An IETF Insider View

TranSwitch
April 15, 2002
Scott Bradner
sob@harvard.edu

- missing
- next gen signaling

Topics

- work being done in IETF working groups
- that seemed possibly related to TranSwitch
- no particular order
- but start out with some history
  - to be sure we are in sync

History

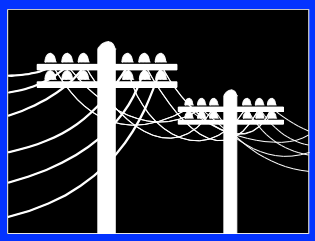
- start with history
- because we have been here before
- because of lessons not learned
- gives hints about IETF (or least my) bias

Background and History

- historical competition between circuit- and packet-based network designs
  - circuit: phone net, SNA, ATM, frame relay, MPLS, switched optical . . .
  - packet: XNS, IPX, AppleTalk, CLNP, IP
- historical competition between smart and stupid networks
  - smart: phone net
  - stupid: Internet
- layers get confusing
  - layers 1, 2, 3 & 8 interact

Circuits

- path through network to destination
- set up before data can be sent
- removed after transfer completed
- all data follows same path through the network
- service requirements can be used in path setup process
  - e.g., bandwidth, reliability, latency ...
- looks like a wire

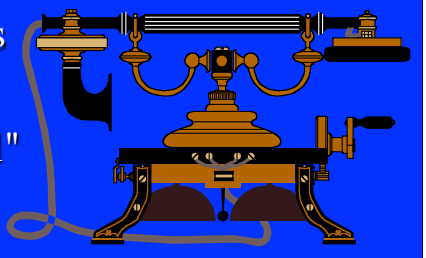


Packets (a.k.a., datagrams)

- self contained chunk of data
- "self contained" in that:
  - it includes delivery & sender addresses
  - may be part of a sequence of chunks
    - but forwarding devices in network needs no knowledge of sequence for proper delivery
  - it can include handling hints
- packets sent to closest forwarder (router)
  - which sends packet to next router in the direction of dest.
  - which sends packet to next router in the direction of dest.
- only state in router is direction to send for each dest.

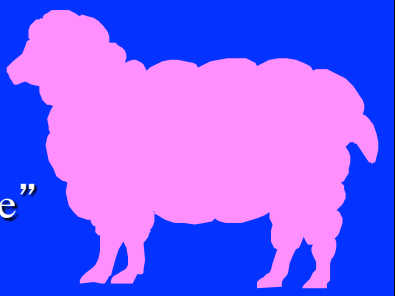
Traditional Phone Network

- circuits
- connection-oriented
- hard state in network devices
- 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

- datagrams
- soft state in network devices
- 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)
- hard to predict developments
  - chaos at "Internet time"



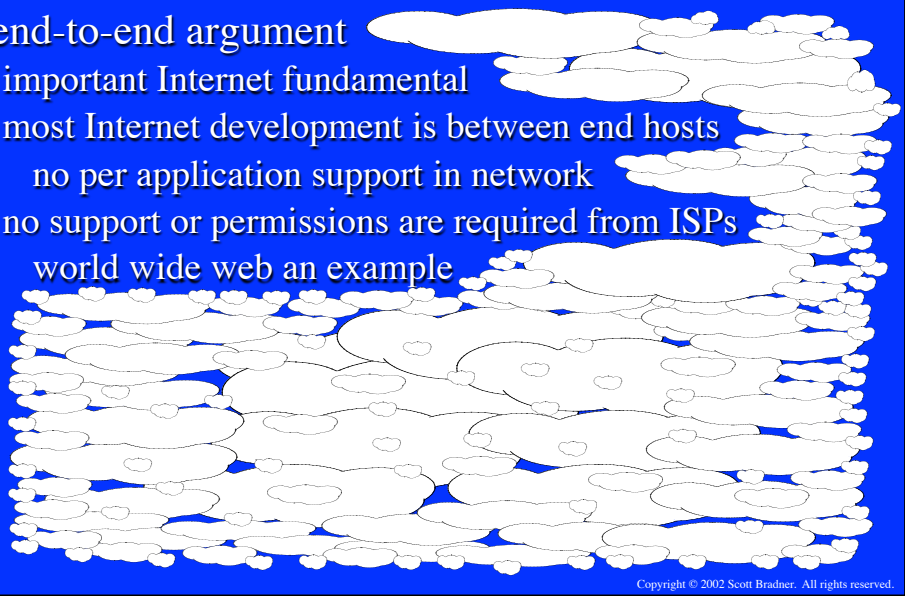
Implications of Packet-Based Networks

- "shortest", rather than "best" path used
- paths through network are not stable
  - they change based on
    - link failure, traffic engineering, routing instability
- impacts QoS
  - can not reserve resources
  - unpredictable QoS
- access control harder
  - e.g. tracking down DoS attacks
- little central control

