

Data Communications

Al Thomas

October 16, 2002

Outline of Presentation

- Mission
- Organization & Scope of Assignment
- **Major plans**
- Video Conferencing
- Websites: supporting docs & background

DCD Mission

“The mission of the Data Communications Department is to provide for the planning, implementation and management of Fermilab’s data communications infrastructure and services – the networks.”

- The primary focus is on the laboratory's research program. However, the responsibilities extend to include the entire laboratory.
- Department activities include: internal & external networking, campus-wide cable & fiber infrastructure, video conference scheduling & consulting, and network monitoring.

Organization:

Data Communications Department

Al Thomas - Head

Keith Chadwick - Assistant Head

Data Communications & Networks

Phil DeMar - Leader planning & strategy

Donna Lamore - Assistant Leader operations

Andrey Bobyshev

Vladimir Bravov

Alden Clifford

Vyto Grigaliunas

Roger Kramme ←

Data entry, ProCard,
VC backup

John McIver

Andy Rader

David Tang

Darryl Wohlt

Data Communications Infrastructure

Ron Cudzewicz - Leader

Chuck Andrews

Orlando Colo'n

Steve Fry

Clif Horvath

Technology

Frank Nagy

PingER

Maxim Grigoriev

Video Conference Coordinator

Shiela Cisko

Providers of Lab-Wide Service

- A slightly worn phrase, perhaps, but true
- Time horizon from the immediate “need it today” to planning infrastructure years in advance
- Project scope from building and inter-building wiring to direct HEP experiment support
 - Very much the data equivalent of “bridge painting” to “regional transportation planning”
- Greater reliance on data networks everyday
 - ES&H card reader access, gates (?)
 - FESS building control & monitoring
 - VoIP is likely not far off
- We must keep the networks running!

... stated another way

- We provide a critical service for the Lab
 - All Divisions & Sections are dependent on our service to accomplish their work
 - Experiments can't store, process, or distribute their data if our service is not operating
 - Other Laboratory services now being based on our underlying service
- Our absolute top priority is to keep the network operating properly
 - Everything else takes lower priority...

Major Plans and Challenges

Run-II Network Challenge (I)

- Run-II network's are top priority...
- Bandwidth needs are increasing
 - Enstore model puts more data on networks:
 - 300 MB/S (CDF peak) and ~175 MB/S (D0 avg)
 - dCache makes data accessible at higher rates and also more distributed across the network
 - 550 MB/S (CDF peak) and 20MB/S (D0 “fuzzy” but looking at growth mostly to offsite)
 - Emerging “private” analysis clusters in portakamps may create additional high volume traffic

Run-II Network Challenge (II)

- Switch fabric & port density needs are increasing rapidly
 - Analysis systems evolving to large farms
 - CDF CAF Farm now 234x100Mb and 66xGE ports
 - Will grow to 648x100Mb 134xGE ports by 2004
 - Gigabit ethernet connections for computer room systems is becoming standard
 - Gigabit to desktop also beginning to appear
 - CDF/PPD (ATOM) cluster includes 24 GE desktops
 - People want support for Cu GE

Run-II Network Challenge (III)

