COLUMBIA UNIVERSITY ADVANCED CONCEPTS DATA CENTER PILOT

Final Report

Prepared For

THE NEW YORK STATE ENERGY RESEARCH AND DEVELOPMENT AUTHORITY
Albany, NY

Prepared by:

COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK
New York, NY
Alan Crosswell (PI)
Associate Vice President and Chief Technologist
Columbia University Information Technology

Victoria Hamilton (Co-PI)
Director, Research Initiatives
Office of the Executive Vice President for Research

Agreement No. ST11145-1

April 18, 2013

NOTICE

This report was prepared by Columbia University in the course of performing work contracted for and sponsored by the New York State Energy Research and Development Authority (hereafter "NYSERDA"). The opinions expressed in this report do not necessarily reflect those of NYSERDA or the State of New York, and reference to any specific product, service, process, or method does not constitute an implied or expressed recommendation or endorsement of it. Further, NYSERDA, the State of New York, and the contractor make no warranties or representations, expressed or implied, as to the fitness for particular purpose or merchantability of any product, apparatus, or service, or the usefulness, completeness, or accuracy of any processes, methods, or other information contained, described, disclosed, or referred to in this report. NYSERDA, the State of New York, and the contractor make no representation that the use of any product, apparatus, process, method, or other information will not infringe privately owned rights and will assume no liability for any loss, injury, or damage resulting from, or occurring in connection with, the use of information contained, described, disclosed, or referred to in this report.

ABSTRACT

Columbia University Information Technology (CUIT) piloted advanced concepts data center techniques that emphasized rigorous before-and-after measurements of a series of recommended best practices and innovative equipment and infrastructure improvements in a real-world setting. The measurements were used to (a) monitor whether changes targeted to improve the energy efficiency and environmental impact of the primarily administrative systems currently in the centralized data center were, in fact, effective, and (b) to simultaneously expand the constrained computational capacity of the facility as a result of those improvements. CUIT included participation and critique from influential skeptics from commencement of the project. Two academic departments piloted a shared high performance computing cluster within the Data Center. As a result of this project, Columbia University continues to realize significant energy and environmental gains while demonstrating economic and operational feasibility in infrastructure improvements, uses of innovative computational equipment for both centralized administrative systems and research computing clusters.

The project objectives were to:

1. Become well versed in data center efficiency design techniques and assessment metrics such as those espoused by Green Grid, ASHRAE and others (PUE, DCeP, etc.).

2. Establish baselines and continuously measure several power and cooling variables to enable the study of historical trends and produce a data set to aid in further analysis of data center electrical and heat loading changes over time.

3. Implement key recommended and advanced best practice improvements and measure the degree of efficiency achieved when these techniques are applied in the CUIT data center. These choices were prioritized based on data center power allocation, cost/benefit analyses and where making space available through the use of high density racks was a priority.

4. Implement a number of data center facility infrastructure, IT best practices and advanced data center techniques and validate claims of the degree of IT capacity, power and cooling efficiency improvements that are achieved.

5. Utilize results to inform subsequent phases of the central data center re-fitting and design of future data center space for Columbia's new Manhattanville campus in West Harlem.

6. Provide training and guidance for other data center operators in New York State and elsewhere, especially for our peer higher education institutions.

ACKNOWLEDGEMENTS

For almost four years, we have received help with this project from a large number of colleagues. We would like to acknowledge the work of the following current or former CUIT staff who have been part of the project working group or assisted over time: Ryan Abrecea, Megan Andersen, Rimma Ashkindaze, Rajendra Bose, Anthony Cirillo, Peter Crosta, Stew Feuerstein, Jeff Fleischman, George Giraldi, Richard Hall, Halayn Hescock, Carol Kassel, Ian Katz, Rob Lane, Don Lanini, Anthony Johnson, Dan Mechanic, Melissa Metz, David Millman, Fran Ovios, Shadhin Rahman, Joseph Rini, Lynn Rohrs, Steven Rosenthal, Tiffany Scott and Kathleen Zeigerson. We would also like to thank our colleagues currently or formerly with Columbia Facilities: Michael Iorii, Coury Revan, Joe Mannino, Francis Fitzgerald, Edward McArthur, as well as the members of our Internal and External advisory groups, listed on page 1-8, below.

We thank Professor Henning Schulzrinne of the Departments of Computer Science and Electrical Engineering, and officials of the Standard Performance Evaluation Corporation (SPEC), for arranging for permission for this project to use SPEC benchmarks for internal evaluation.

We would also like to acknowledge the vital work of Haslyna Watson-Sampson, Marie Tracy and Gregory Culler for helping to arrange and produce various meetings, reports and events, and we thank the staff of Columbia Faculty House (a LEED Gold certified facility) for hosting several successful project events.

Finally, we acknowledge the support of the NYSERDA team, especially Bryan Berry and Joseph Borowiec.

TABLE OF CONTENTS

FIGURES

TABLES

SUMMARY

This report describes the work performed for CUIT's Advanced Concepts Data Center Pilot Project for the duration of the project (April 2009 - February 2013). The proposed goals for this project were to:

1. implement a measurement infrastructure to enable verification and evaluation of various energy conservation practices and policies in an established, 24x365 operating data center,
2. implement several infrastructure and IT improvements and measure those improvements, and
3. communicate both success and failure to peer higher education and other interested data center operators.

The active measurement and verification enabled by the NYSERDA award has fundamentally changed how Columbia operates its data center. Beyond its direct scope, the project catalyzed an aggressive effort to consolidate and virtualize most servers in the University Data Center by the year 2014. For this effort $500,000 has been budgeted on an annual recurring basis, which establishes a steady-state, three-year equipment refresh cycle, leading to the benefits of ongoing capacity and energy efficiency improvements in computing and storage equipment.

The NYSERDA award also led to a $10M ARRA grant from the National Institutes of Health to implement energy-efficient facility electrical and UPS capacity upgrades to create a *Core Research Computing Facility* (CRCF) in the Data Center. This award was furthered by $734K in Empire State Development NYSTAR matching funds and a further Columbia match totaling approximately $11M for this project, which is coordinated with and builds upon the NYSERDA-catalyzed improvements.

Our project team has successfully completed all but one of the *initially* identified detailed goals. We inventoried the Columbia data center, introduced power and temperature measurement instrumentation, collected power usage data at regular intervals, produced an overall data center power consumption profile, replaced older servers and compute clusters with newer and more efficient hardware, explored additional power management features at the server level., and communicated our results through blogs, workshops, and conference presentations. The team was unable to deploy in-row/rack cooling technology after an engineering study revealed the cost to retrofit this capability to be significantly greater than was originally estimated. This goal was revised to further explore the feasibility of several other potential approaches to improving cooling infrastructure energy efficiency and led to the selection of an overhead electrical distribution bus to enable improved under-floor airflow. .

Overall, improvements due primarily to the refresh, consolidation, and virtualization of servers have resulted in significant positive results. Cooling technology improvements have been much more difficult to attain and measure. Over the course of three years we have reduced our data center energy consumption by approximately 17% (707 MWh), which is a carbon footprint reduction of approximately 270 MTCE (metric tons CO_2 equivalent), while simultaneously increasing our computing capacity 20% and our data storage capacity by a factor of three. In addition, data collection has enabled the calculation of a common efficiency metric called Power Usage Effectiveness (PUE). Our data center's annualized average is 2.1, which indicates that every 1 kW of power devoted to IT service requires 1.1 kW of power and cooling infrastructure overhead. We will be able to use the measurement infrastructure and knowledge gained in this study to, on an ongoing basis, understand our PUE, which reflects facilities infrastructure efficiency, as well as improve our Data Center Energy Productivity (DCeP), which reflects "useful work" performed by the IT equipment in terms of energy consumed.

A major objective of this project was to examine achievable improvements in terms of energy, economics and environmental impact in an operational data center, and use tangible measurement results to demonstrate to faculty and administrators the value of shared, centralized research computing resources, such as High Performance Cluster computing (HPC). This report demonstrates our significant progress with this goal, and will support further advances in achieving data center and IT services energy efficiency at Columbia.

1. PROJECT RESULTS, MANAGEMENT AND GOVERNANCE

KEY PROJECT RESULTS AND CATALYSTS

The proposed goals for this project were to implement a measurement infrastructure to enable verification and evaluation of various energy conservation practices and policies in an established, 24x365 operating data center, and to communicate both success and failure to peers. Specific tasks included: server refresh, consolidation and virtualization; and improvements to cooling technology. We can report significant positive results: we now have the ability to continuously measure energy consumption and calculate Power Usage Effectiveness (PUE); we have reduced our data center energy consumption over the past three years by approximately 722 MWh, or a carbon footprint reduction of approximately 305 MTCE (metric tons CO_2 equivalent); and, simultaneously, we have increased our computing capacity 20% and our data storage capacity by a factor of three.

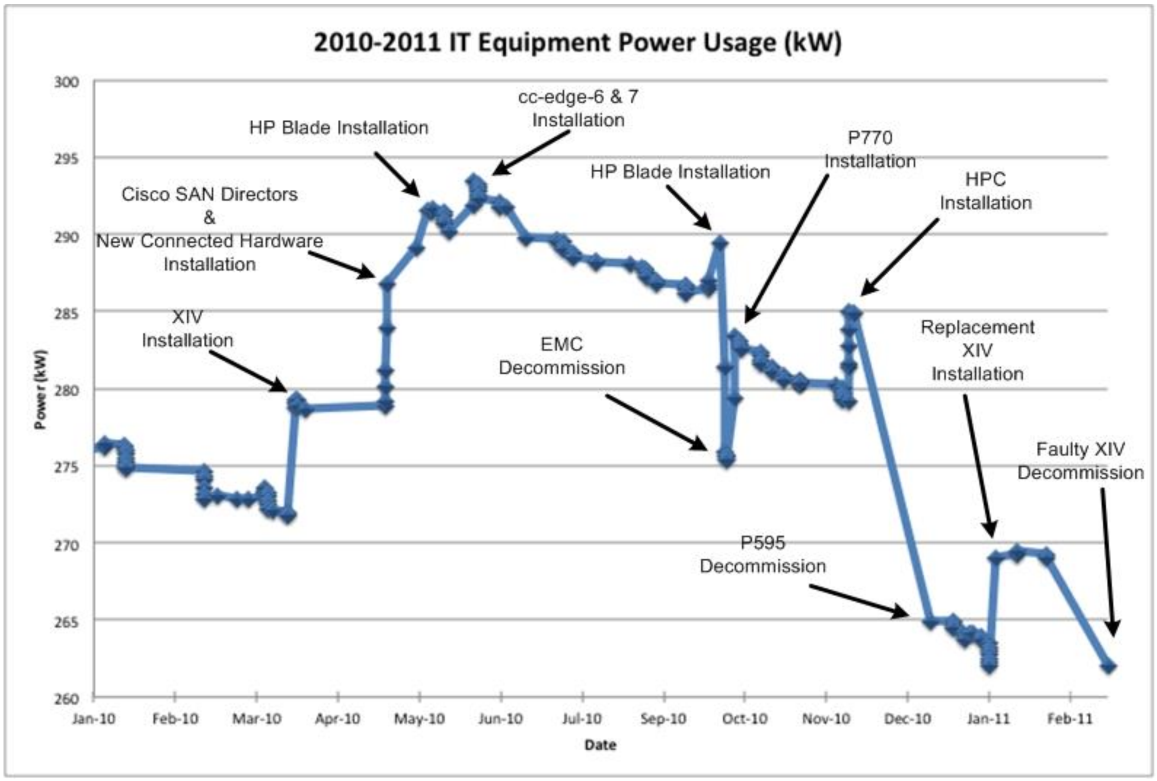


Figure 1-1. IT equipment energy efficiency improvements January 2010 to March 2011

The active measurement and verification that the NYSERDA award catalyzed through our Advanced Concepts Data Center project has fundamentally changed how Columbia operates its data center, as discussed in the following sections.

CONTINUOUS POWER USAGE EFFECTIVENESS AND OTHER MEASUREMENTS

Columbia's data center is now extensively instrumented for real-time measurement of electrical and cooling loads. This has enabled us to calculate our PUE in real time, and to confirm seasonal variations in PUE thanks to use of air- and water-side economizing. See Figures 5-1 through 5-4.

Our comprehensive energy metering has proved an effective tool in measuring energy savings. Industry trends have shown that new equipment is more energy efficient than older equipment. Figure 1-1 illustrates this trend toward lower energy usage for servers and storage in the CUIT data center beginning in 2010 and continuing into 2011. The increased load shows overlap between old and replacement equipment, followed by decreases when old equipment has been removed. Major changes happened when replacing two EMC storage systems with a higher capacity IBM XIV system and when two IBM p595 servers were replaced by two IBM p770 servers. Smaller dips indicate individual old servers being removed. Further savings are being projected and will continue to be documented as we move toward more aggressive server consolidation goals.

				3-yr
	2010	2011	2012	Total
Annual IT Equipment Demand Load Reduction (kW)	**11.01**	**9.19**	**18.96**	**39.16**
Power Usage Effectiveness (PUE)	**2.06**	**2.10**	**2.13**	**2.10**
Total Facility Demand Load Reduction (kW)	**22.69**	**19.29**	**40.39**	**82.37**
Annual kWh saved (kW*24hrs*365days)	**198,754**	**168,986**	**353,790**	**721,529**
Annual $ saved (kWh*$0.185)	**$36,770**	**$31,262**	**$65,451**	**$133,483**
Total Facility Carbon Footprint Reduction (MTCE)*	**83.9**	**71.4**	**149.4**	**304.7**
**Calculated using 0.00042227 MTCE/kWh*				

Table 1-1. Data center demand load reduction 2010-2012

We estimate a cumulative IT power demand load reduction of 39 kW after three years, and that we have reduced our data center energy consumption by approximately 722 MWh, or a carbon footprint reduction of approximately 305 MTCE (metric tons CO_2 equivalent). See Table 1-1[1]. Figure 1-2 shows the IT demand load reduction over three years.

[1] Using coefficients from: NYC Mayor's Office of Sustainability and Long Term Development. Inventory of New York City Greenhouse Gas Emissions. December 2012. http://nytelecom.vo.llnwd.net/o15/agencies/planyc2030/pdf/greenhousegas_2012.pdf

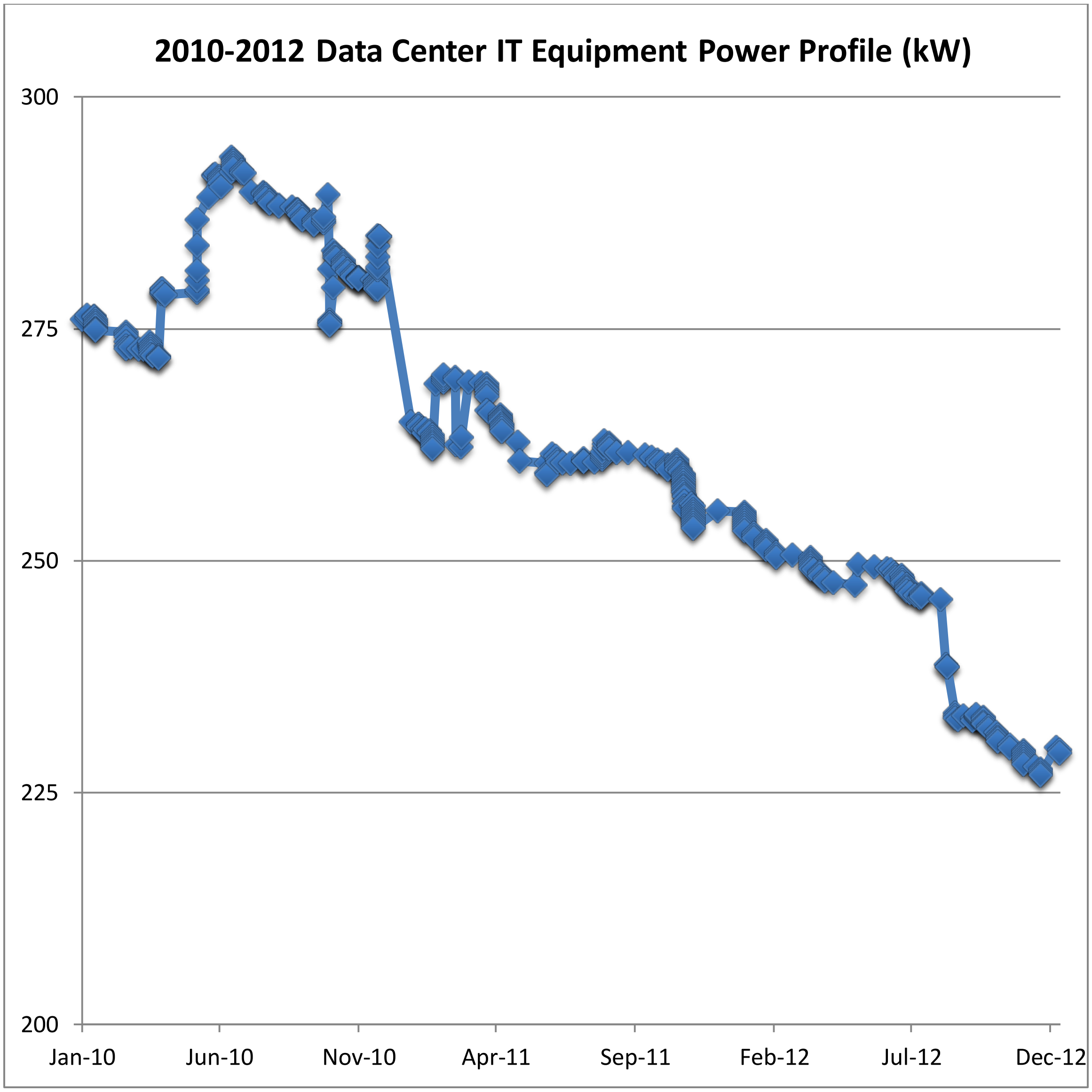


Figure 1-2. IT equipment energy efficiency improvements January 2010 to December 2012

STORAGE REFRESH PROJECT

Our storage refresh project involved installing an IBM XIV to replace our existing EMC DMX 2000 and DMX 800 storage devices. Figure 1-2 shows a 226% increase in the total raw storage capacity while increasing power usage by only 27%. This results in a 61% reduction in the power usage per terabyte of storage provided by the IBM XIV and DS8100 compared to the retired EMC DMX 2000 and 800.

The storage refresh project leveraged our new green data center approaches to (a) require the bidding vendors to document their energy efficiency as one of the selection criteria, and, (b) implement a new overhead cable distribution system to begin the (lengthy) process of removing under floor airflow blockages.

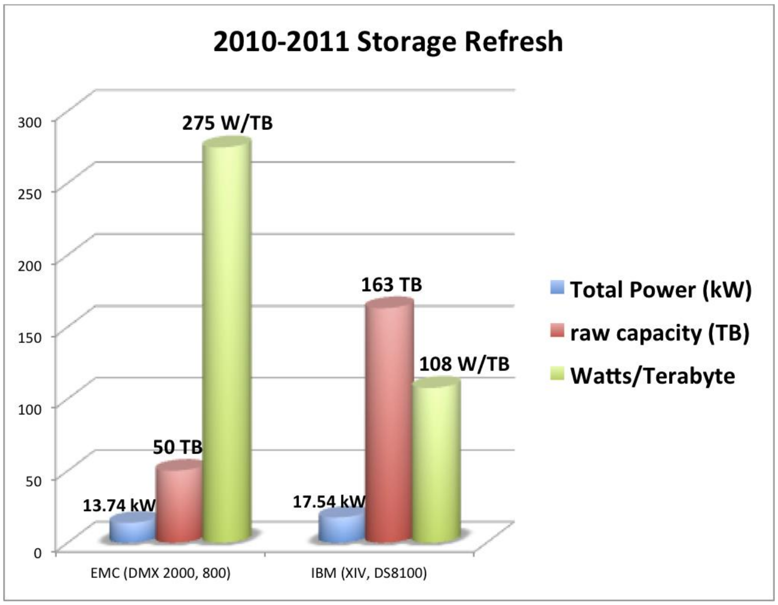


Figure 1-3. Storage systems power consumption

	2010	2011	2012	3-yr Total
Total Machines Retired	**103**	**101**	**80**	**284**
Total Machines Installed	**84**	**67**	**21**	**172**

Table 1-2. Data center servers retired and installed 2010-2012

SERVER CONSOLIDATION PROJECT

The inventory work in project task 2 continued beyond the original project scope, to advance an aggressive server consolidation project. Individual server machines over three years old were replaced with new high-density blade servers or turned into virtual machines (VMs). Most of the minor downward trends shown in Figure 1-1 are from the retirement of these individual old servers. Figure 1-2 shows that, during the first three years of our server consolidation project from 2010 through 2012, CUIT retired 284 machines, which is 112 more servers retired than the total of 172 installed, thus decreasing power consumed by the data center. Not only do the new servers consume less power, they also have more computing capacity. Figure 1-4 shows the estimated increase in compute performance of about 20% for servers in the data center over the course of 2010 into 2011. See Section 7 for the basis of this calculation.

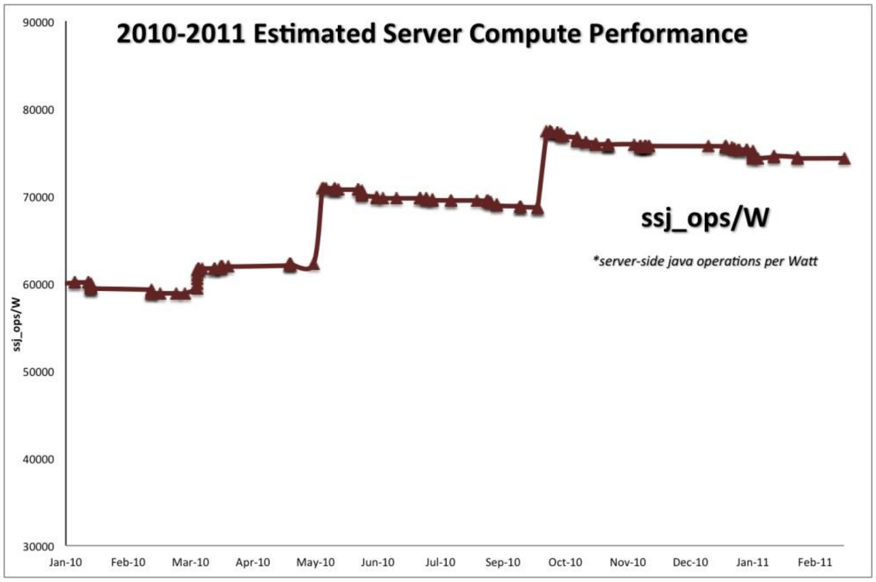


Figure 1-4. Increase in server compute performance

CORE RESEARCH COMPUTING FACILITY

The NYSERDA award led to a successful $10M grant from the National Institutes of Health (NIH Research Facility Improvement Grant 1G20RR030893-01, awarded April 15, 2010. This is supplemented by an additional $1M 10% match from the New York State Foundation for Science, Technology and Innovation (NYSTAR) and Columbia. The NIH grant is in the final construction phase of energy-efficient facility electrical and UPS capacity upgrades to create a *Core Research Computing Facility* (CRCF) to consolidate high performance research computing in a shared, multi-disciplinary, and more energy-efficient approach than is typical at major research universities. We believe it is particularly vital to develop an energy-efficient solution for future research computing needs. In particular, the use of energy-intensive high performance computing (HPC) is experiencing dramatic growth throughout all areas of research, from simulation to extensive scientific data analysis. HPC growth far outstrips that of typical commercial computing workloads, which themselves have continued to grow at a steady pace, while also migrating to external cloud services.