!QoS

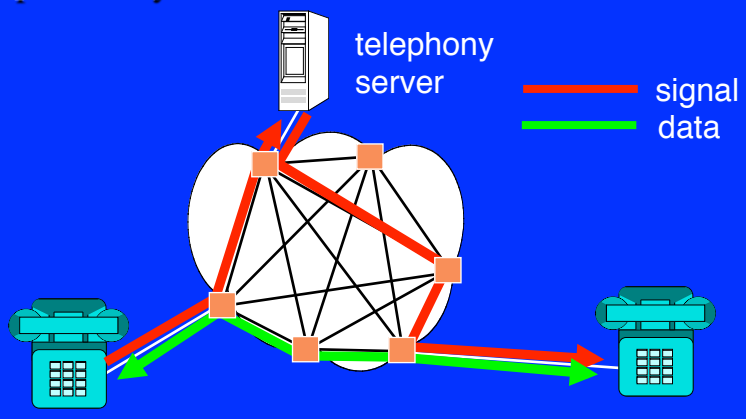
Internet Architecture

- end-to-end argument
  - important Internet fundamental
  - most Internet development is between end hosts
    - no per application support in network
  - no support or permissions are required from ISPs
    - world wide web an example



Internet Architecture, contd.

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

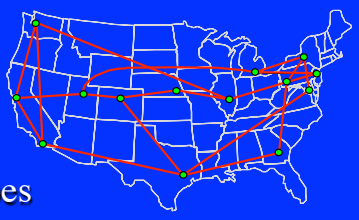


Internet Architecture, contd.

