OSIsoft

Monitoring and Controlling a Scientific Computing Infrastructure

Jason Banfelder
Vanessa Borcherding
Dr. Luis Gracia
Weill Cornell Medical College

Cornell University

Overview

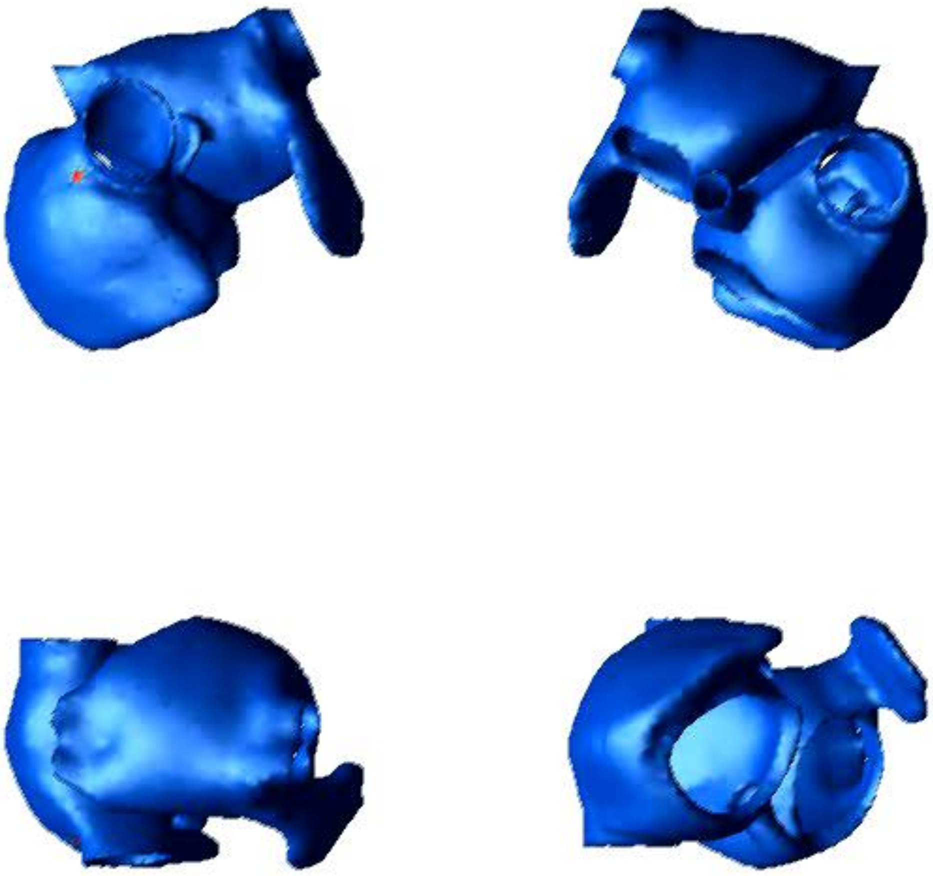
- Overview of operations
  - Scientific research
  - Computational infrastructure
- Building a monitoring framework with IT Monitor and PI
- Conserving power used by compute clusters
- Future challenges

Scientific Research

- Develop, integrate and maintain computational and other technological resources to support…
  - Dept. of Physiology and Biophysics
  - Institute for Computational Biomedicine
  - Computational Genomics Core Facility

…at the Weill Medical College of Cornell University

- Research, educational, and clinical mission
- Roughly 150 people to support



Christini Lab

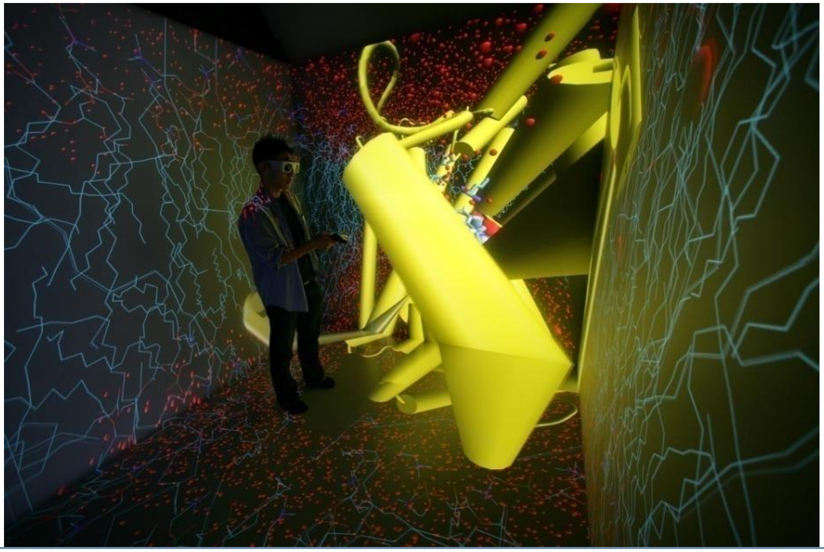
High Performance Computing

- Simulations of biological systems
  - Molecules, tissues, cells, organs

- Analysis of large datasets
  - Genomes, other -omes
  - High-throughput experiments
    - Three DNA sequencers
      - 10-20 TB/yr each
    - Multiple microarray (lab on chip) platforms
  - Corpus of biological literature

High Performance Computing

- Imaging & Viz (MRI, confocal microscopy, ...)
  - Clinical and basic science
  - Immersive visualization (CAVE)
- Other services
  - Desktop, print, videoconference



Desktop Environment

- Not possible or desirable to standardize
- Operating System Distribution
  - 60% Windows XP
  - 35% Mac OS X*
  - 5% Linux (RedHat, Debian)
- Need to support all of them

Compute Resources

- 750+ processors; 2+ Tflop/sec
  - 208 node/416 processor Linux cluster
  - 90 node/186 processor Linux cluster*
  - Approx. 40 other servers
    - 1 – 32 cores; 2 GB – 128 GB memory
- Fairly heterogeneous environment
  - Primarily Dell/Intel (~95%); some SUN
  - Primarily Linux (Red Hat EL 4/5)

Storage Resources

- Mainline storage and cluster storage
  - 75+ TB raw spinning disk
  - 10 RAID arrays
    - Apple FibreChannel (Brocade and QLogic switches)
    - Dell SCSI direct attach
    - Lately favoring iSCSI
- Server Backup is LTO3/4 tape based
  - Four libraries (robots)
  - Seven drives
- Backup is disk-to-disk-to tape based
  - Retrospect for Desktops/Amanda for Servers

Application Resources

- Scientific Software Library
  - 150+ programs/versions
  - Open Source and Commercial
- LAMP+ stack
  - Redundant Apache servers
  - Web app servers (Tomcat/Java)
  - Oracle 10g/11g Enterprise
    - Also MySQL; PostgreSQL

Physical Plant

- Three Server Rooms
  - Cornell University Ithaca Campus
    - 208 node cluster was too power/HVAC intensive to house on NYC campus
    - Fully equipped for remote management
    - Lights out facility, one visit last year
  - NYC Campus
    - 125 kVA server room (10 cabinets) [12.5 kW/cabinet]
    - 300 kVA server room (12 cabinets) [25 kW/cabinet!!!]
- At full load, we can draw over 1 MW to run and cool!

