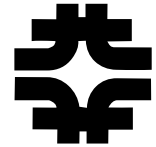


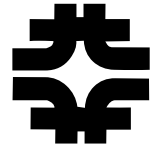
Facility & Network Infrastructure Challenges

R. Tschirhart

Fermilab

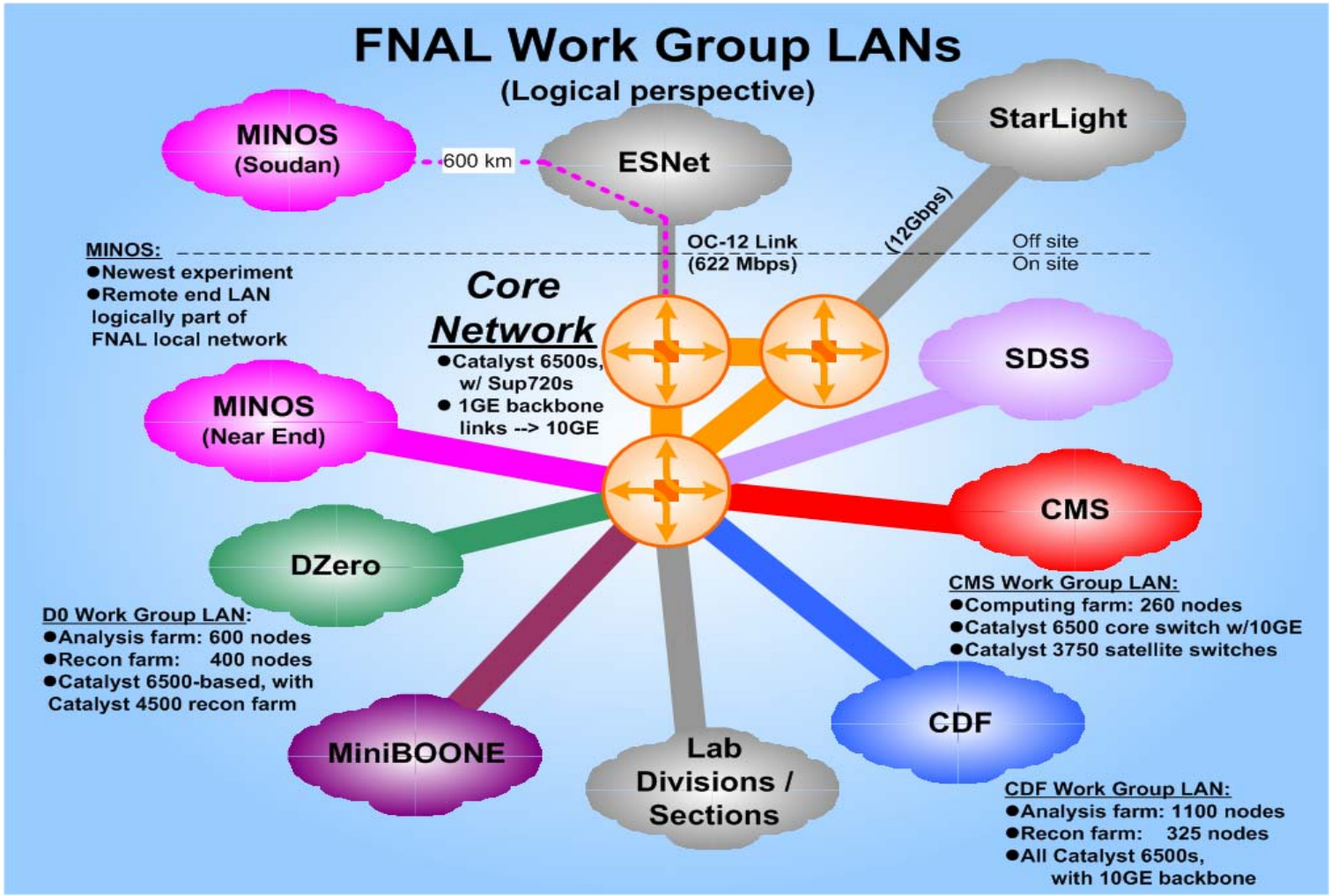
March 30, 2005

Network & Facilities



- Critical to the success of both experimental programs and common services.
- Not a small tax. This has become particularly clear with facility infrastructure in recent years.
- Market and technology trends strongly influence our fate.

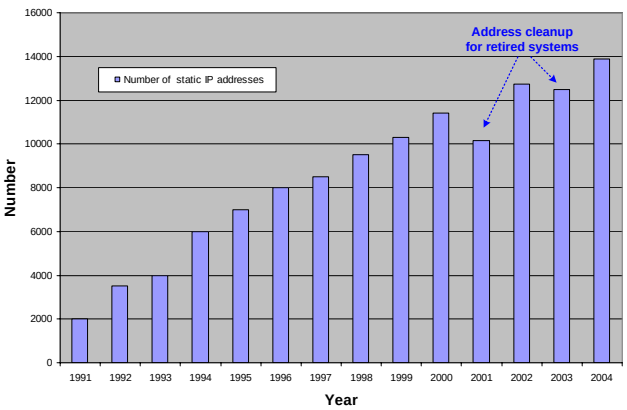
Architecture of FNAL LANs.