- service provided by 3rd parties - not only by ISPs
- different from phone world
- a quote from an IETF 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
```


Circuits in the Internet



- do not seem to go away
- used for traffic engineering
  - city-pair pipes
  - maybe class of service city-pair pipes
- 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
    - would have had to bet the last networking technology!

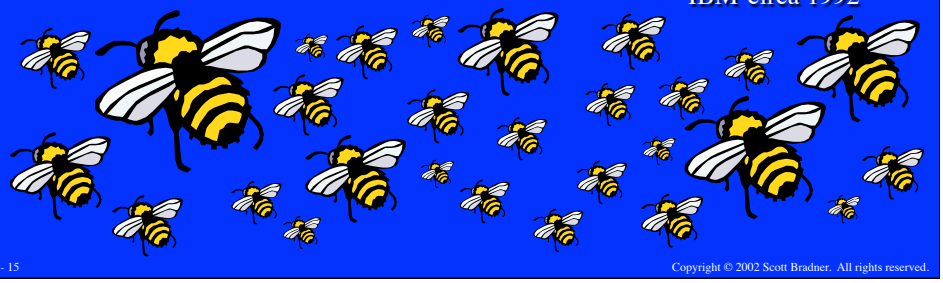
Conceptualization Problem

- 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



Tweaking Circuits

- Internet is getting dynamic underlying circuits

  ATM SVPC, MPLS, switched optical, ...
- how should routing interact?

  which side should be in control
- what is impact of lower layer healing?

  in Internet healing is now at level 3 - but seen as slow

  speed up level-3 healing or use level-2?
- IETF working on a common control plane

  ccamp working group

IP as a Common Bearer Service

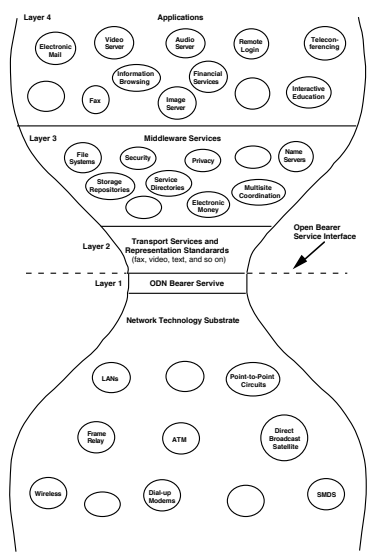


FIGURE 2.1 A four-layer model for the Open Data Network

From: Realizing the Information Future

IP As Common Bearer Service, contd.

*"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

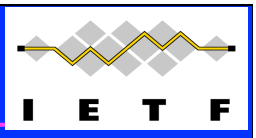
IP As Common Bearer Service, contd.

- but what should it bear?
- just because you **can** get everything to run over IP, **should** you?
- a LAN is a reasonable concept
- a level 2 access network can make sense
- broadcast HDTV over IP may not
- phone calls?
- videoconferences?
- L2 networks (ATM, Frame Relay...)

$$\frac{\text{everything}}{\text{IP}}$$

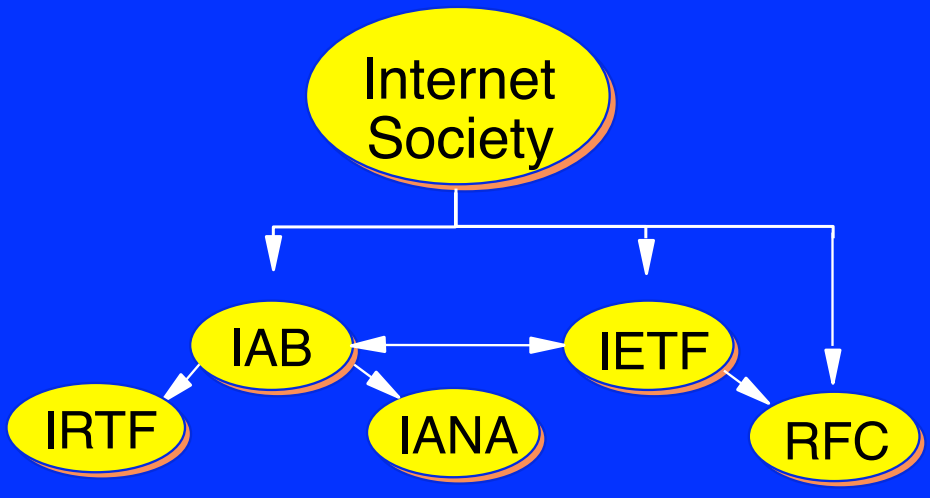
The IETF

IETF www.ietf.org



- Internet Engineering Task Force
  - founded: 1986
  - members: none - but individuals participate
  - decisions: rough consensus of working group to IETF last-call to IESG
  - areas: "on the wire"
    - but now have sub-IP area and technologies
  - structure: working groups grouped into areas managed by IESG with advice from the IAB
  - access: open access to all working documents and RFCs
  - motto: "***rough consensus and running code***"

Top Level View of Organization



Working Groups

- this is where the IETF primarily get its work done
  - on mailing list
  - face-to-face meetings focused on resolving issues (ideally)
- working group focused by charter agreed between chair and area director
  - restrictive charters with milestones
  - working groups closed when their work is done
- working groups organized into Areas
- Areas managed by Area Directors (ADs)
- ADs: Internet Engineering Steering Group (IESG)
  - standards approval body

IETF Areas

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

Sub-IP Area

- area directors

