COMP 790-088
Networked and Distributed Systems

Congestion Control

Jasleen Kaur

October 7, 2009

Network Congestion

Causes

♦ When and where does congestion occur? (And what is congestion?)
  » When outgoing link capacity is a bottleneck (e.g., access links)
  » When sum of incoming traffic exceeds outgoing capacity at large timescales
  ♦ Small timescale bursts are absorbed by queues

♦ How often does congestion occur in the Internet?
  » Don’t really know (have only anecdotal evidence)
  » “Congestion collapse” in the 80s led to design of TCP congestion-control
There are two ways to think about the Internet:

1. From the perspective of the components that make up the Internet. This involves discussing acronyms and terms you might have encountered, such as:
   - Web
   - TCP/IP
   - Inter-network
   - Routers
   - Switches
   - And more...

2. From the perspective of the services and applications that run over the Internet. This includes activities like web-browsing, email exchange, and file sharing. The key point to note is that these applications do not operate solely by themselves; they rely on various services to work together.

In this course, we will explore some of these seams, revealing how the services that support these applications are all part of the Internet. We will spend time discussing how these services function and how they interact with each other.

---

**Congestion Control**

**Conceptual Idea**

- **Why do we need congestion control?**
  - To enable sharing of common network resources by multiple data sources.

- **Goal:** Apply back-pressure to slow down senders if the network is congested.
  - Each host determines how much capacity is available in the network.
  - If the network has available bandwidth changes, adjust the number of packets in transit.

- **TCP sender maintains a new state variable: Congestion Window (cwin).**
  - Used by sender to limit how much data it is allowed to have in transit.
  - Denotes the maximum number of unacknowledged bytes.
  - MaxWin = \( \min(cwin, AdWin) \)
  - EffectiveWin = MaxWin - (LBsem - LBacked)

- **Sender not allowed to send faster than can be accommodated by slowest component (network or destination host).**

- **Challenge:** How to learn the right value for cwin?
  - Unlike destination, network does not explicitly inform sender.

- **Approach:** Set cwin based on the level of congestion perceived.
  - Decrease cwin when congestion increases.
  - Increase cwin when congestion decreases.
There are two ways to think about the Internet:

1. In terms of the components that make up the Internet. This would be a nuts-and-bolts view, where we can talk about several acronyms and terms you may have come across, such as:
   - web
   - TCP/IP
   - intermediate network
   - routers
   - switches
   - and so on...

2. In terms of the services and applications that run over the Internet. So, all of us do web-browsing, exchange emails, share files.

The important point to note is that these applications we care about do not operate solely by themselves. There are numerous services at work together to provide a seamless application.

In this class, we expose some of the seams. The services that enable such applications are all part of the Internet. So, we will spend time talking about how these services work?
There are two ways to think about it: in terms of the components that make up the Internet, and in terms of the services and applications that run over the Internet. We all do web-browsing, exchange emails, share files.

The important point to note is that these applications we care about do not operate solely by themselves. There are numerous services that work together to provide a seamless experience. In this class, we will expose some of the seams. The services that enable such applications are all part of the Internet. We will spend time talking about how these services work.

**Slow Start**

Hastening Up Initial Bandwidth Discovery

- **Two problems with cwin behavior:**
  - Initial additive ramp-up to appropriate cwin may take too long
  - Can we figure out the level of available bandwidth quickly?
  - After recovery from timeout, dumping cwin/2 may be too aggressive
    - Such a "burst" of packets may lead to further losses, even if bandwidth is high

- **Slow-start mechanism:**
  - Increase exponentially (rather than linearly), when cwin is below "SSThreshold"
  - Double the number of packets-in-transit every RTT
  - For every new ACK received:
    - If (cwin > SSThreshold)
      - increment = increment * increment/cwin
      - cwin = min (cwin + increment, AdvWin)
  - After recovery from timeout:
    - set: SSThreshold = cwin/2, cwin = 1
There are two ways to think about the Internet:

1. In terms of the components that make up the Internet. This would be a nuts-and-bolts view in which we can talk about several acronyms and terms that you may have come across, such as:
   - web
   - TCP/IP
   - inter-network
   - routers
   - switches
   - and so on...

2. The second way to think about it is the one that is more common, which is in terms of the services and applications that run over the Internet. So all of us do web-browsing, exchange emails, share files.

The important point to note is that these applications we care about do not operate solely by themselves. There are numerous services that all work together to provide a seamless experience. In this class, we expose some of the seams. The services that enable such applications are all part of the Internet, so we will spend time talking about how these services work.
There are two ways to think about the Internet (and also about what we will cover in this course):

The first is in terms of the components that make up the Internet. So this would be a nuts-and-bolts view in which we can talk about several acronyms and terms that you may have come across, such as:

- web
- TCP/IP
- inter-network
- routers
- switches

and so on...

The second way to think about it is the one that is more common, which is in terms of the services and applications that run over the Internet. So all of us do web-browsing, exchange emails, share files.

The important point to note is that these applications we care about do not operate solely by themselves. There are numerous services that all work together to provide them.

— In this class we expose some of the seams. The services that enable such applications are all part of the Internet. So we will spend time talking about how these services work.

### Congestion Control in High Speed Networks

#### The Sluggishness of TCP

- 10 Gbps network with 100 ms round-trip time
  - Desired $\text{cwin} = 83,000$ packets

- Initial bandwidth discovery:
  - $\text{SSThresh}$ usually set to no more than 32,614 segments
  - Would take hours to achieve a sending rate of 10 Gbps

- Bandwidth rediscovery after timeout:
  - $\text{Cwin}$ reset to 1
  - Additive increase would still take hours to recover 10 Gbps throughput
Congestion Control in High Speed Networks
The Sluggishness of TCP

- Steady-state Congestion
  - Avoidance behavior:
    - If congestion events occur frequently, average throughput will be less than C
  - To achieve 10 Gbps with TCP, only 1 in (2*10^10) packets should be dropped
    - This is past the limits of achievable fiber error rates
    - Packet loss rate of 10^-7 is reasonable to expect

![Figure 1: Traditional TCP scaling properties.](image)

Scalable TCP
Basic Idea

- Multiplicative increase:
  - Increase window more aggressively
    - Standard TCP: cwin - cwin + 1
    - Scalable TCP: cwin - (1+α)cwin

- Multiplicative decrease:
  - Decrease window less aggressively
    - Standard TCP: cwin = min(cwin, c/2)
    - Scalable TCP: cwin = max(cwin - βcwin, c/2)

![Figure 2: Scalable TCP scaling properties.](image)

Average link utilization achieved is independent of link capacity
There are two ways to think about the Internet:

1. In terms of the components that make up the Internet. This would be a nuts-and-bolts view, discussing acronyms and terms such as web, TCP/IP, inter-network, routers, switches, and so on.

2. In terms of the services and applications that run over the Internet. All of us do web-browsing, exchange emails, share files.

Important point to note: These applications do not operate solely by themselves. There are numerous services that work together to provide a seamless experience. In this class, we expose some of the seams. The services that enable such applications are all part of the Internet. We will spend time talking about how these services work.