IEEE 802.3 Flow Control Baseline Requirements

IEEE 802.3 CMSG

July 12, 2004

Siamack Ayandeh
Independent
Siamack@ieee.org

Khaled Amer
Amernet, Inc.
Amer@amernet.net

Ground rules

- This is not a debate about open vs. closed loop congestion control
  - Open loop: relies on packet loss, time out, and retransmissions
  - Closed loop: relies on back pressure
- One or the other or both schemes can be operating on a flow on its e2e journey
- I think we all agree that Pause (802.3x) has limitations
  - That's why we are developing improvements

802.3x Pause FC

Scope: Single link

- 802.1 specification does not advise forwarding of Pause from one port to another in a bridge

Granularity: Per link (i.e. has no granularity)

- A single control loop applies to all traffic on the link
- Head of line blocking is an issue

Control Algorithm: is ON/OFF

- It signals an interval for which transmission should be suspended if XON is not received

So how can we improve on the scope, granularity, and the control algorithm?

Scope of Single Link

Two applications come immediately to mind:

1. Inter chip communications

   where e.g. GE is used in a similar fashion to SPI 3/4

   – Granular flow control would allow buffering and active queue management to be implemented at one place only

Scope of Single Link …

2. Local loop for Metro Ethernet service access (CPE to Service Provider edge)
   - End user often subscribes to a few megabits of long haul bandwidth while using a FE to interface the SP
   - SP would police the traffic based on an SLA
   - Granular flow control would prevent non compliant traffic bursts from CPE
   - Enabler for Metro Ethernet Services

Granularity of Flow Control

- Flow control mechanism would simply offer n x flow control signals
- Meaning can be attached to the grains as a matter of application
- Each signal can be mapped to CoS, VLAN tag, or any other classification criteria locally
- What upper bound should be placed on “n” is a matter of protocol design
  - Not all implementations need to support up to this upper bound
  - Hence the door is open for vendor differentiation

Control Algorithm Improvements

- Pause signals the number of 512 bit intervals for which transmission should be suspended
  - Transmission resumes at full line rate
- Many bandwidth increase/decrease algorithms may be devised with possible gains in throughput & delay
- Any gains however, have to be weighed against the increasing complexity of operation and implementation

802.3 Flow Control Baseline

- Has to work for single link (& not interfere with existing 802.3x)
- Improve granularity to n x control signals
  - Not all implementations need to support the maximum value of “n”
- Work on possible improvements to the on/off control algorithm if any

Feasibility

- The SG should make a list of proprietary implementations

Benefits or Value

- Inter-chip communication
  - Single point of buffering and active queue management reduces silicon cost
- Metro Ethernet Services
  - Control disparity between access and long haul transmission rates
  - Prevent non compliant traffic
- Data Center applications

  (already presented elsewhere)

Discussions