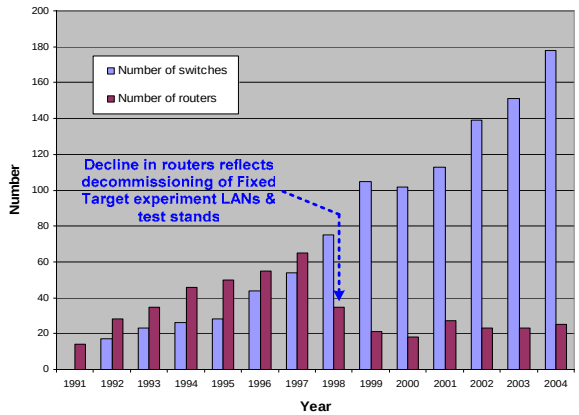
LAN Growth Trends

- Growth in systems continues at ~1000/year (below left)
  - Necessitates corresponding growth in # of switches (below center)
  - System growth rate likely to increase with CMS gearing up
- Upgrades in LAN technologies parallels system growth:
  - Systems now connected at 1000B-T by default
  - New switch uplinks correspondingly deployed at 10 GE

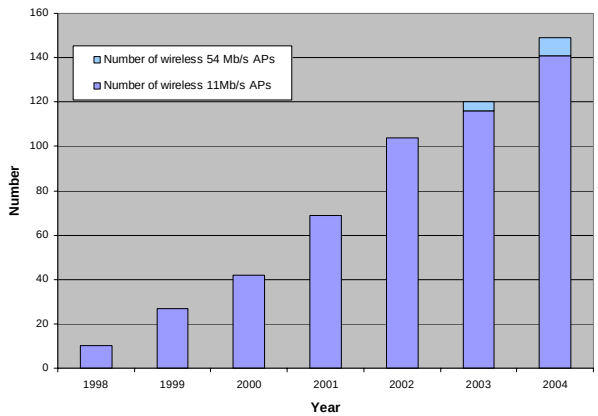
Number of allocated static IP addresses



Campus Network Switches & Routers

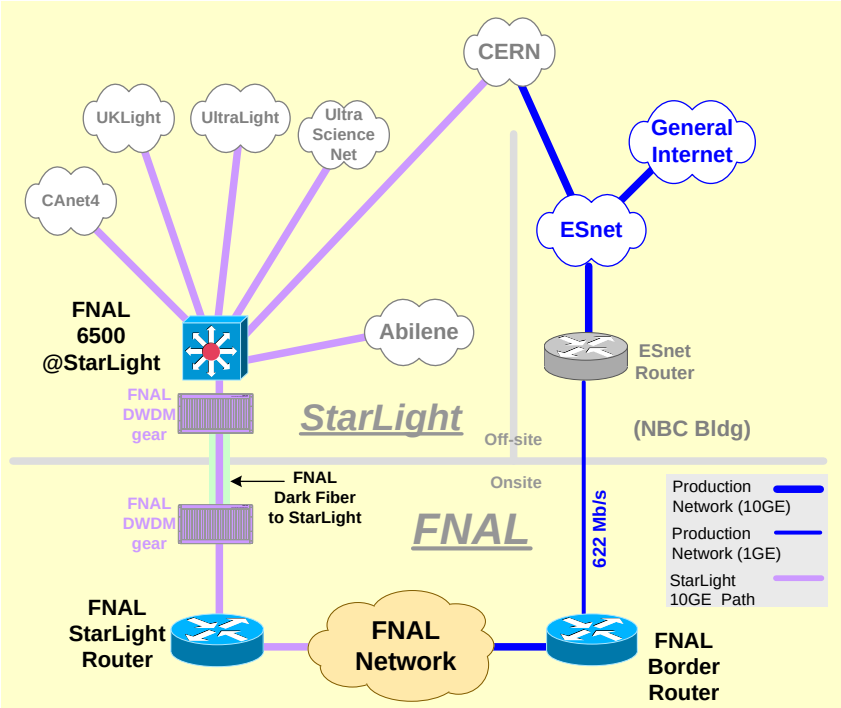


