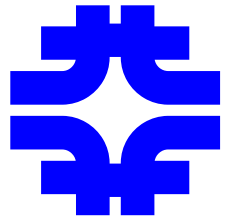


# Fermilab Site Report

## Fall 2010

Keith Chadwick  
Grid & Cloud Computing Department Head  
Fermilab  
[chadwick@fnal.gov](mailto:chadwick@fnal.gov)

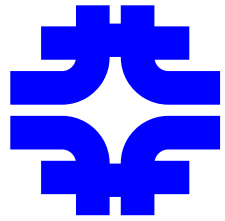


## Some Preliminary Remarks...

---

I am reporting about the work that is being done by the many talented, hardworking and dedicated people at Fermilab.

Please credit those individuals where credit is due and blame me for any mistakes or misunderstandings. 😊

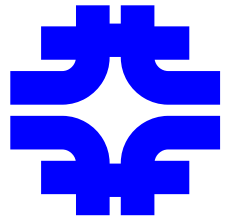


# Fermilab Computer Centers

---

Fermilab operates three computing centers:

- Feynman Computing Center (FCC)
- Grid Computing Center (GCC)
- Lattice Computing Center (LCC)



# Feynman Computing Center (FCC)

---

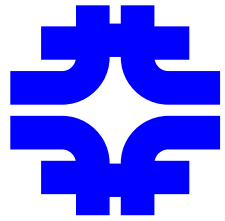
## FCC1 + FCC2 Computer Rooms share:

- Generator, 4 x UPS and chiller cooling infrastructure.
  - UPS-1 576 kVA, UPS-2 60 kVA, UPS-3 100 kVA, UPS-4 100 kVA.
- Historically, the chiller heat exchangers and availability of sufficient quantities of cooling water have been the largest cause of service interruptions.
- This is being addressed using ARRA funds.

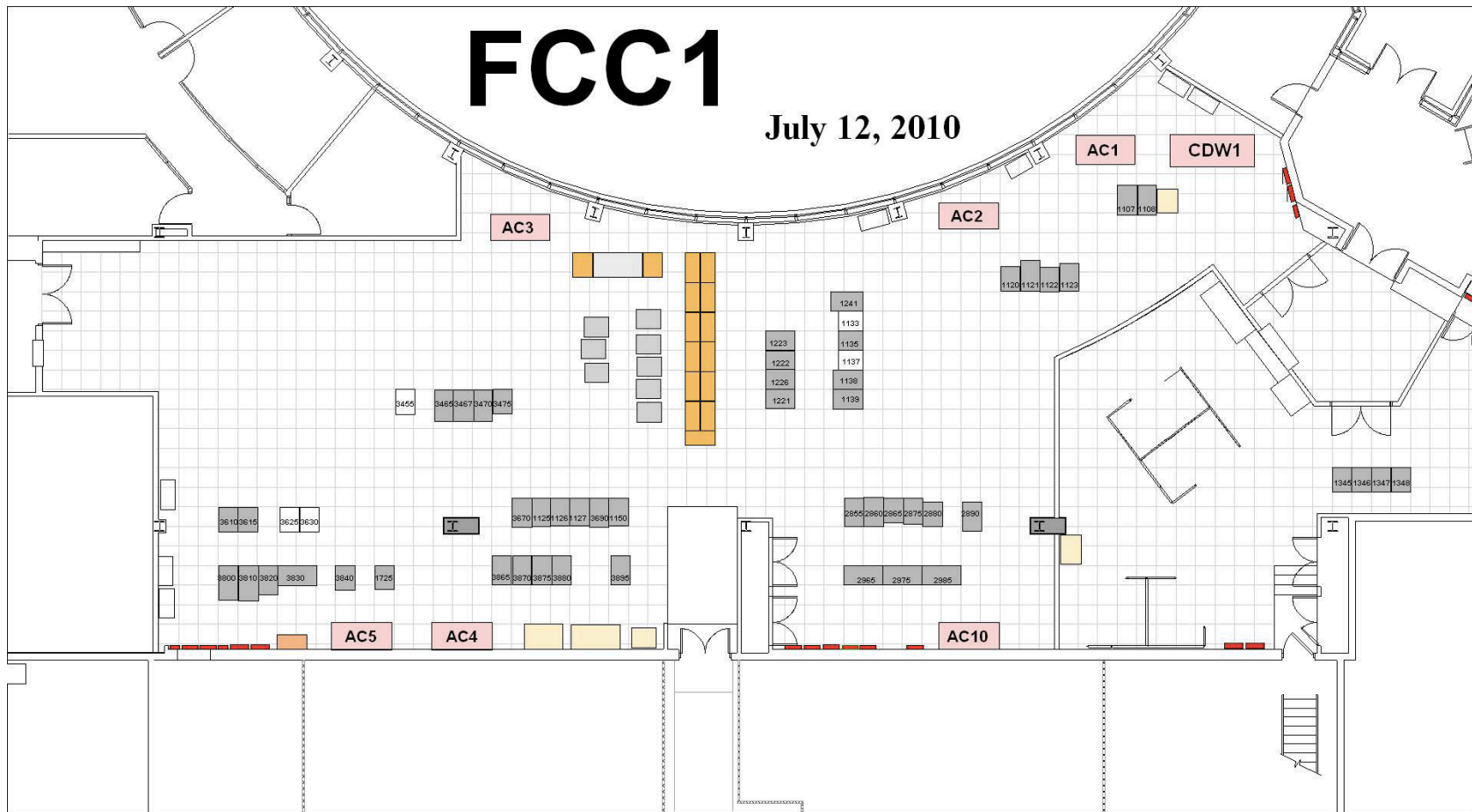
## FCC3 “high availability” Computer Rooms:

- Built using ARRA funds.
- Two new computer rooms with new power, cooling, UPS, generator and feeder infrastructure.
- Target beneficial occupancy ~ 2<sup>nd</sup> or 3<sup>rd</sup> Week of November 2010.

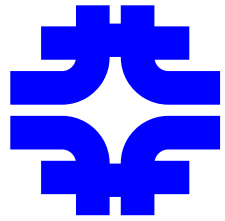




# FCC1 - Present







# New FCC2 Vestibules

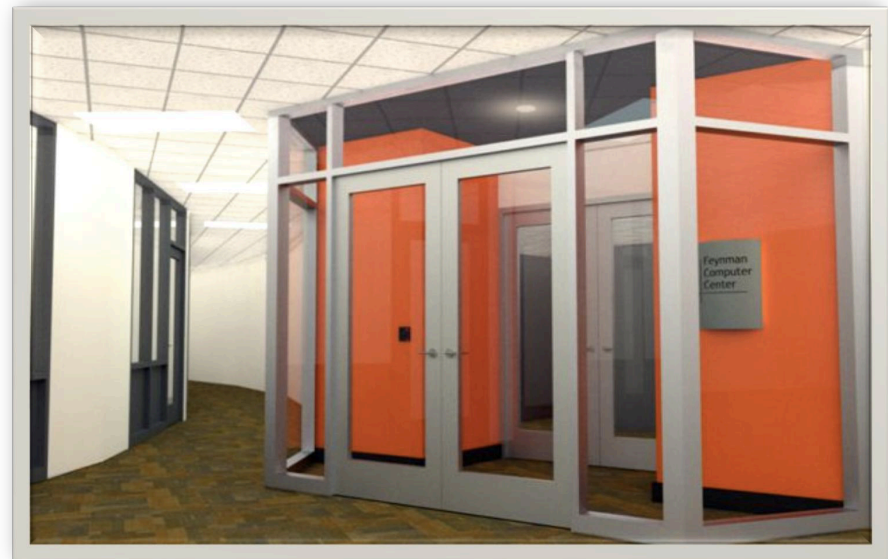
New vestibules will shortly be constructed on the east and west entrances to the FCC2 computer room:

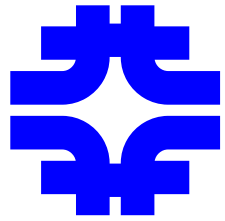
The Goals are:

- To reduce the amount of dust and debris in the computer room.

and

- Minimize the escape of conditioned air when equipment is moved in and out of the computer room.

