow Control - III

Tushar Krishna

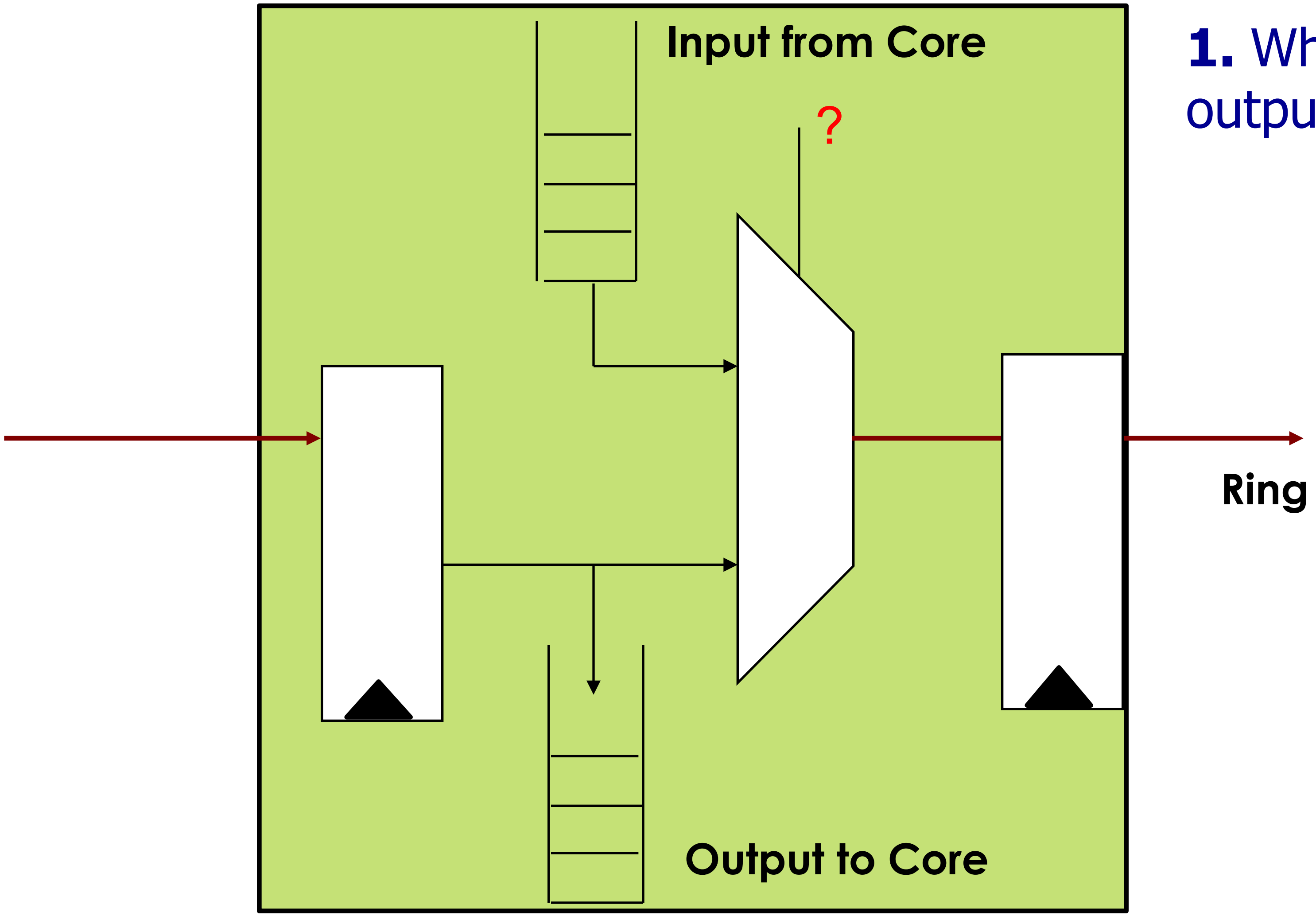
Assistant Professor
School of Electrical and Computer Engineering
Georgia Institute of Technology

tushar@ece.gatech.edu

Acknowledgment: Slides adapted from Univ of Toronto ECE 1749 H (N Jerger)

