Co-Designing HPC Systems and Facilities

Michael Oberg
oberg@ucar.edu

Software Engineer, Computer Science Section
Computational and Information Systems Laboratory
National Center for Atmospheric Research

Robert Strong, P.E.
rstrong@cftrm.com

Partner, Critical Facilities Technology

Context

Objective: deploy a top 50 supercomputer at the University of Colorado to support a broad range of research (Earth system Science, Biotechnology / Life Sciences, Energy / Renewable Energy).

Challenge: original facility plans were **not** viable.

Constraints:

Short time line (< 1 year)

Cost effective

“Green”

Solution: leverage rapid deployment technologies in use by industry to design, build, and test off-site a fully-customized pre-fabricated data center.

Janus Supercomputer Details

342 Dell c6100s (four 2-socket blades in 4u)

- 2.8 GHz Intel Westmere-EP
- 24 GB/node

Fully-nonblocking QDR IB

1.2PB DDN SFA10k

1368 compute nodes = 16,416 cores total

- Estimated peak performance – 184 TFLOP/s, ~91% HPL
- 95% of the system (Top500) – 152.2 TFLOP/s, 86.9% HPL
- #31 on June 2010 Top500 List
- #52 on June 2011 Top500 List

Janus Facility Design

APC / Epsilon Modular Solution

Fabricated to order,
not preconfigured containers

3 months to delivery

Power

2 MW feed (~**600KW** used)

60kW N+1 UPS power

Cooling

337 ton chilled water system

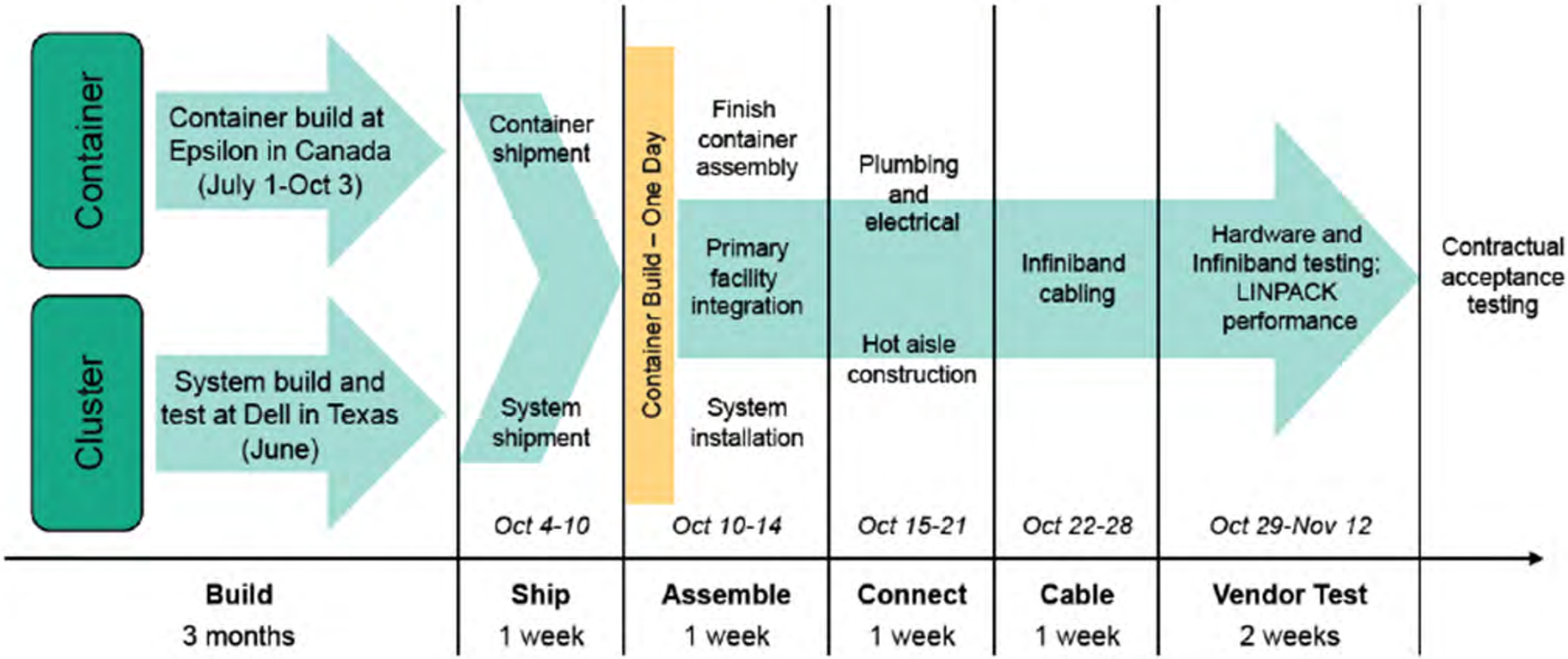
Hot-aisle containment: 24 APC ACRC500 in-row cooling units

