Internet Architectural Philosophy and the New Business Reality

Scott Bradner
Harvard University

Topics

- architecture (as design philosophy)
- key decisions
- architecture (as reality)
- and then there is money

Background

- multiple unrelated efforts (early to mid 1960's)
  - packet switching theory: (Kleinrock) 1961
  - day dreaming: (Licklider's Galactic Network) 1962
  - make use of remote expensive computers: (Roberts) 1964
  - survivable infrastructure for voice and data: (Baron) 1964
- ARPANET (late 1960's)
  - Roberts ARPANET paper 1967
  - RFP for "Interface Message Processor" won by BBN 1968
  - four ARPANET hosts by 1969
  - public demo and email in 1972

Fundamental Goal of Internet Protocols

- multiplexed utilization of **existing** networks
  - different administrative boundaries
  - multiplexing via packets
  - networks interconnected with packet switches
    - called gateways (now called routers)
  - note: international in scope
- did not want to build a new global network
  - too expensive
  - too limiting

Internet Protocols Design Philosophy

- ordered set of 2nd-level goals
  - 1/ **survivability** in the face of failure
  - 2/ support **multiple types** of communications service
  - 3/ accommodate a **variety** of network types
  - 4/ permit **distributed management** of resources
  - 5/ **cost effective**
  - 6/ **low effort** to attach a host
  - 7/ **account** for use of resources
- note: no performance (QoS) or security goals
- not all goals have been met
  - management & accounting functions are limited

Packets!

- basic decision: use packets not circuits
  - Kleinrock's work showed packet switching to be a more efficient switching method
- packet (a.k.a. datagram)

Dest Addr	Src Addr	payload

  - self contained
  - handled independently of preceding or following packets
  - contains destination and source **internetwork** address
  - **may** contain processing hints (e.g. QoS tag)
  - **no delivery guarantees**
    - net may drop, duplicate, or deliver out of order
    - reliability (where needed) is done at higher levels

Routing

- sub parts of the network are connected together by computers that forward packets toward destination
  - these computers are called "**routers**"
- routers use destination address in packet to make forwarding decision
- routers exchange reachability information with other routers to build tables of "next hops" toward specific local networks
  - exchange of reachability information done with "**routing protocol**"

arch - 7

Unreliability can be Important

- basic decision: **offer** an unreliable service
- 1st idea was to only have TCP (a reliable service)
- problems
  - not good for voice & video
    - data has to be delivered in time - retransmission for reliability causes too great a delay
  - not good for all applications
    - e.g. a debugger has to work in lossy environment
    - retransmission algorithm may vary with application
- thus: **split** IP & TCP and add UDP
  - IETF just added SCTP

arch - 8

A Quote

*"the lesson of the Internet is that **efficiency is not the primary consideration**. Ability to grow and adapt to changing requirements is the primary consideration. This makes simplicity and uniformity very precious indeed."*

Bob Braden

Networks as Generic

- design requirement of working over:
  - existing networks & a wide variety of networks
- minimum set of assumptions about network
  - reasonable size packets, reasonable but not perfect delivery reliability, network-wide addressing, way to get error messages back to source, no assumption of in-order packet delivery
- "smart wires" are not much of a help
  - e.g. X.25 (reliable delivery)
  - e.g. ATM (QoS functions)
- thus it is easy to use new types of networks
  - assuming they are not too helpful (feature rich)

End-to-End Argument

- 1981 paper by Saltzer, Reed & Clark
- “smart networks” do not help
  - adding functions into network can be redundant since actual function is end-to-end
    - e.g. encryption, data reliability
  - also harder to change with new technology
    - also see Lampson ***Hints for Computer System Design***
- e2e argument projected to mean
  - no per-session knowledge or state in the network
    - but some “soft-state” (auto refreshed) may be OK
  - network should be transparent to end-to-end applications

Ease of Experimentation With e2e

- easier to experiment in an e2e environment
  - if the network is transparent then only nodes involved are the end nodes
  - note that an end node could be a 3rd party server
  - no need to get permission to experiment
- cheaper to experiment
  - can do much smaller scale experiments - down to 2 nodes
  - than core-based services
- WWW an example of what can be done

Economic Driver?

