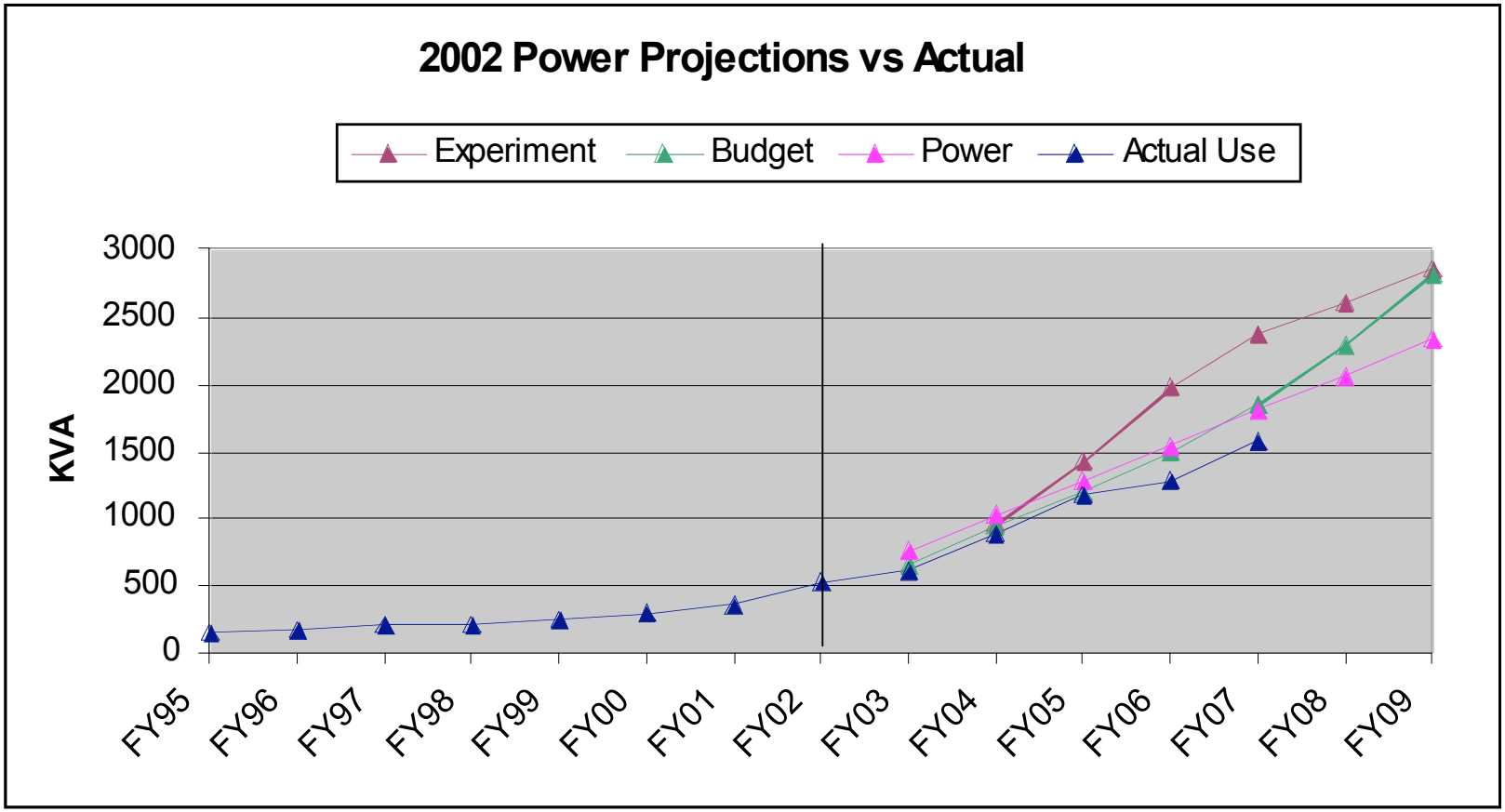
Power and Air Conditioning Challenges in Computer Centers

Amber Boehnlein
For Don Petravick
Fermilab
September 6, 2007

The Challenge

- A common challenge in facilities planning for computing at the scale required for HEP experiments is provisioning for the space, power and cooling required by today's inexpensive commodity hardware
- Facilities are expensive to build or outfit and typically have a much longer lifetime than does the computing hardware.
- Due to the expensive and constraints, there is no single solution to facility provisioning
  - Some experience from FNAL and ANL are presented here.