1.2 PUE – Average Annual Power Utilization Efficiency

Holistic Design Approach

Tight feedback loop between system and facility design teams with overlap of key technical personnel

- Maximizing power efficiency was a direct result of designing the facility **in parallel** to the design of the system hardware



Administrative Coupling

Operational Efficiency:

- Maintaining balance
- Incorporating information about the system's operational status

Automated policy modification: tune for efficiency ***or*** performance

- Application power profiles: BIOS, system and facility
- Temporary power limitations
- Partial system downtimes
- Instantly react to facility or system events

Traditional vs. Holistic Design Philosophies

Traditional Data Center Design starts with asking what is the most I can get out of an individual component.

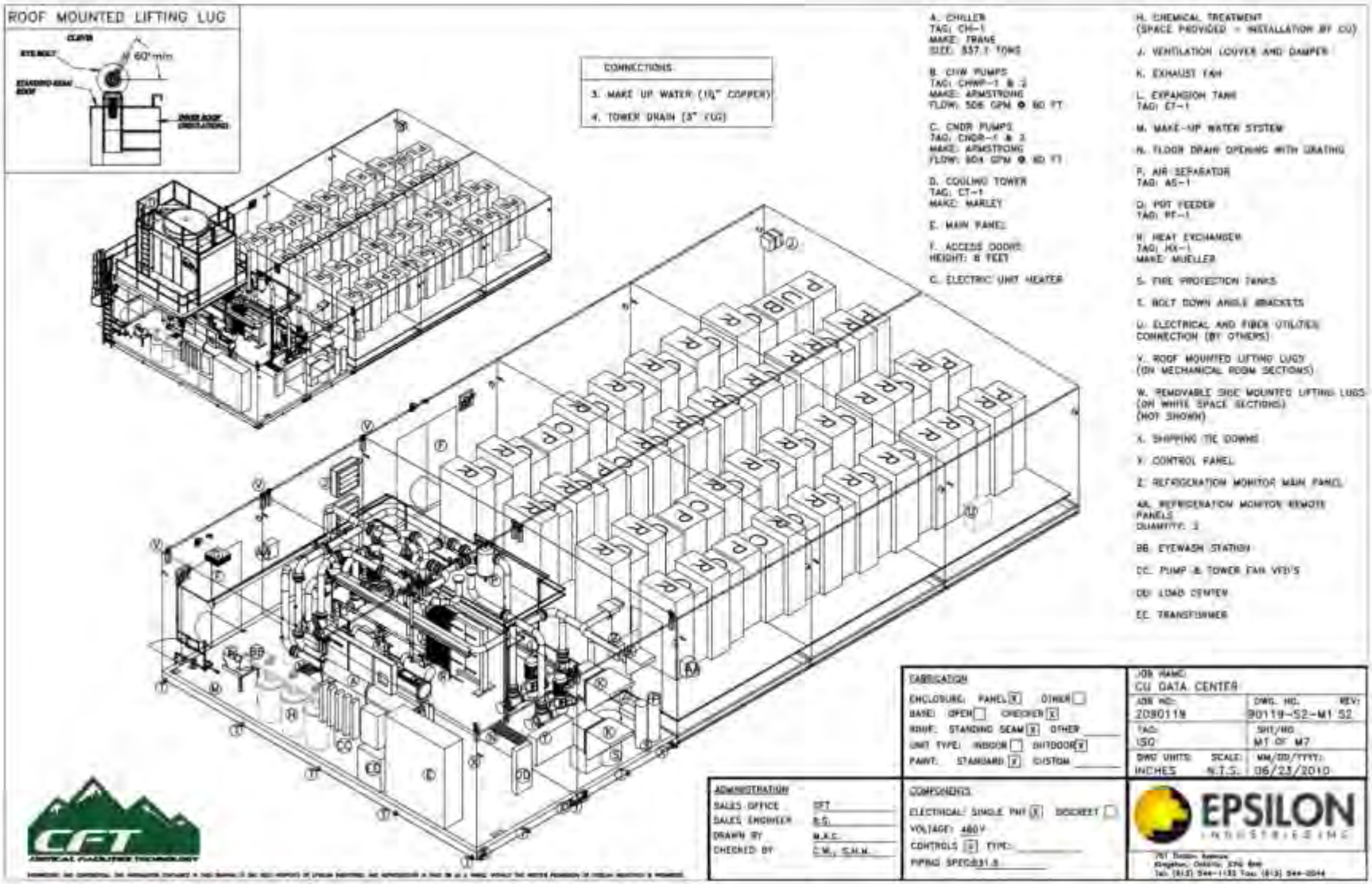
Holistic Data Center Design ask what is the most that I can get out of the system as a whole (including the facility).