- Mark Gaynor Harvard PHD thesis
- define market uncertainty as **MU**
  - how well do you know what the customer wants
- low MU means customer wants are known
  - e.g. "voice service"
  - no opportunity to be "better" than competitor
- high MU means customer wants are not known
  - e.g. future IP-enabled voice service
  - opportunity to better match customer wants than competitor does

Economic Driver, contd.

- low MU
  - commodity service
  - provide most efficient way - frequently centralized
- high MU
  - need to experiment to try to match customer want
  - note: if only one company figures it out they dominate the market
  - easier to experiment on edges
    - i.e. e2e is a innovation friendly model
    - even if its more expensive to provide service to ends

Smart vs. Stupid Networks

- phone network technology: self-named “Intelligent Network” (IN)
  - many network-based services
    - admission control, number translation, accounting, ...
- Isenberg’s ***Rise of the Stupid Network*** compared phone network’s “Intelligent Network” to Internet
  - Isenberg’s basic messages:
    - network (i.e. carrier) -based services slow to change
    - voice is not all there is
    - carrier gets in the way
    - just “deliver the bits” works

But!!

- a “stupid network” is a commodity service
  - the price of a commodity service is driven by the stupidest vendor
- hard to make money delivering commodity services
- new network infrastructure is very expensive
  - fiber optic cables (with installation) & hardware
- access rights can also be very expensive
  - e.g. wireless spectrum licenses
- carriers need something else to make money
  - common dream is that services or content will save the day
    - may be a false dream

$

But!! (2)

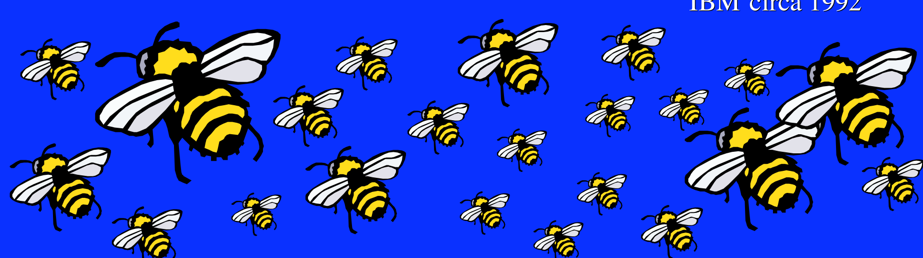
- packets w/o circuits cause problems
  - can not do guaranteed QoS
    - can not control path packets take
    - can not reserve capacity for application
  - security control harder
    - do not have logical "wire" back to source
  - management harder
    - can not see data patterns on the network
    - finding non-catastrophic failures harder
  - service provider interconnections harder
    - no clean interface
- lack of useful formal tools to describe performance

!QoS

Conceptualization Problem

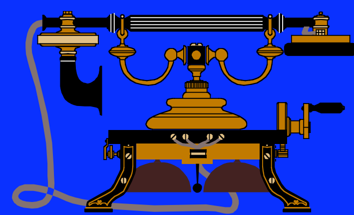
- fundamental disconnect between "Internet" and "phone" people "bell-heads vs. net-heads"
- by their definition the Internet can not work
  - and must be fixed - they will rescue us

"***You can not build corporate network out of TCP/IP.***"
IBM circa 1992



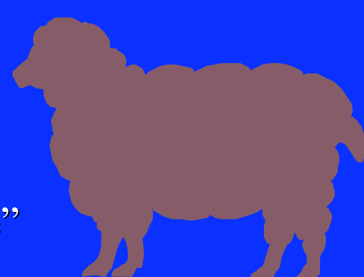
Traditional Phone Network

- circuits & “smart network”
- connection-oriented
- hard state in network devices
- fragile
- central resource control
- socialist? "for the good of all"
- applications in network
  - e.g., phone switch
  - end-to-end touch-tone signaling was a mistake
- predictable development path
  - extended development cycle



