SUNet
the Stanford University Network

A Presentation for Networkers '99
on the Development of
Stanford's Backbone Network

Ron Roberts <rgr@stanford.edu>
Wayne Sung <wsung@stanford.edu>

Session 901

A Brief Overview of SUNet

- **Originally a research project**
  - **Begun by EE, CS and Medical grad students**
- **Networking staff assembled in 1983**
  - **As a unit of Center for Information Technology**
- **Centrally funded since 1984**
  - **Board of Trustees made original funding grant**
- **Cable plant project in 1985**
- **NS responsible for operations since 1986**

Backbone Technology

- 3Mb Ethernet from Xerox PARC, 1979
- Large bridged Ethernet—DEC bridges, 1986
- Campus-wide FDDI ring, 1990
- Multiple FDDI rings from Cisco 7513, 1995
- The “cube” backbone —6 7513 routers, 1998
  - 24 switched 100Fx router interconnects
  - Dual OC3 POS to border GSR router
  - OC12 POS to CalREN-2
- Initial fiber install in 1985—12 MM
- Latest fiber trunks: 96 MM/48 SM

Cable Types

- Underground coax (1/2 inch “hardline”)
  - 3-Mb Ethernet used 75-ohm (CATV) cable
  - 50-ohm cable with custom built N-connectors
- Fiber optic cables
  - First cross-campus fiber, 1985, was 12 MM
  - Interhub went to 36 MM, buildings 12 MM
  - Now most buildings get 24MM/12 SM
  - Interhub runs get 48/24 or even 96/48

In-Building Wiring

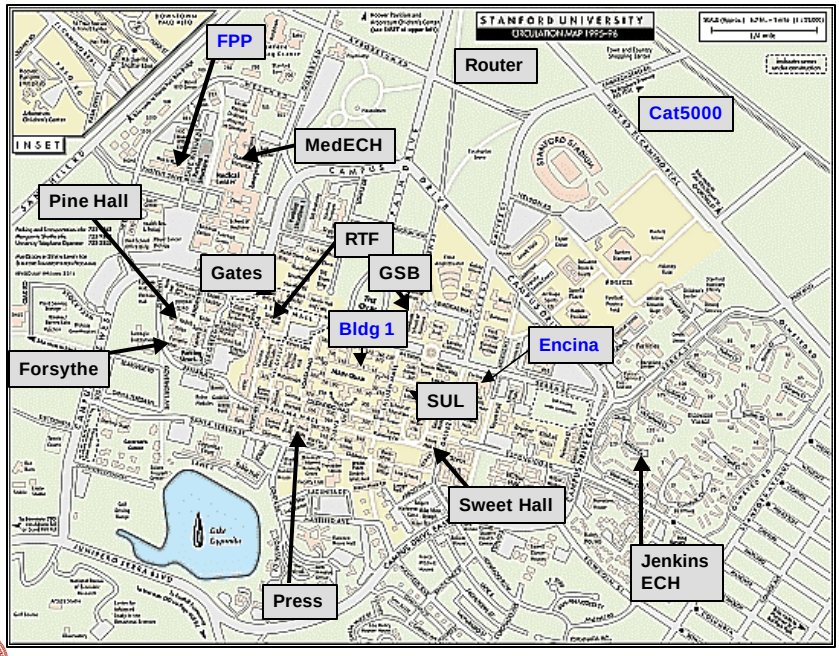
- Complete spectrum down to 3-pair RJ-11
- New buildings have wiring closets
  - At least 2 Cat-5 network jacks + 2 phone jacks
  - Plan to experiment with fiber to the desktop
- Centrally funded wiring improvement
  - Incremental project approved in 1997
  - Requires matching local funds for activation
- Still lots of 10Base5 and 10Base2

Routers aka Gateways

- First “gateway” was PDP-11 based
  - Had homebrew 3Mb Ethernet interfaces
- CS had multibus-based 68000 SUN-1 CPU
  - Installed in “the blue box” multibus chassis
    - Linear power supply made it quite heavy!
  - 3-Mb multibus Ethernet boards from PARC
  - Later 10-Mb boards commercially available
  - Typical router had 4 or 5 interfaces

Internet Access

- Stanford was a node on original ARPAnet
  - Computer Science had IMP-connected hosts
- Stanford hosted BARRNet regional network
  - NSFNET node for Northern California
  - BARRNet bought by BBN in 1994
  - Continued as main commodity connection
- Stanford is charter member of CalREN-2
  - Part of main 4-node Bay Area SONET ring



Research Era Topology (Pre-SUNet)

- Gateways were used between segments
  - A gateway crash would segment the network
  - No bridges, few repeaters
  - But 3-Mb segments could be quite long
- Used the PUP protocol from PARC
  - Routing with GWINFO
  - CS designed MEIS for TOPS-20 Ethernet
  - Access over Ethernet to ARPAnet hosts

Research Era Access (Pre-SUNet)

- Multibus Serial I/O cards designed in CS
  - Built TIPs with same CPU board as gateways
  - Allowed terminals to access hosts over the net
- LOTS (Low Overhead TimeSharing)
  - Several TOPS-20 systems on Ethernet
  - Multiple terminal clusters with TIPs on net
  - Provided ubiquitous student computer access

Router Development

- CS/EE joint project to route TCP/IP
  - Goal to connect Stanford net to ARPAnet
  - 68000 code rewritten
- PUP network numbers mapped into 36/8
  - Second octet was subnet
  - Subnet mask became 255.255.0.0
- Decided they had a good product
  - Left to build a company called Cisco Systems

The SUNet Project

- Board of Trustees approved in 1984
  - To Center for Information Technology (CIT)
  - The Interbuilding Signal System—cable TV
- Prototype (proof of concept)
  - Spiralnet 1/2-inch 75-ohm cable around campus
  - Spare cables were quickly claimed by EE/CS researchers for 3-Mb Ethernet