  Scott Bradner <sob@harvard.edu>

  Bert Wijnen wijnen@vnet.ibm.com

- working groups:

ccamp	Common Control and Measurement Plane
gsmp	General Switch Management Protocol
ipo	IP over Optical
iporpr	IP over Resilient Packet Rings
mpls	Multiprotocol Label Switching
ppvpn	Provider Provisioned Virtual Private Networks
tewg	Internet Traffic Engineering

Transport Area

- area directors

  Scott Bradner <sob@harvard.edu>

  Allison Mankin <mankin@east.isi.edu>

- working groups:

avt	Audio/Video Transport
diffserv	Differentiated Services
enum	Telephone Number Mapping
ieprep	Internet Emergency Preparedness
ippm	IP Performance Metrics
ips	IP Storage
iptel	IP Telephony
issll	Integrated Services over Specific Link Layers
malloc	Multicast-Address Allocation
megaco	Media Gateway Control
midcom	Middlebox Communication
mmusic	Multiparty Multimedia Session Control
nfsv4	Network File System Version 4
nsis	Next Steps in Signaling
pilc	Performance Implications of Link Characteristics
pwe3	Pseudo Wire End-to-End Emulation
rmt	Reliable Multicast Transport
rohc	Robust Header Compression
rserpool	Reliable Server Pooling
seamoby	Context and Micro-mobility Routing
sigtran	Signaling Transport
sip	Session Initiation Protocol
sipping	Session Initiation Protocol Investigation
spirits	Service in the PSTN/IN Requesting InTernet Service
tsvwg	Transport Area Working Group

Multiprotocol Label Switching (MPLS)

- add tags to IP packets at ingress routers
  - tags used by **MPLS** switches in forwarding decision
  - **direct traffic along a path that routing would not take**
  - tags stripped at egress
- started as a traffic engineering (TE) tool
  - direct inter-POP traffic along a path with capacity
  - was performance enhancement idea at one point
- now being seen as a QoS technology and more
- another net-head vs. Bell-head difference
  - net-head: TE using **RSVP**-based signaling
  - Bell-head: MPLS as ATM with variable length cells
    - using LDP & CR-LDP

MPLS, contd.

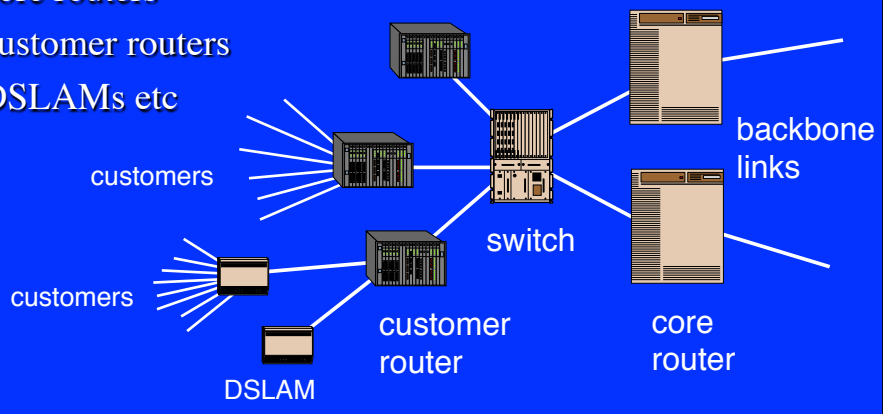
- much confusion over **MPLS** applicability
  - a long way from TE
  - **MPLS** on the LAN?
- seen as a way of converting Internet to circuit base
  - to fix QoS, security, charging, management, . . .
  - "a local gravity well"
  - an IP version of ATM?

MPLS != ATM

- ATM lesson not learned
- remember that datagrams do work

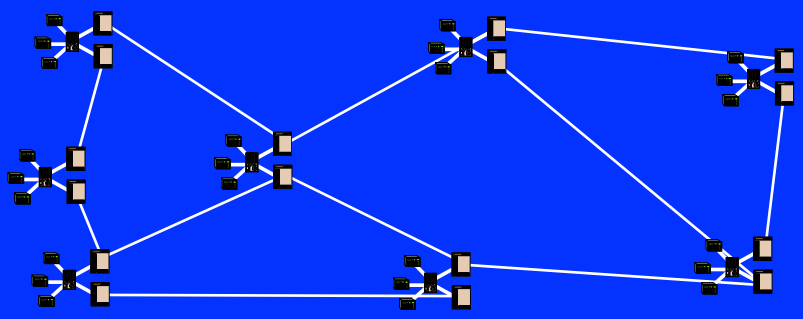
ISP POP

- point of presence (POP)
- parts
  - core routers
  - customer routers
  - DSLAMs etc



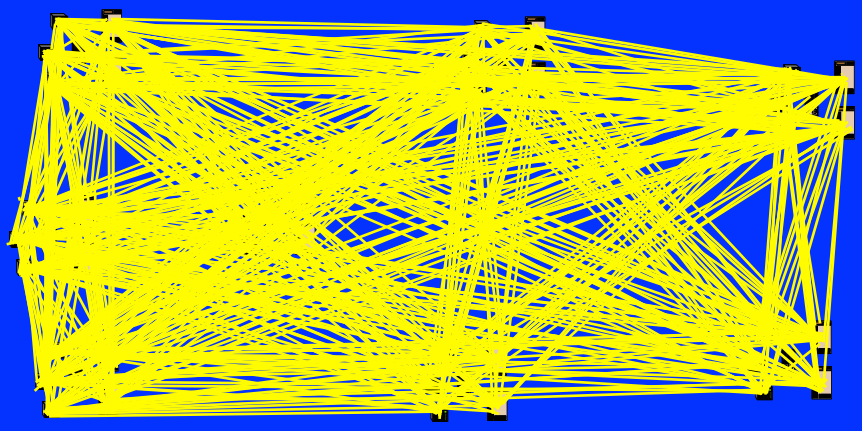
ISP Network

- multiple POPs
- interconnected with backbone links
  - not full mesh



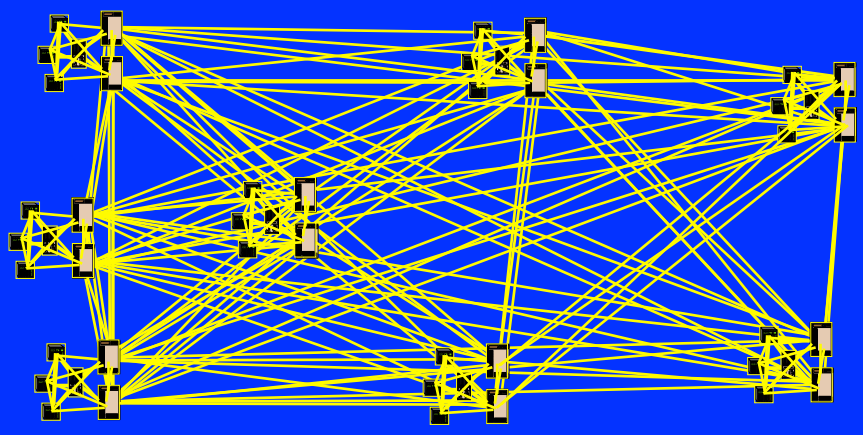
ISP Logical Network

- logical full mesh (BGP requirement)

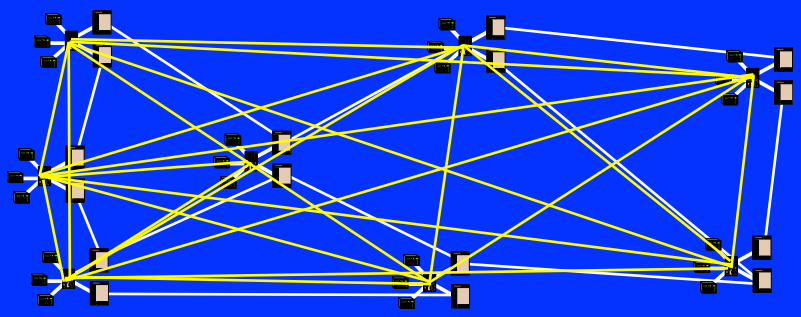


