ECE 1749H:
Interconnection Networks for Parallel Computer Architectures:

Flow Control

Prof. Natalie Enright Jerger

Announcements

- Project Progress Reports
  - Due March 9, submit by e-mail
  - Worth 15% of project grade
  - 1 page
    - Discuss current status of project
    - Any difficulties/problems encountered
    - Anticipated changes from original project proposal

Announcements (2)

- 2 presentations next week
  - Elastic-Buffer Flow Control for On-Chip Networks
    - Presenter: Islam
  - Express Virtual Channels: Toward the ideal interconnection fabric
    - Presenter Yu
- 1 Critique due

Switching/Flow Control Overview

- Topology: determines **connectivity** of network

- Routing: determines **paths** through network

- Flow Control: determine **allocation** of resources to messages as they traverse network
  - Buffers and links
  - Significant impact on throughput and latency of network

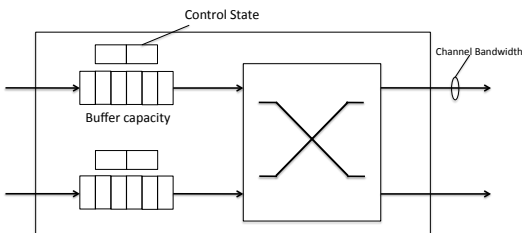Winter 2011   ECE 1749H: Interconnection Networks (Enright Jerger)   4

Flow Control

- Control state records
  - allocation of channels and buffers to packets
  - current state of packet traversing node
- Channel bandwidth advances flits from this node to next
- Buffers hold flits waiting for channel bandwidth

Winter 2011   ECE 1749H: Interconnection Networks (Enright Jerger)   5

Packets

- Messages: composed of one or more packets
  - If message size is <= maximum packet size only one packet created

- Packets: composed of one or more flits

- Flit: flow control digit

- Phit: physical digit
  - Subdivides flit into chunks = to link width

Winter 2011   ECE 1749H: Interconnection Networks (Enright Jerger)   6

Packets (2)

Message

Header Payload

Packet | Route | Seq# |

Head Flit Body Flit Tail Flit

Flit | Type | VCID |

Head, Body, Tail, Phit
Head & Tail

- Off-chip: channel width limited by **pins**
  - Requires phits
- On-chip: **abundant** wiring means phit size == flit size

Packets(3)

Cache line | RC | Type | VCID | Addr | Bytes 0-15 | Bytes 16-31 | Bytes 32-47 | Bytes 48-63 |

Head Flit Body Flits Tail Flit

Coherence
Command | RC | Type | VCID | Addr | Cmd |

Head & Tail Flit

- Packet contains destination/route information
  - Flits may not → all flits of a packet must take same route

Switching

- Different flow control techniques based on granularity

- Circuit-switching: operates at the granularity of messages

- Packet-based: allocation made to whole packets

- Flit-based: allocation made on a flit-by-flit basis

Message-Based Flow Control

- Coarsest granularity

- Circuit-switching
  - **Pre-allocates** resources across **multiple** hops
    - Source to destination
    - Resources = links
    - Buffers are not necessary
  - Probe sent into network to reserve resources

Circuit Switching

- Once probe sets up circuit
  - Message does not need to perform any routing or allocation at each network hop
  - Good for transferring large amounts of data
    - Can amortize circuit setup cost by sending data with very low per-hop overheads

- No other message can use those resources until transfer is complete
  - Throughput can suffer due setup and hold time for circuits
  - Links are idle until setup is complete

Circuit Switching Example

- Configuration Probe
- Data
- Circuit
- Acknowledgement

- Significant latency overhead prior to data transfer
  - Data transfer does not pay per-hop overhead for routing and allocation

Circuit Switching Example (2)

Configuration Probe

Data

Circuit

Acknowledgement

- When there is contention
  - Significant wait time
  - Message from 1 → 2 must wait

Time-Space Diagram: Circuit-Switching

Time to setup+ack circuit from 0 to 8

Time setup from 2 to 8 is blocked

Packet-based Flow Control

- Break messages into packets

- **Interleave** packets on links
  - Better utilization

- Requires per-node **buffering** to store in-flight packets

- Two types of packet-based techniques

Store and Forward

- Links and buffers are allocated to **entire** packet

- Head flit **waits** at router until entire packet is received before being forwarded to the next hop