SUNet Design Parameters

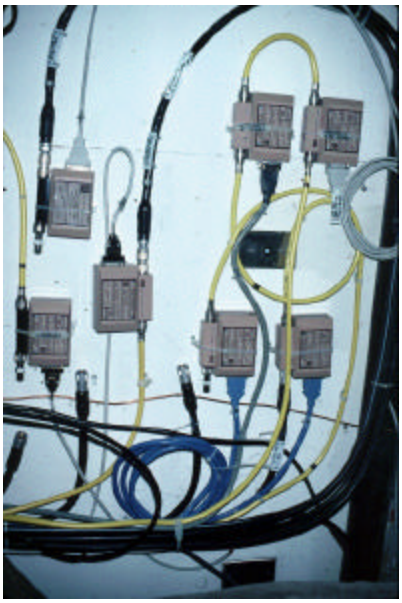
- TCP/IP only supported backbone protocol
- Networking supports to building entrance
  - Local Network Administrators in buildings
  - Network consultants assist LANs
- Other protocols running but unsupporte
  - AppleTalk tunneled over IP (K-box, Fastpaths)
  - Medical Center routes IPX/AppleTalk
  - DECNET (user-supported) almost gone

SUNet Addressing and DNS

- NS to manage IP addresses, names centrally
- NetDB to assign addresses, generate DNS
  - Stanford.edu is flat domain, no sub-domains
  - Distributed access to NetDB for LANs
- Transition from 36/8 to 171.64/14
  - A CIDR developer, Vince Fuller, at BARRNet
  - Before variable subnets, needed more addresses
  - Will return 36/8 to ARIN by July 2000

SUNet Cable Project 1985

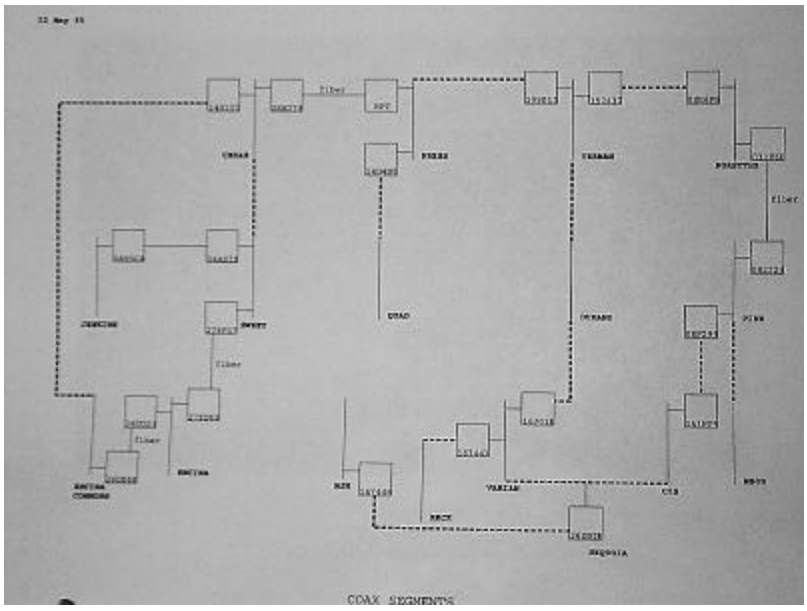
- **Original design: two-way cable TV system**
  - **To provide service to all campus buildings**
  - **Campus computer center wanted remote terminal access from across campus**
- **50-ohm underground cables**
  - **Added by NS just before bid (after design)**
  - **Two cables paralleling most cable-TV trunks**



50-ohm Underground Cables

Bridged Ethernet Backbone

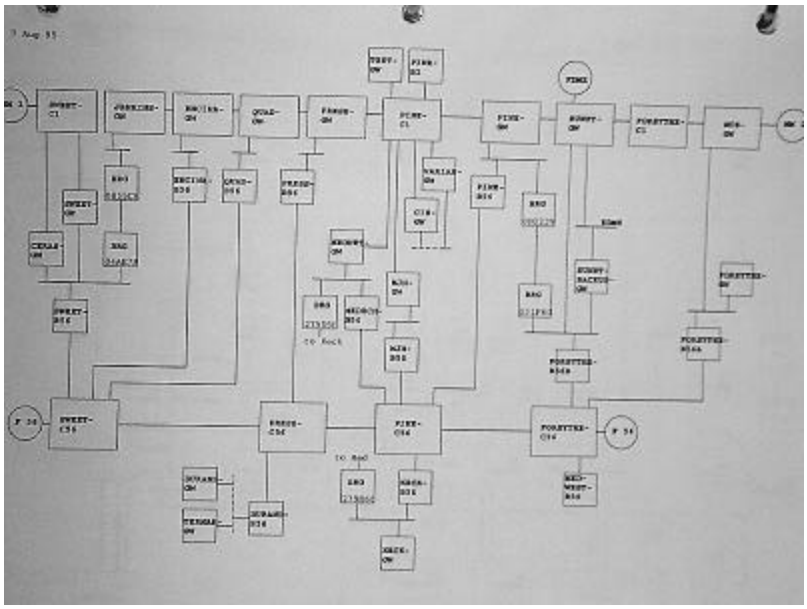
- After cable project completed in 1985
  - Began using the new DEC LanBridge 100
- Multiple loops—relied on spanning tree
- Provided single backbone for all routers
  - But latency from multiple bridge hops
  - Originally called LinkNet, later the EtherSpine
- Continued to use 68000 multibus routers
  - Stanford hardware with software from Cisco



Bridged Ethernet Topology

EtherSpine Enhancement

- **Equipment donation from Digital (DEC)**
  - **FDDI Concentrators**
    - **Including single-mode DAS backbone modules**
  - **FDDI—Ethernet bridges**
    - **Single port and three-port models**
- **Overlayed bridged Ethernet with FDDI ring**
  - **Increased speed/capacity**
  - **Decreased hop count**



