International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 7 Issue 07, July-2018

A Comparative Study of Deadlock Recovery Schemes in Wormhole Routed Networks

Jameel Ahmad
Department of Computer Science & Engineering
Integral University
Lucknow, India

Dr. Mohammadi Akheela Khanum
Department of Computer Science & Engineering
Integral University
Lucknow, India

Professor A. A. Zilli
Department of Computer Science & Engineering
Lucknow University
Lucknow, India

***Abstract*— Issue of deadlocks is a serious problem in wormhole routed networks. A number of deadlock handling algorithms have been proposed in literature. Some are based on deadlock prevention and avoidance and some on deadlock recovery. Since the deadlocks are rare events, only few resources should be dedicated to deal with these rare events [1]. So if a routing algorithm allows the network fully adaptively to form cycles then a deadlock recovery mechanism must be in place to resolve these deadlock cycles. Cost and performance is a major issue of concern in these algorithms. This paper presents a comparative study of three deadlock recovery schemes.**

Keywords—Wormhole routing, Deadlocks, Recovery, Avoidance, Parallel Computers.

INTRODUCTION

The parallel computers on large scientific and commercial scale are preferred candidates for facilitating high power computation characteristics. These computers system are arranged as an ensemble of the processing nodes and connected as well having their own resources like memory, processors, and other peripheral devices. These nodes are inter-connected with different topologies categorized as direct and indirect topology [4].

Massively parallel computers are important category of parallel machines that provide an envisioned computing platform for solving the challenging problems. Different variations of the machines are available commercially today. These machines are assembly of many nodes which can communicate each other with a set of switches and links of communication are collectively known by the interconnection network [12].

The performance of these inter-connection networks is very critical factor that affects the performance of these parallel machines. Wormhole switching is one of the most promising switching technique which has is applied to these inter-connection networks of the parallel multi-computers. The Wormhole switching technique is well suited for the applications in multi-computer inter-connection networks as it allows the design of simple and low cost hardware and router nodes, while provides low latency and high bandwidth communication [6]. The routing algorithms to route the packets in the wormhole networks are characterized by their approach and how they overcome the issue of deadlocks [7,8].

Earlier proposed work in literature has applied the view that the routing algorithm should be deadlock free. By applying sufficient condition to prevent the deadlock occurrence and by implementing a routing algorithm applying this condition during routing, freedom from deadlocks is guaranteed. These algorithms are known as deadlock avoidance algorithms for routing [4,15]. On the other hand a deadlock recovery algorithm allows the packets to route on all the available channels resource without restricting the possibility of deadlocks.

If deadlocks occur, deadlock detection mechanism works and any particular deadlock recovery algorithm is invoked to break the deadlock cycles [5,11]. Both the techniques are almost similar by the fact that in every case the routing algorithm for resolution utilizes some resources, but the amount and cost of these resources differs. The fact has been established in research that deadlock situations are rare events in general. Due to this effect almost all the deadlock avoidance algorithms use expensive resources more than minimum required to avoid the deadlocks. This observation promotes the deadlock recovery techniques to implement more widely than deadlock avoidance techniques.

The routing algorithms are characterized also by the extent of adaptivity they provide.

The Deterministic Algorithms: The deterministic algorithms in the network do not have any adaptivity. The path for routing is determined by the source node, considering the source node and destination node only. The deterministic routers are fast and simple in design. They suffer serious deadlock situations and do not react to the current state of the network condition [12,15].

The Adaptive Algorithms: The Adaptive algorithms of routing on the other hand react to the network conditions and they allow the packets to route along the alternate paths. The routing decisions are taken by each intermediate node, along the path in a distributed manner [13].

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 7 Issue 07, July-2018

The Adaptive algorithms are again classified as partially adaptive and fully adaptive algorithms on behalf of the amount of the adaptivity they facilitate [10].

Partially-Adaptive Algorithms: Partially adaptive algorithms provide limited adaptivity are called and fully adaptive algorithms allow free routing on all the possible paths. If the routing algorithm allows only routing along shortest paths available between any two nodes, then algorithm is called minimal algorithm, otherwise it is called as non minimal routing or misrouting [14].

