Highlights of Early Facility Experience at the New High Efficiency NCAR Wyoming Supercomputing Center

Aaron Andersen
Deputy Director Operations & Services
May 2013
General Facility Information

- 153,000 Gross Sq.Ft.
- 2 - 12,000 Sq.Ft. Data Halls
- 4 – 4MW electrical modules 16MW possible
- 98~99% annual hours Evaporative Water Side Economizer
- Critical Systems are on UPS and Generator Backup
  - Network, Cybersecurity, Storage, Archive
  - HPC Nodes are not on UPS
- 10 ft. Raised Floor System
  - Primary reason was large liquid cooled systems
  - Secondary reason limit copper distances
- Heat Pump System Recover Heat from Computing for Building Use
Facility Highlights

• All about Energy Efficiency
  – Operational cost savings
  – Emissions reduction

• Minimize risk
  – Not increasing inlet temperatures or using outside air
  – Utilize liquid cooling where cost effective

• Efficient use of Capital
  – Efficiency efforts had to show 5 year payback
Big Picture Focus On Efficiency

- Utilize the regions cool, dry climate to minimize energy use
  - Very low pressure drop
    - Minimize bends
    - Oversized pipe
  - Elevated chilled water temp
    - 65 degree
- Utilize liquid cooled computer solutions where practical
  - HPC Systems
- Utilize hot aisle containment for commodity equipment
- Focus on the biggest losses
  - Compressor based cooling
  - Transformer losses
65 F (18.3 C) Degree Chilled Water Evaporative Solution
Cooling Tower

- Very high efficiency tower
- Utilization of outside conditions to optimize performance
  - *Water use reduction* by not running the fans all year
  - *Energy use reduction*

**Figure 4: Cheyenne weather conditions and cooling tower operation**

- **Region 1:** Cooling tower without fans, 46% of the year
- **Region 2:** Cooling tower with fans, 48% of the year
- **Region 3:** Chiller operates, 7% of the year
Hot Aisle Containment
Fan Wall and 10 ft. (3 m.) Raised Floor
Yellowstone Water Cooling
IT Load and Facility Tuning

PUE vs. IT Load

- **PUE** vs. **IT Load**
- Time (month/year) vs. IT Load (kW) and PUE
- Graph showing the relationship between IT load and PUE over time, indicating that IT load improves PUE.
Actual Annual PUE Trending Toward Design PUE

PUE Integral AVG 1.22

PUE 1.09

Time (month/year)

PUE

5/12 6/12 7/12 8/12 9/12 10/12 11/12 12/12 1/13 2/13 3/13
Two Main Sources of Variability

• Computer Load Variability
  – Power saving computer system at scale
  – Peak Demand ~1700W
  – Typical ~700kW – 1500kW
  – Idle ~300kW

• Weather Variability
  – Heating Load increase in winter months affects PUE
  – Real cost savings in part dependent on natural gas prices
NCAR Data-Centric Architecture

Science Gateways
RDA, ESG

Data Transfer
Services

Remote Vis

Partner Sites

XSEDE Sites

1Gb/10Gb Ethernet (40Gb+ future)

NCAR HPSS Archive
100 PB capacity
~15 PB/yr growth

High Bandwidth Low Latency HPC and I/O Networks
FDR InfiniBand and 10Gb Ethernet

DAV Resource
“geyser” & “caldera”

HPC Resource
“yellowstone”
1.50 PF Peak
29 bluefire-equivalents

CFDS Resource
“glade”

>90 GB/sec aggregate
2012: 11 PB
2014: 16.4 PB

Filesystem servers & IO aggregators
Data Collections
Project Spaces
Scratch
User & Home Archive Interface

ft

login

Filesystem servers &
IO aggregators
NCAR HPC Profile

- IBM iDataPlex/FDR-IB (yellowstone)
- Cray XT5m (lynx)
- IBM Power 575/32 (128) POWER6/DDR-IB (bluefire)
- IBM p575/16 (112) POWER5+/HPS (blueice)
- IBM p575/8 (78) POWER5/HPS (bluevista)
- IBM BlueGene/L (frost)
- IBM POWER4/Colony (bluesky)
Historical Power Efficiency on NCAR Workload

Power Consumption (sustained MFLOP per Watt)