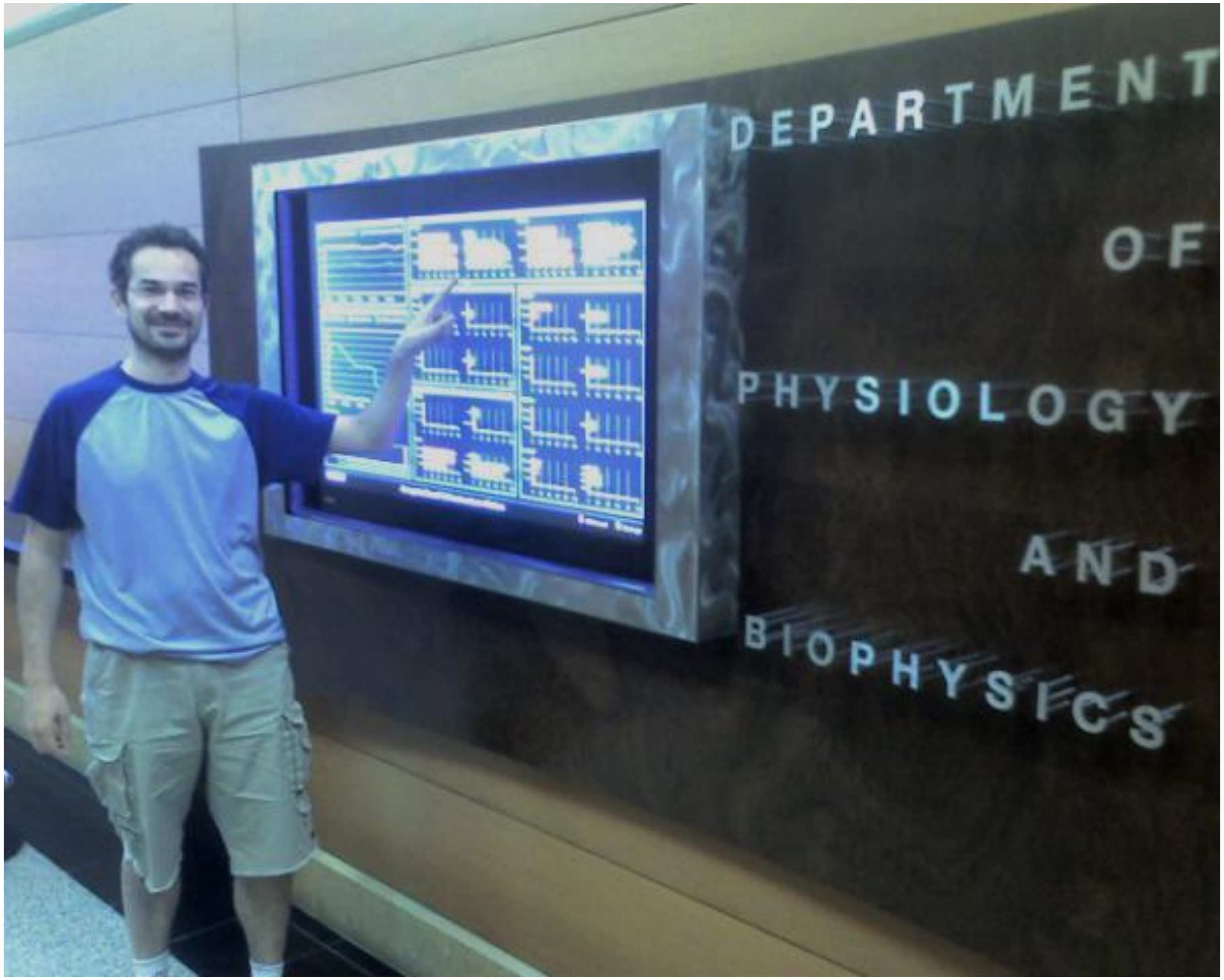
Managing the Infrastructure

- All of the above built and maintained by a group of four people.
  - Automation required
  - Don’t want to standardize too much, so we need to be very flexible

Why IT Monitor and PI?

- PI selected to be the central repository for health and performance monitoring (and control).
  - Able to talk to a diverse set of devices
    - Printers, servers, desktops
    - Cisco switches
    - Building management systems
    - …pretty much anything we care about
  - Pick and choose the parts you want to use, you build the solution
    - Ping, SNMP, HTMP interfaces, ODBC
  - Very strong, proven analytics
- Vendor specific solutions are (expensive) islands

Project 1: The Big Board

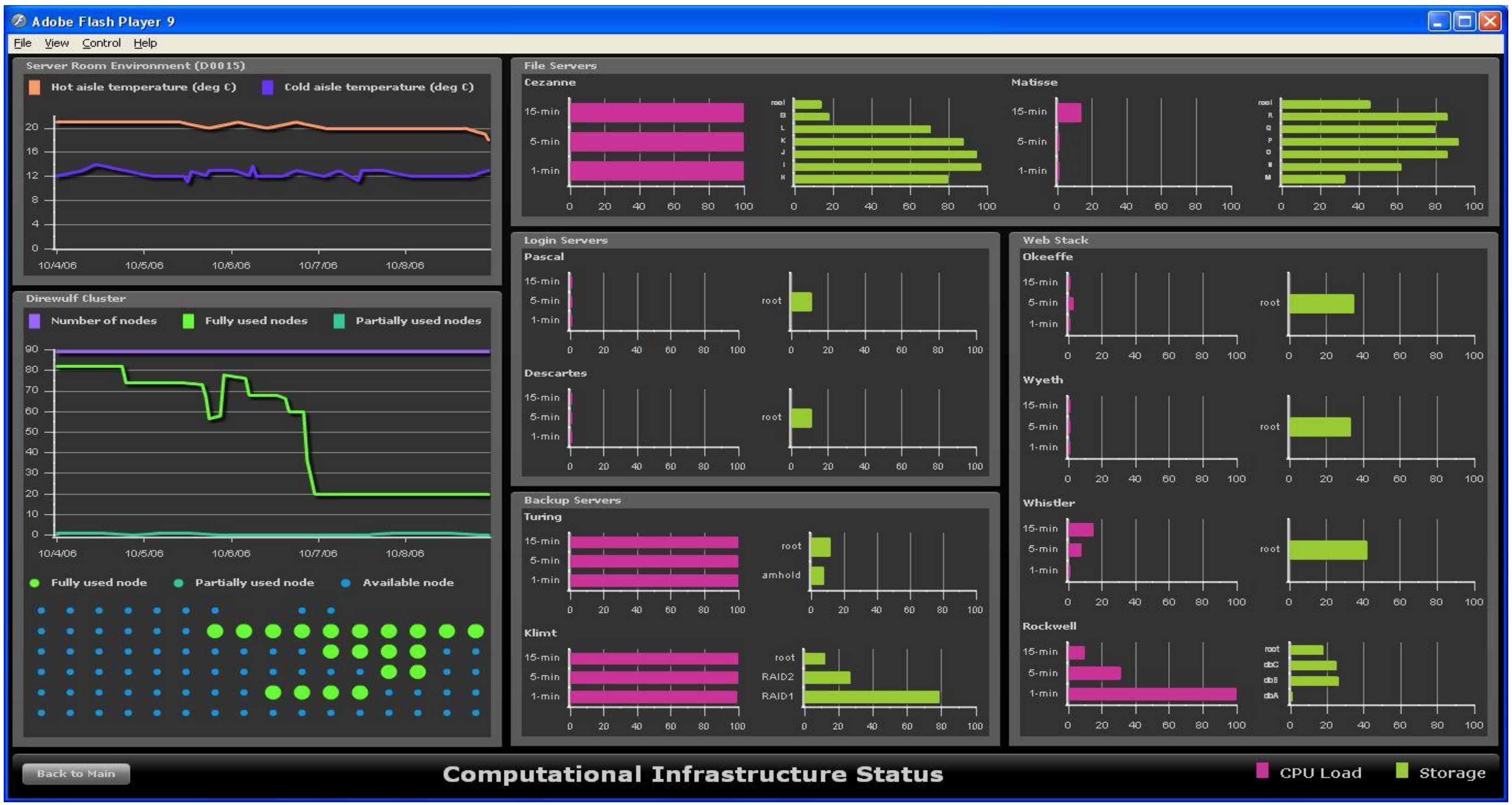


Overall Systems Health

- Want a quick, holistic view of our key resources
  - Core infrastructure
    - File servers, web servers, app servers
    - Backup servers
  - Cluster utilization
    - Node statuses and summary
  - Physical plant
    - Temperature monitoring