Detailed Information on the Server Operating Parameters led to significant mechanical design influences. Airflow, Water Flow, Temperature Deltas

Janus Facility Layout

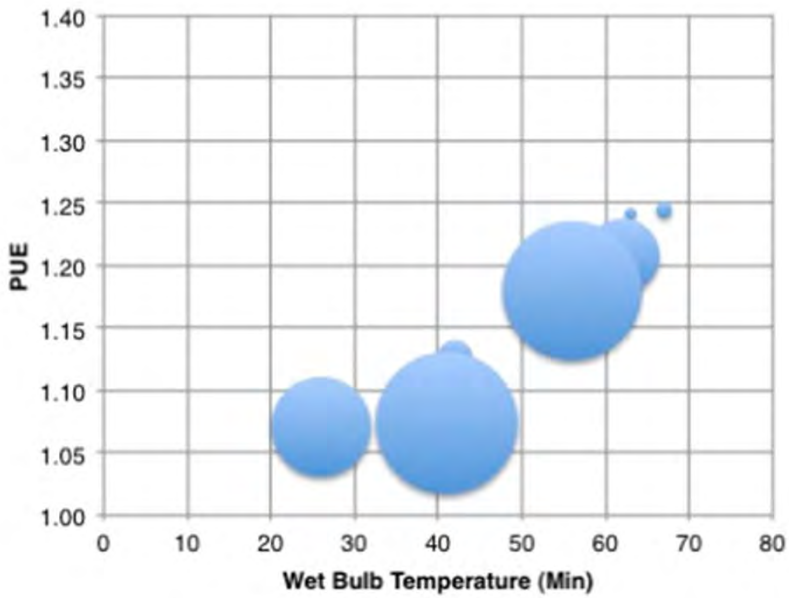
61 x 33.5 x 12.5 feet

APC / Epsilon Modular Data Center

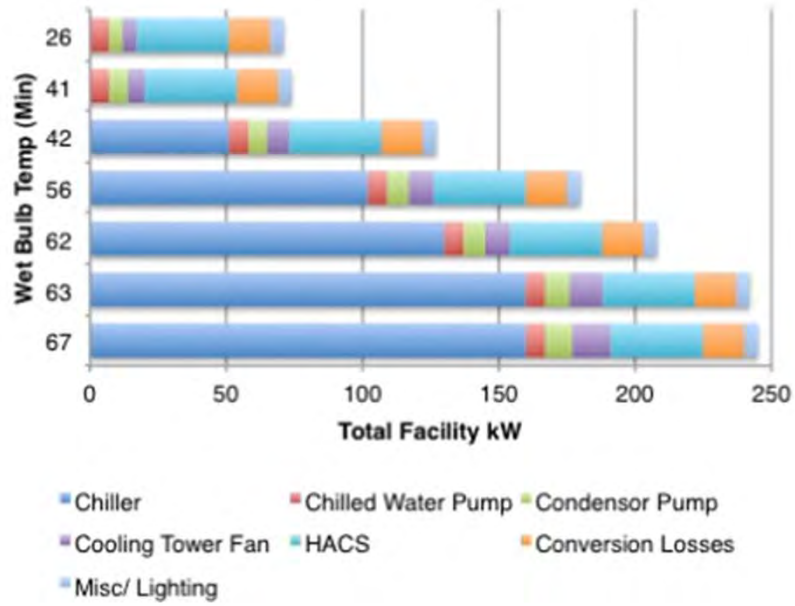


National Center for Atmospheric Research
University of Colorado at Boulder

PUE Analysis



(a) Facility PUE by wet bulb temperature (surface area indicates contribution to yearly average).



(b) Facility power breakdown by component.

