332 Advanced Computer Architecture

Unassessed tutorial exercise 10

Exercise 10 Estimating the cost of communication operations

The latency models for store-and-forward routing and cut-through routing can be used to estimate and compare the cost of communication operations.

For this problem assume:

- A 8 element linear of array of processors numbered from 0 to 7.
- The packets sent are 100KB long.
- The channels in the network have a bandwidth of 250 MB/s bidirectional.
- The time taken for the routing information to cross a switch is 1 μs.
- The sender overhead is 55 μs and the receiver overhead is 60 μs.
- Sends are blocking, *ie* cannot start sending next packet until the current one has been received.
- The simple model for cut-through routing.

10.1 One-to-all broadcast

A one-to-all broadcast can be implemented by:

```
for i = 1 to 7
    processor 0 sends its partition of the array to processor i
end for
```

a. Draw the communication pattern taken by the sends in the algorithm.

b. What is the time taken to complete the operation under store-and-forward routing?

c. What is the time taken to complete the operation under worm-hole routing?

d. Compare the two algorithms.

10.2 Right shift

A right shift by 4 places can be implemented by:

```
for all i < 4
    processor i sends its partition of the array to processor i+4
```

a. What is the time taken to complete the operation under store-and-forward routing?

b. What is the time taken to complete the operation under worm-hole routing with virtual channels for flow control (you may assume an infinite number of virtual channels).

   (You may assume that the bandwidth of all the channels is that of the most heavily loaded channel.)

c. Compare the two algorithms.

Some things to think about

For the calculation of one-to-all broadcast it was assumed that the send was blocked, *ie* the send did not return control to the process until the packet had been received.

How might the sending processor detect that the packet has been received? What are the implications of this to the cost of a blocking send?

What is the time taken to complete the one-to-all broadcast for *non-blocking* sends under the two schemes?

For a number of aggregate communication patterns, such as right shift, the latency difference between using store-and-forward and worm-hole routing are not significant.

Why then have most current MPP designers chosen to use worm-hole routing?

Paul Kelly and Hing Wing To, Imperial College, December 1999

332 Advanced Computer Architecture

Unassessed tutorial exercise 10

Sample solution

Exercise 10 Estimating the cost of communication operations

10.1 One-to-all broadcast

A one-to-all broadcast can be implemented by:

```
for i = 1 to 7
    processor 0 sends its partition of the array to processor i
end for
```

a. Draw the communication pattern taken by the sends in the algorithm.

b. What is the time taken to complete the operation under store-and-forward routing?

c. What is the time taken to complete the operation under worm-hole routing?

d. Compare the two algorithms.

10.1 One-to-all broadcast – sample solution

Recall that the latency for store-and-forward is:

$$l_{sf} = t_s + h(t_{sw} + \frac{m}{bw}) + t_r$$

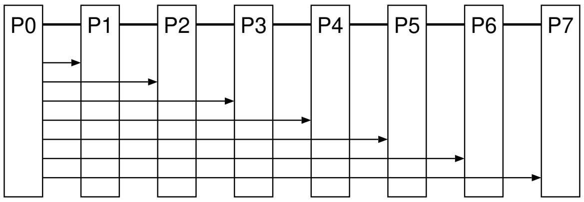
and that the simple model of latency for cut-through (worm-hole) routing is:

$$l_{ct} = t_s + h\, t_s w + \frac{m}{bw} + t_r$$

where

$$\begin{array}{lcl} t_s & = & 55 \times 10^{-6}\text{s} \\ t_r & = & 60 \times 10^{-6}\text{s} \\ t_{sw} & = & 1 \times 10^{-6}\text{s} \\ m & = & 100 \times 1024 \text{ bytes} \\ bw & = & 250 \times 1024^2 \text{ bytes/s} \end{array}$$

a. Communication pattern for one-to-all broadcast is:



b. The sends are blocking so the total latency is the sum of the individual sends:

$$T = t_{\text{P0-P1}} + t_{\text{P0-P2}} + t_{\text{P0-P3}} + t_{\text{P0-P4}} + t_{\text{P0-P5}} + t_{\text{P0-P6}} + t_{\text{P0-P7}}$$