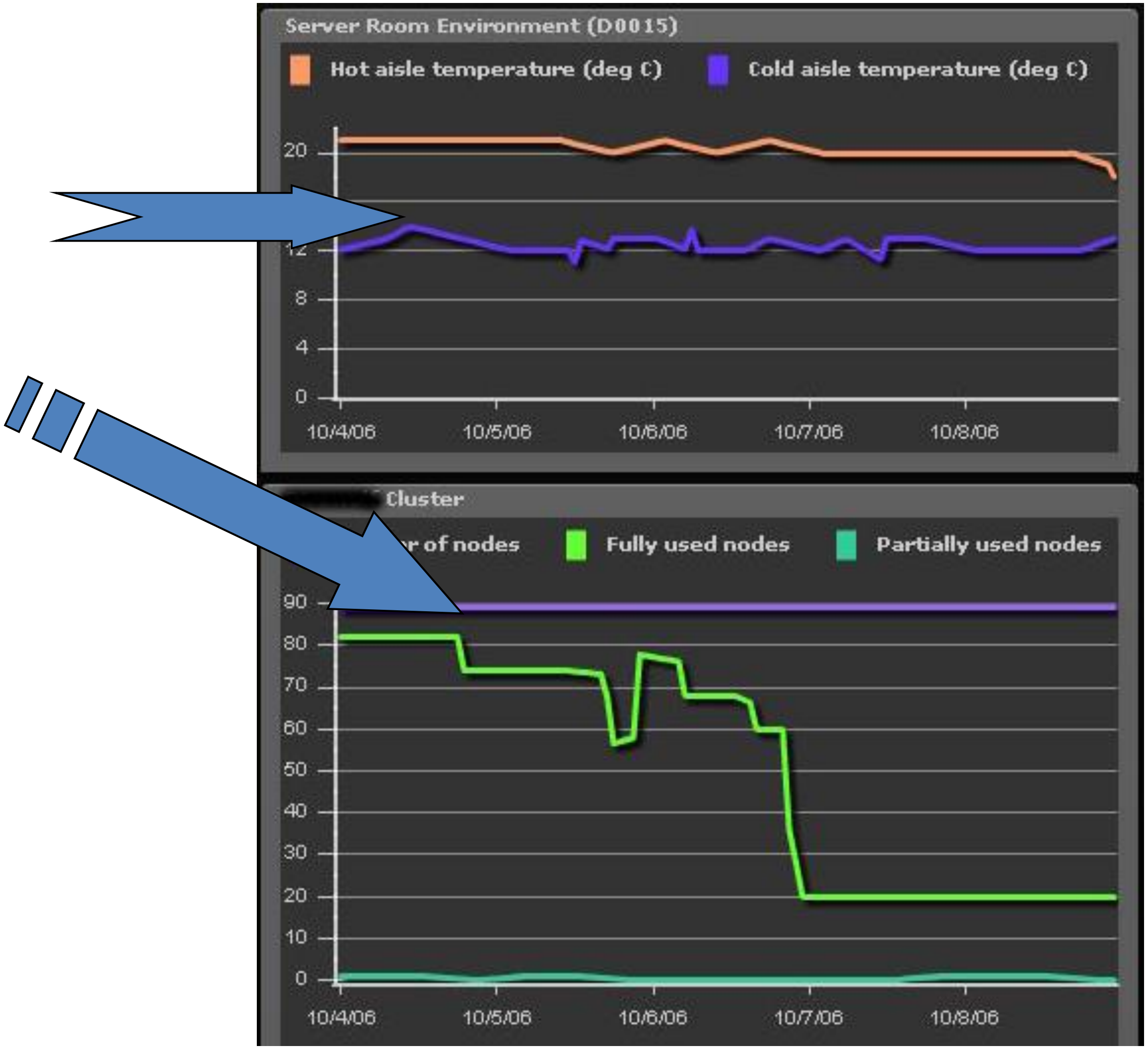
Data is Available to Everyone

- Adobe Flash/Flex used for display



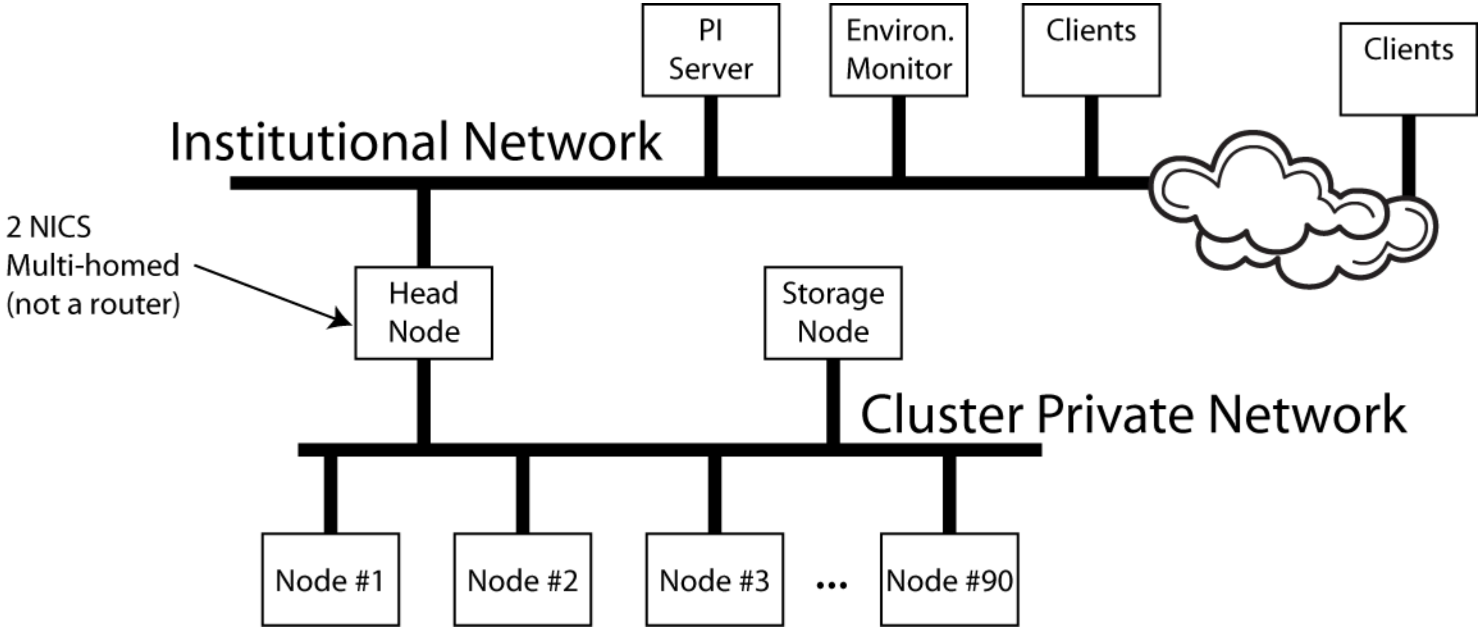
Why PI? (revisited)

Is this affected by that?



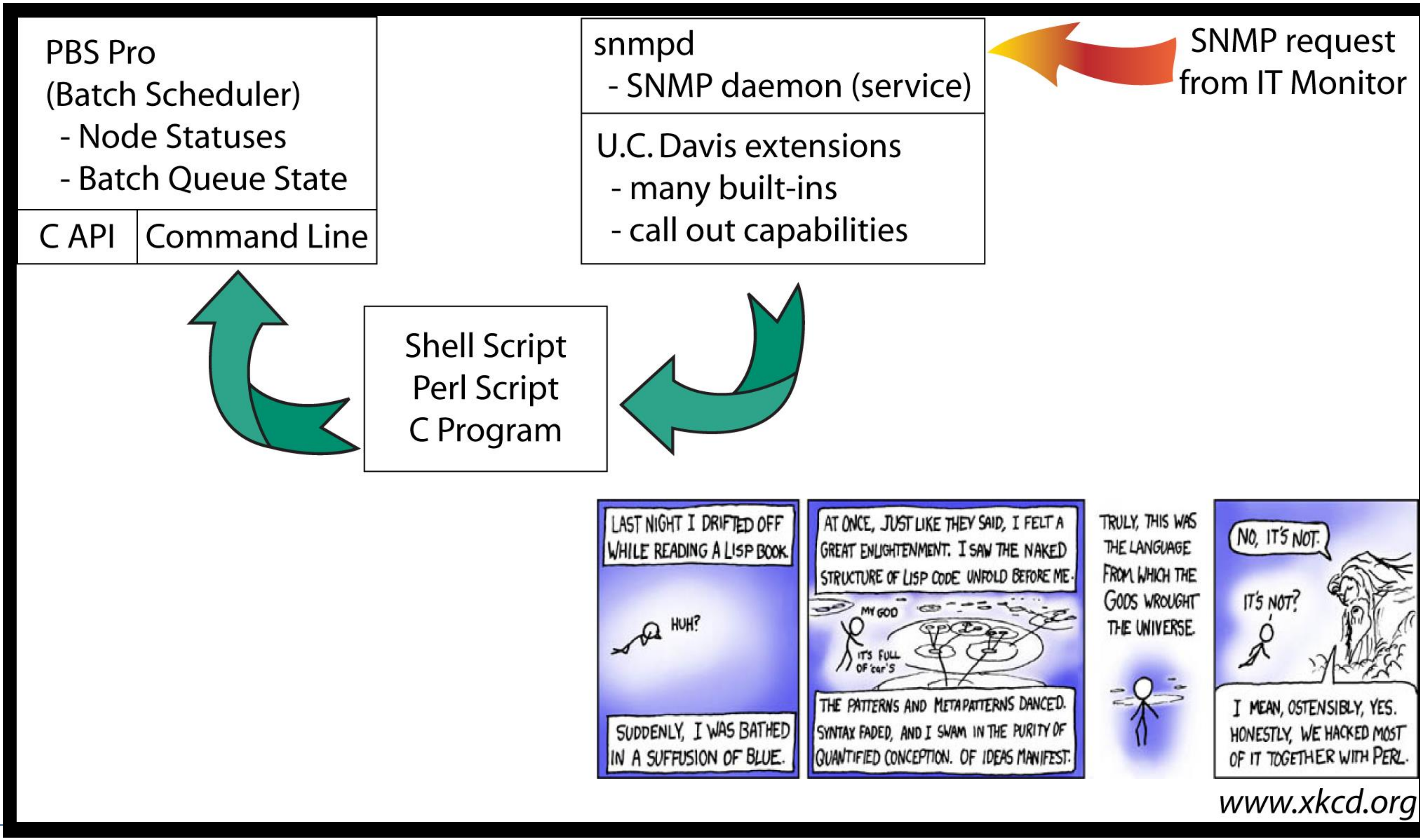
- This can only be answered if all your data is in the same place.

