Quick-Start for TCP and IP

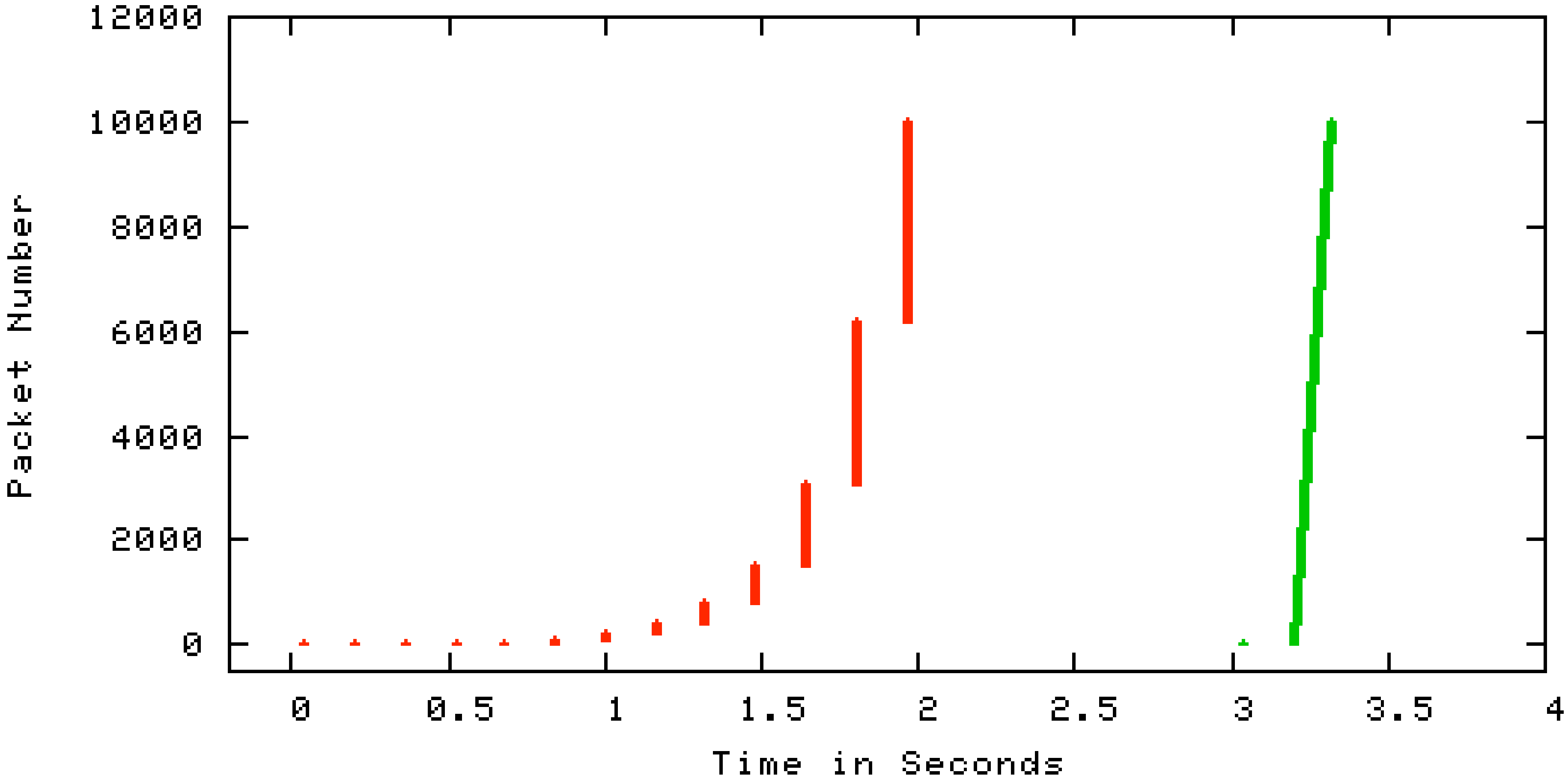
A. Jain, S. Floyd, M. Allman, and
P. Sarolahti
ICSI, April 2006

This and earlier presentations::
www.icir.org/floyd/talks

Congestion control and anti-congestion control:

- Much of my work has been on congestion control:
  - Router algorithms for detecting congestion;
  - Transport protocol responses to congestion:
    - Unicast, multicast
    - TCP, TCP-friendly
  - Detecting misbehaving nodes or aggregates;
  - Network models for evaluating congestion control;
  - Measurement studies of congestion control in the net.
- But Quick-Start is about anti-congestion control.

Slow-Start and Quick-Start in TCP:



QuickStart with TCP, for setting the initial window:

- In an IP option in the TCP SYN packet,

  the sender's desired sending rate:
  - Routers on the path decrement a TTL counter,
  - and decrease the allowed sending rate, if necessary.
- The TCP receiver sends feedback to the sender in the SYN/ACK packet:
  - The TCP sender knows if all routers on the path participated.
  - The sender has an RTT measurement.
  - The sender can set the initial congestion window.
  - The TCP sender continues using normal congestion control..
- From an initial proposal by Amit Jain

Deploying Mechanisms for Explicit Communication between Routers and End-Nodes is Not Easy:

- The only current mechanism is ECN (Explicit Congestion Notification):
  - A paper in 1994.
  - Experimental Standard in 1999.
  - Proposed Standard in 2001.
  - Minimal deployment so far.