Designing a Flow Control Protocol: Managing Buffers and Contention

Flow Control Protocol: Arbitration

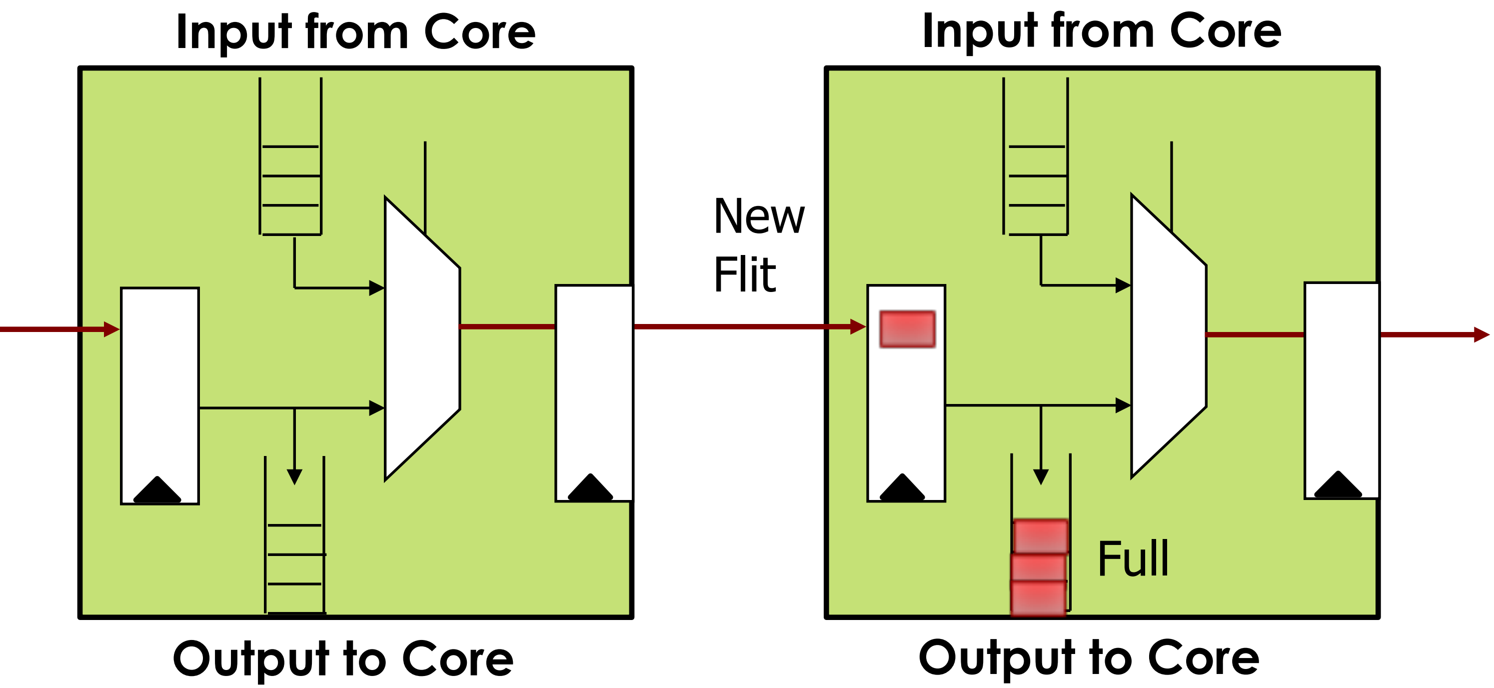


1. Who should use output link?

Flow Control Protocol: Backpressure

2. What to do with the other flit (from ring/core)

3. What should a flit do if its output is blocked?



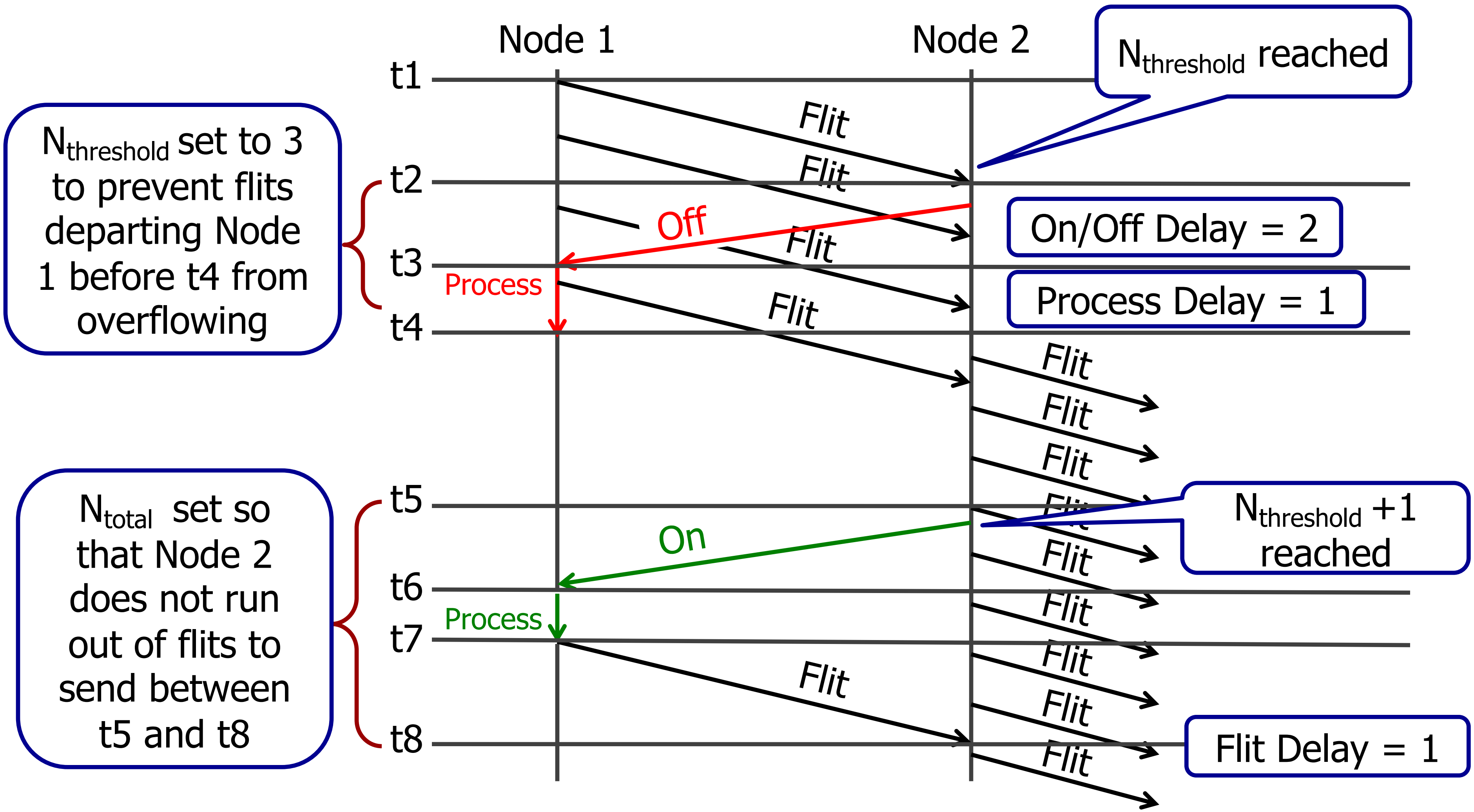
Backpressure Signaling Mechanisms

- **On/Off Flow Control**
  - downstream router signals if it can receive or not
- **Credit-based Flow Control**
  - upstream router tracks the number of free buffers available at the downstream router