Wireless LAN Access Pts



Wide Area Network Overview

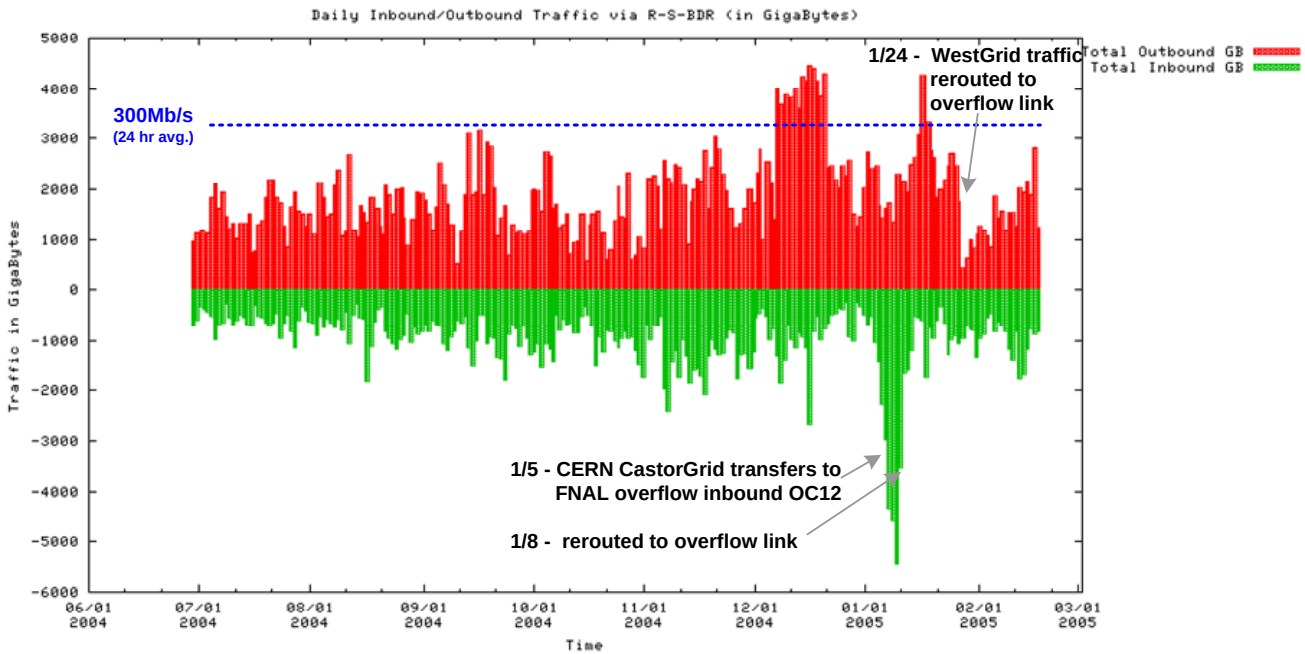
- Production WAN link funded managed by DOE (ESnet)
  - 622 Mb/s
  - Upgrade timing & path unclear
  - CMS challenge: 10,000 Mb/s
- FNAL-funded StarLight fiber
  - Intended for R&D, redundancy, and production overflow traffic
  - Initial configuration 12,000 Mb/s
    - Theoretical capacity = 330 Gb/s



- Soon, FNAL production network rates of 10 Gb/s & higher
  - Good practice = backup link of similar capacity

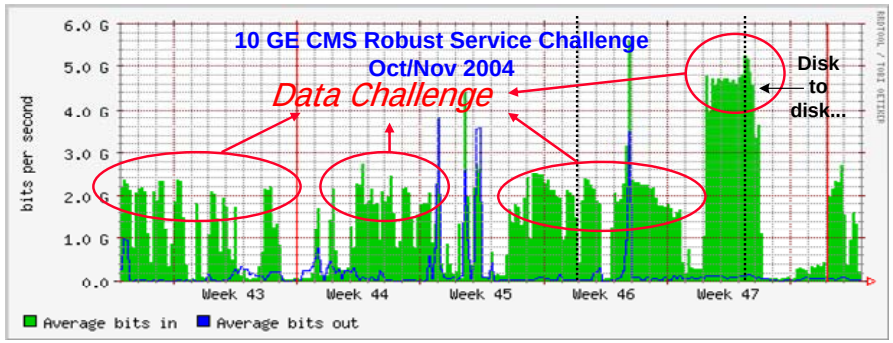
Status of Esnet link.

