Internet Architecture via the Lens of History

Hari Balakrishnan <hari@mit.edu>

6.829 Fall 2016
Lecture #2

Several pictures taken from Wikipedia
It is anticipated that the whole of the populous parts of the United States will, within two or three years, be covered with net-work like a spider's web.

When was this sentence written?
About the *electric telegraph*
In *The London Anecdotes*, 1848
As quoted in *The Victorian Internet*
[Tom Standage]
The Soviet Union launches Sputnik-1 in October 1957
1958

Sputnik led to the creation of ARPA in Feb 1958

Now called DARPA, the Defense Advanced Research Projects Agency.
The Dawn of Packet Switching: 1959-1965

Paul Baran (RAND Corp)
- Early 1960s: New approaches for survivable comms systems; “hot potato routing” and decentralized architecture, paper on packet switching over digital links

Donald Davies (UK), early 1960s
- Coins the term “packet”


J. Licklider & W. Clark (MIT), On-line Man Computer Communication (1962) & Licklider’s vision of a “galactic network”

L. Roberts (MIT then ARPA), first ARPANET plan for time-sharing remote computers [SOSP 1967!]

FIG 1 - Centralized, Decentralized and Distributed Networks
ARPANET

BBN team that implemented the *interface message processor* (IMP)

- 1967: Connect computers at key research sites across the US using telephone lines
- Interface Message Processors (IMP) ARPA contract to BBN
- Sen. Ted Kennedy sends a somewhat confused telegram to BBN on winning the contract
  “Congratulations ... on *interfaith* message processor”
Initial Baby Steps

THE ARPA NETWORK

DEC 1969

4 NODES

FIGURE 6.2 Drawing of 4 Node Network
(Courtesy of Alex McKenzie)
In the Beginning...

- Kleinrock’s group at UCLA tried to log on to SRI computer: His recollection of the event...
- “We set up a telephone connection between us and the guys at SRI...

- We typed the L and we asked on the phone...
  - “Do you see the L?”
  - “Yes, we see the L,” came the response

- We typed the O, and we asked...
  - “Do you see the O?”
  - “Yes, we see the O.”

- Then we typed the G...
  - ...and the system crashed!
1970, ARPANET hosts start using NCP; first two cross-country lines (BBN-UCLA and MIT-Utah) “Hostile overlay” atop telephone network Ran a distance-vector routing protocol
1970s: Packet networks → Internetworking

- 1972: successful ARPANET demo at conference (except it failed when demo’d to skeptics from AT&T!)

- 1972: modified ARPANET email program

- 1972: CYCLADES network (Louis Pouzin et al.): best-effort “datagrams”; *sliding window* protocol; distance-vector routing; time sync – many good ideas

- 1973: Ethernet (MAC protocol inspired by Aloha – CSMA)
- 1973-74: Xerox PUP (internetworking concepts)

- 1973: ARPANET becomes international
- 1973-75: Internetworking effort (Cerf, Kahn, et al.)
  - Developed TCP and IP (originally intertwined) – TCP uses *sliding window*
The Problem

- Many different packet-switching networks
- Different administrations
- Only nodes on the same network could communicate
Kahn’s Rules for Interconnection

- Each network is independent and must not be required to change
- Best-effort communication
- Boxes (then called gateways) connect networks
- No global control at operations level (why?)

Original TCP paper

A Protocol for Packet Network Intercommunication

VINTON G. CERF and ROBERT E. KAIN,

INTRODUCTION

In the last few years, a number of network applications and services have been developed, and many of these have been based on packet switching networks. A packet switching network is a collection of independent computer systems, each responsible for its own part of the network, and connected together by means of a transmission facility. The transmission facility is responsible for delivering packets of data from one computer system to another. Each computer system is capable of generating and receiving packets, and each packet contains a header that specifies the source and destination of the packet.

The earliest packet switching networks were developed in the late 1960s and early 1970s, and were based on the concept of a network of independent computer systems, each responsible for its own part of the network. This concept was later formalized in the model known as the Internet model, which is the basis for the current Internet.

The original TCP paper

Cerf RFC 968

Twas the night before start-up and all through the net, not a packet was moving; no bit nor octet.

The engineers rattled their cards in despair, hoping a bad chip would blow with a flare.