Fully-Adaptive Algorithms: The fully adaptive routing algorithms which apply deadlock recovery techniques, truly utilize all the available resources of buffer for routing the packets without excluding any buffer to prevent deadlock situations. That is why they are called as Truly Fully Adaptive Routing(TFAR) algorithms[5,6,11]. The extent of adaptivity that is provided by the routing algorithm is closely related with the underlying router and hardware complexity, which reduces the clock speed of the inter-connection network and eventually for the over all performance of the multi-computer systems [15].

The Dimension Order Routing algorithm is a deterministic deadlock avoidance algorithm [2] where every packet routes in one dimension at the time when arrives at the direction coordinate in every dimension before moving to next dimension. By incorporating a monotonic order strictly for the dimensions traversed and deadlock free routing is guaranteed.

The Planer Adaptive Routing (PAR) is a partially-adaptive deadlock avoidance algorithm. The algorithm restricts the adaptability by allowing only a packet to route with a sequence of two-dimensional planes, at a time until the packet reaches to destination. This reduces the cost of deadlock prevention. The planer adaptive routing requires three virtual channels for every physical channel independent of the dimension of network [9].

Disha is a truly fully adaptive routing algorithm with implementing deadlock recovery mechanism. The base idea of the deadlock recovery is to equip a central deadlock buffer in every node. This deadlock buffer is connected with the crossbar of router by an input port and this is accessible through all the neighbouring nodes. When any deadlock cycle is detected then one of the packets blocked in the deadlock cycle is shifted to the deadlock buffer lane of next node along with its path and from that node the packet continues to move advance using only the “recovery lane” having the deadlock buffers through the network until reaching its destination.

The hardware required for Disha has a central deadlock buffer which is connected to crossbar with an additional input port at every node. A token of hardware signal wires is connected with all the nodes in to synchronize the process of recovery additionally an extra line of status for selecting the deadlock buffer and a crossbar re-configuration buffer is also needed to store a broken connection temporarily inside the crossbar, when forwarding the deadlocked packet [5]. The additional hardware slightly contributes to increase the complexity of crossbar switch. This deadlock buffer is on the recovery path of routing the packets and so it affects the overall performance of the routing process [6]. Disha has been shown nevertheless to outperform most of the other routing algorithms.

2. The Al-Awwami’s scheme: “A New Deadlock Recovery Mechanism for Fully Adaptive Routing Algorithms” proposed by Al-Awwami is a truly fully adaptive algorithm along suggests a new deadlock recovery mechanism. It suggests a deadlock recovery using minimum hardware resource and also the additional hardware does not lie on the critical path of the routing the normal packets and does not affect the overall processing speed of the routers. The truly fully adaptive algorithm for routing routes the packets on the all available channels and utilize all the buffer resources regardless the possibility of deadlock occurrence. Once any deadlock is detected, the suggested recovery mechanism is invoked. The mechanism for deadlock detection implemented is based on the method described in [11].

The deadlock-recovery scheme proposed is a new approach. Since deadlocks are rare, only as few resources as possible should be dedicated to handle this rare event. The proposed approach takes advantage of the concept behind wormhole routing. Namely the low number of edge buffer requirements per channel, which can be as low as one flit buffer, and the flow control mechanism already in place, and which carries control information in the opposite direction to that of data flow.

2.1 Operation of the Al-Awwami’s mechanism. This mechanism equips a central buffer of same capacity as the edge-buffer which is normally have one flit or two flits in depth. This central buffer is used to break any deadlock cycle at a node where a header of any blocked packet is residing. The mechanism of the flow control has also been extended by an additional signal as the break connect line. When any deadlock situation is detected at any node then the edge buffer which contains the header of a blocked packet is managed to store in the central buffer and then it is cleared to receive the other flit(s). The flow control in opposite direction by using the additional line signals the previous node from where the packet is coming, to break and to free the connection through the crossbar and to enable the other packets to move advance through it. This operation is termed as a break-Mode operation [1,2].