ISP Logical Network

- logical mesh - with confederations

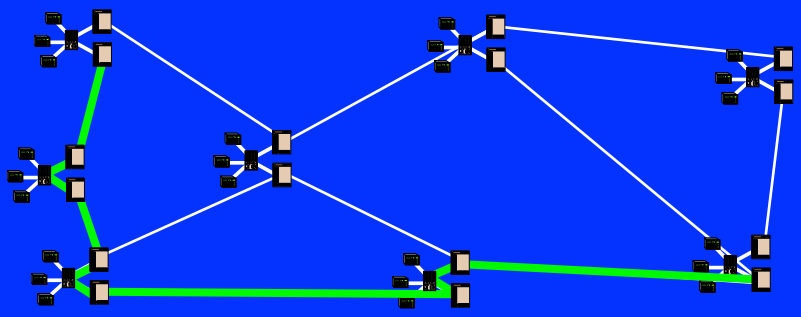


ISP Forwarding Network



Traffic Engineering

- e.g. MPLS

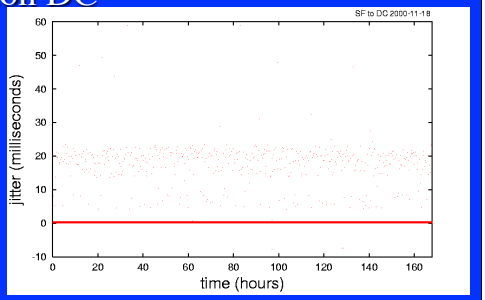


More MPLS

- MPLS can support hierarchical tags
  - e.g. for trunks of individual LSPs
- for some this means MPLS can be used for **MANY** things
  - e.g. VoMPLS - individual phone calls gathered in trunks
    - VPNs (with QoS)
- MPLS == wire
- but do you need wires?
  - is "just IP" good enough?

Data Point

- Steve Casner et al, NANOG presentation
  - http://www.nanog.org/mtg-0105/casner.html
- experiment on active ISP backbone
  - San Francisco to Washington DC
  - POP to POP
  - 1Mbps average data rate
  - 15 5 to 7 day trials
  - results:
    - 99.99% availability
    - jitter < 1ms for 99.99% of packets sent



Optical Network Control: UNI

- **U**ser-**N**etwork **I**nterface
  - ATM terminology
- let customer request optical connection through optical service provider (OSP) cloud
  - e.g. "give me an OC48 to San Francisco"
- multiple approaches
  - new protocol
  - **RSVP**-based signaling
  - **LDP**-based signaling
- is there a business case?

Optical Network Control: NNI

- **N**etwork-**N**etwork **I**nterface
  - ATM terminology
  - between carriers
- no specific standards activity yet
- some "interesting" business issues to deal with first
  - how does the money flow?

Optical Network Control: In Cloud

- how to tell network components to setup or modify an optical path
- IP Optical (ipo) WG
  - optical control plane should be IP-centric, utilizing IP-based protocols for dynamic provisioning and perhaps restoration of lightpaths within and across optical sub-networks
- a number of proposals
  - RSVP, MPLS, OSPF, IS-IS
  - may use common control plane
- under development

Common Control Plane

- develop a common approach to controlling lower layer functionality in IP networks
  - where the lower layer is controllable
    - e.g. ATM, Frame relay, MPLS, switched optical
- deal with interaction with routing system
- GMPLS

Virtual Private Networks (VPNs)

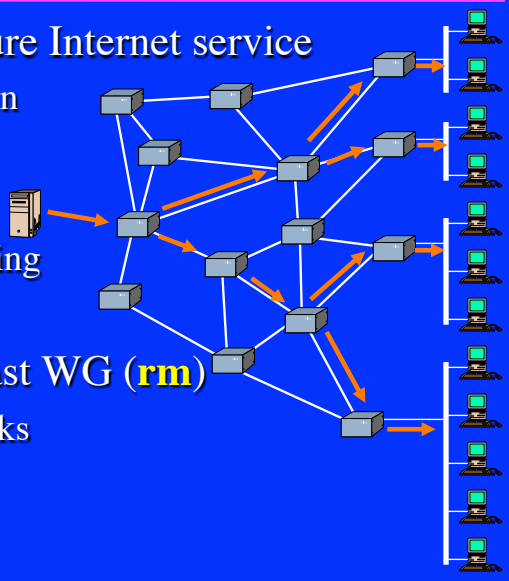
- many meanings for the term
  - end-to-firewall, end-to-end **IPSec**
  - firewall-to-firewall, CPE-to-CPE, POP-to-POP **IPSec**
  - **MPLS, L2TP**
- IETF: **P**rovider-**B**ased **VPN**s (ppvpn)
  - standardize a framework and one or more sets of mechanisms for supporting network-based IP virtual private networks
  - 3 types
    - level-2 VPNs
    - virtual router (VR) VPNs
    - BGP-MPLS VPNs

Pseudo Wires

- Pseudo Wire Emulation Edge to Edge (pwe3) WG
- emulate L2 "wires" over IP & MPLS
  - frame relay, ATM, TDM, SONET, Ethernet, MPLS, ...
- food fight over emulation quality