Anticipated PUE and power breakdown for 1MW load, showing operating regimes and component values

Original videos available at:

http://web.ncar.teragrid.org/~oberg/HPC_CU_Boulder_Short.mov

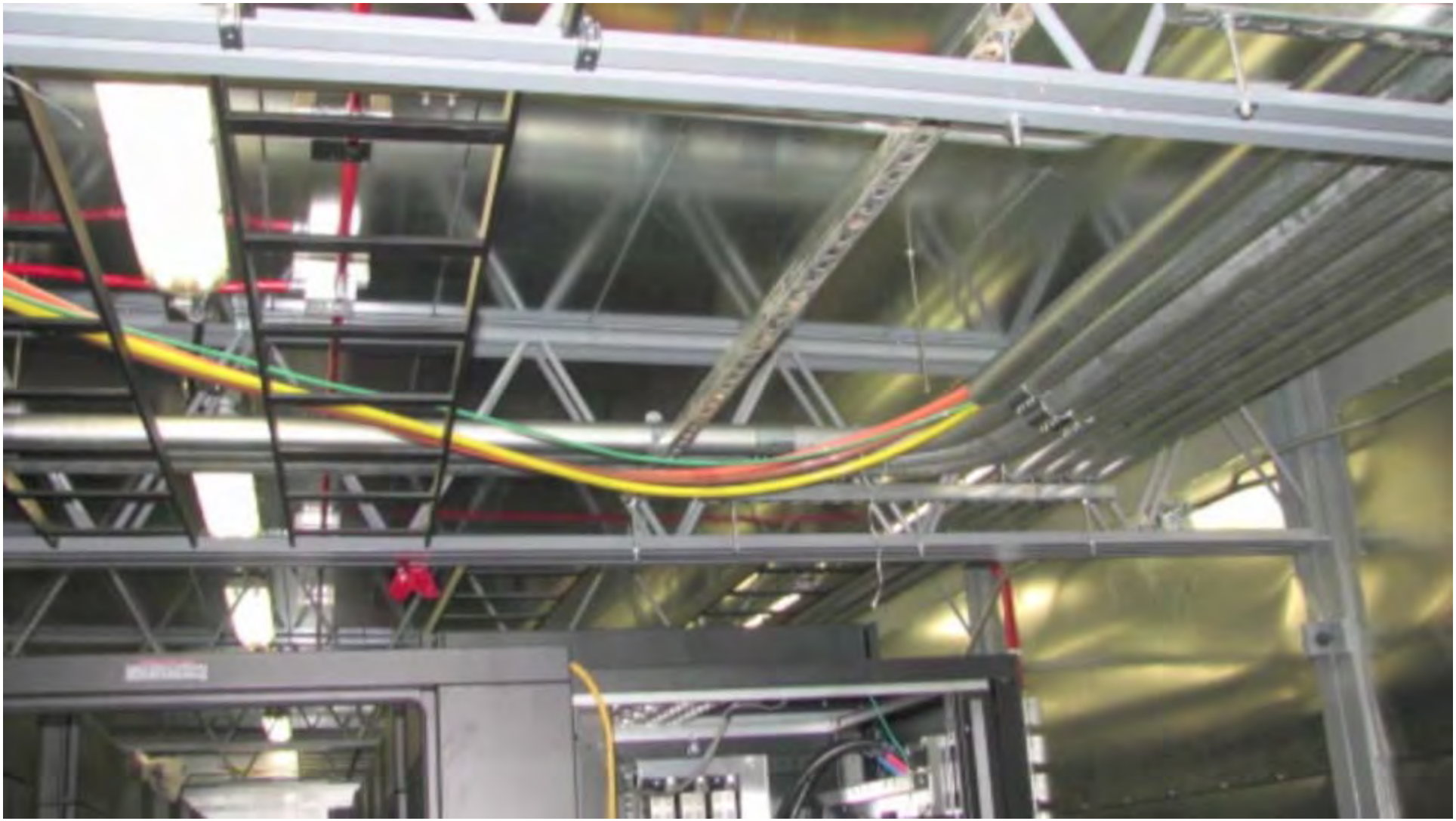
http://web.ncar.teragrid.org/~oberg/20101008_Facility_EmergencyPowerOff-short.mov