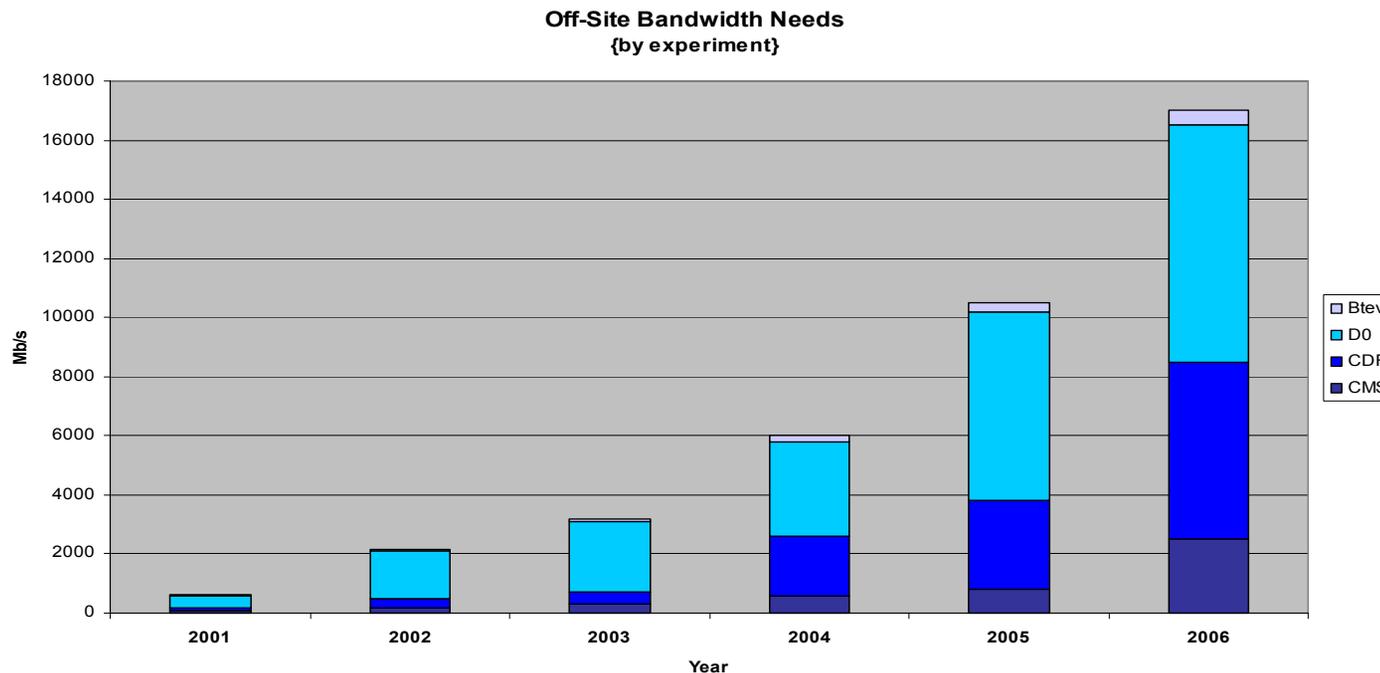
- Advancement in network technology slowing
 - 10 Gigabit Ethernet still costly
 - Link aggregation { $n \times 1000\text{Mb/s}$ } provides a stopgap, but not a long term solution
 - Switch fabric capacity & line card densities not growing as rapidly either:
 - Current generation of switches can't sustain 10 GE
 - Supporting large number of gigabit hosts is difficult/costly
 - Phase 2 of the CAF farm has completely filled a Cisco 6513, meaning additional switches needed for phases 3 & 4

Run-II Network Challenge (IV)

- Meet & stay ahead of CDF & D0's needs by
 - Add new switches to meet new requirements
 - ie., CDF CAB farm; D0 CAB farm
 - Upgrade existing switches with
 - Higher capacity switch fabric (now at 640 Gb/S)
 - Adding line cards, mostly GigE, w/ higher port density
 - Looking at alternative vendors w/higher densities
 - Upgrade trunk links to 10GigE as \$'s allow
 - Cisco interfaces can range as high as 30K\$ each
 - Aggregate GigE links (n x 1000Mb/s) until then
 - Create separate LAN for Run-II off-site bulk data movement (under study & consideration...)

Off-Site Network Challenge (I)

- Projected increase in off-site traffic
 - Collaborations looking for distributed analysis
 - Tier model (D0, CMS) puts emphasis on bulk file transfer capability to select sites



**Data from
TAN
report,
dates are
wrong but
trend
assumed
valid**

Off-Site Network Challenge (II)