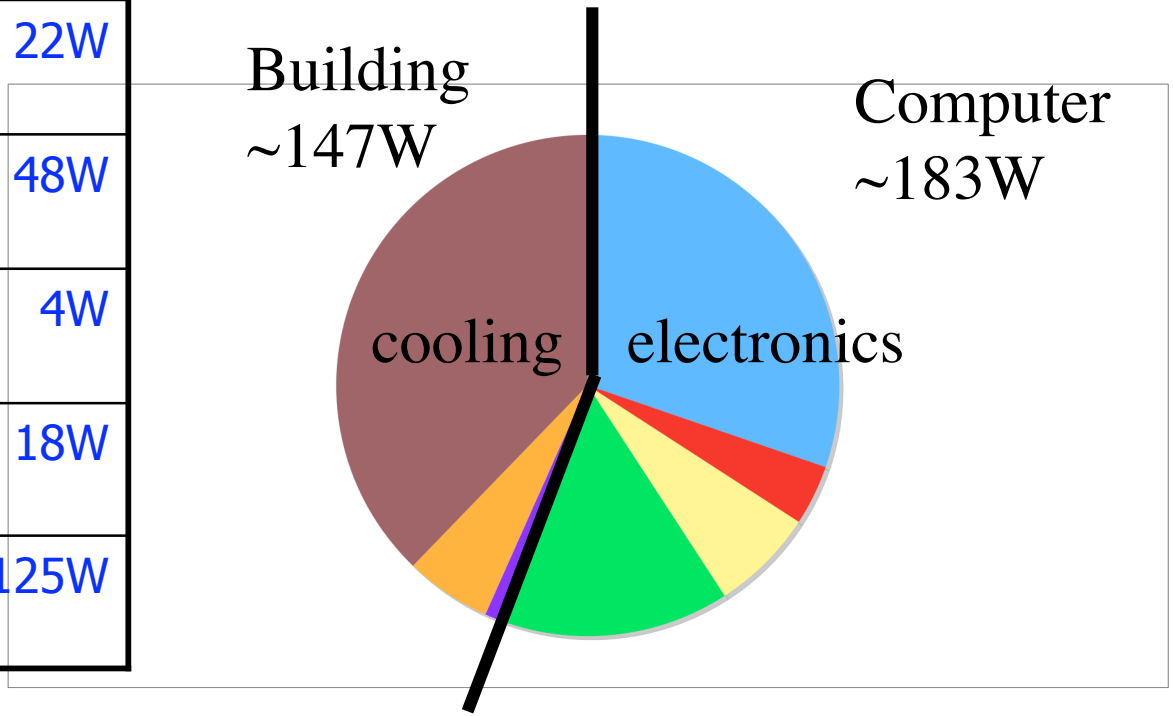
Fermilab 2002 Power Forecast



Power Distribution

Electronics	100W
Server Fans	13W
Voltage Regulation	22W
Case Power Supply	48W
Room Power Distribution	4W
UPS	18W
Room Cooling	125W

