RapidIO®

White Paper

Synopsis of the Multicast Extensions

Based on:

Part XI: Multicast Extensions Specification
Rev. 1.2, 08/2004

Abstract

The goal of the Multicast specification is to add a simple and standard mechanism to the existing RapidIO specifications that provides multicast functionality to a system. The addition of this support into the RapidIO interconnect provides a highly efficient mechanism for sending messages to multiple destinations.

Disclaimer

Do not design to this document. It is intended to provide an overview of the features of the specification. Please refer to the full-length specification for design guidance.

Introduction

The duplication of a single message and sending it to multiple selected destinations is known as 'multicast'. It is a useful mechanism in many computing and communications systems and can be accomplished by a variety of means. The most efficient method is to have hardware provide the mechanism for the duplication and forwarding of messages. This preserves processor resources at the end points for conducting other operations.

Currently, it is possible to multicast from an end point within the RapidIO standard. However, the *RapidIO Multicast Extensions* has defined a standard interface for implementing multicast from the switch devices within the RapidIO fabric. With a standard interface and definition for behavior, it allows for different components within the system to inter-operate. It also facilitates the existence of a common software driver. The interface has been defined with some level of abstraction in order to maximize implementation flexibility.

A RapidIO switch that does not support multicast can co-exist in a RapidIO fabric with other switches that do support multicast. The only requirement is that the switch be capable of routing the destination IDs used for multicast transactions.

The multicast mechanism fulfills the following goals:

- Simple – streamlined implementation
- Compact – does not cost excessive silicon area in a switch
- Robust – same level of protection and recovery as the rest of RapidIO technology
- Scalable – can extend to multi-layer switch systems
- Compatible – works with all physical layers

Multicast Operation

A RapidIO multicast operation consists of the replication of a single packet so that it can be received by multiple end points. This replication is performed by the switch devices in the fabric rather than by the end point itself. This is designed in this manner in order for the capability of packet replication to scale up with the number of switches (and hence possible end points) in a system. The end points in the system are largely unaffected by the multicast function.

Each switch may be individually programmed to control which ports the replicated packets are sent to and, therefore, the specific set of end point devices receiving the replicated packet. The replication process transmits the packets out of the appropriate ports without modifying the packets.

The multicast specification defines multicast for transactions that do not require responses (such as NWRITE and SWRITE transactions). This greatly simplifies multicast support for RapidIO switches, which will therefore have no need to aggregate responses from other types of RapidIO operations. Multicasting transactions that require responses are more complex and less common. However, use of multicast for transactions that require a response is not precluded. As the specification does not define the behavior of an implementation in this case, support of this capability is implementation specific.

Multicast operations are defined by configuring "multicast masks" and "multicast groups". The set of target end points that are to receive a particular multicast packet is known as a multicast group. Each multicast group is associated with a unique destination ID. The destination ID of a received packet allows a RapidIO switch to determine that a packet is to be replicated for a multicast.

A multicast mask is a value that controls which egress ports one or more multicast groups are associated with. Conceptually, a multicast mask is a register with one enable bit for each possible switch egress port. There is one set of multicast masks for the entire switch.