- Not suitable for on-chip
  - Requires buffering at each router to hold entire packet
    - Packet cannot traverse link until buffering allocated to entire packet
  - Incurs high latencies (pays **serialization** latency at each hop)

Store and Forward Example

Total delay = 4 cycles per hop x 3 hops = 12 cycles
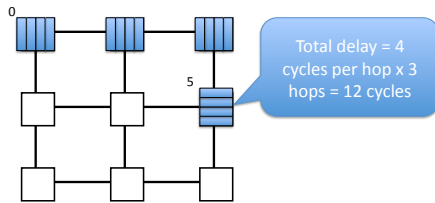
- High per-hop latency
  - Serialization delay paid at each hop
- Larger buffering required

Time-Space Diagram: Store and Forward
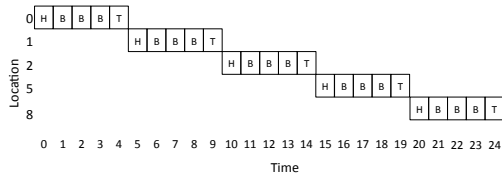
Packet-based: Virtual Cut Through

- Links and Buffers allocated to **entire** packets

- Flits can proceed to next hop before tail flit has been received by current router
  – But only if next router has enough buffer space for entire packet

- Reduces the latency significantly compared to SAF

- But still requires **large** buffers
  – Unsuitable for on-chip

Winter 2011   ECE 1749H: Interconnection Networks (Enright Jerger)   19

Virtual Cu[...]

Allocate 4 buffers bef[...] proce[...]

Allocate 4 flit-sized buffers before head proceeds

0

5

Total delay = 1 cycle per hop x 3 hops + serialization = 6 cycles
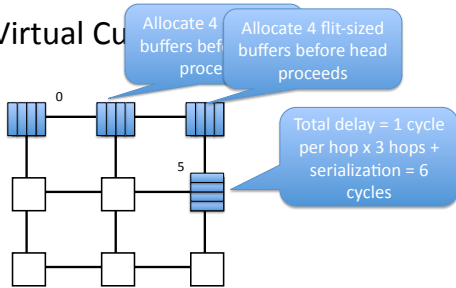
- Lower per-hop latency
- Large buffering required

Winter 2011   ECE 1749H: Interconnection Networks (Enright Jerger)   20

Time-Space Diagram: VCT

Location									
0	H	B	B	B	T				
1		H	B	B	B	T			
2			H	B	B	B	T		
5				H	B	B	B	T	
8					H	B	B	B	T

Time: 0 1 2 3 4 5 6 7 8
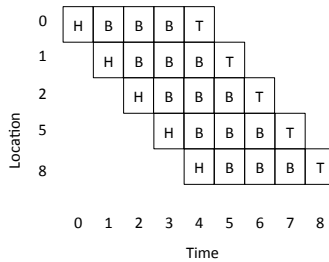
Winter 2011   ECE 1749H: Interconnection Networks (Enright Jerger)   21

Virtual Cut Through

Cannot proceed because only 2 flit buffers available

- Throughput suffers from inefficient buffer allocation

Time-Space Diagram: VCT (2)

Location												
0	H	B	B	B	T							
1		H	B	B	B	T						
2			H	B	B	B	T					
5			Insufficient Buffers			H	B	B	B	T		
8							H	B	B	B	T	

Time: 0 1 2 3 4 5 6 7 8 9 10 11

Flit Level Flow Control

- Help routers meet tight area/power constraints

- Flit can proceed to next router when there is buffer space available for that **flit**
  - Improved over SAF and VCT by allocating buffers on a flit-basis

Wormhole Flow Control

- Pros
  - More efficient buffer utilization (good for on-chip)
  - Low latency

- Cons
  - Poor link utilization: if head flit becomes blocked, all links spanning length of packet are idle
    - Cannot be re-allocated to different packet
    - Suffers from **head of line** (HOL) blocking

Wormhole Example

- 6 flit buffers/input port

Time-Space Diagram: Wormhole

Virtual Channels

- First proposed for deadlock avoidance
  – We'll come back to this

- Can be applied to any flow control
  – First proposed with wormhole

Winter 2011   ECE 1749H: Interconnection Networks (Enright Jerger)   28

Virtual Channel Flow Control

- Virtual channels used to combat **HOL** blocking in wormhole

- Virtual channels: **multiple** flit queues per input port
  – Share **same** physical link (channel)