- The 622Mb/s link saturated:
  - Outbound averaged over 300Mb/s in Dec (24x7 basis)
    - Inbound link saturated in January
  - Migrating very large flows to StarLight overflow link

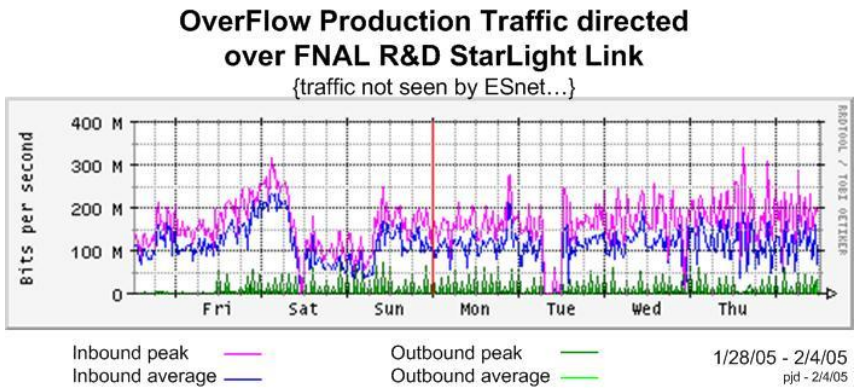


StarLight Link Usage

- R & D projects:
  - CMS robust service challenge sustained 2.5 Gb/s for weeks
  - SC2004 = 7.5 Gb/s sustained

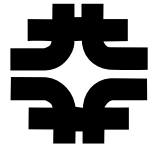


- Overflow production traffic:
  - CERN CastorGrid traffic
  - Westgrid traffic
  - Working on McGill, UCL



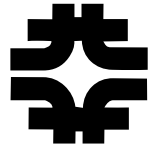
- Redundant off-site link:
  - Automated failover for ESnet link utilized 2-3 times already
  - Reliability still a concern; two extended outages last year

LAN Technology Risks



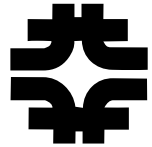
- LAN bandwidth capacity becoming insufficient for high performance computing farms
  - Exacerbated by growing size & geographic distribution of farms
  - Mitigation: deploy switches having capacity to aggregate 10GEs
- Capability to selectively route specific high volume data traffic to available high bandwidth WAN paths
  - Mitigation: LambdaStation research project to facilitate per-flow forwarding capability
- LAN technology beyond 10GE is unclear:
  - Mitigation: Track technology directions; deploy sufficient fiber to aggregate 10 GEs

LAN Budgetary Risks



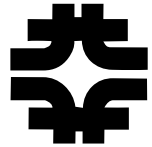
- Networking cost for farms has historically been factored in to total cost (15%...) of system
  - Moore's Law price/performance curve continues to hold for network switch infrastructure at the 1GE-level…
- Costs for 10GE capacity & opto-electronics remains high

WAN Technology Risks



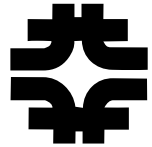
- Insufficient bandwidth for our physics program:
  - Mitigation: Cooperative effort with ESnet to work toward sufficient bandwidth & adequate connectivity to remote sites of interest, leveraging Starlight link
  - Mitigation: Ensure HEP funded transatlantic link is adequately funded and useful to Run II experiments as well as LHC expts
- Developing the capability to utilize high bandwidth WAN paths effectively:
  - Participating in advanced data movement demonstrations, including fast transport protocol implementations
- Development of WAN optical network light path technology unclear