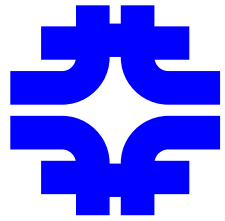




# Delivery of new FCC2 CRACs







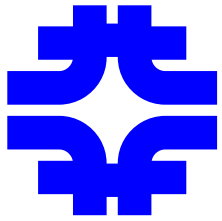
# New FCC2 CRACs

---

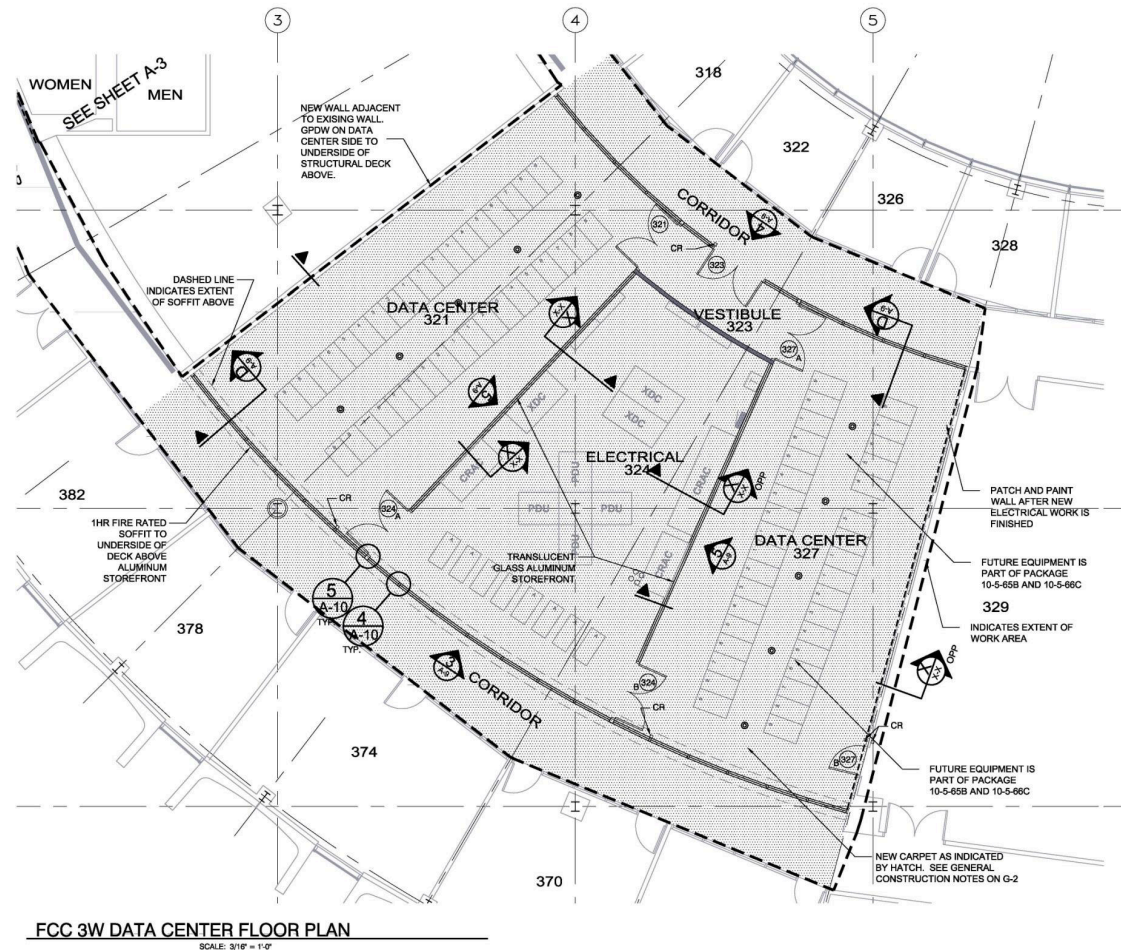
8 x 30 ton capacity CRACs

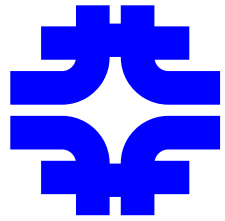
210 (n+1) tons of AC.





# New FCC3 Computer Rooms





# Overview of FCC3 Computer Rooms

---

## High Availability low power density capacity

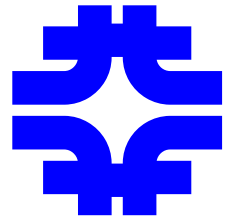
- UPS and standby generator.
- Racks 5–6 kW.
- Cooling & electric decoupled from existing FCC.
- No raised floor, no dropped ceiling.
- Overhead cooling only.

## 2 rooms with 2 rows of 18 racks (71 racks total)

- Dual cording capability in each rack position.
- One rack position occupied by building support beam.

## Lots of window glass on north and south sides

- Expect less tours inside the room.
- People will be able to view into the rooms.

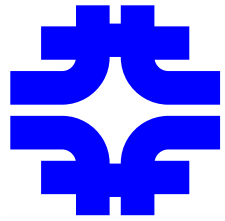


# FCC3 Computer Rooms – 1

---

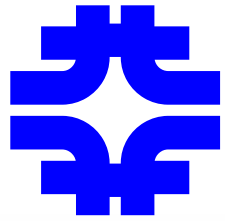






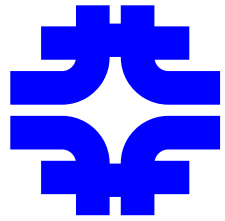
# New Substation and Generators





# New Rooftop Condensers



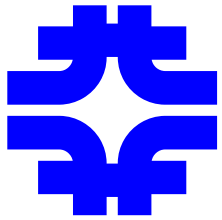


# All roads lead to ~~Rome~~ FCC (before going anywhere else)

---

Since the FCC Computer Rooms are supported by generator + UPS, it is the “mother ship” of the Fermilab computing infrastructure.

- Many critical services
- Central compute servers
- Central disk servers
- Centrally managed databases
- Laboratory financial services
- Site network hub
- Offsite network access point via Chicago MAN.



# Grid Computing Center (GCC)

---

## Three computer rooms:

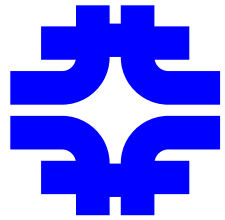
- Grid Computer Room A (GCC-A).
  - 80 relay racks, ~3,000 systems, ~750 KW power/cooling.
- Grid Computer Room B (GCC-B).
  - ~2,500 systems, ~840 KW power/cooling.
- Grid Computer Room C (GCC-C).
  - ~1,000 systems, ~840 KW power/cooling.

## Two network rooms:

- GCC Network Room A
  - GCC Hub & Serves GCC-A and GCC-TRR
- GCC Network Room B
  - Satellite of GCC Network Room A
  - Serves GCC-B and GCC-C.