- Link utilization improved
  – Flits on different VC can pass blocked packet

Winter 2011   ECE 1749H: Interconnection Networks (Enright Jerger)   29

Virtual Channel Flow Control (2)

In1	AH	A1	A2	A3	A4	A5		AT
	1	1	2	2	3	3	3	3	3	2	2	1	1

| In2 | BH | B1 | B2 | B3 | B4 | | B5 | | BT |
| | 1 | 2 | 2 | 3 | 3 | 3 | 3 | 3 | 3 | 2 | 2 | 1 | 1 |

Out: AH BH A1 B1 A2 B2 A3 B3 A4 B4 A5 B5 AT BT

A Downstream: AH A1 A2 A3 A4 A5 AT

B Downstream: BH B1 B2 B3 B4 B5 BT

Winter 2011   ECE 1749H: Interconnection Networks (Enright Jerger)   30
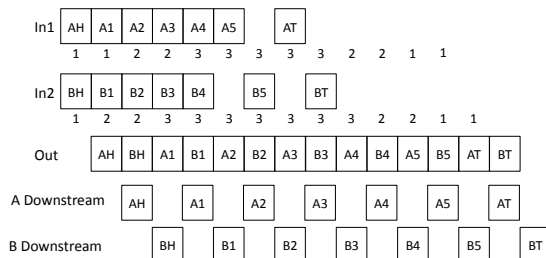
Virtual Channel Flow Control (3)

- Packets compete for VC on flit by flit basis

- Example: on downstream links, flits of each packet are available every other cycle

- Upstream links throttle because of limited buffers

- Does not mean links are idle
  - May be used by packet allocated to other VCs

Winter 2011　　　　ECE 1749H: Interconnection Networks (Enright Jerger)　　　　31

Virtual Channel Example

- 6 flit buffers/input port
- 3 flit buffers/VC

Winter 2011　　　　ECE 1749H: Interconnection Networks (Enright Jerger)　　　　32

Summary of techniques

	Links	Buffers	Comments
Circuit-Switching	Messages	N/A (buffer-less)	Setup & Ack
Store and Forward	Packet	Packet	Head flit waits for tail
Virtual Cut Through	Packet	Packet	Head can proceed
Wormhole	Packet	Flit	HOL
Virtual Channel	Flit	Flit	Interleave flits of different packets

Winter 2011　　　　ECE 1749H: Interconnection Networks (Enright Jerger)　　　　33

Deadlock

- Using flow control to guarantee deadlock freedom give more **flexible** routing
  - Recall: routing restrictions needed for deadlock freedom

- If routing algorithm is not deadlock free
  - VCs can break resource cycle

- Each VC is time-multiplexed onto physical link
  - Holding VC implies holding associated buffer queue
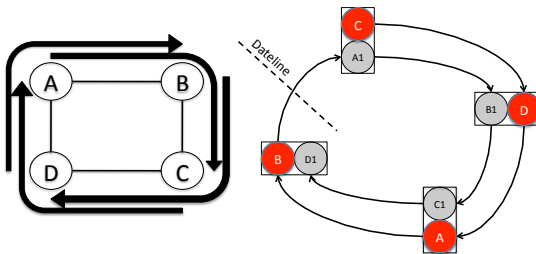 - Not tying up physical link resource

- Enforce order on VCs

Deadlock: Enforce Order

- All message sent through VC 0 until cross dateline
- After dateline, assigned to VC 1
  - Cannot be allocated to VC 0 again

Deadlock: Escape VCs

- Enforcing order lowers VC utilization
  - Previous example: VC 1 **underutilized**

- Escape Virtual Channels
  - Have 1 VC that is deadlock free
  - Example: VC 0 uses DOR, other VCs use arbitrary routing function
  - Access to VCs arbitrated fairly: packet always has chance of landing on escape VC

- Assign different message classes to different VCs to prevent protocol level deadlock
  - Prevent req-ack message cycles

Buffer Backpressure

- Need mechanism to prevent buffer **overflow**
  - Avoid dropping packets
  - Upstream nodes need to know buffer availability at downstream routers