WAN Budgetary Risk Issues



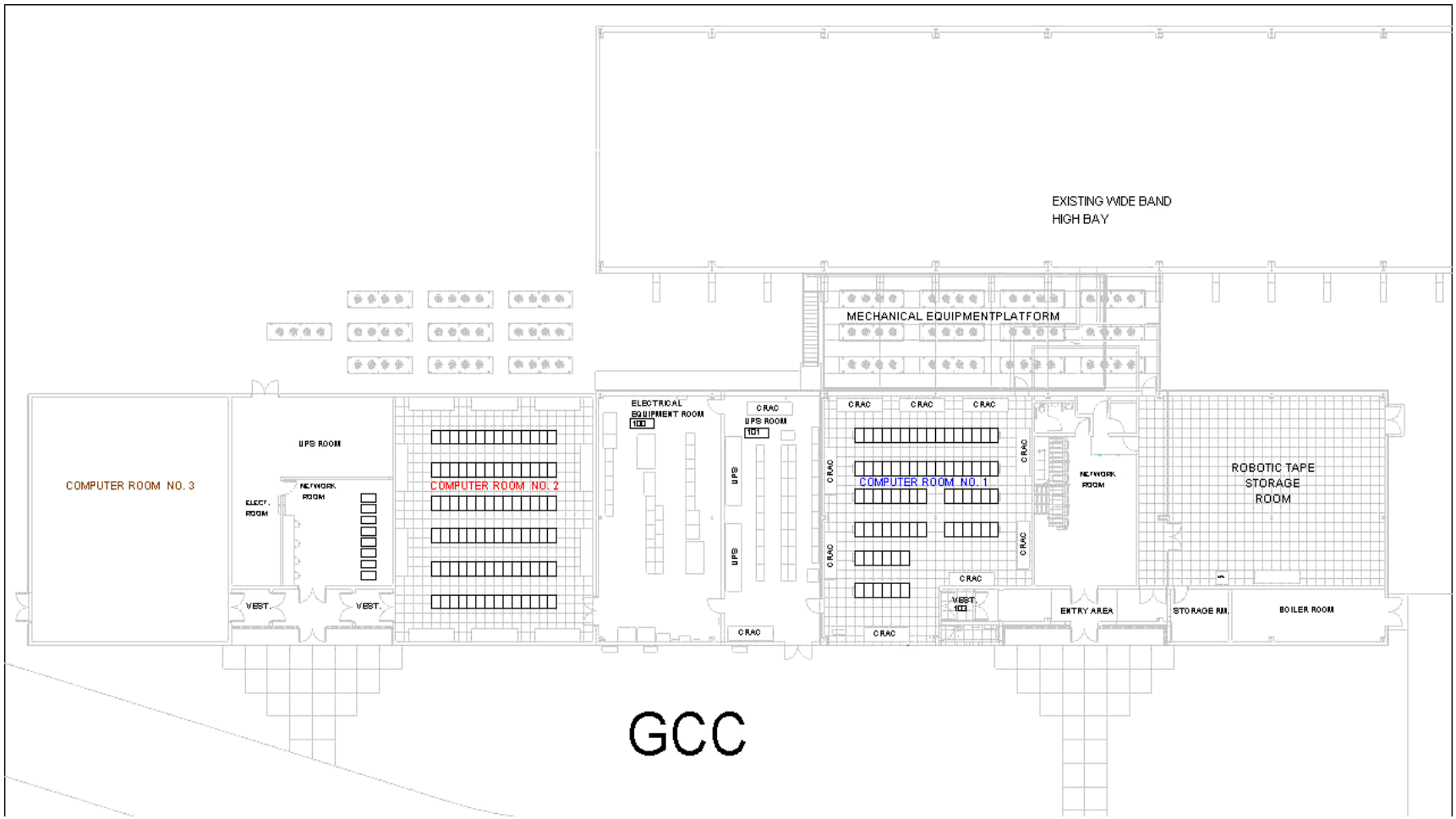
- DOE/ESnet funding for FNAL tail circuit upgrade not forthcoming
  - Mitigation: Could pursue metro area fiber initiatives with regional partners for alternate fiber path connectivity to StarLight, but not clear who would pay
- Cost of additional 10GE channels to StarLight fiber infrastructure is $80k each
  - Mitigation: Pursuing potential cost-sharing opportunities of our existing StarLight infrastructure with regional partners
  - Mitigation: Investigating lower cost per 10GE channel alternatives using different (CWDM) technology
- DOE/HEP funding for Transatlantic networking

Facility Challenges.