## Tape Robot Room (GCC-TRR):

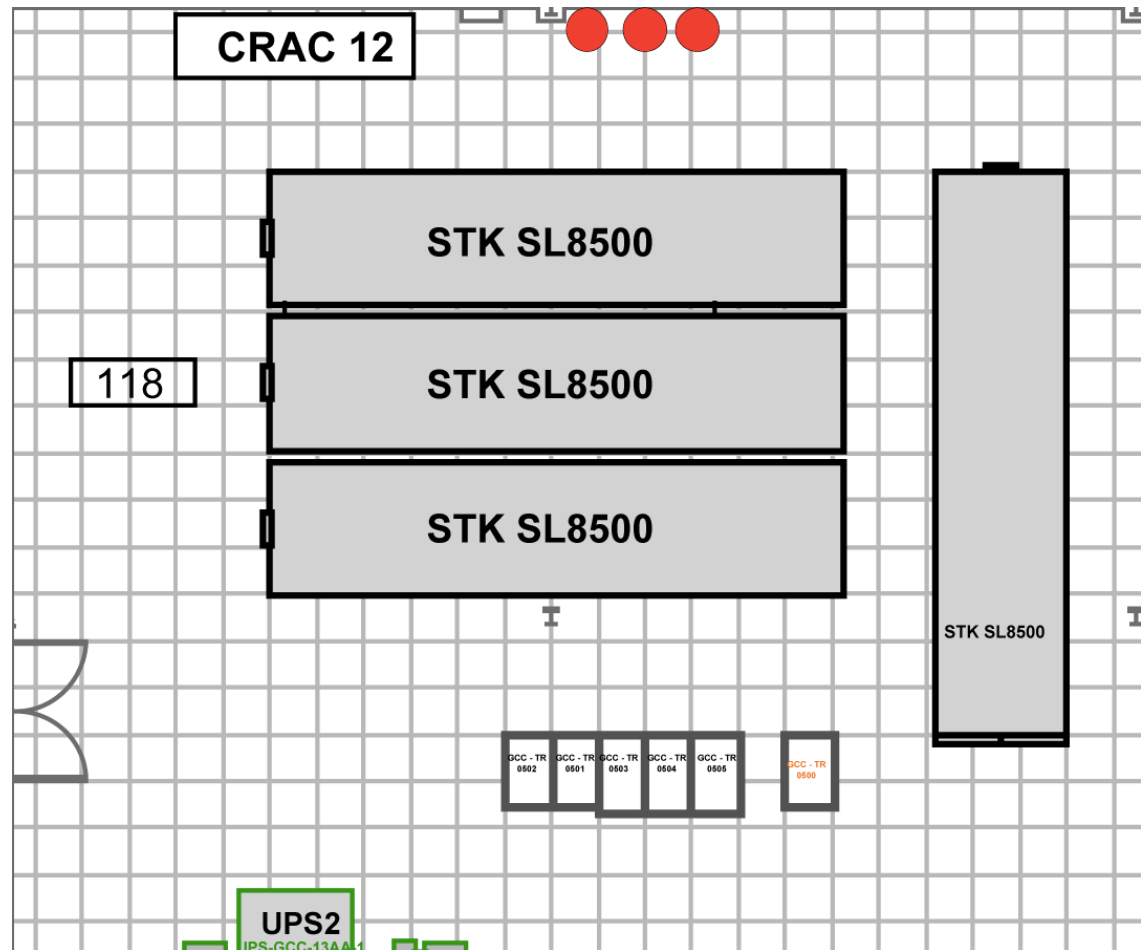
- Four STK SL8500 Robots

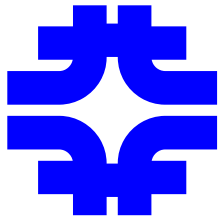


# GCC-TRR

Qty 4 STK SL8500  
tape robots:

- Qty 3 for CMS data,
- Qty 1 for anyone else.





# GCC-A Computer Room

In operation since late 2004.

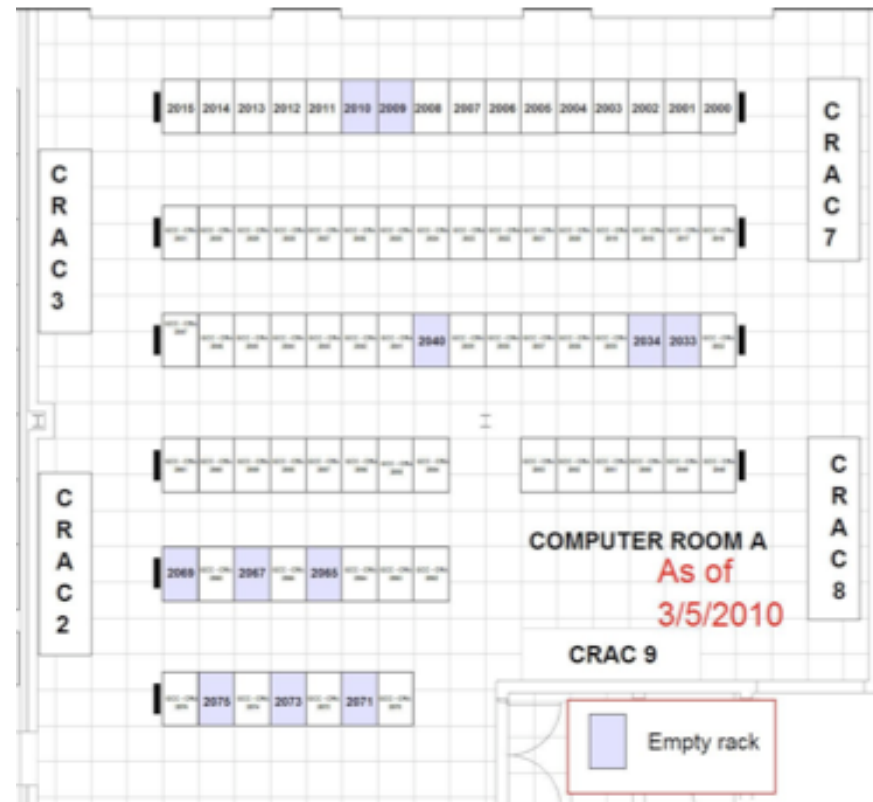
- Designed for 10 KW racks,
- 120 VAC Power Distribution.

Approximately 3,000 systems

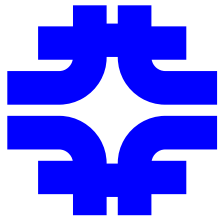
- Mix of dual and quad core systems,
- Systems have 2 to 4G of memory,
- Average of ~ 2GHz/cpu.

Was under consideration to be stripped down and re-engineered, but wound up being repurposed due to:

- Lack of budget
- FCC power issues







# GCC-B Computer Room

Commissioned in early 2007.

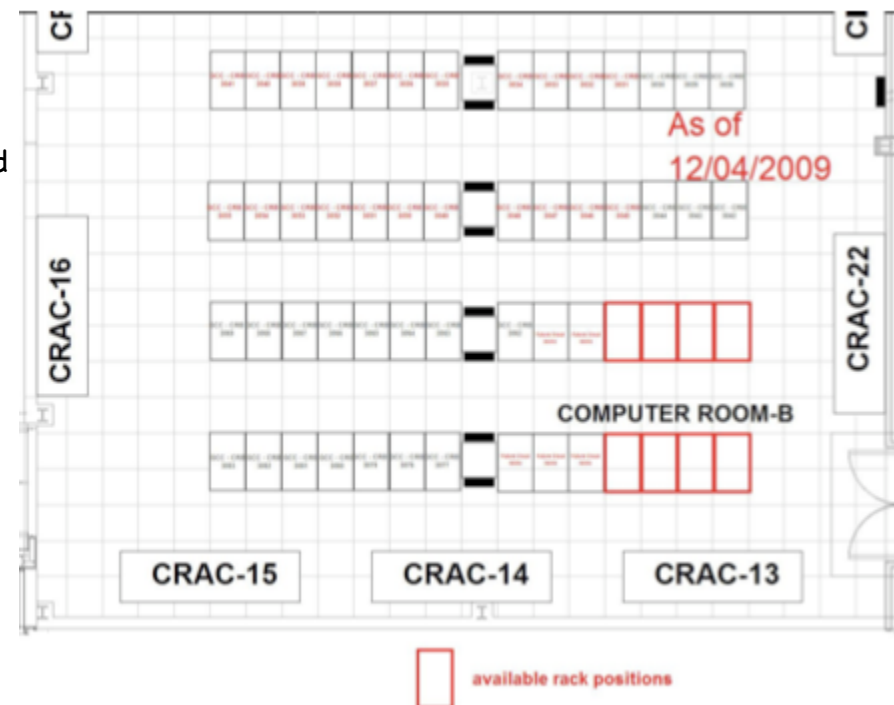
- Designed for 10 KW racks,
- 120 VAC Power Distribution.

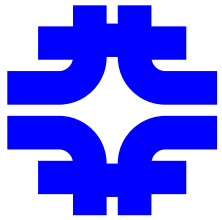
Approximately 2,500 systems:

- Systems are a mix of quad (dual-dual) or dual-quad core.
- Oldest have 16 GB of memory,
- Newer have 24 GB of memory.

Has (just) been filled by FY2010 acquisitions.

- Supermicro H8DGU-F - AMD SR5670/SP5100 Chipsets, Dual Socket 6100 series.
- 8 core AMD Opteron 6134 2.3GHz,
- 2 x 8 core = 16 cores/system
- 32 GB of memory/system.
- 36 systems/rack.
- 576 CPUs/rack.