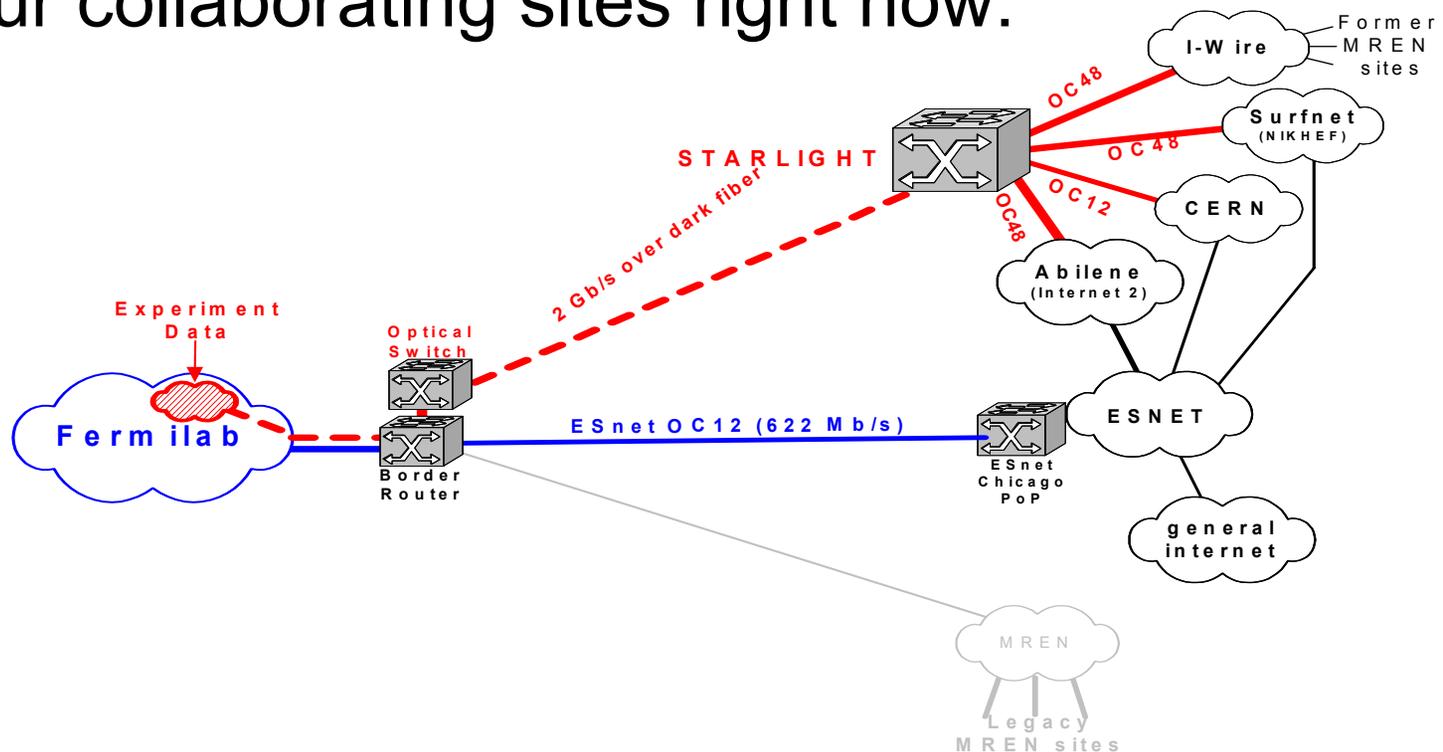
- Some sites have upgraded capacity to FNAL
 - CERN: OC12 (prod'n) & OC48 (DataTag)
 - NIKHEF: OC48 (prod'n) & OC192 (research)
 - Abilene/I2: OC192 backbone; planning OC768
 - I2 sites (universities) starting to upgrade to OC48
 - CA*net: OC192 backbone w/ OC48 to US
 - Planning on optical network switching by 2004
 - I-Wire: Dark fiber, OC48 now, w/ OC192 looming
- All these networks appear at StarLight
 - <http://www.startap.net/starlight/>

Off-Site Network Challenge (III)

- ESnet not keeping up in terms of bandwidth
 - Backbone is just now upgrading to OC48 though there are plans for OC192
 - The site tail circuit situation is of great concern
 - FNAL's OC12 upgrade has taken ~18 months so far
 - No funding requests for upgrades thru FY04 (FY05?) for site tail circuit upgrades
- ESnet not aggressively pursuing optical network technologies
 - Seem to have an “IP service provider mentality”