Network Layout

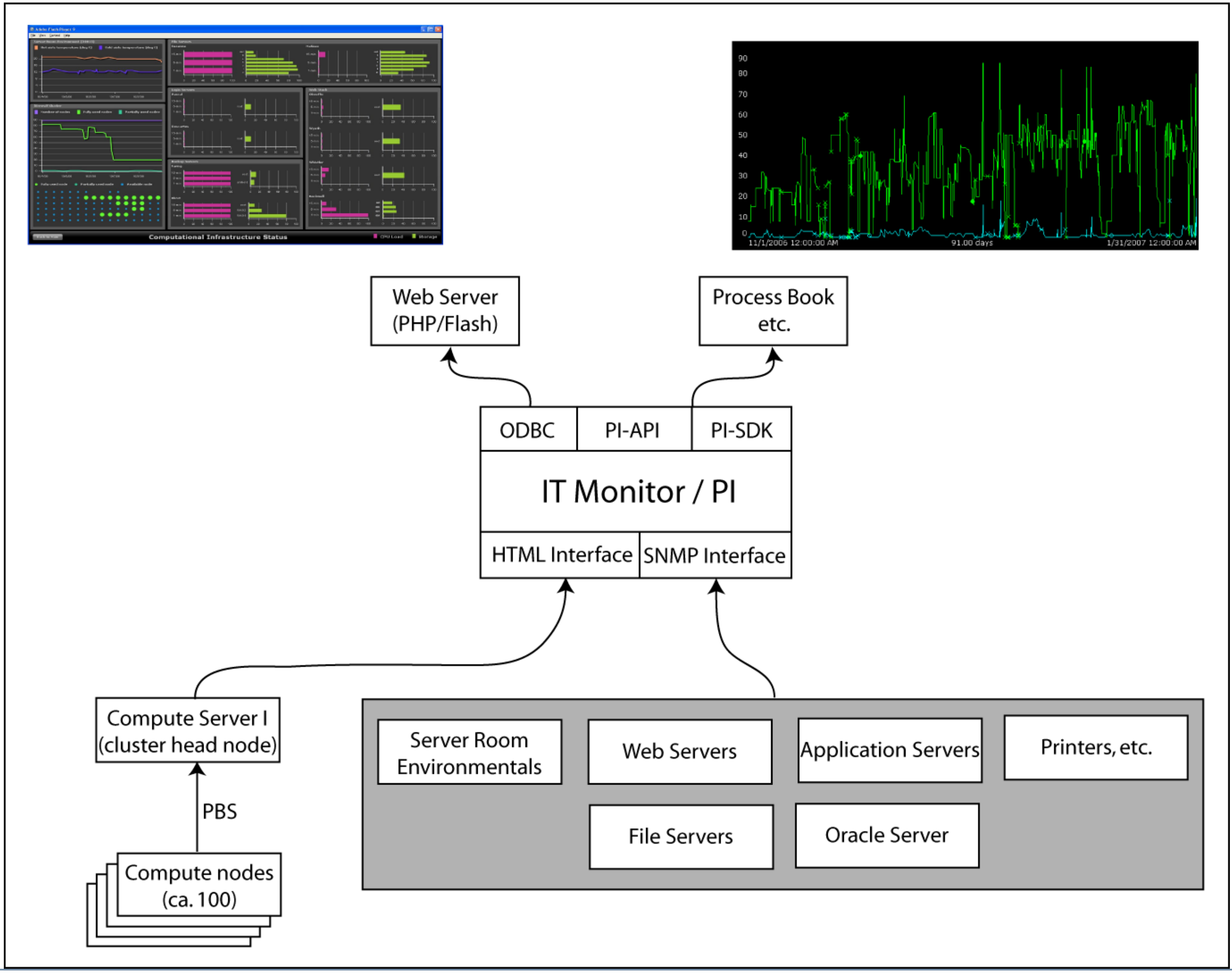


- PI Server can only see head node
  - OK; head node knows what's going on anyway
- How does PI Server read the data we are interested in?
  - Node statuses and summary statistics

PI Speaks SNMP Fluently



Data Collection and Dissemination Architecture



U.C. Davis SNMP (aka Net-SNMP)

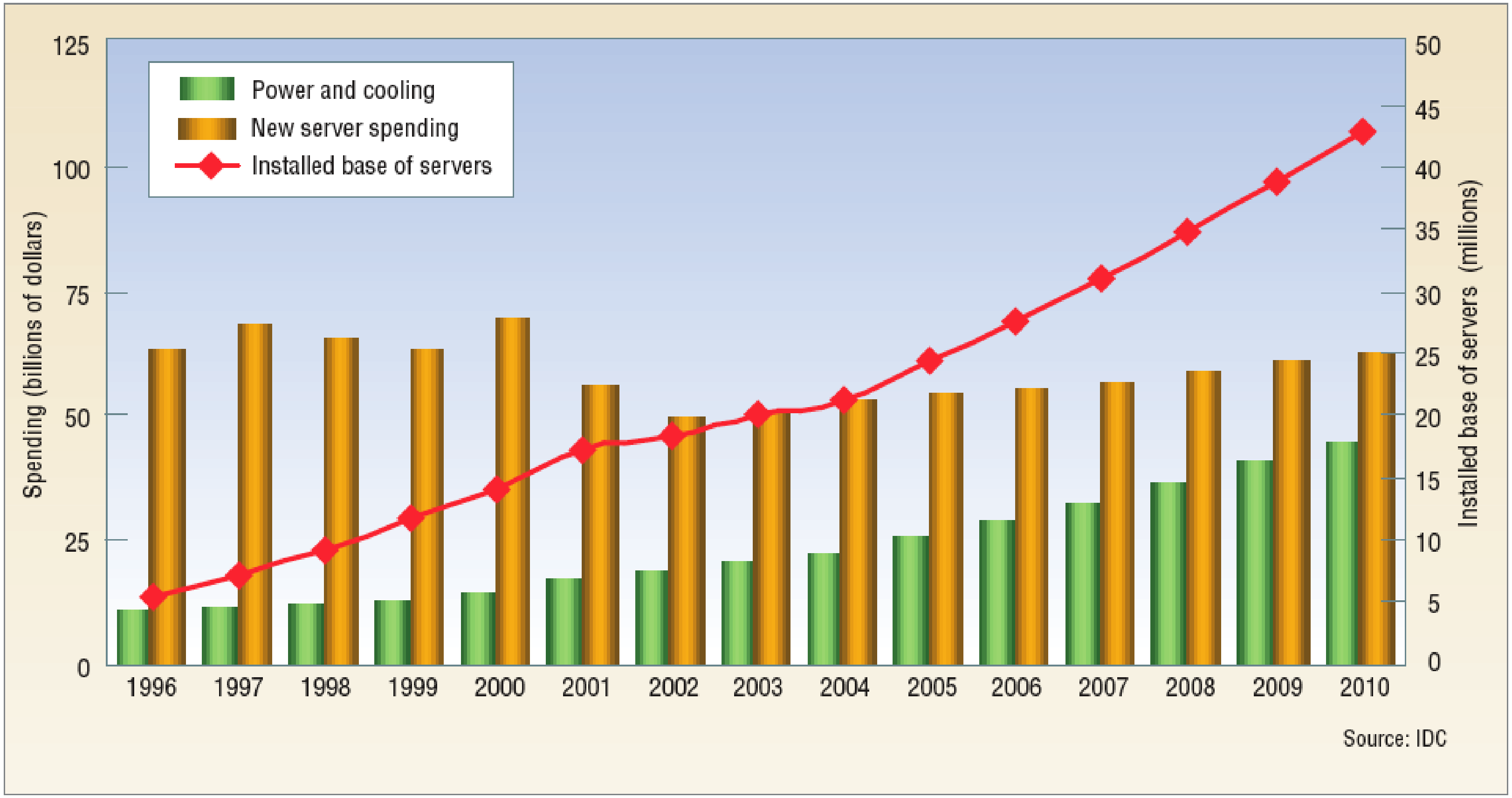
- Built-ins
  - System information
  - System load
  - NICs/network activity
  - Running processes
  - Disk usage
  - Log files

- Extensibility Options
  - Return one or many lines of output
  - Return a single value or a whole subtree of a MIB
  - One-shot or stay-resident invocation
  - Embedded Perl support
  - Proxy support
- Getting the data is easier then writing the MIB!

Project 2: Cluster Power Management

- Green computing…
  - Save power (and $$$) by shutting down nodes that are not in use.
- …but minimize impact on performance
  - Maintain a target number of stand-by nodes ready to run new jobs immediately.
  - Reduce latency perceived by end-users

The Cost of Power and Cooling



Lawton, *IEEE Computer*, Feb 2007

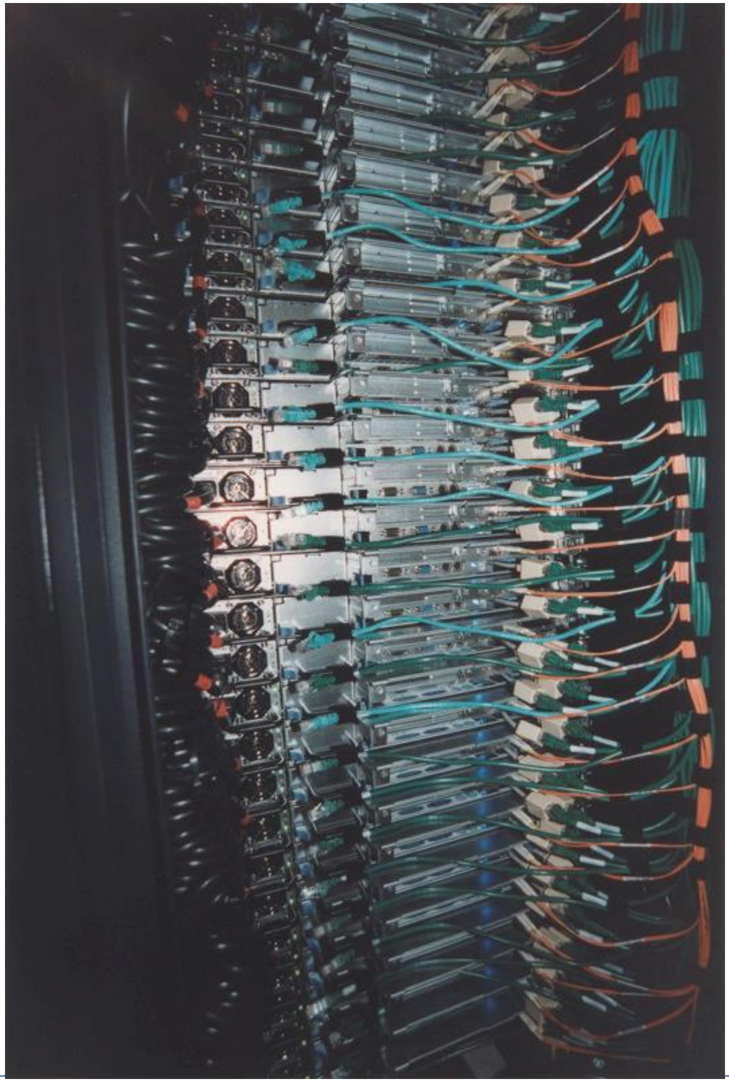
Powering HPC

- Density is increasing
  - 20+ kW per cabinet (standard designs were 2-4 kW only a few years ago)
  - Localized heat removal is a problem
  - HVAC failures leave very little time for response
- Load is highly variable
  - Harder to control