It requires ~330 Watts to provision
For 100 Watts of electronics

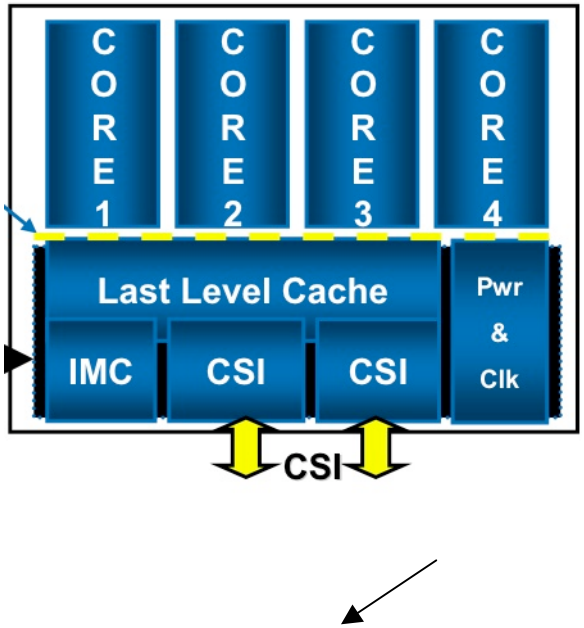


Four layers of integration

- In considering the power profile, it is useful to consider four layers of integration corresponding to the power profile
  - The processors themselves
  - Processors integrated into boxes
  - Boxes integrated into racks
  - Racks integrated into a building

Power and processor performance

- $P \sim {}^1/_2 C V_0^2 \omega^3$
  - $P \sim \omega^3$ –> multi-core processor.
  - $P \sim V_0^2$ -> low voltage techniques.
- Some optimizations speed up a fixed computation, but also speed up the power needed to perform the computation.
  - E.g. speculative execution



Server level

- Power supplies of increasing efficiency.
- Power more efficient power transformations within the box.
- Potential for card level integrations **
- An increasing number of cores, at relatively constant core speed.
  - Power to the memory system becomes significant.
  - Memory exists in its own commodity marketspace.
- Continued air cooling of commodity servers seems most likely.

Rack level integrations

- Correct integration with heat flow in the building.
  - High ΔT makes cooling more efficient.
  - Power saved in cooling is available for computing.
  - Computer Room B@FNAL -- currently 10 kW/rack
  - Air cooled racks over 20 kW are possible, with designs such as chimneys.
  - Examples of water cooled racks at GridKa, BNL
- Some consider a single, more efficient DC supply in racks.

Buildings/facilities(1)

- Space considerations can be as relevant as power/cooling
- Swing spaces
  - Scientific computing -- bulk purchases.
  - There is a rhythm to these retirements
    - FIFO
    - a 3 or 4 year duty cycle is common
  - When procuring annually, an upper bound on swing space is the fraction retired.

Buildings/facilities(2)

- Increasing average power density of computing equipment implies re-assessment and possible refurbishing of computing spaces.
- However, playing with infrastructure in an active space carries potential for service disruptions.
- *Fallowing* is the notion that one systematically empties a space of computing, to refurbish the power and HVAC.

Buildings/facilities (3)

- Advantages of fallowing
  - Power/cooling can be built to specific requirements layout of systems.
  - Need to get it right for the next few purchasing cycles, not once and for all.
    - Purchase cooling and power late, just in time.
    - Great potential for high ΔT, hence efficiency and greenness.

Type of Space

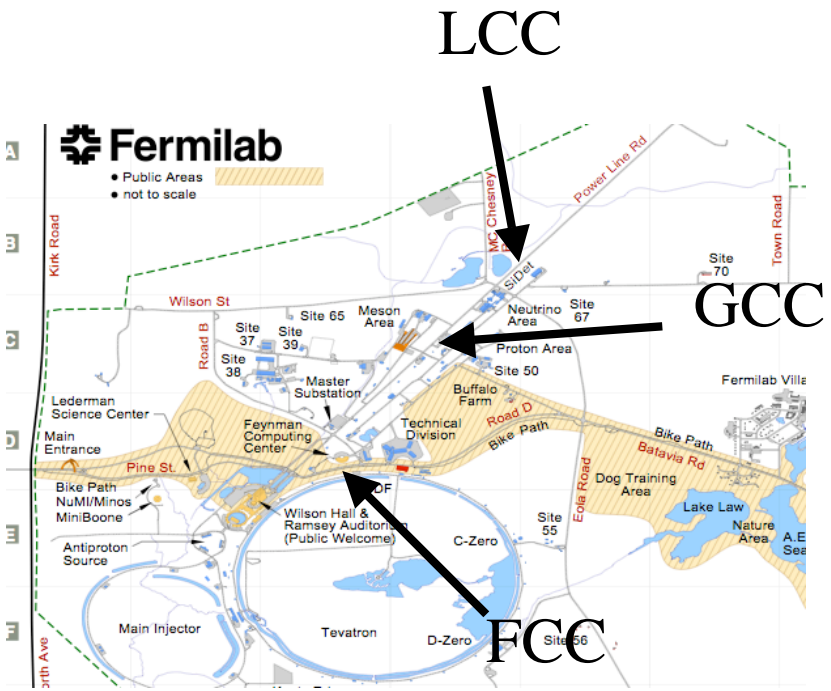
- Computing Centers typically have systems with different up-time and accessibility profiles.
  - Space for state-full computing systems.
    - UPS, UPS + generator
  - Space for new, nearly stateless computing systems.
    - No UPS, UPS and roll up generator
  - Space for long-lived, “tape” storage.
    - UPS, generator, minimal cooling demands