- Significant impact on throughput achieved by flow control

- Two common mechanisms
  - Credits
  - On-off

Winter 2011   ECE 1749H: Interconnection Networks (Enright Jerger)   37

Credit-Based Flow Control

- Upstream router stores credit counts for each downstream VC

- Upstream router
  - When flit forwarded
    - Decrement credit count
  - Count == 0, buffer full, stop sending

- Downstream router
  - When flit forwarded and buffer freed
    - Send credit to upstream router
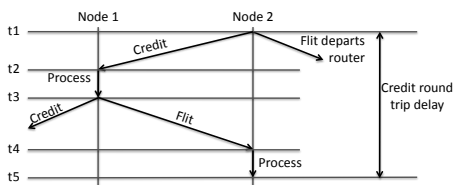   - Upstream increments credit count

Winter 2011   ECE 1749H: Interconnection Networks (Enright Jerger)   38

Credit Timeline

- Round-trip credit delay:
  - Time between when buffer empties and when next flit can be processed from that buffer entry
  - If only single entry buffer, would result in significant throughput degradation
  - Important to size buffers to tolerate credit turn-around

Winter 2011   ECE 1749H: Interconnection Networks (Enright Jerger)   39

13

On-Off Flow Control

- Credit: requires upstream signaling for every flit

- On-off: decreases upstream signaling

- Off signal
  - Sent when number of free buffers falls below threshold F_{off}

- On signal
  - Sent when number of free buffers rises above threshold F_{on}

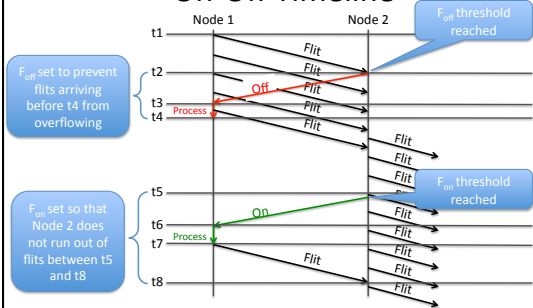Winter 2011   ECE 1749H: Interconnection Networks (Enright Jerger)   40

On-Off Timeline

- Less signaling but more buffering
- On-chip buffers more expensive than wires

Winter 2011   ECE 1749H: Interconnection Networks (Enright Jerger)   41

Buffer Utilization

Cycle	1	2	3	4	5	6	7	8	9	10	11	12
Credit count	2	1	0	0	0	0	0	1	2	1	0	0
Head Flit	VA/SA	ST	LT	BW	VA/SA	ST						
Body Flit 1		SA	ST	LT	BW	SA	ST					
Credit (head)						C	C-LT	C-Up				
Body Flit 2									SA	ST	LT	
Credit (body 1)						C	C-LT	C-Up				
Tail Flit									SA	ST	LT	

Winter 2011   ECE 1749H: Interconnection Networks (Enright Jerger)   42

14

Buffer Sizing

- Prevent backpressure from limiting throughput
  - Buffers must hold flits >= turnaround time

- Assume:
  - 1 cycle propagation delay for data and credits
  - 1 cycle credit processing delay
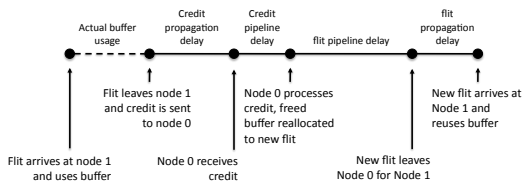 - 3 cycle router pipeline

- At least 6 flit buffers

Winter 2011   ECE 1749H: Interconnection Networks (Enright Jerger)   43

Actual Buffer Usage & Turnaround Delay

Winter 2011   ECE 1749H: Interconnection Networks (Enright Jerger)   44

Flow Control and MPSoCs

- Wormhole flow control

- Real time performance requirements
  - Quality of Service
  - Guaranteed bandwidth allocated to each node
    - Time division multiplexing

- Irregularity
  - Different buffer sizes

Winter 2011   ECE 1749H: Interconnection Networks (Enright Jerger)   46

Flow Control Summary

- On-chip networks require techniques with lower buffering requirements
  - Wormhole or Virtual Channel flow control

- Avoid dropping packets in on-chip environment
  - Requires buffer backpressure mechanism

- Complexity of flow control impacts router microarchitecture (next)

Winter 2011   ECE 1749H: Interconnection Networks (Enright Jerger)   47