SHARED HIGH PERFORMANCE COMPUTING CLUSTER

High Performance Computing (HPC) for research has benefitted through our demonstrated ability to save energy and other costs through sharing of a research computing cluster. This cluster, initially consisting of 32 compute servers with 256 cores, started as a pilot for two departments, Astronomy and Statistics, has since been expanded to include several other research groups and now has 62 compute servers with 616 cores – while still reducing the overall data center power consumption. This shared facility now serves 200 users and has resulted in 62 research publications to date, including four PhD theses.[2]

[2] https://wikis.cuit.columbia.edu/confluence/display/rcs/Research+Products

In addition to the $10M National Institutes of Health (NIH) grant and supplemental funds to provide increased centralized research computing electrical capacity with improved energy efficiency, our NYSERDA-sponsored work enabled us to submit proposals in 2011, 2012 and coming in February 2013 to the National Science Foundation's Major Research Instrumentation (MRI) program to further increase cross-disciplinary collaboration and again upgrade the capacity and number of researchers sharing the HPC cluster[3].

PROJECT MANAGEMENT

Project Tasks

The project was broken down into twelve major tasks. These tasks, comprising the contracted scope of work, were as follows:

1. Project Management
2. Inventory
    a. Create detailed physical inventory of existing in-scope servers
3. Instrument server power consumption
    a. Install network monitored power monitors for each server
    b. Perform data collection at 5-minute intervals
4. Instrument server input air temperature and overall data center chilled water
    a. Install server input ambient air temperature measurement for each server cabinet
    b. Install energy metering for data center chilled water supply and return lines
    c. Perform data collection at 5-minute intervals
5. Establish overall data center profile
    a. Use equipment load results to establish baseline energy consumption measurements
    b. Determine the Power Usage Effectiveness (PUE) ratio for entire data center
6. Investigate alternatives for HVAC efficiency improvements[4]
    a. 9 server racks outfitted for high power density
        i. Solicit, review and select vendor product to implement 9 racks of high power density in-row cooling.
        ii. Develop feasibility study, engineering design and budget estimates to interconnect those 9 racks of in-row cooling into the existing facility HVAC systems.
        iii. If feasible, implement above.
    b. Improvements to existing forced air cooling
        i. Perform a base Computational Fluid Dynamics (CFD) analysis consisting of surveys of existing conditions, including under-floor air flow blockages.

[3] This was not awarded 2011 or 2012 but was scored sufficiently favorably to again resubmit.
[4] The scope of this task was revised approximately 12 months into the project.

ii. Produce a projected CFD analysis incorporating several options:
   1. Removal of under floor blockages.
   2. Addition of CRAC return air ducting via the hung ceiling.
   3. Addition of hot aisle curtain containment.

iii. Develop energy savings projections for options (ii), above.

iv. Develop conceptual drawings and budget estimates for:
   1. Overhead electrical bus distribution.
   2. HVAC CRAC return ducting.
   3. HVAC curtain containment.
   4. Coordinated HVAC CRAC control system.

v. Develop budget estimates for the options (iv) above.

vi. If feasible, implement one or more of the above options (iv).

7. Replace 30 "old" servers and measure efficiency improvement
   a. Consolidate the replacement servers into high density racks and re-implement the same IT services
   b. Take measurements of before-and-after power consumption
   c. Document expected and actual efficiency improvement
8. Compare old and new research high performance computing clusters
   a. Run benchmark applications on new Astronomy/Statistics HPC cluster
9. Implement server power management
   a. Implement server BIOS and/or Operating System power management features on servers identified in Task 2
10. Increase chilled water set point and measure
    a. Document measured before-and-after energy consumption
11. Communicate results
    a. Share results with key stakeholders

Project Duration

The contracted timeline for this project was originally 18 months, from April 1, 2009 until October 1, 2010. Columbia University requested and was granted two no-cost extensions to February 2013. The purpose of the extensions was to ensure the completion of all agreed deliverables, improve the collection of the data set for the project's Final Report, and to change the scope of task 6 to investigate several alternatives for HVAC efficiency improvements, in addition to in-row/rack cooling which was determined to be not practicable after detailed engineering studies were performed.

Project Governance

Figure 1-5 outlines the project governance structure per the standard project management methodology employed by CUIT. Highlights of the governance structure, which includes executive oversight and a number of advisory committees is outlined below.

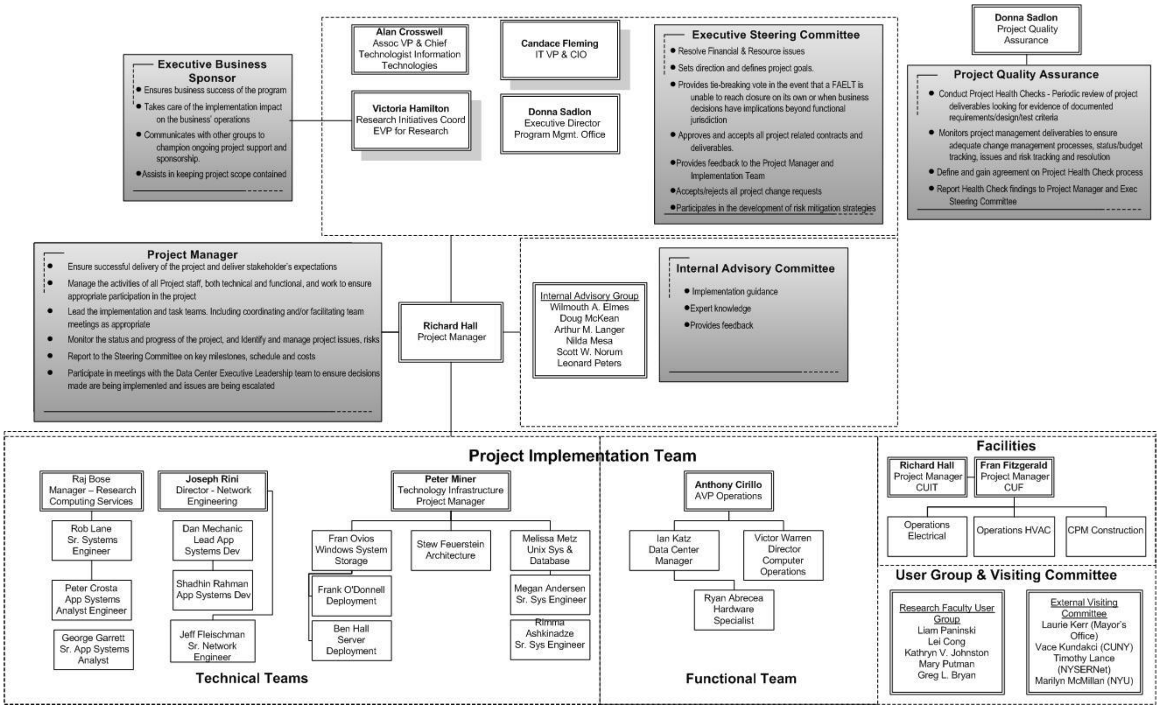

Executive Business Sponsor
Ensures business success of the program
Takes care of the implementation impact on the business' operations
Communicates with other groups to champion ongoing project support and sponsorship.
Assists in keeping project scope contained
Alan Crosswell
Assoc VP & Chief Technologist Information Technologies
Victoria Hamilton
Research Initiatives Coord
EVP for Research
Candace Fleming
IT VP & CIO
Donna Sadlon
Executive Director
Program Mgmt. Office
Executive Steering Committee
Resolve Financial & Resource issues
Sets direction and defines project goals.
Provides tie-breaking vote in the event that a FAELT is unable to reach closure on its own or when business decisions have implications beyond functional jurisdiction
Approves and accepts all project related contracts and deliverables.
Provides feedback to the Project Manager and Implementation Team
Accepts/rejects all project change requests
Participates in the development of risk mitigation strategies
Donna Sadlon
Project Quality Assurance
Project Quality Assurance
Conduct Project Health Checks - Periodic review of project deliverables looking for evidence of documented requirements/design/test criteria
Monitors project management deliverables to ensure adequate change management processes, status/budget tracking, issues and risk tracking and resolution
Define and gain agreement on Project Health Check process
Report Health Check findings to Project Manager and Exec Steering Committee
Project Manager
Ensure successful delivery of the project and deliver stakeholder's expectations
Manage the activities of all Project staff, both technical and functional, and work to ensure appropriate participation in the project
Lead the implementation and task teams. Including coordinating and/or facilitating team meetings as appropriate
Monitor the status and progress of the project, and Identify and manage project issues, risks
Report to the Steering Committee on key milestones, schedule and costs
Participate in meetings with the Data Center Executive Leadership team to ensure decisions made are being implemented and issues are being escalated
Richard Hall
Project Manager
Internal Advisory Group
Wilmouth A. Elmes
Doug McKean
Arthur M. Langer
Nilda Mesa
Scott W. Norum
Leonard Peters
Internal Advisory Committee
Implementation guidance
Expert knowledge
Provides feedback
Project Implementation Team
Raj Bose
Manager – Research Computing Services
Rob Lane
Sr. Systems Engineer
Peter Crosta
App Systems Analyst Engineer
George Garrett
Sr. App Systems Analyst
Joseph Rini
Director - Network Engineering
Dan Mechanic
Lead App Systems Dev
Shadhin Rahman
App Systems Dev
Jeff Fleischman
Sr. Network Engineer
Peter Miner
Technology Infrastructure Project Manager
Fran Ovios
Windows System Storage
Frank O'Donnell
Deployment
Ben Hall
Server Deployment
Stew Feuerstein
Architecture
Melissa Metz
Unix Sys & Database
Megan Andersen
Sr. Sys Engineer
Rimma Ashkinadze
Sr. Sys Engineer
Technical Teams
Anthony Cirillo
AVP Operations
Ian Katz
Data Center Manager
Victor Warren
Director Computer Operations
Ryan Abrecea
Hardware Specialist
Functional Team
Facilities
Richard Hall
Project Manager
CUIT
Fran Fitzgerald
Project Manager
CUF
Operations Electrical
Operations HVAC
CPM Construction
User Group & Visiting Committee
Research Faculty User Group
Liam Paninski
Lei Cong
Kathryn V. Johnston
Mary Putman
Greg L. Bryan
External Visiting Committee
Laurie Kerr (Mayor's Office)
Vace Kundakci (CUNY)
Timothy Lance (NYSERNet)
Marilyn McMillan (NYU)

Figure 1-5. Project governance

Implementation Team

The Implementation Team has day-to-day responsibility for the planning, execution and measurement required by the proposal, supplemented by consultants to perform feasibility studies, engineering design work and to advise on best practices and effective technological innovations.

Research Faculty User Group

The Research Faculty User Group is responsible for vetting the results of the combined HPC cluster for the Department of Astronomy and the Department of Statistics. Representing the two departments are individual professors who have been involved in the local departmental clusters, and are therefore ideally placed to compare the status quo to the new shared HPC cluster.

Department of Statistics

Liam Paninski, Associate Professor. Collaborating with Biological Science and Neuroscience Professor Rafael Yuste, has been working to combine new experimental and analytical methods to reverse engineer large neuronal circuits. Specifically, they optically measure the spontaneous and evoked activity of neuronal populations in cortical brain slices and then use statistical methods to estimate the network connectivity from the observed correlated neuronal firing patterns. The necessary computations turn out to be extremely amenable to parallelization using HPC.

Department of Astronomy

Kathryn V. Johnston, Associate Professor. Uses HPC to run thousands of small-scale simulations of the disruption of purely dark matter halos, and subsequently ``painted'' these simulations with analytic descriptions of the embedded stellar distribution and overlaid them to model the diffuse stellar distribution around galaxies.

Mary Putman, Associate Professor. Research interests include galaxy formation and evolution, intergalactic medium, halo gas and star formation.

Greg L. Bryan, Associate Professor. Has developed an adaptive, highly parallel hydrodynamics code, Enzo, and used it to explicitly model the complex baryonic physics of star formation and feedback. Bryan's focus has been on understanding how to accurately follow the physics that shapes galaxies.

Internal Advisory Group

The Internal Advisory Group serves two primary functions. Not only are they a valuable sounding board for the execution of the pilot, they are also expected to pose the hard questions about issues surrounding scale-up. To extend the results of the pilot throughout the institution, we need the active support of multiple constituencies. Moreover, many in this group belong to influential external groups and can thus share the results of our program. While each is available for informal consultation and updates, a formal meeting will be held upon the commencement of the project and at six month intervals thereafter.

Wilmouth A. Elmes, Associate Vice President of Engineering/Technical Services, Manhattanville Development Project. Responsible for providing mechanical and electrical technical input, and guidance to in house engineers, project managers and all outside consultants currently engaged on the University's Manhattanville Project to assure that the systems being developed comply to standards set by the University's Facilities Group, Information Technology Group and other university stake holders throughout the campus. In particular, challenges the engineering design team to provide cost effective energy efficient designs that conform to the New York State Energy Conservation Code, the USGBC LEED Guidelines, and Columbia University's Sustainability Framework Guidelines, including review and approval of all energy-related studies for onsite cogeneration, thermal storage, fuel cells, heat recovery and other energy conservation opportunities that may be available for all new building projects on the Manhattanville Project. Member of the US Green Building Counsel (USGBC).

Arthur M. Langer, School of Engineering and Applied Science, Senior Director of the Center for Technology, Innovation, and Community Engagement and Faculty & Associate Director, Executive Masters of Science in Technology Management at the School of Continuing Education. Responsible for the design and faculty coordination of the masters program in executive technology management. Created mentors program that provides students with an executive mentor from industry. Dr. Langer is the author of Analysis & Design of Information Systems (2007), Information Technology & Organizational Learning (2005), Applied Ecommerce (2002), and The Art of Analysis (1997) and has published numerous articles and papers. Member of the Board of Directors, V.P., Academics, Society for Information Managers, New York Chapter, and serves on the Editorial Board, International Refereed Journal of Reflective Practice, Carfax Publishing and Advisory Board, CIOZone among other organizations.

Nilda Mesa, Assistant Vice President, Environmental Stewardship. Founded Columbia University sustainability office to develop programs and policies to lessen the University's environmental footprint. Develops and implements initiatives and policies on behalf of the President and Senior Executive Vice President and oversees University greenhouse gas emissions inventory and action plan development, including energy strategy. Works with NYC Mayor's Office of Long-Term Planning and Sustainability on a major university initiative to decrease greenhouse gas emissions 30% in 10 years. Member, Manhattan Borough President's Go Green Standing Committees. Established energy and air quality partnerships with the Sierra Club and Environmental Defense Fund. Steering Committee, NECSC. Member, AASHE, USGBC, Ivy Plus Sustainability Working Group; Adjunct Professor, Columbia University School of International and Public Affairs.

Scott W. Norum, Chief Administrative Officer for Arts and Sciences and Vice President, Office of the Vice President for Arts and Sciences. Reporting to the Vice President and Dean of Faculty for Arts and Sciences, responsibilities include organizational and strategic planning, financial and budget management. Responsible for establishing frameworks for planning and coordination among the six schools and 29 academic departments of Arts and Sciences, forecasting and providing for the operating and capital requirements of these organizations. Also prepares the annual operating plan and capital budget for Arts and Sciences, monitors against plan during the year, prepares quarterly variance analysis, oversees adherence to NY State endowment statutes, and all other funding mechanisms and policies.

Leonard Peters, School of Business, Associate Dean and Chief Information Officer – Information Technology. As a member of the Business School's senior management team, responsibilities include strategic direction and management of all aspects of technology related to teaching, faculty research, students and administration. Functional areas managed are network services, computer labs, faculty research computing, software development, IT training and all technology assets. Member of numerous organizations including Educause-Higher Education Technology Organization, Gartner's Business & Technology Decision-Maker Panel and Business Computing Directions – IT leaders from top business schools.

External Visiting Committee

Analogous to the Internal Advisory Group, the External Visiting Committee is charged with maintaining skepticism about our plans and our accomplishments, but from the perspective of external institutions. To the extent we demonstrate feasibility and potential impact, they will serve as powerful ambassadors to other institutions. Each member represents important external constituencies. While each is available for informal consultation and updates, a formal meeting will be held upon the commencement of the project and at six month intervals thereafter.

Laurie Kerr, Senior Policy Advisor for Energy and Green Buildings, NYC Mayor's Office of Long-Term Planning and Sustainability. The New York City Mayor's Office of Long Term Planning and Sustainability was created to coordinate and institutionalize the implementation of the 127 sustainability initiatives outlined in Mayor Bloomberg's PlaNYC 2030 (www.nyc.gov/planyc). Several initiatives in this long term plan aim to ensure that New York City attains the cleanest air quality of any major U.S. city by the year 2030.

Vace Kundakci, Assistant Vice President for Information Technology and Chief Information Officer, City College of New York/City University of New York**.** Responsible for all telecommunications and networking infrastructure and services; data center operations including several research clusters, communication systems including email and web; student computing facilities; classroom and special events A/V; help desk administrative computing; desktop support; and IT training for The City College of New York (CCNY). CCNY is the first college of The City University of New York (CUNY), and a comprehensive teaching, research, and service institution dedicated to accessibility and excellence in undergraduate and graduate education. It has 15,000 students and thirteen doctoral programs. Member of a number of CCNY and CUNY committees such as CUNY High Performance Advisory Committee and CUNY IT Strategic Planning Committee.

Timothy Lance, President and Board Chair, NYSERNet. NYSERNet is a private not-for-profit corporation created to foster science and education in New York State. Its mission is to advance network technologies and applications that enable collaboration and to promote technology transfer for research and education. An internet pioneer, NYSERNet has delivered next-generation Internet services to New York State's research and education community for more than twenty years. NYSERNet members include New York State's leading universities, colleges, museums, healthcare facilities, primary and secondary schools, and research institutions. NYSERNet's Board of Directors is composed of CIO's and other senior personnel drawn from and representing New York's leading research universities and institutions.

Marilyn McMillan, Associate Provost and Chief Information Technology Officer, New York University. Leads the delivery and evolution of University-wide services, infrastructure, policies, and plans for information technology and related activities. Her responsibilities include leadership of NYU Information Technology Services (ITS), coordination with providers of IT-related services in schools and departments, and facilitation of planning and policy development for information technology. In these matters, she works closely with the deans, the vice presidents, and other senior officers. She convenes the Faculty Working Group on IT Direction and Services, the CIO Council, and the HIPAA Working Group.

2. INVENTORY

INTRODUCTION

The project scope of work was to conduct an inventory of existing server racks and their contents within the data center and to identify the following hardware:

- 30 servers more than three years old;
- An existing 100-node Electrical Engineering research cluster in the central data center;
- An old Astronomy computing cluster in a departmental server room;
- A new shared Astronomy/Statistics 32-node HPC cluster in the central data center.

This project task served to accelerate plans for a more comprehensive inventory of the data center, and it spurred an aggressive effort to consolidate and virtualize most servers through 2014. For this, $500,000 has been budgeted annually, which allows for the maintenance of a steady-state, three-year equipment refresh cycle. More information about this consolidation project is provided in Section 1 of this report.

OLD SERVERS

During May and June 2009 we created an initial physical inventory of over forty servers more than four years old in the data center and in an ancillary machine room (Philosophy Hall). (A summary of selected servers is provided in Table 2-1, the complete inventory is included in Appendix A). Inventory work continued beyond the original project scope, however, to advance an aggressive internal server consolidation project. Using the inventory results, individual server machines were replaced with new high-density blade servers, and the adoption of virtual machines enabled server consolidation. Section 1 provides details about our progress in increasing data center efficiency.

Qty	Year purchased	Make/Model	CPU speed	Max power (W)
10	2001-02	Sun Netra T1	500 MHz	120
20	2002	Sun Fire V100	500 MHz	200
9	2003-04	Sun Fire 280R	900 MHz	560
3	2004	HP DL380g3	2.8 GHz	460

Table 2-1. Original inventory, summary of selected servers

ELECTRICAL ENGINEERING HPC CLUSTER

At the time of inventory, a Dell cluster owned by an Electrical Engineering professor was a tenant in our data center and consisted of 100 dual quad core PowerEdge 1955 2.0GHz compute nodes and two dual quad core PowerEdge 2950 2.66GHz master nodes. It was supported by five 10kVA Liebert GXT UPSes, which were carefully balanced to keep them from shutting down when the cluster was 100% in use and drawing maximum power. These servers were purchased in October 2007. The cluster was shut down in January 2013 when it had passed its useful life. Appendix A provides the detailed inventory.

OLD ASTRONOMY HPC CLUSTER

An old Astronomy computing cluster, called Beehive, was a 16-core Linux cluster that was built in 2005 by the Astronomy department at Columbia University (but during our measurements only 14 cores were operational). At the time of inventory Beehive consisted of a master server, file server, and eight compute nodes. Each compute node had dual-core AMD Opteron CPUs rated at 2.19 GHz. Two nodes had 8GB of RAM while the remaining six had 2 GB of RAM. The NFS file server supported 10 TB of SATA-attached storage. The detailed inventory is provided in Appendix A.

Beehive's servers ran GNU Linux and supported research applications written in C, IDL, and FORTRAN. Cluster scheduling was managed by the Open Portable Batch System (OpenPBS: http://www.openpbs.org). This cluster resided in the Physics department server room in an academic building on campus.

