



Infiniband VL Arbitration Overview Pat Thaler

Data structures

- Two tables
 - Two elements per row: VL and weight
 - Weight is in 64 byte units
 - High priority table
 - at least one row and up to 64
 - Low priority table
 - at least as many rows as supported VLs and up to 64
- Limit of high priority value
 - Limits max duration of high priority transmission
 - 8 bit value
 - 255 means no limit

Arbitration between tables

- **HighPriCounter**
 - Decrement for size of high priority packet sent
 - Expires when negative
 - Reset when no high priority VLS have packets to transmit or when expired
 - When expires, next packet sent from low priority
 - Loaded with 4K bytes * limit of high priority value
- **At completion of sending a packet, choose table for next packet**
 - Use high priority table if
 - At least one packet for a high priority table entry is available for transmission, and,
 - HighPriCounter has not expired
 - Otherwise use low priority table

Arbitration within a table

- Table has a pointer to the current row and available weight
- Transmit from current row VL if
 - Packet available to transmit (including flow control credit available) and
 - Available weight is positive
- Otherwise transmit from next row with available packet and load weight into available weight
- Reduce available weight by packet length
- Note that available weight is not checked vs packet length.
 - Weighting is approximate
 - Could maintain deficit in row for deficit round robin