

1. (Router architecture) On-chip routers use pipeline to enhance performance. In a pipelined virtual channel router, Consider a four-stage virtual channel router, namely, RC, VA, SA, ST,

- a) Why only head flits need to perform the RC and VA?
- b) The pipeline may be stalled due to different reasons. What kinds of stalls may occur?
- c) Draw a pipeline diagram (with a three-flit packet) to show the following stalls 1) VA stall for 4 cycles; 2) SA stall for 4 cycles; 3) Buffer empty stall for 4 cycles.

a) Only head flits perform RC (routing computation)  
then the result of RC is input to VAC virtual-channel allocator). If successful, the allocator assigns a single output virtual channel, then the body flits can follow the head flit.

b) packet stalls:  
can occur if the virtual channel cannot advance to its R (routing), V (waiting for output VC) or A (active) state.

- ① head flit is not in the first buffer slot to be serviced
- ② Routing is not completed.
- ③ VA is not successful

Flit stalls:  
if a virtual channel is in Active state and the flit cannot successfully complete switch allocation due to

- ① Lack of flit (buffer empty)

② lack of credit

③ losing arbitration for the switch time slot

c) D UA stalls for 4 cycles

RC

Cycle 1 2 3 4 5 6 7 8

Head flit (packet A) 

RC	stall	stall	stall	stall	VA	SA	ST
----	-------	-------	-------	-------	----	----	----

Body flit (packet A)

SA	ST
----	----

Tail flit (packet A)

SA	ST
----	----

② SA stall for 4 cycles

Cycle 1 2 3 4 5 6 7 8

Head flit (packet A) 

RC	VA	SA	ST
----	----	----	----

Body flit (packet A)

stall stall stall stall 

SA	ST
----	----

Tail flit (packet A)

SA	ST
----	----

③ buffer empty stall for 4 cycles

	Cycle	1	2	3	4	5	6	7	8
Head flit (packet A)		RC	VA	SA	ST				
Body flit (packet A)		stall	stall	stall	stall	SA	ST		
Tail flit (packet A)						SA	ST		

2. (Router architecture) Answer the following questions:

- Explain the credit round-trip latency. What is channel duty factor? What is the minimum buffer (virtual channel) size in order to achieve 100% duty factor?
- Assume a single cycle router and zero delay on channel wires. A single-cycle router means that a flit travels the router from an input link to an output link in one cycle, irrespective of if it is a head, body or tail flit. What is the minimum buffer size in order for the router to reach 100% duty factor?

a) credit round-trip latency <sup>(loop)</sup>: expressed in flit times, gives a lower bound on the number of flit buffers needed for the channel to operate with full bandwidth.

$$t_{\text{crt}} = t_f + t_c + \geq T_w + | \quad \begin{matrix} \uparrow & \uparrow & \uparrow \\ \text{flit pipeline delay} & \text{credit pipeline delay} & \text{one-way wire delay} \end{matrix} \quad \begin{matrix} \leftarrow \\ \text{switch allocation} \end{matrix}$$

number of buffers available per virtual channel

$$\text{channel duty factor: } d = \min(1, \frac{F}{t_{\text{crt}}})$$

whether there are sufficient flit buffers to cover the round trip latency

in order to achieve 100%:  $F = t_{\text{crt}}$

assume: 3-flit in a packet '(head, body, tail)

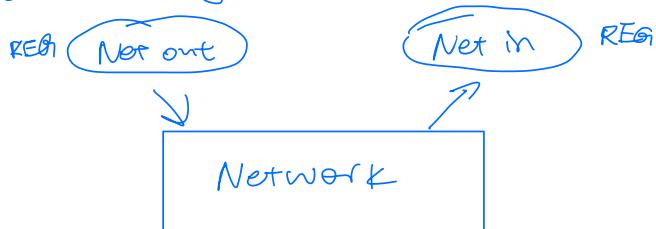
$$\begin{aligned} b) t_{\text{act}} &= t_f + t_e + 2T_w + 1 \\ &= t_f + t_e + 1 \Rightarrow t_c + 2 \\ &= 3 + 1 = 4 \end{aligned}$$

$$F = 4$$

3. (Network Interface) Answer the following questions:

- Explain how the *two-register* and *descriptor-based* network interface for a message-passing machine works.
- Discuss the pros and cons of the two-register network interface.
- Draw a block diagram of MSHR and briefly explain how it works.
- Draw a block diagram of TSHR and briefly explain how it works.

