The Internet and Optical Networking at the IETF

COIN 2002

Scott Bradner
Harvard University

Syllabus

- a short technical history of the Internet
- Internet architectures
- the IETF
- the IETF Sub-IP area
- IETF & optical networks

Why History

- important to know where the Internet people are coming from
- specifically packets vs. circuits
- effects IETF view of optical nets

Fundamental Goal of Internet Protocols

- multiplexed utilization of **existing** networks
  - different administrative boundaries
  - different network types
  - 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, same format end-to-end
    - not link dependent
  - 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

End Node Function

- assumption:
  - end nodes do more than one thing at a time
    - not all with the same other end
  - multiple concurrent sessions on each end node
  - e.g., a WWW session is actually multiple sessions
    - to multiple servers
- i.e., multiplex at packet level
  - not single circuit
  - not even a few VCs (e.g., ATM)

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**"

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

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

IP as a Common Bearer Service

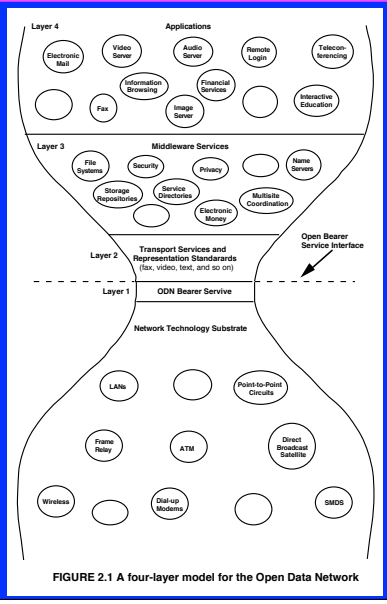


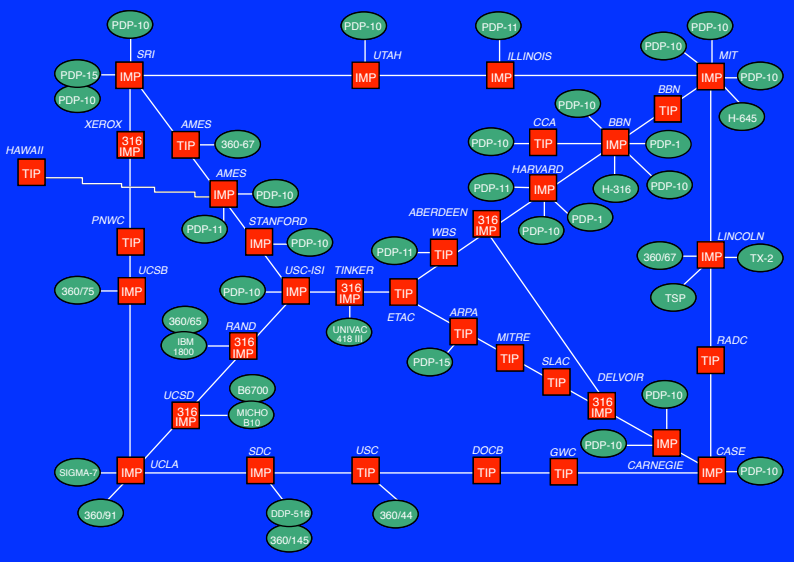
FIGURE 2.1 A four-layer model for the Open Data Network