FDDI Etherspine Overlay

Router Consolidation

- Move to FDDI backbone—early 1990s
  - Built with AGS+ routers
  - Etherspine continued as backup “safety net”
  - More subnets from each router
  - Fiber feed using 10FL to building bridge
  - Network nodes with better power and access
- Could invest in UPS and air conditioning
- Ever-increasing need for fiber plant

Multiple FDDI Rings

- Ring load monitoring showed peaks >60%
- Considered/tested DEC giga-switch
  - Worked well, but concern over multicast limits
- Decided on routed solution
  - Using 7513 with multiple FIPs
  - Backed up by AGS+ routers with dual FDDIs
  - Ethernet back-up net partitioned as well

IGP Transition

- Began using IGRP as IGP in late 1980s
  - GWINFO kept breaking in each Cisco IOS release
- OSPF was deployed by BARRNet
  - After initial debugging, seemed to work well
- Wanted quick failover and variable subnets
  - Computer Science moving into Gates building
  - To move out of 36/8, wanted efficient subnets
  - Problems with failover to Etherspine

Conversion to OSPF as IGP

- Done during Christmas break in 1995
- Accomplished in stages
  - Single ring change during maintenance window
  - Would run OSPF and IGRP simultaneously
    - Then change the "distance" so OSPF preferred
    - If problems could back out in reverse
    - Some strangenesses since IGRP using route table
- Worked well to survive FDDI problems
  - Transition to Etherspine usually transparent

Network Components in 1997

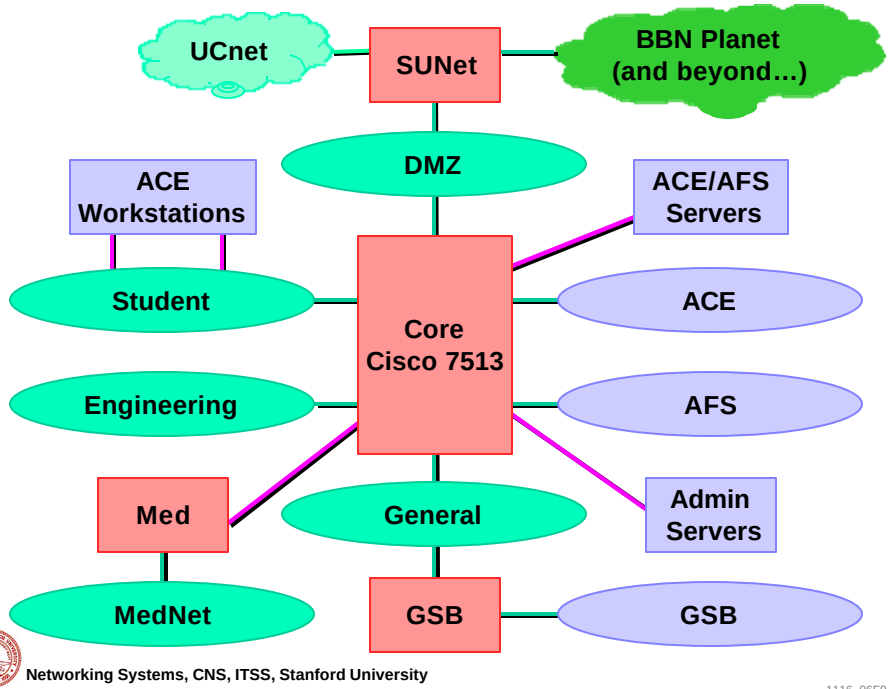
- 49 routers from Cisco
  - 3 7513, 2 7507, 6 7000, 4 7206, 5 4700
  - 7 2501, 1 2514, 1 3000
  - And still 9 AGS+, 10 AGS, 1 CGS
- FDDI concentrators from DEC
- FastEthernet switches: Catalyst® 5000/5500
- ATM switches: LS1010

SUNet Statistics 970914

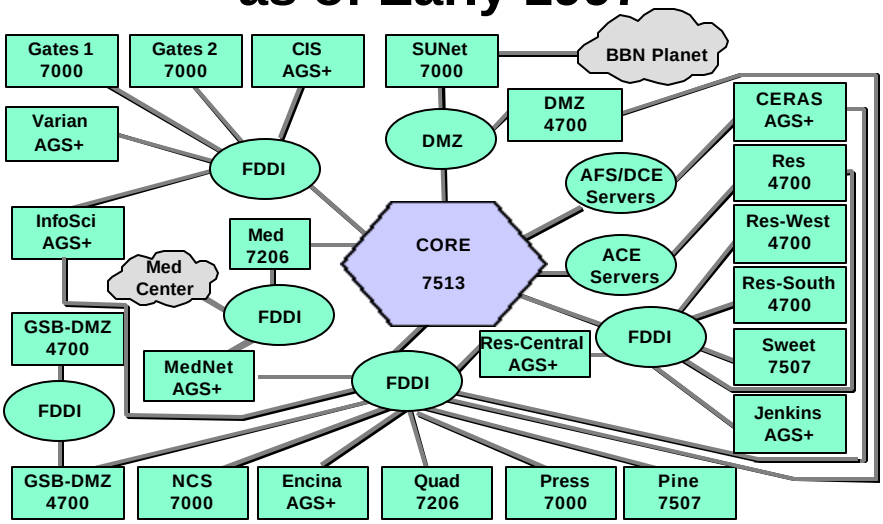
- From NetDB, the network database
  - Generates DNS and BootP tables
- 442 subnets defined, 342 active
- 327 physical subnets
- 240 main campus, 87 Medical Center
- 33866 hosts, 35445 addresses
- 24749 campus hosts, 9117 Medical Center