# GCC-C Computer Room

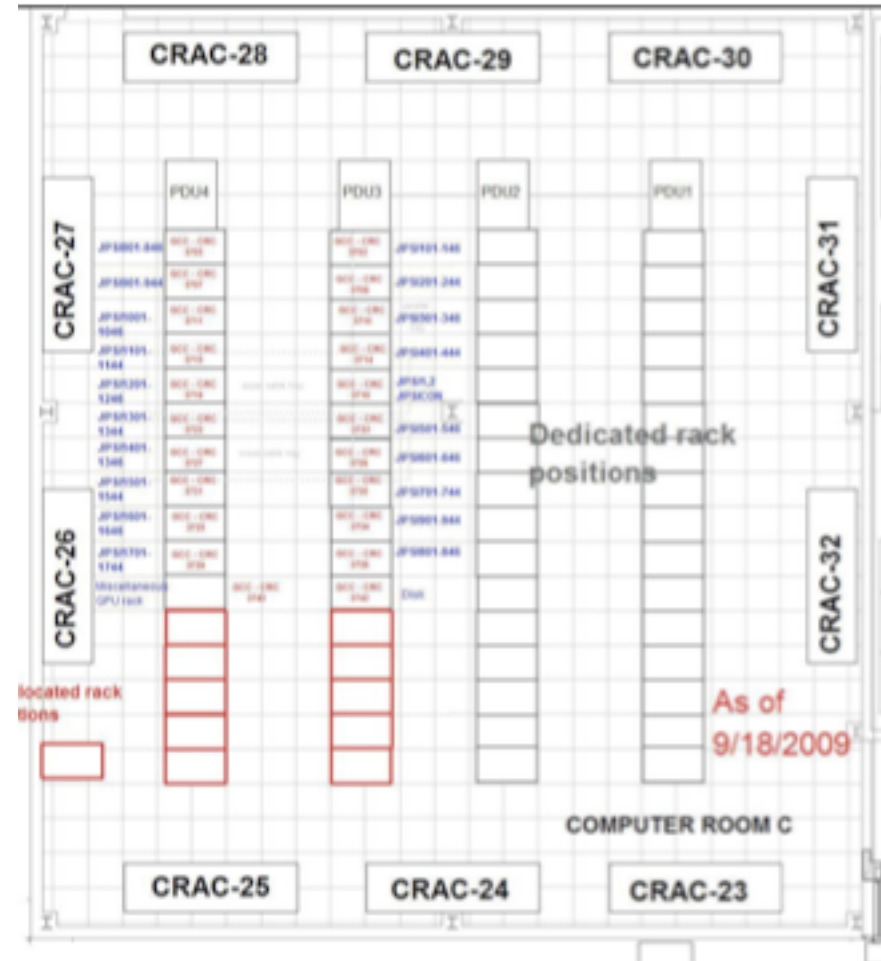
Commissioned in the summer of 2008.

- Designed for 14 KW racks,
- 208 VAC power distribution.

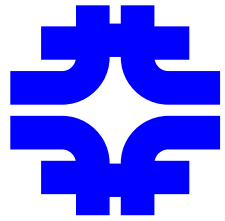
Hosts the Lattice QCD J/Psi cluster.

Has been filled by the FY2010 Lattice QCD acquisition  $D_s$  cluster.

- Super Micro H8QGi-F - AMD SR5690/SP5100 Chipsets, Quad Socket G34
- 8 core AMD Opteron 6128 2.0GHz
- 4 x 8 = 32 cores/system
- 16 x 4GB DDR3-1333 = 64 GB/system



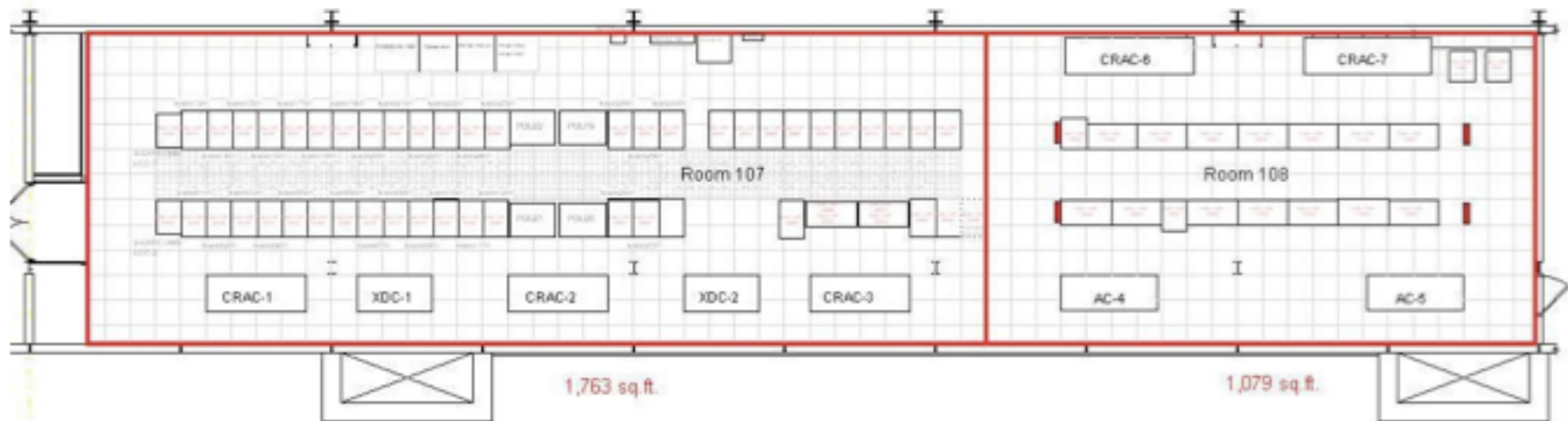


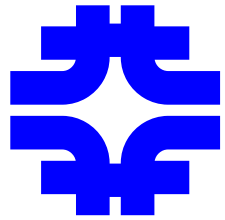


# Lattice Computing Centre

Lattice QCD “pion” cluster systems have been removed from room 108.

Test and development systems have been relocated from FCC to LCC.





# Tape & Robots

---

Space has been identified for two additional SL8500 Robots in FCC.

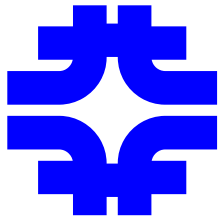
- We will likely be ordering one of these robots later this (fiscal) year.

Redundant “bots” have been installed in the CMS SL8500 Robots.

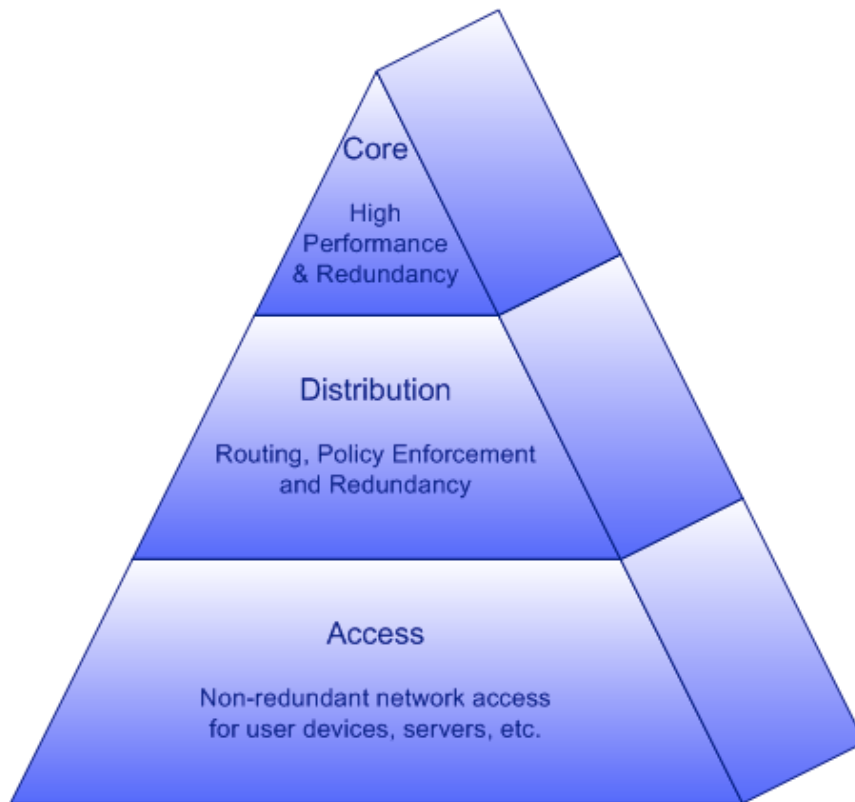
- Goal is to reduce downtime due to “bot” failure.

