

Result Analysis of Overcoming Far-end Congestion in Large-Scale Networks

Shalini Vyas ¹, Richa Chauhan ²M.Tech Scholar, Department of ECE, Oriental Institute of science and technology, Bhopal, M.P., India¹Professor, Department of ECE, Oriental Institute of science and technology, Bhopal, M.P., India²

ABSTRACT: Interconnection networks enable fast digital communication between parts of a digital system. Today, interconnections networks are utilized in a range of applications like switch and router materials, processor-memory interconnect, I/O interconnects, and on-chip networks, to call a few. The planning of an interconnection network has 3 aspects—the topology, the routing rule used, and also the flow management mechanism used. Previous work have targeted on understanding near-end congestion – i.e., congestion that happens at this router – or downstream congestion – i.e., congestion that happens in downstream routers. However, most previous work doesn't evaluate the impact of far-end congestion or the congestion from the high channel latency between the routers. During this work, we've a tendency to refer to far-end congestion as phantom congestion because the congestion is not "real" congestion. Attributable to the long inter-router latency, the in-flight packets (and credits) lead to inaccurate congestion information and will cause inaccurate adaptive routing selections. We tend to propose a history window primarily based approach to remove the impact of phantom congestion. We have a tendency to conjointly show however using the typical of native queue occupancies and adding an offset significantly remove the impact of transient congestion.

KEYWORDS: Interconnect, topology, Adaptive Routing.

I. INTRODUCTION

Interconnection networks are the communication infrastructure of any digital system and have a big impact on the system performance and price [1]. The interconnection network could be a key component of a tightly coupled digital computer system. It provides low latency and high bandwidth communication for a range of workloads. Because the normal microchip is being replaced by multithreaded ones or by chip multiprocessors, each the amount of injectors and therefore the total offered load per node has increased considerably. Interconnection networks are an important component of latest pc systems. From giant scale systems to multicore architectures [16], the interconnection network that connects processors and memory modules significantly impacts the final performance and price of the system. As processor and memory performance continues to increase, multicomputer interconnection networks are becoming even a lot of crucial as they principally verify the bandwidth and latency of remote access. A decent interconnection network is intended around the capabilities and constraints of accessible technology. Increasing material router pin bandwidth, for example, has motivated the utilization of high-radix routers [15] during which the increased bandwidth is used to extend the amount of ports per router, instead of maintaining a little range of ports and increasing the bandwidth per port. The Cray black widow system, one in all the primary systems to use a high-radix network, uses a variant of the folded-Clos topology and radix-64 routers a big departure from previous low-radix 3D torus networks. Recently, the appearance of economical optical signaling permits topologies with long channels.

Although congestion could be a common drawback to all varieties of networks, up to currently it's not been a crucial issue in interconnection network style. The explanations for this are multiple. Firstly, the dimensions of most systems (real or modeled) ranged from 64 to 512 nodes, all having single injection queues, whose HOLB (head-of-line blocking) was providing enough hidden management to stop the node from flooding the network. Secondly, the software system overheads were high, so it absolutely was rare for a true system to supply loads near the 100% utilization. Thirdly, networks with few resources (2 or three virtual channels and buffers of a couple of phits) unbroken congestion low as a result of blocked messages unfold on their methods, limiting network output because of contention.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 10, October 2016

Thus, it's not shocking that almost all of the effort went into reducing contention by increasing adaptivity and handling the issues this brought in, like deadlock.