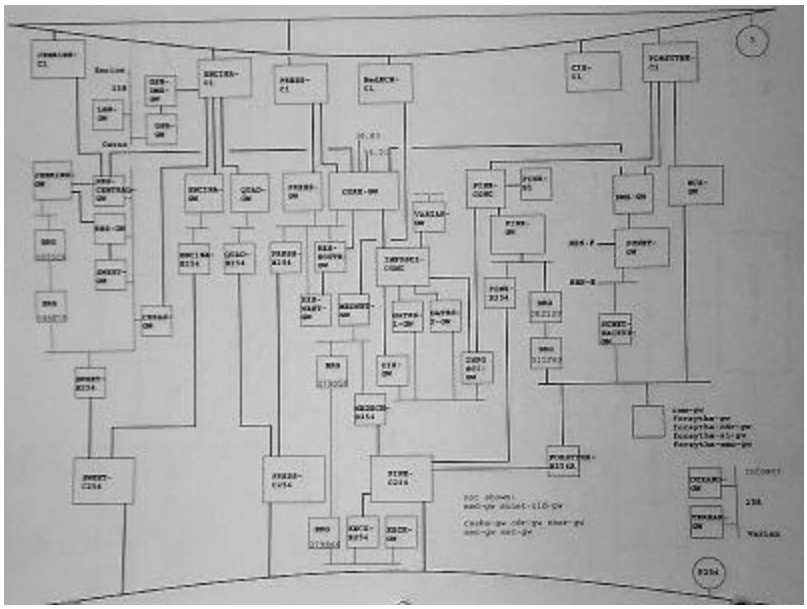
SUNet Statistics

- From NetDB on 1-Jun-99
  - 520 subnets defined, 420 active
  - 48712 hosts, 50323 addresses
- Statistics on 981117
  - 508 subnets defined, 378 active
  - 981117 was: 44034 hosts, 45331 addresses
  - 970914 was: 33866 hosts, 35445 addresses

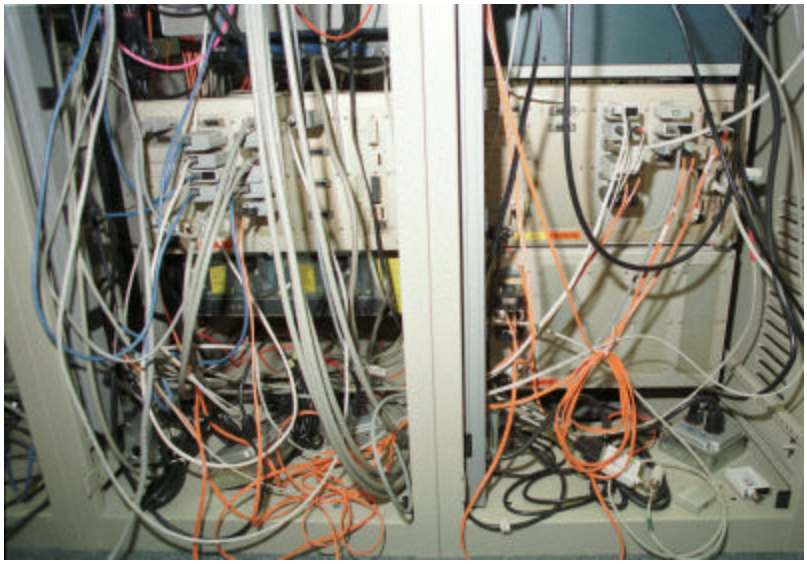


SUNet Backbone as of Early 1997





FDDI/Ethernet Router Topology



AGS+ Routers in the FDDI/Ethernet Era

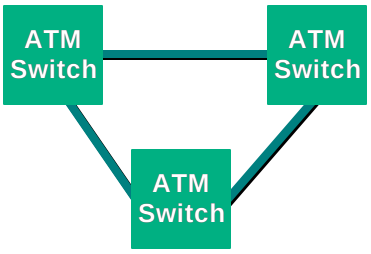
Next Generation Driving Factors

- Migration from Mainframe to distributed computing model for Admin systems
- High-speed (fast Ethernet) user networks
- Higher bandwidth for image transfer, etc.
- Increased reliability as well as capacity
- Old technology no longer supported (AGS)
- Internet-2 and CalREN-2; NSF grant

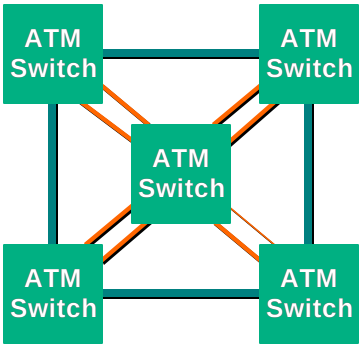
Next Generation Choices

- ATM seemed likely as late as 1996
- New routers were being ordered with ATM
- New Computer Science building provided early implementation experience for both LANE and FastEthernet VLAN trunking
- Sought advice from Cisco Consulting Engineers as part of partnership

Proposed ATM Backbones



3 x OC12



4 x OC12 + 8 x OC3

Concerns with ATM Design

- Redundancy vs. Interface expense
  - ATM interfaces more expensive, less dense
- Configuration complexity
  - Need to build VCs as well as IP nets
- Total throughput
  - Trunks between switches limit total bandwidth
- Staff education on a new technology
  - Many staff years of Ethernet experience

N-Dimensional Mesh Network

- Proposed by Cisco's Roger Beeman
- Designed to have well defined expansion
- Multiple paths provide aggregate capacity
- Significant redundancy
  - Makes individual components less critical
- Networks on corners

N-Dimensional Mesh Routers

- Smaller dimensions can be visualized
  - Cube (3D) and hyper-cube (4D)
- Multiple ways to configure routers
  - 2-way connected routers on edges
  - 4-way connected routers on faces
  - 8-way connected routers in cubes

Router Connections

- 4-way connected routers on faces
  - 3-D (cube): 8 networks, 6 routers
  - 4-D (hyper-cube): 16 networks, 24 routers
  - 5-D: 32 networks, 80 routers
- 8-way connected routers in cubes
  - 4-D (hyper-cube): 16 networks, 8 routers

The SUNet Choice: A Cube

- Attracted by the combination of redundancy and expandable capacity
- Purchased 6 7513s
  - 4-VIP2s with 100FX port adapter each
  - Added 5-port 10FL port adapters on 2 VIPs
    - For legacy Ethernet and services
  - Additional VIPs for FE hotspot connections

