

Proposal to improve expedited forwarding (...continuation from May interim meeting...)

Paul Congdon
IEEE 802 Plenary – San Francisco
July 20, 2005

Last Time We...

- Reviewed current 802.1 expedited forwarding capabilities
- Reviewed current 802.1 metering specifications
- Reviewed SBM and DiffServ models
- Discussed scheduling algorithms to provide guarantees and avoid starvation
 - Rate Controlled Priority Queuing
 - Deficit Weighted Round Robin
- Discussed possible improvements to 802.1 specifications
 - Minimum bandwidth guarantees for Egress classes
 - Maximum bandwidth limits for Egress classes
 - Configuration of Ingress meters
- Made the following recommendations...

Recommendations

(from last time...)

Time to do something

Need to keep it simple (don't define algorithm variables, just results)

Work in support of current internet models

Acknowledge shipping solutions and avoid obsolescence

Evaluate scheduling algorithm work done in MEF, IETF, DSL Forum TR-059, ITU

Consider the following three modifications to clause 8.6.7 and 8.6.8:

1. Minimum bandwidth guarantees for Egress classes
2. Maximum bandwidth limits for Egress classes
3. Configuration of Ingress meters

This time we will...



- Consider a very basic proposal for Minimum bandwidth guarantees for Egress classes
- Provide example configurations that represent existing scheduling algorithms in products today
- Discuss the changes to 802.1Q that would be required to support the proposal

Proposal Description



Objective

1. Provide a simple mechanism to provide minimum bandwidth guarantees by egress queue scheduling
2. Provide flexibility in a single mechanism to support multiple scheduling algorithms
3. Support Rate Controlled Priority Queuing
4. Support Weighted Round Robin Schemes

Mechanism

1. Define 8 parameters to select bandwidth percentages per traffic class
2. Define 8 parameters to select a round robin scheduling groups priority

Definitions

Traffic Class

The 8 traffic classes defined by 802

Weight

A value representing the minimum guaranteed bandwidth for a traffic class. Potentially a percentage of bandwidth (0-100) or an absolute value in Kbps or Mbps (Note: implementations may need to adjust internally)

Priority Group

Frames from traffic classes with a higher priority group are scheduled before frames in a lower priority group only if the amount of traffic in the higher priority group remains below the guaranteed minimum. Frames from multiple traffic classes in the same priority group are scheduled in a round robin fashion within the group.

Traffic Class	Weight	Priority Group
0	0	0
1	0	1
2	0	2
3	0	3
4	0	4
5	0	5
6	0	6
7	0	7

Default: Strict Priority

Example Scenarios

Traffic Class	Weight	Priority Group
0	12.5	0
1	12.5	0
2	12.5	0
3	12.5	0
4	12.5	0
5	12.5	0
6	12.5	0
7	12.5	0

Deficit Weighted Round Robin

Traffic Class	Weight	Priority Group
0	5	0
1	50	1
2	5	2
3	5	3
4	5	4
5	5	5
6	5	6
7	20	7

Rate Controlled Priority Queuing

More Scenarios

Traffic Class	Weight	Priority Group
0	0	0
1	0	0
2	0	0
3	0	0
4	0	0
5	100	1
6	0	0
7	0	0

Modified Deficit Round Robin (MDRR)
Strict Priority Mode, PQ = Traffic Class 5

Traffic Class	Weight	Priority Group
0	5	0
1	50	1
2	5	2
3	5	2
4	5	2
5	5	2
6	5	2
7	20	3

A mixed scenario: 4 scheduling groups

Subtle notes



How minimum guarantees are implemented is not specified

- Could use a deficit scheme
- Internal counters could be packet based, octet based, rate based

Time quantum for measuring bandwidth isn't specified, but could be suggested

It should be possible to have restricted configuration capability of these parameters, or no configuration at all

- Can implement weights but without priority groups
- Priority groups of more than 1 might not be supported

Changes to support this proposal



In 802.1Q, at least the following:

1. add an additional item to 8.1.6 that explains that queue scheduling configuration supports traffic expediting
2. rewrite 8.6.8 to describe new parameters and scheduling algorithm
3. Update 12.6.3 to include parameters as managed objects
4. Update A.5, A.14 and A.16 PICs for new optional capabilities
5. Update Annex G to incorporate new scheduling algorithms
6. Consider new sub-clause in Annex G to describe use cases