Multipath Fault-Tolerant Routing Policies to deal with Dynamic Link Failures in High Speed Interconnection Networks

Gonzalo Zarza†
Computer Architecture and Operating Systems Department,
Universitat Autònoma de Barcelona, Spain
Advisor: Dr. Daniel Franco‡
†gonzalo.zarza@uab.es, ‡daniel.franco@uab.es
Supported by the MICINN Spain, under contract TIN2007-64974.

Computer science has become an indispensable tool and a valuable source of knowledge for modern societies, especially in the last few decades. Along these years, computing systems have opened a trend in daily behavior and lifestyle of many people by becoming the engine of an increasing number of essential applications and services. Since then, relations between human societies and computing systems have become remarkably strong and the demand for even more computing power has never stopped. Clusters of computers, together with massively parallel processing systems, have become the two prevailing approaches for addressing this increase in computing power demand. Regardless of the approach being used, these High Performance Computing (HPC) systems are made of thousand of components, including processing nodes, memory banks, disks, and other peripherals, where the high-speed interconnection network plays a key role by allowing these systems to work as large coherent entities. Under these circumstances, it is imperative to keep the interconnection network up and running as long as possible since networks failures have serious impacts on the overall computing system. This provides the motivation of the thesis.

In this thesis, we present fault-tolerant routing policies based on concepts of adaptability and deadlock freedom, capable of serving interconnection networks affected by a large number of dynamic link failures. The strongest point of this thesis is that it provides a simple but complete solution to the problem of dynamic fault tolerance in interconnection networks. The proposed solution does not require any information about network faults when the system is started or restarted. Throughout the thesis, we present the conception, design, implementation and evaluation of two contributions. The first of these contributions is the adaptive multipath routing method *Fault-Tolerant Distributed Routing Balancing* (FT-DRB) [2], [4], [5]. This method has been designed to exploit the communication path redundancy available in many network topologies, allowing interconnection networks to perform in the presence of a large number of faults. The second contribution is the scalable deadlock avoidance technique *Non-blocking Adaptive Cycles* (NAC) [3], [6], specifically designed for interconnection networks suffering from a large number of failures. This technique has been designed and implemented with the aim of ensuring freedom from deadlocks in the proposed fault-tolerant routing method FT-DRB.

Fault-Tolerant Distributed Routing Balancing (FT-DRB)

Conceptually, the proposed fault-tolerant routing method is based on the state information of source-destination paths. This information includes latency values of the path and the links state information of the communication path. If there are no link failures along the path, each application message records latency information about the path it traverses. Once the message reaches the destination, this terminal node sends the latency information of the path back to the source terminal node, using an ACK message. If there is at least one link failure along the source-destination path, it is discovered when a packet tries to use the faulty link. In the fault tolerance theory, this first action would correspond to the *error detection phase*. After this phase has been completed, *damage confinement* and *error recovery* must be provided. To this end, the network node which discovers the failure sends back a special ACK packet, in order to alert the source node about the failure in the path. The latter action corresponds to the phase of *damage confinement*. Almost simultaneously, those messages that have been already sent through the path where the link failure has been discovered are rerouted towards the destination node. This corresponds to the *error recovery* phase of the fault tolerance theory. As this rerouting action is intended to be a fast and temporary response to link failures, it may not be the optimal solution. For this reason, the proposed method includes a third and last action, which represents the phase of *fault treatment and service continuity*. At this point, the source node disables the faulty path and reconfigures new paths for the following messages, in order to avoid faults, ease routing paths, and improve performance. Once those new paths have been configured, their latency values are recorded and then sent back from the destination to the source node. Counting on this information, the source node is able to calculate the number of alternative paths that must be used and can distribute messages among them, according to the network traffic burden. Using one or more alternative paths, the method is able to avoid and/or circumvent link failures, while improving the system performance by means of distributing and balancing communications among the alternative paths. The entire set of actions of FT-DRB is summarized in the node diagram of Fig. 1 .