SUNet Cube Facilities

- Chose to locate in four facilities
  - Sweet, Press, Forsythe and Pine
    - Spaces under organizational control, 7x24 access
    - Already fiber capacity, UPS and air conditioning
  - Expectation of generator power backup
    - Exists in Press and Forsythe
    - Project to install for Pine by late 1999
    - Sweet's external connection capability installed
      - Generator itself in approval process
      - Mainly aesthetic concerns from University Architect

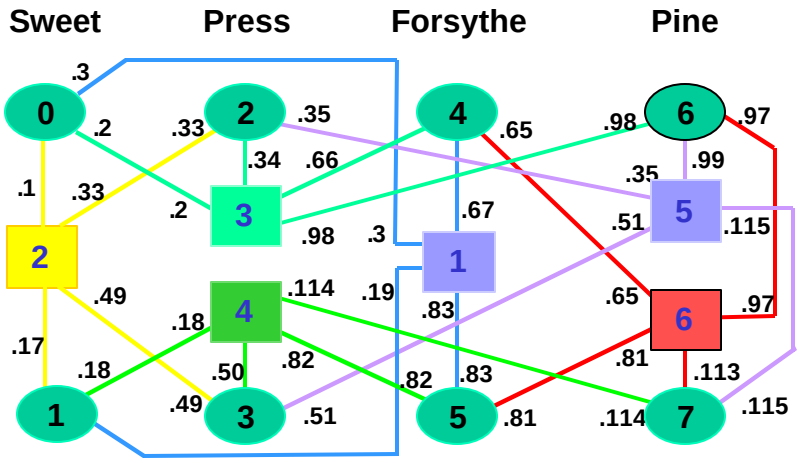
SUNet Cube Corners

- Using Catalyst 5000 for the corners
- First purchased 4 Catalyst 5000s
  - With two 12 port 100FX boards each
  - Each was two corners (using VLANs)
    - One VLAN per 100Fx board
- now using 8 Catalyst 5505s
  - Each with one 100Fx blade
  - Some with 10FL for management net

SUNet Phase 4 Deployment

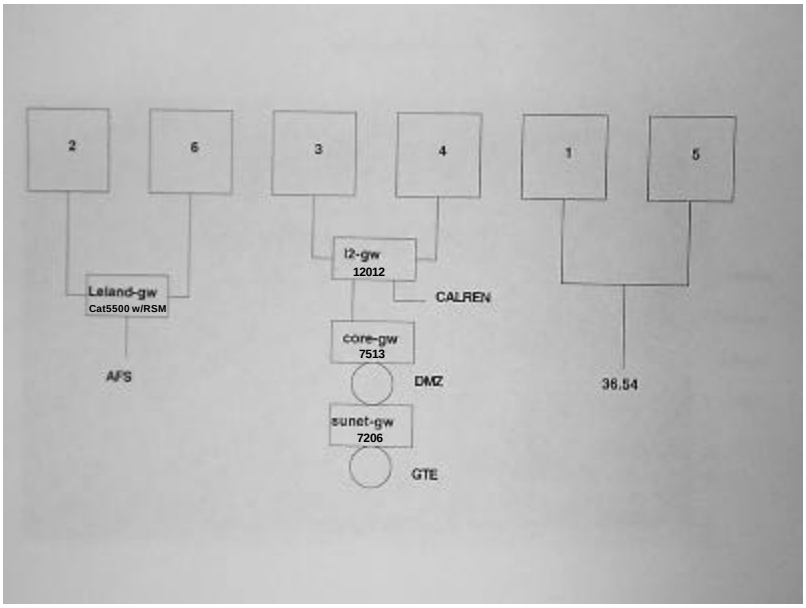
- Core1…Core6 installed Fall 1997
- Moved student residence network and "hotspots" during Christmas quarter break
- First deployed Jenkins, MedECH, Pine and Sweet-RSM (RTF delayed by construction)
- Completion target was end of Summer 1998
- Actually complete in November 1998

SUNet Cube Backbone Layout and Addressing

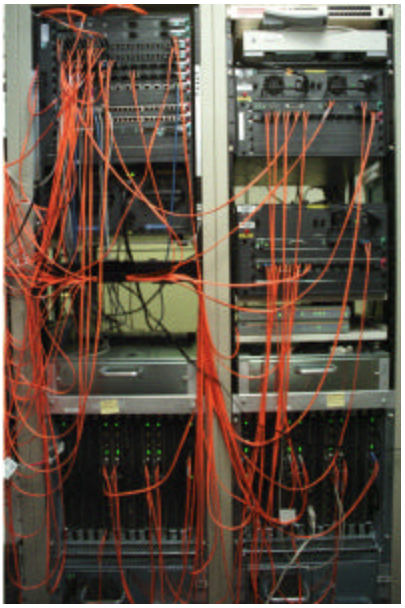


Connections to the Cube

- User routers connected to at least two corner networks, usually opposite corners
- Server nets directly connected to two backbone routers using HSRP <or>
- Server router directly connected to two opposite face backbone routers
- Three “hotspots”—ACE servers, Admin servers and Off-campus/Internet



Core Router—Hot Spot Connections



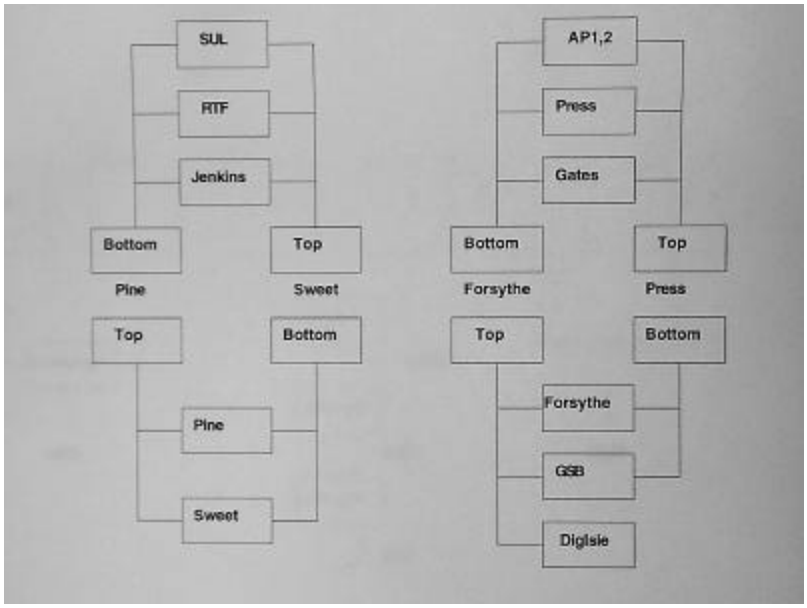
A Cube Corner
2 x 7513
2 x Cat5505
(+ User Router)