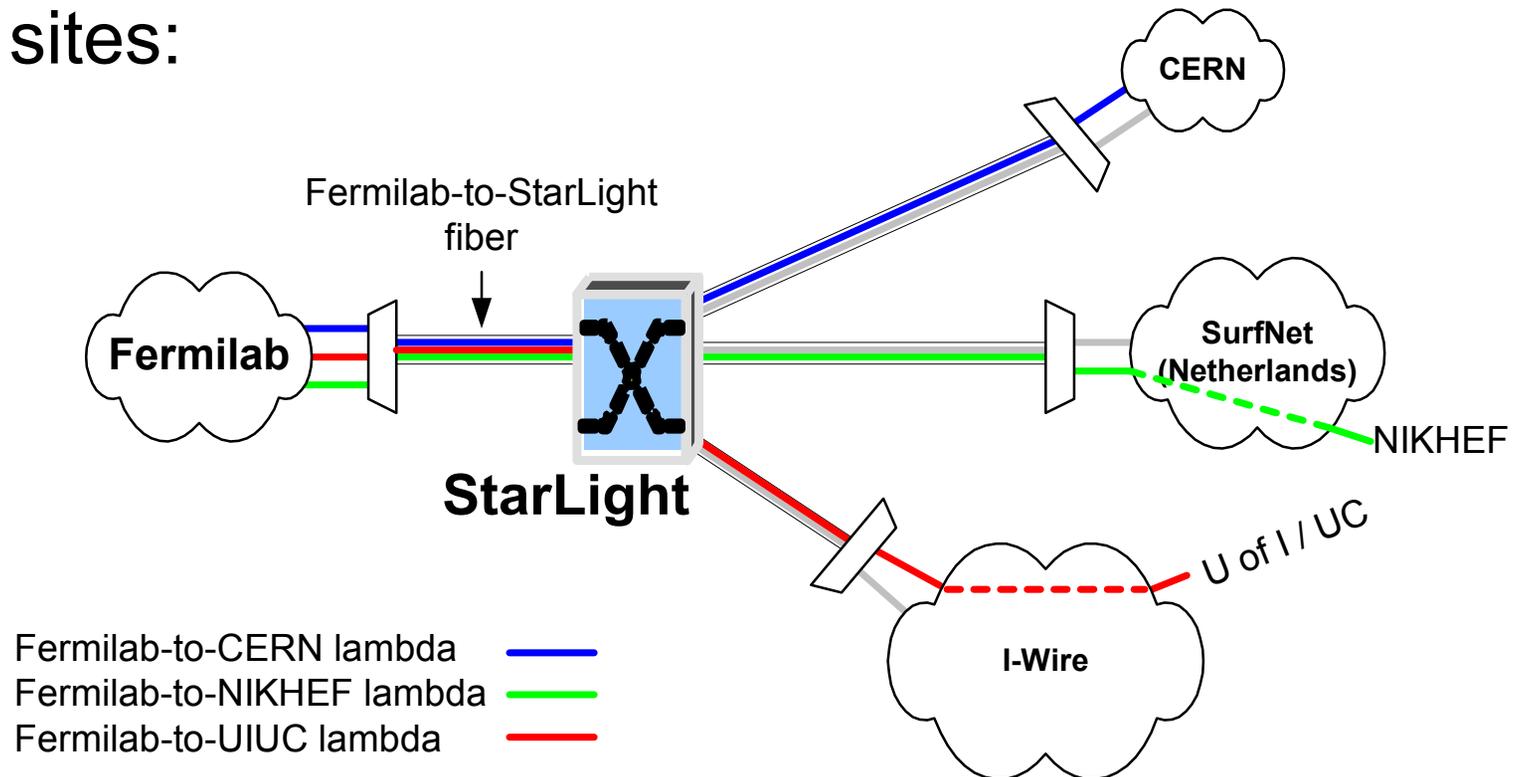
Off-Site Network Challenge (IV)

- StarLight (710 N. Lake Shore Dr., Chicago) beckons
 - The optical network exchange point in the U.S.
 - Could support a multi-gigabit IP path to many of our collaborating sites right now:



Off-Site Network Challenge (V)

- StarLight (cont):
 - Longer term potential for testing optical network technologies with select collaborating sites:



Off-Site Network Challenge (VI)

- actively pursuing 1 dark fiber pair to StarLight
 - Costs are the major issue
 - ComEd fiber costs 380K\$ upfront + ~100K\$/year
 - Optical electronics also not cheap
 - Trying to capitalize on the I-Wire project efforts:
 - not part of the initial state-funded I-Wire project
 - I-Wire will let us use their optical ports & 6509 ports
 - FNAL might offer λ 's to nearby interested parties
 - IMBCA, TRECC, NIU via I-Wire,
 - Potential: good will, cost sharing, foster connectivity,
 - Trade dedicated λ 's through I-Wire to our sites