II. LITERATURE SURVEY

Ted Jiang et. al. [1] "Overcoming Far-end Congestion in Large-Scale Networks" Accurately estimating congestion for correct global adaptive routing selections (i.e., verify whether or not a packet should be routed minimally or non-minimally) contains a important impact on overall performance for high-radix topologies, like the dragonfly topology. Previous work have targeted on understanding near-end congestion – i.e., congestion that happens at the present router or downstream congestion – i.e., congestion that happens in downstream routers. However, most previous work doesn't measure the impact of far-end congestion or the congestion from the high channel latency between the routers. In this work, we have a tendency to refer to far-end congestion as phantom congestion because the congestion isn't "real" congestion. We tend to determine the impact of far-end congestion that occurs in large-scale networks due to long latency between neighboring routers and therefore the completely different length channels within the topology. The congestion at the far-end of the channel isn't accurately represented at the near-end since in-flight packets (or credits) that are being transmitted don't represent true congestion. W. J. Dally et. al. [2] "Technology-Driven, Highly-Scalable Dragonfly Topology" Developing technology and increasing pin-bandwidth encourage the utilization of high-radix routers to reduce the diameter, latency, and expenditure of inter-connection system. High-radix networks, however, require longer cables than their low-radix corresponding item. Because cables control network expenditure, the number of cables, and particularly the amount of long, global cables must be decrease to realize a capable network. In this paper, we have a tendency to introduce the dragonfly topology that uses a group of high-radix routers as a virtual router to increase the effective radix of the network. Baba Arimilli et. al. [3] "The PERCS High-Performance Interconnect" The PERCS arrangement was intended by IBM in response to a DARPA challenge that desirable a high-productivity high-performance divides arrangement. A major modernization in the PERCS design is that the network that's designed using Hub chips that are integrated into the calculate nodes. Each Hub chip is about 580 mm² in size, has over 3700 signal I/Os, and is put together in a component that also include LGA-attached optical electronic procedure. John Kim et. al. [4] "The BlackWidow High-Radix Clos Network" In this work describe the radix-64 fold over Clos arrangement of the Cray Black-Widow scalable vector multiprocessor. We describe the Black-Widow network which extent to 32K progression with a worst case diameter of seven hops, and therefore the underlying high-radix router microarchitecture and its implementation. By using a high-radix router with several narrow channels which may be capable to take benefit of the higher pin concreteness and faster signaling rates available in modern ASIC technology. The BlackWidow router is an 800 MHz ASIC with sixty four 18.75 GB/s bidirectional ports for an aggregate off chip bandwidth of 2.4Tb/s. every port consists of three 6.25 GB/s differential signals in each direction. The router supports deterministic and adaptive packet routing with separate buffering for request and reply virtual channels. YARC may be a high-radix router utilized in the network of the Cray black widow multiprocessor. Paul Gratz et. al. [5] "Regional Congestion Awareness for Load Balance in Networks-on-Chip" Interconnection networks-on-chip (NOCs) are quickly replacement other types of interconnect in chip multiprocessors and arrangement on-chip design. Reachable interconnection system use either oblivious or adaptive routing algorithms to observe the way in use by a packet to its goal. Regardless of somewhat higher implementation complexity, adaptive routing benefit from better fault easiness characteristics, will enlarge network throughput, and reduces latency compared to insensible strategy when appearance with non-uniform or bursty traffic.

III. THEORY

Congestion management is one in all the most critical and difficult issues interconnect designers face nowadays. While not appropriate congestion management, network output might degrade dramatically once the network or a part of it reaches saturation. 2 major approaches exist to tackle congestion things. These 2 approaches are preventive and reactive techniques. Every of those approaches has its own set of benefits and drawbacks for explicit congestion situations explored. Congestion management supported path or load distribution alters original communication methods supported the standing of the network, so as to avoid network areas underneath congestion. These techniques inject