Issues with Quick-Start:

- Other approaches to faster startups.
- Impact of Quick-Start on competing traffic.
- Sender algorithms for sizing requests.
- Router algorithms for processing requests.
- Attacks on Quick-Start.
- Misbehaving senders or receivers.
- Real-world problems:
  - Packets with IP options dropped.
  - IP tunnels, MPLS.
  - Switches in layer-two networks.
  - Router incentives to use Quick-Start

Other Approaches to Faster Start-ups:

- Reservations
  - and other Quality-of-Service mechanisms.
- Information from previous connections.
- Faster start-up without modifying routers:
  - Packet-pair and extensions.
- Less-than-best-effort for the initial window.
- Other forms of feedback from routers:
  - Free buffer size, available bandwidth.
- New congestion control mechanisms.
  - E.g., XCP, AntiECN.

Sender Algorithms for Sizing Requests:

- The sender doesn’t necessarily know the amount of data to be transmitted.
- The sender knows more after an idle period.
- End-hosts might know:
  - The capacity of last-mile hop.
  - The size of the local socket buffer.
  - The receiver’s advertised window.
  - Information from the application.
  - Past history of Quick-Start requests.

Minimal Router Algorithm for Processing Requests:

- T: Configured QuickStart threshold (in Bps).
  - Requires knowledge of output link bandwidth.
- L: Current link utilization (in Bps).
  - Maximum link utilization over a recent sub-interval.
- R: Recent granted QuickStart requests (in Bps).
  - Requires state of aggregate of granted requests.
- Max request to grant: T - L - R Bps

“Extreme” Router Algorithms:

- “Extreme Quick-Start” in routers:
  - Maintains per-flow state for Quick-Start flows.
  - Estimate potential Quick-Start bandwidth more accurately.
  - Apply local policy:
    - About fairness;
    - About which Quick-Start requests to approve.
  - Check for senders that send requests that are never used.

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.
  - Extreme Quick-Start in routers can help.
  - 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?

The 30-bit QS Nonce:

Bits	Purpose
Bits 0-1:	Rate 15 -> Rate 14
Bits 2-3:	Rate 14 -> Rate 13
Bits 4-5:	Rate 13 -> Rate 12
Bits 6-7:	Rate 12 -> Rate 11
Bits 8-9:	Rate 11 -> Rate 10
Bits 10-11:	Rate 10 -> Rate 9
Bits 12-13:	Rate 9 -> Rate 8
Bits 14-15:	Rate 8 -> Rate 7
Bits 16-17:	Rate 7 -> Rate 6
Bits 18-19:	Rate 6 -> Rate 5
Bits 20-21:	Rate 5 -> Rate 4
Bits 22-23:	Rate 4 -> Rate 3
Bits 24-25:	Rate 3 -> Rate 2
Bits 26-27:	Rate 2 -> Rate 1
Bits 28-29:	Rate 1 -> Rate 0

One-way Hash Function as an Alternate QS Nonce:

- "An alternate proposal for the Quick-Start Nonce from [B05] would be for an n-bit field for the QS Nonce, with the sender generating a random nonce when it generates a Quick-Start Request. Each route that reduces the Rate Request by r would hash the QS nonce r times, using a one-way hash function such as MD5 [RFC1321] or the secure hash 1 [SHA1]. The receiver returns the QS nonce to the sender."
- "Because the sender knows the original value for the nonce, and the original rate request, the sender knows the total number of steps s that the rate has been reduced."
- From Bob Briscoe.

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?

Routers using the Report of Approved Rate:

- If Report of Approved Rate reports a higher rate than router recently approved:
  - Router could deny future requests from this sender.
- If router sees Report of Approved Rate, and didn’t see an earlier Quick-Start Request:
  - Either path changed, or sender is cheating.
  - In either case, router could deny future requests from this sender.

Routers using the Report of Approved Rate, continued:

- If router sees a Quick-Start request, but doesn't see a Report of Approved Rate:
  - The QS Request was denied and dropped downstream; OR
  - The sender didn't send a Report of Approved Rate; OR
  - The Report was dropped; OR
  - The Report took a different path in the network.
- In any of these cases, the router could deny future QS Requests from this sender.

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.
  - Is the segments underutilized?
  - Are other nodes on the layer-2 network also granting Quick-Start requests?

Possible Initial Deployment Scenarios:

- Intranets:
  - Centralized control over end nodes and routers.
  - Could include high-bandwidth, high-delay paths to remote sites.
- Paths over satellite links:
  - High bandwidth, high delay
- 2G/3G wireless networks:
  - RTTs of up to one second

Questions:

- Is something like this really needed?
- Would the benefits of Quick-Start be worth the added complexity?
- Would Quick-Start be deployable?
  - Even if only in restricted scenarios?
- What would be the relationship between Quick-Start and new router-based congestion control mechanisms (e.g., XCP)?

What else does Sally work on?

- Internet Research Needs Better Models:
  - We need to improve the models that we use in simulations, experiments, and in analysis for evaluating congestion control mechanisms.
- DCCP: a new transport protocol for unreliable transfer:
  - How do we adapt congestion control for best-effort audio traffic that sends frequent small packets?