TCP/IP Architecture

Brighten Godfrey
cs598pbg August 31 2010
Today

Cerf and Kahn: TCP / IP

Clark: TCP / IP design philosophy

One more thing...
Interconnection challenges

- Different ways of addressing, supported packet lengths, latency, status information, routing
- Must let each network operate independently
- Solution:

![Diagram of network with "unacceptable alternative" note]
• Gateways sit at interface between networks

![Diagram of networks interconnected by gateways](image)

Fig. 2. Three networks interconnected by two gateways.

• ...and speak an Internetworking protocol

![Diagram of internetwork packet format](image)

Fig. 3. Internetwork packet format (fields not shown to scale).
IP packet fragmentation

- Allow maximum packet size to evolve
- Protocol mechanisms to split packets in-transit (byte-level sequence numbers)
- Reassemble at end-hosts
  - Why not gateways?
Unreliable datagrams

• No need for reliability support from underlying network
  • Greatly simplifies design
  • Exception handling always adds complexity

• Any problem? Just drop the packet.

• What’s not a stated reason for datagrams?
  • Statistical multiplexing.
• Routing unspecified—but constrained!
  • Picked address format
  • Hierarchical (network, host)
  • Route computed within network
  • 8 bits for network. “This size seems sufficient for the foreseeable future.”
  • Later: 32 bits in three size classes (A,B,C), and then CIDR.

• Many proposed new routing algorithms need to change the address format.
Ports

- Associated with a process on a host
- Identify endpoints of a connection (“association”)
- Rejected design: connection at host level; packet may include bytes for multiple processes
- What’s the difference between a port and an address?
What we now call TCP

- Window-based scheme
- Provides reliability, ordering, flow control
  - And now, congestion control too
  - Note you might want only some of these

![Diagram of window concept](image-url)
What we now call TCP

“It is our expectation that the host level retransmission mechanism ... will not be called upon very often in practice. Evidence already exists [ARPANET] that individual networks can be effectively constructed without this feature.”

- Why do you think they wrote this? Is it true now?
- No congestion control in this early version!
  - Congestion control introduces losses intentionally
- (Also missing: 3-way handshaking)
Today

Cerf and Kahn: TCP / IP

Clark: TCP / IP design philosophy

One more thing...

delayed to next time
Cerf and Kahn: TCP / IP
Clark: TCP / IP design philosophy
One more thing...
What happened on Friday?

- About 1% of Internet destinations disrupted for about 30 minutes
- How did this happen?
Internet had a bad Friday

BGP update rate at London Internet Exchange

[Plots by Brighten based on raw update feeds from Route Views]
Internet had a bad Friday

[Plots by Brighten based on raw update feeds from Route Views]
~1% of prefixes affected

[Earl Zmijewski, Renesys]
Brewing a storm

[insert whiteboard here]
Brewing a storm

1. An unusual announcement
2. Propagation from router to router
3. Buggy software mangles announcement
4. BGP session dropped upon receipt of mangled message
5. BGP session reestablished and process repeats
Lessons

• Many unsavory BGP announcements can be contained, but this one wasn’t
  • Spread geographically because it was an entirely valid announcement
  • Spread to many prefixes because BGP spec lets one bad announcement affect all traffic

• Widespread correlated failures from similar software