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 10, October 2016

packets destined to a node through a series of other methods, to stay latency values low while not reducing traffic load levels. Honest traffic load distribution is accomplished using this type of algorithms as a result of traffic is shipped over idle or less burdened nodes within the network, so avoiding congestion and as a consequence rising performance by avoiding congested spots within the network. Preventive techniques involve solutions based on the conception of increasing resources offered within the system (over dimension) so avoiding congestion by permitting communications to perform while not interruption. A necessary condition for this method to figure is that application behavior is to be known beforehand, so resources to satisfy the task is properly calculable. One disadvantage of this methodology is that the inability to continuously verify a priori resources needed for an application. Also, reservation of resources isn't an optimum resolution because it will result in inefficient network usage and additionally cause congestion, caused by the method of search and reservation of resources itself, that is performed through packets injection among the network. Reactive techniques embrace mechanism to use offered resources properly, adapting to adverse traffic using no quite the minimum needed resources. To accomplish this task, a method of analyzing network standing throughout communications is performed, and once network situation is approaching congestion state, notification procedures are executed so as to alert injecting nodes regarding the case, and allow them to know that some strategy to manage this drawback has got to be done. So, 3 main phases is extracted to perform the congestion control:

- Network monitoring and detection of congestion situations
- Notification concerning the recently found congestion to injecting nodes
- Execute corrective procedures so as to effectively control the congestion situation

IV. METHOD

A. ADAPTIVE ROUTING

Adaptive routing is Routers might not be directly connected to international channels. Every router should choose a global channel using only local data. local data depends only indirectly on the state of the global channel. Adaptivity refers to however routing algorithms choose a path between the set of possible methods for every source-destination try. Deterministic routing algorithms always selected a similar path between a source-destination combine, although there are multiple possible methods. Oblivious routing algorithms select a route while not considering any data regarding the network's gift state (note that oblivious routing includes settled routing). Finally, adaptive routing algorithms adapt to the state of the network, using this state data for creating selections. Adaptive routing algorithms will be classified as a progressive or backtracking. Progressive always make selections at every routing operation, whereas backtracking will return and deallocate resources previously allotted. At lower level, routing algorithms will be classified as according to their minimality as profitable or misrouting. Profitable only provide channel that brings the packet nearer to its destination. Misrouting algorithms may provide channel that sends a packet away from its destination. At the lowest level will be classified in step with the quantity of different methods as fully adaptive (also called fully adaptive) or partially adaptive.

B. DRAGONFLY TOPOLOGY

A dragonfly has narrow wings and wide body. Similarly, this too has large number of groups whereas channels connecting the groups are less and narrow. The dragonfly topology [2] was planned as a high bandwidth topology for large-scale networks in HPC systems. Hierarchical network with 3 levels

- Routers
- Groups and
- Systems

Group radix $k' \gg k$. Ultimate desire is to achieve the global diameter of 1. The following symbols are used in our description of the dragonfly topology:

N Number of network terminals

p- Number of terminals connected to each router

a -Number of routers in each group

k -Radix of the routers

k' - Effective radix of the group (or the virtual router)

h- Number of channels within each router used to connect to other groups

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 10, October 2016

- g- Number of groups in the system
- q- Queue depth of an output port
- q_{vc} - Queue depth of an individual output VC
- H- Hop count
- Out_i -Router output port i

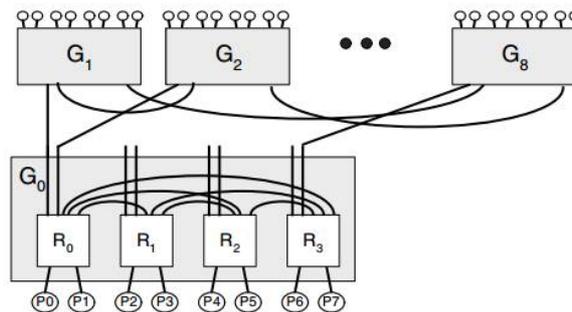
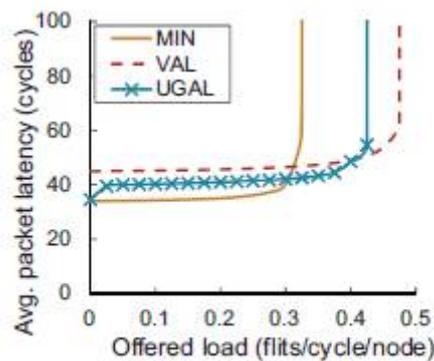


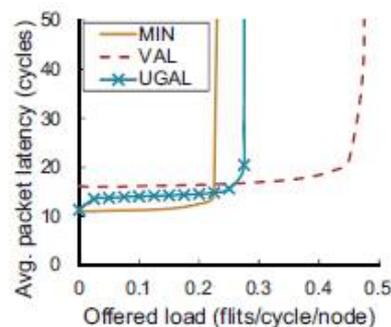
Fig.1 An example block diagram of a dragonfly topology with $N = 72$.

