Priority-based Flow Control: Draft PAR & 5 Criteria

Joe Pelissier, Brocade
Mike Ko, IBM
Pat Thaler, Broadcom
Manoj Wadekar, Intel

Title (2.1)

- Amendment to 802.1Q

- Standard for Local and Metropolitan Area Networks: Virtual Bridged Local Area Networks – Amendment 9: Priority-based Flow Control.

PAR Scope (5.2)

- This standard specifies protocols, procedures and managed objects that support operation of per priority (VLAN tag encoded priority values) flow control within congestion-managed short range networks. This mechanism is intended to eliminate loss due to congestion.

- This mechanism enhances support for and co-existence of higher layer protocols that are highly loss sensitive. This is achieved by a pause mechanism similar to the Ethernet PAUSE operating on individual priority levels in the Forwarding Process allowing each priority level to operate independently, including whether or not the flow control mechanism is enabled.

PAR Scope (5.3)

- Is the completion of this document contingent upon the completion of another document? No

PAR Purpose (5.4)

- Data center networks employ higher level protocols that depend on the delivery of data packets with a much lower probability of packet loss than is typical of IEEE 802 VLAN bridged networks. This amendment will support the use of these applications as well as their co-existence with traditional LAN applications on a single VLAN bridged network.

Need for the project (5.5)

- There is significant customer interest and market opportunity for Ethernet as a consolidated Layer 2 solution in high-speed short-range networks such as data centers, backplane fabrics, single and multi-chassis interconnects, computing clusters, and storage networks. These applications currently use Layer 2 networks that offer very low frame loss (e.g. Fibre Channel, Infiniband). Use of a consolidated network will realize operational and equipment cost benefits. This project brings to Ethernet the lossless characteristics comparable to the networks that currently host these applications. By applying flow control based on VLAN tag encoded priority values, this project allows a consolidated network to have priorities with current Ethernet characteristics along with lossless flow-controlled priorities.

Stakeholders for the Standard (5.6)

- Developers and users of networking for data center environments including networking IC developers, switch and NIC vendors, and users.

5C

Broad Market Potential

a) Broad sets of applicability

❑ Mechanisms to avoid frame loss are essential for support of the highly loss sensitive higher layer protocols, which are prevalent in the important applications E.g. data storage, clustering, backplane fabrics. Back-end data storage networks, clustering networks and backplane fabrics are typically limited in size, making them amenable to a flow control mechanism that operates over a limited number of hops.

❑ The data traffic to be controlled by the proposed congestion notification mechanism will be segregated using the Ethernet priority mechanism, thus ensuring that traffic types that are not amenable to flow control may efficiently co-exist with those that are, thus enabling network consolidation.

b) Multiple vendors and numerous users

❑ Multiple equipment vendors have expressed interest in the proposed project. There is strong and continued user interest in converting existing networks to Ethernet and in the realization of operational and equipment cost savings through use of a consolidated network. Further there is strong interest in increased use of data storage networks, provided that they can be realized with familiar technology and a consolidated network. In addition, INCITS T11 has begun a standardization project (FC-BB-5), the usability of which will be greatly enhanced by this project.

c) Balanced costs (LAN versus attached stations)

❑ The introduction of this flow control mechanism is not expected to materially alter the balance of costs between end stations and bridges. Significant equipment and operational costs savings are expected as compared to the use of separate networks for traditional LAN connectivity and for loss/latency sensitive applications.

Compatibility

- The proposed standard will be an amendment to 802.1Q, and will interoperate and coexist with all prior revisions and amendments of the 802.1Q standard.

- The data traffic to be controlled by the proposed flow control mechanism will be segregated using a priority-based technique, thus ensuring that traffic types already supported by VLAN Bridges are not affected.

- The proposed amendment will contain MIBs, or additions to existing MIBs, to provide management operations for any configuration required together with performance monitoring for both end stations and bridges.

Distinct Identity

a) Substantially different from other IEEE 802 standards.

IEEE Std 802.1Q is the sole and authoritative specification for priority aware Bridges and their participation in LAN protocols. No other IEEE 802 standard addresses priority based flow control by bridges.

b) One unique solution per problem (not two solutions to a problem)

The need to subject certain classes of traffic to flow control mechanisms while allowing others to operate without has not been anticipated by any other IEEE802 specification; consequently, this proposal is the only solution to the problem of allowing a coexistence of such traffic types.

c) Easy for the document reader to select the relevant specification.

IEEE Std 802.1Q is the natural reference for priority based handling of traffic flows, which will make the capabilities added by this amendment easy to locate.

Technical Feasibility

a) Demonstrated system feasibility.

Similar techniques are widely deployed in other networking technologies, such as Fibre Channel and InfiniBand. The proposal is a natural extension of the expedited forwarding capability defined in IEEE Std. 802.1Q and widely deployed in bridge products.

b) Proven technology, reasonable testing.

The technique of this proposal has been proven in real world deployments of Fibre Channel, InfiniBand, and other networking technologies. These techniques have been shown to be reasonably testable.

c) Confidence in reliability.

The techniques of the proposal have been proven reliable in real-world deployments of Fibre Channel, InfiniBand, and other networking technologies.

d) Coexistence of 802 wireless standards specifying devices for unlicensed operation.

Not applicable.

Economic Feasibility

a) Known cost factors, reliable data.

The proposed amendment will retain existing cost characteristics of bridges including simplicity of queue structures and will not require maintenance of additional queues or queue state beyond the existing per traffic class (priority) queues for conformance to either its mandatory or optional provisions. In particular per flow queuing will not be required.

The proposed amendment may require some functions, specifically the generation of per-priority flow control frames, at a rate and within a time not practical for some existing and otherwise conformant bridge implementation architectures. However these functions can be performed by some existing bridges with known implementation costs.

b) Reasonable cost for performance.

The proposed technology will reduce overall costs where separate networks are currently required by enabling the use of consolidated network. The proposed solution allows network to avoid packet loss without significant throughput reduction.

c) Consideration of installation costs.

Installation costs of VLAN Bridges or end stations are not expected to be significantly affected; any increase in network costs is expected to be more than offset by a reduction in the number of separate networks required.