From: Realizing the Information Future

Architecture of the Internet

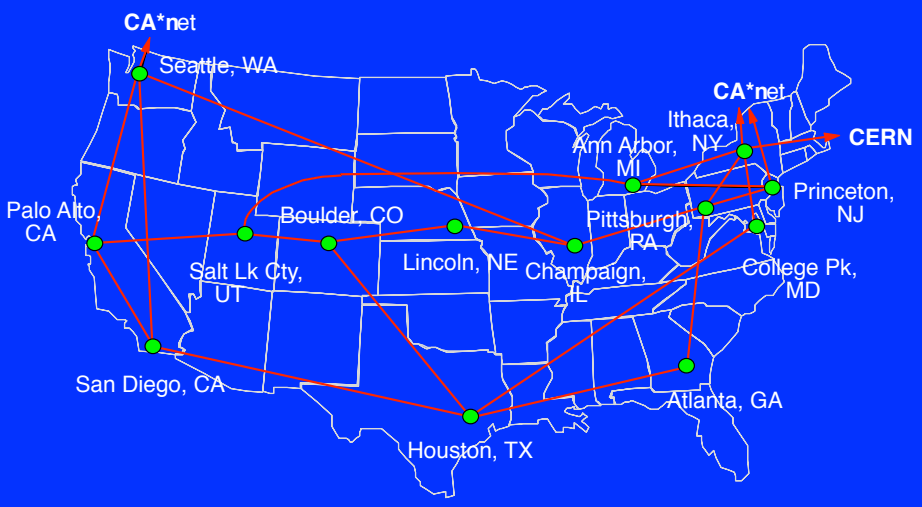
- history of nets that created the Internet
- traffic engineering