Internet

- packets & e2e
- soft state in network devices
- resilient
- competitive resource control
- capitalist? "individual initiative”
  - but too much selfishness hurts all
  - must play by the same rules - but no enforcement
    - **the tragedy of the commons**
- applications in hosts at edges (end-to-end)
  - and in 3rd party servers anywhere on the net
- hard to predict developments
  - chaos at the rate of “Internet time”

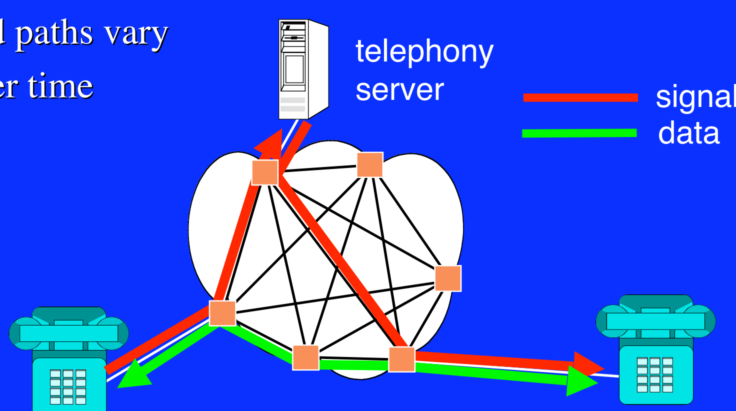


More Conceptualization Problems

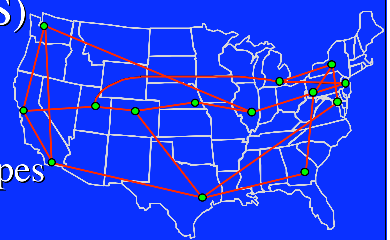
- service provided by 3rd parties - not only by carriers
  - different from phone world
- a quote from an IETF telephony mailing list
  - Hi Roy,
    - I still don't understand why it is a "users" choice where the "services" are executed - I would have thought that this would be networks choice

Disjoint Control and Data Paths

- signaling and data paths in Internet may not coincide
  - and paths vary over time



Circuits in the Internet



- do not seem to go away (MPLS)
- used for traffic engineering
  - city-pair pipes
  - maybe class of service city-pair pipes
- and customer connections
- finer grain (instance of application) use still pushed
- remember the fate of ATM
  - circuit - used for trunks not flows
  - QoS - ignored (ATM not end-to-end)
  - link sharing - may make sense
  - as the bearer service - did not make it

IP as a Common Bearer Service

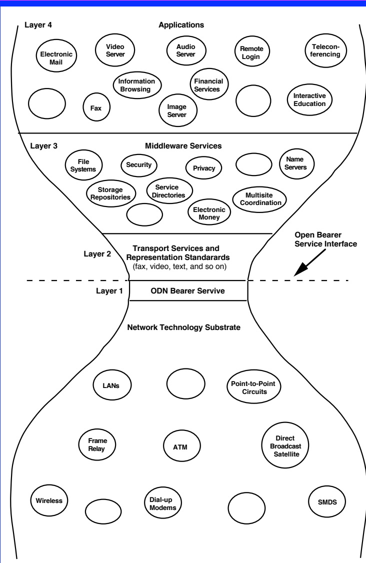


FIGURE 2.1 A four-layer model for the Open Data Network

From: Realizing the Information Future

Net is No Longer Transparent

- end-to-end argument says the net should be transparent
  - i.e. packet not modified in transit (other than TTL)
  - global-scope internetwork address
  - i.e., packet goes to address in destination address field
- transparency now gone in some cases
  - NATs, firewalls, proxies, content caches, TCP reshapers
  - replace addresses, intercept traffic, insert traffic
- other issues
  - wiretapping, taxation, content filtering