NEW SHARED HPC CLUSTER

At the time of inventory, Hotfoot was a 256-core Linux cluster, built in 2009 by CUIT. It consisted of two head nodes (HP DL360 servers), one NFS server (HP DL360) managing 30 TB of SATA storage, and 16 blades (HP BL2x220c G5) each holding two servers with dual quad-core Intel Xeon CPUs rated at 2.66 GHz. All of the hardware resided in the Columbia University computer center in one full-size rack. Appendix A provides the detailed inventory.

Hotfoot's servers ran Red Hat Enterprise Linux and a suite of applications for research use, including Matlab, R, and IDL. Cluster scheduling was managed by Condor software (http://www.cs.wisc.edu/condor).

3. INSTRUMENT SERVER POWER CONSUMPTION

INTRODUCTION

Task 3 called for installing power monitoring instrumentation in the data center and measuring baseline power consumption for each group of machines in the Task 2 inventory.

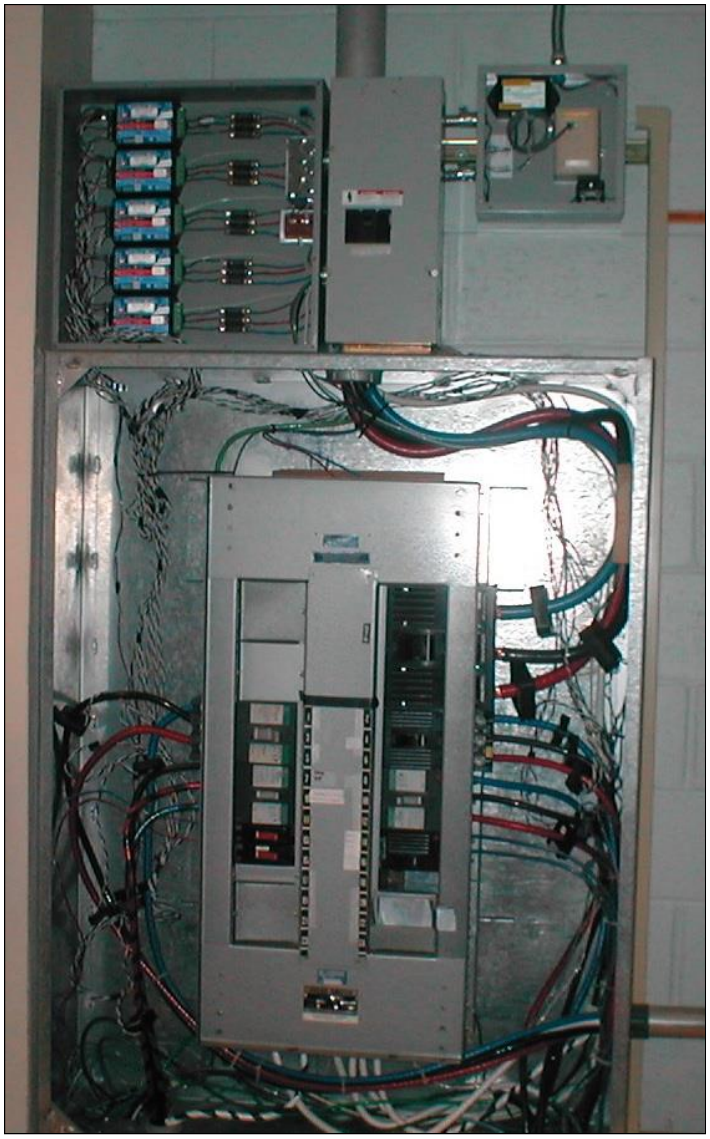


Figure 3-1. Wattnode power meters (above) wired to branch circuit current transducers (CT)

INSTRUMENTATION

We evaluated several options and chose to install Wattnode power meters throughout the data center as well as in the mechanical room to meter our power panels (Figure 3-1). To meter our inventoried server machines, we installed either Raritan power distribution units (PDUs) or used the power measurement features of some of our individual Liebert uninterruptible power supplies (UPS).

Power Panel Metering

We installed Wattnode meters in 20 power panels, including 17 panels in the data center and 3 main feeder panels in the mechanical room directly below the data center: automatic transfer switches (ATS) 2 and 3 which carry the HVAC loads and ATS 4 which carries the IT load. See Figure 3-2.This enabled us to track the data center IT load by measuring all of the main feeder panels inside the data center or by taking the sum of ATS 4 and power panels (PP) 26 and 27 (which are fed from a different source). We were also able to track the data center HVAC load by summing the load of ATS 2 and 3.

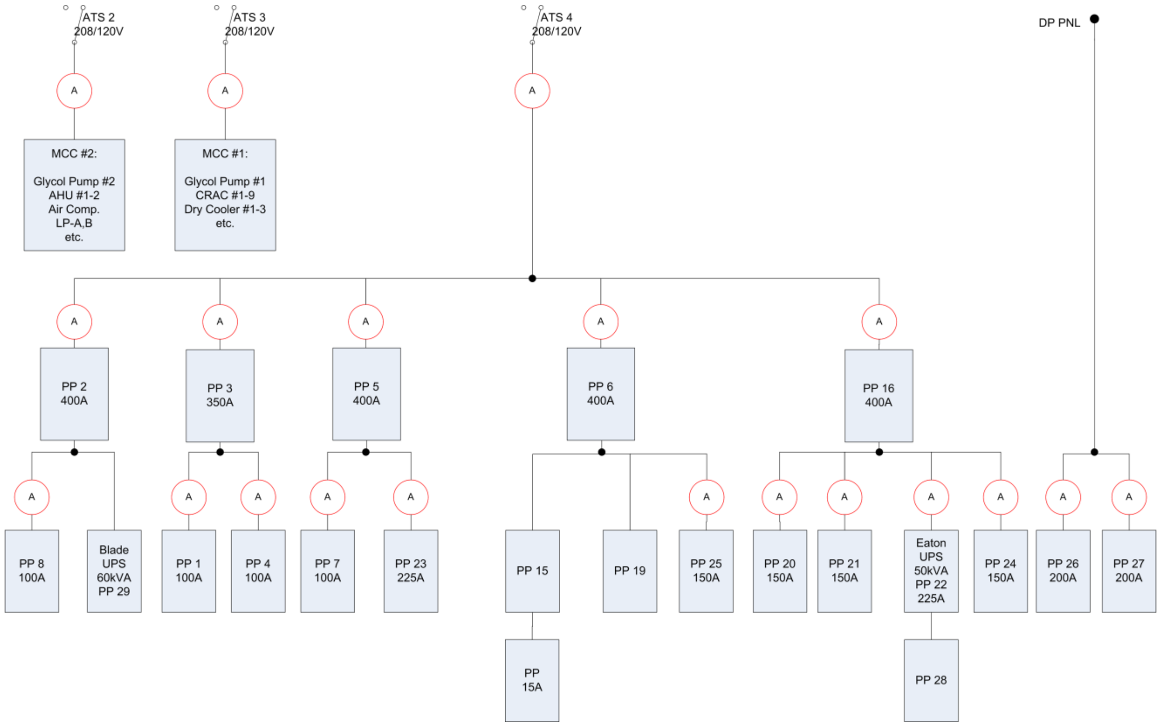


Figure 3-2. Electrical distribution showing locations of Wattnode meters

While digital energy/power measurements on the Facilities side often use the ModBus protocol for data transmission, IT server monitoring often uses the Simple Network Management Protocol (SNMP). As we wanted to correlate IT server and facilities infrastructure monitoring, we used Babel Buster SPX devices, which translate from the ModBus protocol to SNMP. Use of SNMP allowed us to easily integrate data collection with our existing IT monitoring infrastructure.

Server Level Metering

The Sun hardware (including models NetraT1, V100, V210, V240, 280R, V880, T2000) and HP hardware (including models DL360G4p, DL360G5p, DL380G5) from the extended inventory of machines included in Appendix A, were plugged in to Raritan power distribution units (PDUs) to enable individual server power supply load monitoring via SNMP. See Figure 3-3.

Figure 3-3. Raritan SNMP-monitored PDUs

About 30 servers were identified to establish idle and at-load power usage as well as changes in power usage after equipment upgrades (See Section 7).

The newer blade chassis (HP c7000) and blade servers (HP BL460c) were metered using built-in instrumentation.

MEASUREMENT DATA COLLECTION

SNMP data from all power meters was polled at 5 minute intervals by two existing CUIT systems that are used for several purposes: Nagios and Cricket. Nagios (nagios.org) is open source software used to monitor IT infrastructure operation. Cricket (cricket.sourceforge.net), originally developed to monitor network traffic, is a general system that can be used for monitoring trends in time-series data.

We discovered that querying power meter measurements in our Nagios system impacted the performance of the operational monitoring system role of this software. Because of this, we created an external MySQL measurement database that was separate from our Nagios server monitoring systems, and we performed regular imports of real-time data collected in the Nagios system into this external database.

BASELINE POWER CONSUMPTION

The installation of networked power meters allowed us to make initial, baseline power consumption measurements for each group of machines in the Task 2 inventory.

Old Servers

Power consumption measurements for the old servers in the Task 2 inventory are discussed in Section 7.

Electrical Engineering HPC Cluster

At the time of measurement, the Dell cluster accounted for over 12% (36kW) of the total IT load (290kW) in our data center, and drew 20kW (7%) even when idle.

Old and New Shared HPC Clusters

When idle, the old astronomy cluster, Beehive, drew 2.7 kW and the new shared Astronomy/Statistics cluster, Hotfoot, a significantly more powerful system, drew 4.2 kW. See Section 8 for more details.

4. INSTRUMENT SERVER INPUT AIR TEMPERATURE AND OVERALL DATA CENTER CHILLED WATER HEAT LOAD

INTRODUCTION

Server air intake temperature was measured at server rack cabinets. The chilled water supply and return was also measured in order to track the heat load of the data center. We installed chilled water flow meters and measured an overall data center heat load of approximately 120 tons.

INSTRUMENTATION

Server Intake Air Temperature

Server air intake temperature was measured by Raritan temperature sensors, which are installed at the front of the server rack cabinets and connected to the Raritan PDU's environmental sensor port to allow for SNMP data collection. This enhanced our other server monitoring efforts.

Chilled Water Heat Load

We installed Flexim Fluxus ADM 7407 chilled water meters in the data center's mechanical room (located directly below the server room).

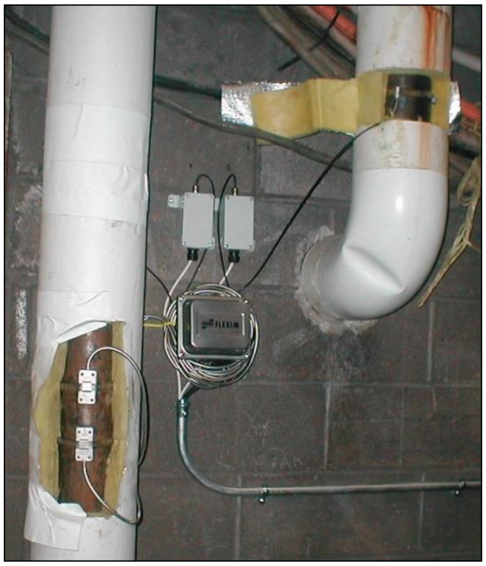


Figure 4-1. Flexim chilled water meter

The sensors associated with the meters measured flow rate and temperature. Using the delta-T and flow rate, we arrived at the BTUs or tons of heat transferred through the various branches of the chilled water distribution network.

The sensors were installed in three locations:

1. At the heat exchanger between the primary campus chilled water loop and the secondary chilled water loop feeding the Liebert computer room air conditioning (CRAC) units within the data center,

2. On another connection to the campus chilled water loop feeding the comfort cooling air handling units (AHU) that provide overhead cooling within the data center, and

3. On the chilled water loop that feeds the data center's rooftop dry coolers (the backup cooling system).

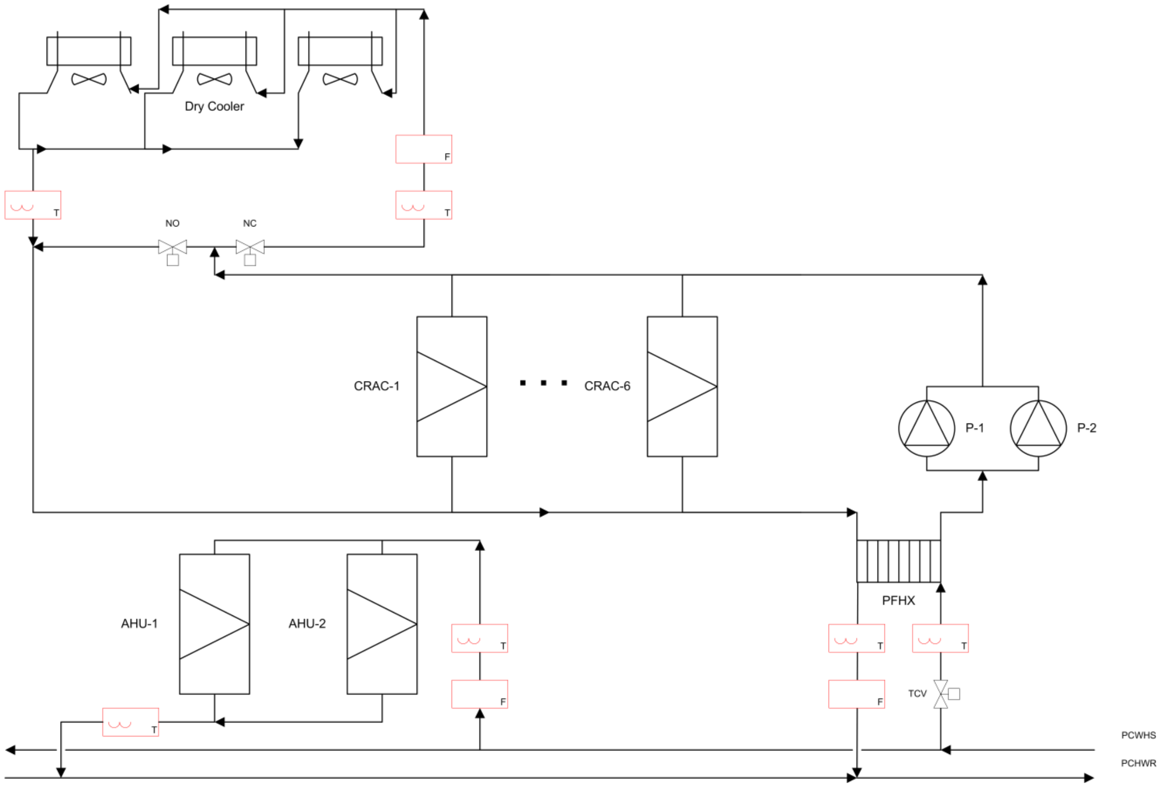


Figure 4-2. Chilled water distribution showing locations of added Flexim meters

All Flexim chilled water meters were tied into the same Modbus network, polled by the existing Nagios system, in the same fashion as the Wattnode power panel meters.

5. ESTABLISH OVERALL DATA CENTER PROFILE

INTRODUCTION

The instrumentation of the data center in this project, described in Sections 3 and 4, enabled us to measure and calculate the data center Power Usage Effectiveness (PUE) in real time. We also used standard data center analysis software supplied by the Department of Energy (DOE). Our measurements confirmed the existence of seasonal variations in PUE due to the data center mechanical systems which include the use of air- and water-side economizing.

CALCULATING PUE

PUE is defined as the ratio of total facility power usage to IT equipment power usage. Facility power includes cooling, lighting, IT equipment, and any other data center overhead. IT equipment power is limited to servers, storage devices, and other components that are directly involved with IT services. The lowest and asymptotically best possible value for PUE is 1.0, which represents a situation where all energy entering a data center is used for powering IT equipment.

Facility and IT power usage data were collected using the metering equipment and Nagios software described in Section 3. Nagios polled the metering equipment for power usage and thermal data at five-minute intervals, and wrote these values to a MySQL database. After collecting data for twelve months (May 2010 to May 2011), we were able to calculate an annualized measure of PUE that spans all four seasons.

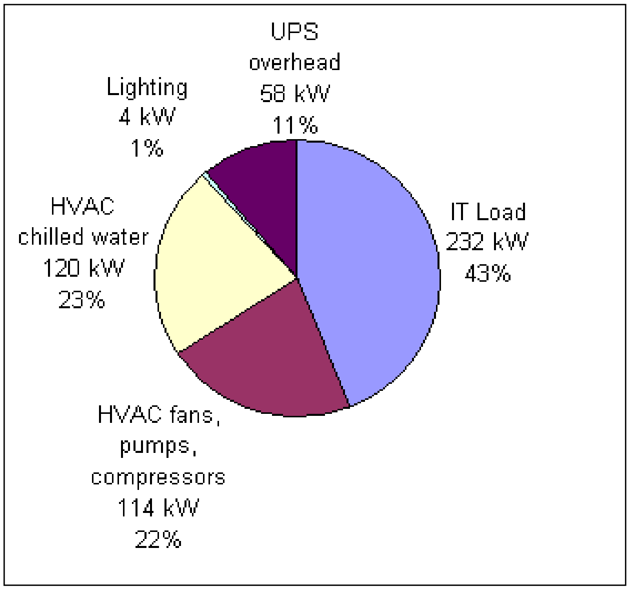


Figure 5-1. Summer 2010 PUE=2.27

RESULTS

We have calculated the data center's PUE at various points in time, as well as an annualized average. Several assumptions had to be made beyond what was directly measured: First, the energy cost of the central chilled water plant was assumed to be 1 kW per ton, based on averages provided by the Facilities department. Second, UPS efficiency was estimated at 82%, based on sampled power measurements up and down-stream of several UPSes. Based on these estimates and directly measured data, we estimated an annualized 2010 average PUE of 2.06 and a median PUE of 2.1.

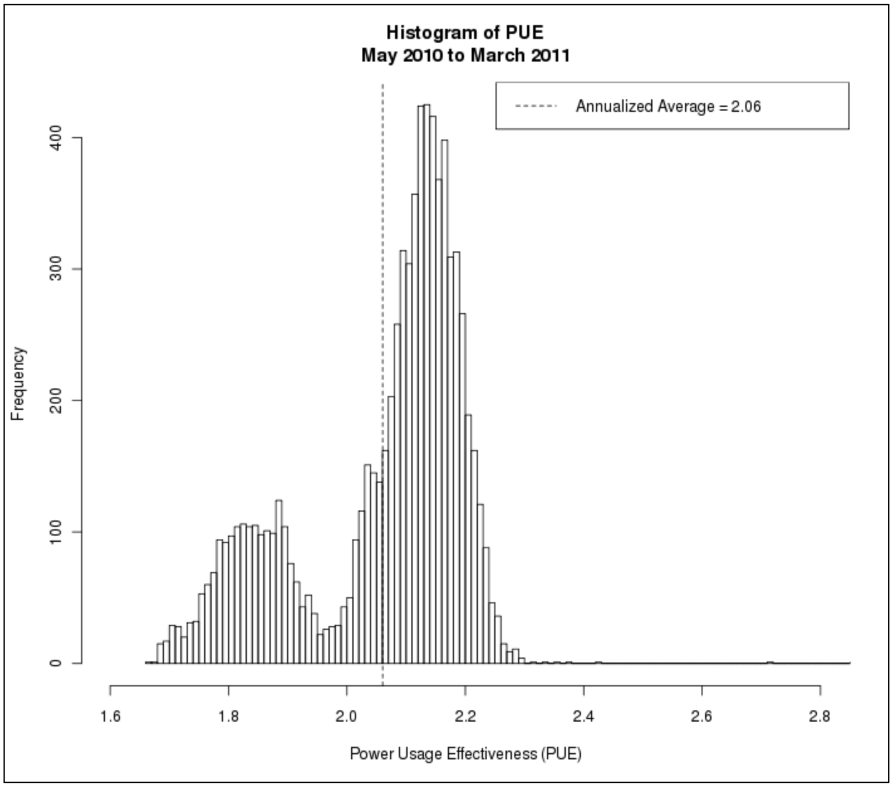


Figure 5-2. PUE distribution

During a summer observation, we measured an IT load of 232 kW (290 kW measured at power panels de-rated by estimated UPS efficiency of 0.82). As shown in Figure 5-1, the total power usage was 528kW (232 kW IT load + 58 kW UPS overhead + 4 kW lighting + 120 kW estimated central chilled water load + 114 kW HVAC electrical load). The PUE was calculated as 2.27 = 528 kW/ 232 kW. In contrast, we measured total power usage during the winter at around 425 kW with a winter PUE of 1.83.

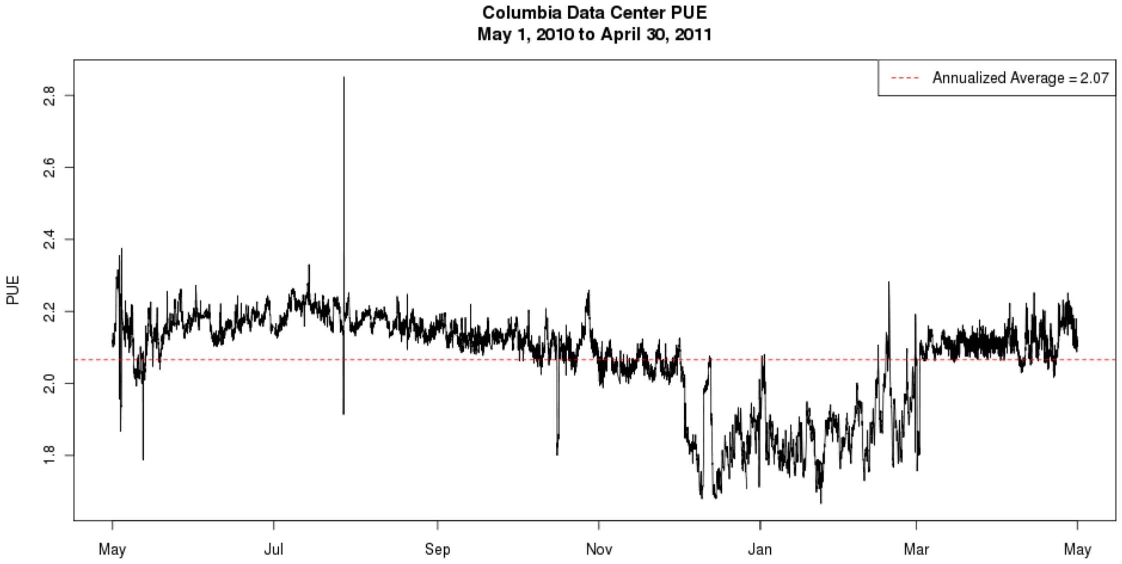


Figure 5-3. Change in PUE over time at 5-minute intervals

As expected, the PUE varies significantly by season. Figure 5-2 presents a histogram of PUE values created hourly from May 2010 through April 2011. The bimodal distribution of the hourly PUE values shows clustering of the winter peak value around 1.85 and the summer peak value around 2.17. Figure 5-3 graphs PUE measured at 5-minute intervals over time for a one-year period. Figure 5-4 details the spike in PUE experienced on July 27, 2010 when a campus chilled water outage occurred. As can be seen, both the cooling and IT loads increased, by approximately 150 kW and 8 kW, respectively. The cooling load increase can be attributed to the CRAC compressors and dry cooler fans operating to compensate for the lost central chilled water. The IT load increase is illustrative of a problem with accurately calculating PUE: As the temperature in the data center increased, the variable-speed server cooling fans sped up and increased their energy consumption. In an ideal measurement scenario, the fan and power supply energy consumption of the IT servers would be attributed to facility infrastructure overhead, not IT load. However, these loads are “built in” to the servers and are lumped in with the “useful work” energy load of the server CPU, memory, disks and so on. This is one of many reasons why we feel the importance of PUE minimization is overemphasized by the industry; what we really want to minimize is overall energy required per unit of “useful work.” This is addressed in sections 7 through 9 of this report.