1973: ARPANET

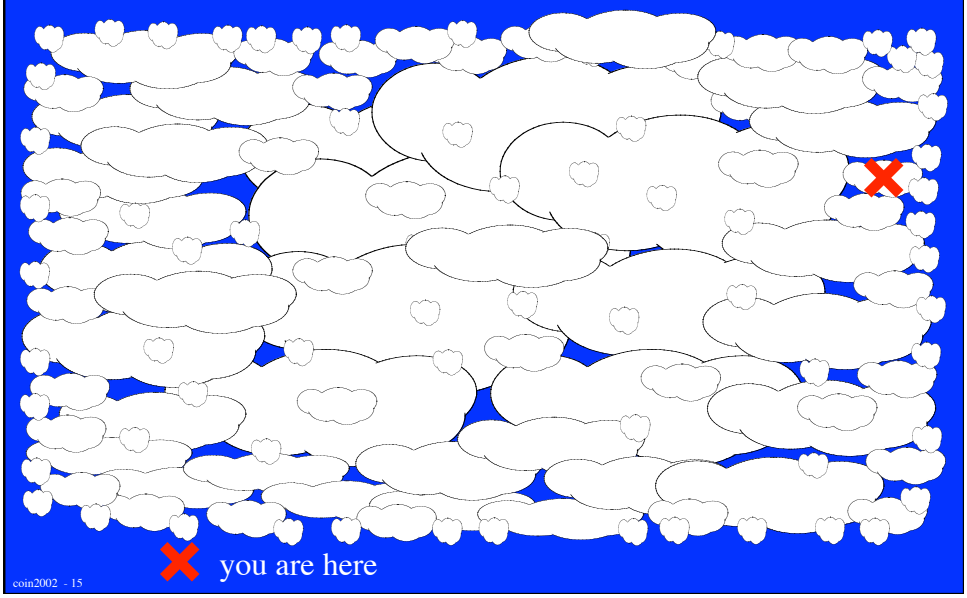


ARPA NETWORK, LOGICAL MAP, JANUARY 1973

1990: NSFNet T1 network



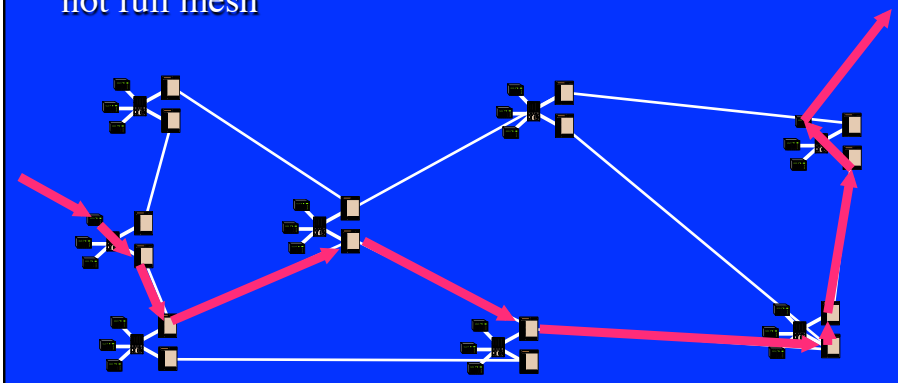
2002: Current Internet Architecture



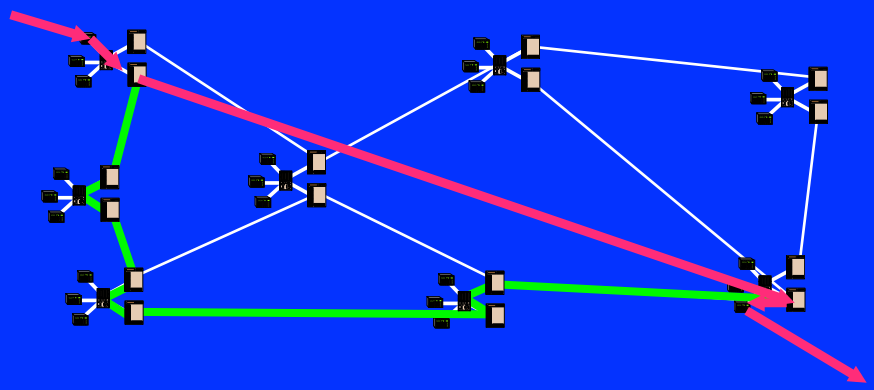
Traditional ISP Architecture

- multiple POPs
- interconnected with backbone links

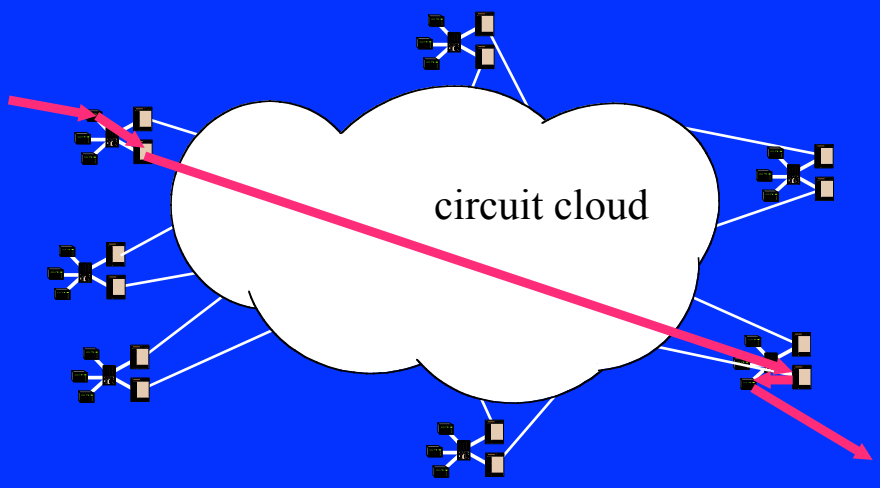
  not full mesh



Traditional ISP Arch.: with MPLS



Alternate ISP Architecture



Core Circuits

- multiple options for technology for core circuits
  - frame relay
  - ATM
  - MPLS
  - SONET/SDH
  - optical
- (full) mesh connections between ISP POPs
  - "core" routers are L3 adjacent
  - note that "customer" routers not connected to core

Circuits

- IP ISPs use circuits to connect POP core routers together
- can redirect circuit path to do traffic engineering
- currently few or no QoS-specific "sub-pipes"
- others see MPLS as a way to do other types of circuits
  - Virtual Private Networks (VPNs)
  - application-specific circuits (e.g. phone calls)

The IETF

- The Internet Engineering Task Force
- standards development for the Internet
- since 1986
- international
  - most recent meeting - last week in Yokohama
- individuals not organizations
- no defined membership
- scale: about 2,000 attendees in Yokohama
  - thousands more on mailing lists (from 100s of companies)
- under umbrella of the Internet Society (ISOC)