NAT/Firewall/Cache Issues

- can not trust IP address as end-to-end
  - breaks IPSec, not sure who you are talking to
- applications with addresses in data
  - have to have application-specific support (ALG) in devices
  - deploying new application requires approval of net manager
- dynamic port usage
  - ALG must snoop on application traffic
  - ALG must understand application logic
- new IETF effort to develop generic signaling
  - may help some
  - but will not make these devices transparent

Trust-Free Environment

- original Internet architecture assumed a trustworthy environment
- no longer the case
  - mistrust net itself (eavesdropping, reliability etc)
  - mistrust that you are talking to the right end point
    - e.g., proxy, redirect, spoofing (MAC & IP address)
  - unsolicited correspondence (spam)
  - anonymity hard to get
  - mistrust own hardware and software
  - 3rd parties insist on being in the middle
    - filters, wiretapping, …

Summary of Architectural Points

- datagram-based network
  - not circuit switched
- network of networks
  - different parts under different management
- minimize per-session state in network
  - some auto-refreshed state is OK
- end-to-end model maximizes flexibility
  - network does not need to know what you are doing
- "smart wires" can get in the way
  - e.g., nested control loops
- reliable delivery is an option
  - not a requirement

Key Decisions

- a few key decisions brought us here
  to the Internet of today
- but there was no way to predict where we are now
- **unplanned parenthood**

10 Decisions That Made a Difference

- support existing network**s**
- **datagram**-based
- creating the **router** function
- split TCP **and** IP
- DARPA fund Berkeley to add TCP/IP to **UNIX**
- CSNET and **CSNET/ARPANET** deal
- NSF **require TCP/IP** on NSFnet
- ISO **turn down** TCP/IP
- NSF Acceptable Use Policy (**AUP**)
- **minimal** regulation

Internet Architecture

- #1 goal of original Internet protocols was to deal with a **network of networks**
  - not a single type of network
  - not under one management
- networks interconnected at datagram level
  - no session-aware logic at interconnections
- bi-lateral interconnection agreements
  - "customer" - buy transit service to "the Internet"
  - "peer" - cost sharing connection to a network and its customers

Customer Interconnection

- one network pays another for access to "the Internet"
  - paying network can be Internet service provider (ISP) or enterprise
  - only as useful as resulting coverage
    - "Metcafe's Law": value of network increases by square of the number of reachable nodes
- customer can move business to another network if they do not like the service
  - may have to renumber to preserve addressing topology

Peering

- business decision
  - no current regulations
- it can be cost effective for two networks to interconnect sharing the costs of the links
  - interconnection can be at "public peering points" or using dedicated links between networks
- but only "see" other network and their customers
  - not the other network's other peers
- must peer with all large networks to get "the Internet"
  - or be a customer to another network (or networks)

Multi-Homing

- one network (ISP or enterprise) can connect to more than one other network
  - for redundancy and reliability
  - called "multi-homed"
- causes some complexity in the routing setup

Public Peering Points

- 3 originally designated by National Science Foundation (NSF) as part of the breakup of the NSFnet
- now many local peering points around the world
  - but telcom costs can discourage use in some countries
    - cheaper to get lines to US than within country
- level-2 interconnect
  - like an local area network (e.g. an Ethernet)
  - i.e. not involved in IP-level routing

Private Peering

- two ISPs can agree to interconnect sharing costs
  - “you buy and run one line, I’ll buy and run another”
  - peering list normally private
- ISPs have minimum criteria before peering will be considered
  - some publish the criteria
- criteria normally include
  - minimum level of interconnect traffic, traffic balance,
  - backbone size, geographic scope,
  - competent network operations center