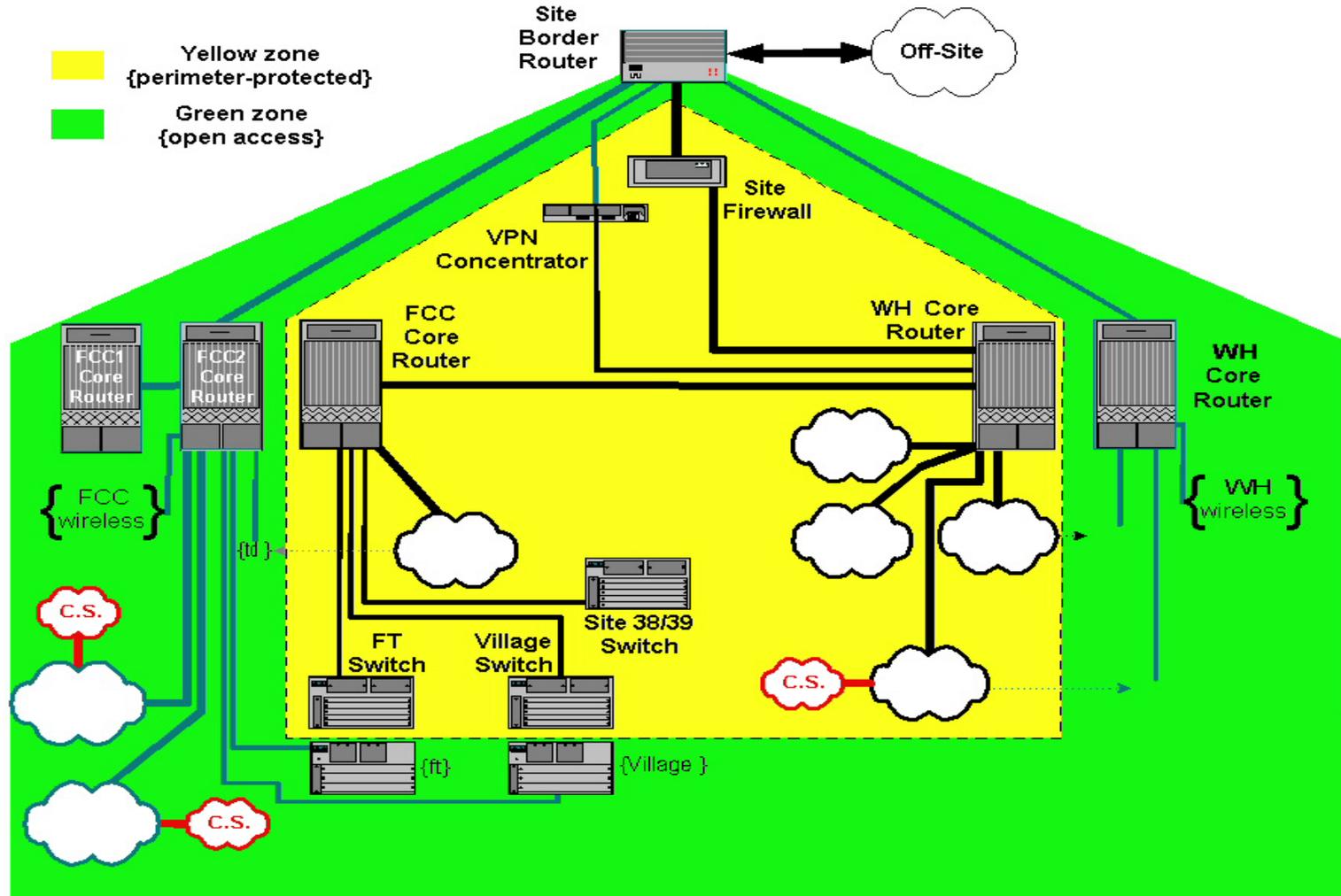
Perimeter Protection Challenge (I)

- The Lab supports open Internet access, but
 - The Internet is an increasingly hostile place
 - Significant effort spent on incident investigation
 - Much of the site network doesn't need open access
- A project to draft a multi-level security zone (green, yellow) implementation plan is under way
 - Now testing firewall protection in the CD LAN
 - VPN support for “secure” off-site access to the protected zone is part of that pilot