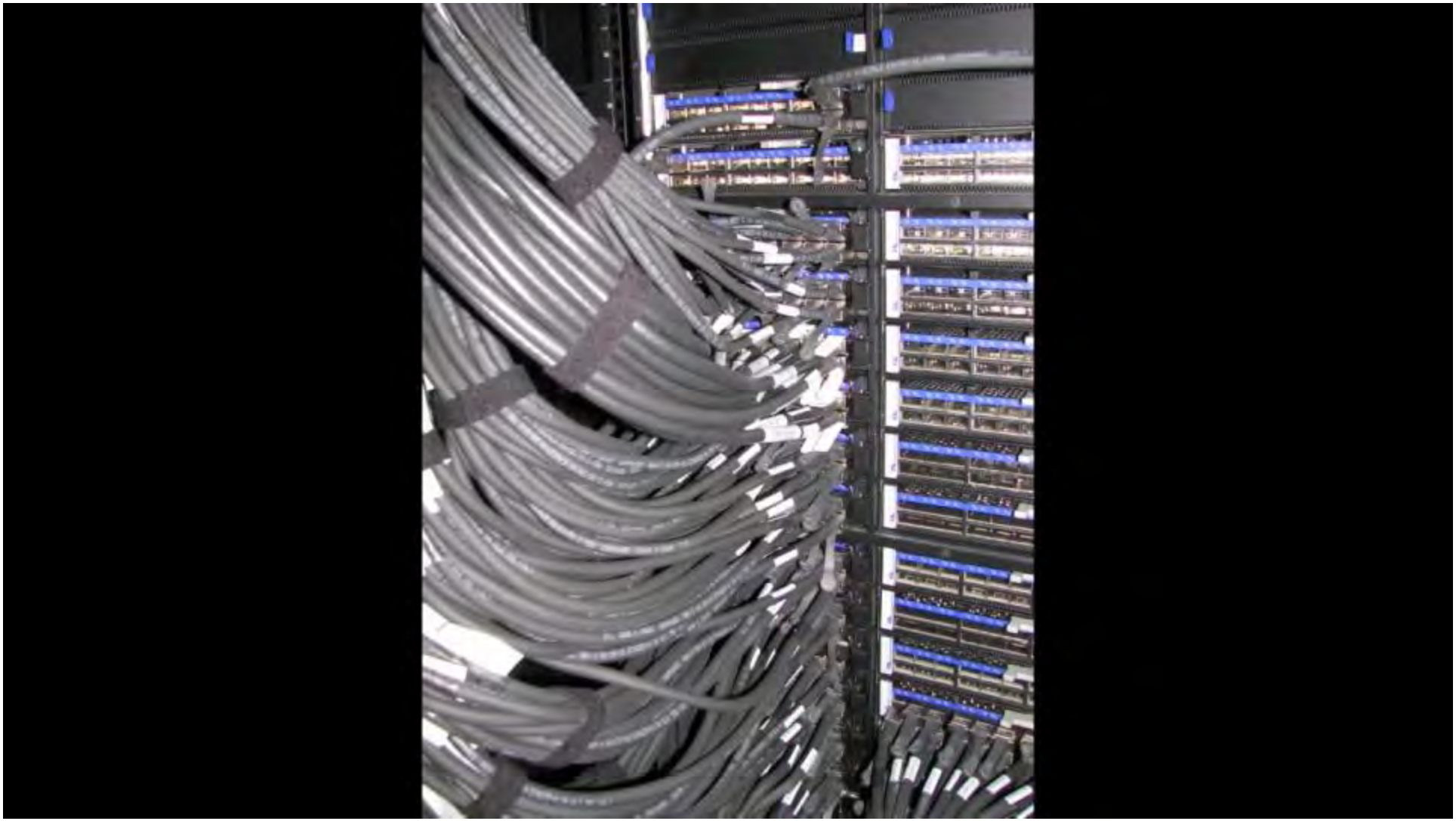
July 2010

October 2010

BAC

OVERSIZE LOAD





CHILLED WATER RETURN
CHILLED WATER SUPPLY
CHILLED WATER SUPPLY

Key Points

Top 50 HPC system

Data center was designed to order, fabricated, and tested off-site in three months (~= time for supercomputer alone)

Target annual PUE of 1.2 (2011 measured 1.14)

Leveraged rapid deployment technologies in use by industry

Facility and system design processes matched - feedback between system and facility engineers critical to minimizing costs and maximizing efficiency.

15+ year lifespan with ability to expand system by 50%

Low cost (approximately $2.5K to $5K per kW)

Questions?

Michael Oberg
oberg@ucar.edu

Software Engineer, Computer Science Section
Computational and Information Systems Laboratory
National Center for Atmospheric Research

Robert Strong, P.E.
rstrong@cftrm.com

Partner, Critical Facilities Technology

Some Design Decisions

By coupling heat rejection with cooling in an in-row form factor in a hot-aisle containment system, we were able to eliminate mixing and thus maximize our coil differential temperatures on the air side and minimizing fan horsepower.

Using elevated cold aisle temperatures allowed us to increase the supplied chilled water temperatures and extend the differential on the water side to 16 deg F (which also allowed us to reduce system pipe sizes and pumping horsepower due to reduced flow).

The elevated chilled water temperatures also allowed us to maximize the number of free cooling hours of operation.

However, this also necessitated an increase in coil surface area ➔ 43% increase in the number of CRAC units (8).