We are starting the process to migrate to LTO5:

- LTO5 drives are being installed in SL8500 robots in GCC;
- New “movers” are being commissioned;
- LTO5 tapes are on order;
- The plan is to run 24x7 tests on the new LTO5 drives and tapes through January;
- Once the drives and tapes pass the acceptance criteria, additional drives and tapes will be ordered;
- Goal is to have production LTO5 deployed for CMS by 1-Mar-2011.



# Network Architecture



## Modular Architecture

- Performance/Redundancy Pyramid

## Core Layer

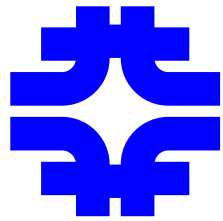
- Multiple individual Layer 3 switches or routers, optimized for data forwarding.
- Diverse fiber-optic paths between computer rooms and strategic locations.
- Currently Catalyst 6506; plan to upgrade to Nexus 7009.

## Distribution/Aggregation Layer

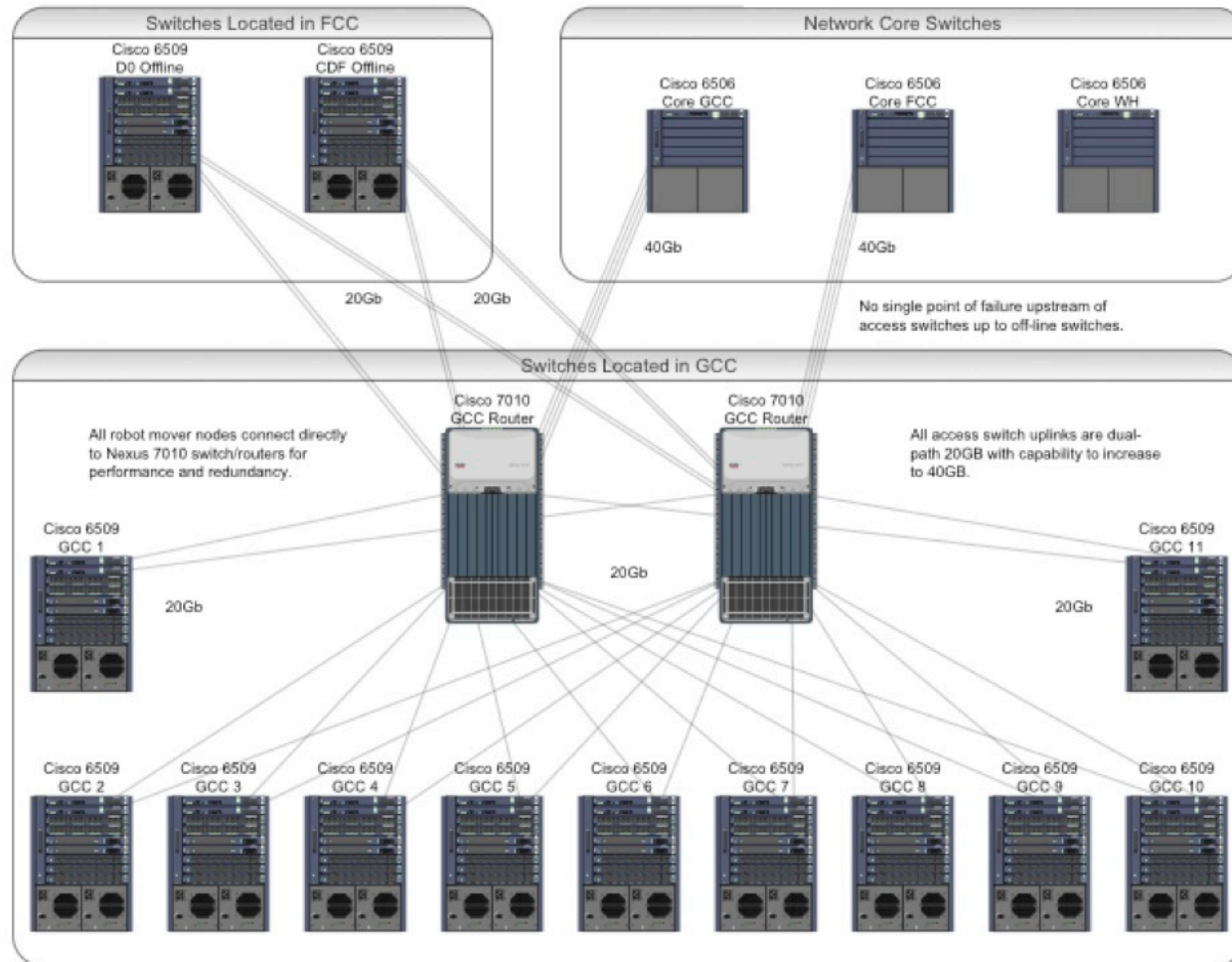
- Pairs of Layer 3 switches or routers, each with uplink redundancy to Core switches.
- Configured for redundant access-layer routing to function as the “default gateway” for end systems.
- Configured to control uplink and downlink access via Access Control Lists or Firewall features.
- Currently Catalyst 6509; upgrade to Nexus 7010 in progress.

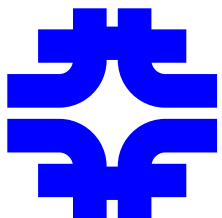
## Access Layer

- Individual Layer 2 switches, each with uplink redundancy to Distribution switches.
- End system network access redundancy up to end system manager.
- Currently “anything/everything”; upgrade plans being developed.