On/Off Flow Control

- Downstream router sends a 1-bit on/off if it can receive or not
  - Upstream router sends only when it sees on
- Any potential challenge?
  - Delay of on/off signal
  - By the time the on/off signal reaches upstream, there might already be flits in flight
  - Need to send the off signal *once the number of buffers reaches a threshold* such that all potential in-flight flits have a free buffer

On/Off Timeline with N buffers



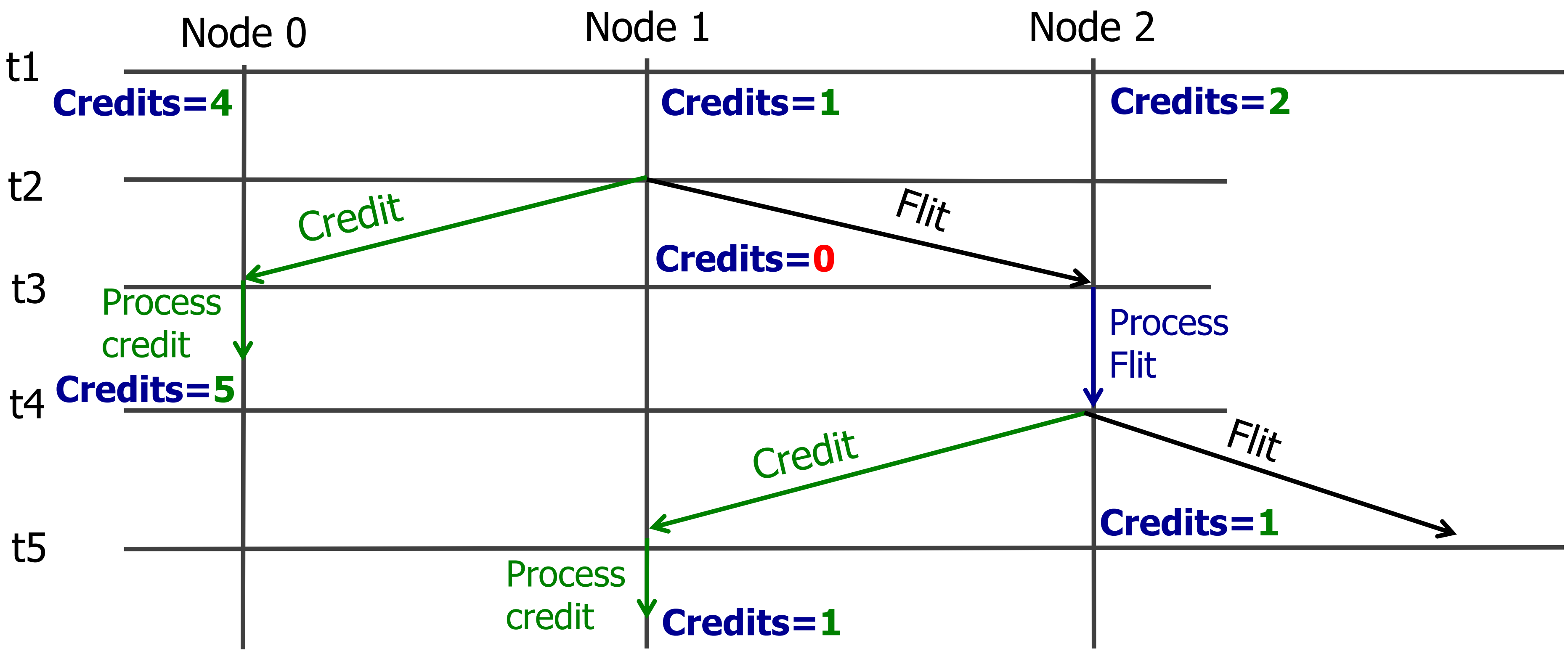
Backpressure Signaling Mechanisms

- **On/Off Flow Control**
  - Pros
    - Low overhead: one-bit signal from downstream to upstream node, only switches when threshold crossed
  - Cons
    - Inefficient buffer utilization – have to design assuming worst case of $N_{threshold}$ flights in flight

Credit-based Flow Control

- **Upstream router** tracks the **number of free buffers available at the downstream router**
  - Upstream router sends only if credits > 0
- When should credit be decremented at upstream router?
  - When a flit is sent to the downstream router
- When should credit be incremented at upstream router?
  - When a flit leaves the downstream router

