Whither Congestion Control?

Sally Floyd

E2ERG, July 2006

www.icir.org/floyd/talks

Topics:

- Congestion control:
  - Router algorithms for detecting congestion;
  - Transport protocol responses to congestion:
    - Unicast and multicast.
  - Detecting misbehaving nodes or aggregates;
  - Difficult issues for unreliable transport.
- Explicit communications with routers:
  - For congestion control (e.g., XCP)?
  - For anti-congestion control (e.g., Quick-Start)?
  - For communicating with layer two (e.g., corruption)?
- The role of the IETF?
- Models for evaluating congestion control.

Issues I am not talking about:

- Transport:
  - E.g., HighSpeed TCP, BIC/CUBIC, HTCP, STCP, FAST TCP, etc.
- Router Mechanisms:
  - For congestion notification using packet drops or ECN.
  - E.g., RED, REM, Blue, etc.
- Misbehaving nodes or aggregates:
  - E.g., RED-PD, ACC, etc.

Difficult Issues for Unreliable Transport (e.g., DCCP):

- Applications that send frequent small packets:
  - Network bottleneck in bytes per second or packets per second?
  - Routers treat small and large packets the same, or not?
  - Would recommendations to router designers be useful?
- Applications that want to more than double their sending rate from one RTT to the next (video).
- Applications that want to start up fast after an idle period (audio).

Forms of Explicit Communication:

- QoS-related.
- New congestion control mechanisms based on explicit feedback from routers (e.g., XCP).
- "Anti-congestion control" mechanisms based on explicit feedback from routers (e.g., Quick-Start).
- Explicit communication including layer two:
  - Packet corruption;
  - Path changes;
  - Link changes;
  - Interactions with layer-two congestion control?
  - Etc.

Forms of Explicit Communication:

- How to proceed?
  - Top-down, exploring the space, and also bottom-up, exploring specific mechanisms.
  - Keeping the long time horizon in mind, and also exploring real-world obstacles.
  - Exploring positives and negatives.
- E.g., for communication involving layer two:
  - Whole space, and scecific mechanisms both.
  - Thinking about both future and current layer-two mechanisms.
  - Communication to and from layer two.
  - Communication involving the whole path, or a single link.

Problems with explicit communication with routers (from Quick-Start):

- Attacks from others (e.g., SYN floods).
- Misbehaving senders or receivers.
- Real-world problems:
  - Problems with middleboxes:
    - Packets with IP options dropped.
    - Packets dropped or "normalized", etc.
  - IP tunnels, MPLS, etc.
  - Switches in layer-two networks.
  - Router incentives to play.
  - And more…

The Future of the IETF and Congestion Control?

- Or instead, let a hundred flowers bloom?
  - Linux.
  - Microsoft.
  - Etc.

Research Internet Needs Better Models.

- We need better models to use in simulations, experiments, and in analysis for evaluating congestion control mechanisms.
- Typical scenarios should include:
  - two-way traffic, and
  - a range of round-trip times, and
  - a range of connection sizes, and
  - a range of receive windows, and
  - a range of access link bandwidths.
  - And maybe a range of applications, including audio and video with variable bandwidth demands.

Extra viewgraphs:

Attacks on Quick-Start:

- Attacks to increase router's processing load:
  - Easy to protect against -

    routers ignore Quick-Start when overloaded.

- Attacks with bogus Quick-Start requests:
  - Similar to Quick-Start requests denied downstream.
  - Harder to protect against.
  - It doesn't cost a sender anything to send a bogus Quick-Start request.

The Problem of Cheating Receivers: the QS Nonce.

- Initialized by sender to a random value.
- If router reduces Rate Request from K to K-1, router resets related bits in QS Nonce to a new random value.
- Receiver reports QS Nonce back to sender.
- If Rate Request was not reduced in the network below K, then the lower 2K bits should have their original random value.

- Do receivers have an incentive to cheat?

Protection against Cheating Senders:

- The sender sends a "Report of Approved Rate" after receiving a Quick-Start Response. The Report might report an Approved Rate of zero.
- Routers may:
  - Ignore the Report of Approved Rate;
  - Use Report to check for misbehaving senders;
  - Use Report to keep track of committed Quick-Start bandwidth.

- Do senders have an incentive to cheat?

Real World Problems: Misbehaving Middleboxes:

- There are many paths where TCP packets with known or unknown IP options are dropped.
  - **Measuring Interactions Between Transport Protocols and Middleboxes**, Alberto Medina, Mark Allman, and Sally Floyd. Internet Measurement Conference 2004, August 2004.
  - For roughly one-third of the web servers, no connection is established when the TCP client includes an IP Record Route or Timestamp option in the TCP SYN packet.
  - For most web servers, no connection is established when the TCP client includes an unknown IP Option.

Real-World Problems: IP Tunnels.

- IP Tunnels (e.g., IPsec) are used to give a virtual point-to-point connection for two routers.
- There are some IP tunnels that are not compatible with Quick-Start:
  - This refers to tunnels where the IP TTL is not decremented before encapsulation;
  - Therefore, the TTL Diff is not changed;
  - The sender can falsely believe that the routers in the tunnel approved the Quick-Start request.
  - This will limit the possible deployment scenarios for Quick-Start.

Real-World Problems: Layer-2 Networks

- Multi-access links, layer-2 switches:
  - E.g., switched Ethernet.
  - Are the segments underutilized?
  - Are other nodes on the layer-2 network also granting Quick-Start requests?