VITESSE

Ensuring Delivery of Broadcasts In Zoned Expanders

T10/05-258r0 SAS 2.0
8 July 2005

YOUR PARTNER FOR SUCCESS

- From:
  - Ed D'Avignon (davignon@vitesse.com)
  - Phillip Roberts (robertsp@vitesse.com)
- Revision History
  - Revision0 (8 July 2005) first revision
- Related Documents
  - sas1r09 - Serial Attached SCSI 1.1 revision 9
  - T10/05-144r2 SAS-2 Zoning

Problem Statement

VITESSE

Loss of BROADCAST Frame

- SAS 1.1
  - BROADCAST Primitive is Redundant
    - 6 Primitives Transmitted, 3 Consecutive Primitives must be Received
    - Requires 2 bit Errors in 240 Bits to Lose Broadcast Primitive
- SAS 2.0 Zoning Proposal
  - BROADCAST Frame with No Handshake or Acknowledgement
    - 10 d-words (400 bits) Transmitted
    - Single Error in 400 Bits to Lose Broadcast Frame

Probability of Losing Broadcast

VITESSE

- Probability of *e* errors in *n* bits is:
  - $p^e(1-p)^{n-e}$
- Broadcast Frame (with probability of error 10^{-12})
  - $\left(10^{-12}\right)^1\left(1-10^{-12}\right)^{400-1} = 9\,9999999601\times10^{-12}$
  - Error Probability of 10^{-12} at 3Gb is 1 Bit Error Every 55 Minutes
- Broadcast Primitive (with probability of error 10^{-12})
  - $\left(10^{-12}\right)^2\left(1-10^{-12}\right)^{240-2} = 9\,9999999762\times10^{-23}$
  - Error Probability of 10^{-23} at 3Gb is 1 Bit Error Every 10 Million Years

Proposed Solution

VITESSE

Loss of BROADCAST Frame

- Create New Primitive BROADCAST_RECEIVED
  - Sent From Zoning Expander to Zoning Expander ONLY
  - Sent when Valid Broadcast Frame Received
  - Must Be Received from Each PHY Broadcast Frame is Transmitted On
  - If not Received within 1ms of Transmitting Broadcast Frame, Retransmit the Frame to Ports that did not respond
    - 1ms Timeout can be reduced since BROADCAST Frames are sent between expanders that are directly connected
- Risk of Loss of Broadcast Frame Eliminated
- Efficient Use of Hardware
  - Can Share Some HW used for Open Address Frame Processing

Problem Statement

VITESSE

BROADCAST Frame Can Overrun Hardware

- SAS 1.1
  - 8 BROADCAST Primitives (6 Reserved)
  - BROADCAST Primitive Sent to Every Port
    - Bit vector can be used to indicate Ports Needing BROADCAST
      - 288 (8*36) Bits for a 36 PHY Expander
  - Multiple BROADCAST primitives of same type arriving in same cycle generate a SINGLE BROADCAST primitive to be sent on each non-receiving PHY
  - If a new BROADCAST primitive of the same type is received before the existing primitive has been transmitted to all ports, the bit vector is reset and the new broadcast sent to all ports.
    - Broadcasts of the same type do not stack up

Problem Statement

VITESSE

BROADCAST Frame Can Overrun Hardware

- SAS 2.0 Zoning Proposal
  - 8 BROADCAST Primitives times 128 Groups -> 1024 Unique BROADCAST Frames
  - BROADCAST Frames to Every Port that Communicates with Source Group
    - One PHY Bit Vector Needed Per Group
      - 36 864 (8 * 128 * 36) bits required for 36 PHY Expander
  - MULTIPLE Broadcasts from DIFFERENT groups arriving in the SAME cycle generate one Broadcast Frame for each Broadcast received
    - 36 unique broadcasts could be received in a single cycle on a 36 PHY expander
      - Causes up to 12600 dwords to be transmitted
    - Possible to generate more traffic than sending BROADCAST Primitives to all groups

Proposed Solution

VITESSE

BROADCAST Frame Can Overrun Hardware

- BROADCAST Frame Queue
  - Used to Hold Outgoing Broadcast Frames
  - Only Contains Broadcasts from Groups other than 127
    - Broadcasts from Group 127 are Transmitted as Primitives
    - Broadcasts of Same Type from Other Groups Removed when Broadcast from Group 127 is Received
  - When Queue Fills
    - All Broadcast Frames are translated to Broadcast Primitives
    - Queue is Emptied
- Efficient Use of Hardware
  - Queue can be sized for Typical Case not Theoretical Worst Case
  - No Broadcast Event is Lost due to Lack of Queue Space
    - Lack of Queue Space Only Causes Broadcast Primitive to be Sent