The IETF Organization

- most work done on mailing lists
- 3 times a year face-to-face meetings
- individuals or groups request BOFs
  - exploratory meeting - may lead to working group
- working groups for specific projects
  - about 135 working groups
  - restrictive charters with milestones
  - working groups closed when their work is done
- working groups gathered together into Areas
  - each area has 1 or 2 Area Directors - managers

IETF Areas

- Applications Area
- General Area
- Internet Area
- Operations and Management Area
- Routing Area
- Security Area
- Sub-IP Area
- Transport Area

IETF Standards Process

- "rough consensus and running code"
  - rough consensus required not unanimity
  - interoperable implementations needed to advance standard
- multi-stage standards process
  - Proposed Standard: good idea, no known problems
  - Draft Standard: multiple interoperable implementations
  - Standard: market acceptance

Above and below

- traditionally the IETF has been:
  - "above the wire and below the application"
  - not (often) defining user interfaces
  - not defining physical wire types
- while doing "IP over foo"
  - "foo" has been types of networks
    - Ethernet, Token Ring, ATM, SONET/SDH, ...
  - but foo has been changing

IP over "trails", "circuits", "paths", ...

- what looks like wires to IP may not be physical wires
  - may instead be something where paths can be configured
  - where a path looks like a wire to IP
    - e.g. ATM VCs & optical networks
  - might also be routed datagrams another layer down
    - e.g. IPsec tunnels
- and then there is MPLS
  - a progressively more important "foo"

A New Area

- a systematic approach to sub-IP issues would be nice
  - but exact scope is not clear
- IESG created a temporary area for sub-IP
  - like what was done for IPng
- to be short lived (1-2 years)
  - 2 current ADs were appointed to run the area

Non-Objectives

- the IETF is not expanding into standards for physical or virtual circuit technologies
  - no new circuit switch architecture from IETF
  - e.g., the IETF is not working on optical switches
  - leave them to others
- need to communicate with other standards organizations on what we are actually doing
- significant overlap with some other SDOs
  - OIF, ITU-T, etc

Area Work

- "Layer 2.5" protocol: MPLS
- protocols that monitor, manage or effect logical circuit technology
  - e.g. IP Over Optical, Traffic Engineering, Common Control and Management Protocols
- protocols that create logical circuits over IP
  - e.g. Provider Provisioned VPNs
- protocols that interface to forwarding hardware
  - General Switch Management Protocol

IETF Optical Work

- technologies for Internet service providers (ISPs)
  - not necessarily anyone else - but may be useful to others
- i.e, IETF works on technology for the Internet (including private IP networks), the technology may be useful for networks not carrying IP but it's not a design goal
- ways to control optical networks from IP point of view
  - based on IETF traffic engineering technologies
  - i.e., intelligent IP-based control plane for optical networks

IETF Sub-IP Basic Architecture

- for all sub-IP network types
  - not just pure optical nets
- two main components
  - topology discovery
  - control signaling
- development work being done in IETF ccamp working group

Topology Discovery

- run link-state routing protocol
  - e.g., OSPF or IS-IS
- report on L2 links
  - OSPF & IS-IS normally report on L3 links
- add link parameters
  - e.g., bandwidth, type of link, restoration functions
- can support link bundling
  - multiple parallel L2 links treated as one
  - bundle characteristics are an aggregate of link characteristics

Control Signaling

- used to establish paths through sub-IP network
- uses labels to identify paths

Traffic Engineering

- aim: combat congestion at a reasonable cost
  - networks w/o congestion are not a problem
  - other than speed of light issues
- decide paths through network rather than letting routing do its thing
  - paths could be in infrastructure: ATM PVCs, Frame relay PVCs, optical (SONET, Ethernet & other)
  - paths could be IP-ish: MPLS
- note - tail circuits a common congestion point
  - but can not be traffic engineered around
- same issue with servers

MPLS

- Multiprotocol Label Switching
- basic functions:
  - direct packets in a way that routing would not have
    - but not required feature
  - enable packet forwarding based on things other than IP destination address
  - simplify network core (e.g., no routing needed)
  - aggregate traffic with some common characteristics
  - can provide traffic matrix data
  - apply QoS to specific traffic group

MPLS, contd.