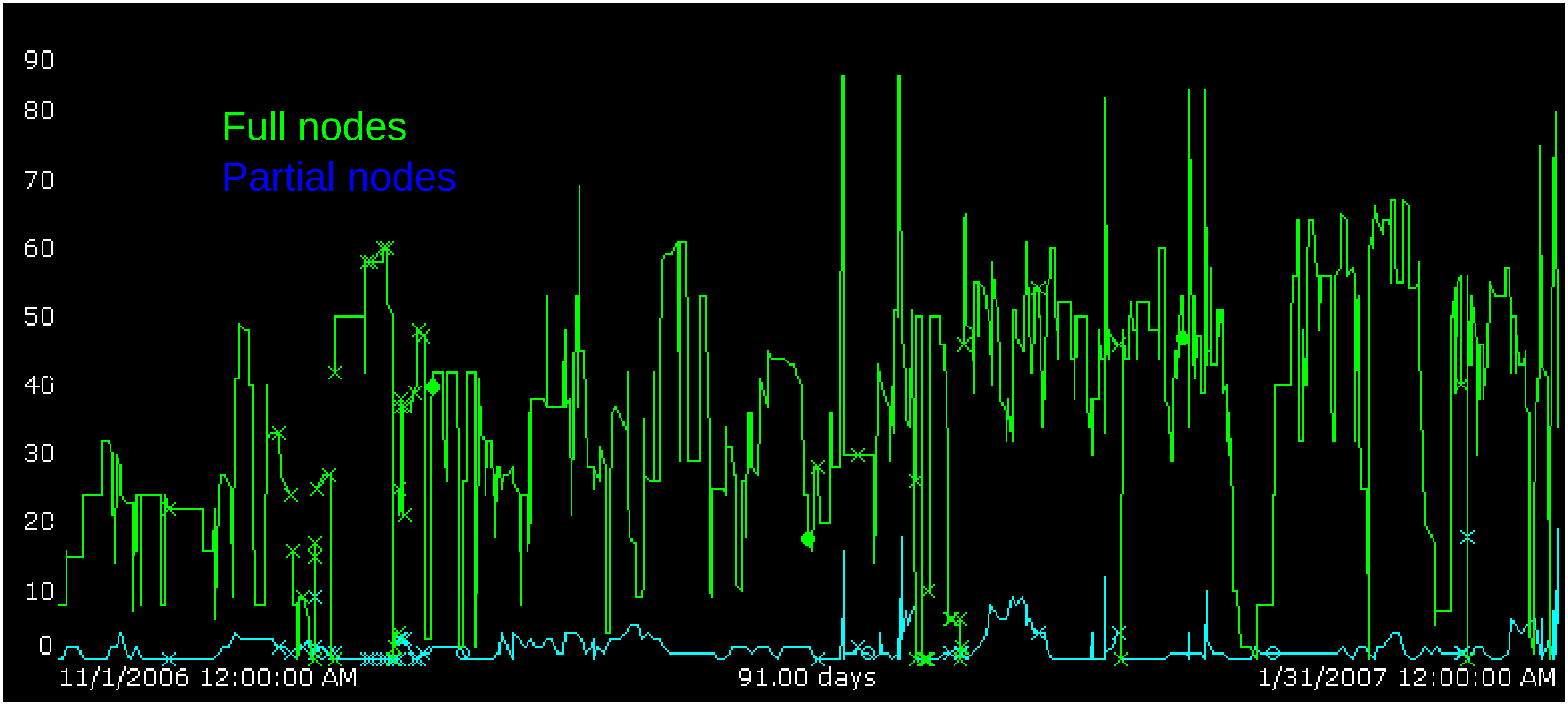
Our Cluster

- 90 compute nodes
  - 3 ‘fat’ nodes; 2 ‘debug’ node
  - 85 nodes under CPM
- Power used:
  - In Use node: 250 W
  - Stand-by node: 125 W
  - Power Save nodes: 0 W
- With 50% usage and no standby nodes:
  power savings is 32%
- With 66% usage and 16 standby nodes:
  power savings is 11%

Dense Computing



Historical Cluster Usage



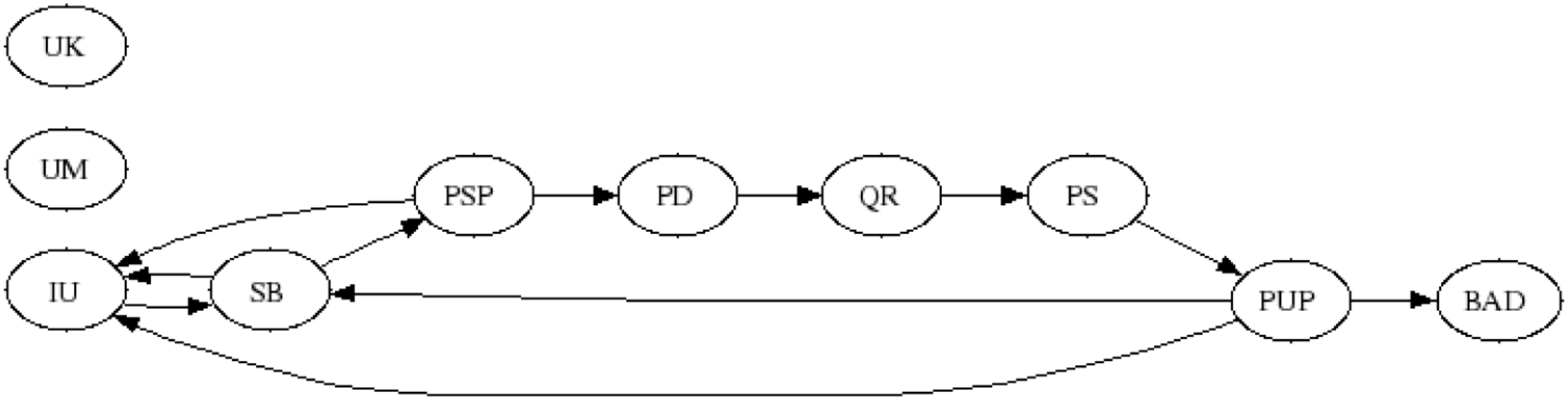
Hardware Requirements

- Hardware Requirements
  - Chassis Power Status
  - Remote Power Up
  - PXE is a plus for any large system
- Dell Servers do all of this standard (and much more!)
  - Baseboard Management Controller
  - Dell Remote Access Card

IPMI + SNMP
= Data + Control

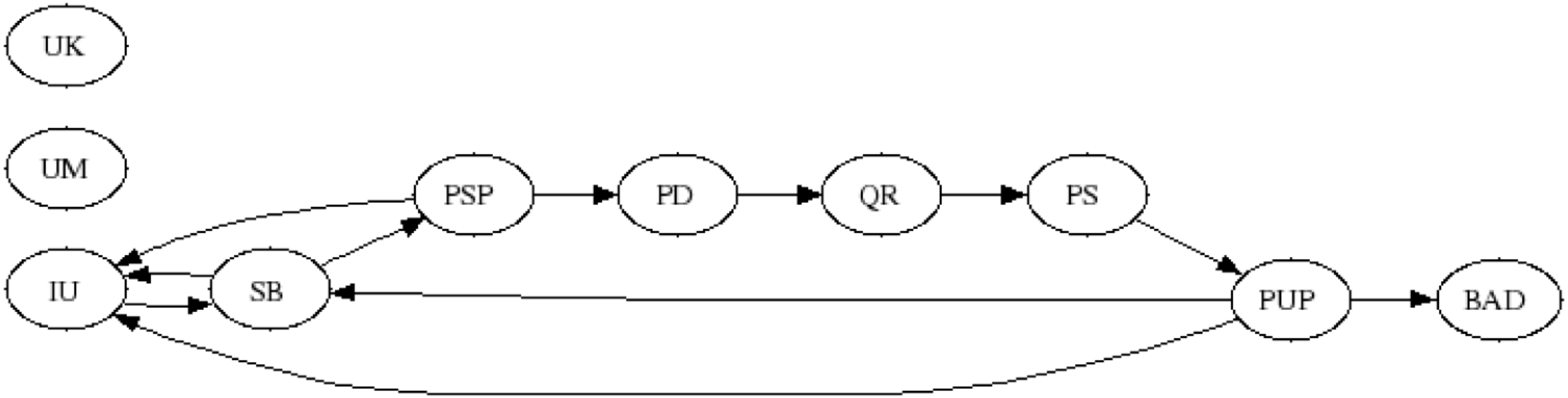
```
[root@cluster clusterfi]# ipmitool -I lan -H 10.1.12.190 -U root -f
   passfile sensor list

Temp             | 21.000   | degrees C | ok | 120.0 | 125.0 | na
Temp             | 20.000   | degrees C | ok | 120.0 | 125.0 | na
Temp             | 23.000   | degrees C | ok | 120.0 | 125.0 | na
Temp             | 23.000   | degrees C | ok | 120.0 | 125.0 | na
Temp             | 40.000   | degrees C | ok | na    | na    | na
Temp             | 61.000   | degrees C | ok | na    | na    | na
Ambient Temp     | 16.000   | degrees C | ok | 5.000 | 10.0  | 49.0 | 54.0
Planar Temp      | 20.000   | degrees C | ok | 5.000 | 10.0  | 69.0 | 74.0
CMOS Battery     | 3.019    | Volts     | ok | 2.245 | na    | na   | na
```