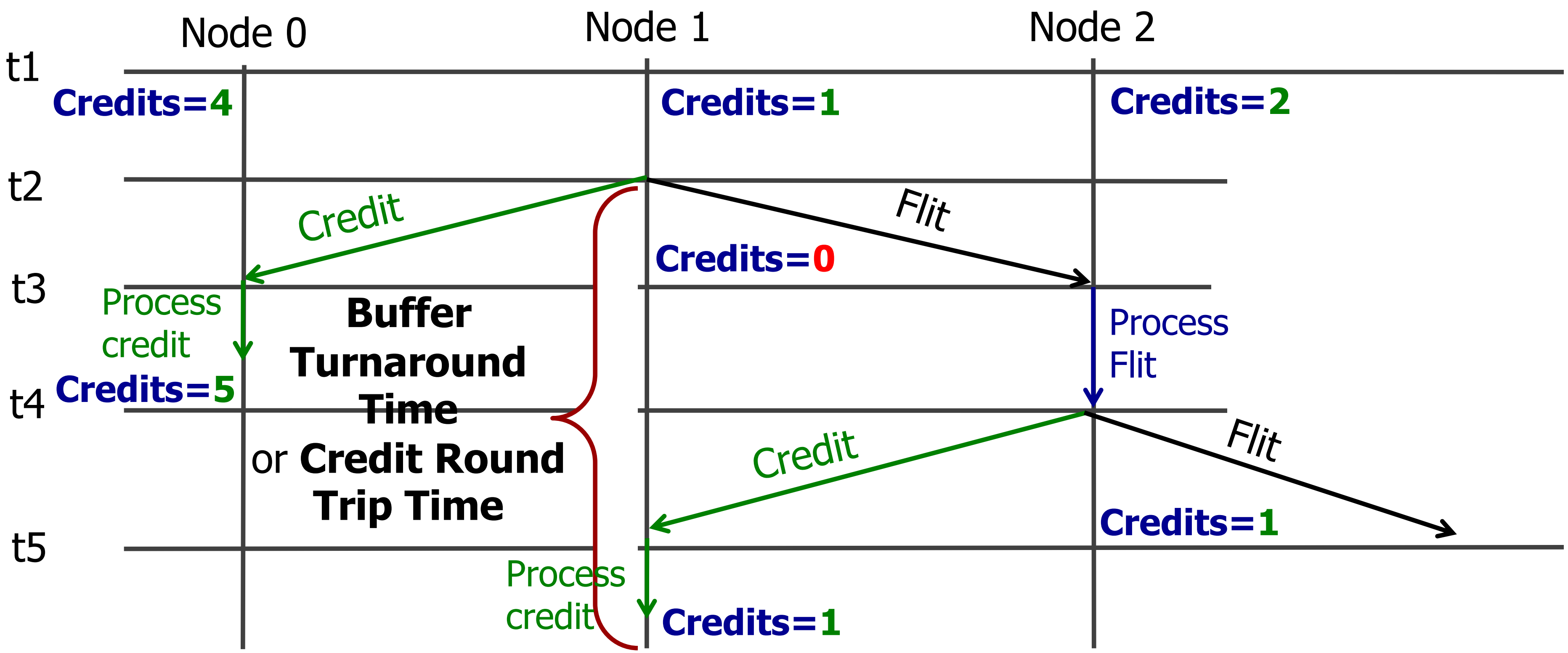
Credit Timeline



Backpressure Signaling Mechanisms

- **On/Off Flow Control**
  - Pros
    - Low overhead: one-bit signal
  - Cons
    - Inefficient buffer utilization – have to design assuming worst case of $N_{threshold}$ flights in flight
- **Credit Flow Control**
  - Pros
    - Each buffer fully utilized - an keep sending till credits are zero (unlike on/off)
  - Cons
    - More signaling – need to signal upstream for every flit

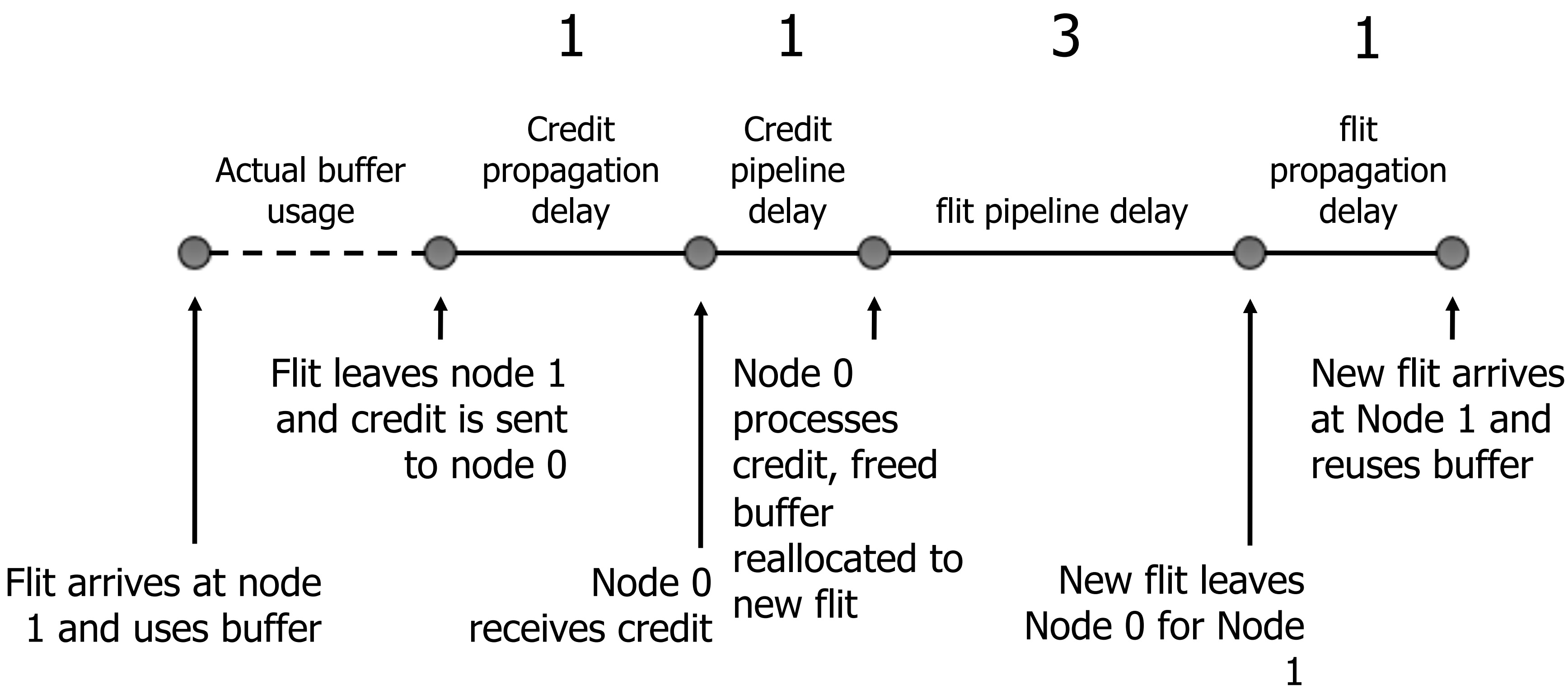
Backpressure and Buffer Sizing



No flit can be sent into this buffer during this delay

To prevent backpressure from limiting throughput, number of buffers >= turnaround time