The salesmen were nestled all snug in their beds, while visions of data nets danced in their heads.

And I with my datascope tracings and dumps prepared for some pretty bad bruises and lumps.

When out in the hall there arose such a clatter, I sprang from my desk to see what was the matter.

There stood at the threshold with PC in tow, An ARPAFILE hacker, all ready to go.

I could see from the creases that covered his brow, he'd conquer the crisis confronting him now.

More rapid than eagles, he checked each alarm and scrutinized each for its potential harm.
The Internetworking Vision

• Bob Kahn & Vint Cerf imagined there would be only a few networks and thus only a few gateways
  – “The choice for network identification (8 bits) allows up to 256 distinct networks. This size seems sufficient for the foreseeable future.”
  – They were a little wrong!

• Gateways would “translate” between networks
  – Evolved in the 1974 Cerf/Kahn paper as a universal network layer, later called the Internet Protocol, or IP
  – We now think of it as all routers supporting IP
Handling Heterogeneity

• Make it easy to be a node or link on the network (best-effort)

• Universal *network layer*: standardize addressing and forwarding

• “IP-over-everything”

• Switches: *no per-connection state* on behalf of end points

• Original addressing model (then called a “TCP address” because IP and TCP hadn’t been split into different layers):

  ![TCP Address Diagram](image)

  *Fig. 4. TCP address.*
1970s: Internetworking

- **1978: Layering!** TCP and IP split; TCP at end points, IP in the network
  - IP network layer: simple best-effort delivery
- In retrospect: Packet switching (& TCP/IP) won because it is good enough for almost every application (though optimal for nothing!)
- Competitor to TCP/IP: ISO, standardizing 7-layer OSI stack

---

Classic Internet layering

"hourglass" model
Most Useful Lesson

One should architect systems for flexibility – you’ll almost never know what apps make it succeed.

(Even if it means sacrificing some performance!)

Il semble que la perfection soit atteinte non quand il n'y a plus rien à ajouter, mais quand il n'y a plus rien à retrancher.
Perfection is achieved, not when there is nothing more to add, but when there is nothing left to take away

-- Antoine de Saint-Exupery

Or,

When in doubt, leave it out
1980s: Handling Growth with Topological Addressing

• 1978-79: ARPANET moves to *link-state routing*

• Per-node routing entries don’t scale well

• Solution: Organize network hierarchically
  – Into “areas” or “domains”
  – Similar to how the postal system works
  – Hide detailed information about remote areas

• For this approach to work, node addresses must be *topological*
  – Address should tell network *where* in the network the node is
  – I.e., address is a *location* in the network
1980s: Rapid Growth

- 1981-89: Dave Clark of MIT is Internet’s “Chief Architect”
  - Co-author of the end-to-end arguments (w/ Saltzer/Reed)
  - Ensures consistency of design and vision
  - “We reject kings, presidents, and voting. We believe in rough consensus and running code.”

- Saltzer, Reed, Clark: End-to-end arguments in system design [1984]

- 1982: US DoD standardizes on TCP/IP
  - Berkeley’s computer systems research group produces BSD & sockets

- 1983: MIT Project Athena – large-scale campus-area networking

- 1984: Domain Name System (DNS) introduced

- 1985: NSFNet picks TCP/IP as standard
Growth Problems: Congestion

- 1986: Congestion collapse episodes on the Internet
  - Problems with bad timeout settings
  - Window size not appropriate for network state
  - Athena network file system congestion problems (bad timeout settings)

- Congestion avoidance and control
  - RTT estimation using EWMA + new timeout method
  - TCP congestion control by Van Jacobson (concurrent work on DECBit scheme by Ramakrishnan & Jain)
  - Adapt the window size to congestion: If congested, decrease window; else increase. Use exponential back-offs on timeouts
  - By the end of the 1980s, essentially all running TCPs had congestion control

*Jacobson*
1990s

- 1990: no more ARPANET
- Mid-1990s: NSFNet backbone ends
  - Commercial ISPs take off
- “Classless” addressing for scale
  - And the rise of NATs
- BGP4: *Path vector protocol* between competing ISPs, who must yet cooperate
1990s: Handling Growth with CIDR
IPv4 Addresses & Address Prefixes