### ① two-register interface



for sending: the processor writes to a specific net-out register

for receiving: the processor reads in a specific net-in register

### ② Descriptor-based network interface

the processor composes a message in a set of dedicated message descriptor register.

Cache descriptor contains:  
an immediate value / a reference to a  
processor register / a reference to a block of memory  
a co-processor steps through the descriptors and composes  
the messages.

b) pros of 2-register interface:  
efficient for short messages

Cons:

Inefficient for long messages

Processor acts as DMA controllers

Not safe, because it does not prevent the  
network from misbehaving SW running on the processor

c) MSHR (miss status holding register)

load/store requests are stored in

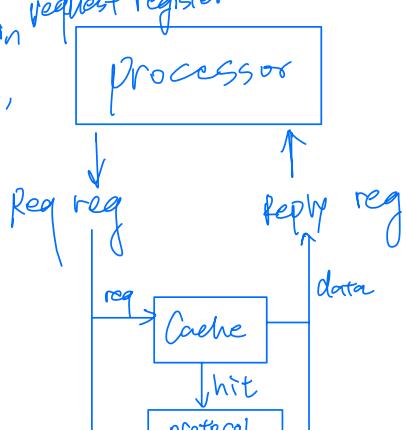
In case of a cache miss,  
requests are stored in

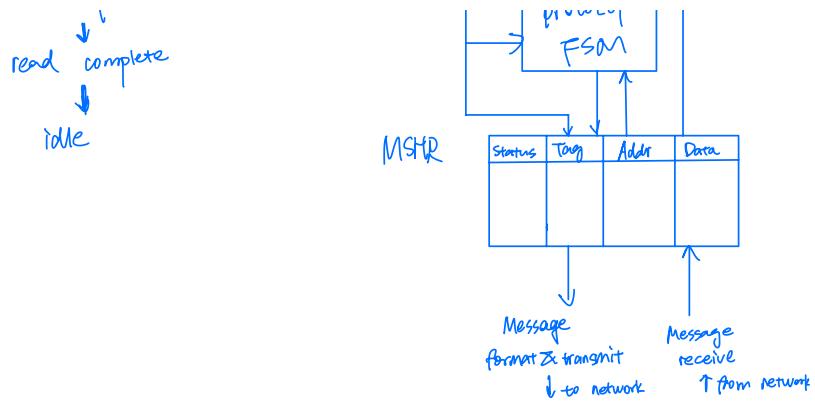
MSHR

read operation:

