Impact of expanding maximum frame size to overhead, aggregation

Kevin Daines
World Wide Packets

30 September 2004

Outline

- Review overhead calculations
  - Current frame sizes
  - Expanded frame vs. reduced data field
- Overhead summary

- Review constriction examples
- Aggregation summary

Frame/data rates

- Calculations below are for Gigabit Ethernet:

		Frame size	Data size	Overhead[1]	Frame/s	Data rate (octet/s)
Existing	Min	64	46	45.2%	1,488,095	68,452,370
	Max	1,518	1,500	2.5%	81,274	121,911,000
Increase max frame	Add 357[2]	1,875	1,500	20.8%	65,963	98,944,500
	Add 530[2]	2,048	1,500	27.5%	60,444	90,666,000
Decrease max data field	Data=1143	1,518	1,143	25.7%	81,274	92,896,182
	Data=1024	1,518	1,024	32.1%	81,274	83,224,576

[1]Overhead equation = $\frac{\text{IPG + PRE + SFD + DA + SA + L/T + FCS}}{\text{IPG + PRE + SFD + DA + SA + L/T + DATA + FCS}}$

[2]Sample larger frame size

Overhead Summary

- Current min frame yields worst efficiency
  - 45% (38 / 84)
- Current max frame provides best efficiency
  - 2.5% (38 / 1538)
- Increasing max frame size hurts efficiency
  - 21-28%
- …but better than shrinking max data field
  - 26-32%

- Conclusion: increase frame size to allow for new optional envelope fields

Outline

- Review overhead calculations
  - Current frame sizes
  - Expanded frame vs. reduced data field
- Overhead summary

- **Review constriction examples**
- **Aggregation summary**

Aggregation

- Ethernet's traditional 10x speed improvement provides clean aggregation
  - (10) 100 Mb/s ports can be easily aggregated into (1) Gig port
  - Caveat:
    - Assumes 100 Mb/s ports and the Gig port are either all untagged or all tagged
  - If each 100 Mb/s port is untagged and the Gig port is tagged, congestion in the form of *constriction* (term borrowed from Hugh Barrass/ July 04 CMSG) occurs

Tagged constriction

- (10) 100 Mb ⇒ (1) tagged Gig
  - Minimum sized frames
    - 10 * 148,809 ⇒ 1,420,454 *(adds 4B)*
  - Difference of 67,641 min frames

- Solutions
  - Count on sustained rate being less than 95.5%; buffer some line rate bursts; discard excess frames (widely used)
  - Source is throttled back to “tagged” frame rate using PAUSE or unspecified egress rate control mechanisms

Add 2nd tag constriction

- (10) tagged 100 Mb ⇒ (1) double tagged Gig
  - Minimum sized fps:
    - 10 * 142,045 ⇒ 1,358,695 *(adds 4B)*
  - Difference of 61,759 min tagged frames

- Solutions
  - Count on sustained rate being less than 95.7%; buffer some line rate bursts; discard excess frames (widely used)
  - Source is throttled back to "double tagged" frame rate using PAUSE or unspecified egress rate control mechanisms

Add two tags constriction

- (10) untagged 100 Mb ⇒ (1) double tagged Gig
  - Minimum sized fps:
    - 10 * 148,809 ⇒ 1,358,695 *(adds 8B)*
  - Difference of 129,400 min frames

- Solutions
  - Count on sustained rate being less than 91.3%; buffer some line rate bursts; discard excess frames (widely used)
  - Source is throttled back to “double tagged” frame rate using PAUSE or unspecified egress rate control mechanisms

10G WAN constriction

- (1) 10 Gig LAN ⇒ (1) 10 Gig WAN
  - Minimum sized fps:
    - 14,880,952 ⇒ 14,262,857
  - Difference of 618,095 min frames

- Solutions
  - Count on sustained rate being less than 95.8%; buffer some line rate bursts; discard excess frames (widely used)
  - Source is throttled back to "WAN" frame rate using PAUSE or unspecified egress rate control mechanisms

MACSec constriction

- (1) Gig ⇒ (1) MACSec Gig
  - Minimum sized fps:
    - 1,488,095 ⇒ 822,368 *(adds 4B SecTag, 64B ICV)*
  - Difference of 665,727 min frames

- Solutions
  - Count on sustained rate being less than 55.3%; buffer some line rate bursts; discard excess frames
  - Source is throttled back to “MACSec” frame rate using PAUSE or unspecified egress rate control mechanisms

PWE3 constriction

- (1) Gig ⇒ (1) PWE3 Gig
  - Minimum sized fps:
    - 1,488,095 ⇒ 1,096,491 *(adds 30B)*
  - Difference of 391,604 min frames

- Solutions
  - Count on sustained rate being less than 73.7%; buffer some line rate bursts; discard excess frames (widely used)
  - Source is throttled back to "PWE3" frame rate using PAUSE or unspecified egress rate control mechanisms

Clock tolerance constriction

- (1) +100ppm Gig ⇒ (1) –100ppm Gig
  - Minimum sized fps: *(1,488,095 nominal)*
    - 1,488,244 ⇒ 1,487,946
  - Difference of 298 min frames

- Solutions
  - Count on sustained rate being less than 99.98%; buffer some line rate bursts; discard excess frames (widely used)
  - Source is throttled back to "-100ppm" frame rate using PAUSE or unspecified egress rate control mechanisms

802.3 FESG constriction

- (1) legacy Gig ⇒ (1) FESG Gig
  - Minimum sized fps:
    - 1,488,095 ⇒ 203,583 (*assumes 530B additional header*)
  - Difference of 1,284,512 min frames

- Solutions
  - Count on sustained rate being less than 13.7%; buffer some line rate bursts; discard excess frames (*Study Group phase*)
  - Source is throttled back to "FESG" frame rate using PAUSE or unspecified egress rate control mechanisms

Aggregation Summary

- Ethernet's traditional 10x speed improvement provides clean aggregation
  - Caveat: Assuming aggregated ports are configured identically to aggregation port and aggregation link clock running at –100ppm
- Many applications exist where aggregation port adds field(s) to frame
- Expanding the frame size only worsens an existing problem
  - *It does not create a new problem*