Load distribution with the Proximity Congestion Awareness in a Network on Chip

Erland Nilsson, Mikael Millberg, Johnny Öberg, and Axel Jantsch
Laboratory of Electronics and Computer Systems / Royal Institute of Technology (KTH)
Email: {erlandn, micke, johnny, axel}@imit.kth.se

Abstract

In Networks on Chip, NoC, very low cost and high performance switches will be of critical importance. For a regular two-dimensional NoC we propose a very simple, memoryless switch. In case of congestion, packets are emitted in a non-ideal direction, also called deflective routing. To increase the maximum tolerable load of the network, we propose a Proximity Congestion Awareness, PCA, technique, where switches use load information of neighbouring switches, called stress values, for their own switching decisions, thus avoiding congested areas. We present simulation results with random traffic which show that the PCA technique can increase the maximum traffic load by a factor of over 20.

1 Introduction

On-chip communication becomes a challenge as the number of transistor functions increases on a single silicon die. Clock and data distribution over large distances is impossible to accomplish in a simple manner. Several research groups propose packet switched Network on Chip [1] [2] to address the problem with communication between Intellectual Properties, where each IP-block is a Resource or a part of a Resource. We propose the use of Nostrum, a two-dimensional Network on Chip using hot-potato routing algorithm as switching policy for datagram distribution [3].

Other ways to organise this mesh are for example the flattened torus model [4], or a plain two-dimensional mesh, the latter is also the model we have chosen to implement Nostrum.

In the plain two-dimensional mesh, the number of packets passing through the centre of the mesh is significantly higher compared to packets traveling along the edges. As every Resource may transmit to any other Resource in the mesh with equal probability, most packets will pass through the centre, which becomes a hot-spot. Such a hot-spot is not desirable.

However, hot-spots can be avoided by distributing control information over the network. We call this concept Proximity Congestion Awareness, PCA. The simulations made in this paper are based on random traffic where each Resource has the same probability of communication to any other Resource in either way.

2 Switch load distribution

The load can spread over a larger area by using different routing rules. For example Round Robin [5], local active deflection, and non-local deflection.

PCA can be used to make the load distribution more uniform. Information to help the Switches in their routing decision is sent between the Switches. The informative value PCA is using is called stress value and is sent from one Switch to its neighbours in all directions. The stress value relates to the load level in that Switch. The surrounding Switches receives four such values from its closest neighbours; this helps each Switch to get a picture of the surroundings.

To avoid oscillations, i.e. a situation when two neighbouring switches get a high stress value every other cycle, causing a packet to bounce between the two switches, the stress value should be averaged over a number of switch cycles.

The incoming packets are sorted in priority order, which in the current implementation is the number of switch cycles the packet has been traveling for; the packet with the highest priority gets to make the first choice of output and the following packets in descending order. The packets which preferred output is occupied by a higher priority packet must be forced in a direction that contradicts to the desired.

3 Simulation results & Conclusion

Packets sent from the Resource are first put in a FIFO-buffer. The study of all FIFOs in the network mesh is a method to see how the load is distributed around a hot-spot.

