Power & Cooling Impact on Data Center
Reliability and Availability

The Data Center Perspective
J. Ray Scott
Director, Systems and Operations
scott@psc.edu

© 2008 Pittsburgh Supercomputing Center
Machine Room Facility

• Westinghouse Electric Company
  – Monroeville, PA
  – 16 miles East of Pittsburgh

• 11,000 sq ft
Facility Comparison – LeMieux and XT3

- Peak gigaflop
- Cab count
- Space used ft^2
- Area of floor cut-outs
- Circuit count
- Cooling load (tons)

© 2008 Pittsburgh Supercomputing Center
Current Building – Power

• Vendor HPC Power Estimates
  – Exaggerated to be Safe
  – Duty cycle rather than peak
• Current usage is about .9 MW
• Two 23,000 volt underground lines serving an underground automatic substation providing alternate pathing to the site
• Enable service:
  – During routine line maintenance
  – Single-line interruptions.
• Two 23,000-4160 volt transformers
• Automatic switching incoming lines or transformers
• 4160 volt power is distributed to 7 power centers
• Each power center has two 1000/1333 kVA 4160-480/277 volt transformers
Current Building - Power

• **UPS**
  - One 500 kVA UPS unit
  - Two 650 kVA UPS units
  - Static Maintenance Bypass
  - 15 minutes of run time

• **Generator**
  - Five 1 MW diesel generators
  - Three day, 20,000 gallon fuel tank
  - Generators automatically start after power interruption
  - ~ Five minute sync time
  - Shed building load to maintain power to the computer systems and computer room HVAC equipment
  - Upon return of utility power, transfer the load from generator to utility without interruption
Current Building – Power, cont.

- 13 PDU’s in the Computer room
Current Building – Power, cont.

• Grounding
Cooling Planning

• Clearing out under floor space
Cooling Planning

- Clearing out under floor space
Cooling Planning

- Eliminating Vortexes
Cooling Planning

- Temperature Monitoring
  - LeMieux Charts
  - XT3 Charts
- Over-temperature service contract issues
Current Building – Cooling

• Heat Recovery = $60K for 1MW
• Four 300-ton chillers that feed a common manifold.
• Three chillers for capacity and one for redundancy
• Five pumps
  – Two parallel sets
  – 10 HP, 400 GPM
  – Only one pump in each set is required
• 14 Airflow/ADP 30 Ton Air Handling Units
Difficulties in predicting disk failures

- Bianca Schroeder
  Garth A. Gibson
  "Disk failures in the real world: What does an MTTF of 1,000,000 hours mean to you?"

- High percentage of NPF disks
- Power-cycle can “fix”
- Cold is better
  - Google paper adds age

- Power failures
Effects of Power Outages

Lemieux Hardware MTBI

© 2008 Pittsburgh Supercomputing Center
Expanding to Petascale Systems

• Building Regulatory Issues
  – Cooling towers on the Roof
• Component Lead Times
• Power metering
  – Separate meters for functions
• 480 Volts to Racks
  – 80% Savings in infrastructure
  – Includes direct wiring cabinets – no connectors
• Water cooling racks save ~10%
  – Takes the 40% factor to 30%
Expanding to Petascale Systems, cont.

- Used / Uninstalled Equipment
Facility Storage Trends

• Spindle reductions on NSF Program Systems (Courtesy Phil Andrews, NICS and Chris Jordan, TACC)
  – 95 spindles per TFLOP for SDSC DataStar
  – 5 for TACC Ranger
  – 2.5 for NICS Kraken

• Storage scheduling
  – Jobs requesting I/O needs

• Increased Spindle Efficiency
  • “Zest” talk by Paul Nowoczynski, Tuesday, 1:30pm
The End – Thank You