Every node keeping part of any blocked packet constantly propagates a break mode signals to its previous node and continues till the tail flit or source node of packet is reached. A break-mode operation performs many steps. It starts by moving the flits of a deadlocked packet which is at the edge buffer to a central buffer. Next this saves input and the output ports number used by the packet through crossbar of the node. Now it clears that connection of the crossbar to so allow other packets to utilize it. In this way the packets remaining which were involved in deadlocked cycles could advance and ultimately free up the occupied resources. Once a node with

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 7 Issue 07, July-2018

diverted header in the central queue finds a free edge-buffer then the header is again rerouted.
If this routing is successfully completes then a flow control re-connect signal is now sent back to that very node which is containing the remainder packets just to re-establish the connections through the crossbar and to resume the operation interrupted for routing the packet. This is the signal which is referred as a connect mode operation and propagates back through the same path followed by the break-mode signal and then reverses which latter has done by using the trails which were previously saved.

Figure 1 shows the operation of this mechanism. If the packet-1 is marked as deadlocked packet then the header Hl of the packet moves to the central buffer of the node. A break mode signal then is sent back to the previous node having the data flits of the packet Dl. The break mode operation moves the data flits to the central buffer and clears the crossbar connection that was established by the packet. In this way the packet-4 which was blocked by the packet-1 can now advance.

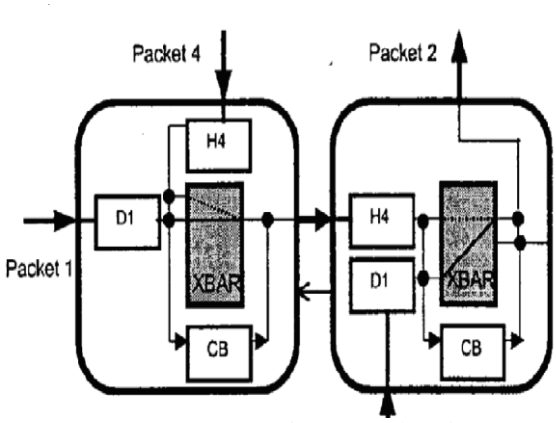

Figure 1: Operation of Al-Awwami's Scheme [1]

In the wormhole routing if the header of a packet is accepted by the channel, then the rest all the flits of a packet are also accepted in pipeline fashion by the channel. This approach doesn't allow the interleaving of the flits between the packets. But in this approach, limited interleaving of the flits to break the deadlock cycles has been allowed.

2.2 Hardware Requirements
The hardware requirement for this approach is kept moderate and also this doesn't increase the complexity either of the crossbar switch and the controller of the virtual channel. It also doesn't lie on the critical-path of the routing decisions for the normal packets. Both the factors increase the cost and also decrease the speed of the routers [4]. Every node requires a central buffer. The capacity of the central buffer and the edge buffer should be same. This central buffer connects with all input ports of that switch and can give output to all the output ports. Two 7 bits registers are needed per node to store the inputs and outputs switch port. A break-connect signal wire also required to break and re-connect packets. This wire is too much similar as the flow- control wires that are used for the wormhole switching. The main function of this is to signal to the previous node to performing either a break-mode or a connect-mode operation. Here only one wire has been used instead of two wires for every mode with a bit. When any break signal is received at a node then it toggles the bit, so that when receiving next time it can be interpreted as a connect-signal. This is sufficient hardware is to perform the sequential deadlock-recovery. That means one deadlocked packet is treated at a time.

3. Deadlock Recovery Scheme 'Disha Enhanced'
This scheme aiming the optimization of the performance for routing of an inter-connection network on optimized cost. This has been devised by reducing the deadlock buffers 40% from the alternate columns of two dimensional mesh inter-connection network against the deadlock recovery scheme 'Disha' and the cost of the crossbar ports are also has been reduced to optimize the cost and the complexity of the router and so speeding up the routing to enhance the overall performance of the proposed deadlock recovery mechanism. The cost of routing and complexity of the routers also has been further optimized by removing the number of virtual channels in the interconnection network from the alternate columns [3].

The scheme is based on the established assumption that the situations of deadlocks are although severe but rare, in the inter-connection network and it doesn't make sense to engage resources permanently, just for rare events. This scheme allows a fully-adaptive routing without restrictions on the path selection and if deadlocks occur then the recovery scheme in place resolves the cycles of deadlock and resumes the inter-connection network again back to normal.

