Rapid PHY Selection (RPS):
Emulation and Experiments using PAUSE

Ken Christensen
Department of Computer Science and Engineering
University of South Florida
Tampa, FL 33620
christen@cse.usf.edu
(813) 974-4761

This material is based upon work funded by
the National Science Foundation under grant
CNS-0520081.

Acknowledgments

• Most of this work was done by Francisco Blanquicet
  (graduate student at USF)
Motivation

- Question…

What are the effects of a link disruption as may be caused by RPS on higher-layer protocols and applications?

Link disruption by RPS

- RPS mechanism could be a MAC frame handshake

No packets can be sent in either direction during this time

The link is down (or blocked)

There are no in-flight packets to be lost
Link disruption by RPS continued

- Link disruption may cause packet loss
  - Packets lost due to buffer overflow in switch

![Diagram showing link disruption by RPS]

Worst case packet loss

- Worst case is a burst of packets at full data rate
  - And, an RPS rate switch during the burst

1 millisecond of 1 Gb/s = 122 KBytes
1 millisecond of 10 Gb/s = 1.2 MBytes

What is the probability of this occurring?

What is impact of this if/when it occurs?
Packet loss from RPS

- **Calculating packet loss**
  - Assume RPS occurs during a burst

- **Inputs are**
  - $R = $ Burst rate (bits/sec)
  - $B = $ Burst size (bits)
  - $S = $ Switch buffer size (bits)
  - $T = $ RPS switching time (sec)

$$Loss = R \cdot \min\left(\frac{B}{R}, T\right) - S$$

- If negative result, no loss
- Divide by mean packet length to get (roughly) number of packets lost

Packet loss graph (theory)

- **For a 10 Gb/s link**
  - Assume $T = 1$ millisec and $B = 10$ Mbytes
  - Assume $R = 2000, 4000,$ and $8000$ Mb/s
Preventing packet loss with PAUSE

- Loss can be prevented/minimized with PAUSE
  - Switch PAUSEs server during RPS switching time

Did some experiments at 1 Gb/s

- Idea: Use PAUSE to *emulate* RPS switching time

- PAUSE can emulate RPS to some extent
  - Blocks traffic in one direction
  - Can select duration of blocking (to emulate RPS switching time)
  - Note that link returns to same data rate as before
**Experiment set-up**

- To study packet loss and effects on applications
  - Used a rawsend program to send PAUSE frames to emulate RPS
  - TCP and UDP server throughput is about 350 Mb/s
  - UDP streamer sends packets at 144 kb/s (emulates Skype)

- Note: Sends PAUSE frames

**Experiments**

- **Experiment #1 – TCP file downloads**
  - Download files while RPSing

- **Experiment #2 – UDP bulk data transfer**
  - Download while RPSing

- **Experiment #3 – TCP download + UDP stream**
  - Download + streaming while RPSing

Emulated RPS once per second

Emulated RPS switch times were 1, 10, and 20 milliseconds
Experiment observations

- TCP downloads are always successful
  - Download time increase is 2x to 5x total RPS switching time

- UDP bulk data packet loss is as expected
  - Packet loss proportional to total RPS switching time per second

- UDP streaming packet loss
  - No packet loss detected
    - Due to low bit rate resulting in very low probability of packet loss

Experiment observations continued

- For all of the experiments...
  - If switch sends PAUSE to server then no packet loss occurs
Conclusions

- Do not want to introduce something into network that causes packet loss
- PAUSE flow control can automatically prevent (or at least reduce) packet loss
- Not clear that packet loss will be “bad” for low utilization links
- Can think of packet loss as a trade-off
  - Energy saved versus packets lost
- In any case, RPS is not intended for all links

Questions?

Ken Christensen
Department of Computer Science and Engineering
University of South Florida
Tampa, FL 33620
christen@cse.usf.edu
(813) 974-4761