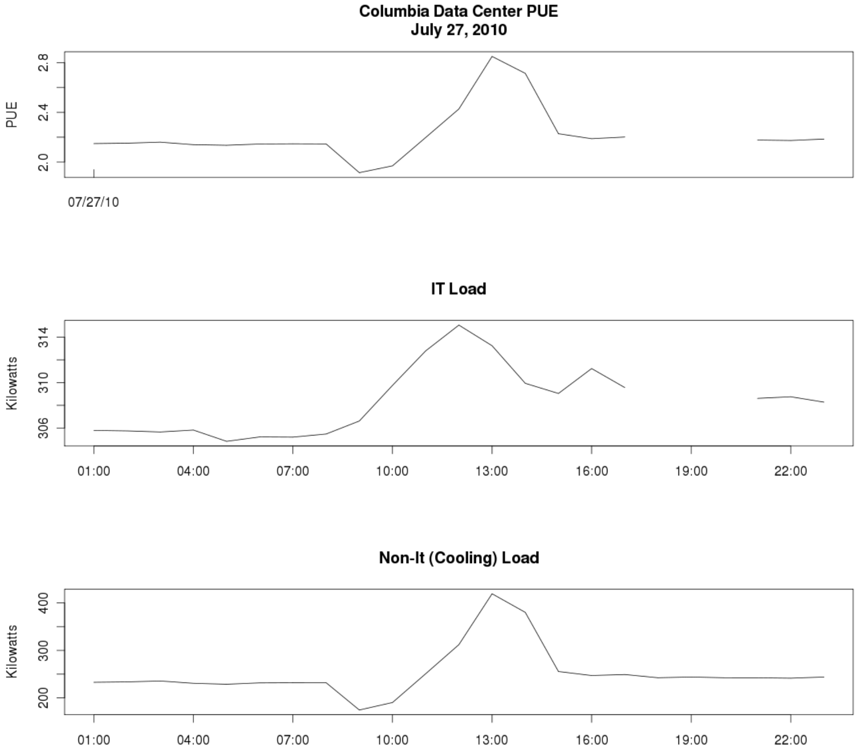


Figure 5-4. July 27, 2010 central chilled water outage

DOE DC PRO SOFTWARE

In addition to using our own approach to measuring and calcultating PUE, we used the Data Center Energy Profiler (DC Pro) software (http://dcpro.ppc.com/) supplied by the DOE to help identify how energy is used in Columbia's data center and to identify potential energy and cost savings. A DC Pro 2.0 report (Appendix B) calculates our *Source PUE* at 2.2 which is a close match to our computed PUE of around 2.1. We had some trouble understanding the basis of the calculations made by this software, including the distinction between *Site PUE* and *Source PUE*. Our calculated PUE is close to the Source PUE identified in the DC Pro report but is nowhere near the Site PUE. In general, the report was not entirely useful in terms of establishing meaningful targets. The rather generic suggested next steps, which of course do not come with quantifiable expected energy savings, are a good guide and essentially summarize recommendations available elsewhere such as in the ASHRAE Best Practices book. Key recommendations from that report (pages 5-11) include a number of improvements to the categories of Air Management, Cooling, Environmental Conditions, Global, IT Equipment, IT Equipment Power Chain, and Lighting, several of which we have already implemented or plan to pursue as part of our long-term data center improvement strategy that has been informed by this study.

6. INVESTIGATE ALTERNATIVES FOR HVAC EFFICIENCY IMPROVEMENTS

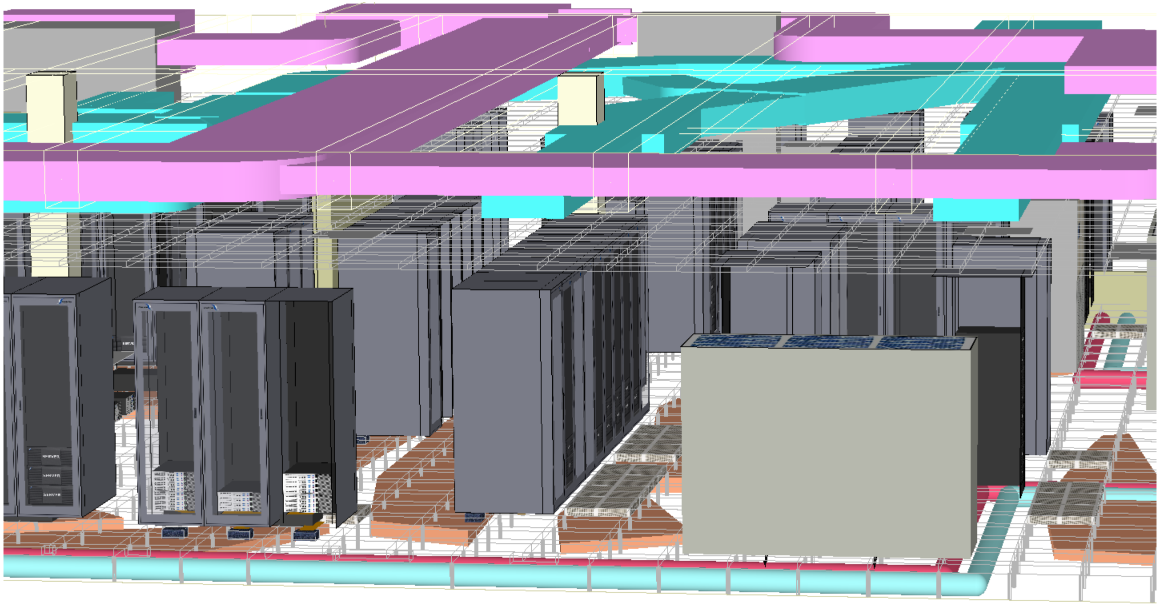


Figure 6-1. CFD baseline model - isometric

INTRODUCTION

The original goal of this task was to experiment with one specific approach to improving the cooling infrastructure efficiency, to complement the approach taken in tasks 7, 8 and 9, which each attempt to improve efficiency of the IT equipment. After extensive planning studies, we determined that the initially proposed approach, of implementing nine racks of high power density in-row cooling, was not feasible within the estimated budget. We requested and were granted a change to project scope to investigate several options, including the original in-row cooling plan, and to implement one or more of those options, if feasible within the project budget.

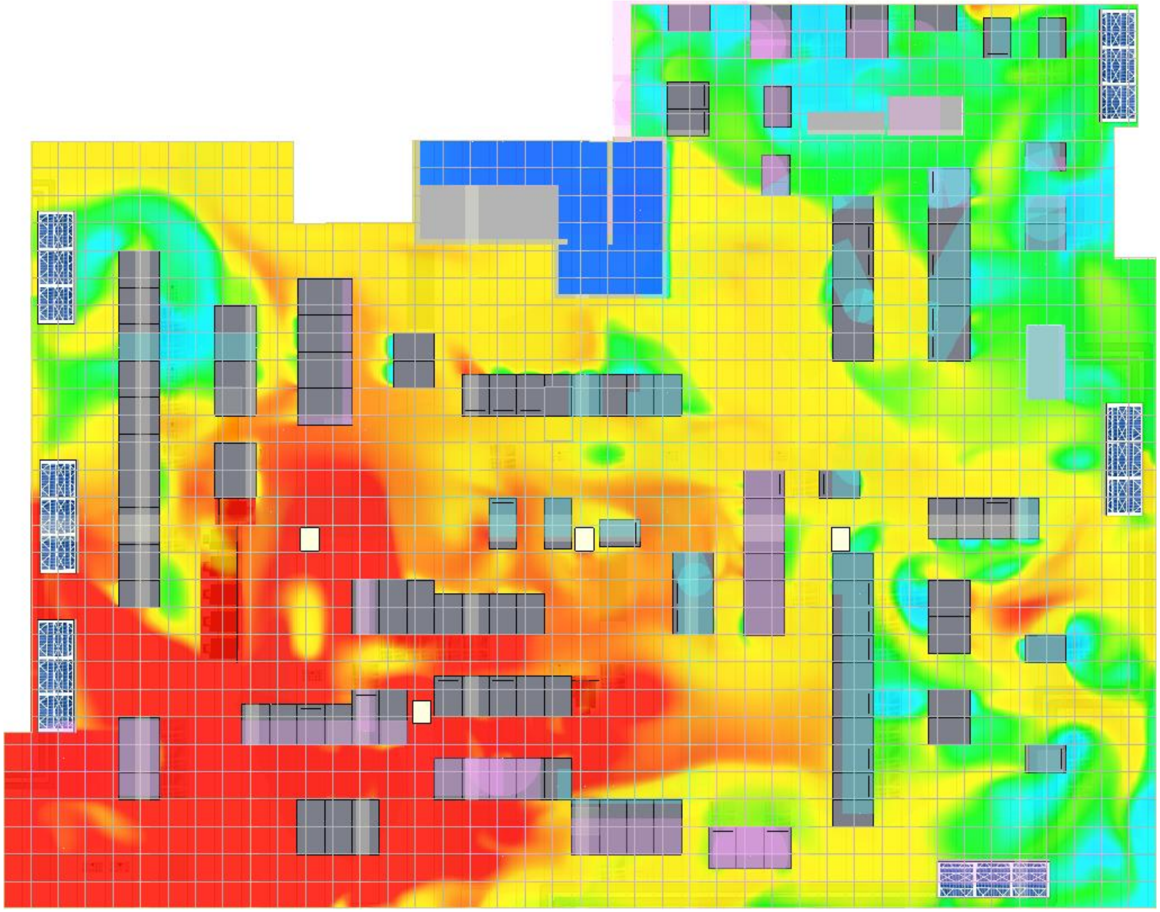


Figure 6-2. CFD baseline model temperature profile - plan

RESULTS

The project had proposed experimenting with a small pilot deployment of in-row/rack cooling technology for high power density equipment. The goal was to test typical industry claims that this technology is 30% more energy-efficient than traditional data center cooling approaches. Columbia funded detailed RFPs, engineering studies and peer reviews to determine the costs and procedures needed to retrofit this technology into our data center. Approximately $109,000 toward these studies was funded as part of Columbia's cost-share of the Agreement. An NIH-funded Core Research Computing Facility (CRCF) construction grant has further developed full Schematic Design, Design Development, and 100% Construction Documents for the electrical, facility UPS and cooling needs for that equipment, giving us a much better understanding of the challenges we faced in our 50-year old facility. The net result of the engineering studies is that our original estimated budget for the in-row/rack cooling technology task was significantly lower than the required costs of the retrofit, due, in large part, to the unusual bimodal operation of our cooling plant and the lack of available in-row cooling products that will operate over both a typical 45° F and unusual 100° F cooling water temperature range (when dry coolers and CRAC compressors are operating).

After revising the scope, we proceeded to explore other potential cooling improvements, starting with CFD models (baseline and three iterations: reconfigured hot/cold aisles, CRAC ducting to the ceiling plenum, and cold aisle containment) which led to several key recommendations:

- Clearing underfloor of all cabling will result in the most dramatic improvements to air flow.
- Ducting CRAC return air from the overhead ceiling plenum will improve hot/cold aisle separation and minimize recirculation in the cabinets.
- Vinyl cold aisle containment will help but less so than earlier recommendations.

Based on these recommendations and available budget, we prioritized creating proper hot and cold aisles and clearing out electrical and network cables from under the floor. The project has installed a Starline overhead electrical distribution bus system which has been connected to the new NIH grant-funded power distribution units that will come online in mid-2013. At this point we can begin the difficult task of removing underfloor power cables and continue removing network cables. As such, this task has not resulted in specifically measurable results yet, but we are confident it is moving the facility toward greater efficiency and reliability.

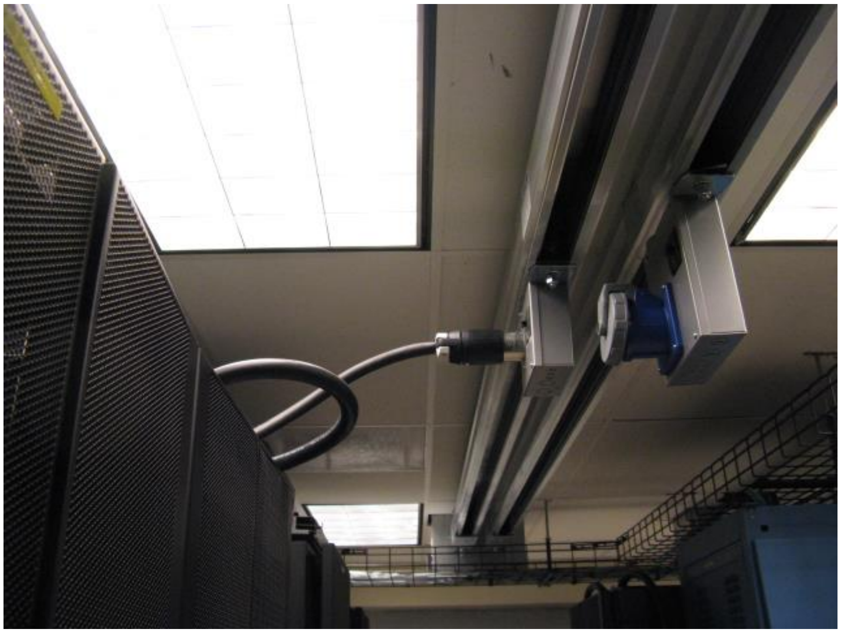


Figure 6-3. Overhead electrical bus

7. REPLACE OLD SERVERS AND MEASURE EFFICIENCY IMPROVEMENT

INTRODUCTION

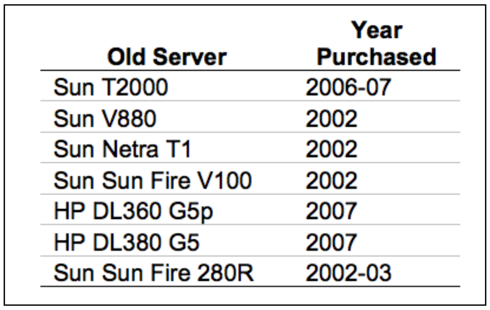
The purpose of this task was to replace old servers inventoried in Task 2 with newer hardware and to investigate changes in power consumption.

The ultimate goal was to measure the power consumed by running a selection of IT services on old hardware and compare it to the power consumed by running the same IT services on newer hardware. We collected data from eight identified IT services to help us understand the relationship between hardware changes and power consumption and to guide our plans for server replacement.

REPLACING OLD SERVERS

Identification of Old Servers

As noted in Section 2, our data center contained over forty servers that were purchased between 2001 and 2004. Although all of these servers had scheduled retirement dates, not all were part of the Task 7 replace-and-measure plan. Instead, we selected a subset of older servers that were marked for upgrade to new or newer hardware. The cohort was chosen to be representative of a variety of planned hardware changes so that extrapolation of data to unmeasured hosts would be possible. A summary of the old servers is provided in Table 7-1.



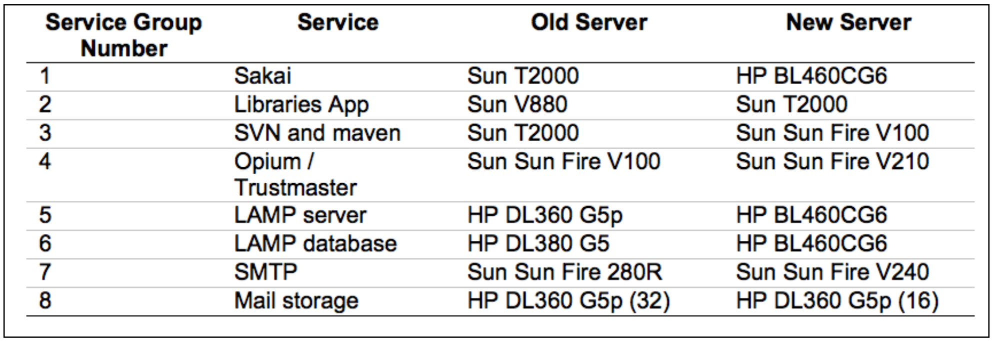
Old Server	Year Purchased
Sun T2000	2006-07
Sun V880	2002
Sun Netra T1	2002
Sun Sun Fire V100	2002
HP DL360 G5p	2007
HP DL380 G5	2007
Sun Sun Fire 280R	2002-03

Table 7-1. Task 7 old servers and purchase dates

Consolidation of Replacement Servers

All old servers were not replaced with new high-density blades. Rather, a subset of old servers received blade replacements, and the remainder were replaced with relatively newer servers that had previously been used for other purposes.

MEASUREMENT STRATEGY



Service Group Number	Service	Old Server	New Server
1	Sakai	Sun T2000	HP BL460CG6
2	Libraries App	Sun V880	Sun T2000
3	SVN and maven	Sun T2000	Sun Sun Fire V100
4	Opium / Trustmaster	Sun Sun Fire V100	Sun Sun Fire V210
5	LAMP server	HP DL360 G5p	HP BL460CG6
6	LAMP database	HP DL380 G5	HP BL460CG6
7	SMTP	Sun Sun Fire 280R	Sun Sun Fire V240
8	Mail storage	HP DL360 G5p (32)	HP DL360 G5p (16)

Table 7-2. Representative IT service groups

We implemented a service-oriented approach to measuring power consumption and began by identifying eight different IT services scheduled to move from older hardware to newer servers. Examples of these IT services included email storage, online collaboration tools (Sakai), and Linux-Apache-MySQL-PHP (LAMP) application development and production environment hosting. These eight services, listed in Table 7-2, were representative of the large number of services running on machines in the data center. Each will be addressed in the following sections.

We measured power consumption while servers were at various load levels, from idle to maximum. Measurements were structured so that we could compare power consumption of old and new hardware types independent of the service, but also could compare the power consumption when running the services.

The first measurement strategy employed the Standard Performance Evaluation Corporation's (SPEC) SPECpower_ssj2008 benchmark (http://www.spec.org/power_ssj2008/). This benchmark is an industry-standard that uses a Java program to put machines at various load states while simultaneously monitoring power consumption. It provides a hardware-dependent (and service-independent) way of comparing two computer systems using a DCeP performance metric called *server-side java operations per second per watt* (ssj_ops/W). The benchmark runs for 74 minutes and starts with a calibration phase before loading the server at 100% load and then, every four minutes, reducing by 10% increments down to no load, ending in an active idle state. It should be stressed that we used the benchmark software out-of-the-box, and did not introduce any Java tuning. In addition, power measurements were taken with the infrastructure described in Task 3 of this report. These devices are not on the list of approved measurement devices for valid SPECpower benchmarks but are functionally equivalent. Therefore, *our benchmarks are not valid for external comparisons with published SPEC benchmarks* and may only be used for our internal relative server comparison purposes.

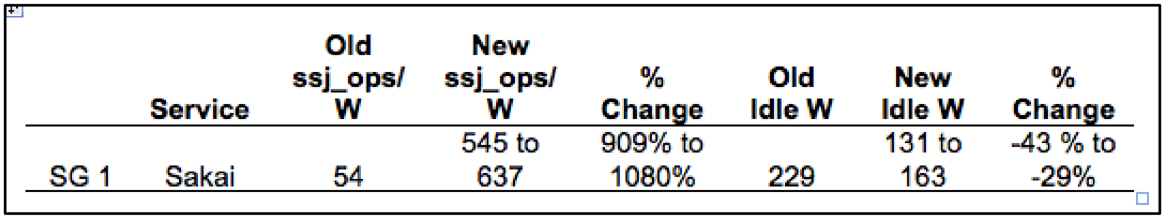
The second measurement strategy measured power consumption during a typical week of the year while the server was hosting its particular services. This performance per watt metric provides a realistic view of how much power a server draws to perform its function over a generic time period but is subject to the vagaries of actual user demand, leading to results that are not as predictable or reproducible as the SPECpower benchmark.

SERVICE GROUP MEASUREMENTS

For each Service Group (SG) in Table 7-2, we measured a typical week's activities, the SPECpower ssj_ops/W, and idle power usage both on the older hardware and the new hardware. Some hardware changes resulted in power savings (service groups 2, 3, 7, 8) because the service was moved from less efficient to more efficient hardware. One service group (4), however, resulted in an increase in power consumption since the newer hardware consumed more power than the older hardware and the application load was minimal. For the following discussion, items measured in Watts (power consumption) should decrease from old to new hardware and items measured in ssj_ops/W (efficiency) should increase if the hardware change was beneficial.

Service Group 1

SG 1 considered the migration of the Sakai service from Sun T2000 servers to HP BL460 G6 blades. Sakai is a collaboration tool used for course and research group management. The service was scheduled to be moved to blades, but the timeline was such that the migration occurred outside the timeline of this phase of the project. However, we were able to run the SPECpower benchmark on each host, which enabled a partial view of how the migration would affect power consumption. Table 7-3 summarizes the results.



	Service	Old ssj_ops/ W	New ssj_ops/ W	% Change	Old Idle W	New Idle W	% Change
SG 1	Sakai	54	545 to 637	909% to 1080%	229	131 to 163	-43 % to -29%

Table 7-3. Service Group 1 power measurement results

As seen in Table 7-3, the Sakai application was slated to move from a server that achieves 54 ssj_ops/W to one that achieves at least 545 – a marked increase in performance per watt[5]. In addition, idle power consumption of the blade was lower than the T2000, so that just having the blade plugged in instead of the T2000 would reduce the power draw by about 66 Watts (229 W – 163 W = 66 W).