  - "Turing test"?
    - or
  - define what user will get
- IP networks can have low jitter
  - see Scott Shanker's NANOG presentation
  - < 1ms between Washington DC and San Francisco

Multicast

- seen as most basic future Internet service
  - audio & video distribution
  - news services
  - stock quotes
  - audio & video conferencing
  - general data distribution
- IETF **R**eliable **M**ulticast WG (**rm**)
  - developing building blocks
  - no one technology



Multicast, contd.

- traditional multicast is multi-to-multi
  - long been a future, but many deployment, business and operations issues
- IETF **S**ource-**S**pecific **M**ulticast (**SSM**)
  - one-to-many
  - change multicast group label to be 64-bit "**S,G**"
    - 32-bit sender IP address "**S**"
    - 32-bit multicast group relative to that sender "**G**"
    - reserve 232/8 for **SSM**

sender	group

SSM Advantages

- address allocation **16,777,216 groups**
  - a /8 per sender, no synchronization required
- finding the sender
  - part of group "name" **mcast.cnn.com, h-news**
  - just send packets via unicast to sender
- manageability
  - router-enforced single sender model
  - protects network
  - understandable billing model (sender pays)
- i.e. might actually make sense

SSM, contd.

- uses modified version of **IGMPv3** on LAN
  - changes some processing rules - ignore some messages for addresses in **SSM** address range
- host uses (**S,G**) pair
  - how host knows is outside the scope of the protocol
- uses modified version of **PIM-SM** off LAN
  - change processing rules for addresses within 232/8

Quality of Service (QoS)

- QoS controls seen as critical (by some) for future converged Internet
  - a big net-head vs. Bell-head difference
  - over-provision vs. complex controls
  - should there be busy signals on the Internet?
- QoS requirements coming from many places
  - ITU-T, TIA, QoS Forum, ETSI, IEPS, . . .
- too much focus??

QoS Technology: per-flow

- IETF Integrated Services (intserv) WG
  - Resource Reservation Protocol (RSVP) signaling
  - intserv services: Guaranteed & Controlled Load Service
    - renamed by the ITU-T Y.iptc to:
      - "delay sensitive statistical bandwidth capability"
      - "delay insensitive statistical bandwidth capability"
  - intserv offers link-level per-flow QoS control
  - RSVP offers signaling for intserv
    - also used as a general signaling protocol - e.g. MPLS
    - new RSVP extensions WG

QoS Technology: class-based

- IETF Differentiated Services (diffserv) WG
  - class-based QoS
  - packets marked at network "edge"
  - routers use markings to decide how to handle packets
  - four services
    - best effort - normal Internet traffic
    - 7 precedence levels - prioritized classes of traffic