- 18.31.0.82 is actually the 32 bit string 00010010001111100000000001010010
- Routers have forwarding table entries corresponding to an address *prefix* (a range of addr w/ common prefix bitstring)
- 18.0.0.0/8 stands for all IP addresses in the range 00010010 00...0 to 00010010 11...1 (i.e., $2^{24}$ addresses of the form 00010010*)
- 18.31.0.0/17 stands for a range of $2^{15}$ consecutive IP addresses of the form 00010010001111100* (1st 17 bits are the same for each address in that range)
- Hence, *subnetworks* may be of size 1, 2, 4, 8, ... (maxing out at $2^{24}$ usually), and may be recursively divided further
- Forwarding uses *longest prefix match*
  - At each router, routes are of the form “For this range of addresses, use this route”
  - Pick the route that has the longest matching prefix w/ dest addr
1990s

- 1990: no more ARPANET
- Mid-1990s: NSFNet backbone ends
  - Commercial ISPs take off
- “Classless” addressing for scale
  - And the rise of NATs
- BGP4: *Path vector protocol* between competing ISPs, who must yet cooperate
- IP multicast
- IP quality-of-service (QoS)
1990s (cont.)

- 1993: search engines (Excite)
- Mid-1990s: E-commerce starts
- Late 1990s: news on the web
- Mid-1990s: some misplaced fears of the Internet melting down
- 1998: Google reinvents search
- 1998: Content distribution networks (Akamai)
- 1996-2001: .com bubble starts & bursts
2000-2010: The Internet Matures
(Top 5 List)

• 2000-2001: .com bust
  – And 9/11 happens
  – Power of content distribution networks to handle load
1. The rise of peer-to-peer networks
   – Napster, Gnutella, Freenet, distributed hash tables (e.g., Chord), BitTorrent
2. Security threats and defenses
   – 2000: Large-scale distributed denial-of-service (DDoS) attacks start
   – 2003: SQL slammer worm
   – Spam → phishing and pharming → complex ecosystem
   – Route hijacking by errors or malice
3. User-generated content & social networks
   – Blogs, Youtube (2004), Facebook (2003), Twitter
4. The rise of wireless and mobile data
5. Cloud computing and large-scale datacenters (Amazon, Google, Microsoft, Facebook, etc.)
• Almost everything moves to the Internet: telephony, video, entertainment
What You Said

“An unforeseen consequence of the IPv4 "narrow waist" was massive reliance on this shared protocol. As a result, deployment of IPv6 has faced monumental challenges, and furthermore it has become almost impossible to make changes ("ossification") to the Internet architecture.” [Akshay Narayan]

“The end-points of the network were implicitly assumed to be machines at the beginning. Later on this is being questioned, and extended to services and information objects.” [Qinxuan Pan]

“Initially, security was considered in the context of availability: what if a gateway fails, accidentally or due to hostile intervention. ... Thus all kind of security solutions were developed: white listing of trusted machines to fight spam, firewalls, and more specific techniques such as taint-tracking.” [Sitan Chen]

“The question of ... how to align the economic incentives of individual agents to the goals of the network itself is an important concern.” [Oleksandr Chaykovskyy]
Architectural Weaknesses & Warts

1. Designed with trusted endpoints in mind
   Unwanted traffic, lack of accountability, poor level of security
2. Mobility not a first-class goal
3. Middleboxes an afterthought
4. Management, measurement, and fault diagnosis are hard
5. Higher availability for those that need it: in architecture or as an overlay?
6. Design for economic viability: resource accounting is hard
7. Multi-homing an afterthought
8. Should data and content be first-class objects?
9. Better-than-best-effort services and support for application objectives
10. The architecture isn’t evolvable
11. Cross-layer feedback is hard
This Decade and Next

- Greater growth (video, entertainment, collaboration), internationalization
- Better traffic control for application performance (human engineering, machine learning)
- Combating complexity: new methods to make things simpler (control plane: “software-defined networks”)
- Programmable data planes: doing more with network hardware
- Mobility and Wireless: anytime/anywhere access, the spectrum crunch, high variability
- Physical embedding & embodiment: sensors & actuators over the network, mobile robots and autonomous agents, vehicles
- Network security & privacy, anti-censorship, surveillance