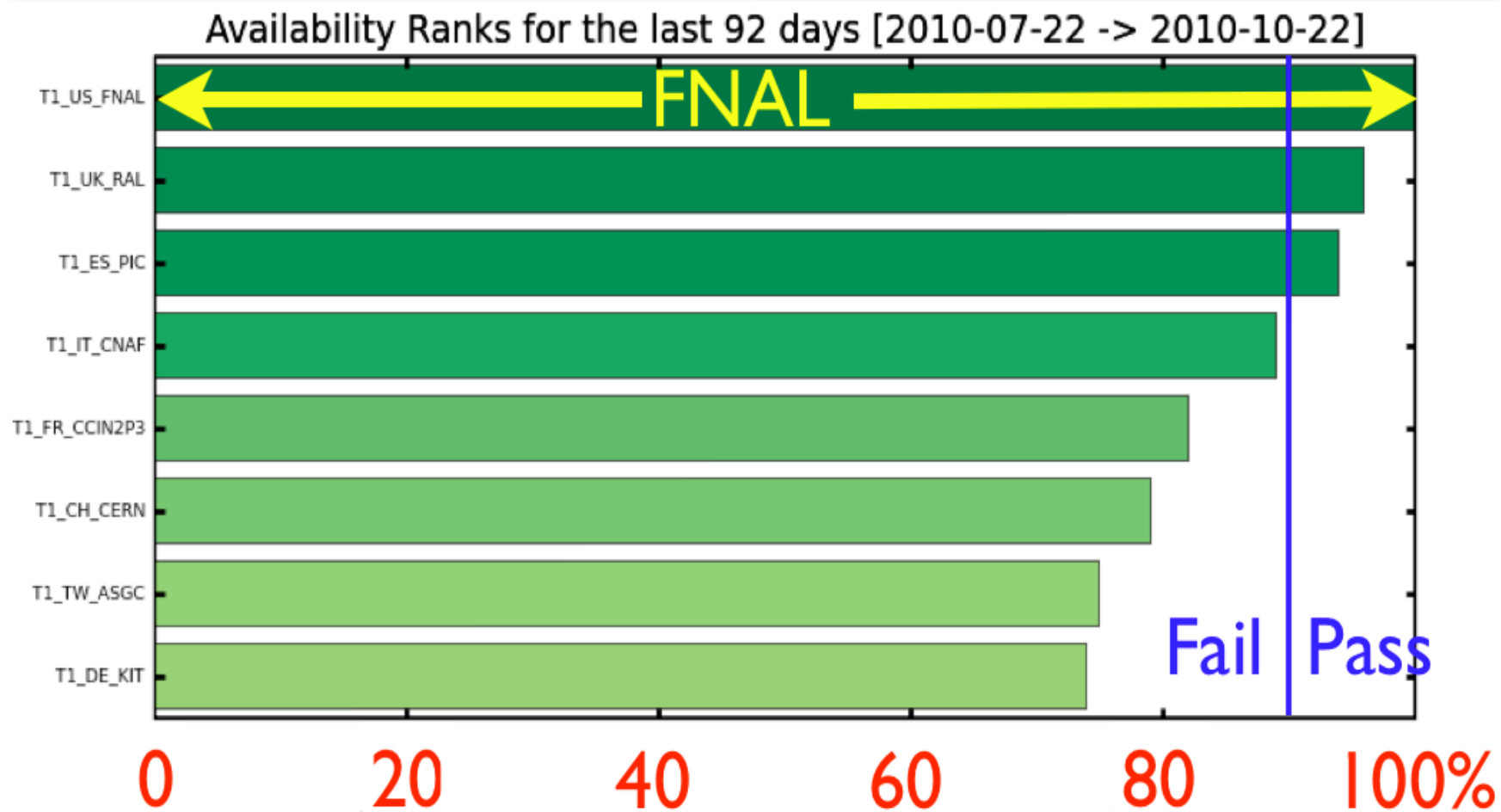


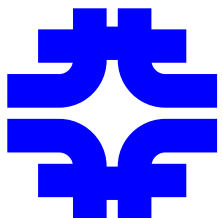
# GCC Network Topology



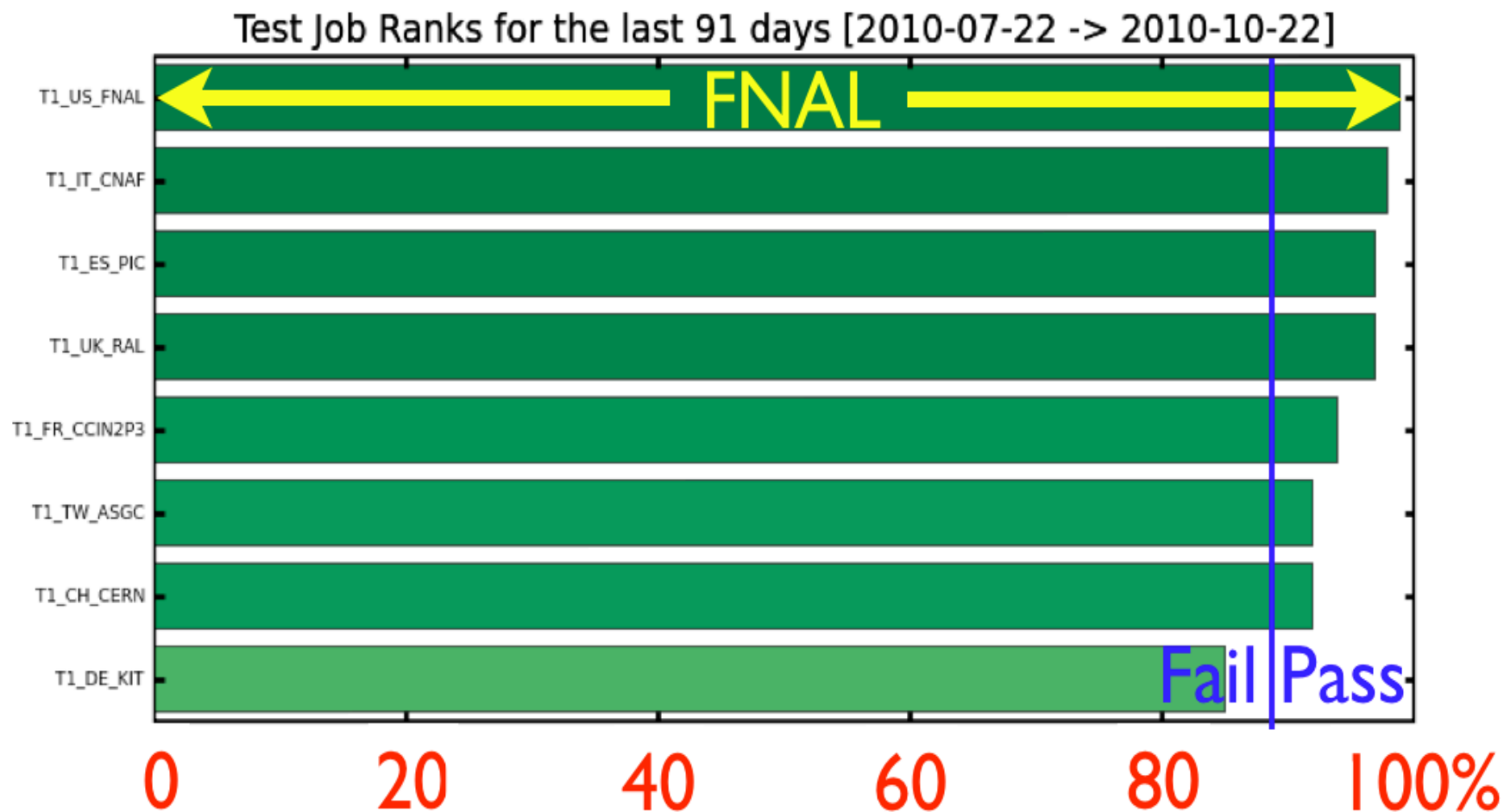


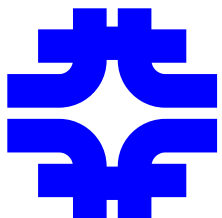
# CMS Availability Rank



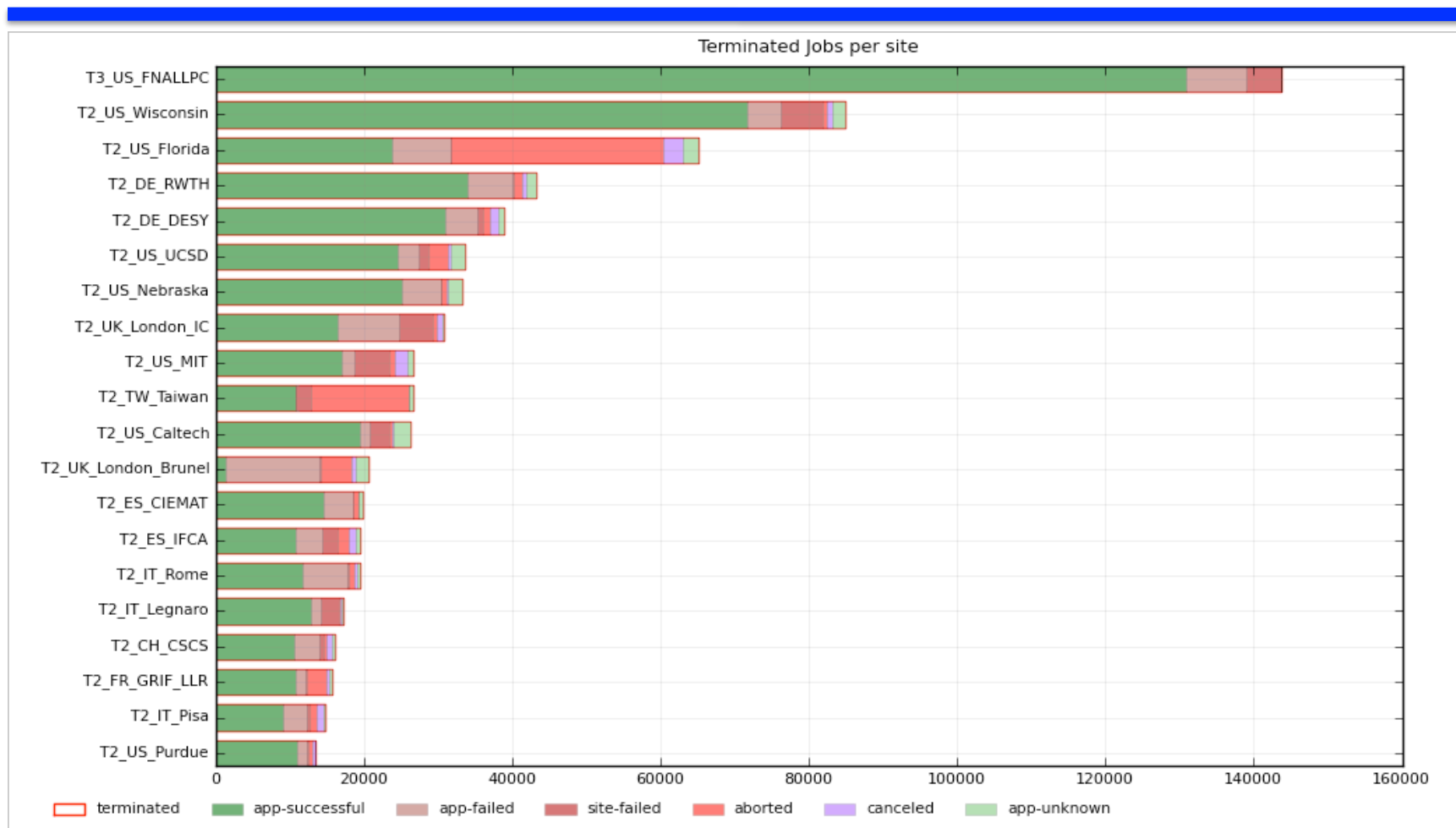


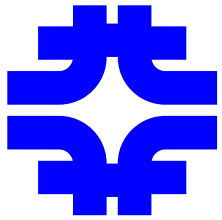
# CMS Test Job Rank





# CMS Analysis Jobs

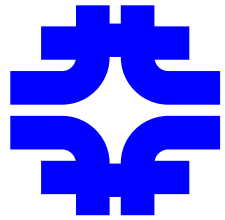




## Measured FermiGrid Service Availability for the Past Year\*

Service	Availability	Downtime
VOMS-HA	100%	0m
GUMS-HA	100%	0m
SAZ-HA (gatekeeper)	100%	0m
Squid-HA	99.988%	59.97m
MyProxy-HA	99.985%	78.80m
ReSS-HA	99.979%	107.96m
Gratia-HP	99.616%	2,014.32m
Database-HA	99.867%	697.73m





# ASHRAE Guidance

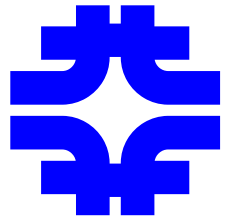
---

## 2004 Recommended Environmental Envelope

- 20 to 25°C (68 to 77°F) (dry bulb temperature)
- relative humidity (RH) range of 40 to 55%
- ...refers to the inlet air entering the equipment

## 2008 TC9.9

- 18 to 27°C (64.4 to 80.6°F)
- Data center humidity levels should now be measured by dew point and fall within 5.5 degrees Celsius to 15 degrees (41.9 degrees Fahrenheit to 59 degrees). The previous range was narrower and measured by relative humidity, which ASHRAE decided wasn't as accurate a metric as dew point.



# Computer Room Optimization

---

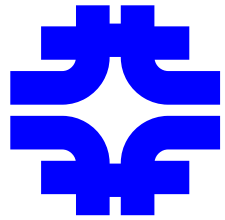
Return Air set points are generally 73F (23C)