[5] All figures for blade servers are presented as a range. Each blade server sat in a chassis. The blade consumed its own power which we measured directly, but the chassis also consumed power that was shared among the 16 blades. Since we could only measure the chassis power consumption as a whole, we expressed blade power consumption as a range.

Service Group 2

SG 2 considered a library management system application that migrated from a Sun V880 to a Sun T2000. Table 7-4 summarizes the power changes from this move.

	Service	Typical Week Old W	Typical Week New W	% Change	Old ssj_ops/ W	New ssj_ops/ W	% Change	Old Idle W	New Idle W	% Change
SG 2	Libraries App	1674	539	-68%	12	54	350%	1218	229	-81%

Table 7-4. Service Group 2 power measurement results

For SG 2, we have both the SPECpower measurement as well as measurement from a typical week under load. The Sun V880 was one of the most power consuming servers in the data center, drawing an average of 1.6 kW over a week. Replacing it with a Sun T2000 resulted in a 68% reduction in power consumption – 81% when idle. The SPECpower results confirmed that the T2000 is generally more efficient that the V880 as well. Figure 7-1 presents a plot of the SPECpower results for the V880 and Figure 7-2 shows the plot for the T2000. Both plots show that server power consumption was basically constant regardless of the system load, but the T2000 saw a slight decrease in power usage over the course of the benchmark.

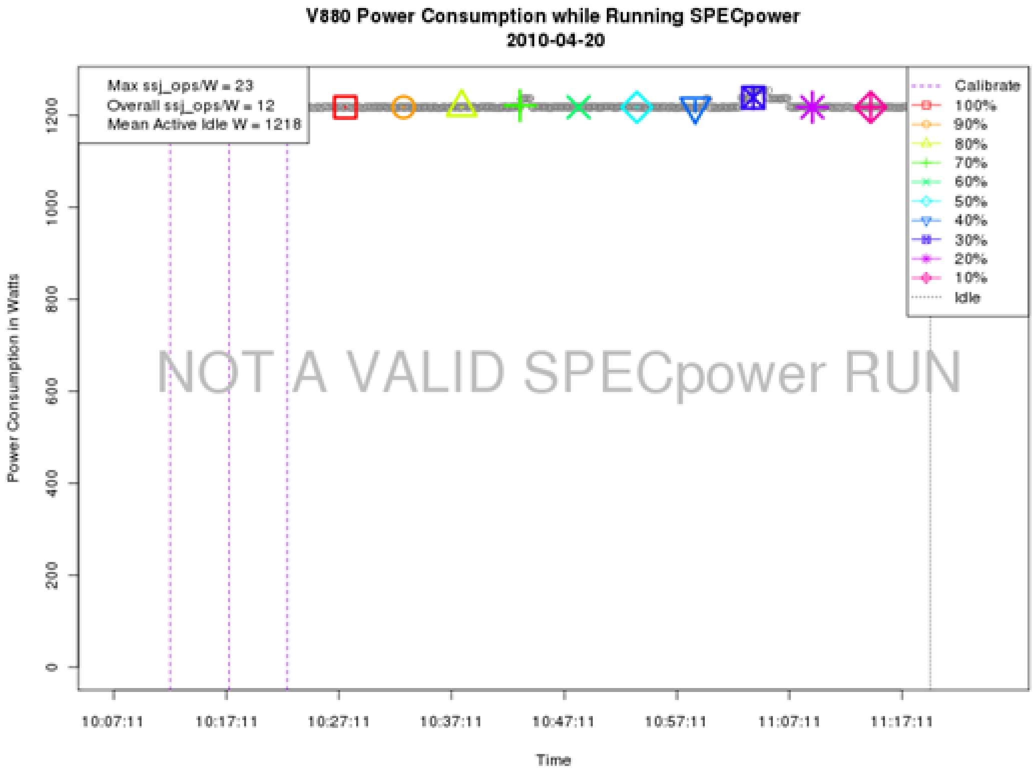


Figure 7-1. SPECpower benchmark for Sun V880

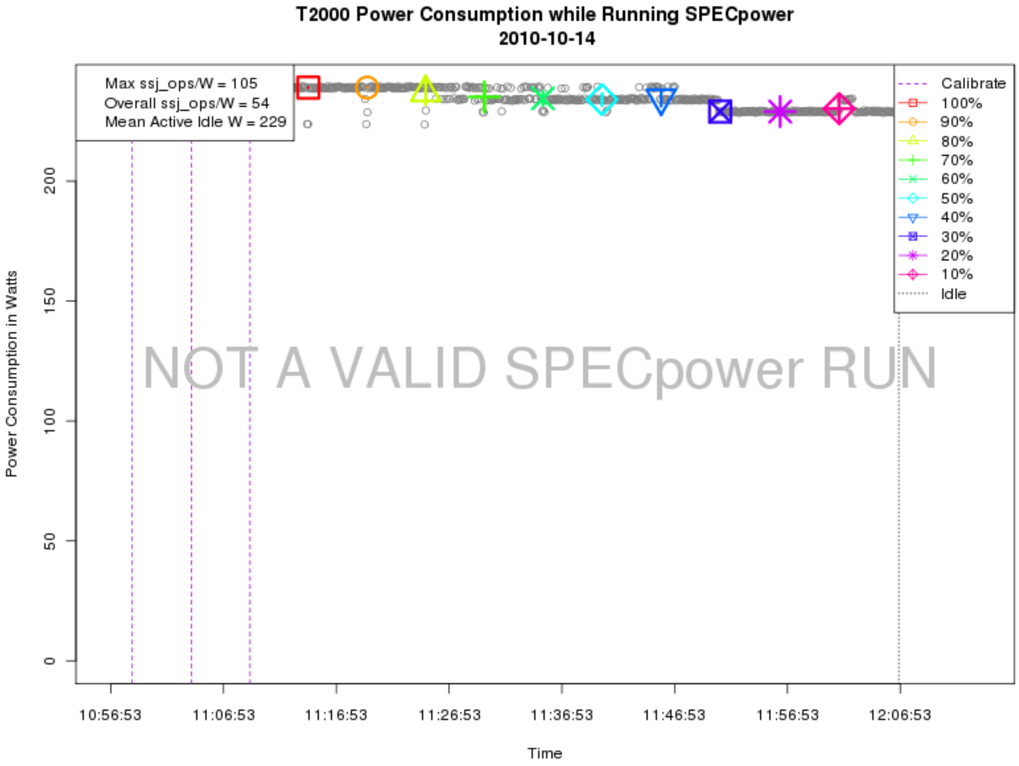


Figure 7-2. SPECpower benchmark for Sun T2000

Service Group 3

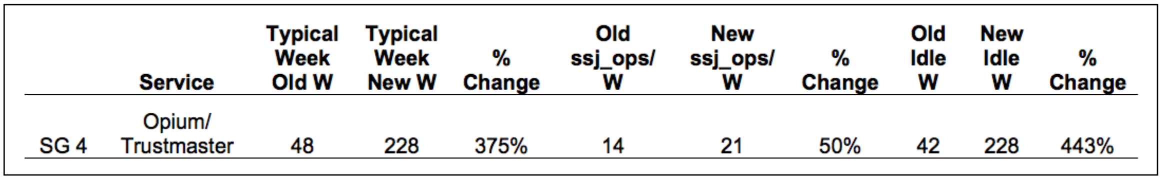
SG 3 involved the migration of a Sun Netra T1 to a Sun Sun Fire V100. The servers ran Subversion and Maven, software for version control and project management. Table 7-5 presents a summary of how the server change affected power consumption. As the V100 was not much newer than the Netra T1, it is not surprising that the change in power consumption and efficiency was somewhat small. However, from an overall energy perspective, the migration resulted in lower consumption to run this service.

		Typical	Typical		Old	New		Old	New	
		Week	Week	%	ssj_ops/	ssj_ops/	%	Idle	Idle	%
	Service	Old W	New W	Change	W	W	Change	W	W	Change
SG 3	SVN/Maven	65	42	-35%	12	14	17%	57	42	-26%

Table 7-5. Service Group 3 power measurement results

Service Group 4

SG 4 considered the Opium/Trustmaster service that was migrated from a Sun Fire V100 to a Sun Fire V210. This was another example of old hardware that was upgraded to relatively newer hardware, but this newer hardware was still quite old. Table 7-6 summarizes the resulting power consumption change.



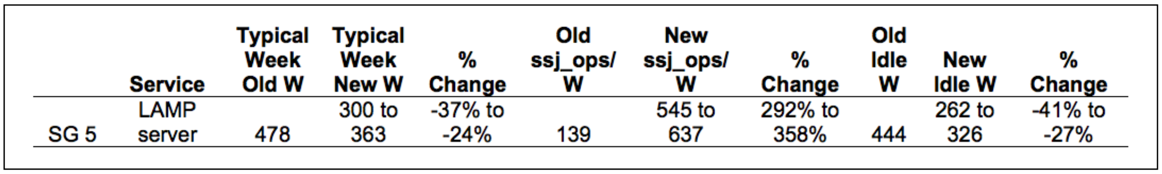
	Service	Typical Week Old W	Typical Week New W	% Change	Old ssj_ops/ W	New ssj_ops/ W	% Change	Old Idle W	New Idle W	% Change
SG 4	Opium/ Trustmaster	48	228	375%	14	21	50%	42	228	443%

Table 7-6. Service Group 4 power measurement results

The move from a V100 to a V210 increased the SPECpower performance per watt measure, but did so at a significant cost. Running the service on the newer hardware used about 375% more power on average. This service group was notable because it showed that newer hardware may not always be the best choice from a power consumption perspective. More specifically, we conclude that the new server is likely significantly oversized for the workload: This is an intermittently-used application , a server management tool, that distributes software and configuration updates to other servers on an infrequent basis and otherwise remains idle.

Service Group 5

SG 5 considered two load-balanced LAMP production hosts that were migrated from HP DL360 G5p standalone servers to HP BL460C G6 high-density blades. Table 7-7 presents the before-and-after power consumption information.



	Service	Typical Week Old W	Typical Week New W	% Change	Old ssj_ops/ W	New ssj_ops/ W	% Change	Old Idle W	New Idle W	% Change
SG 5	LAMP server	478	300 to 363	-37% to -24%	139	545 to 637	292% to 358%	444	262 to 326	-41% to -27%

Table 7-7. Service Group 5 power measurement results

Shifting our LAMP infrastructure to high-density blades resulted in a 24% to 37% reduction in power consumption, on average. The SPECpower benchmark showed the very large increase in efficiency of the blade servers in comparison to the standalone DL360 – a 300% increase in performance per watt of power consumed. We present graphs of a typical week of power usage for the DL360 and BL460. Figure 7-3 shows one of the DL360s that hosted LAMP during a typical week and Figure 7-4 shows one of the BL460 blades that hosted LAMP during a typical week. Both graphs show the cyclical nature of power consumption, as the CPU speeded up and slowed down depending on the load. The DL360 appears to have two peaks and two troughs each day, however, whereas the BL460 has one peak during mid-day and one trough overnight. The patterns could have been calendar differences (February versus September) or differences in the CPU speed-stepping algorithm.

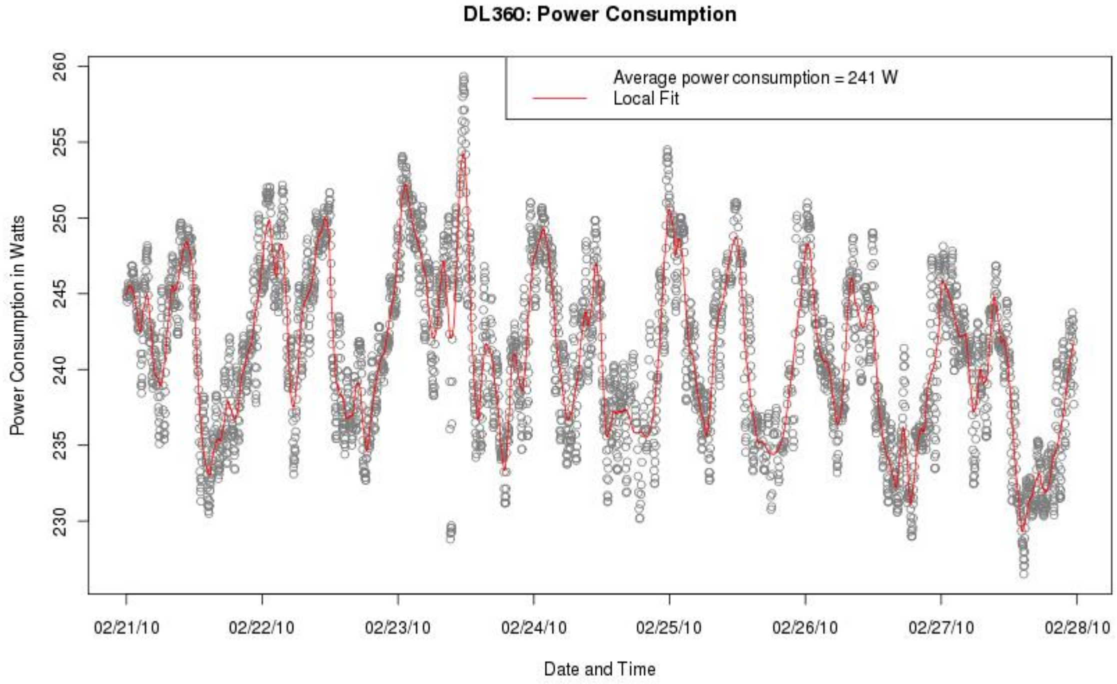


Figure 7-3. Power consumption of HP DL360 LAMP server during a typical week

We also compared the SPECpower results between the DL360 and the BL460. Figure 7-5 shows a smoothed version of the plots for visual clarity. The advantage of the blade server was clear in its ability to reduce power consumption at a faster rate than the DL360 as the load on the server decreased. Even though

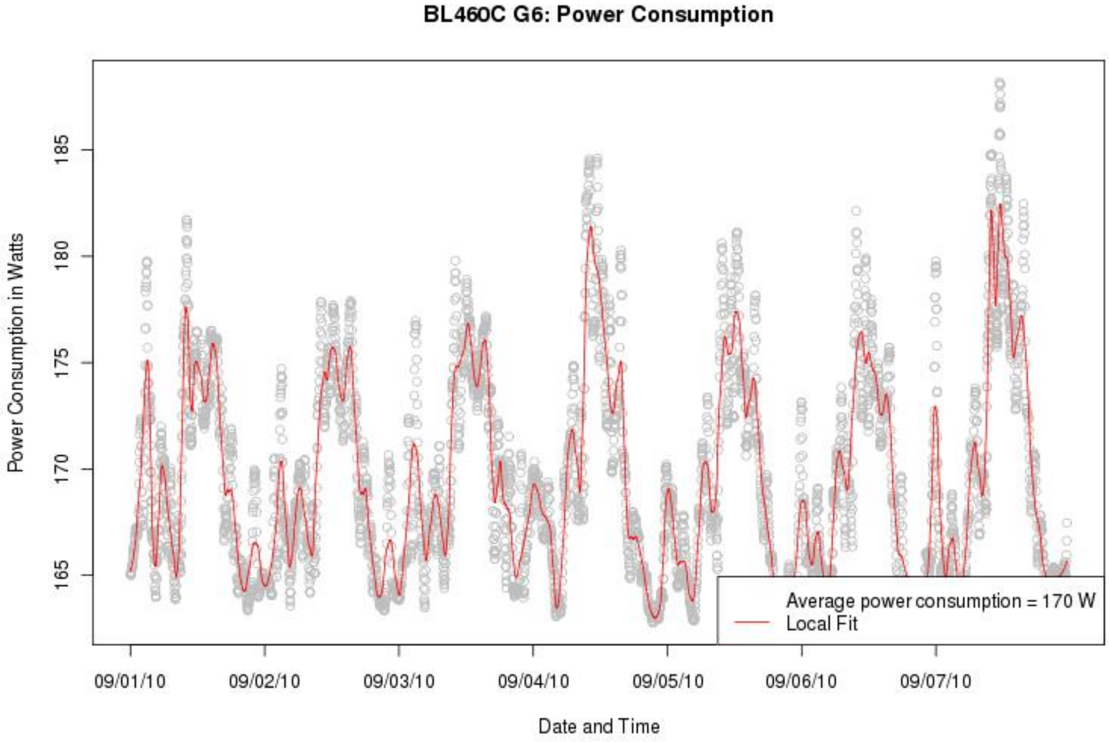
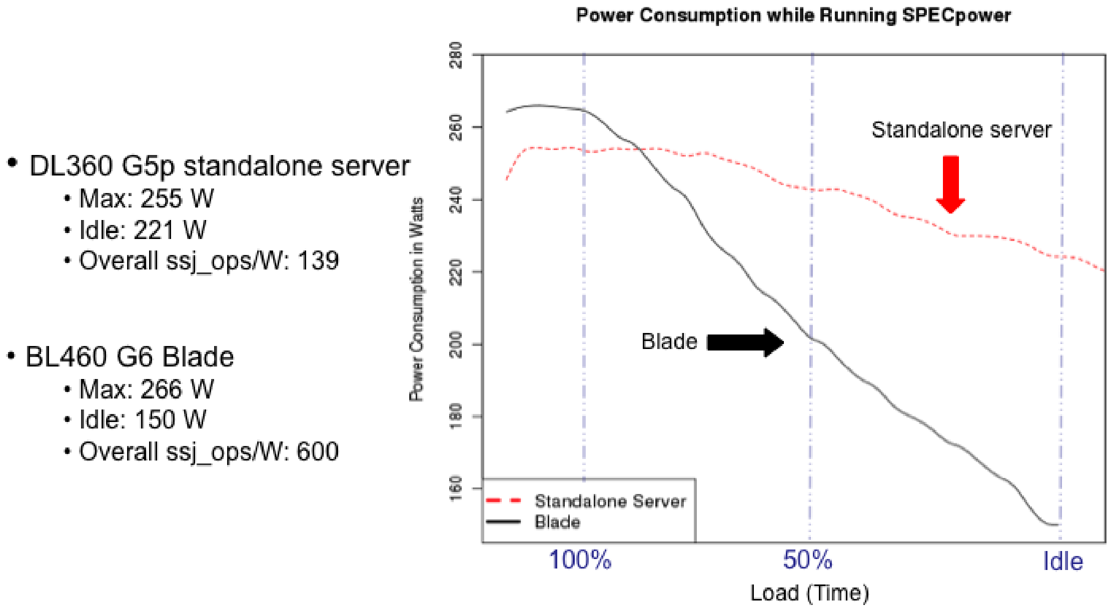


Figure 7-4. Power consumption of HP BL460 LAMP server during a typical week

the blade consumed more power at the peak load (as noted at the leftmost side of the plot), its rapid decrease in consumption made it more desirable than the DL360 – especially for services which often ran below peak load.



- DL360 G5p standalone server
  - Max: 255 W
  - Idle: 221 W
  - Overall ssj_ops/W: 139
- BL460 G6 Blade
  - Max: 266 W
  - Idle: 150 W
  - Overall ssj_ops/W: 600

Figure 7-5. SPECpower comparison of DL360 and BL460

Service Group 6

SG 6 involved the migration of a production LAMP database host from an HP DL380 G5 standalone server to an HP BL460C G6 high-density blade. As shown in Table 7-8, we reduced typical power consumption by 44% to 31% and replaced the LAMP database host with a far more efficient one.

	Service	Typical Week Old W	Typical Week New W	% Change	Old ssj_ops/ W	New ssj_ops/ W	% Change	Old Idle W	New Idle W	% Change
SG 5	LAMP server	478	300 to 363	-37% to -24%	139	545 to 637	292% to 358%	444	262 to 326	-41% to -27%

Table 7-8. Service Group 6 power measurement results

Service Group 7

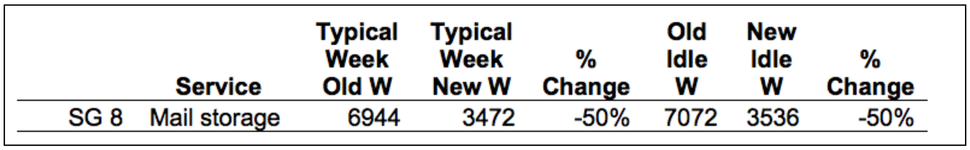
SG 7 followed the migration of Simple Mail Transfer Protocol (SMTP) service from a Sun Sun Fire 280R to a Sun Sun Fire V240. The SPECpower benchmark suggested that the V240 is somewhat more efficient that the 280R. We noted a 32% reduction in power for running SMTP on the relatively newer hardware compared to the older hardware as shown in Table 7-9.

	Service	Typical Week Old W	Typical Week New W	% Change	Old ssj_ops/ W	New ssj_ops/ W	% Change	Old Idle W	New Idle W	% Change
SG 7	SMTP	401	272	-32%	20	22	10%	390	278	-29%

Table 7-9. Service Group 7 power measurement results

Service Group 8

Our last service group example considered a change in the mail storage configuration that reduced the number of servers required to run the service from 32 to 16 HP DL360 G5s. Cutting the number of servers in half resulted in an average power savings of 3.4 kW. This example showed that improvements in power consumption can occur in other ways besides upgrading hardware or moving to high-density blade infrastructure. Table 7-10 summarizes the results for this group.



	Service	Typical Week Old W	Typical Week New W	% Change	Old Idle W	New Idle W	% Change
SG 8	Mail storage	6944	3472	-50%	7072	3536	-50%

Table 7-10. Service Group 8 power measurement results

DISCUSSION OF RESULTS

When reviewing the specific results for Service Groups 2 through 8, it is apparent that the SPECpower benchmark and actual usage improvement percentages do in fact correlate fairly well. As such, we are able to conclude that the SPECpower benchmark is a useful tool for estimating server energy efficiency in actual use.

On average, the hardware changes in Task 7 resulted in a reduction in power consumption and an increase in energy efficiency. We used simple examples that could be extrapolated easily to larger numbers of servers. Task 7 originally intended to replace standalone servers with high-density blades, but we only had a limited blade infrastructure in place at the time of the measurements. However, with Service Groups 1, 5, and 6 we did analyze services migrating to blades, and there were measured improvements in power consumption. All of the blade scenarios involved the movement of only one service. As we upgrade our data center, we have been using virtualization such that a single VMware server typically hosts ten or more services. Simplistically stated, one VMware server can perform the equivalent services of ten or more standalone servers. A challenge with measuring energy consumption of virtual machines (VM) is that they all reside within a single physical server, making it difficult to clearly apportion energy consumption to individual VMs and the application services they support.