Lifecycle of a Compute Node



- CPM uses a finite state machine model
- Tunable parameters
  - Target number of standby nodes
  - Global time delay for shutdowns
    - Prevent churn of nodes

Lifecycle of a Compute Node

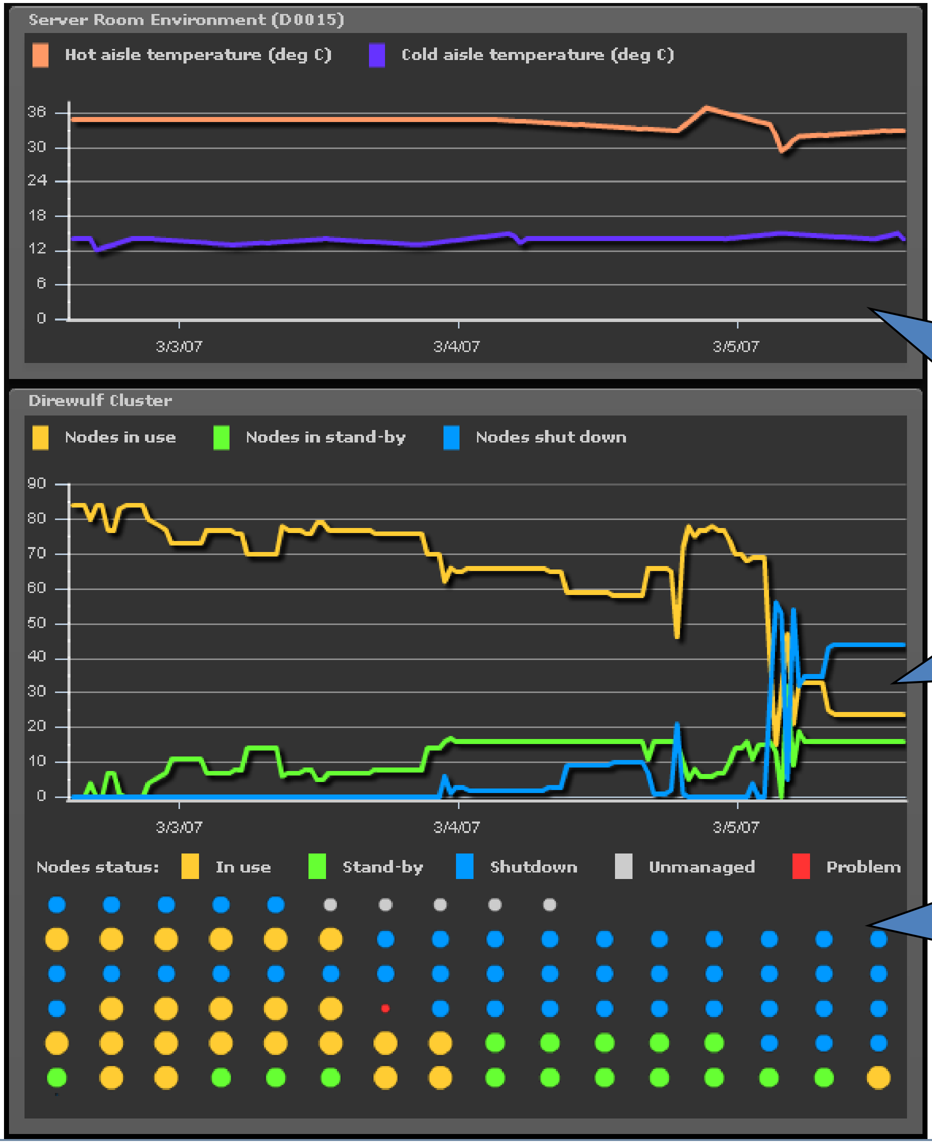


- IU: In Use
- SB: Standing by
- PSP, PD, QR:
  Shutting down
- PS: Power Save
- PUP: Powering Up
- BAD: Problems
- UK: Unknown
- UM: Unmanaged

	1	2	3	4	5	6	7	8	9	10
Nodes 081-090	PS	PS	PS	PS	PS	UM	UM	UM	UM	UM
Nodes 071-080	QR	QR	QR	PS	PS	PS	PS	PS	PS	PS
Nodes 061-070	PS	PS	PS	PS	IU	IU	IU	IU	IU	IU
Nodes 051-060	PS	PS	PS	PS	PS	PS	PS	PS	PS	PS
Nodes 041-050	SB	SB	SB	SB	PS	PS	PS	PS	PS	PS
Nodes 031-040	SB	SB	IU	IU	IU	IU	IU	IU	Bad	SB
Nodes 021-030	IU	IU	IU	IU	SB	SB	SB	SB	SB	SB
Nodes 011-020	IU	IU	IU	IU	IU	IU	IU	IU	IU	IU
Nodes 001-010	SB	IU	IU	SB	SB	IU	IU	IU	IU	IU

BAD	1
IU	33
PD	0
PS	32
PSP	0
PUP	0
QR	3
SB	16
UK	0
UM	5

Cluster Power Management In Action



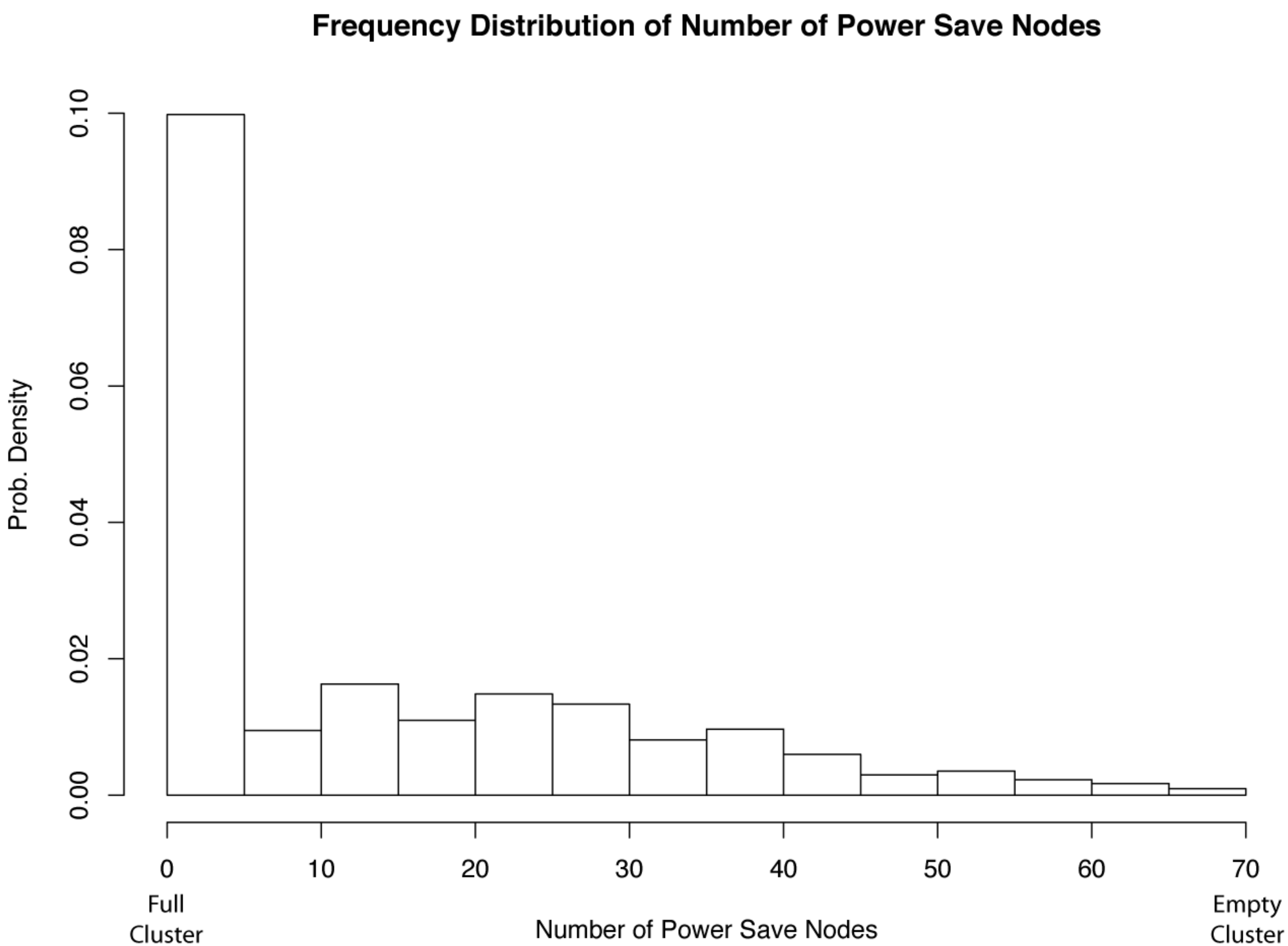
Note correlation between temperature and cluster usage