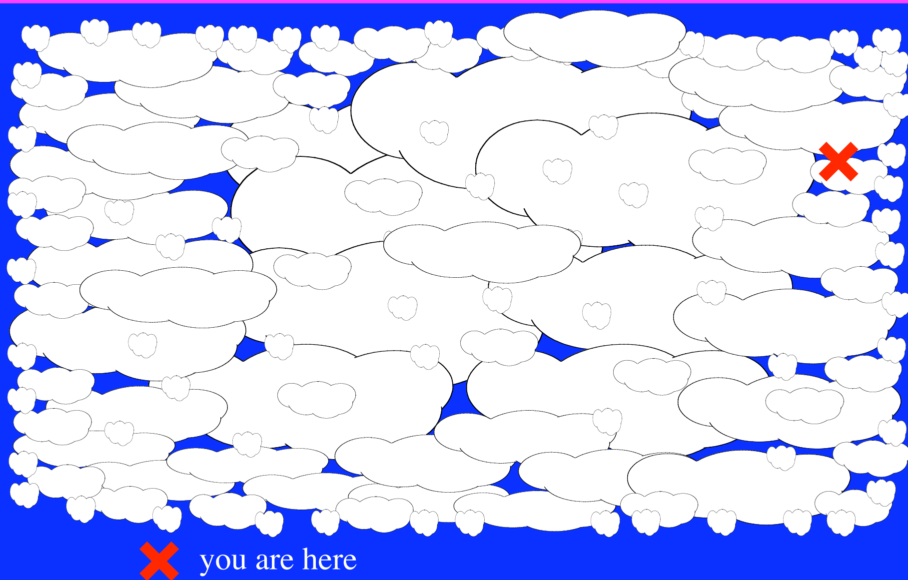
Tier 1 ISPs

- some big ISPs are referred to as "Tier 1 ISPs"
- no real externally verifiable definition
- general concept:
  - "an ISP that gets most of not all of its connectivity from peering, not by being a transit customer"
  - i.e. a Tier 1 ISP is one that is connected to the other Tier 1 ISPs

Interconnection Pattern

- no explicit network hierarchy assumed
- no specific pattern to ISP interconnections
  - other than that peering tends to be between networks of the same basic size
  - but not always - can have business reasons for mismatch
- peering and transit connections can appear random
- **notes:**
  - most traffic does not flow through Tier 1 ISPs
    - many "lower-level" interconnections
  - hard (impossible) to know relative sizes of ISPs

Current Internet Architecture



Money

- “but who is going to make money at that?”

  John McQuillan

- how is the carrier supported?

  “we do not know how to route money”

  Dave Clark

- carrier wants a piece of the action

  e.g., WAP, AT&T proposal

- is content king?

  factoid: total US movie revenue ~= 2 weeks of US phone charges

iMode: A Model?

- DoCoMo's iMode service
  - more than 30 million subscribers
  - 9.6Kb data service
  - 50,000 iMode compatible sites
  - DoCoMo works with less than 10% of them
    - does billing, runs servers etc
    - rest are on their own
- key decision: open access (NOT WAP!)
  - makes service more attractive
  - DoCoMo charges monthly fee and for data transferred

More on Money

- QoS does not seem to be a useful charging base
- differentiated by application is an intelligence test
  - railroads in US used to do this (Rhode Island Line)
  - not enough will fail the test
- and then there is all that fiber
- do municipalities have a role?

A Bit More on Money

- what happened to that $ trillion anyway?
  - few infrastructures pay for themselves
  - the Internet is not an exception
- is there a difference now that the fiber is “free”?

Last Word

- Internet “too important to fail” (?)
  - what about ISPs (can you say “KPNQuest”?)
- will there be anyone left standing other than the telcos?
  - what can they see from their point of view?
- will you be able to say “Internet” and “business model” in the same sentence?
  - without a “no” in between

What’s Next?

arch - 45