pending read

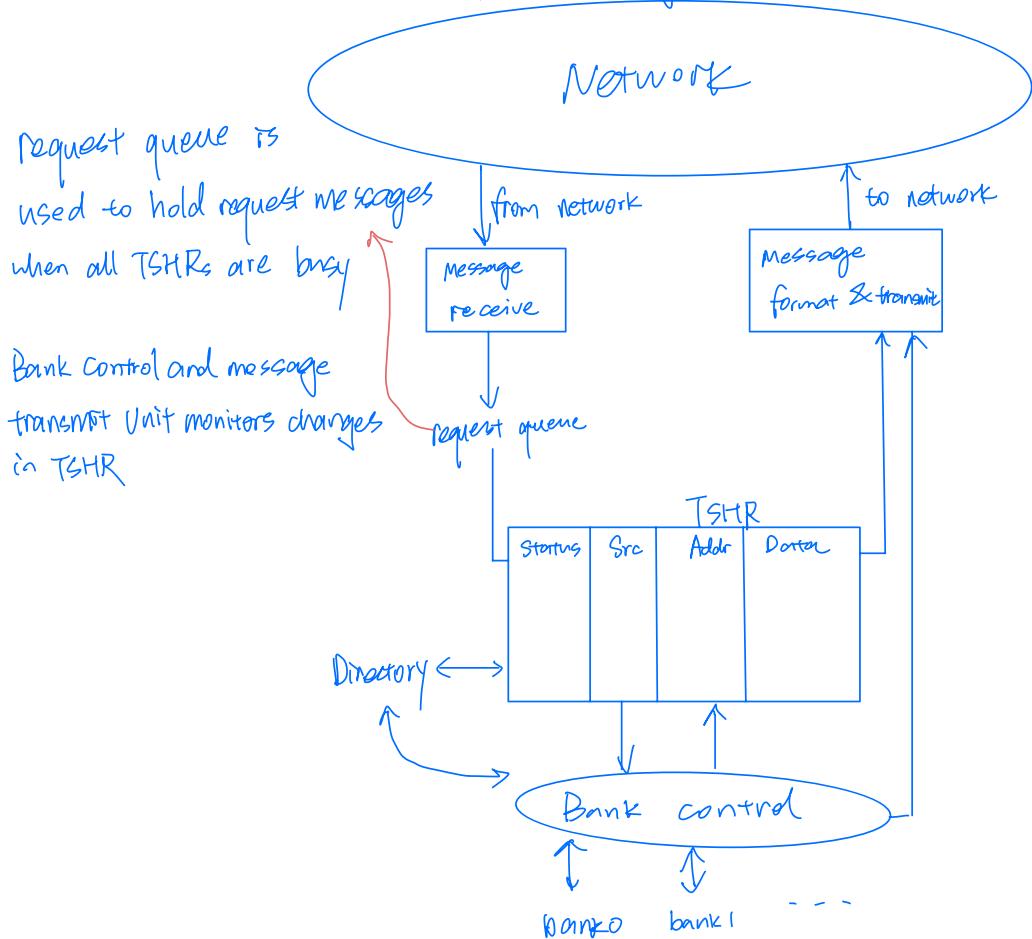
read requested





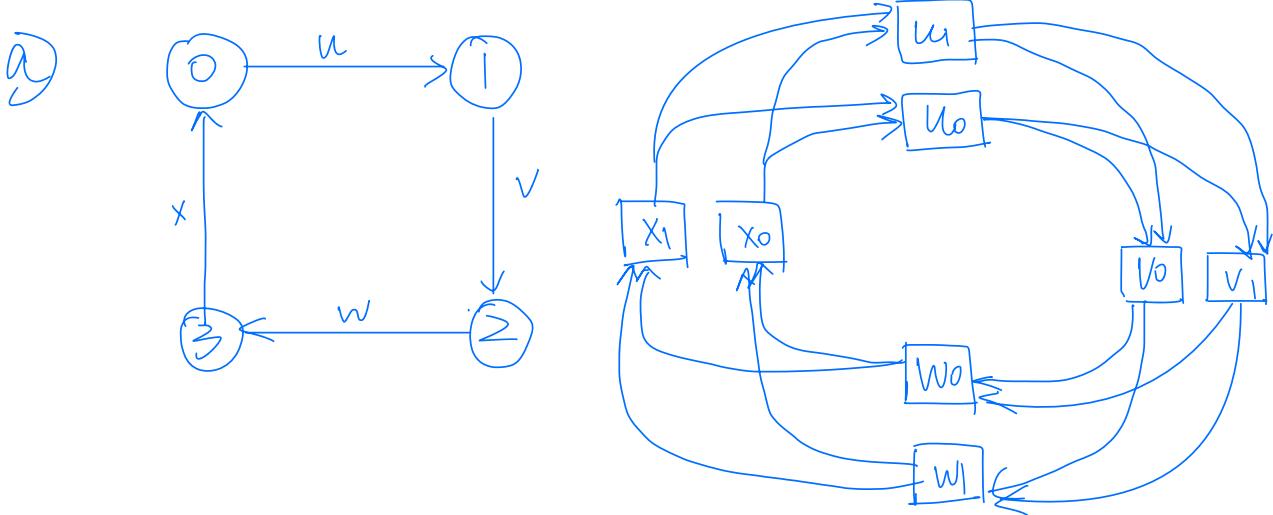
### TSHR (transaction status holding register)