User Router Choices

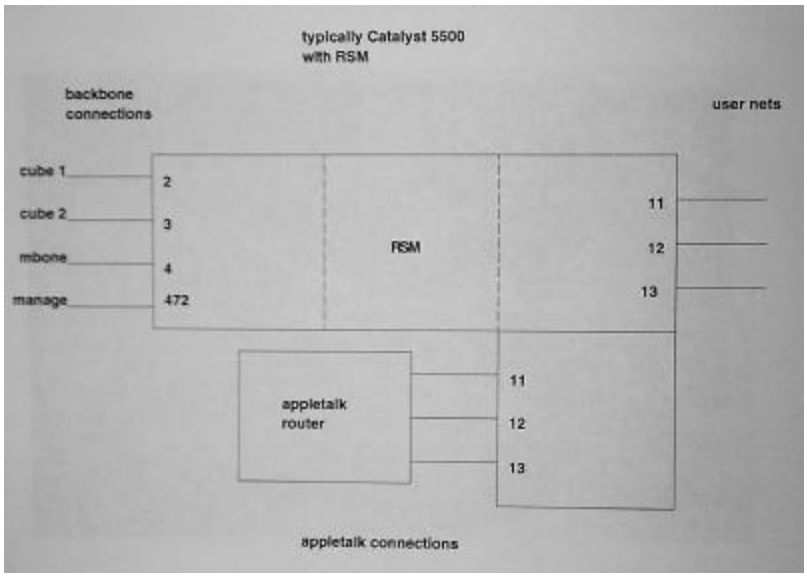
- Planned to serve multiple networks per router interface using VLANs and Catalysts
- Instead using new router module (RSM) in Catalyst 5500, feeding Catalyst 5000s
- All networks moved off of AGS+ routers
- Limited use of Catalyst 3200
  - Replaces AGS “departmental routers”

Catalyst 5500 w/RSM

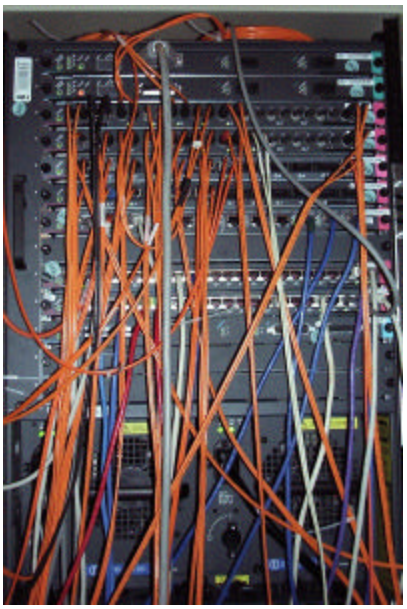
- Dual power supplies and switch controllers
- 9 slots for 12x100Fx <or> 12x10FL <or> 24x10/100Tx
- New subnets easy using VLANs
- VLAN trunks used to remote switches
- Multiple RSMs can be installed
  - For reliability or load sharing reasons



User Router—Cube Corner Connections



VLAN Usage in RSM Routers



5500 w/RSM
as a User Router

Dual User Routers

- Some organizations want more reliability
  - Graduate School of Business (GSB)
  - University Libraries (SUL)
- Dual Catalyst/RSM configurations
  - Using HSRP for every network
  - Each has one corner connect plus interconnect
- They bought their own hardware
  - To be installed and configured by Networking

Routing Configuration

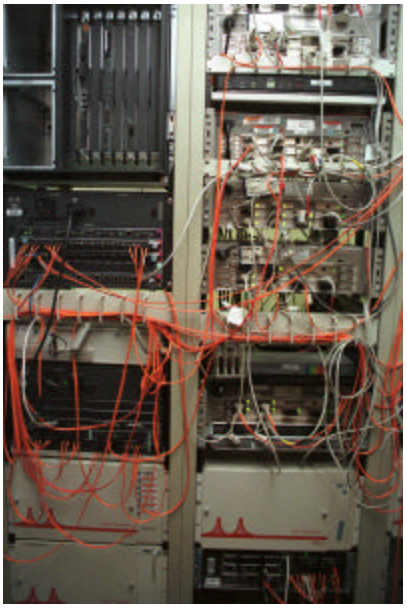
- Using OSPF
  - Most routers in area 0, limited other areas
    - Only a handful of off-campus connections
      - Still a small number of 25xx routers in use
    - 171.64.0.0/21 reserved for backbone use
- BGP with CalREN-2 and GTEI
  - Soon SLAC and PAIX as well
- OSPF-OSPF redistribute in Med-gw
  - Between 171.65 and 171.64