- Air delivered to front of servers in lower 60sF (16C)

Currently the Fermilab computer rooms are operating outside of ASHRAE guidance.

To address this, the facilities personnel are proposing an increase of temperature by 5-10F (3-5C).

- Goal: Operate the computer rooms more efficiently
- Energy savings



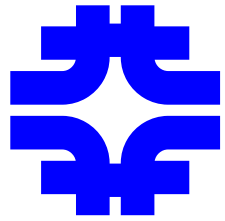
## About those UPS systems...

---

Comment from a recent DOE/Federal Energy Management Program (FEMP):

- “Eliminate UPS for scientific computing – currently all is on UPS but there is no generator backup. Utility power is reliable so UPS is not often needed and its conversion loss can be avoided.”

If this is adopted, running the computer centers will be a significant challenge for both the facility and the service operators...



# Planned FCC Changes

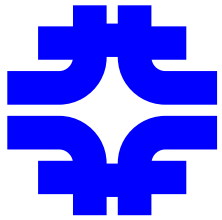
---

## FCC2 Computer Room:

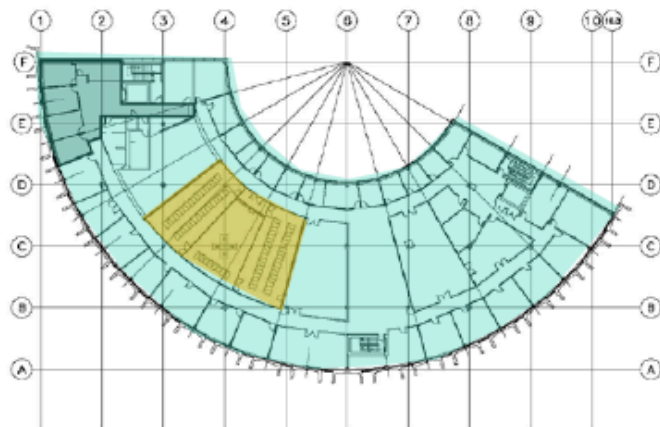
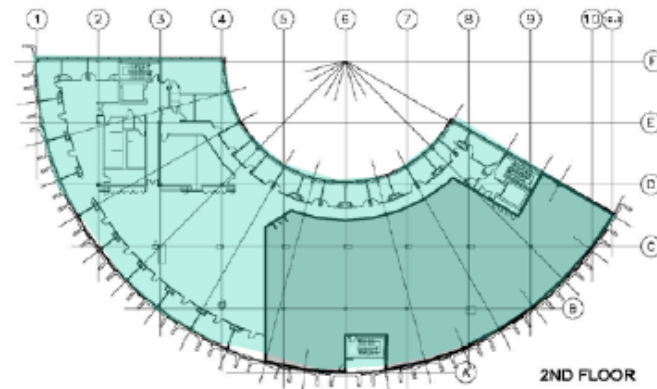
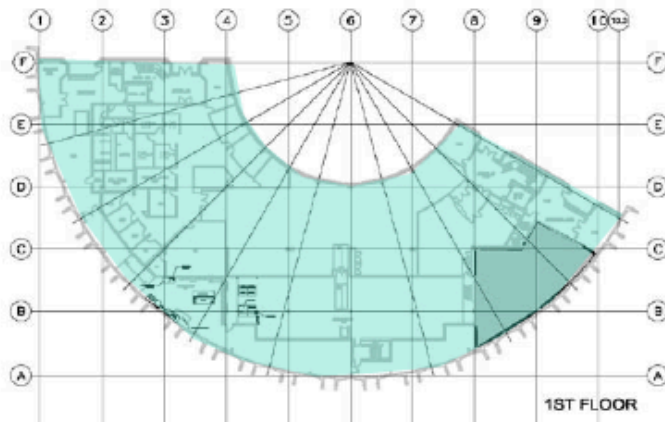
- Switch from chiller based CRACs to refrigerant based CRACs (in process);
- Replace 4x UPS with 1 larger UPS;
- Install “new” higher capacity Generators;
- Upgrade electrical distribution system.



## 3<sup>rd</sup> and 4<sup>th</sup> Computer Room on FCC3

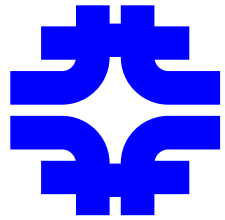
- Not yet funded;
- “Duplicate” of the two new Computer Rooms on FCC3.