The base mechanism 'Disha' has used an additional deadlock buffer on every router, but the proposed mechanism suggests the deadlock buffer at alternate columns one after one to optimize the cost and routing complexity at the routers. A deadlock recovery lane to support the recovery mechanism has been provided. Few additional d buffers have been equipped at some nodes to complete the recovery lane as alternative path. If some of the packets are blocked in the deadlock cycle then any one of the packet is moved to the path of the deadlock buffer and ultimately reaching to its destination, thereafter breaking the deadlock cycle and resuming back the network to move advance other packets towards their way. Since only one packet during shifting is allowed to move along the recovery path, it is sure that the recovery lane is deadlock free. So the suggested scheme here is using fewer resources to recover deal with deadlocks and improving the performance of the proposed scheme. In this way enhanced performance and optimized cost framework has been suggested [1,5].

3.1 Recovery Path
In case of deadlock situations a recovery path has been provided to shift a packet on recovery path to ensure deadlock free recovery scheme. Seven additional flit buffers have been taken to complete the recovery path. The recovery path is acyclic to guarantee that the packet on recovery path can not further stuck in cycle. The figure below shows the recovery path.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 7 Issue 07, July-2018

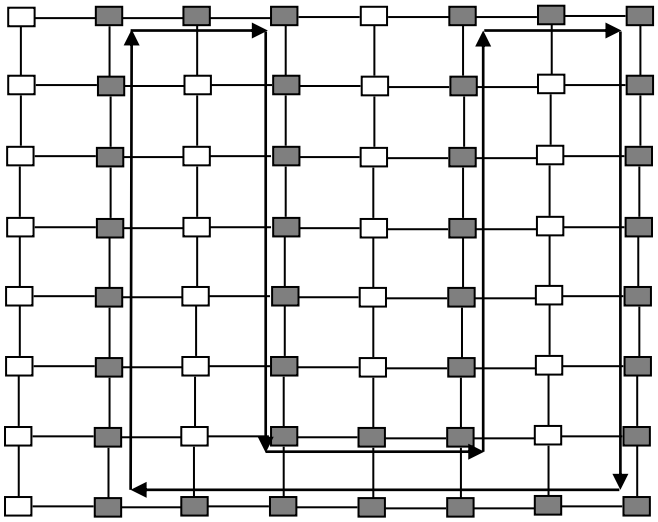

Figure2: Recovery Path

As shown in the above figure, every node on the recovery path is equipped with deadlock buffer to facilitate deadlock free routing on the recovery path.

4. THE COST MODEL

Data-Through Latency

$T_{DT} = T_{FC} + T_{CB} + T_{VC}$ [6]

T_{FC}: Flow Controller Delay

T_{CB}: Crossbar Switch Delay

T_{VC}: Virtual Channel Controller Delay

For gate array technology 'CMOS'

$T_{FC} = 2.2$

$T_{CB} = 0.4 + 0.6 \log P$ ns where P = number of input ports

$T_{VC} = 1.24\ 0.6 \log V$ (V=3)

Data-through latency for 'Disha' $T_{DISHA} = 7.1$ ns

Data-trough latency for Al-Awwami's scheme = 7.0 ns

Data-trough latency for the proposed scheme = 6.7 ns

So, Al-Awwami's scheme is 1.4% faster than 'Disha'

The proposed scheme is 5.63% faster than Disha on reduced resources

The proposed scheme is 4.28% faster than Al-awwami's scheme.