interfaces receive memory request messages and sends replies

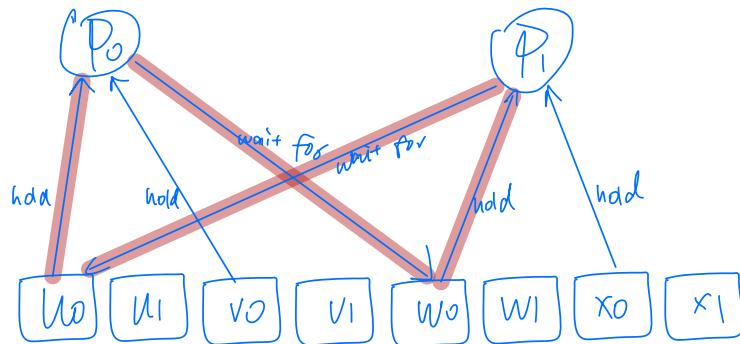


4. (Deadlock configuration, Dally & Towles Exercise 14.1) Consider a 4-node ring network using flit-buffer flow control with 2 virtual channels per node, with the resource request criteria that a packet may use either virtual channel at each hop, whichever becomes free first.

- Draw a flit-buffer dependency graph for the network.
- Can this resource request criteria result in deadlock? If so, draw a wait-for graph for a deadlocked configuration of the network.



b) Yes



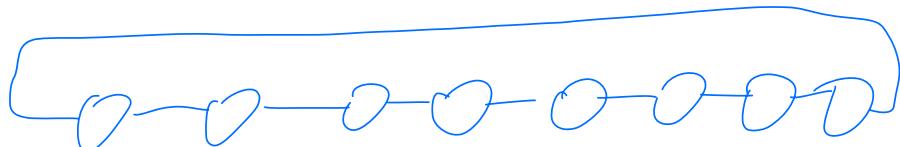
5. (Deadlock avoidance) Deadlock can be avoided using buffer classes. Answer the following questions:

- Consider a unidirectional ring with 8 nodes, what is the minimum number of buffer classes in order to be deadlock free?
- Consider a bidirectional ring with 8 nodes, what is the minimum number of buffer classes in order to be deadlock free, using distance classes?

- c) Consider a 4-by-4 mesh network with adaptive routing, what is the minimum number of buffer classes in order to be deadlock free, using distance classes?
- d) Consider a 8-by-8 mesh network with XY routing, what is the minimum number of buffer classes in order to be deadlock free?

*if use distance classes:*

a)  $7+1$



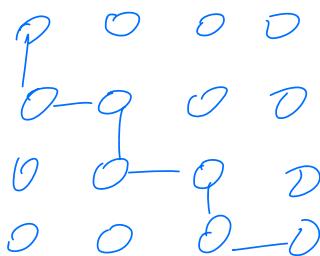
*but! if use dateline class: 2 is enough*

b) using distance classes:  $5$   
5

c) using distance classes:

$$b+1 = 7$$

7

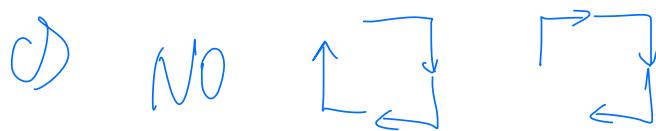
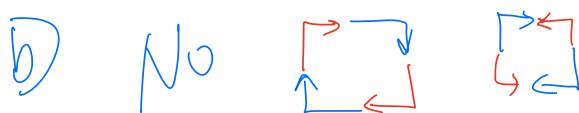
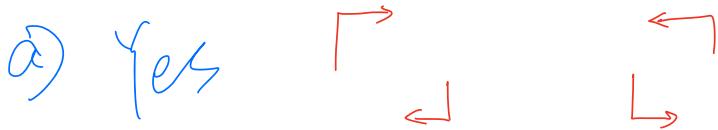