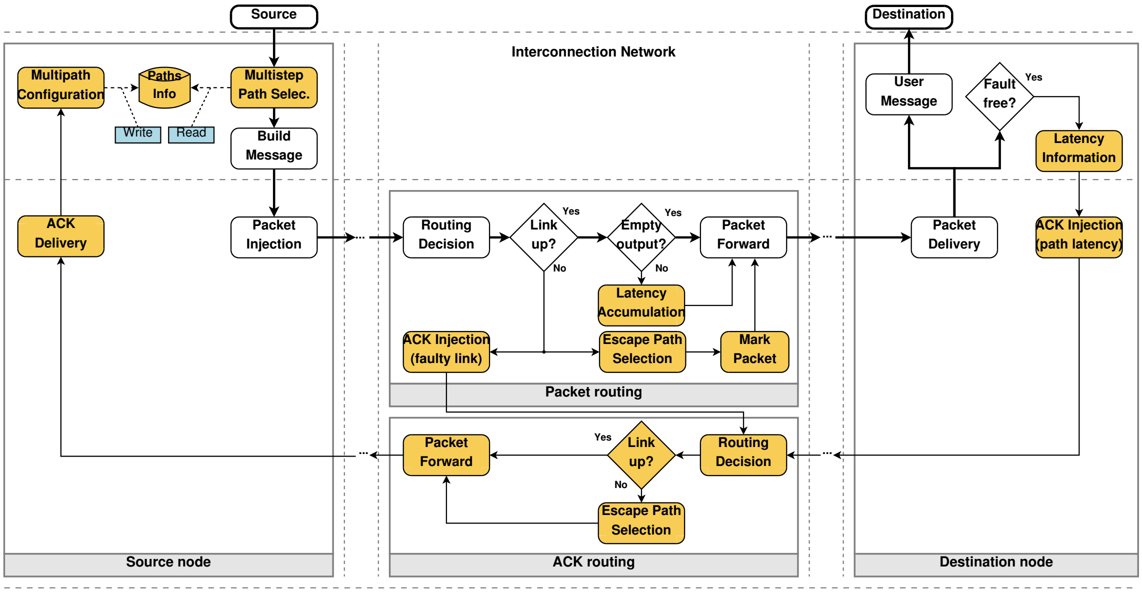

Figure 1. FT-DRB functionality.

Non-blocking Adaptive Cycles (NAC)

The proposed scalable deadlock avoidance technique has been designed with the aim of avoiding deadlocks in FT-DRB. NAC is a three-stage approach that covers the key aspects of deadlock avoidance. These aspects include the detection of deadlock prone situations, the identification of the routing cycles involved in these situations, and the application of predefined protocols to ensure the normal functioning of the system (recovery of packet forwarding) under these circumstances.

The aim of NAC is to prevent the saturation of some buffers, to avoid deadlocks caused by blocked packets that cannot advance because the buffers requested by them are full. This is equivalent to ensure that the *hold-and-wait* and *circular wait* conditions proposed by Coffman et al. [1] are not satisfied. A *hold-and-wait* is a situation where a packet requests resources held by other packets while holds resources requested by other packets. On the other hand, under a *circular wait*, circular chains are formed along the network where packets hold resources that are being requested by the next packet in the chain. NAC avoids these two Coffman conditions by means of adding an one-slot deadlock avoidance buffer to each input buffer, and applying a simple set of actions when accessing output buffers with no free space. These actions are only applied under specific circumstances directly related to the free space left in the buffers of the local router as well as in the next router in the path along source-destination pairs. It is worth noting that each router already knows its buffer space availability as well as the available space at the input buffer of all its neighbors by means of the flow control mechanism (i.e. credit-based systems). This is part of the information that our technique needs for avoiding deadlock occurrences.

References

[1] E. G. Coffman, M. Elphick, and A. Shoshani. System Deadlocks. *ACM Computing Survey (CSUR)*, 3(2):67--78, 1971. ISSN 0360-0300. doi:10.1145/356586.356588.

[2] G. Zarza, D. Lugones, D. Franco, and E. Luque. A Multipath Fault-Tolerant Routing Method for High-Speed Interconnection Networks. In *Euro-Par 2009: Proceeding of the 4th International Euro-Par Conference on Parallel Processing*, volume 5704 of *LNCS*, pages 1078--1088, August 2009. doi:10.1007/978-3-642-03869-3_99.

[3] G. Zarza, D. Lugones, D. Franco, and E. Luque. Deadlock Avoidance for Interconnection Networks with Multiple Dynamic Faults. In *18th Euromicro International Conference on Parallel, Distributed and Network-Based Computing (PDP)*, pages 276--280, Feb 2010. doi:10.1109/PDP.2010.82.

[4] G. Zarza, D. Lugones, D. Franco, and E. Luque. FT-DRB: A Method for Tolerating Dynamic Faults in High-Speed Interconnection Networks. In *18th Euromicro International Conference on Parallel, Distributed and Network-Based Computing (PDP)*, pages 77--84, Feb 2010. doi:10.1109/PDP.2010.65.

[5] G. Zarza, D. Lugones, D. Franco, and E. Luque. Fault-tolerant Routing for Multiple Permanent and Non-permanent Faults in HPC Systems. In *Proceedings of the International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA)*, pages 144--150, Las Vegas, NV, USA, July 2010.

[6] G. Zarza, D. Lugones, D. Franco, and E. Luque. Non-blocking Adaptive Cycles: Deadlock Avoidance for Fault-tolerant Interconnection Networks. In *IEEE International Conference on Cluster Computing Workshops and Posters (CLUSTER WORKSHOPS)*, pages 1--4, Sep 2010. doi:10.1109/CLUSTERWKSP.2010.5613085.