RapidIO Trade Association

Suite 325, 3925 W. Braker Lane

Austin, TX 78759

512-305-0070 Tel.

512-305-0009 FAX.

TWG Second Showing

Item 11-01-00001.006

Subject: Per Port Reset Specification

Background: Error recovery when one link partner has been reset is now unbearably complicated in RapidIO. This showing proposes a new capability, Per-Port Reset, which will make clean up after an asymmetric reset easier.

Contributor: Barry Wood, IDT

Comment Expiration Date:

Distribution: RapidIO TA Technical Working Group members

1.2 Introduction

The RapidIO specification does not allow simple recovery in the case that one link partner is reset and the other is not. Reset of one link partner is a common occurrence in systems. For example, if an endpoint experiences a software failure, that endpoint must be reset in order to recover it. Support of hot swap requires the ability to recover link communication when one link partner has been reset.

1.3 Discussion

When one link partner is reset after packets have been exchanged on the link, it is unlikely that the ackIDs tracked by the link partner which was not reset will match those of a newly inserted link partner. It is therefore necessary to resynchronize ackIDs before packets can be exchanged on the link once a link partner reappears.

When a link partner is removed, it is likely that error conditions (input-error stopped, output-error stopped, port-error) are detected on the link. These error conditions must also be cleared before packets can be exchanged on the link.

The Gen 2 specification supports the LP-Serial Extended Feature Block with Software Assisted Error Recovery. These are insufficient to clear error conditions as there is no mechanism to force the link partner to issue a link-request, which is necessary to clear input error-stop conditions on the near end of the link. They also rely on a relatively short period for the Link Timeout period. Clearing error conditions and sync'ing ackIDs requires knowledge of which port is connected on the link partners device. There is no way to learn this information with the current scheme.

One idea is to trigger automatic link recovery when a "Fatal Error" condition is detected i.e. illegal ackID is returned in a link-response control symbol.

Recovery requires ackID synchronization and clearing error states/error recovery. This is similar to the behavior invoked by a reset. It is proposed that a new "Port Reset" reset request control symbol be defined which will cause the following actions when received:

- Clear received, transmitted and outstanding AckID values to 0.
- Cause the discard of all currently outstanding unacknowledged packets.
- Clear all input, output, and retry error states.
- Cancel all outstanding link-requests.

Once the "Port Reset" reset request is transmitted, the above actions must occur on the transmitting link partner. This could be triggered automatically by the transmission of the "Port Reset" request, or by a newly defined control bit.

Note that the "Port Reset" request does not require packet discard by the port receiving the "Port Reset" request. While it is possible to define two varieties of "Port Reset", one which forces the discard of packets and another which requires packet retention, for simplicity

packet discard is controlled by the mechanism agreed on as a result of the discussion in the Hot Swap Showing, 10-06-00000.

1.4 Proposed Changes to Part 6 LP-Serial Physical Layer Specification

In order to support error state clearing, ackID synchronization and the resumption of packet exchange when a link partner has been reset, Part 6 control symbol support is enhanced to include a "Reset Port" command for a Link Request control symbol.

1.4.1 Proposed Changes for Section 3.5

Add following definition of the new "Reset-Port" command value to table 3-8 in section 3.5.5:

Cmd value '0b0101', cmd Function "Reset-Port", Packet Delimiter indication of "*", Reference Section 3.5.5.3.

1.4.2 New Section 3.5.5.3 Reset Port Command

Create a new section 3.5.5.3 as follows:

"A reset-port command is intended to allow packet exchange to resume after an unrecoverable link error condition has been detected and system software has handled this condition. Examples of the use of per-port reset are link recovery after a field replaceable unit has been inserted, and when one link partner has failed and/or been reset but not the other.

Scenarios which require a reset-port command for recovery also require packet discard to prevent packets which are undeliverable due to the unrecoverable link error condition from creating a cascade congestion failure of the entire system. Packet discard mechanisms which are part of the RapidIO Error Management/Hot Swap extensions may be activated by the unrecoverable link error condition. Implementation specific packet discard mechanisms may also be activated by the unrecoverable link error condition. System recovery from packet discard is vendor specific, and outside the scope of this specification.

A device which receives a reset-port command shall perform the following:

- The design shall reset all ackID tracking logic for packets received, transmitted and unacknowledged to a state consistent with a power-up reset.
- Clear all input-error, output-err, input-retry, output-retry and port error states.
- Clear the tracking of link-request/input status control symbol requests received or transmitted.
- The port's initialization state machine shall be reset.
- Packet discard mechanisms shall be deactivated.

- Register values shall reflect updated status based on the above changes.
- Registers which are not affected by the above changes shall retain their values.

The reset-port command shall not generate a link-response control symbol.

The timing relationship between deactivation of packet discard mechanisms and the arrival of packets is not deterministic. For this reason, no assumptions shall be made about the effect of a reset-port command on packet storage for transmission or reception. The effect of a reset-port command on packet storage is implementation specific behavior and outside the scope of this specification.

Note that transmission and reception of a reset-port request may trigger additional functionality defined in Part 8: Error Management/Hot Swap Extensions Specification.

After a port has transmitted a reset-port request, if the port's initialization state machine transitions to the SILENT state within one link-response timeout period the port shall behave as if it has received a per-port reset request. The timeout period shall be tracked only for the most recently transmitted per-port reset command.