d) XY routing, deadlock free



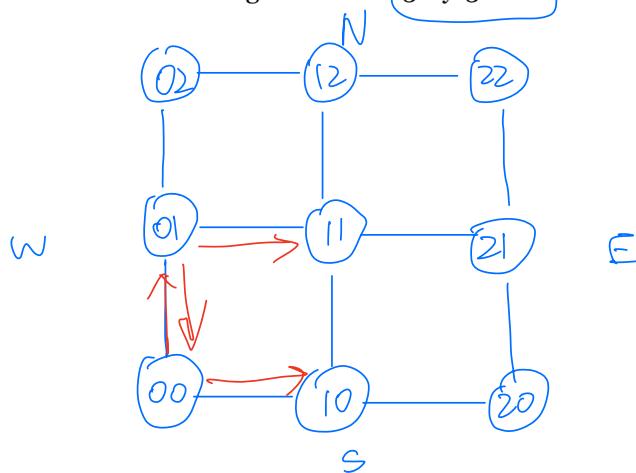
6. (Deadlock and turn model) Determine whether the following oblivious routing algorithms are deadlock-free for the 2D mesh.

- Deterministic Y-X routing, first route along Y then X dimension.
- Randomized dimension-order: All packets are routed minimally. Half of the packets are routed completely in the X dimension before the Y dimension and the other packets are routed Y before X.
- Limited-turns routing for the 2D mesh: All routes are restricted to contain at most three right turns and no left turns.



7. (Routing and Deadlock-livelock construction) (**Optional**) Routing table is a general approach to implement *various* routing algorithms. When construct routing tables, care must be taken to avoid deadlock and livelock. Construct routing tables for a 3-by-3 mesh network with oblivious routing by which

- Deadlock may occur
- Livelock may occur



*A)*

b)

From

00 → 11

to      00      10      20      01      11      21      02      12      22

00      XX      WN      WN      SE

10      EN      XX      WN      SE      -      -      -

20      EN      EN      XX      SE

01      NE      WN      WN      XX

11      NE      NW      WN      ES      XX

21      NE      NE      NE      ES      XX

02      NE      WN      NW      NE      XX

12      NE      NW      NW      NE

22      NE      NE      NE      NE      -      -      XX

XX

8. (Routing and Deadlock, Livelock) (**Optional**) A student argues that fully adaptive routing is not a good option because it may result in both deadlock and livelock. Do you agree or not? Why? Give an example.

No

adaptive routing network may be deadlock free even in  
the presence of dependence cycles

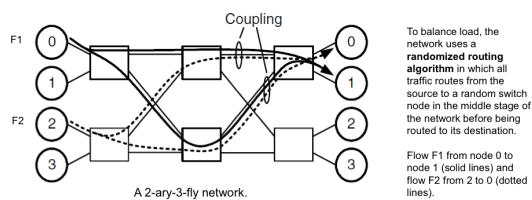
9. (QoS and QoS mechanisms) (Optional) Answer the following questions:

- What is QoS? Is there a QoS requirement for best-effort networks?
- Explain Aggregate resource allocation, and exemplify guaranteed resource reservation schemes?
- What is non-interfering network? What is its purpose? Discuss also its advantages and possible drawbacks.
- Explain and illustrate the tree saturation problem.