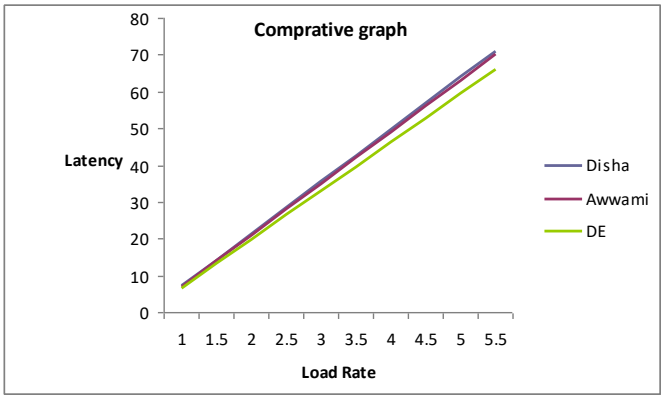

Figure3: Comparative Graph

5. CONCLUSION

In view of the study and analysis made in this paper, we reach to the conclusion that the deadlock recovery approach after the occurrence and detection is far better than deadlock avoidance and prevention. This approach not only optimizes the resources, but reduces the complexities of the routing and improves the performance as well. The "Disha" scheme was the beginning of this approach. Following this approach the Al-Awwami's scheme, added moderate resources and slightly improved performance
than "dishsa". The third recent approach is based on the "Disha" approach that further reduces the resources and achieved better performance than both the earlier approaches. The 'Disha Enhanced' scheme has sustained the basic mechanism of 'Disha' but optimized the resources and complexity up to great extent and eventually achieved improved performance than 'Disha' and Al-Awwami's scheme.

REFERENCES:

[1] Al-Awwami, M.S. Obaidat and M. Al-Mulhim, "A New Deadlock Recovery Mechanism for Fully Adaptive Routing Algorithms", Department of Information and Computer Science, KFUPM, Dhahran 31261 Saudi Arabia, Published in Performance, Computing and Communications Conference, IPCCC, Conference Proceeding of the IEEE International, 2000.

[2] Al-Awwami, M.S. Obaidat and M. Al-Mulhim (2001), Preemptive Deadlock Recovery Mechanism for Wormhole Networks, Manmouth University and KFUPM, 2001 Simulation Councils Inc. ISSN 0037-5497101 USA.

[3] Jameel Ahmad, Mohammadi Akheela Khanum and A.A. Zilli, "A Cost Effective and Performance Enhanced Deadlock Recovery Scheme for Wormhole Routed Networks",Published in AISC Book Series of Springer, Indexed in ISI Proceedings, EI-Compendex, DBLP, SCOPUS, Google Scholar and Springer Link, 2017.

[4] Prashant Mohapatra,"Wormhloe Routing Techniques for directly Connected Multicomputer Systems, 201 Cover Hall Lowa State University Ames, IA 50011, 2001.

[5] Anjan K. V., Trimothy Mark Pinkston, "An Efficient, Fully Adaptive Deadlock Rccovery Scheme, DISHA," In Proceedings of the 1995 International Symposium on Cornputer Architecture, pp. 201-210, June 1995.

[6] Andrew A. Chien, "A Cost and Speed Model for k-ary n-cube Wormhole Routers," IEEE Transactions on Parallel and Dislributed Systems, Vol. 9, No. 2, pp. 150-162, February 1998.

[7] Trimothy Mark Pinkston, Sugath Wamakulasuriya, "On Deadlocks in Interconnection Networks," In Proceedings of the 1997 International Symposium on Computer Architecture, pp. 38-49, June1997.

[8] Sugath Wamakulasuriya, Trimothy Mark Pinkston, "Characterization of Deadlocks in Interconnection Networks," In Proceedings of the 1997 International Conference on Parallel Processing, pp. 80-86, 1997.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 7 Issue 07, July-2018

[9] Jae H. Kim, "Planar-Adaptive Routing (PAR): Low-Cost Adaptive Networks for Multiprocessors," Master Thesis University of Illionosis at Urbana-Comeign, 1993.

[10] Kazuhiro Aoyama, "Design Issues in lmplemcnting an Adaptive Router," Master Thesis, University of Illinois at Urbana-Champaign, 1993.

[11] P. Lopez, J. M. Martinez, J. Duato, "A Very Efficient Distributcd Deadlock Detection Mechanism for Wormhole Networks," In the International Symposium on High- Performance Computer Architecture, pp, 57-66, 1998.

[12] Lionel M. Ni, Philip K. McKinley, "A Survey of Wormhole Routing Techniques in Direct Networks," W E Computer, Vol. 26, No. 2, pp. G2-76, February 1993.

[13] Kazuhiro Aoyama, "Design Issues in Implementing an Adaptive Router," Master Thesis, University of Illinois at Urbana-Champaign, 1993.

[14] Kevin Bolding, "Chaotic Routing - Design and Implementation of an Adaptive Multicomputer Network Router," Doctoral Dissertation, University of Washington, 1993.

[15] William J. Dally, "Performance Analysis of k-aty n-cuhc Interconncction Netwarks," EEE Traiislrclioris on Conputem, Vol. 39, No. 6, pp. 775-785, Junc 1990.