8. COMPARE OLD AND NEW HIGH PERFORMANCE COMPUTING CLUSTERS

INTRODUCTION

Task 8 required comparing a legacy computer cluster (Beehive) built by the Astronomy department to a new cluster (Hotfoot) built by CUIT for the Astronomy and Statistics departments. Our analysis confirmed the hypothesis that the new cluster is significantly more energy efficient than the old cluster when performing research computing tasks.

ANALYSIS AND RESULTS

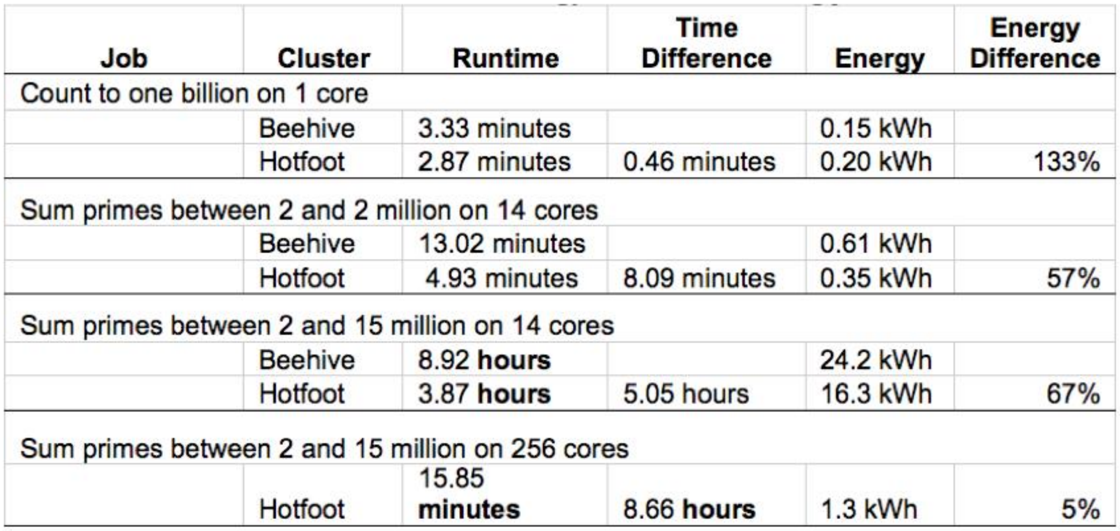
We selected three programs to run that would place the clusters under load while we measured power. These programs were chosen because they were simple in structure and easy to understand. All were written in C and compiled with GNU gcc. The first program counted from 1 to 1,000,000,000 and prints each number on a separate line, creating a 9.3 GB output file. The second program, designed to be parallelized with a message passing interface (MPI) protocol, found the sum of prime numbers between 2 and 2,000,000. The third program was a longer version of the second, summing primes between 2 and 15,000,000.

At baseline, Beehive drew fewer watts than Hotfoot—2721 W vs. 4151 W, respectively. This was to be expected, however, as Beehive consisted of fewer computers (14 cores) than Hotfoot (256 cores). A different metric that is useful for analyzing idle power consumption is to adjust power usage for the amount of potential processing power, or peak Flops (floating point operations per second). Beehive had 14 cores and a theoretical peak Flops of 61.32 GFlops while Hotfoot had 256 cores and a theoretical peak of 2723.84 GFlops, where theoretical peak Flops = (number of cores) * (clock speed) * (floating-point operations per cycle)[6].

Dividing the watts drawn while idle by the theoretical peak flops gives us an idea of how many watts each system draws per GFlops. The ratios for Beehive and Hotfoot at idle were 44.37 and 1.52 watts per GFlops, respectively, suggesting that Hotfoot consumed significantly less power for each floating point operation per second that it was able to compute.

Table 8-1 provides a summary of energy usage for the two clusters while under load. The first two rows compared the energy required from the program that counts to one billion. It was run on one core on each cluster, with Beehive taking about half a minute longer than Hotfoot to run it. For this short, one-core program, Hotfoot was less efficient than Beehive (199 W·h vs. 152 W·h, respectively).

[6] Hotfoot had 256 cores, each rated at 2.66 GHz, and each core can perform four floating point operations per clock cycle. 256 cores * 2.66 billion cycles per second * 4 floating point operations per clock cycle = 2723.84 GFlops (billion floating point operations per second). See http://technet.microsoft.com/en-us/library/ee146531%28WS.10%29.aspx and http://www.intel.com/support/processors/xeon/sb/CS-020863.htm and http://www.theinquirer.net/inquirer/news/1039779/woodcrest-will-outperform-all-other-cpus-on-the-market



Job	Cluster	Runtime	Time Difference	Energy	Energy Difference
Count to one billion on 1 core					
	Beehive	3.33 minutes		0.15 kWh	
	Hotfoot	2.87 minutes	0.46 minutes	0.20 kWh	133%
Sum primes between 2 and 2 million on 14 cores					
	Beehive	13.02 minutes		0.61 kWh	
	Hotfoot	4.93 minutes	8.09 minutes	0.35 kWh	57%
Sum primes between 2 and 15 million on 14 cores					
	Beehive	8.92 **hours**		24.2 kWh	
	Hotfoot	3.87 **hours**	5.05 hours	16.3 kWh	67%
Sum primes between 2 and 15 million on 256 cores					
	Hotfoot	15.85 **minutes**	8.66 **hours**	1.3 kWh	5%

Table 8-1. Summary of energy usage from cluster comparison

This was not surprising, as Hotfoot drew more power at baseline than Beehive, and this program was not big enough for Hotfoot to be efficient, leaving 255 cores idle. Figure 8-1 and Figure 8-2 present graphs of power usage versus time for Beehive and Hotfoot, respectively. (Note: The graphs showing power usage on Beehive have two lines, each representing information from separate PDU components. The total power output is the sum of these two components.)

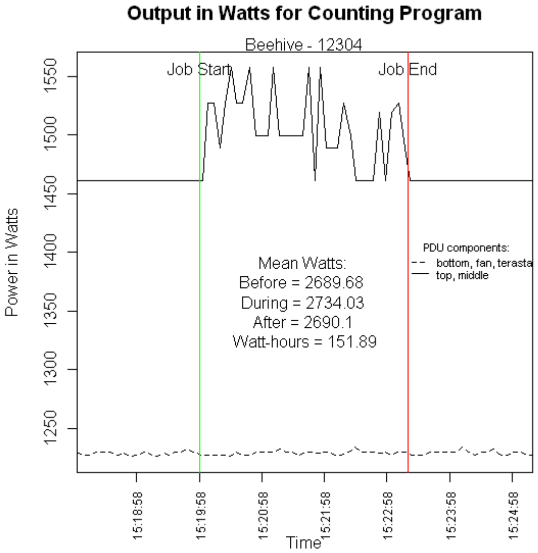


Figure 8-1. Old cluster power consumption: counting to 1 billion

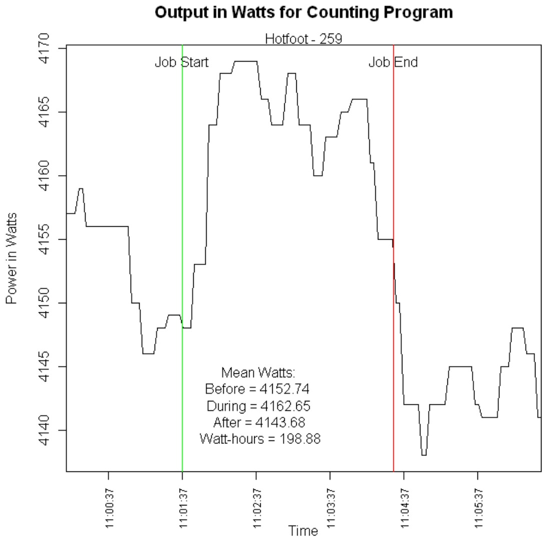
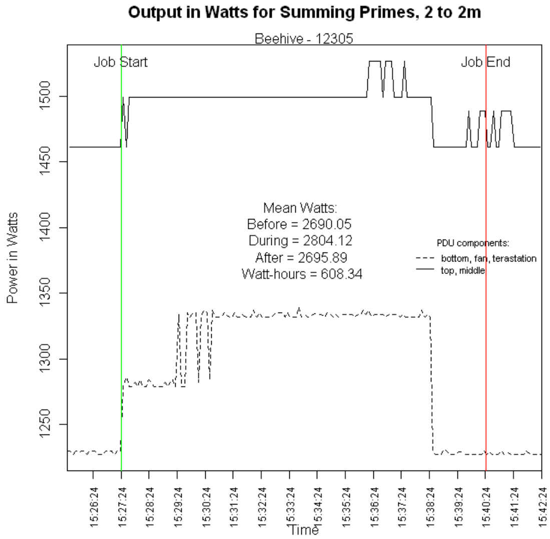


Figure 8-2. New cluster power consumption: counting to 1 billion

The next two rows in Table 8-1 compare energy usage for an MPI program that summed prime numbers between 2 and 2,000,000. The program was run on 14 cores on each cluster, since Beehive's maximum number of cores was 14. Beehive took nearly three times longer and used almost twice as much energy than

Hotfoot to run this program (608 W·h vs. 347 W·h, respectively). Figures 8-3 and 8-4 graphically compare the energy used during these jobs.



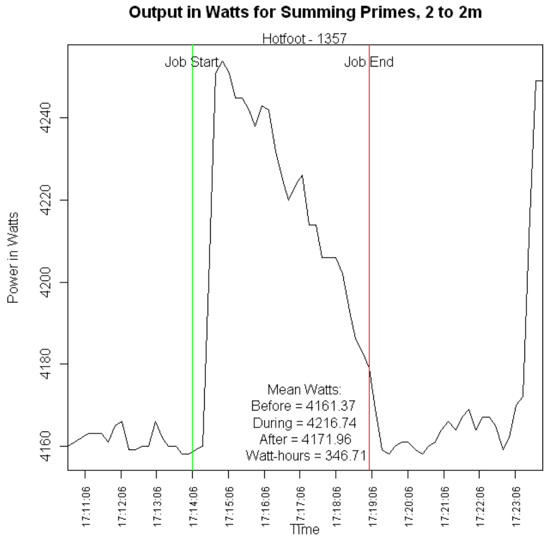
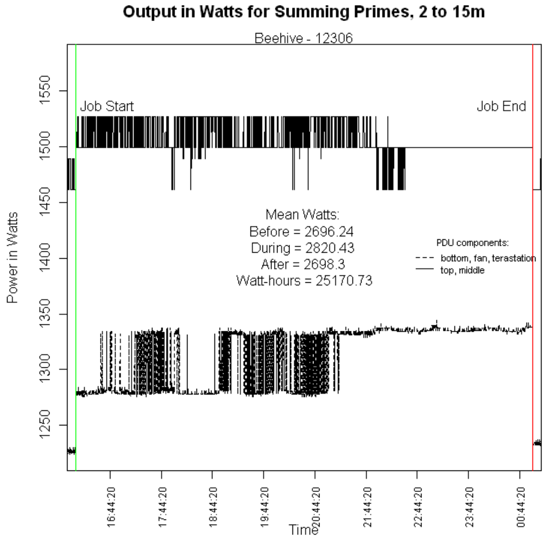


Figure 8-3. Old cluster power consumption: sum of primes 2 to 2m

Figure 8-4. New cluster power consumption: sum of primes 2 to 2m

The last three rows in Table 8-1 report energy usage for the program that summed prime numbers between 2 and 15,000,000. On 14 cores, Beehive took nearly nine hours to complete the task, while Hotfoot required about 4 hours. The total energy consumed by Beehive was roughly 50% greater than Hotfoot (24,171 W·h vs. 16,307 W·h, respectively). For comparison purposes, we also ran this job on the entire 256-core Hotfoot cluster. The job took just less than 16 minutes and used only 1,304 W·h, a mere fraction



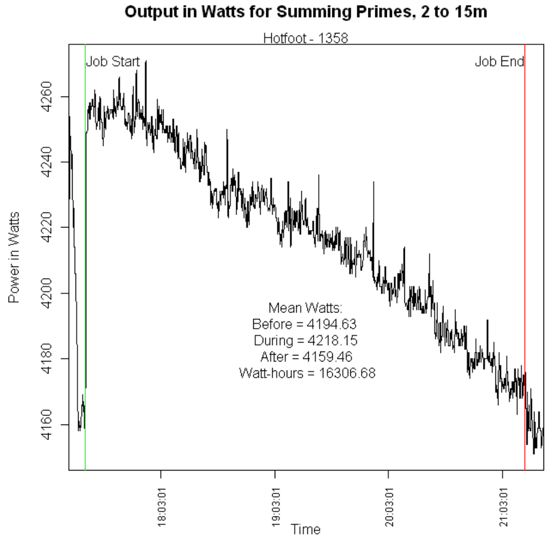


Figure 8-5. Old cluster power consumption: sum or primes 2 to 15m

Figure 8-6. New cluster power consumption: sum or primes 2 to 15m

(5%) of the energy that Beehive used at maximum capacity. Figure 8-5, Figure 8-6 and Figure 8-7 present power output vs. time graphs for Beehive and the two Hotfoot jobs, respectively. Figure 8-8 summarizes how Hotfoot uses less energy than Beehive to run jobs. In addition to energy savings, the wall clock time savings facilitate greater research productivity, for example, reducing a 9-hour job to 16 minutes.

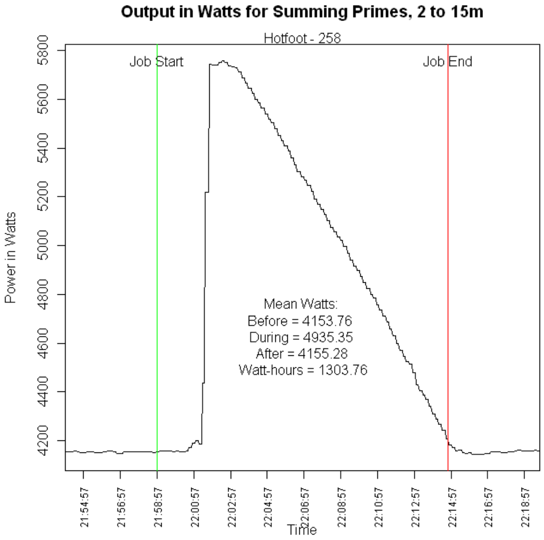


Figure 8-7. New cluster power consumption using all available cores: primes 2 to 15m

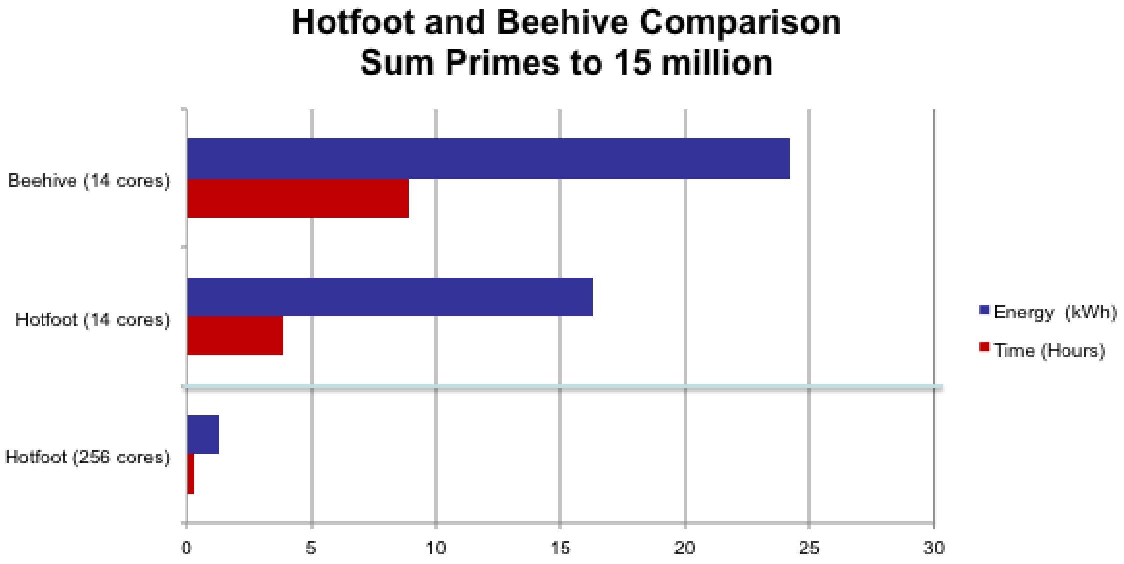


Figure 8-8. Old vs. new clusters: power consumed and elapsed time

9. IMPLEMENT SERVER POWER MANAGEMENT

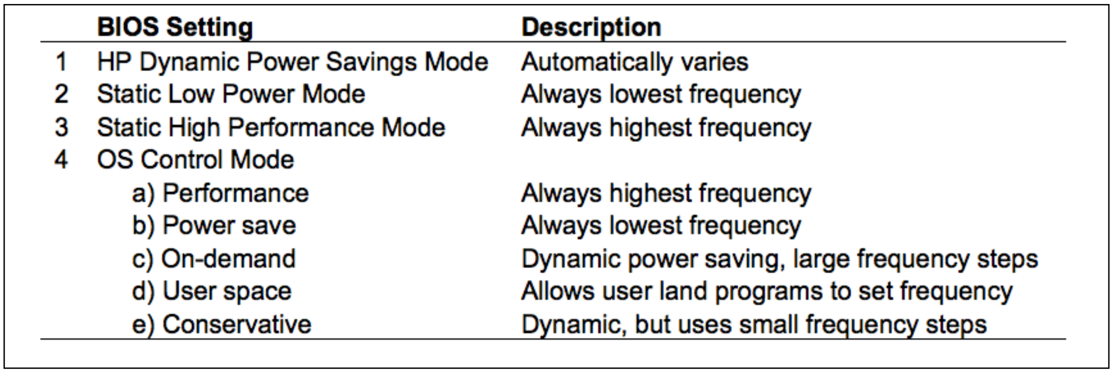
INTRODUCTION

This task calls for the implementation of power management features that are available at the BIOS- and OS-level on the servers identified in the Task 2 inventory. We used the power consumption data collection methods and benchmark programs described in Tasks 3, 7, and 8 to complete this task.

RESULTS

Servers

Of the various Sun and HP servers identified in Task 2 and 7 that could serve as replacement or "newer" hardware, very few of them provided any power management or power tuning options. A thorough review of manuals and related documentation revealed that only the standalone server HP DL380g5 and the blade server HP BL460cg6 were capable of power tuning. A summary of power tuning options is provided in Table 9-1. Options 1-3 denote power tuning that is specifically controlled by the BIOS. The fourth BIOS option enabled the OS control mode, which allowed five additional options for power management that are fine-tuned using system files. We selected a subset of these tuning options for testing.



	BIOS Setting	Description
1	HP Dynamic Power Savings Mode	Automatically varies
2	Static Low Power Mode	Always lowest frequency
3	Static High Performance Mode	Always highest frequency
4	OS Control Mode	
	a) Performance	Always highest frequency
	b) Power save	Always lowest frequency
	c) On-demand	Dynamic power saving, large frequency steps
	d) User space	Allows user land programs to set frequency
	e) Conservative	Dynamic, but uses small frequency steps

Table 9-1. Power tuning settings for select HP servers

On both servers, the default power mode was HP Dynamic Power Savings Mode, which automatically varies processor speed and power usage based on processor utilization and is controlled by the BIOS. We compared SPECpower benchmarks with this default power setting to benchmarks run in other power-saving modes. For the DL380, we used three OS-level controls: On-demand, Power save, and Performance. (We were able to confirm that BIOS modes 1, 2 and 3 are equivalent to OS modes c, b, and a, respectively. OS Control Modes were selected for tests because they are modifiable without rebooting the server.) For the BL460, we compared the default setting to the OS-level Performance and Conservative modes.

Table 9-2 summarizes the SPECpower benchmarks and active idle power measurements for these two systems. Both servers were shown to have the best SPECpower benchmark results under the default Dynamic Power Savings setting. Compared to the default setting, power measurements at active idle levels remained the same for the DL380 regardless of how it was tuned. The BL460, however, had higher active idle consumption when in the Performance mode, as expected, and had standard consumption levels when conservatively tuned to Power Save mode.