    - Expedited Forwarding (EF) - leased line like service
    - Assured Forwarding (AF) - 4 queues with 3 drop classes
  - requires edge policing - technology not yet defined

QoS Technology: Other Ideas

- a number of similar ideas from traditional telcom
- map flow-based QoS into a circuit of some type
  - MPLS Label Switched Paths
  - ATM VCs
  - optical lambdas
- the old circuits vs. packets fight
- could make sense for trunks

IPv6

- IETF ipngwg working group
- technology standards done - many implementations
- waiting on uncle Bill
- cell phones and China may show the way

  but routing is not any better

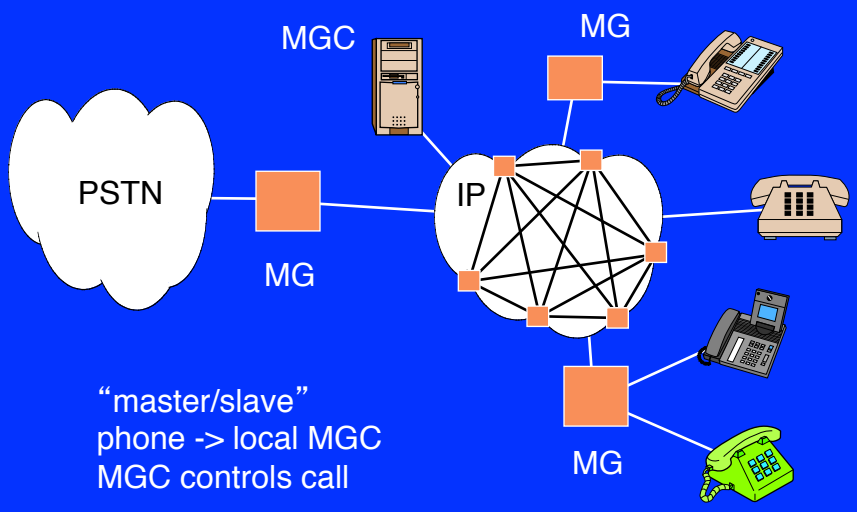
Telephony

- telephony cntrl: MGCP, megaco/H.248. H.323, SIP
- phone number resolution: enum
- wireless: WAP, SeaMoby, 3G, rohc
- settlements: ITU-T
- PSTN/IN control: pint, spirits
- finding PSTN gateways: trip
- lawful interception: raven, ETSI, T1

Telephony Control: Phone Model

- two protocols
  - **MGCP** - Media Gateway Control Protocol - RFC 2705
    - informational RFC: **not an IETF standard**
    - well supported in industry - including cable modems
  - **megaco/H.248** - joint IETF/ITU-T effort
    - in RFC Editor's queue (Aug '00), also ITU-T publication
    - MGCP was an input to the effort
- break up phone switch into controller and gateways
  - "looks" like phone switch
  - a.k.a. softswitch (but softswitches can often do much more)
  - MGC is in control

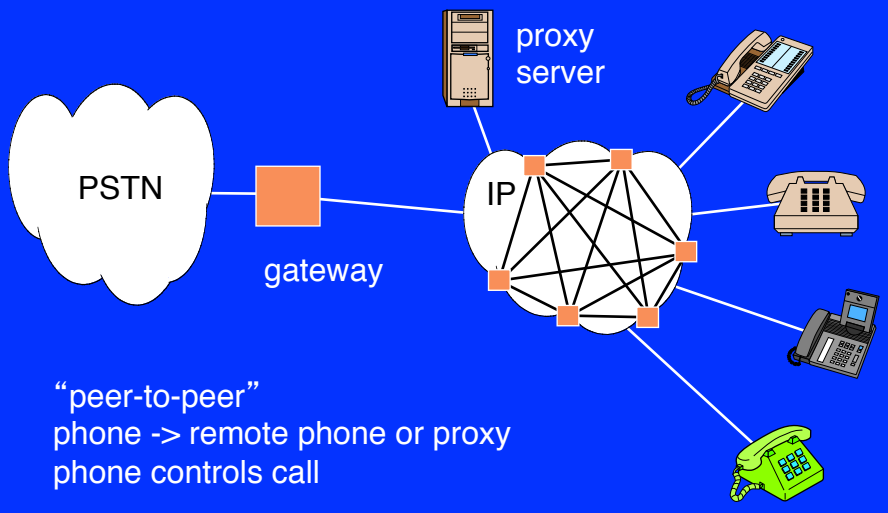
Telephony Cntrl: Phone Model, contd.



Telephony Cntrl: Internet Model

- two protocols
  - **H.323** - ITU standard
    - e.g. net meeting
  - **SIP** - **S**ession **I**nitiation **P**rotocol - IETF Proposed Standard
    - RFC 2543 (new version just Oked)
- interworking effort underway
- Internet model of smart edges
  - light-weight servers in network (proxy, forwarding)
  - do not have to be run by connectivity provider

Telephony Cntrl: 'Net Model, contd.



"peer-to-peer"
phone -> remote phone or proxy
phone controls call

The Importance of Phones

- big issue in IETF development of telephony technology for IP networks
- phone people assumed that phone traffic would have precedence over all other use
  - IETF did not agree
- particular issue in responding to congestion
  - everyone thinks the other guy should back off

I'm more important! I'm more important!

I'm more important!

I'm more important!

Emergency Response

- much interest in the 9/11 aftermath
- lots of 'make the Internet do what the phone net does'
  - call special area code & enter credit card # gets priority processing (but not preemption)
  - but the Internet does not block, just degrades
  - also - how about other Internet-based services?