V. RESULT

In this paper results focused on the Dragonfly topology but phantom congestion is applicable to other topologies where global adaptive routing is used to determine between minimal and non-minimal routes. In Figure 2, in this paper evaluate alternative topologies including a simple 8-node ring topology, as well as 16-node 1D flattened butterfly (FBFLY) and 64-node 2D flattened butterfly topologies, both FBFLY with a 4-way concentration.



(a)

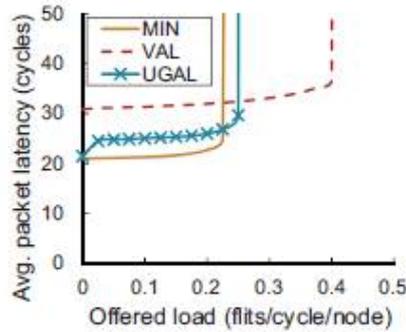


(b)

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 10, October 2016



(c)

Fig. 2 Impact of phantom and transient congestion on alternative topologies, including (a) ring, (b) 1D flattened butterfly, and (c) 2D flattened butterfly

This paper present results for larger packet size and alternative traffic pattern. In Figure 3, we evaluate the proposed adaptive routing algorithm on longer packets and results follow similar trend.

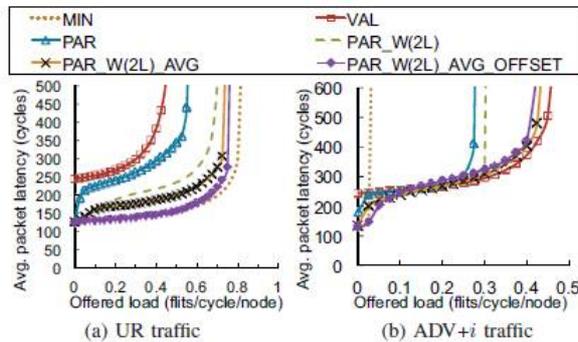


Fig. 3 Evaluation with large packets

This paper also focused our evaluation on uniform random (UR) traffic and adversarial (ADV) traffic pattern since the two traffic patterns represent the two extremes network traffic behavior. Results for random permutation traffic pattern are shown in Figure 4.