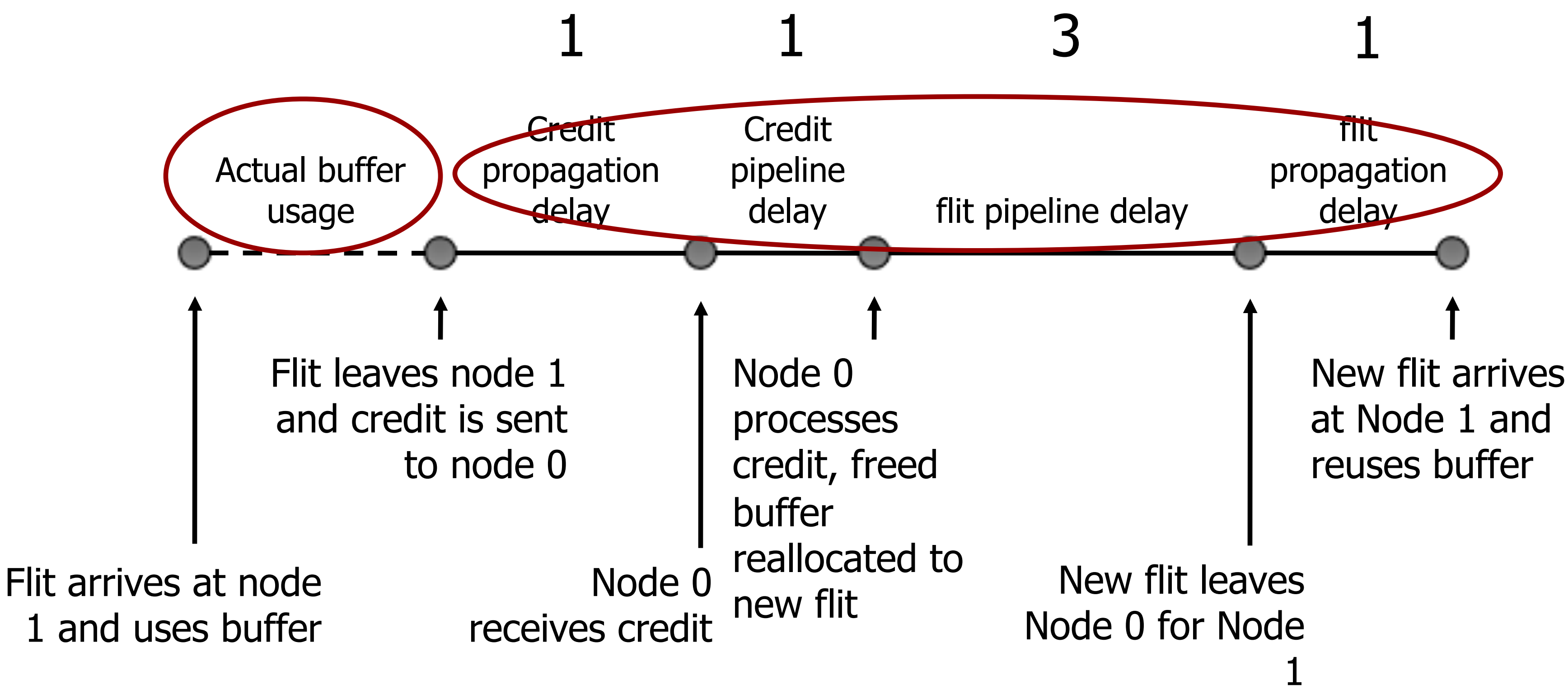
Buffer Turnaround Delay



How many buffers needed? 1+1+3+1 = 6

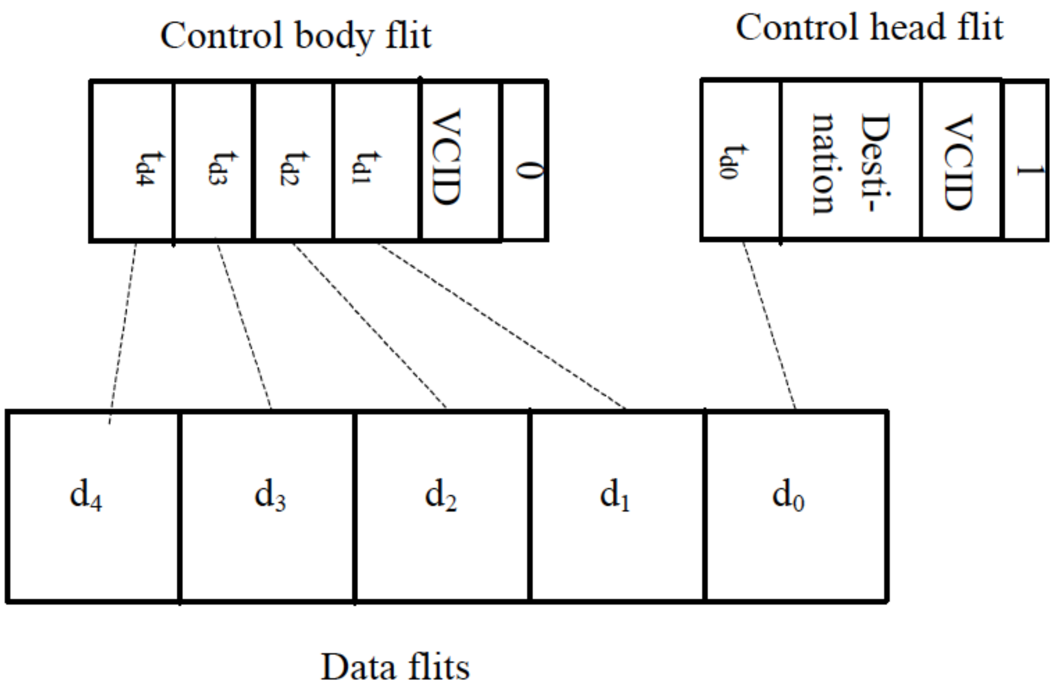
How many buffers needed in on/off flow-control? 6 + 2 = 8 (off propogation + processing)