Powered down

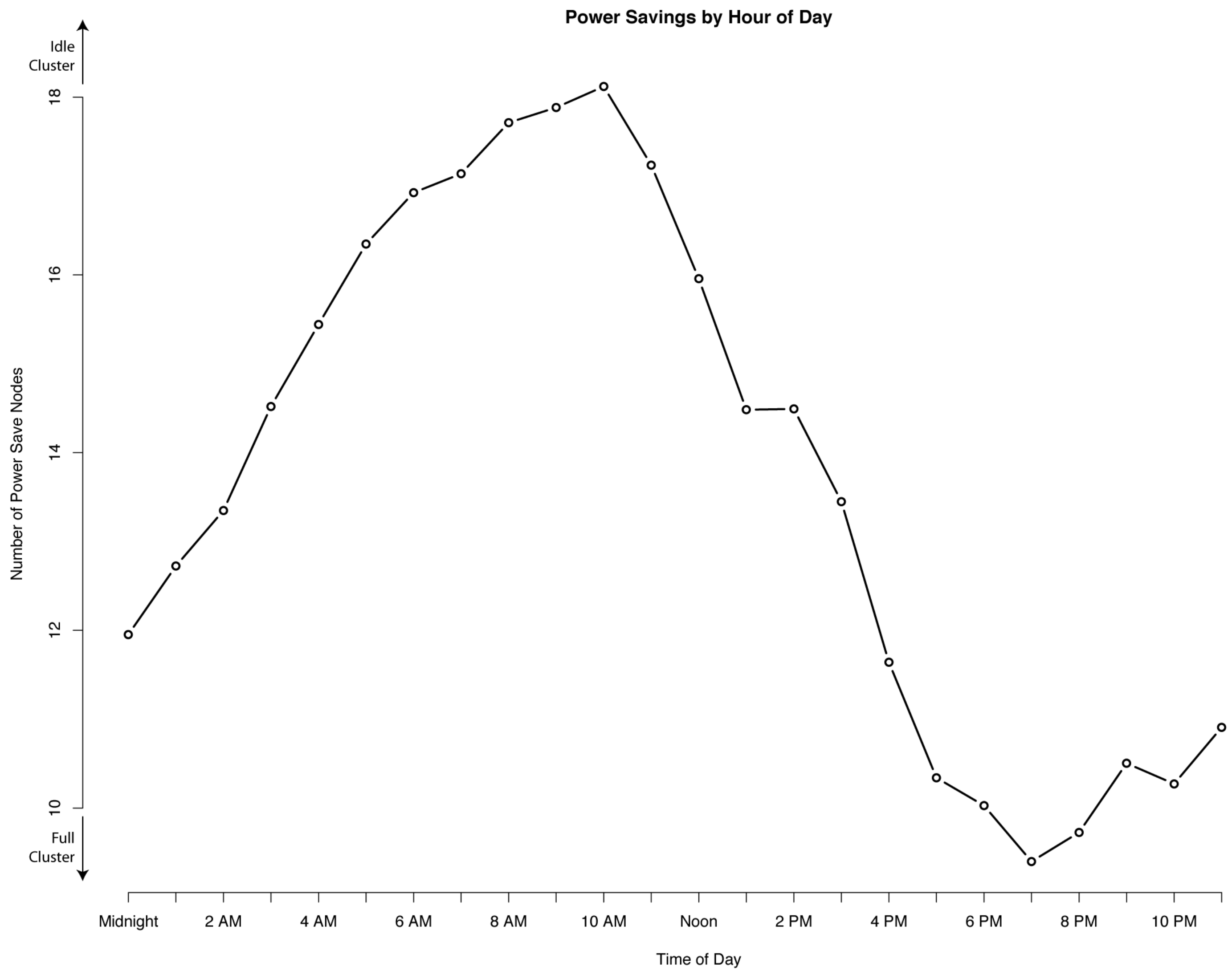
Results: Six Months of CPM

- Six Month Average:
  - 13.6 nodes shut down for power savings.
  - 16% of managed nodes
  - 8% power savings
- 15.06 MW*h annual savings
- $3,000 annual power savings ($0.20/kW*h)
- Probably double when considering HVAC
- Equipment life