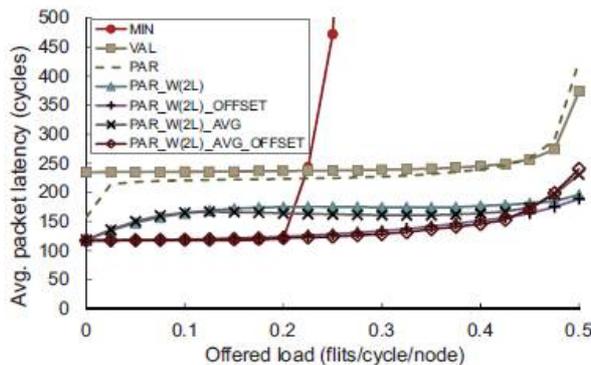


Fig. 4 Performance comparison with random permutation traffic in the Dragonfly

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 10, October 2016

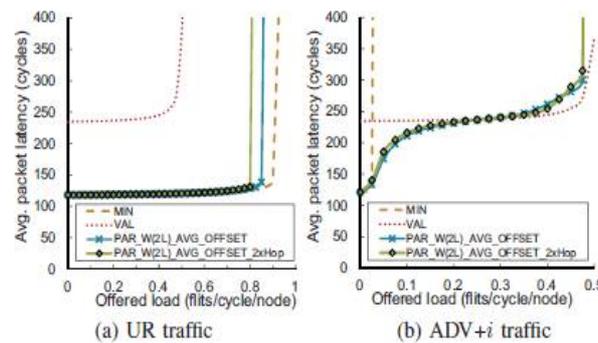


Fig.5 Impact of approximating hop count in the adaptive routing decision

VI. CONCLUSION

In this paper result analysis of Overcoming Far-end Congestion in Large-Scale Networks which provides above results, that shows the complication of the previous ways ought to be increased. We tend to determine the impact of far-end congestion that occurs in large-scale networks due to long latency between neighboring routers and also the totally different length channels within the topology. Our results show that the planned techniques considerably improve the latency and output of the PAR adaptive routing on the dragonfly network.

REFERENCES

- [1] Jongmin Won, Gwangsun Kim, John Kim, S. K., anTed Jiang, Mike Parker, Steve Scott. "Overcoming Far-end Congestion in Large-Scale Networks" IEEE 2015
- [2] J. Kim et al., "Technology-driven, highly-scalable dragonfly topology," in Proceedings of ISCA'08, Beijing, China, June 2008, pp. 77–88.
- [3] B. Arimilli, Ravi Arimilli, Vicente Chung Scott Clark, Wolfgang Denzel, Ben Drerup , Torsten Hoef "The percs high-performance interconnect," in High Performance Interconnects (HOTI), 2010 IEEE 18th Annual Symposium on, Mountain View, CA, August 2010, pp. 75–82.
- [4] S. Scott et al., "The blackwidow high-radix clos network," in Proceedings of ISCA'06, Boston, MA, June 2006, pp. 16–28.
- [5] P. Gratz et al., "Regional congestion awareness for load balance in networks-on-chip," in Proceedings of HPCA'08, Salt Lake City, UT, Feb 2008, pp. 203–214.
- [6] N. Jiang *et al.*, "Indirect adaptive routing on large scale interconnection networks," in *Proceedings of ISCA'09*, Austin, TX, June 2009, pp. 220–231.
- [7] W. Dally, "Virtual-channel flow control," *Parallel and Distributed Systems*, IEEE Transactions on, vol. 3, no. 2, pp. 194–205, 1992.
- [8] D. Helbing, "Traffic and related self-driven many-particle systems," *Rev. Mod. Phys.*, vol. 73, pp. 1067–1141, Dec 2001.
- [9] J. Bell et al., "Boxlib users guide," 2013, Center for Computational Sciences and Engineering, Lawrence Berkeley National Laboratory.
- [10] H. Dong et al., "Quasi diffusion accelerated monte carlo," 2011, Los Alamos National Laboratory.
- [11] D. F. Richards et al., "Beyond homogeneous decomposition: Scaling long-range forces on massively parallel systems," in Proceedings of SC'09, Portland, Oregon, November 2009, Article 60.
- [12] P. F. Fischer et al., 2008, <http://nek5000.mcs.anl.gov>. H. Adalsteinsson et al., "A simulator for large-scale parallel computer architectures," *Int. J. Distrib. Syst. Technol.*, vol. 1, no. 2, pp. 57–73, Apr. 2010.
- [13] "Characterization of the doe mini-apps," National Energy Research Scientific Computing Center. [Online]. Available: <http://portal.nersc.gov/project/CAL/doe-miniapps.htm>
- [14] J. Kim et al., "Flattened butterfly: A cost-efficient topology for highradix networks," in Proceedings of ISCA'07, San Diego, CA, June 2007, pp. 126–137.
- [15] J. Kim et al., "Microarchitecture of a high-radix router," in Proceedings of ISCA'05, Madison, Wisconsin, June 2005, pp. 420–431.
- [16] J. Ahn et al., "Hyperx: topology, routing, and packaging of efficient large-scale networks," in Proceedings of SC'09, Portland, Oregon, November 2009, Article 41.