But this is inefficient



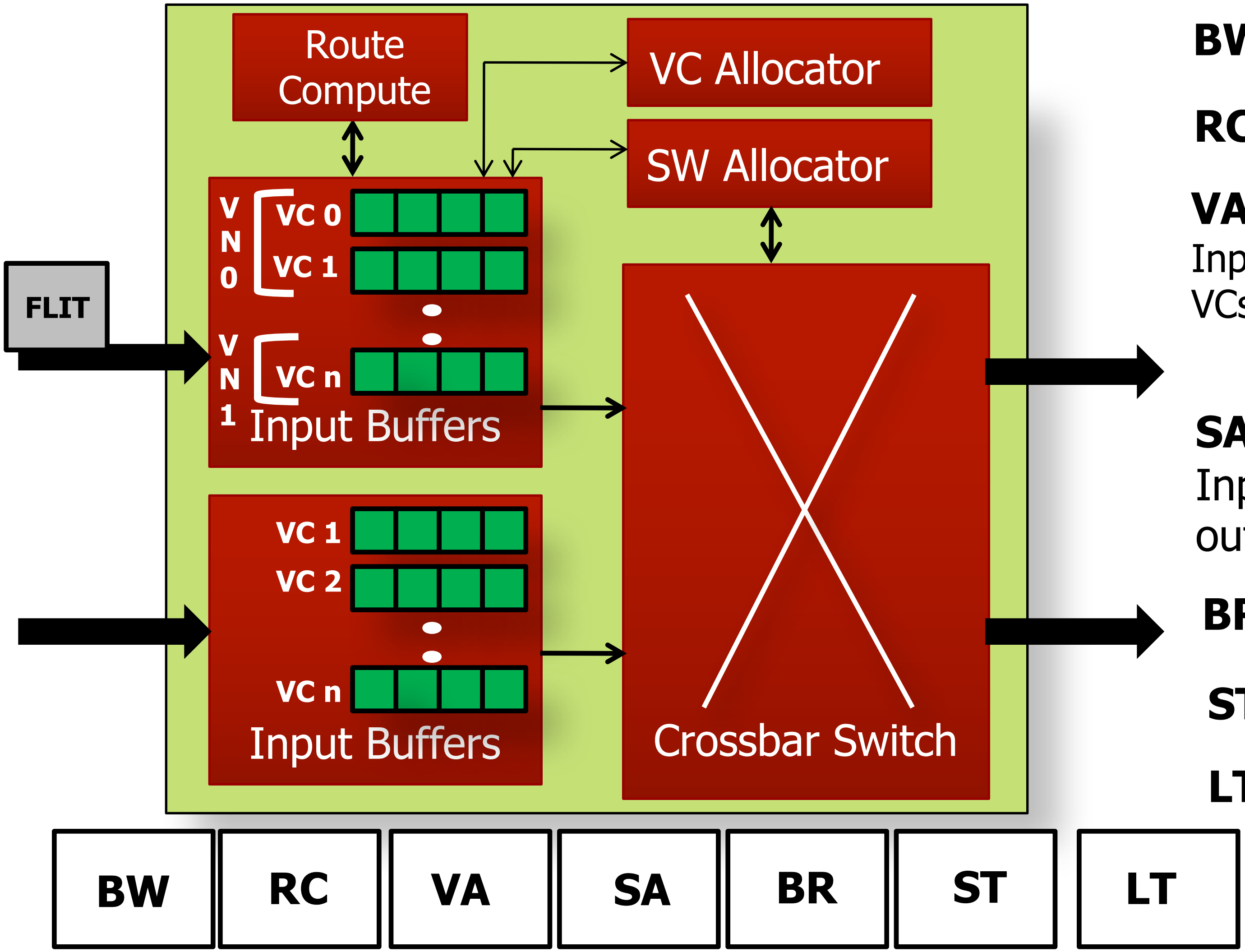
Flit Reservation Flow Control (Peh et al., HPCA 2000)

- What is the key idea (and benefit)?



- Why can't we just do static scheduling?

Conventional Virtual Channel Router



BW: Buffer Write

RC: Route Compute

VA: VC Allocation
Input VCs arbitrate for "output" VCs (Input VCs at next router)