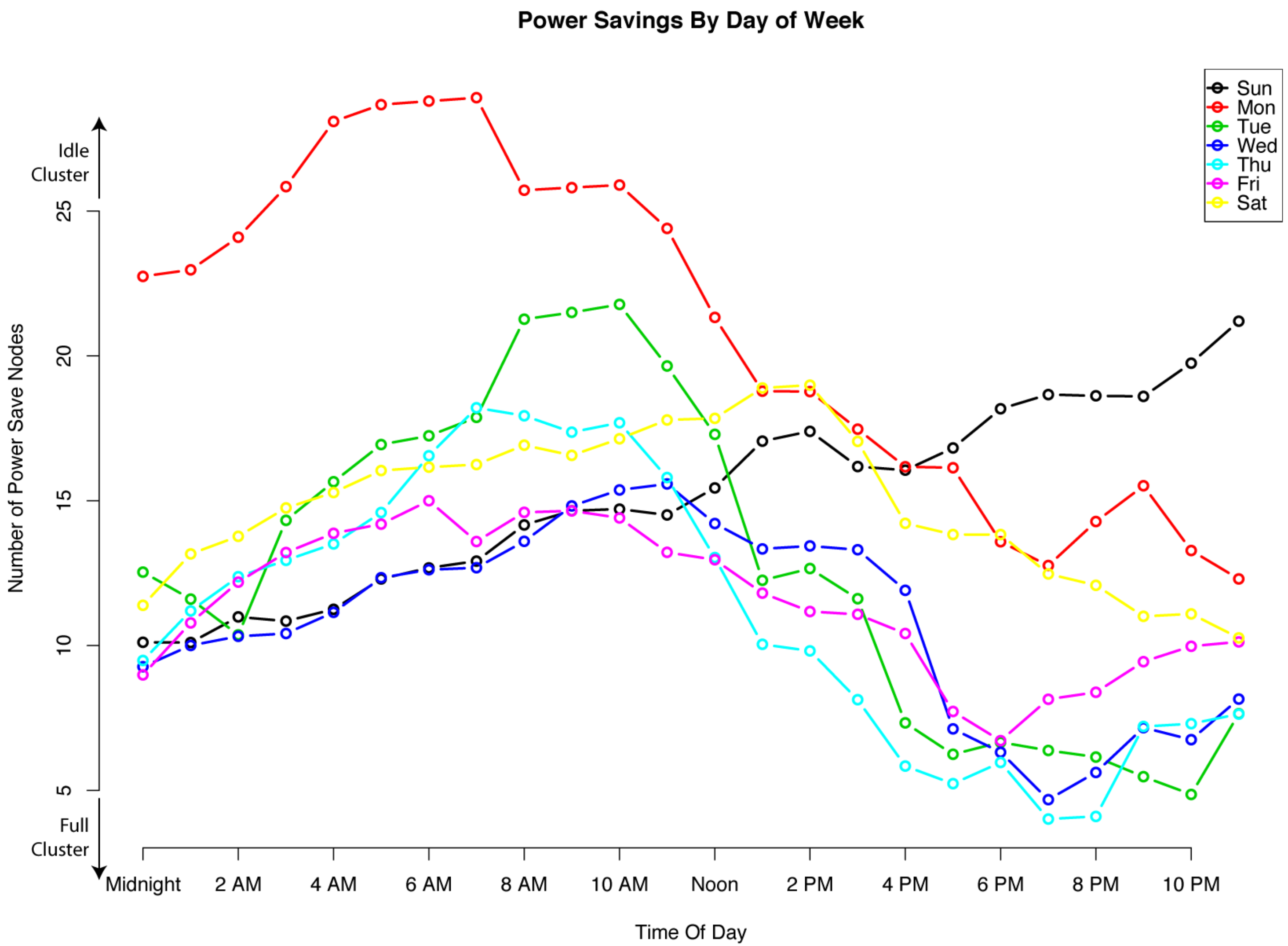
Results: Six months of CPM



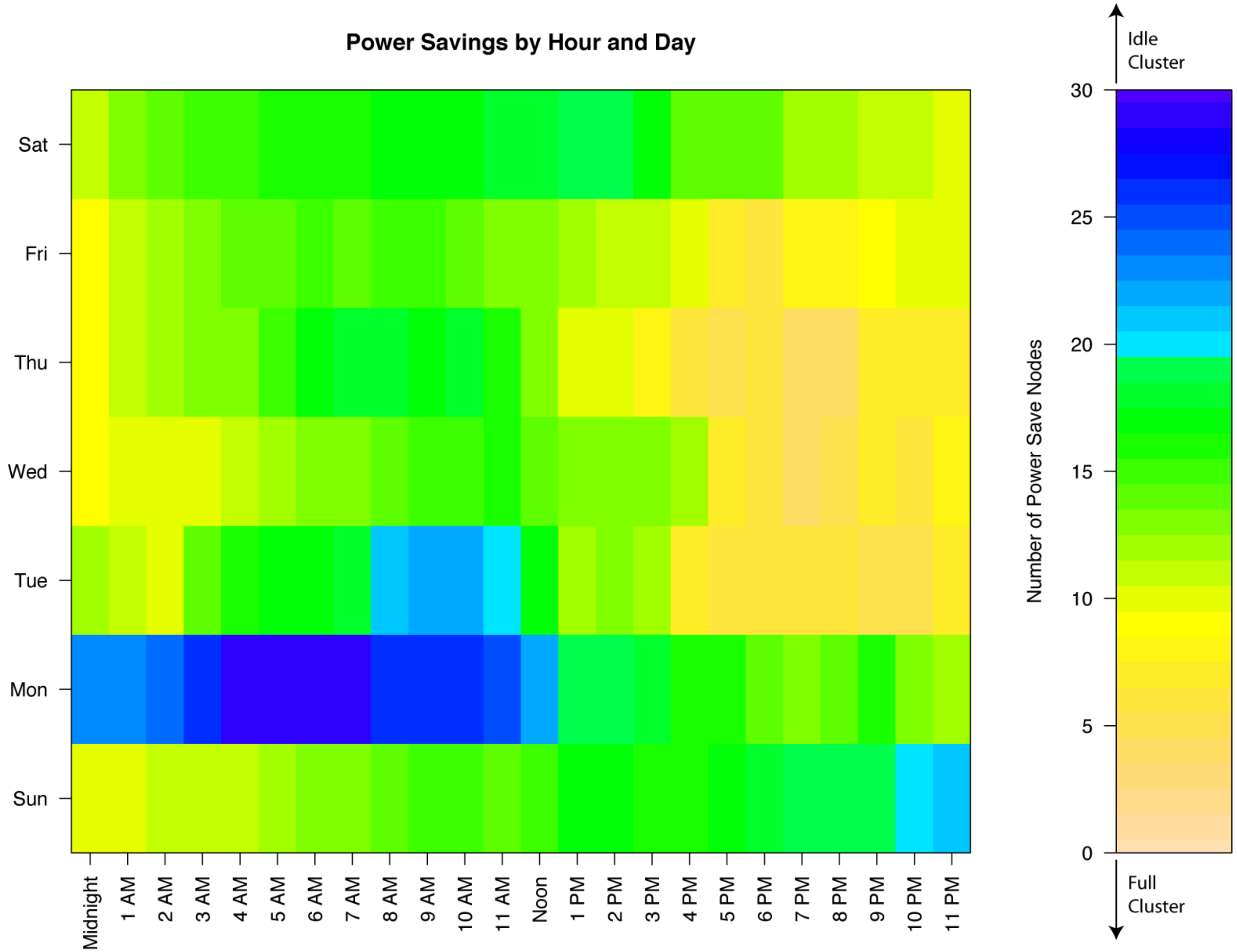
Results: Six Months of CPM



Results: Six Months of CPM



Results: Six Months of CPM



3 Years of PI

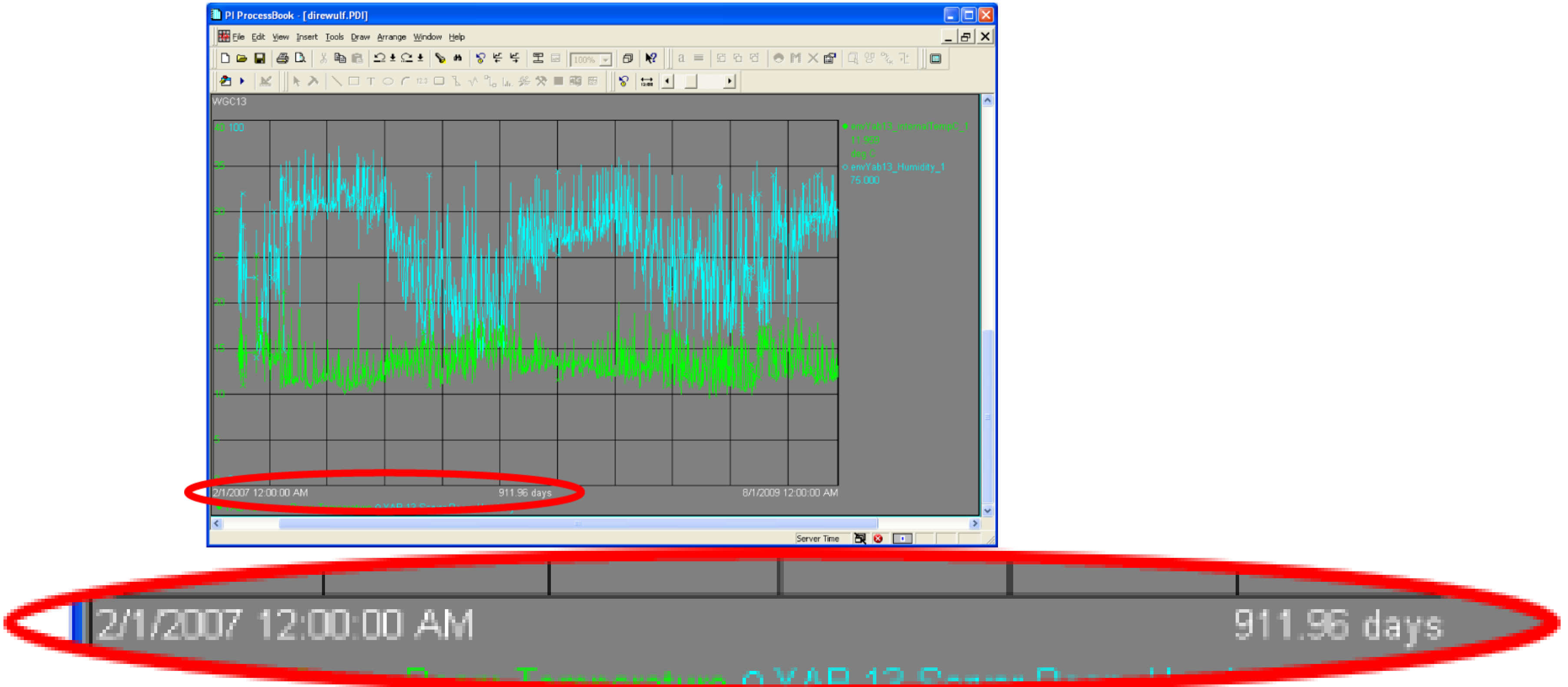
March 2007

Sept 2009

Full, complete detailed operating history of this facility is retained indefinitely…

3 Years of PI

- Data since date of commissioning is available
  - Day-to-day and seasonal variation
- The value of this cannot be overstated



3 Years of PI

- Data to Support of Federal Grant Applications

Like the compute infrastructure, the current core storage infrastructure for the ICB (designed, commissioned, and maintained by the TMT) has proven very reliable. At the time of writing, the four primary file servers that comprise this infrastructure have been running continuously for 772 days, 429 day, 807 days, and 529 days (the latter figure representing 100% uptime since it was originally commissioned). Although there have been occasional component failures, system redundancy as well as thoughtfully designed backup and restore procedures have proven reliable, and there has not been a single instance of data loss since the inception of any of these servers.

D.2.a.i. Recordkeeping
To maximize the effectiveness of the diverse resources under its administration, the TMT has instituted and continues to expand a program for real-time monitoring and recording of the operating state of technology infrastructure components. Representative plots from this system (**Figure 7**) show the capacity utilization of several file systems across two of the fileservers currently maintained by the TMT (and demonstrate the need for additional storage capacity within the ICB). The system incorporates information on processor utilization, environmental conditions, etc., and is invaluable for troubleshooting, capacity planning, and for preparing quantitative quarterly reports for the Management and Advisory Committee (see **Section E.1.c**). This resource will be particular useful to the TMT in predicting when users will be likely to exceed their storage allocations, so that they can be notified and plan for additional capacity or take remedial action well before limits are reached and progress of research is affected.

Excerpt from an application for a $3.5MM grant for a ~1 petabyte research data storage system.

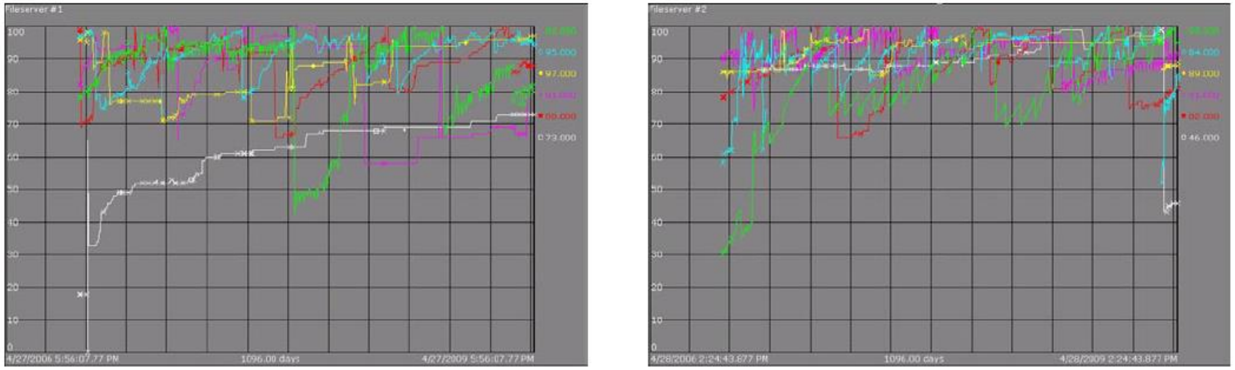


Figure 7 Three year history of file system utilization on two representative fileservers maintained by the TMT. Curves show percent capacity for several 375 GB file systems. Data are recorded every 5 minutes, and are permanently retained.

Challenges/Direction

- Tighter Integration with IT Hardware
  - We are beta-testing the IPMI interface
    - More scalable than our Perl hacks
  - Automatic Point Creation
    - Need to integrate with our asset management database
    - Some preliminary success with Oracle triggers and PI-OLEDB…
      - …but we are looking forward to the JDBC interface
    - Additional support forthcoming from server vendors

Challenges/Direction

- Integration with building systems
  - Very fast temperature rises (20-30 min)
  - Detecting nascent problems is critical
    - Differentiating from transient disturbances (e.g., switch from free cooling to absorption chillers)
  - We are currently implementing the BACnet interface to collection 150 points from our BMS.

Q & A

Jason Banfelder

jrb2004@med.cornell.edu

OSIsoft®

Thank you