# Electrical on FCC1, FCC2, FCC3 “~Today”



-  System 1 – Existing
-  System 2 - New FCC3



# FCC2 Electrical Upgrade

---

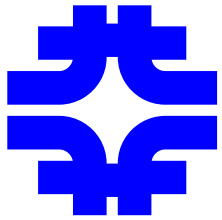
Independent System from Existing FCC and FCC3

Backed by 750 kVA UPS (replaces 400 kVA UPS)

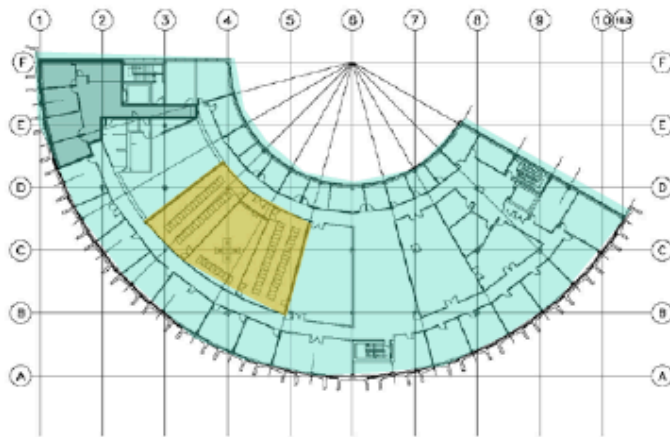
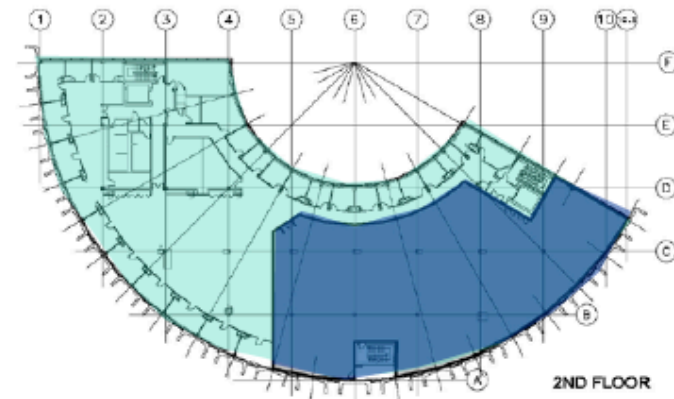
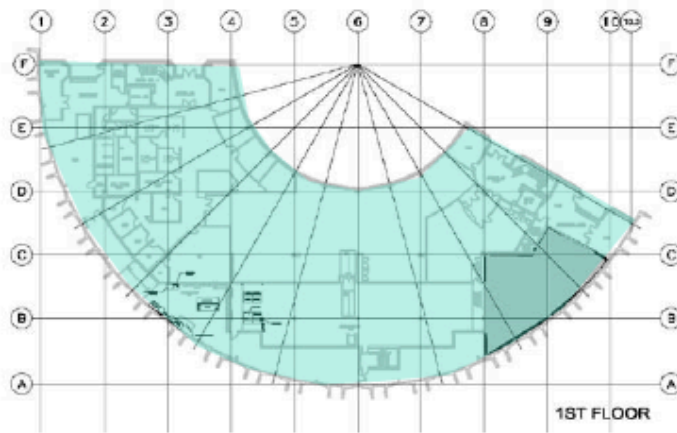
New Standby Diesel Generator Set (system 3):


- Dual 800 kVA units
- Existing “building wide” (system 1) unit will also remain.

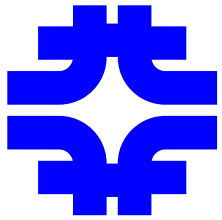
Includes backup power for CRAC units



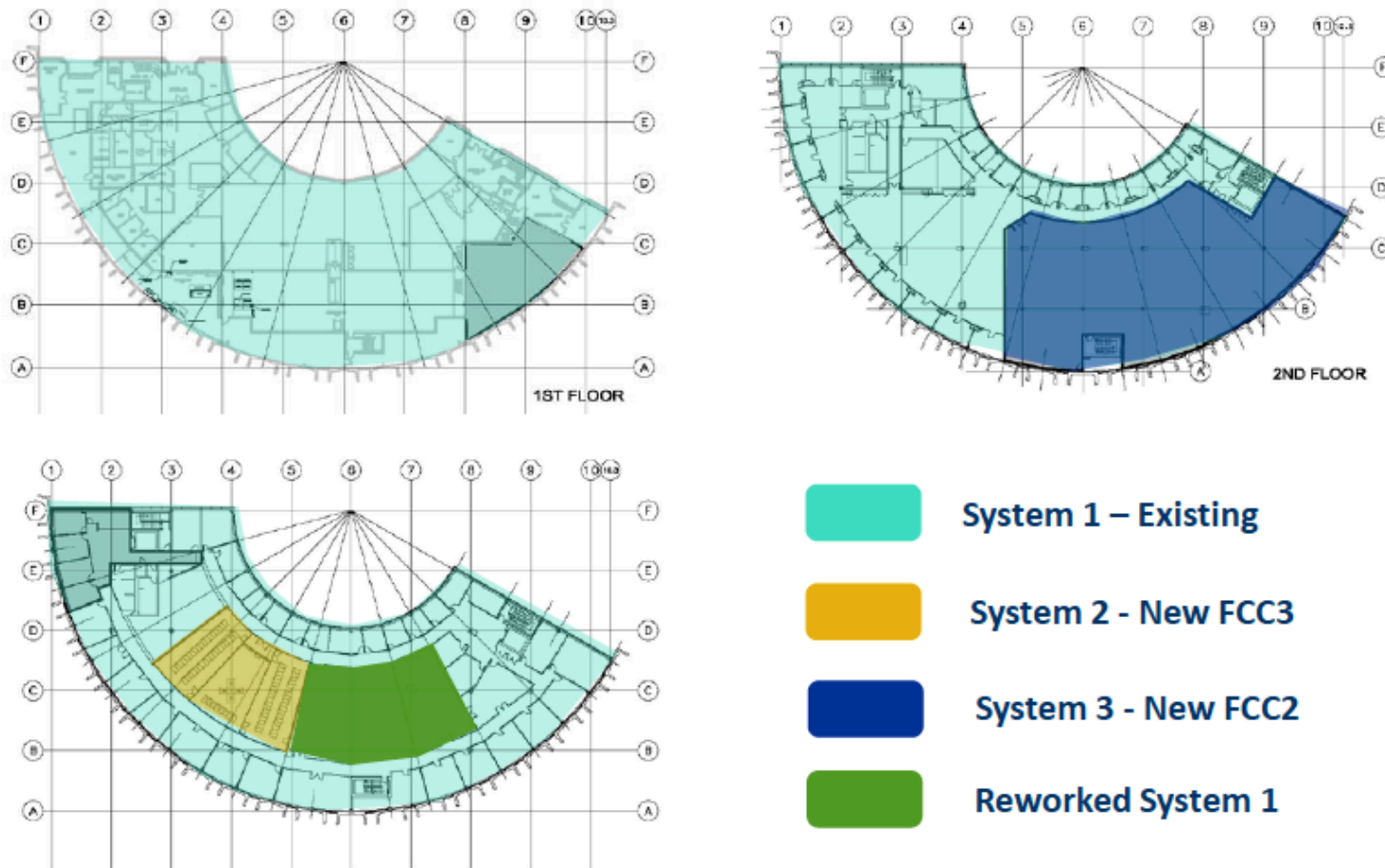
# Electrical on FCC1, FCC2, FCC3 “Today+~1 Year”



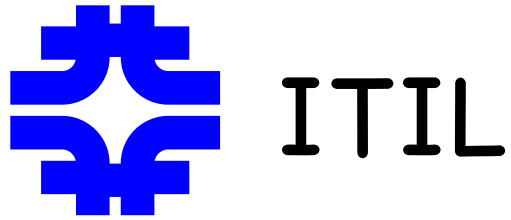
-  System 1 – Existing
-  System 2 - New FCC3
-  System 3 - New FCC2



# Electrical on FCC1, FCC2, FCC3 “Today+???” (if funded).







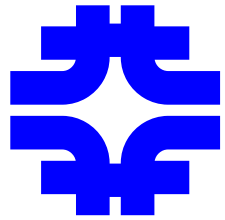
---

CD has commissioned the following ITIL processes:

- Incident Management
- Problem Management
- Change Management
- Release Management

The following ITIL processes are in the process of being commissioned:

- Service Level Management
- Configuration Management
- Continuity Management

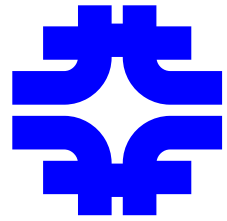


# Conclusions

---

The personnel of Fermilab Computing Division are meeting the challenges of today, and making plans to deal with the challenges of tomorrow:

- We have just (about) filled the GCC computer rooms,
- We have just (about) taken delivery of the new FCC3 computer rooms,
- We are starting the work to increase the capacity of the FCC2 computer room,
- We have preliminary plans to add two additional computer rooms on FCC3 (if needed and if budget is available),
- The Fermilab network is evolving to offer even higher levels of service, reliability and availability,
- The Fermilab service providers are adopting ITIL practices and delivering improvements to operations.



Fin

---

Any Questions?