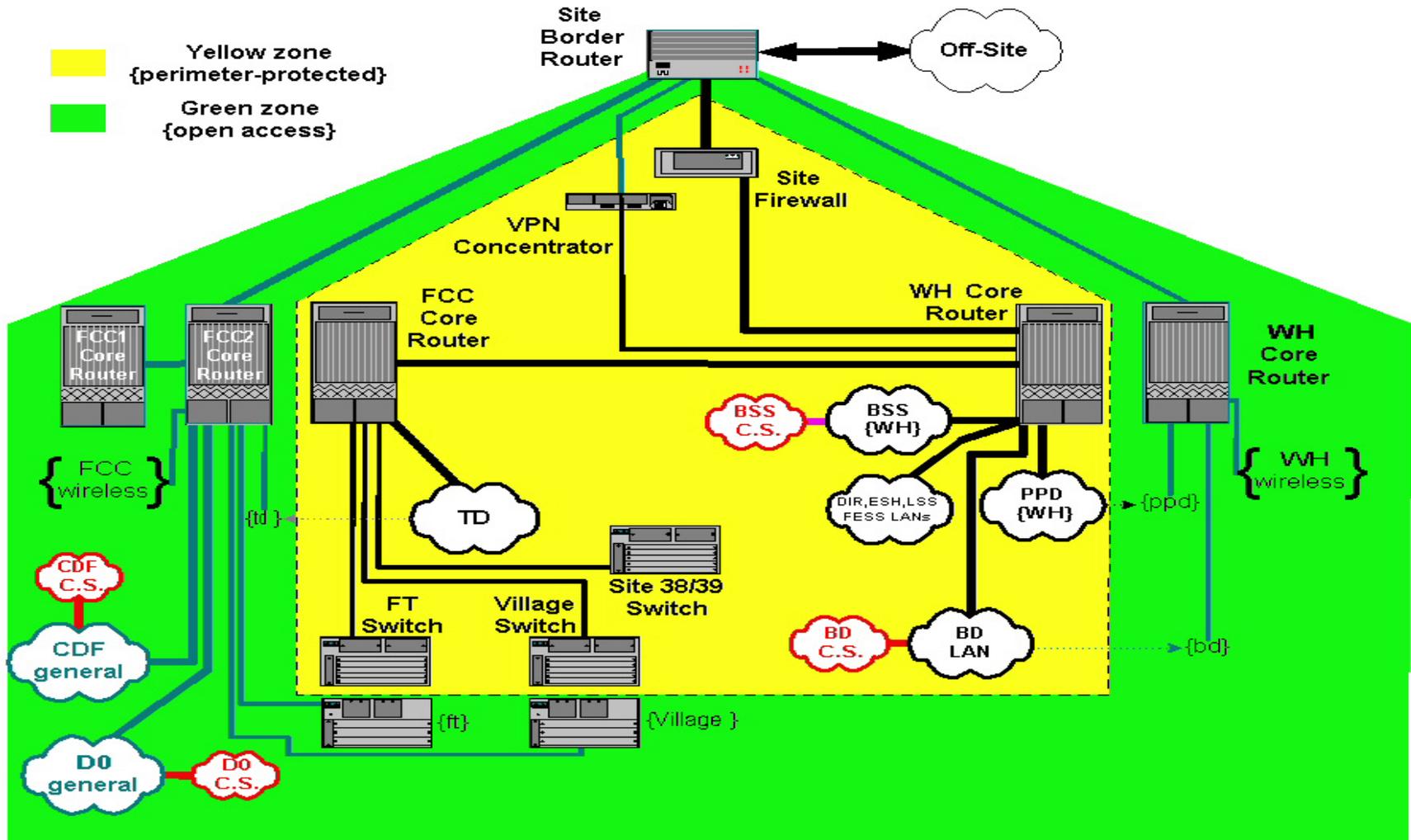
Perimeter Protection Challenge (II)

- Cornerstones of the plan
 - Continue to support open access for collaboration & experiment resources that need it
 - Provide protection for parts of the network that don't require completely open access
 - Not hinder very high performance networking
- Significant architectural change to the facility network would be necessary:
 - Apparently no \$\$ from DOE, funding's an issue
 - Major effort; competing with other priorities

A Generic Protection Scenario (III)



A Possible Protection Scenario (IV)



Core Network Redundancy (I)

- Little redundancy in the core network today
 - Redundancy adds significant expense & complexity
 - Core network devices have been highly reliable
 - Low levels of down time have been acceptable
- Increasing reliance on nets makes impact of downtime increasingly painful
 - distributed computing and web-based monitoring
 - 7x24 services starting to use the network
 - Metasys for bldg controls; PegaSys for key card system
 - Can VOIP be far off?