OSPF on User Routers

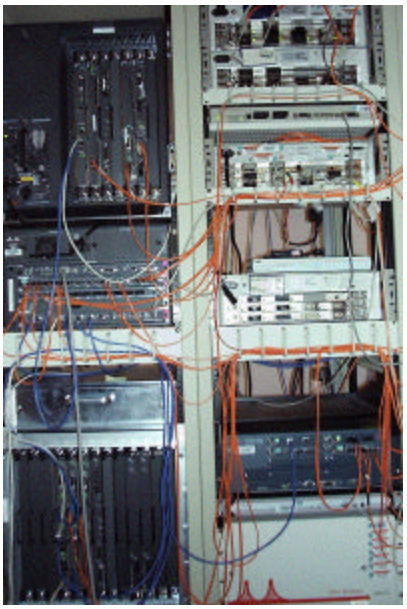
- Set high OSPF cost on corner connections
  - To keep core from routing across
- Currently doing “redistribute connected”
  - For convenience
    - Plus no other random peers
  - May want to transition to “authentication”
    - No address summarization on type External routes
    - Addresses had been too random until renumbering

OSPF-OSPF Boundary

- **Actually doing OSPF-OSPF redistribution**
  - **171.65 addresses in Med Center**
    - **But still some legacy 36/8 addresses in use**
  - **In Med-gateway, on border with Med Center**
  - **Hospital networking is separate group**
  - **But Medical School is supported by SUNet**
    - **Former MedNet group—now Med School only**
    - **Share configuration of Med School equipment**



A Variety of Legacy Technologies in Pine Hall



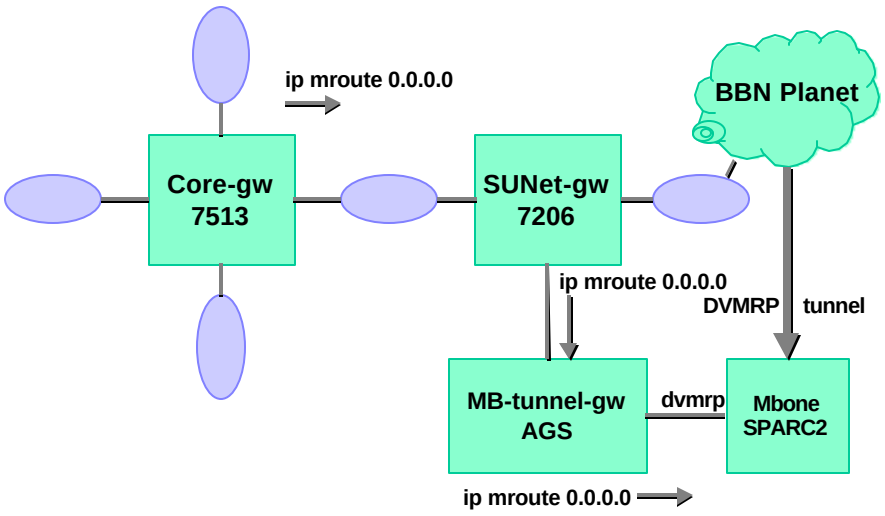
Early Multicast

- Computer Science
  - Championed by grad student, Steve Deering
  - Grad students set up MBone DVMRP routers
- Later one added one in his dorm room
- Decided we should have one SUNet tunnel
  - Requested BARRNet to install in their core
  - NS ran SUNet end of tunnel
    - mrouted/DVMRP tunnels for MBONE on campus

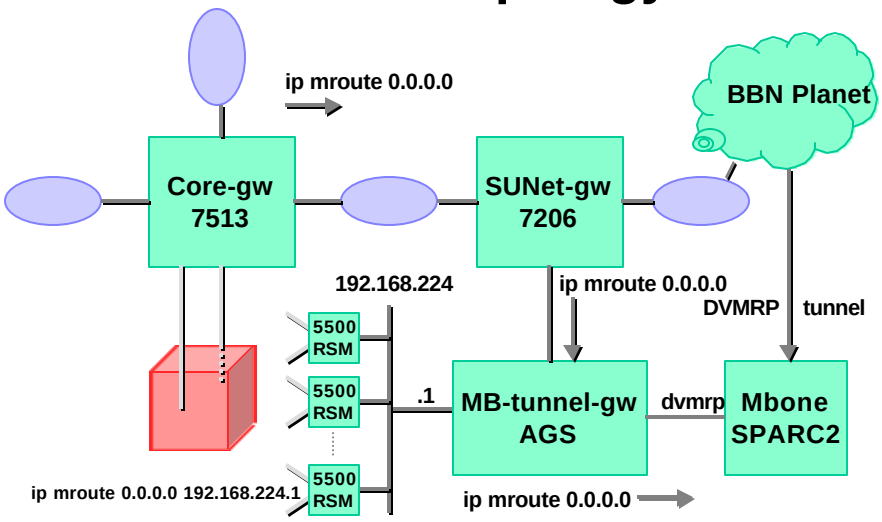
Multicast

- Originally activated used PIM dense mode
  - Worked in 11.0 AGS routers (if no loops)
  - Not configured in the cube backbone; Instead deployed a dedicated multicast network connected to each user router (5500 w/RSM)
- Expected multicast users
  - SCPD (SITN) Engineering class broadcast
  - Library—supplement class videotape checkout
  - Conferencing with remote sites