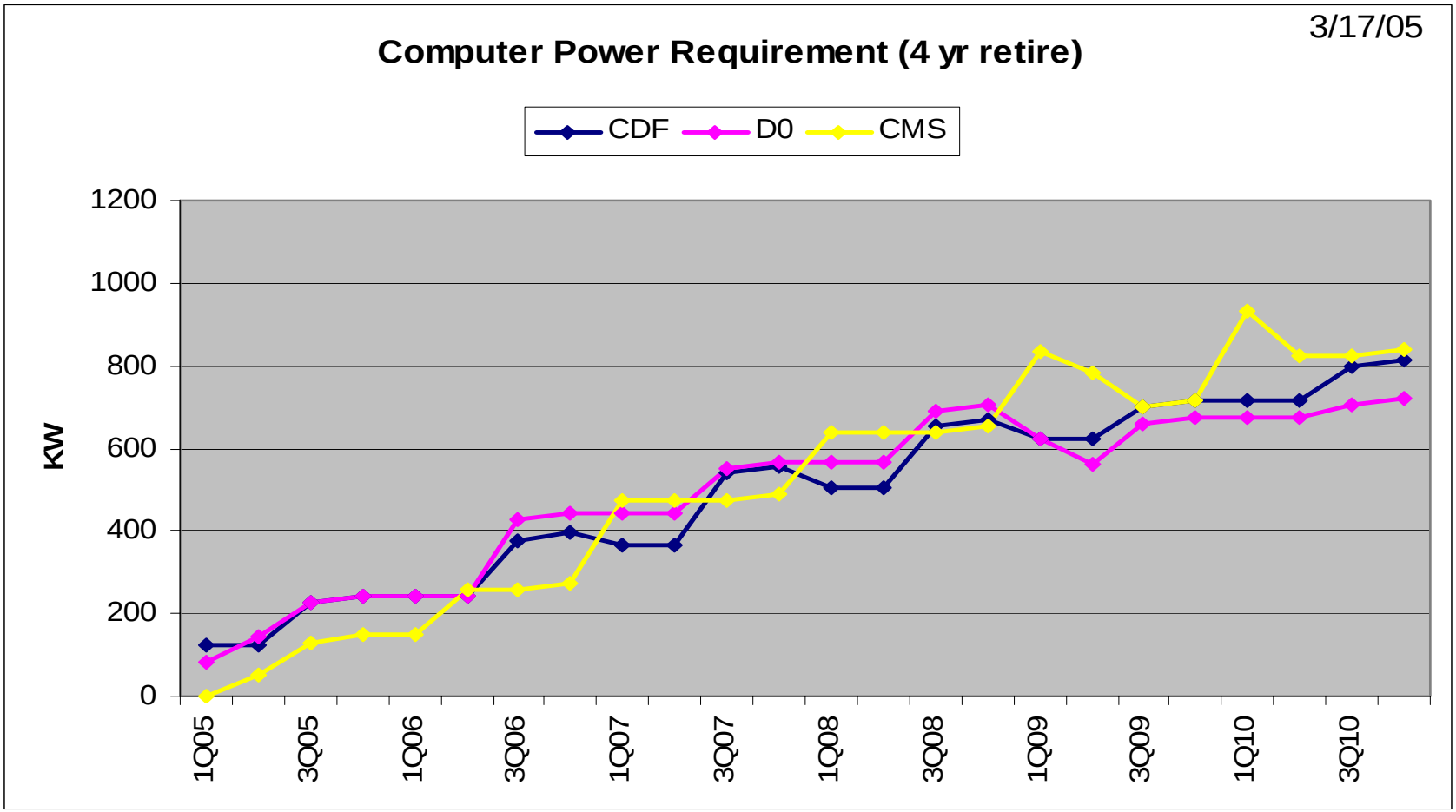


- Providing rack space, power & cooling.
  - Needs considerable investment of GPP funds
- Understand and address the risks associated with exclusively centralized data storage.
- Uncertainty in commodity computing trends, e.g. Blade computing, retirement cycles.
- Uncertainty in projections of computing need. Formal review processes are in place, but do not fully capture the developing story.

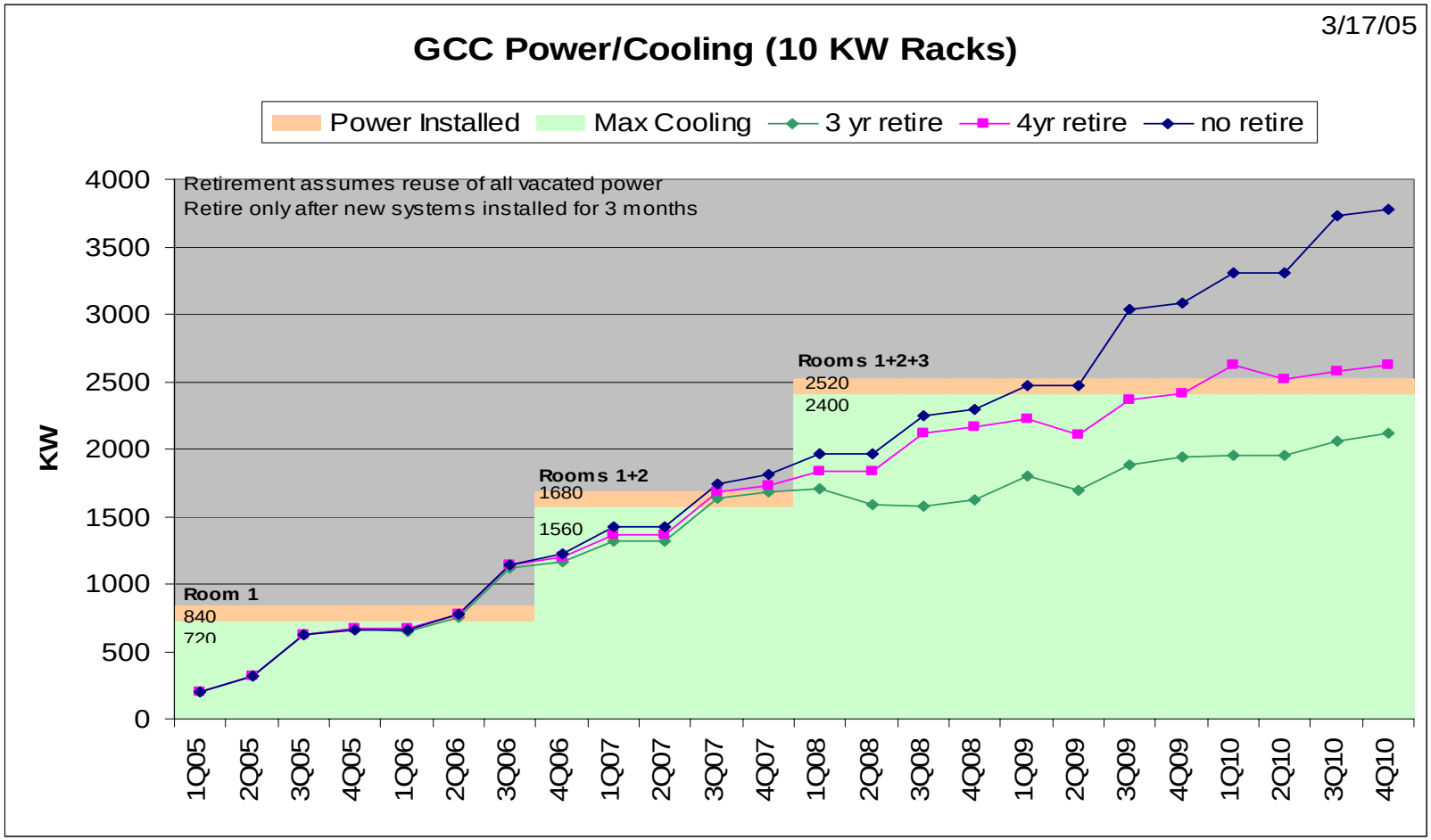
The Grid Computing Center: Reuse of Retired FT experiments