	Default Setting		OS On-demand		OS Power save		OS Performance	
		Idle		Idle		Idle		Idle
Server	ssj_ops/W	W	ssj_ops/W	W	ssj_ops/W	W	ssj_ops/W	W
HP DL380 G5	230	263	226	265	218	265	219	263
HP BL460c G6	764	100	-	-	539	99	687	141

Table 9-2. Power tuning power measurements

Graphs in Figure 9-1, Figure 9-2 and Figure 9-4highlight the performance differences when the BL460 is running the SPECpower benchmark in default, Performance, and Power save modes, respectively. In

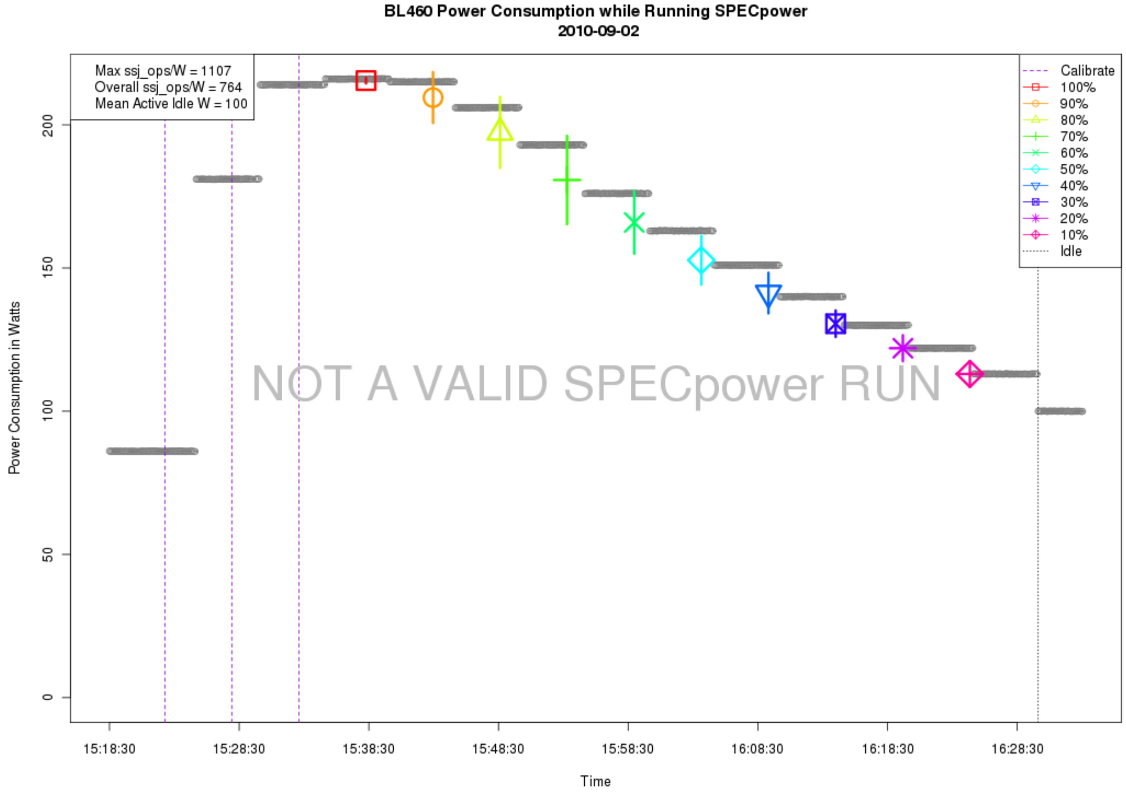


Figure 9-1. BL460 SPECpower: HP Dynamic Power Savings BIOS power setting

Figure 9-1 power consumption decreased steadily towards 100 W as the machine was put at lower loads, whereas in Figure 9-2 power consumption decreased more slowly towards 140 W. Figure 9-4 shows that when tuned to Power save mode, the maximum power consumption was around 150 W (compared to 200

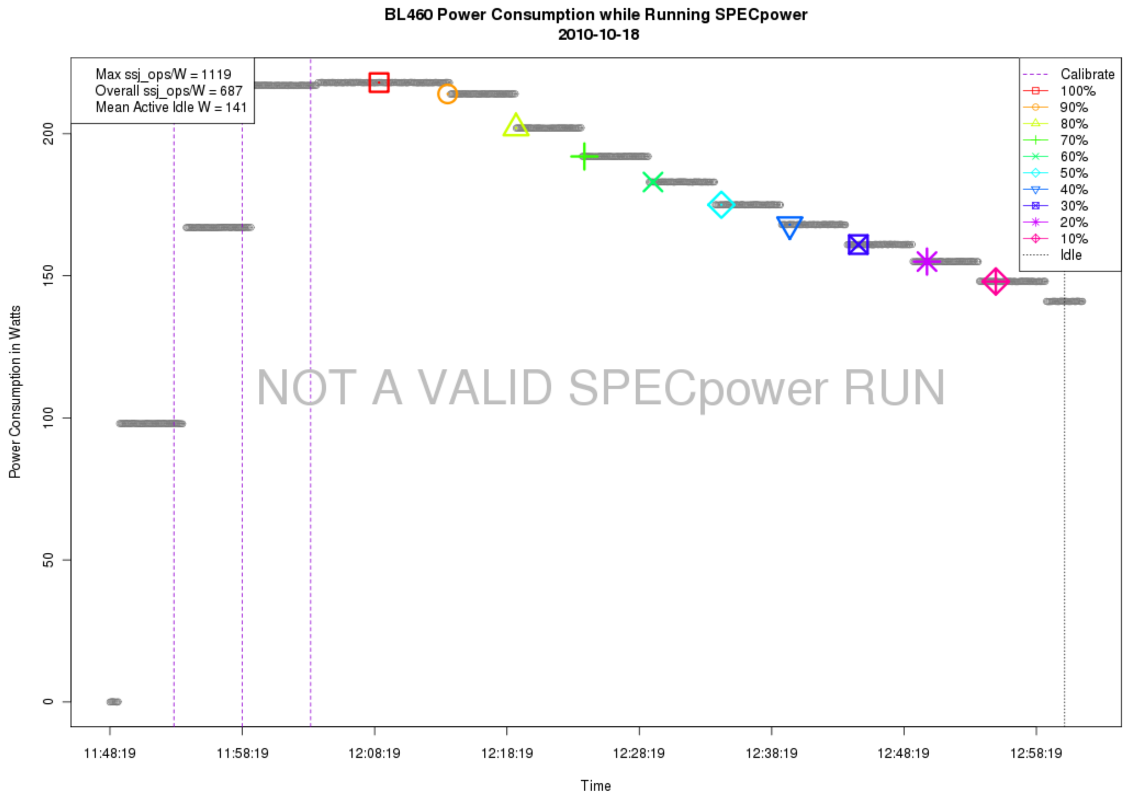


Figure 9-2. BL460 SPECpower: Performance OS power setting

W in the other modes), and the consumption decreased to 100 W during the benchmark. These three examples summarized the power tuning range for the BL460, providing information useful for future tuning decisions.

In addition to using SPECpower to evaluate power tuning comparisons, we implemented the OS Power Save Mode on the BL460 blade in Service Group 6 (LAMP database host) and monitored power before and before and after the change. In Figure 9-3 we present the power consumption profile of a BL460 blade server before and after it was power tuned. In default mode, the server consumed about 95 W. After being put in low power mode, the average was 92 W. (Note: the BL460 measured with the SPECpower benchmark had idle power consumption of about 100 W. We do not understand why this particular server consumed less at idle.) The change is shown by the dashed black line and the corresponding step down in power consumption.

Plots like Figure 9-3 show more than just what happened on the power tuning date. The tick marks on the horizontal axis are drawn at the beginning of each day (midnight), and we also indicate that October 25 is a Monday by an elongated tick. The daily pattern for this server included power spikes in the first few hours of the day, and a small increase around mid day during the work week (note its absence on Sunday the 24th). During this typical week mid-semester, it was clear that the power consumption never approached its

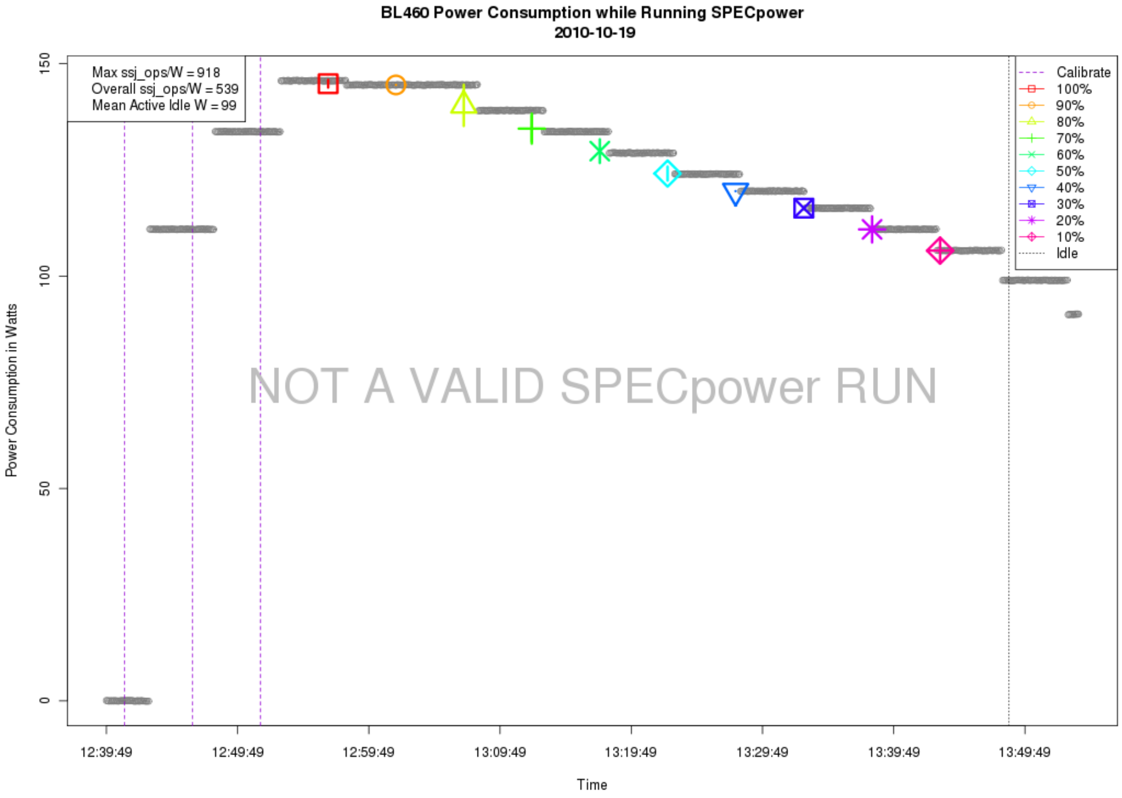


Figure 9-4. BL460 SPECpower: Conservative OS power setting

200 W maximum shown in the SPECpower graphs. This LAMP database host was a good candidate for a permanent low power setting, or for multiple services or virtualization.

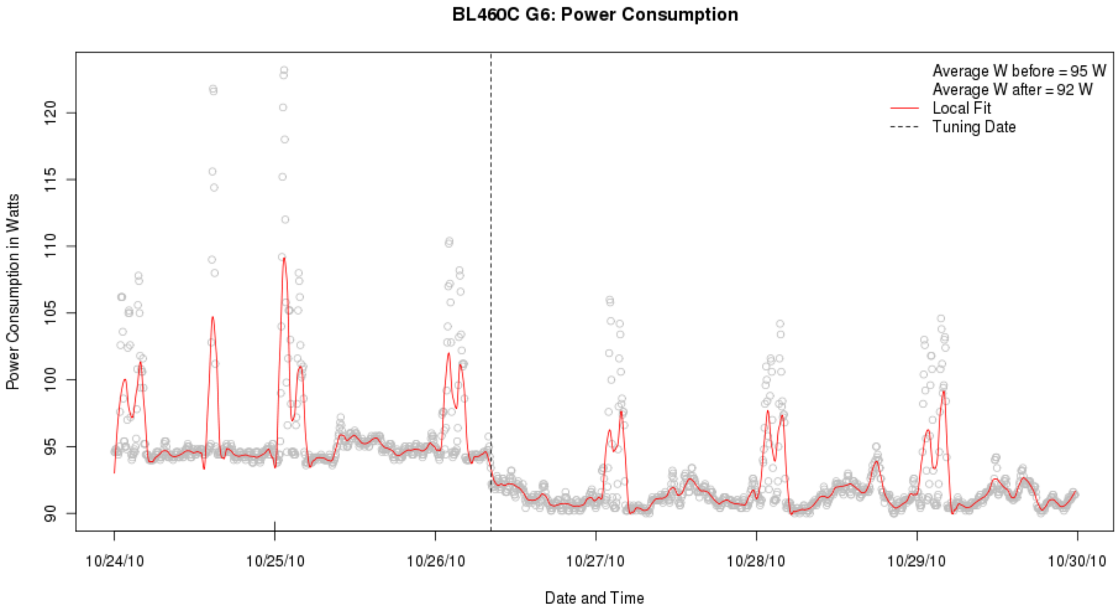


Figure 9-3. BL460 power tuning impact over time

New High Performance Computing Cluster

The 32 HP BL2x220c blade servers on the HPC system (Hotfoot) were capable of being power tuned (but the two HP DL360 head nodes and NFS server were not). The default power mode was HP Dynamic Power Savings Mode as defined above. We ran the SPEC benchmark and monitored active idle power usage on the entire HPC system with the default setting and with the OS-level Conservative power tuning setting. The results of the SPECpower benchmark were the same in both runs, and this was largely due to the additional overhead of the three standalone servers and the storage shelf that used the same monitored power supply as the blades and chassis. Since the OS-level conservative power tune option stepped power up and down more gradually, the power usage during the active idle state after the blades were tuned was greater than usage before they were tuned – 4209 W vs. 4153 W, respectively.

Electrical Engineering High Performance Computing Cluster

We discovered that the 100-node Dell cluster defined in Task 2 could not be power tuned. In order to enable power tuning, four components of the host must be capable of supporting it: the CPU, the motherboard and chipset, and BIOS, and the operating system. After communicating with Dell, we learned there was no power tuning capability on this system that had been selected for highest performance at lowest purchase cost, without consideration for ongoing operating cost. This is typical of how many server systems are evaluated and selected within the University funding model, something we hope to influence with this work.

10. INCREASE CHILLED WATER SET POINT

INTRODUCTION

The purpose of Task 10 was to investigate the concept of saving energy by implementing a 5-degree increase in the baseline temperature of chilled water delivered to the newly installed high power density racks. Theoretically, this should produce energy savings. Originally, we planned to ask our chosen vendor for the "self-cooled" racks in Task 6 about the feasibility of raising the chilled water temperature within the racks for increased energy savings.

Also in line with this task, CUIT has investigated the feasibility of increasing the temperature of the chilled water delivered to the entire data center.

RESULTS

CUIT worked with CU Facilities to perform a feasibility study. The following was determined:

The campus chilled water system is a primary water system (without heat exchangers) that feeds multiple campus buildings for both comfort and process cooling. The system nominally operates at 43° F supply temperature (CHWS) and 55° F return temperature (CHWR). The proposal to raise the CHWS temperature by 5 degrees is not possible. Where the latent cooling load is high due to occupancy of classrooms, auditoriums, cafeterias and the like, psychometrically the humidity levels within the spaces would become uncomfortably high and beyond recommendations of ASHRAE Standard 55 "Thermal Environmental Conditions for Human Occupancy". Similar conditions would occur in our 100 percent outdoor air lab buildings in the summer. The chilled water simply would not be cold enough to "wring out" the moisture.

While many data centers are standalone facilities with their own dedicated chilled water plant, Columbia's is not; the chilled water serving the data center is a campus-wide shared service, used for both comfort and process cooling. We benefit from the energy efficiency of this shared resource but are also constrained by that.

11. COMMUNICATE RESULTS

INTRODUCTION

Successful dissemination of knowledge was one of the key themes of our proposal. We have engaged in multiple opportunities to transfer the knowledge we gained from this project to constituencies within Columbia University and to other New York State and national institutions facing similar challenges. The emphasis on rigorous measurement, and the inclusion of the Research User Group, Internal Advisory Group and External Visiting Committee, are examples of our effort to create a culture of "green thinking" when it comes to university data center planning.

PRESENTATIONS AND EVENTS

Project Blog

Shortly after this project was awarded, we created a publically available blog to provide updates on our progress. All presentation materials discussed below are available on our project blog:

http://blogs.cuit.columbia.edu/greendc

Presentations and Publications

- 7/19/11 Ian Katz presented details on data center metering at the Global Strategic Management Institute's Green Data Center Conference in Boston, MA.

- 3/24/11 Ian Katz participated on a panel at the Extreme Data Center Efficiency Summit in New York, NY.

- 10/19/10 Alan Crosswell and Rich Hall presented at the 2010 EDUCAUSE Annual Meeting, Anaheim, CA.

- 06/14/10 Alan Crosswell was an invited speaker at the ACM SIGMETRICS GreenMetrics2010 Workshop at Columbia University, New York, NY.

- 05/03/10 Rajendra Bose, Alan Crosswell and Victoria Hamilton participated in the NSF Workshop on Sustainable Cyberinfrastructure, Cornell University, Ithaca, NY.

- 04/15/10 This project was cited in the EDUCAUSE Center for Applied Research Green IT study (See Sheehan and Smith, 2010, pp. 50, 52, 65, 67, 91, 97, 105).

- 03/03/10 Alan Crosswell participated in the Datacenter Dynamics conference panel sponsored by NYSERDA, New York, NY.

- 10/20/09 Alan Crosswell presented at the Association of IT Professionals, Long Island Chapter (AITP-LI) meeting.
- 10/06/09 Alan Crosswell presented at the Internet2 Member Meeting, San Antonio, TX.
- 5/07/09 This project was cited in the newly published book *The Greening of IT: How Companies Can Make a Difference for the Environment* by John Lamb (IBM Press).
- 03/04/09 Alan Crosswell participated in Canada's Advanced Research and Innovation Network (CANARIE) Green IT workshop, Ottawa, Canada.

Open House Workshop

A public "Winter Workshop" on Green Data Centers was held at the Columbia Faculty House on January 7, 2011, with an audience of roughly 50 information technology and facilities professionals and other higher education staffers from the New York metro area and elsewhere, including representatives from City University of New York (CUNY), Albert Einstein College of Medicine, Rockefeller University, New York University, Pennsylvania State University, Princeton University, Yale University and the University of Chicago. The agenda and presentations for the workshop are available on the project blog .

REFERENCES

Anderson, D., et. al (2008). A Framework for Data Center Energy Productivity. The Green Grid. http://www.thegreengrid.org/~/media/WhitePapers/WhitePaper13FrameworkforDataCenterEnergyProductivity5908.pdf?lang=en

ASHRAE Technical Committee 9.9 (2009). Best Practices for Datacom Facility Energy Efficiency, Second Edition, ASHRAE.

Balakrishnan, S. and D. Z. Spicer (2008). IT and Campus Climate Change. Boulder, CO. http://www.educause.edu/ecar

Bose, R., A. Crosswell, V. Hamilton, and N. Mesa (2010). Piloting sustainable HPC for research at Columbia. In Proceedings of Sustainable Funding and Business Models for Academic Cyberinfrastructure (CI) Facilities workshop, Cornell University, Ithaca, NY. http://www.cac.cornell.edu/srcc/

Crosswell, A. (2009). Columbia's Green Data Center Program. In proceedings of Internet2 Member Meeting, San Antonio, TX. http://www.internet2.edu/presentations/fall09/20091006-green-crosswell.pdf

Crosswell, A. (2010). Green Data Center Program. In proceedings of ACM Sigmetrics 2010 Green Metrics Workshop, New York, NY. http://www.sigmetrics.org/sigmetrics2010/greenmetrics/AlanCrosswell.pdf

Crosswell, A. (2012). Measuring and Managing a Data Center Cooling Shutdown. http://blogs.cuit.columbia.edu/greendc/files/2013/01/MeasuringandManagingDCShutdown.pdf

Crosta, P. and R.D. Hall (2010). Measuring and Validating Attempts to Green Columbia's Data Center. In proceedings of EDUCAUSE 2010 Annual Conference, Anaheim, CA. http://www.educause.edu/sites/default/files/library/presentations/E10/SESS074/MeasuringAndValidatindCUDC_Educause_2010-10-19.pptx

DOE and EPA (2008). Energy Efficiency in Data Centers: Recommendations for Government-Industry Coordination. October 16, 2008. http://www.energystar.gov/ia/partners/prod_development/downloads/energy_eff_data_centers_rec.pdf

Environmental Protection Agency (2007). Report to Congress on Server and Data Center Energy Efficiency: Public Law 109-431. 2 Aug 2007. http://www.energystar.gov/ia/partners/prod_development/downloads/EPA_Datacenter_Report_Congress_Final1.pdf

Esser, A. (2008). Best Practices for Unlocking Your Hidden Data Center. http://www.dell.com/downloads/global/power/ps1q08-20080198-Esser.pdf

Lamb, J.P. (2009). The Greening of IT: How Companies Can Make a Difference for the Environment. Pearson Education. ISBN 9780137150830.

Sheehan, M. C. and S. D. Smith (2010). Powering Down: Green IT in Higher Education. Boulder, CO. http://www.educause.edu/ecar

APPENDIX A: INVENTORY DOCUMENTATION

OLD SERVERS TO BE REPLACED

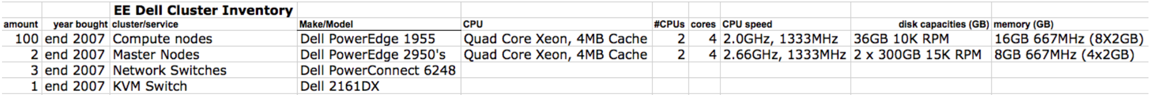
Hostname	Make/Model	Rack Location	Power Supply #	PDU Name	PDU Outlet
tepin	Sun Fire 280R	already retired			
cayenne	Sun Fire 280R	philorack9	1	unix103rack10-pdu1	7
cayenne	Sun Fire 280R	philorack9	2	unix103rack11-pdu1	8
caraway	Sun Fire V100	philorack10	1	unix103rack10-pdu1	3
ginger	Sun Fire V100	philorack10	1	unix103rack10-pdu1	4
mustard	Sun Fire V100	philorack10	1	unix103rack10-pdu1	5
thyme	Sun Fire V100	philorack10	1	unix103rack10-pdu1	6
serrano	Sun Fire 280R	philorack10	1	unix103rack10-pdu1	1
serrano	Sun Fire 280R	philorack10	2	unix103rack11-pdu1	7
cashew	Sun Netra T1	philorack11	1	unix103rack11-pdu1	5
thunderhead	Sun Netra T1	philorack11	1	unix103rack11-pdu1	6
mint	Sun Fire V100	philorack12	1	unix103rack11-pdu1	3
nutmeg	Sun Fire V100	philorack12	1	unix103rack11-pdu1	4
parsley	Sun Fire V100	philorack12	1	unix103rack11-pdu1	2
sage	Sun Fire V100	philorack12	1	unix103rack11-pdu1	1
boprod1	HP DL380	maltsrackA8	1	maltsracka7-pdu2	1
boprod1	HP DL380	maltsrackA8	2	maltsracka7-pdu1	1
boprod2	HP DL380	maltsrackA8	1	maltsracka7-pdu2	2
boprod2	HP DL380	maltsrackA8	2	maltsracka7-pdu1	2
boprod3	HP DL380	maltsrackA8	1	maltsracka7-pdu2	3
boprod3	HP DL380	maltsrackA8	2	maltsracka7-pdu1	3
bodev	HP DL380	maltsrackA8	1	maltsracka7-pdu2	4
bodev	HP DL380	maltsrackA8	2	maltsracka7-pdu1	4
bostage	HP DL380	maltsrackA8	1	maltsracka7-pdu2	5
bostage	HP DL380	maltsrackA8	2	maltsracka7-pdu1	5
botest	HP DL380	maltsrackA8	1	maltsracka7-pdu2	6
botest	HP DL380	maltsrackA8	2	maltsracka7-pdu1	6
funnel	Sun Netra T1	unixrack11	1	unixrack12-pdu1	1
peanut	Sun Netra T1	unixrack11	1	unixrack12-pdu1	2
coconut	Sun Netra T1	unixrack12	1	unixrack12-pdu1	3
filbert (to be retired)	Sun Netra T1	unixrack12	1		
hazelnut	Sun Netra T1	unixrack12	1	unixrack12-pdu1	4