Many ways to a good facility

- Because of the capital costs involved, the optimal way forward varies from site to site.
- At Fermilab
  - Opportunity -- fixed target area with copious power.
  - Constraint -- reuse buildings, limited funds each year.

FNAL Computing Centers

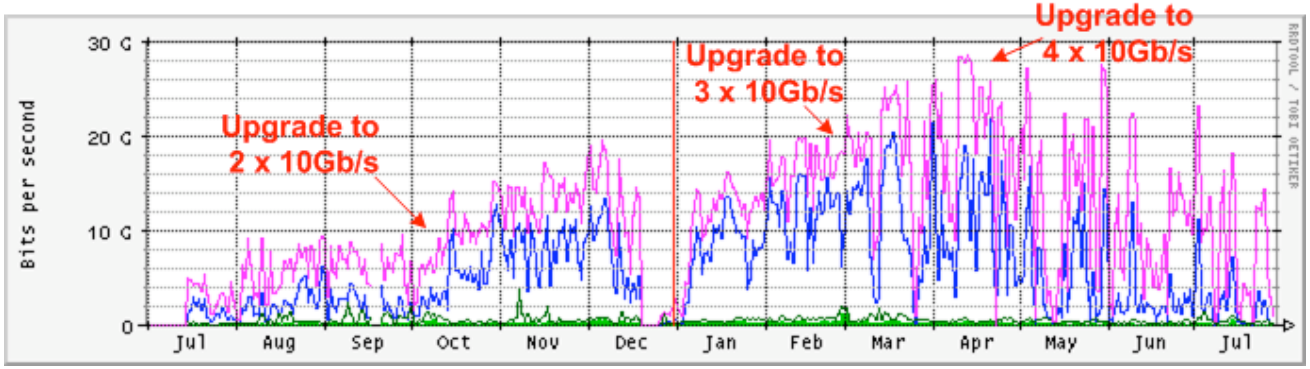
- Three computing buildings.
- FCC, -- > 20 year old purpose built
- LCC, GCC: built on former experimental halls w/ substantial power infrastructure.



