

Trevor Bekolay, 6796723, umbekol0

Assignment 4, Question 1

Virtual cut-through (VCT) switching is an interconnection network switching method that is similar to wormhole switching. Like wormhole switching, in VCT a message is broken up into smaller units called flits. The first flit contains routing information, and other flits follow the flit that came before it. Unlike wormhole switching, however, when a flit cannot be sent to the next node, the flits that follow it will continue to travel along the network, filling up the buffer of the node that is blocked.

It is more efficient than store-and-forward switching because we don't have to wait for the entire message to arrive at each node before forwarding the message on to the next node.

VCT switching is different from wormhole switching in that nodes in VCT can buffer more than one flit. That is, with VCT, when a node is blocked, subsequent flits can continue along the network, gathering together at the blocked node.

VCT switching seems like a good switching method for a cluster interconnect. The only challenge would be that you would have to build larger buffers for each routing node, whereas with wormhole switching, you can get away with small buffers. On the other hand, VCT switching would be able to route a large number of messages at the same time easier, and the overhead for avoiding deadlock would be negligible. One study found that using an adaptive VCT switching algorithm resulted in a higher network saturation point, better performance for networks with heavy traffic, and little hardware overhead in implementing the VCT switching over wormhole switching. (Source: http://cal.postech.ac.kr/slee1/journal_pdf/Adaptive%20Virtual%20Cut-Through%20as%20a%20Viable%20Routing%20Method.pdf)

VCT switching might be a better choice for clusters with very high network traffic. Wormhole switching will give better performance for low network traffic loads. Since messages in VCT can continue moving even if the head of the message is blocked, there will be fewer possibilities for blockage in the interconnection network, resulting in more routing options for each message going through the network at any one time. This results in lower latency when there is high network traffic.