② QoS: Quality-of-Service  
guaranteed services  
best-effort services  
best-effort services: latency fairness, throughput fairness

b) aggregate resource allocation: Flow coupling when sharing  
the channels

- Flow coupling when sharing the channels



Two flows under aggregate resource allocation. Because the flows share channel resources, there is coupling between the two flows, which affects their delay and jitter.

guaranteed resource reservation schemes

Virtual circuit: reservation in space

TDM: (time-division multiplexing): reservation in  
space and time

② non-interfering network: isolating traffic classes

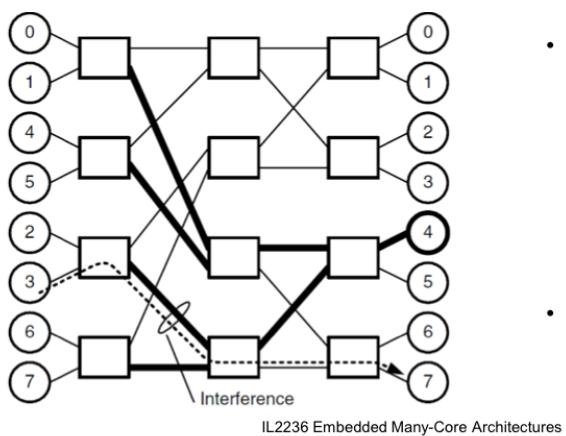
there cannot be any resource shared between A and B that can be held for an indefinite amount of time by A such that B cannot interrupt the usage of the resource

PURPOSE: interference isolation

advantage: no interference between classes

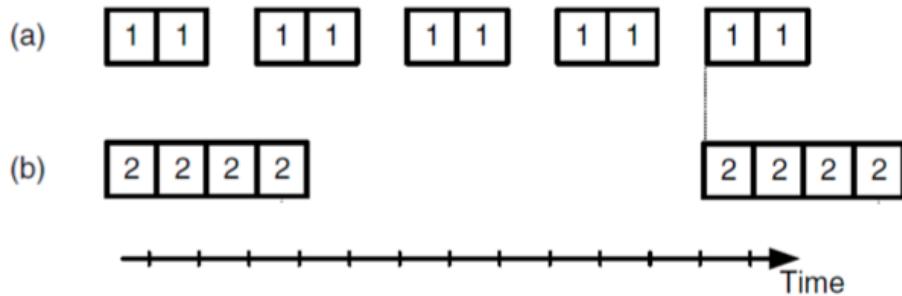
drawbacks: costly in hardware implementation

## d) tree saturation



- Tree-saturation in a 2-ary 3-fly. Destination 4 is overloaded, causing channels leading toward it to become blocked. These blocked channels in turn block more channels, forming a tree pattern (bold lines).
- Interference then occurs for a message destined to node 7 because of the tree-saturated channels.

10. (Traffic specification for QoS) Based on the definition of  $(\sigma, \rho)$  regulated flow, given the following source traffic injection behavior, calculate its  $(\sigma, \rho)$  values.



- a) 2 packets are injected consecutively every 3 cycles.
- b) 4 packets are injected consecutively every 12 cycles.
- c) 1 packet is injected every 3 cycles (not shown above).

a) rate =  $\frac{2}{3}$   $(\sigma, \rho) = \left(\frac{2}{3}, \frac{2}{3}\right)$

$$\text{burstiness} = 2 - 2 \times \frac{2}{3} = \frac{1}{3} - \frac{4}{3} = \frac{2}{3}$$

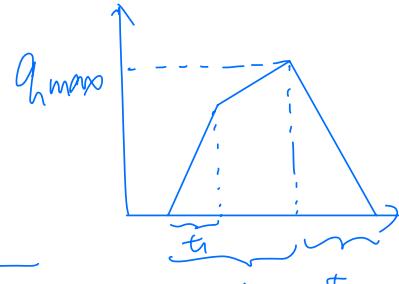
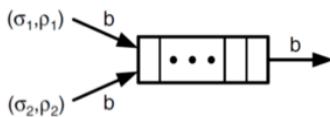
b) rate =  $\frac{4}{12} = \frac{1}{3}$   $(\sigma, \rho) = \left(\frac{8}{3}, \frac{1}{3}\right)$

$$\text{burstiness} = 4 - 4 \times \frac{1}{3} = \frac{12}{3} - \frac{4}{3} = \frac{8}{3}$$

c) rate =  $\frac{1}{3}$   $(\sigma, \rho) = \left(\frac{2}{3}, \frac{1}{3}\right)$

$$\text{burstiness} = 1 - 1 \times \frac{1}{3} = \frac{2}{3}$$

11. (FIFO multiplexing) Consider the two-flow multiplexing example. The multiplexer has a first-come, first-served (FIFO) service discipline. Assuming the two inputs to the multiplexer are a  $(\sigma_1, p_1)$  characterized flow and a  $(\sigma_2, p_2)$  characterized flow, respectively, what is the maximum packet delay of this multiplexer?