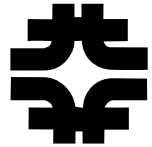
GCC Experiment Projections



GCC Power/Cooling



Facility Risks & Mitigation



- Excessively centralized data storage.

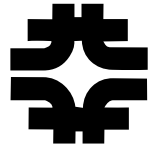
  Disperse Robots, investigate new technology.
- Rapidly evolving computing requirements.

  Greater reliance on Grid and off-site computing.
- Rapidly evolving commodity technology.

  Tracking computing *and* infrastructure trends critical.
- Out-year facility budgets.

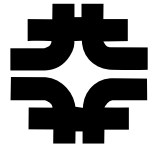
  Continue to communicate trends & requirements to Lab and community.

Cyber Security



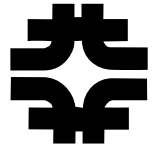
- We are re-writing our CSPP and overhauling our Computer Security Program to
  - Do a better job (actual and paperwork wise)
  - Add more formality to several processes
  - Go from one Enclave (whole campus) to two enclaves – General Computing + Open Science with different authentication and controls in place
  - Lot of work to do and increasingly vigilance in operations needed
- We are watching and waiting to see what PIV (Personal Identity Verification) actually will mean and who will pay
  - Responding to data calls
  - Working through SLCCC
  - Not sure what we can do to stop this train wreck
  - Note: We do have a Kerberos infrastructure in place with Cryptocard one-time-passwords as an option

DOE's Consolidated Networking?



- SLCCC has responded to the proposal by DOE CIO to lump all networking investments at all lab's into a single OMB-300 investment, presumably managed out of DOE HQ?
  - Strongly worded letter sent to DOE-CIO.
  - Word "embarassing" used
  - If such a thing were to actually happen we believe it could be crippling to our science and an unimaginable mess.