- Yellowstone ~43 sus MFLOPs/Watt
- Lynx ~14 sus MFLOPs/Watt
- Bluefire ~6 sus MFLOPs/Watt
- Bluevista / Blueice ~1.4 sus MFLOPs/Watt
- Frost ~8 sus MFLOPs/Watt
- Cray J90 ~0.2 sus MFLOPs/Watt
Beyond PUE
Highly Dynamic Compute System Loads
System Utilization

Yellowstone Availability & Utilization
4518 nodes, 16 CPU/node

- Avg Availability: 98.02%
- Avg Utilization: 84.82%
- Avg AdvRes: 0.97%

Geyser Availability & Utilization
16 nodes, 40 CPUs/node

- Avg Availability: 78.95%
- Avg Utilization: 11.63%
- Avg AdvRes: 0.00%

Caldera Availability & Utilization
16 nodes, 16 CPUs/node

- Avg Availability: 87.50%
- Avg Utilization: 17.59%
- Avg AdvRes: 0.00%

Erebus (AMPS) Availability & Utilization
84 nodes, 16 CPUs/node

- Avg Availability: 95.57%
- Avg Utilization: 29.67%
- Avg AdvRes: 0.00%

Racks H01 & H02 being tested with cgroups (special queue)
Input queue emptied
Racks H01 & H02 returned to production 04/10 15:00

'Sacks' AdvRes geyser09
Input queue emptied

Yellowstone
Avg Availability 98.02%
Avg Utilization 84.82%
Avg AdvRes 0.97%

Erebus
Avg Availability 95.57%
Avg Utilization 29.67%
Avg AdvRes 0.00%
Summer Mechanical Trend
An Initial Fall Surprise
(Mechanical Load (kw) vs. WetBulb (F))
Looking Closer at October Data

Mechanical Systems Power Use, October 2012

Outside Temperature

Outside Relative Humidity

KW
Comparison of Daily Cost Heat Pumps vs. Boiler

<table>
<thead>
<tr>
<th>Energy Cost per Day</th>
<th>Jan-13</th>
<th>13-Feb</th>
<th>13-Mar</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cost Per Day Gas</strong></td>
<td>$14.14</td>
<td>$122.95</td>
<td>$90.76</td>
</tr>
<tr>
<td><strong>Cost Per Day Elec</strong></td>
<td>$2,870.19</td>
<td>$2,638.93</td>
<td>$2,665.76</td>
</tr>
</tbody>
</table>
NWSC as a Teaching Laboratory
Emerson Hannon
Villanova
Electrical Engineering

• Energy Efficiency
• Modeling of NWSC during construction
• Developed energy model NWSC based on outside conditions
Jared Baker
University of Wyoming
Mechanical Engineering

- Computational Fluid Dynamics
- Validated TileFlow During Construction
- Simulated multiple computer vendor configurations during procurement selection process
• Compared Jared’s Model to Actual Conditions
• Further Tileflow analysis
• Comparison of Hot Aisle Containment
• Temperature pressure measurements
Jason Jones
Clemson University
Mechanical Engineering

- Energy Efficiency Analysis & Measurement
- Comparison of Actual to Model
- Fan wall energy analysis
- Mechanical response to ambient conditions
Ademola Olrinde
Texas A&M Kingsville
Mechanical Engineering

• Energy Efficiency Optimization of NWSC
• Detailed Analysis of 12 month operating Data
• Optimization Efforts
  – UPS Systems
  – Heat Pumps
• Graduate Work in Pipe Flow and Pumping Power
Now Focused on the Left Side of PUE Decimal
"Stochastically Robust Resource Allocation for Thermal-Aware Heterogeneous Computing”

NSF Funded Proposal

- Colorado State University
  - Sudeep Pasricha
  - HJ Siegel
  - Mark Oxely

- NSF Funded Project
- Looking at energy aware resource scheduling
"Optimized Energy Aware Scientific Throughput at Scale"

NSF Proposal Stage

- NCAR
  - Aaron Andersen
  - Stephen Sain
- Colorado State University
  - Sudeep Pasricha
  - HJ Siegel
- Georgia Institute of Technology
  - Yogendra Joshi
- Power Aware Scheduling
- Power Aware Data Movement
- Statistical Approaches to Scale