$$bt_1 = g_1 + p_1 t_1 \Rightarrow t_1 = \frac{g_1}{b - p_1}$$

$$t_2 = \frac{g_2}{b - p_2}$$

$$\text{max queue size: } q_{\max} = bt_1 + p_1(t_2 - t_1) = g_1 + \frac{p_1 g_2}{b - p_2}$$

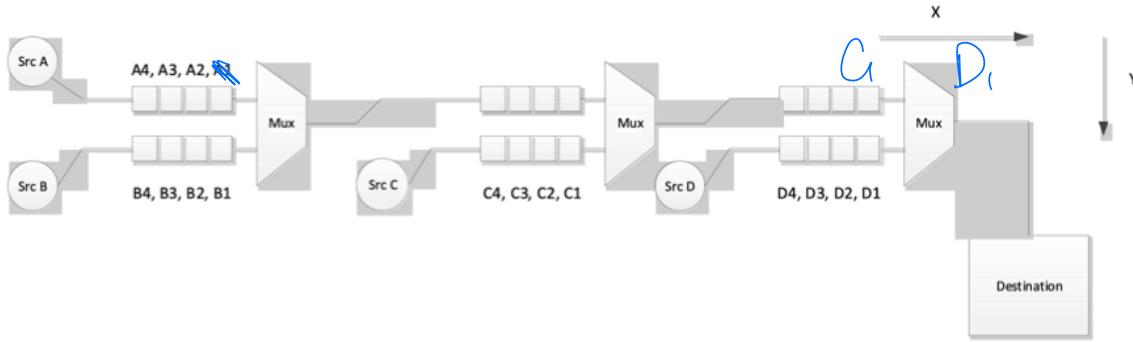
$$t_{\text{drain}} = \frac{q_{\max}}{b - p_1 - p_2}$$

$$\text{max packet delay } D_{\max} = t_2 + t_{\text{drain}}$$

$$= \frac{g_2}{b - p_2} + \frac{q_{\max}}{b - p_1 - p_2}$$

$$= \frac{g_2}{b - p_2} + \frac{\frac{g_1 + p_1 g_2}{b - p_2}}{b - p_1 - p_2} = \frac{g_1 + g_2}{b - p_1 - p_2}$$

12. (Latency fairness) Latency fairness) On a mesh network using X-Y routing. Assume each source continuously injects packets into its queue as long as the queue has space. All packets shall be routed along the X dimension first and then go to a destination on the Y dimension. See figure below. Suppose that a packet takes one cycle to leave the node through the multiplexer once it is allowed to use the shared output channel. Determine the output sequence to the Y dimension till the fourth packet A4 of packet source A, and calculate the time for the fourth packet A4 of packet source A to reach the Y dimension.



- a) Assume that the multiplexers use a locally fair round-robin policy.  
 b) Assume the multiplexers use an age-based arbitration policy.

D<sub>1</sub> D<sub>2</sub> D<sub>3</sub> D<sub>4</sub>, D<sub>5</sub> ---  
 ↗ doesn't stop here  
 ② Round Robin

D<sub>1</sub> C<sub>1</sub> D<sub>2</sub> B<sub>1</sub> D<sub>3</sub> G<sub>2</sub> D<sub>4</sub> A<sub>1</sub>

D<sub>5</sub> G<sub>3</sub> D<sub>6</sub> B<sub>2</sub> D<sub>7</sub> G<sub>4</sub> D<sub>8</sub> A<sub>2</sub> -