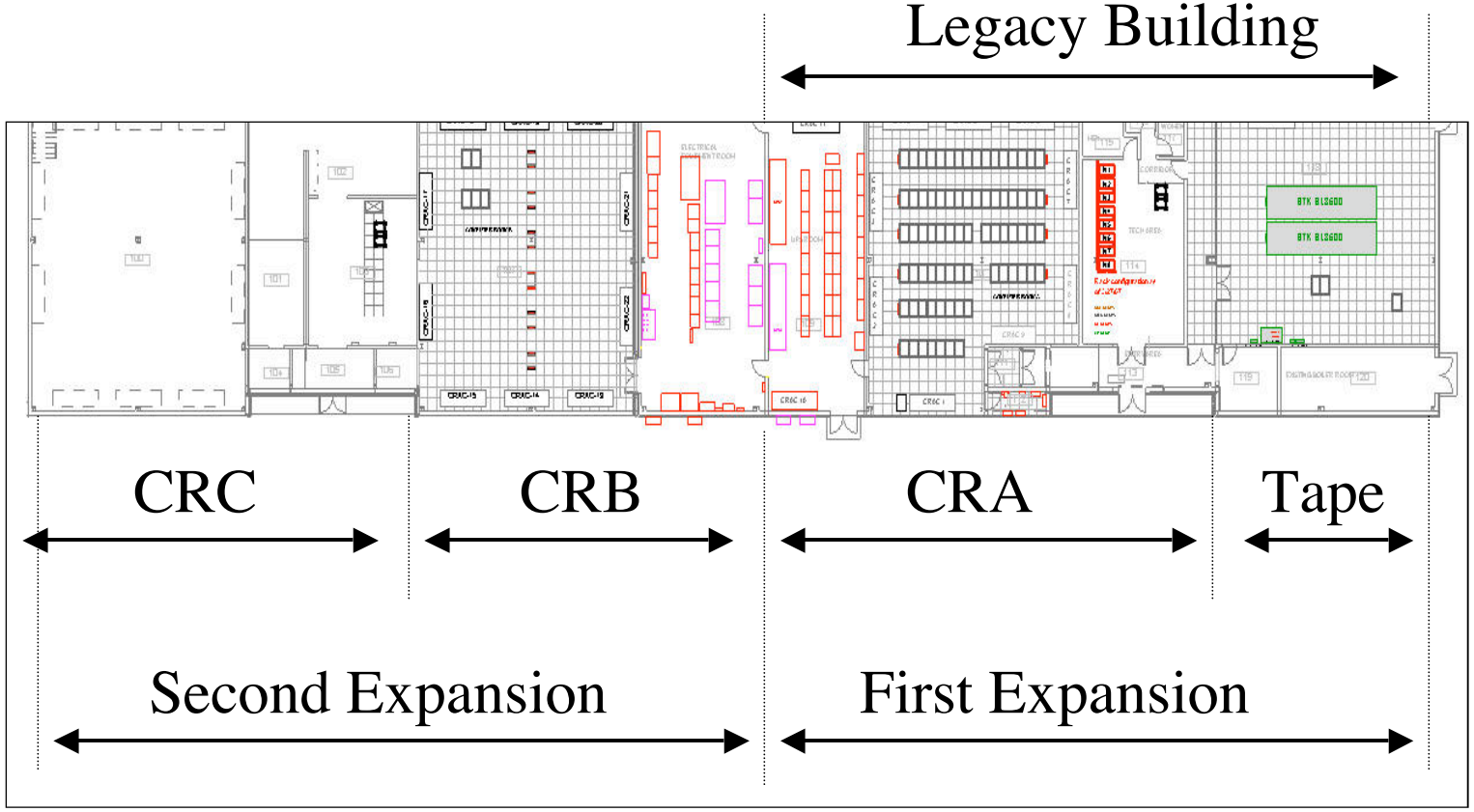
Network Underpinnings

- Ample campus networking has helped Fermilab realize a fallowing strategy
  - Appropriate computers in appropriate buildings.
    - For example, only worker nodes in GCC CRx.
  - Ample provisioning
    - 192 fiber pairs between building.
    - Ethernet is affordable as an interconnect.