Core Network Redundancy (II)

- Core network redundancy will require:
 - Architectural design changes in the core network
 - \$\$\$
 - Effort (manpower)
- We're planning on building in redundancy as part of core network upgrades
 - It will occur gradually, probably starting in FY04
 - How to make it compatible with perimeter protection and bulk offsite data transfers ?

Infrastructure Development (I)

- Bring a clarification of rolls to the internal operation of network infrastructure
- Nagy & Gregoriev serving as focus for OS and applications
- Trying to regularize systems and allow network experts to concentrate on the networks rather than OS, scripting, backups, etc.
- Put procedures in place appropriate to a small software development or “customizing” operation

Infrastructure Development (II)

- Common admin for PingER, IEPM-BW, DNS, MRTG, DHCP, NetFlow archive, DCG0
- Starting a framework for resource monitoring and control tools (AutoBlocker)
- Want to build that framework on a well defined base of support platforms with attention to proper admin tools & techniques
- Support installation & operation of a DHCP registry system in conjunction with CST

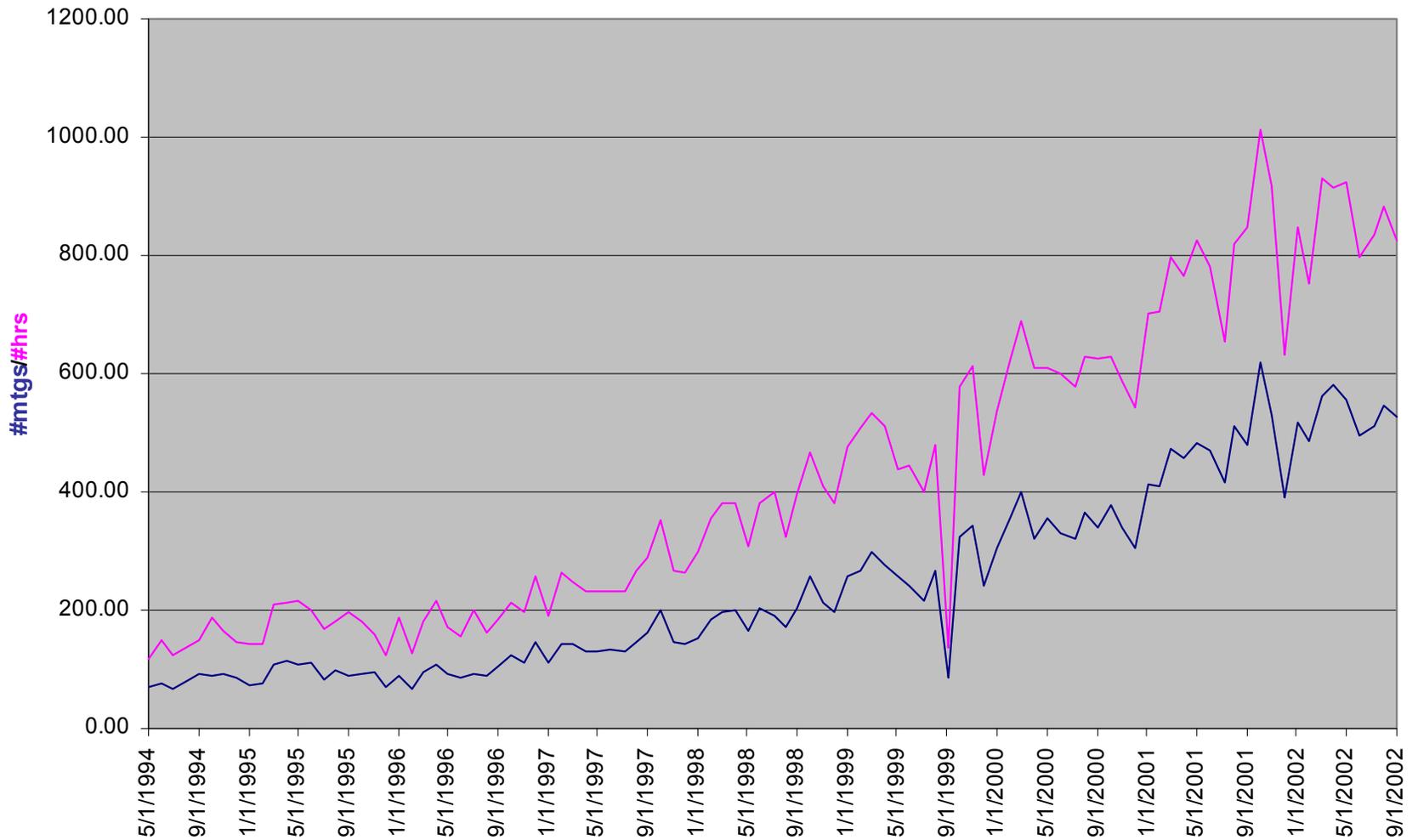
Plant Upgrades

- Fiber route to Kautz Road Substation
- RF coverage fix-ups in WH, DAB, CDF
- Aerial fiber runs in the village
- Migrate most on-site DSL lines to LRE
- Complete FCC mezzanine connectivity
- Complete connection to IMBCA system
- FCC1 & FCC3 cabling upgrades
- Gigabit support for FT & village areas

Video Conferencing Status

- One FTE for scheduling, minimal backup
- Coordinator schedules rooms and resources for ~190 standing video conferences
 - 72 are Pt2pt (~15 are IP-based)
 - 115 are multipoint (~12 could move to IP).
 - ~18 standing audio conferences
- Involved with ~16 rooms
 - Advise and assist in equipment upgrades, room improvements, troubleshooting and consulting
- New conference room involvement
 - Two new rooms created each year
 - ~6 months involvement each