1530-1591/03 $17.00 © 2003 IEEE

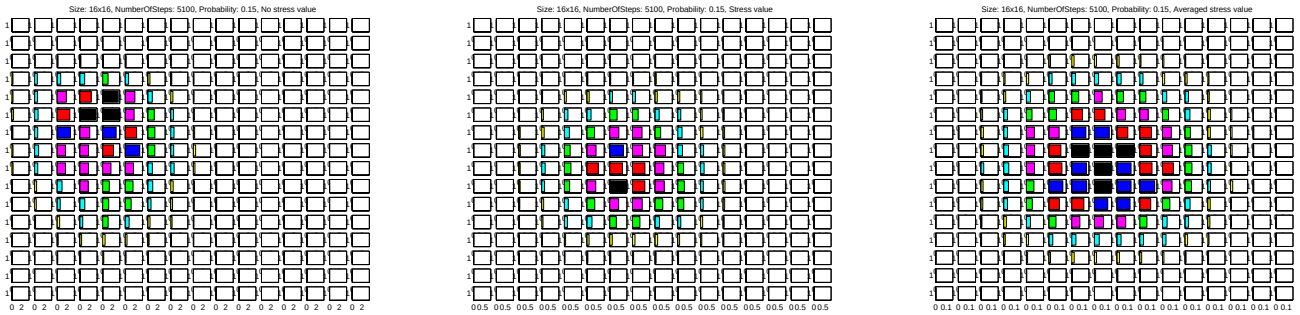

Figure 1. a) Average load i FIFOs without using stress value, largest average number is 3.2 packets. b) Same simulation using stress value, largest average number is 0.9 packets. c) Average load using four cycles averaged stress value, 0.1 packets waiting in FIFO.

The first model acts as a reference where PCA is not used.

The second model is to use PCA with stress values. The stress values are updated every switch cycle.

The third model is one where the stress values are averaged over four cycles. Although, the third model is the most advanced, the stress value averager is fairly small, about 200[1] gates.

In the three cases, the same input data has been used with a mesh size of 16×16. The packet probability is for the first model close to the maximum possible on purpose, shown by simulations [6], to create as much congestion as possible.

An average load of 0.01 tells us that the FIFO is occupied by one packet every 100 switch cycle, which is fairly low. The intention is to show the influence of PCA between the three cases presented.

In figure 1.a, the hot-spot is pushed to the north-west corner, if north is up in the figure. The reason for the non-centered load is a result of the routing decisions in the Switch. The output for the deflected packet is fixed, in this case primarily to the West, secondary North, etc.

The stress value notifies the surrounding Switches about how many packets the Switch handles during that cycle. Using stress values in this manner increases the performance of the Switch since the outputs are ordered in the most preferable order. Compare the maximum value on the x-axis in figure 1.b, which has a maximum value[2] of 0.9, to the value in figure 1.a, which is 3.2.

In figure 1.c, it can be seen that load is distributed over a larger area than before. With experience of the visualised data, the maximum average load is estimated to be 0.15. Compared to the implementation using stress value with no averaging, the average load now achieved is six times less compared to the non averaged stress value. In relation to the most basic implementation, it is enhanced with a factor of more than 20.

It can clearly be seen from the previous discussion that the implementation of a more balanced load using stress values increases the network throughput and decreases the packet delivery time. All simulations in this paper has been made using random traffic. However, it is shown in further experiments in preliminary results that the positive effect of using PCA is also valid for a traffic model where the probability for communication between nearby Resources is higher compared to Resources on far distance.

Synthesis of the third model was made using lsi10k-technology. Optimised on size resulted in an area of 13 964 with a critical path of 79. When optimised on speed, an area of 21 029 and a gate depth of 48 was achieved.

[1] 218 gates using the lsi10k-library.

[2] Observe the scaling of the x-axis since every bar in the whole figure is scaled from the Switch with the maximum average load.

References

[1] L. Benini and G. De Micheli. *Network on Chips: A new SoC Paradigm*. Stanford University, IEEE, 2002.

[2] W. Dally and B. Towles. *Route Packets, Not Wires: On-Chip Interconnection Networks*. Design Automation Conference, IEEE, Las Vegas, USA, 2001.

[3] S. Kumar, A. Jantsch, J-P. Soininen, M. Forsell, M. Millberg, J. Öberg, K. Tiensyrjä, and A .Hemani. *A network on chip architecture and design methodology*. In Proceedings of IEEE Computer Society Annual Symposium on VLSI, April 2002.

[4] W. Dally and C. Seitz. *The Torus Routing Chip*. California Institute of Technology, 1986.

[5] W. Stallings. *Data and Computer Communications*. Prentice Hall International Editions, 5^{th} edition, 1997.

[6] E. Nilsson. *Design and Implementation of a hot-potato Switch in a Network on Chip*. Master's thesis, Royal Institute of Technology, IMIT/LECS 2002-11, Sweden, June 2002.