SA: Switch Allocation
Input ports arbitrate for output ports

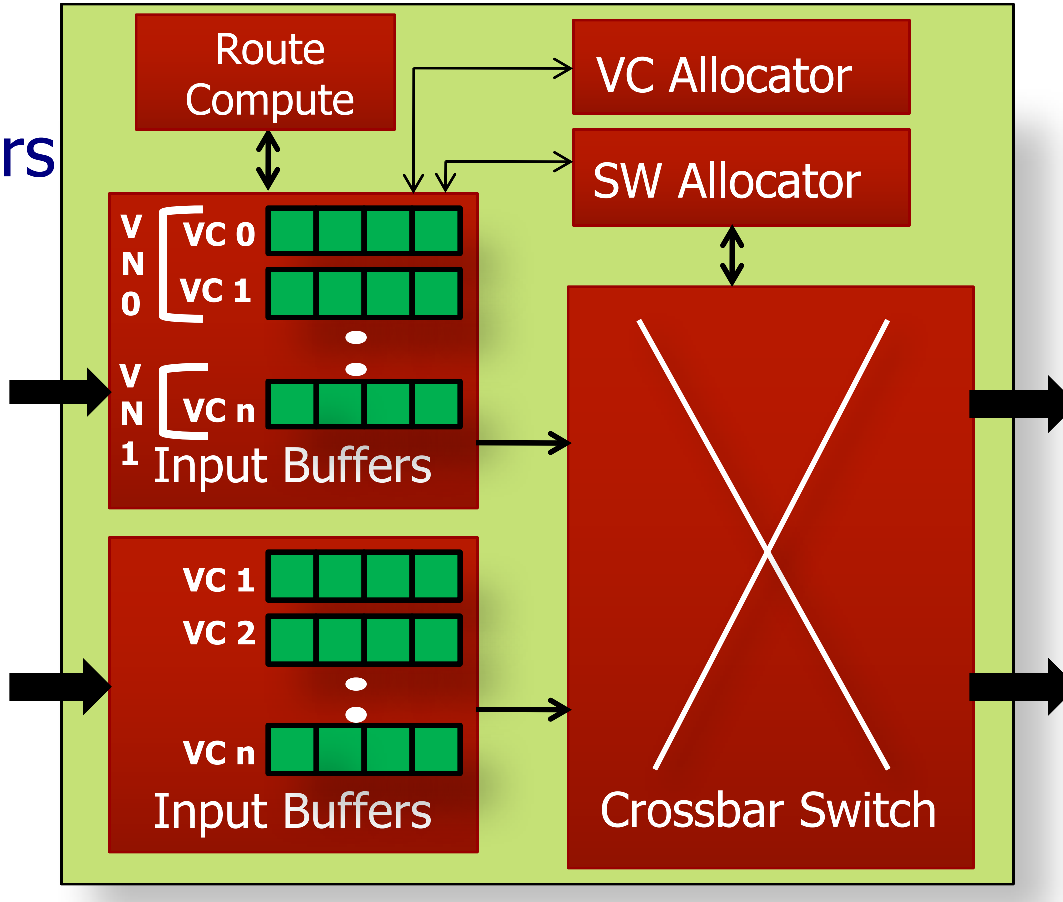
BR: Buffer Read

ST: Switch Traversal

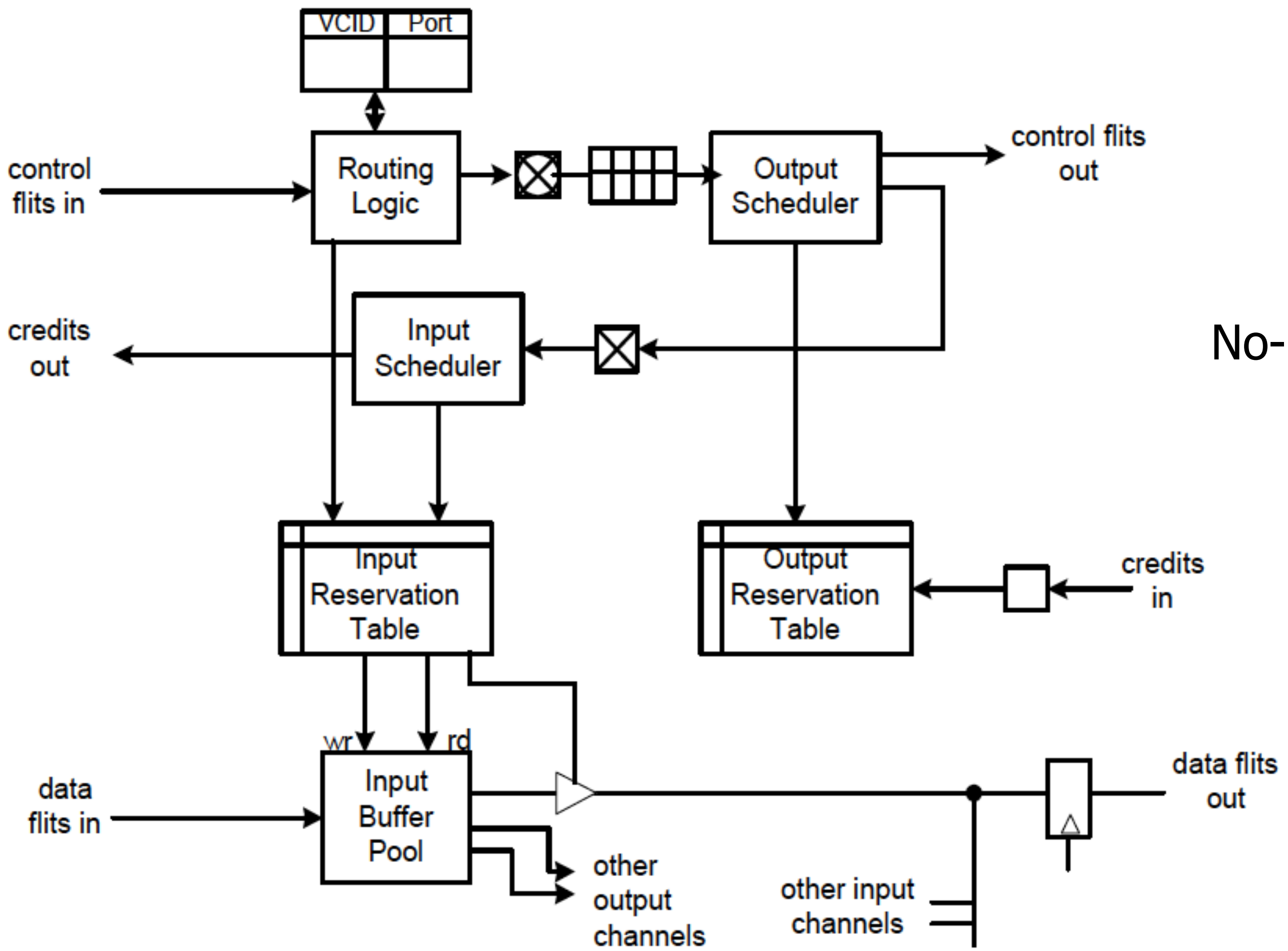
LT: Link Traversal

Router Microarchitecture

- Components
  - Virtual Channel Buffers
  - Routing Logic
  - Allocation
    - Switch Allocation
    - VC Allocation
  - Crossbar Switch
  - Link



Flit Reservation Router



No-load delay?