Under store-and-forward routing:

$$t_{\text{P0-P}_i} = t_s + i(t_{sw} + \frac{m}{bw}) + t_r$$

therefore:

$$\begin{aligned} T_{sf} = \; & t_s + 1(t_{sw} + \tfrac{m}{bw}) + t_r + \\ & t_s + 2(t_{sw} + \tfrac{m}{bw}) + t_r + \\ & \vdots \\ & t_s + 7(t_{sw} + \tfrac{m}{bw}) + t_r \end{aligned}$$

simplifying:

$$T_{sf} = 7(t_s + t_r) + (t_{sw} + \frac{m}{bw}) \sum_{i=1}^{7} i$$

substituting in values for the constants:

$$\begin{aligned} T_{sf} &= 7(55 \times 10^{-6} + 60 \times 10^{-6}) + (1 \times 10^{-6}) + \tfrac{100 \times 1024}{250 \times 1024^2})28 \\ &= 8.05 \times 10^{-4} + (1 \times 10^{-6} + 3.90625 \times 10^{-4})28 \\ &= 0.01177\text{s} \end{aligned}$$

c. Under cut-through routing:

$$t_{\text{P0-P}_i} = t_s + i\ t_{sw} + \frac{m}{bw} + t_r$$

therefore:

$$\begin{aligned} T_{ct} = \; & t_s + 1t_{sw} + \tfrac{m}{bw} + t_r + \\ & t_s + 2t_{sw} + \tfrac{m}{bw} + t_r + \\ & \vdots \\ & t_s + 7t_{sw} + \tfrac{m}{bw} + t_r \end{aligned}$$

simplifying:

$$T_{ct} = 7(t_s + t_r + \frac{m}{bw}) + t_{sw} \sum_{i=1}^{7} i$$

substituting in values for the constants:

$$\begin{aligned} T_{ct} &= 7(55 \times 10^{-6} + 60 \times 10^{-6} + \tfrac{100 \times 1024}{250 \times 1024^2}) + (1 \times 10^{-6})28 \\ &= 7(3.90625 \times 10^{-4}) + 28 \times 10^{-6} \\ &= 0.002762\text{s} \end{aligned}$$

d. The cut-through routing version performs better than the store-and-forward routing version. Long packets are sent therefore the latency of the individual sends in the cut-through routing scheme are relatively independent of the distance the packets are sent. In contrast the latency of the distant sends is considerably higher in the store-and-forward version.

10.2 Right shift

A right shift by 4 places can be implemented by:

```
for all i < 4
    processor i sends its partition of the array to processor i+4
```

a. What is the time taken to complete the operation under store-and-forward routing?

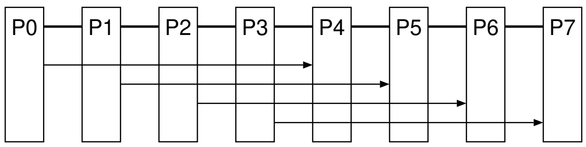
b. What is the time taken to complete the operation under worm-hole routing with virtual channels for flow control (you may assume an infinite number of virtual channels).

   (You may assume that the bandwidth of all the channels is that of the most heavily loaded channel.)

c. Compare the two algorithms.

10.2 Right shift – sample solution

The communication pattern for right shift by four places is:



a. All four sends occur simultaneously, therefore the total latency is that of the longest send. In this case all the sends have the same cost.

   Under store-and-forward routing there is no contention assuming all the sends begin at the same time, since the packets are being stored one switch ahead of each other.

   Therefore the time for one send is:

$$T_{sf} = t_s + 4(t_{sw} + \frac{m}{bw}) + t_r$$

   substituting in the values for the constants:

$$\begin{aligned} T_{sf} &= 55 \times 10^{-6} + 4(1 \times 10^{-6} + \tfrac{100+1024}{250 \times 1024^2}) + 60 \times 10^{-6} \\ &= 0.0016815\text{s} \end{aligned}$$

b. Assume worm-hole routing with at least four virtual channels. The bandwidth of the channel between P3 and P4 will have four packets traversing it by time-division multiplexing. Therefore the bandwidth of the channel is only a fourth of the raw channel bandwidth.

Assume that this is a bottleneck and that the bandwidth of all channels is therefore the same as the most heavily loaded channel.

Therefore the time for one send is:

$$T_{ct} = t_s + 4t_{sw} + \frac{m}{bw/4} + t_r$$

simplifying:

$$\begin{aligned} T_{ct} &= t_s + 4t_{sw} + \tfrac{4m}{bw} + t_r \\ &= t_s + 4(t_{sw} + \tfrac{m}{bw}) + t_r \\ &= T_{sf} \end{aligned}$$

c. The performance under both routing schemes is the same. The advantage of the cut-through routing scheme when sending messages over long distances is offset by the loss in available bandwidth caused by sharing the channel between multiple messages.

Paul Kelly and Hing Wing To, Imperial College, December 1999