Multicast Topology ’97



Multicast Topology ’98



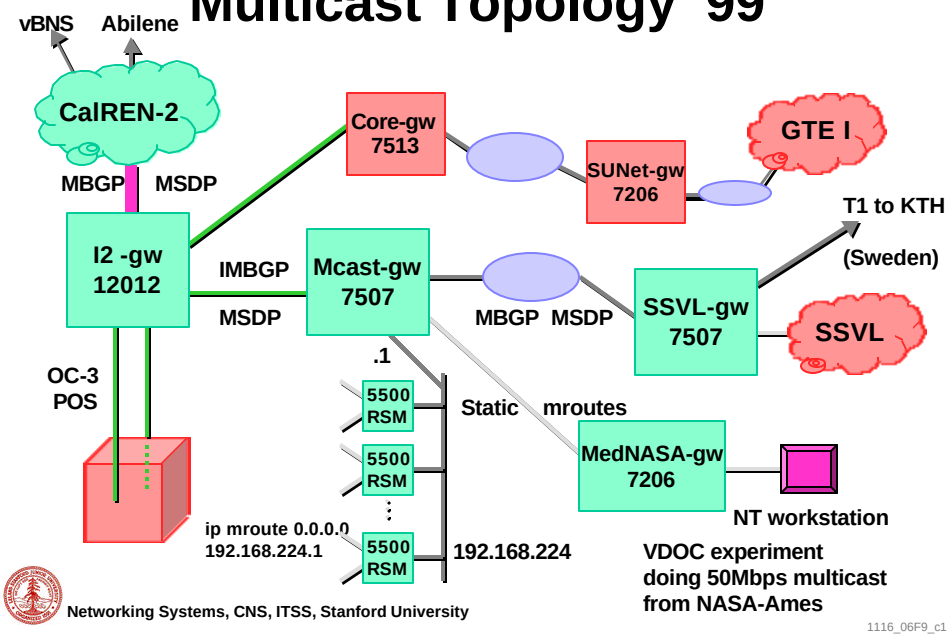
Multicast ’99

- Transitioned to multicast from CalREN-2
- 7507 as dedicated multicast router
  - Via OC-3 POS from border GSR (I2-gw)
- Still separate 10-Mb multicast backbone
  - Still using static Mroutes, PIM Sparse-Dense
- Border GSR doing MBGP/MSDP (12.0S)
- Traffic from both vBNS and Abilene
  - vBNS is main feed to AS10888 (old Mbone)

Transition to I2

- Added Mcast-gw (7507) direct POS to GSR
- Moved 192.168.224.1 address
  - From mb-tunnel-gw (AGS) to mcast-gw (7507)
  - Still a 10-Mb Ethernet
- Turned off mrouted tunnel to GTEI

Multicast Topology '99



Network Monitoring

- Stanford Networking's kiss of death:
  - DEC MSU/Remedy Health Profiler
  - Kaspia/IBM Netview 6000??
- NAT/Tech Elite: Ethermeters/RMON
- HOWIS, a Perl script, is still main tool
- And What's Up
- Recent Distributed Systems Monitoring project now provides a WWW view

Traffic/Usage Data

- Have been trying Cisco's Netflow Export
  - Netflow Collector looks very promising
  - Immense amounts of data!!
  - Installing collector server(s) per site
    - Approx. 2 days of uncompressed 5-min. data files
  - Archive data w/tar and gzip to central system
- Other products being examined
  - Other vendors' Netflow analysis software
  - Other data gathering hardware

The Perils of SNMP

- Traffic analysis showed >30% was SNMP
- Have installed a separate management net
  - Using RFC1918 (“private”) addressing
- Trying some alternate tools
  - Expect script for ARP collection
- Would like “SNMP proxy” system
  - Collect once
  - Other tools then pointed at proxy

AppleTalk Migration

- Shiva no longer supporting Fastpath
  - Hardware is aging as well
- Recycled AGS+ routers as AppleTalk net connected via additional 5500 10Mb ports
- Now two 75xx routers w/100Fx ISL trunks
- Not planning to migrate easy NetDB support for new zones—hand configuration
- Policy is “no new zones”

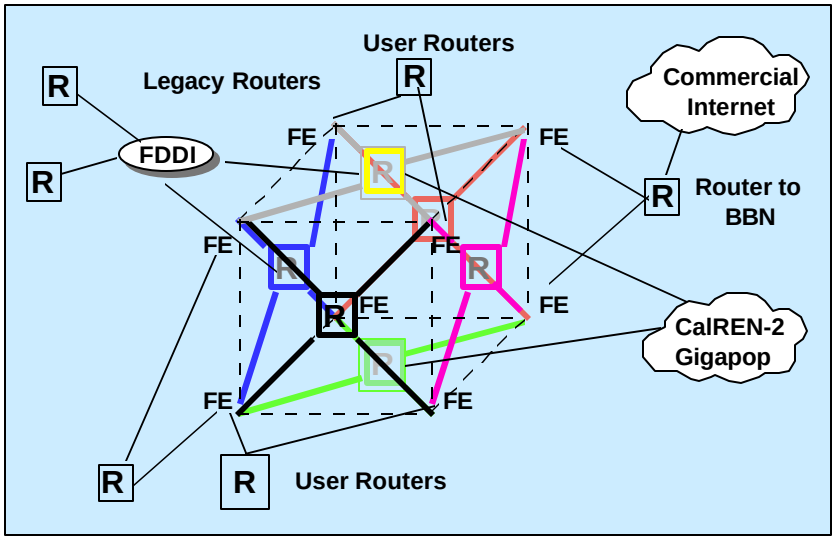
So, What's Next?

- Gigabit Ethernet was due in '98, now '99
  - Expected to be mainly a backbone technology
  - Still, some users will demand it…
- Projects to enhance building wiring ongoing
- Fast Ethernet to the desktop
  - Now very cost effective—departments buying
  - Switches now the cost of repeaters just 18 months ago

More Next...

- More bandwidth to off-campus sites
  - Hopkins Marine Station in Pacific Grove
  - Overseas Studies centers
    - Europe, Japan, Central America
    - Connect through local Internet-2 affiliates
- QoS testbed with CalREN-2, Internet2
- More diverse routing to commodity internet
  - Connection at PAIX makes peering possible

SUNet Cube Backbone



Organization

- ITSS (Information Technology Systems and Services)
  - CNS—Communication and Network Services
    - Communication Services—fee for service
      - University telephone and cable TV systems
    - Networking Systems—centrally funded
      - SUNet Operations
      - SUNet Systems
      - Monitoring and Management
      - Network Consultants
      - LAN Operations

Operations' Responsibilities

- Backbone routers
- End-to-end TCP/IP protocol
- Network to the building entrance
- Support network consultants/LANs

Support Profile

- Onsite staff—7am-11pm, Monday-Friday
- Weekends—8 hours/day
- On-call support 7 by 24
- Off-hour trouble calls taken by Computer Operations staff who page on-call person

Staffing Level

- Operations has 5 Network Specialists
- Systems has 3 programmers, 2 SysAdmins
- One network specialist/one sysadmin work the late shift (3pm-11pm)