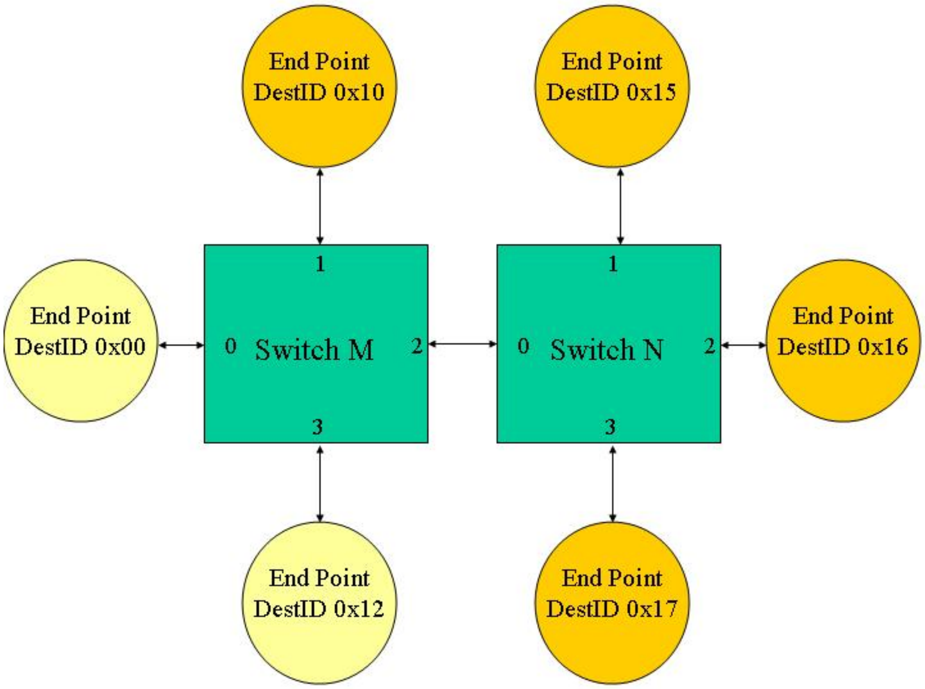

Figure 1: *Multicast System Example*

In this example, a multicast group is defined for end point 0x00 that is comprised of end points 0x10, 0x15, 0x16 and 0x17. This multicast group is given the destination ID 0x80. Software configures the switch devices in the fabric to associate the destination ID that represents the particular multicast group with the multicast mask. For Figure 1, switch M associates destination ID 0x80 with egress ports 1 and 2, and switch N associates destination ID 0x80 with ports 1, 2, and 3.

Configuring a RapidIO switch to replicate packets for a multicast group is a two-step process. First, a list of egress ports is set in a multicast mask list. Second, one or more destination IDs, which represent the multicast groups, are associated with the multicast mask in the switch. During normal system operation, any time a switch receives a packet with a destination ID that has been associated with a multicast mask, it will send that packet to all egress ports enabled by that multicast mask.

Configuration of the multicast parameters is performed using CSRs (Control and Status Registers) and CARs (Capability Registers). The defined CSRs allow a switch to associate destination IDs with multicast masks, using a small number of write operations. The CARs are used to configure which multicast features are to be invoked and the associated parameters the switch will support.

While each destination ID is associated with a unique multicast group, the programming model allows a destination ID to be mapped to a different multicast mask for each port on the switch. However, for each port, a destination ID can be associated with at most one multicast mask. It is also possible to map a given destination ID to the same multicast mask for all ports.

A RapidIO switch may be capable of supporting a large number of multicast groups by dedicating a sequential range of destination IDs to an equal number of sequentially numbered multicast masks. This method, called block association, provides a quick method to configure a series of masks by using only a couple of writes to the CSRs. A switch may also be designed that does not require all multicast destination IDs to be sequential. The programming model supports both of these implementations.

A packet will never be multicast back out of the port it was received on even if it is included in the multicast mask for that destination ID. This allows a group of end points, which need to multicast to each other to share the same multicast mask, reducing the complexity and amount of configuration needed.

RapidIO packets that are in the same multicast group and flow will be multicast on the egress ports in the same order that they were received. By maintaining ordering within the flow, it allows an application to return a message indicating that a transfer has completed. It also keeps the complexity to a minimum, since a system would not need to use overhead to re-order packets.

Conclusion

The Rapid IO Multicast Extensions Specification adds a simple and highly efficient mechanism to the RapidIO framework that provides multicast functionality to a system. Multicast is provided within the hardware switches, which saves device resources at the end points and scales easily with the number of end points in a system.

The advantage of this approach of multicasting is that the packet replication does not have to occur until the very last point in the system where the traffic forks. This cuts down on the overall system traffic necessary. And, since RapidIO fabric uses destination addressing, there is no need to provide an additional field to indicate a multicast packet. This keeps the packet overhead to a minimum.

Configuration of this system is both straightforward and flexible.

RapidIO is a registered trademark of the RapidIO Trade Association.

www.rapidio.org