 - Coordination between DCI, FESS, requestors and vendors
 - Customers: BTeV, CD, CDF, CKM, CMS, D0, DoE FAO, KTeV, LHC, MICE, Miniboone, NLC, NuMI, NuTeV, PPD, Neutrino Factory, and MuCollider
- Need help for backup and load sharing with Sheila

FNAL ISDN Video Meetings



Video Conf Room: WH10NW



Video Conf Room: SiDet



Video Conferencing Futures

- Migration of ISDN video conferences to IP
 - Involves changes to conference room procedures and providing information to users
 - Desktop H.323 hardware clients can now be included in IP meetings
 - Per ESnet as IP user base expands FNAL (plus other labs & universities) will need to manage an IP gatekeeper. ESnet warns that this is the most time consuming aspect of H.323 video conferencing.
 - New technologies include data collaboration with teleconferences, both are services provided by ESnet.

Video Conference Room Plans

New installations (inc room prep)

DAB, off high bay area, by PPD request

WH8X, by PPD request

Upgrades planned

CDF Theater – new sound and video gear

FCC1 – new projector, preview audio leveller(s)

WH7X – new projector & some cleanup

Major Consulting effort

By request of D0 Video Conference Task Force to assist regarding improvements to 3 portakamp video rooms.

Interesting Items & Further Reading

- Network Weather Map <http://www-dcg.fnal.gov/~netadmin/nwm/cgi-bin/temp/core.html>
- Top 20 Hourly and Daily Statistics <http://www-dcg.fnal.gov/~netadmin/topn.cgi>
- Access to more FNAL network statistics <http://www-dcg.fnal.gov/stats.html>
- H.323, T.120 and audio services <http://www.ecs.es.net/>
- Ad-Hoc bridge and H.323 <http://www-staff.es.net/%7Emikep/adhoc/index.html>
- New PingER website (Apache, MYSQL, Perl) <http://pinger.fnal.gov/>
- PingER participants list <http://pinger.fnal.gov/participants.html>
- IEPM-BW homepage <http://www-iepm.slac.stanford.edu/bw/>
- IEPM-BW SLAC-FNAL http://dmzmon0.deemz.net/~cottrell/html/slac_wan_bw_tests.html
- Starlight <http://www.startap.net/starlight/>