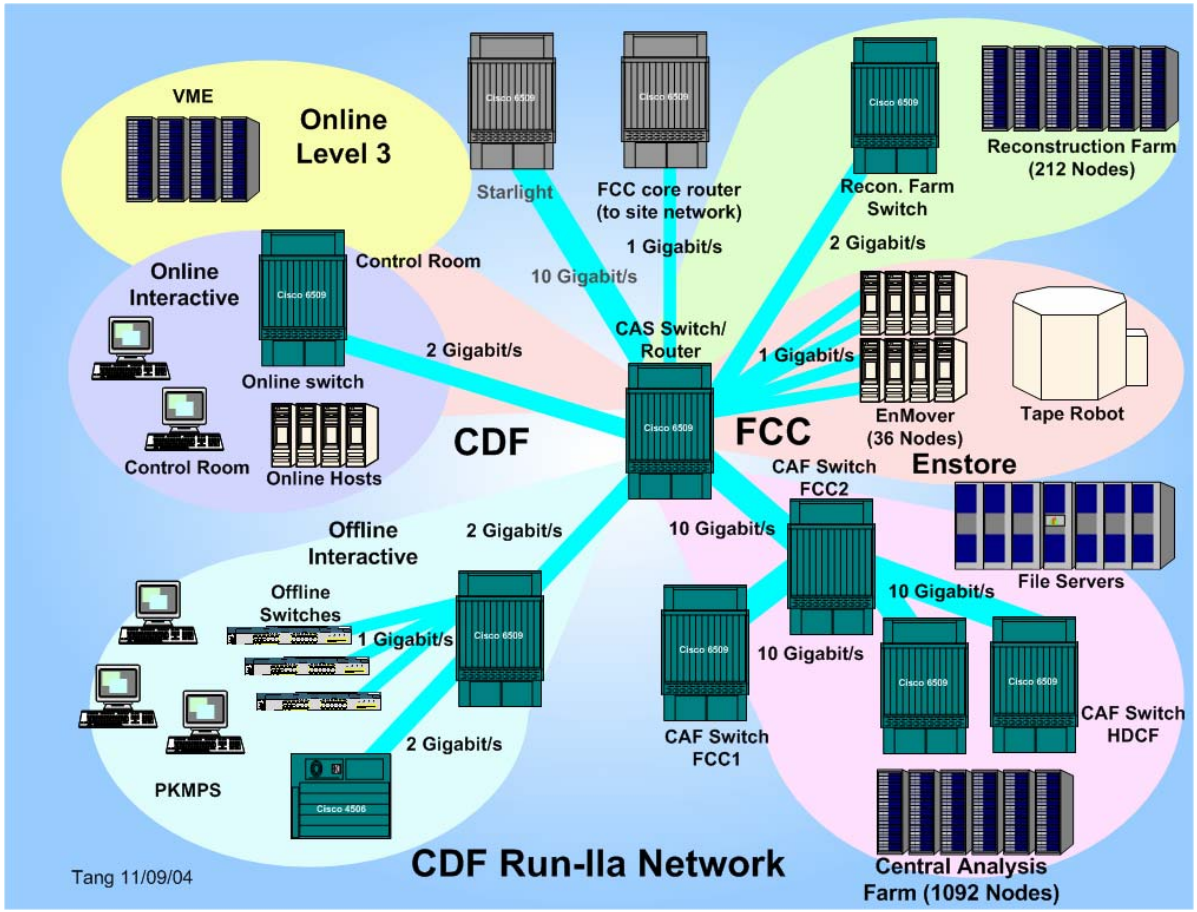
Conclusions



- There are plenty of risks and uncertainties in infrastructure (buildings and networking)
- There are some uncertainties in needs
- We do/will keep on top of projections and adjust the plan as needed
- The major investment in buildings appears to be on track
  - Grid Computing Center – additional rooms
- Wide Area Networking and Transatlantic Networking is still a large budgetary risk

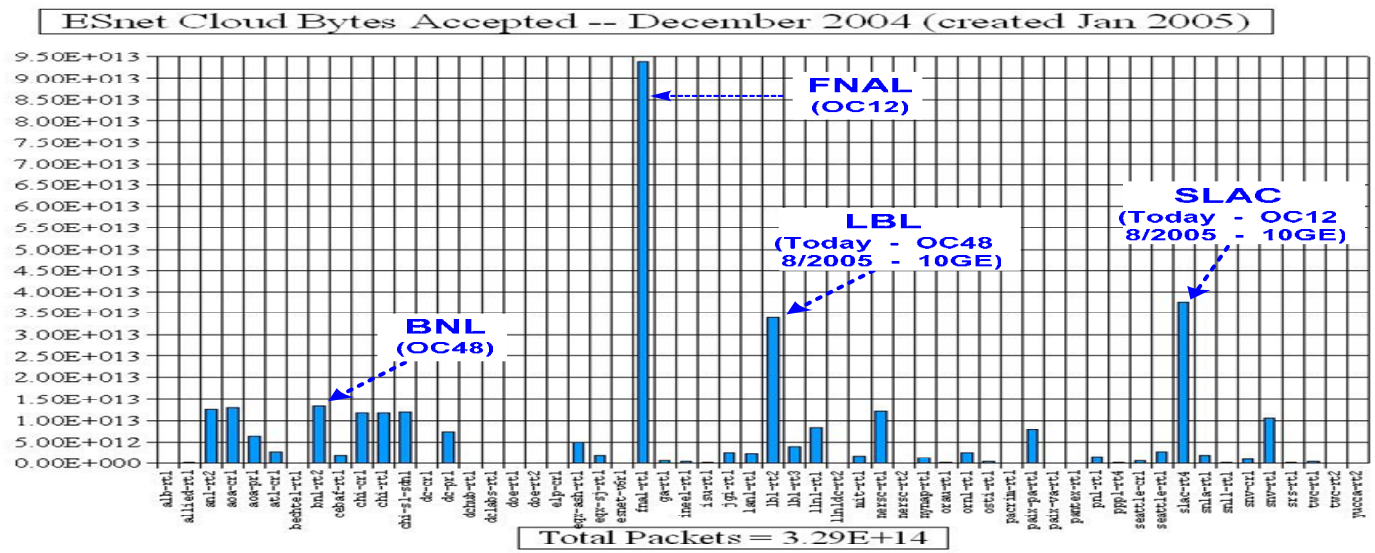
Spares

CDF: One of several major LANs



Esnet futures…

- ESnet link upgrade unclear:
  - Proposed ESnet Chicago metro area network (MAN) in limbo:
    - Funding is the bottom line issue
    - But ESnet Bay area MAN deployment is proceeding…
  - FNAL investigating 2nd fiber to StarLight for ESnet link



GCC Power Fractions