    - web servers, emal etc?
- IETF ieprep WG
  - mostly proposing solutions w/o defining problems

Finding Things Using Phone Numbers

- **T**elephone **Num**ber Mapping (enum) - IETF WG
- IETF working group - RFC 2916
  - input: an e.164 style phone number
  - output: one or more URIs
- uses domain name (DNS) system
  - for phone number of **+ 46 8 9761234**
  - look up **4.3.2.1.6.7.9.8.6.4.e164.arpa**
- significant political issues
  - who controls per-country mappings?
  - who controls or runs the mappings for a user
- is privacy a problem?

PSTN / IN Control - IETF WGs

- **P**STN and **I**nternet **In**ternet**t**working (pint)
  - Internet control of PSTN services
  - e.g. click-to-call

Call Scott

- **S**ervice in the **P**STN / **I**N **R**equesting **I**n**t**ernet **S**ervice (spirits)
  - notification of PSTN events to Internet servers
  - e.g. Internet call-waiting
- call processing language: CPL
  - tell phone switch what to do
- interesting security and accounting issues

Scott is calling
- hang up on him
- take message
- voice mail
- forward to joe
- accept call

Finding PSTN Gateways

- **T**elephony **R**outing over **IP** (TRIP) - IETF WG
- Internet routing protocol to find PSTN gateways
  combination of **BGP**, **IS-IS** and **OSPF**
- **TRIP** is used by location servers (LSs) to exchange phone reachability information
  LS advertises phone numbers it can reach
  e.g. country, local area, or organization
- telephony signaling protocol independent
  i.e. supports **SIP** & **H.323**

Lawful Interception

- IETF www.ietf.org/mailman/listinfo/raven
  "raven" discussion in IETF resulted in a decision to not mandate intercept features
  technical and logical reasons
  e.g. no consistent international definition
  RFC 1804

IP Storage

- **IP S**torage (ips) Working Group
  - iSCSI - run SCSI over IP networks
  - FC over IP - run Fiber Channel over IP networks
- original idea was for storage area networks
  - connect servers and storage systems
  - restricted geography
- but once something runs over IP it is hard to restrict
  - WG required to address IPS in all environments
  - pushback on security requirements

Signaling In the Internet

- end system signaling to request network services
  - QoS, security, traffic engineering
- RSVP is the current IETF signaling protocol
  - soft state
  - used for QoS & MPLS (so far)
  - in Windows to request difserv codepoint
- nsis (next steps in signaling) WG looking at what to do next
  - could be a revision of RSVP
    - e.g., remove multicast complexity

Summary

- the IETF is a busy place
  - we have only looked at a few of the 130ish WGs
- many other SDOs
- lots going on that looks like it could be of interest to TranSwitch