Network Utilization Between CMS Tier-1 Computing Facility Core Switches



GCC Grid Computing Center

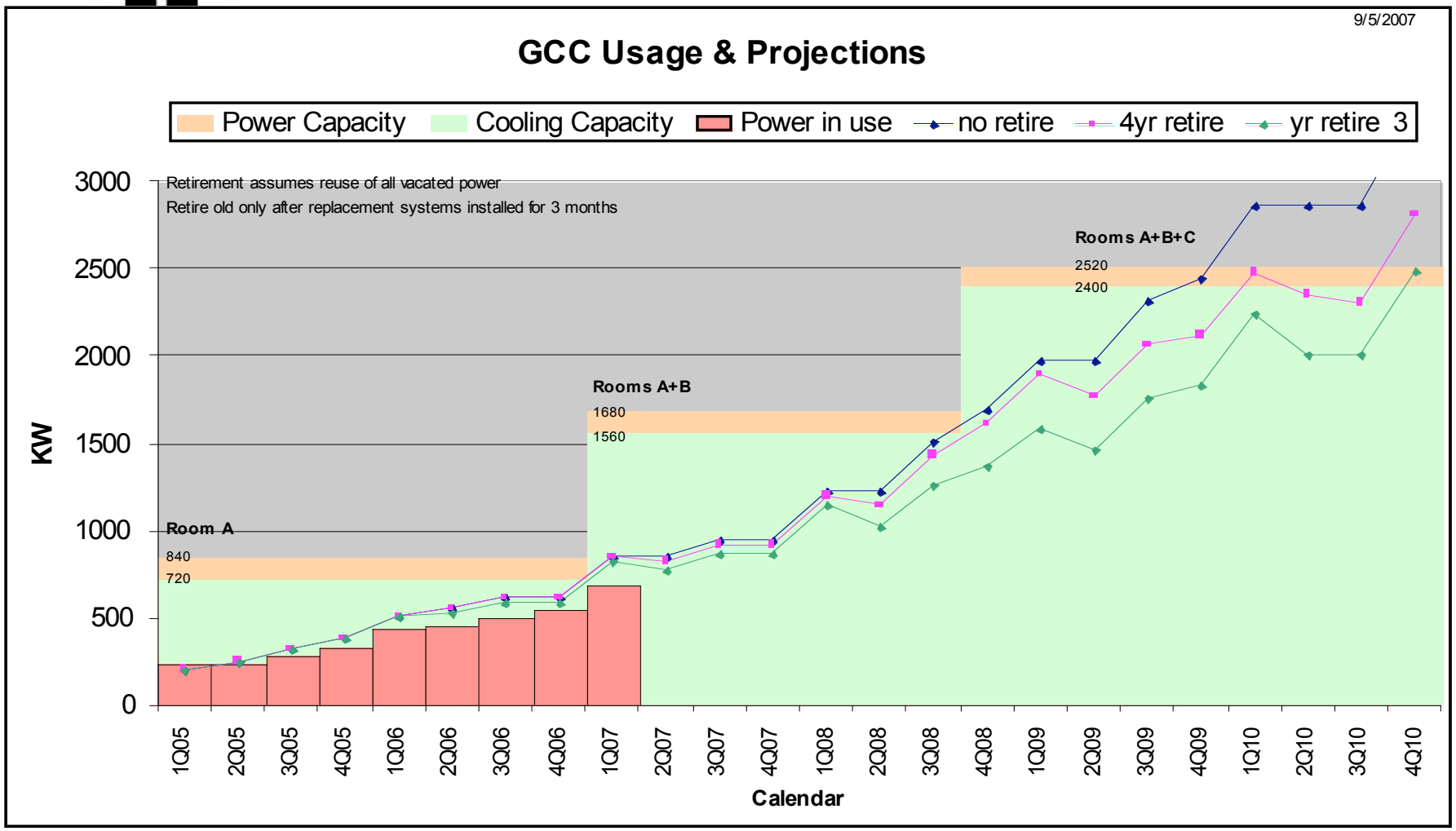


FNAL Power Footprint

	FCC	600 KVA
	LCC	710 KVA
GCC	CRA	840 KVA
GCC	CRB	840 KVA
GCC	CRC	840 KVA
GCC	Tape	45 KVA
	Total	3875 KVA

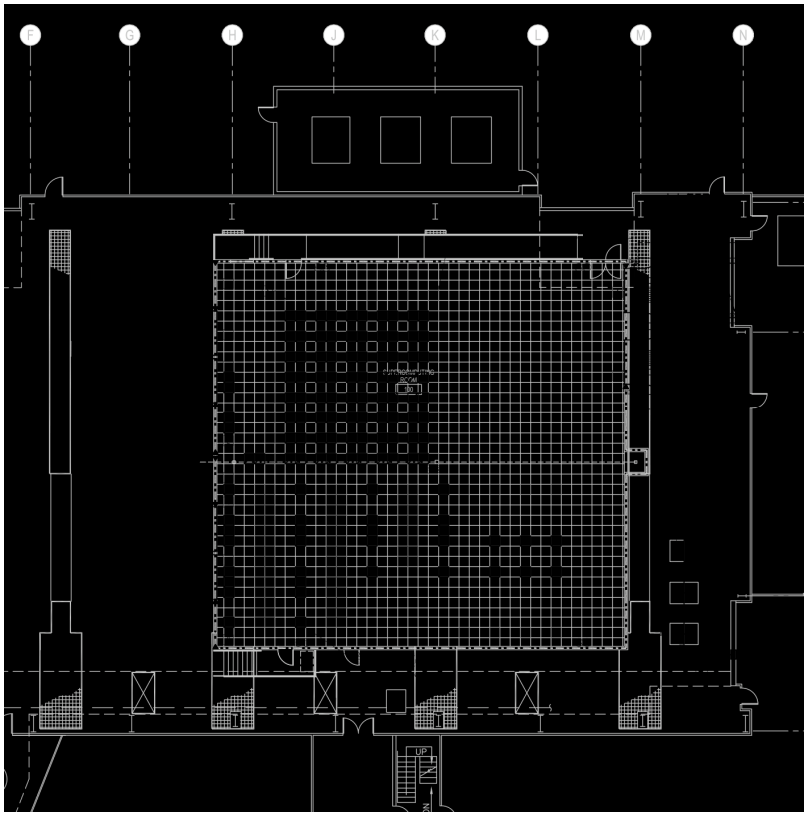
Does not include cooling

Just in Time Delivery



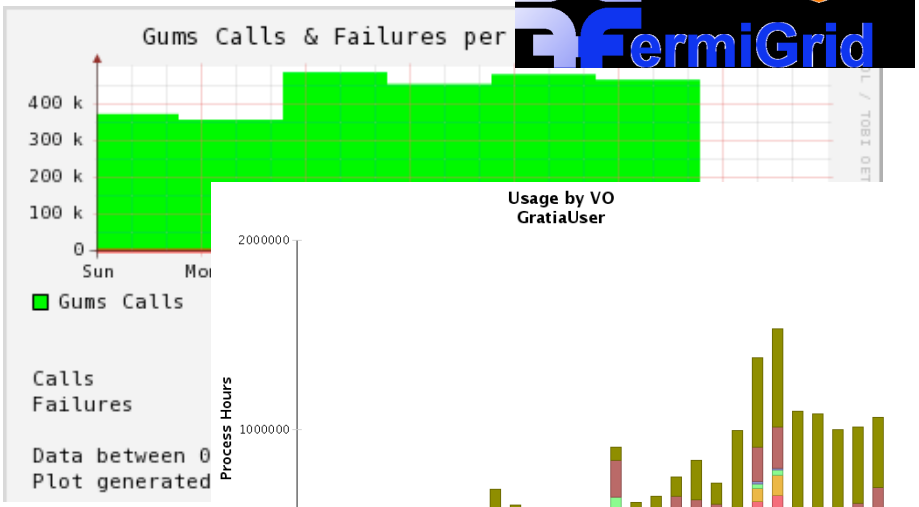
Leadership Computing at ANL

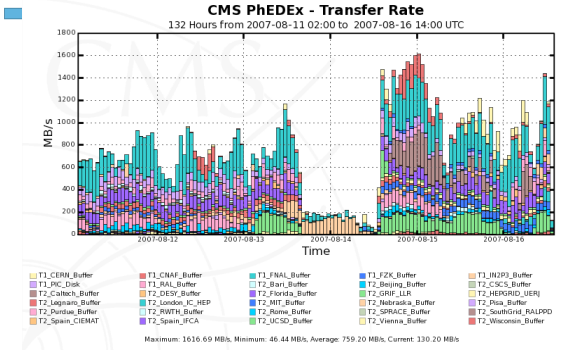
- In the US, DOE has funded several "Exascale" computing centers.
  - O(10**6) cores.
- One such new center is at Argonne.
  - Permanent building is planned for
  - Early start by re-using a building with substantial legacy cooling and power.



HEP capacity, adapted…

- HEP has adopted grid techniques as the basis for a world wide computing infrastructure.
- Allow for contingency in provisioning.





Summary

- Forecasting has shown a need to continually expand computing facilities for space, power and cooling
  - Fermilab has expanded its capacity incrementally
  - Other facilities are/will construct purpose built computing centers
- Solving the facilities challenges requires
  - Forecasting needs and technologies
    - Grid computing can provide a buffer
    - Understanding of possible power savings at each level of integration
  - Using local advantages and working around local disadvantages
  - Funding
- Providing for continuous level of service across many stakeholders for many years might require techniques as fallowing and other uses of swing space.