- not really routing (was in IETF routing area)
- circuit-based path setup
- original purposes:
  - traffic engineering & forwarding speed
- moving into QoS
  - circuit per QoS class -> circuit per flow
- some treating MPLS like packet-based ATM

GMPLS

- generalized MPLS
- assumes sub-IP links can be controlled with tags
  - extension of MPLS concepts
- routing algorithms do not need to be standardized
  - used to compute explicit routes - not hop-by-hop routing
- can do link bundling for scaling
  - parallel links between switches can be treated as a bundle
- data and control planes do not need to be the same

The Internet & Optical Networks

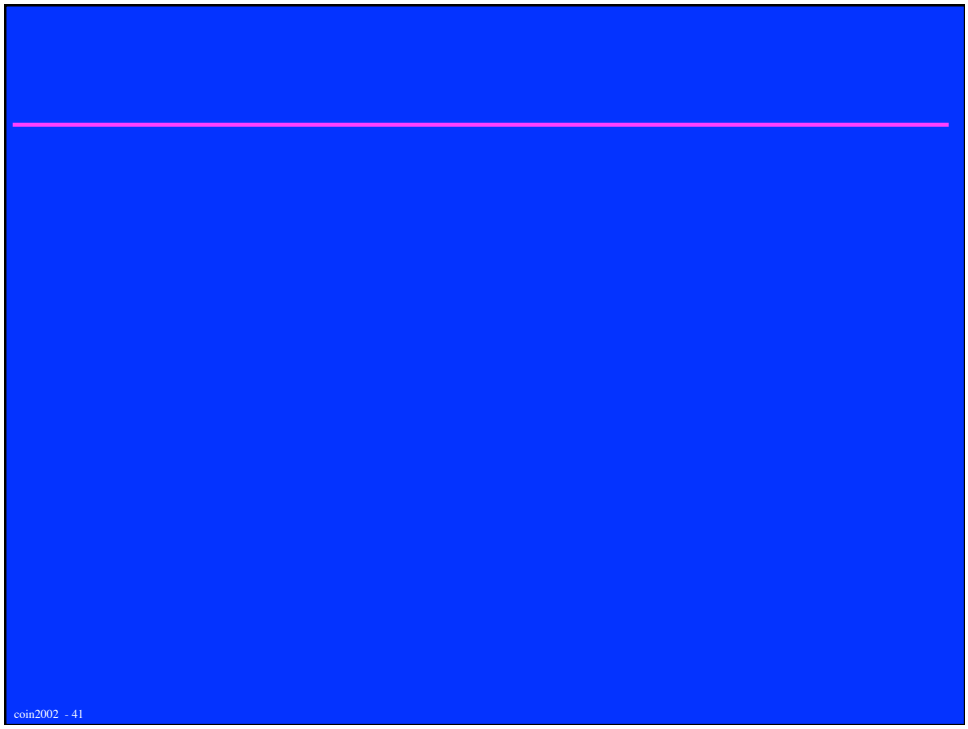
- to the Internet a lambda switched optical network is another link layer
  - not an end-to-end circuit
- could be a point-to-point link between routers
- different case for optical packet switched networks
  - lots of that at this conference
  - not "tomorrow" but I'd like to install some before I retire

IP Packets & Link-Level Frames

- IP end-to-end
  - an IP connection may go over multiple link types
    - Ethernet, DSL, wireless, token ring, POS, TDM etc
- can not assume single e2e link technology
  - one of ATM's problems
  - not the last networking technology
  - need to maintain the isolation that layering provides
- the IP packet format can not be changed to support just one link type (or even multiple link types)
  - this has been proposed for wireless and some other links

Summary

- the Internet is a switched packet inter-network
  - rides above L2 networks
  - including switched ones
- optical networks are key L2 support networks
  - including switched L2 frame optical networks
- e2e circuits are not the Internet
  - e.g., switched lambdas
  - can be very useful to support data services but not Internet
- COIN + PS 2002 is showing us the future
  - and it is exciting


coin2002 - 41