Due to the undefined reliability of system designs it is necessary to put a safety lockout on the reset-port function of the link-request control symbol. A port receiving a reset-port command in a link-request control symbol shall not perform the reset-port function unless it has received four reset-port commands in a row without any other intervening packets or control symbols, except status control symbols. Such a sequence is known as a reset-port request. Reset-port requests are intended to prevent spurious reset-port commands from inadvertently resetting a port.

Note that it may be necessary to transmit more than four per-port reset commands to overcome transmission errors."

1.5 Proposed Changes to Part 8 Error Management/Hot Swap Extensions

The Error Management/Hot Swap specification must be changed to clear error conditions which can prevent packet exchange with the link partner. Specifically, the Output Failed-encountered condition and the associated Error Rate Counter field must be cleared.

1.5.1 Changes to Section 2.2 Additions to Existing Registers

Change Table 2-2 from:

'

Table 1-1. Bit Settings for Port *n* Error and Status CSRs

Bit	Name	Reset Value	Description
5	Output Packet-dropped	0b0	Output port has discarded a packet. Once set remains set until written with a logic 1 to clear.
6	Output Failed-encountered	0b0	Output port has encountered a failed condition, meaning that the port's failed error threshold has been reached in the Port *n* Error Rate Threshold register. Once set remains set until written with a logic 1 to clear.
7	Output Degraded-encountered	0b0	Output port has encountered a degraded condition, meaning that the port's degraded error threshold has been reached in the Port *n* Error Rate Threshold register. Once set remains set until written with a logic 1 to clear.

"

to:

"

Table 1-2. Bit Settings for Port *n* Error and Status CSRs

Bit	Name	Reset Value	Description
5	Output Packet-dropped	0b0	Output port has discarded a packet. Once set remains set until written with a logic 1 to clear.
6	Output Failed-encountered	0b0	Output port has encountered a failed condition, meaning that the port's failed error threshold has been reached in the Port *n* Error Rate Threshold register. Once set remains set until written with a logic 1 to clear. Receipt of a reset-port request shall clear this bit to 0. State machines associated with this bit shall be reset to their power-up state.
7	Output Degraded-encountered	0b0	Output port has encountered a degraded condition, meaning that the port's degraded error threshold has been reached in the Port *n* Error Rate Threshold register. Once set remains set until written with a logic 1 to clear.

1.5.2 Changes to Section 2.3.2.17 Port n Error Rate CSR

The description of the register fields changes from:

"

Table 1-3. Bit Settings for Port *n* Error Rate CSR

Bit	Name	Reset Value	Description
0-7	Error Rate Bias	0x80	This field specifies the rate at which the Error Rate Counter is decremented (the error rate bias value) 0x00 - do not decrement the error rate counter 0x01 - decrement every 1ms (+/-34%) 0x02 - decrement every 10ms (+/-34%) 0x04 - decrement every 100ms (+/-34%) 0x08 - decrement every 1s (+/-34%) 0x10 - decrement every 10s (+/-34%) 0x20 - decrement every 100s (+/-34%) 0x40 - decrement every 1000s (+/-34%) 0x80 - decrement every 10000s (+/-34%) other values are reserved
8-13	—		Reserved
14-15	Error Rate Recovery	0b00	The value of this field limits the incrementing of the Error Rate Counter above the failed threshold trigger. 0b00 - only count 2 errors above 0b01 - only count 4 errors above 0b10 - only count 16 error above 0b11 - do not limit incrementing the error rate count
16-23	Peak Error Rate	0x00	This field contains the peak value attained by the error rate counter since the field was last reset.
24-31	Error Rate Counter	0x00	This field contains a count of the number of physical layer errors that have been detected by the port, decremented by the Error Rate Bias mechanism, to create an indication of the physical layer error rate.

“

to:

“

Table 1-4. Bit Settings for Port *n* Error Rate CSR

Bit	Name	Reset Value	Description
0-7	Error Rate Bias	0x80	This field specifies the rate at which the Error Rate Counter is decremented (the error rate bias value) 0x00 - do not decrement the error rate counter 0x01 - decrement every 1ms (+/-34%) 0x02 - decrement every 10ms (+/-34%) 0x04 - decrement every 100ms (+/-34%) 0x08 - decrement every 1s (+/-34%) 0x10 - decrement every 10s (+/-34%) 0x20 - decrement every 100s (+/-34%) 0x40 - decrement every 1000s (+/-34%) 0x80 - decrement every 10000s (+/-34%) other values are reserved
8-13	—		Reserved

Table 1-4. Bit Settings for Port *n* Error Rate CSR

Bit	Name	Reset Value	Description
14-15	Error Rate Recovery	0b00	The value of this field limits the incrementing of the Error Rate Counter above the failed threshold trigger. 0b00 - only count 2 errors above 0b01 - only count 4 errors above 0b10 - only count 16 error above 0b11 - do not limit incrementing the error rate count
16-23	Peak Error Rate	0x00	This field contains the peak value attained by the error rate counter since the field was last reset.
24-31	Error Rate Counter	0x00	This field contains a count of the number of physical layer errors that have been detected by the port, decremented by the Error Rate Bias mechanism, to create an indication of the physical layer error rate. Receipt of a reset-port request shall clear this field to 0x00. State machines associated with this field shall be reset to their power-up state.