1-cycle

Key Unit: I/O Reservation Tables

Output Reservation Table

output channel	time	8	9	10	11	12	13	14	15	16	17
East Channel	Channel busy			X							
	Free buffers on next node	2	1	1	0	1	2	3	4	4	4

(a)

Input Reservation Table

input channel	time	8	9	10	11	12	13	14	15	16	17
West Channel	Flit Arriving?			X							
	Buffer in			5							
	Departure Time			+2							
	Buffer out					5					
	Output Channel					E					

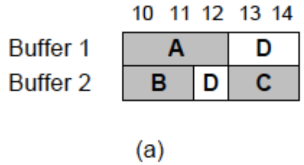
	0	1	2	3	4	5	6	7
Buffer occupancy bits						X		

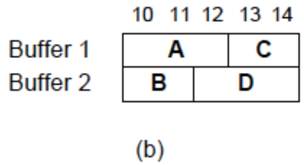
(c)

Credit Flow?

Design Details

- How is the crossbar switch driven?
- When is buffer ID allocated? Why?
- What if data arrives before output allocated?
  - Why could this happen?





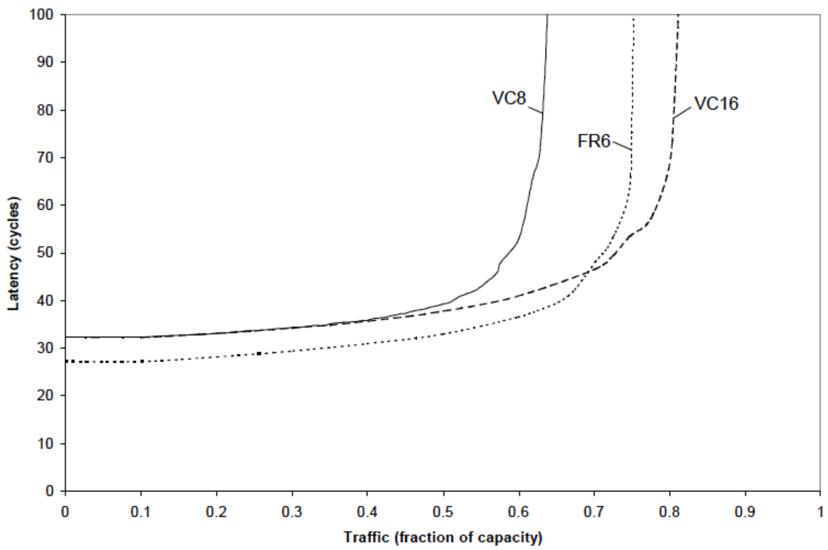
- What if 2 control flits for the same output port?
  - Not discussed in paper. They probably allocate serially, but this would require buffering
- Buffer organization – shared pool vs. per VC
  - Head-of-line blocking?
- Scheduling horizon?

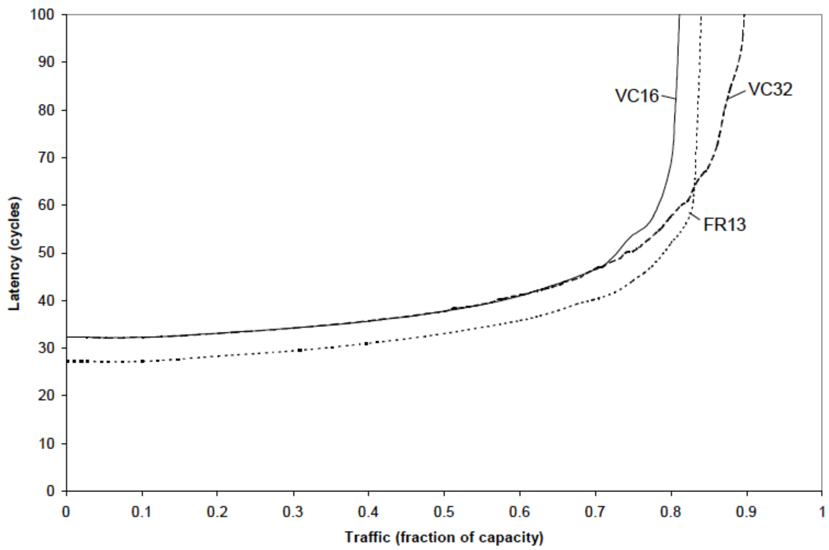
Overheads

	Virtual-Channel Flow Control				Flit-Reservation Flow Control		
	General	VC8	VC16	VC32	General	FR6	FR13
		b_d=8 v_d=2	b_d=16 v_d=4	b_d=32 v_d=8		b_d=6 v_c=2 b_c=6	b_d=13 v_c=4 b_c=12
Data buffers	$(f + \log_2 v_d + t) \times b_d \times 5$	10360	20800	41760	$f \times b_d \times 5$	7680	16640
Control buffers	-	-	-	-	$(\log_2 v_c + t + (d \times \log_2 s)) \times b_c \times 5$	240	540
Queue pointers	$2 \times \log_2 b_d \times v_d \times 5$	60	160	400	$2 \times \log_2 b_c \times v_c \times 5$	60	160
Output reservation table (Status bits and buffer counts)	$(1+ \log_2 b_d) \times 4 \times v_d$	32	80	192	$(1 + \log_2 b_d) \times s \times 4$	512	640
Input reservation table	-	-	-	-	$[(1 + \log_2 s + 2 + 2 \times \log_2 b_d) \times s + b_c] \times 5$	2270	1980
Bits per node		10452	21040	42352		10762	19960
Flits per input channel		8.17	16.44	33.09		8.40	15.59

Evaluations

What is the key benefit of the scheme?





FR6 has same throughput as FR16

Evaluation Methodology

- What if only single-flit packets
- How representative is random traffic for real life applications
  - Does that take away from the merits of the paper?