One packet from source A will be sent to the Y dimension every 8 cycles

$$4 \times 8 = 32 \text{ cycles}$$

b) age-based

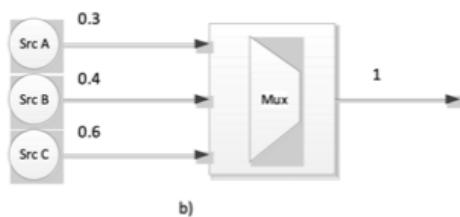
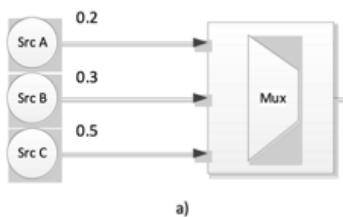
D<sub>1</sub> C<sub>1</sub> B<sub>1</sub> A<sub>1</sub> D<sub>2</sub> C<sub>2</sub> B<sub>2</sub> A<sub>2</sub>

D<sub>3</sub> C<sub>3</sub> B<sub>3</sub> A<sub>3</sub> D<sub>4</sub> C<sub>4</sub> B<sub>4</sub> A<sub>4</sub>

One packet from source A will be sent to the Y dimension every 4 cycles

$$4 \times 4 = 16 \text{ cycles}$$

13. (Throughput fairness) Calculate the allocated bandwidth for each flow for each scenario in a) and b) considering both the following arbitration/bandwidth allocation strategies:



- a) Max-min fairness bandwidth allocation
- b) Rate proportional bandwidth allocation

for flow a

$$\textcircled{2} \quad R_0 = b$$

$$a_0 = \min\left[b_0, \frac{R_0}{3-0}\right] = \min[0.2b, \frac{b}{3}]$$

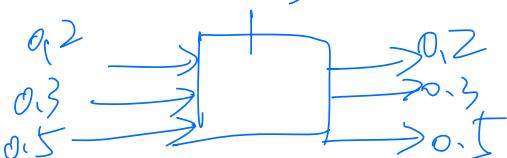
$$= 0.2b$$

$$R_1 = R_0 - a_0 = b - 0.2b = 0.8b$$

$$a_1 = \min\left[0.3, \frac{0.8b}{3-1}\right] = \min[0.3, 0.4b] = 0.3b$$

$$R_2 = R_1 - a_1 = 0.8b - 0.3b = 0.5b$$

$$a_2 = \min\left[0.5, \frac{0.5b}{3-2}\right] = 0.5b$$

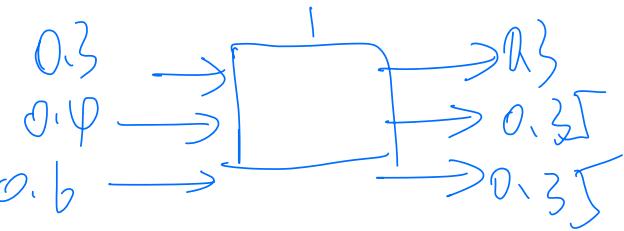


for flow b

$$R_0 = b \quad a_0 = \min[0.3b, \frac{b}{3}] = 0.3b$$

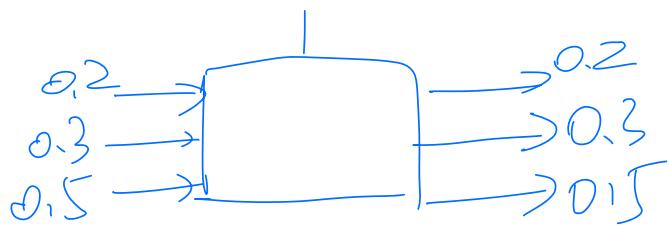
$$R_1 = b - 0.3b = 0.7b \quad a_1 = \min[0.4b, \frac{0.7b}{2}] = 0.35b$$

$$R_2 = 0.7b - 0.35b = 0.35b \quad a_2 = \min[0.6b, \frac{0.35b}{1}] = 0.35b$$



b) rate proportional



$$0.3 + 0.4 + 0.6 = 1.3$$