pecan	Sun Netra T1	unixrack12	1	unixrack12-pdu1	7
pistachio	Sun Netra T1	unixrack12	1	unixrack12-pdu1	8
cobnut	Sun Netra T1	unixrack13	1	unixrack13-pdu1	1
hickory	Sun Netra T1	unixrack13	1	unixrack13-pdu1	2
hickory-disks	Sun D130	unixrack13	1	unixrack13-pdu1	3
chili	Sun Fire 280R	unixrack14	1	unixrack14-pdu1	7
chili	Sun Fire 280R	unixrack14	2	unixrack14-pdu2	7
chili-raid	Sun StorEdge t3	unixrack14	1	unixrack14-pdu1	8
chili-raid	Sun StorEdge t3	unixrack14	2	unixrack14-pdu2	8
datil	Sun Fire 280R	unixrack14	1	unixrack14-pdu1	3
datil	Sun Fire 280R	unixrack14	2	unixrack14-pdu2	3
mirasol	Sun Fire 280R	unixrack14	1	unixrack14-pdu1	1
mirasol	Sun Fire 280R	unixrack14	2	unixrack14-pdu2	1
pimento	Sun Fire 280R	unixrack14	1	unixrack14-pdu1	2
pimento	Sun Fire 280R	unixrack14	2	unixrack14-pdu2	2
sausage	HP DL360g3	unixrack20	1	unixrack20-pdu1	6
sausage	HP DL360g3	unixrack20	2	unixrack20-pdu2	1
spam	HP DL360g3	unixrack20	1	unixrack20-pdu1	7
spam	HP DL360g3	unixrack20	2	unixrack20-pdu2	2
allspice	Sun Fire V100	unixrack20	1	unixrack20-pdu1	1
cardamom	Sun Fire V100	unixrack20	1	unixrack20-pdu1	2
poppy	Sun Fire V100	unixrack20	1	unixrack20-pdu1	3
saffron	Sun Fire V100	unixrack20	1	unixrack20-pdu2	4
sesame	Sun Fire V100	unixrack20	1	unixrack20-pdu2	5
bacon	HP DL360g3	unixrack22	1	unixrack20-pdu2	3
bacon	HP DL360g3	unixrack22	2	unixrack20-pdu2	4
dill	Sun Fire V100	unixrack2	1	unixrack2-pdu1	1
tayberry	Sun T2000	unixrack2	1	unixrack2-pdu1	2
tayberry	Sun T2000	unixrack2	2	unixrack2-pdu1	3
brazilnut (to be retired)	Sun Netra T1	unixrack5	1		
basil	Sun Fire V100	unixrack5	1	unixrack4-pdu1	1
cilantro	Sun Fire V100	unixrack5	1	unixrack4-pdu1	2
cumin	Sun Fire V100	unixrack5	1	unixrack4-pdu1	3
lovage	Sun Fire V100	unixrack5	1	unixrack4-pdu1	4
oregano	Sun Fire V100	unixrack5	1	unixrack4-pdu1	7
rosemary	Sun Fire V100	unixrack5	1	unixrack4-pdu1	8
strawberry	Sun T2000	unixrack34	1	unixrack34-pdu1	1
strawberry	Sun T2000	unixrack34	2	unixrack34-pdu2	1
jalapeno	Sun Fire 280R	unixrack63	1	unixrack63-pdu1	1
jalapeno	Sun Fire 280R	unixrack63	2	unixrack63-pdu1	2
casaba	Sun Fire v880	unixrack66	1	unixrack66-pdu1	3
casaba	Sun Fire v880	unixrack66	2	unixrack66-pdu1	4
casaba	Sun Fire v880	unixrack66	3	unixrack66-pdu2	3
casaba-raida	Sun StorEdge t3	unixrack66	1	unixrack66-pdu2	2
casaba-raida	Sun StorEdge t3	unixrack66	2	unixrack66-pdu2	8
casaba-raidb	Sun StorEdge t3	unixrack66	1	unixrack66-pdu1	2
casaba-raidb	Sun StorEdge t3	unixrack66	2	unixrack66-pdu2	7
chipotle	Sun Fire 280R	unixrack66	1	unixrack66-pdu1	1
chipotle	Sun Fire 280R	unixrack66	2	unixrack66-pdu2	1

squid	HP DL380	unixrack70	1	unixrack70-pdu1	1
squid	HP DL380	unixrack70	2	unixrack70-pdu2	1
cockle	HP DL360g5	unixrack70	1	unixrack70-pdu1	2
cockle	HP DL360g5	unixrack70	2	unixrack70-pdu2	2
ormer	HP DL360g5	unixrack70	1	unixrack70-pdu1	3
ormer	HP DL360g5	unixrack70	2	unixrack70-pdu2	3

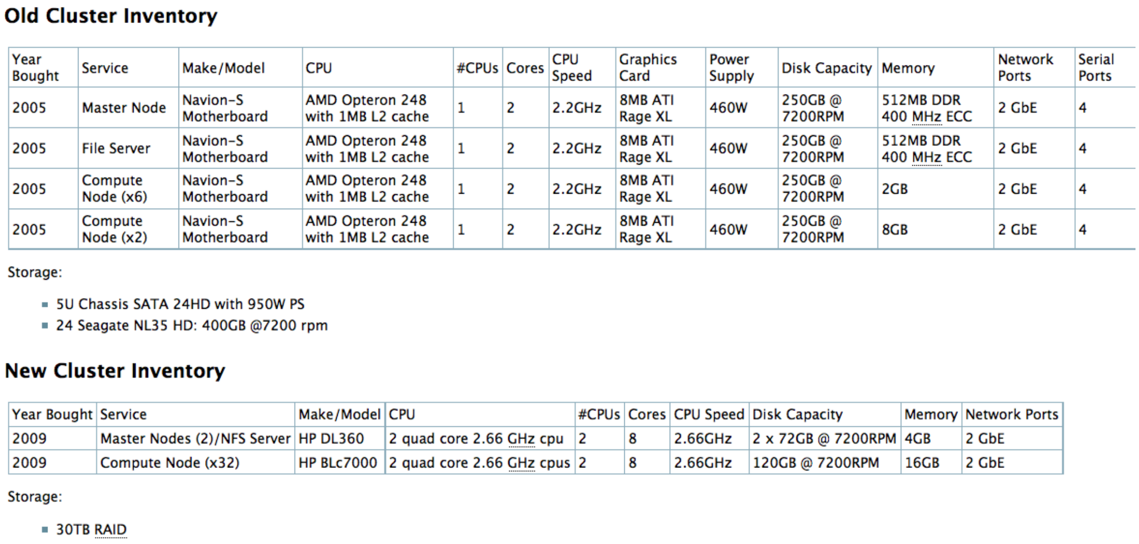
🟥 no pdu necessary
🟩 computer center
🟦 103 philosophy

ELECTRICAL ENGINEERING CLUSTER



		EE Dell Cluster Inventory							
amount	year bought	cluster/service	Make/Model	CPU	#CPUs	cores	CPU speed	disk capacities (GB)	memory (GB)
100	end 2007	Compute nodes	Dell PowerEdge 1955	Quad Core Xeon, 4MB Cache	2	4	2.0GHz, 1333MHz	36GB 10K RPM	16GB 667MHz (8X2GB)
2	end 2007	Master Nodes	Dell PowerEdge 2950's	Quad Core Xeon, 4MB Cache	2	4	2.66GHz, 1333MHz	2 x 300GB 15K RPM	8GB 667MHz (4x2GB)
3	end 2007	Network Switches	Dell PowerConnect 6248						
1	end 2007	KVM Switch	Dell 2161DX						

OLD AND NEW HIGH PERFORMANCE COMPUTING CLUSTERS



Old Cluster Inventory

Year Bought	Service	Make/Model	CPU	#CPUs	Cores	CPU Speed	Graphics Card	Power Supply	Disk Capacity	Memory	Network Ports	Serial Ports
2005	Master Node	Navion-S Motherboard	AMD Opteron 248 with 1MB L2 cache	1	2	2.2GHz	8MB ATI Rage XL	460W	250GB @ 7200RPM	512MB DDR 400 MHz ECC	2 GbE	4
2005	File Server	Navion-S Motherboard	AMD Opteron 248 with 1MB L2 cache	1	2	2.2GHz	8MB ATI Rage XL	460W	250GB @ 7200RPM	512MB DDR 400 MHz ECC	2 GbE	4
2005	Compute Node (x6)	Navion-S Motherboard	AMD Opteron 248 with 1MB L2 cache	1	2	2.2GHz	8MB ATI Rage XL	460W	250GB @ 7200RPM	2GB	2 GbE	4
2005	Compute Node (x2)	Navion-S Motherboard	AMD Opteron 248 with 1MB L2 cache	1	2	2.2GHz	8MB ATI Rage XL	460W	250GB @ 7200RPM	8GB	2 GbE	4

Storage:

- 5U Chassis SATA 24HD with 950W PS
- 24 Seagate NL35 HD: 400GB @7200 rpm

New Cluster Inventory

Year Bought	Service	Make/Model	CPU	#CPUs	Cores	CPU Speed	Disk Capacity	Memory	Network Ports
2009	Master Nodes (2)/NFS Server	HP DL360	2 quad core 2.66 GHz cpu	2	8	2.66GHz	2 x 72GB @ 7200RPM	4GB	2 GbE
2009	Compute Node (x32)	HP BLc7000	2 quad core 2.66 GHz cpus	2	8	2.66GHz	120GB @ 7200RPM	16GB	2 GbE

Storage:

- 30TB RAID

APPENDIX B: DC PRO REPORT

A DC Pro 2.0 report is attached below. See the discussion on page 5-4.

Data Center Profiler Case Results

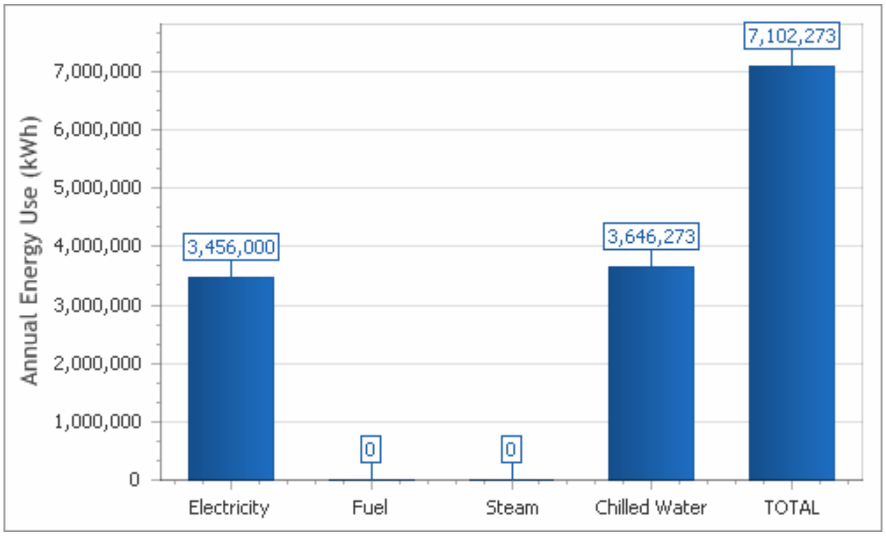
This is your customized DCPro Summary Report. The report is broken into five basic sections. If you wish to go back and edit any of your values or add more data click the previous button at the bottom of the page to navigate to the desired screen.

Case Information

Case Name	CUIT Data Center assited by NYSERDA PON 1206
Company	Columbia University
County	New York City
State	New York

Annual Energy Use

	Site Usage	Unit	Site Cost	Unit Cost
Electricity	3,456,000	kWh	$656,640	$0.19
Fuel	0	kWh	$0	$0.00
Steam	0	kWh	$0	$0.00
Chilled Water	3,646,273.4	kWh	$362,880	$0.10
TOTAL	7,102,273.4	kWh	$1,019,520	$0.14

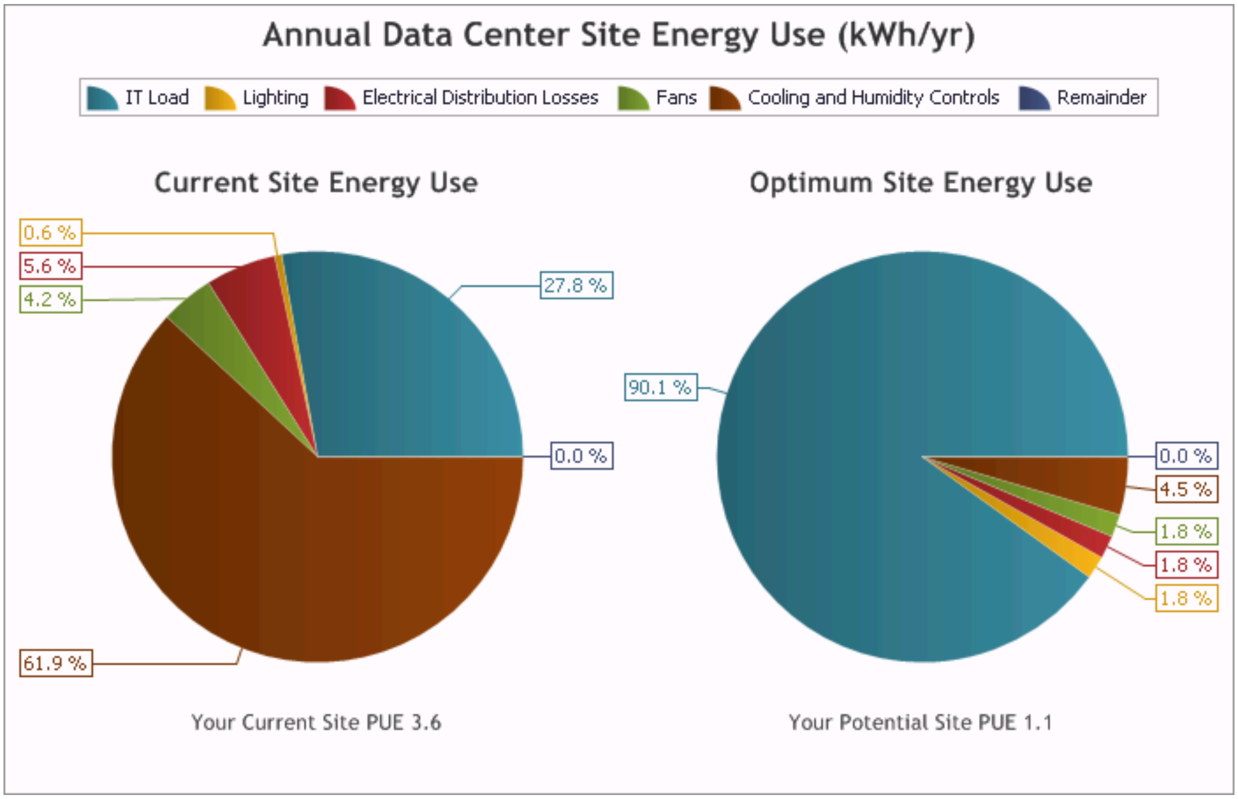


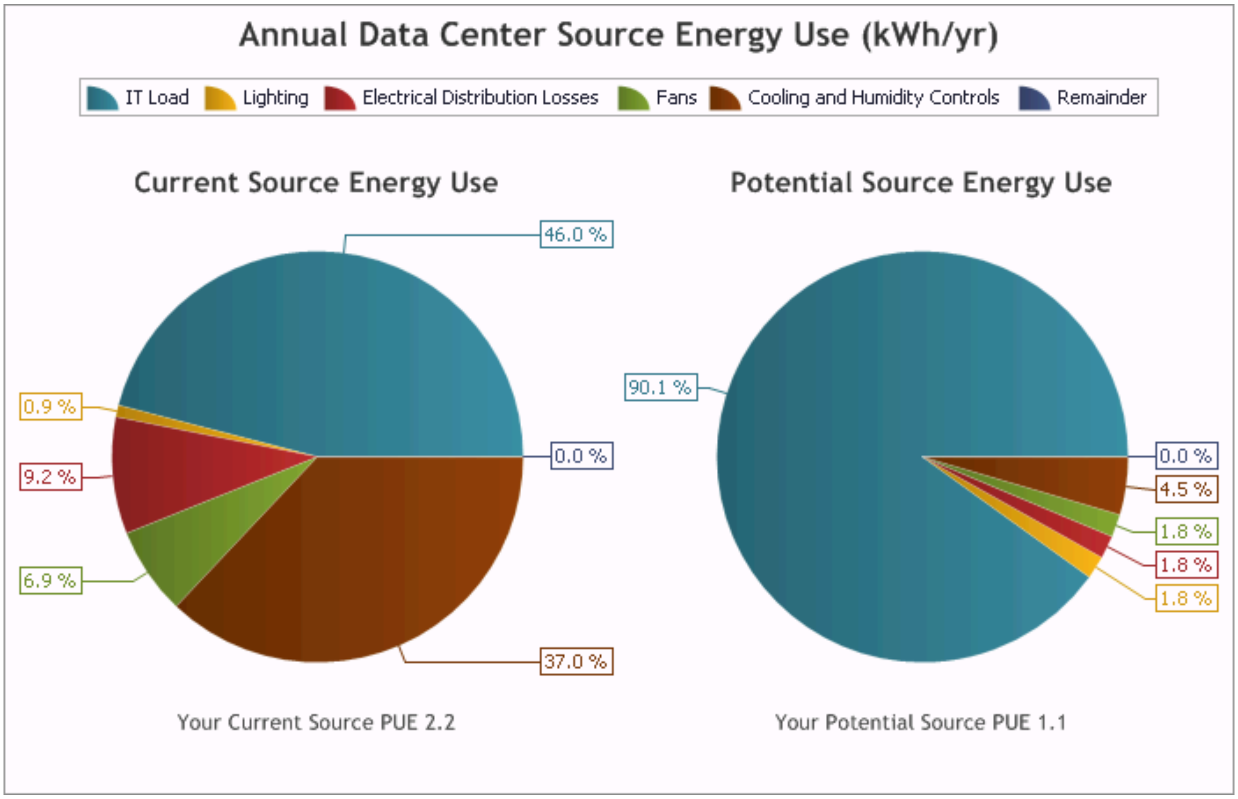
Potential Annual Energy Savings

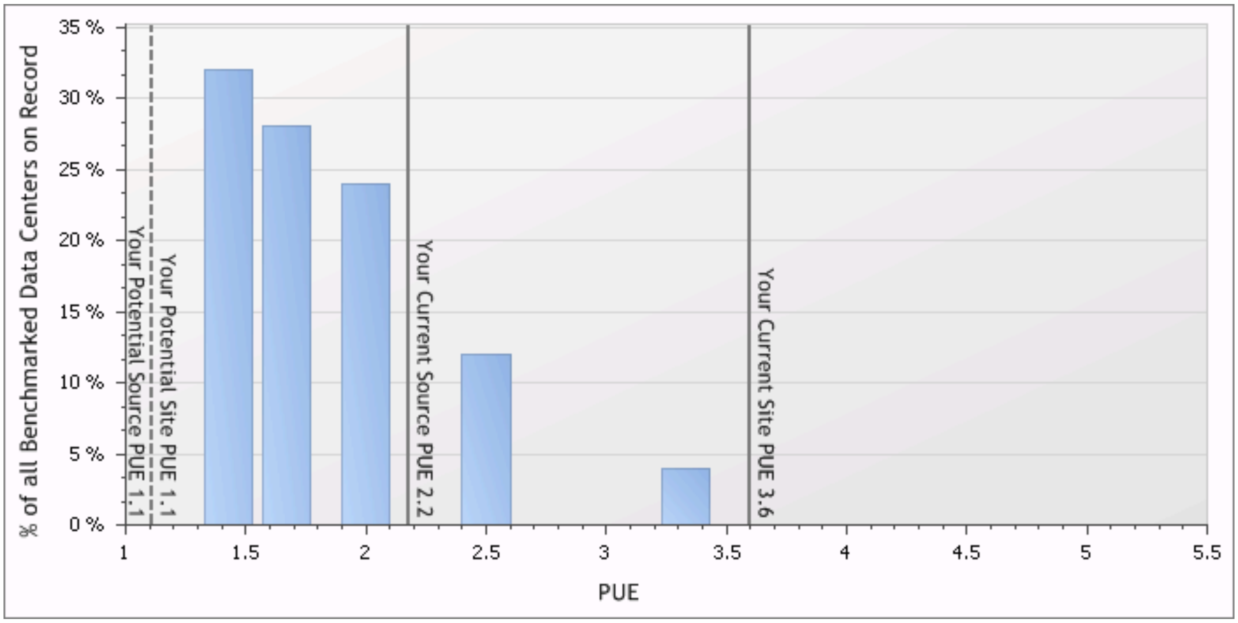
The following chart and data table summarize your data center's potential annual energy savings by breakout category. NOTE:The energy and money savings listed below are only estimates based on the data you entered and the estimated costs associated with the data center suggested improved. Your actual savings will vary.

	Current Energy Use				Optimum Energy Use				Potential Savings (Site Energy)		
	Site Energy		Source Energy		Site Energy		Source Energy				
Breakout Category	**kWh/yr**	**%**	**kWh/yr**	**%**	**kWh/yr**	**%**	**kWh/yr**	**%**	**kWh/yr**	**% ***	**$**
IT Load	1,974,758.8	27.8 %	6,595,694.4	46.0 %	1,974,758.8	90.1 %	6,595,694.4	90.1 %	0	0.0 %	0
Lighting	39,398.2	0.6 %	131,590	0.9 %	39,495.2	1.8 %	131,913.9	1.8 %	-97	0.0 %	-14
Electrical Distribution Losses	394,675.1	5.6 %	1,318,214.8	9.2 %	39,495.2	1.8 %	131,913.9	1.8 %	355,179.9	5.0 %	49,725
Fans	296,179.1	4.2 %	989,238.2	6.9 %	39,495.2	1.8 %	131,913.9	1.8 %	256,683.9	3.6 %	35,936
Cooling and Humidity Controls	4,396,570.5	61.9 %	5,313,623.9	37.0 %	98,737.9	4.5 %	329,784.7	4.5 %	4,297,832.6	60.5 %	601,697
Remainder	691.2	0.0 %	2,308.6	0.0 %	691.2	0.0 %	2,308.6	0.0 %	0	0.0 %	0
Total	7,102,272.9		14,350,669.9		2,192,673.5		7,323,529.4		4,909,599.4	69.1 %	687,344
PUE	3.6		2.2		1.1		1.1				

*Potential Savings % (Site Energy) displayed in the table above show the percent of your data center's total current energy consumption that can be saved (i.e. potential savings % = 100 * [potential savings / current total energy use]).







This chart compares your data center to a peer group of 25 other data centers.

Potential Annual CO_2 Savings

Based on the potential energy savings identified above, your data center may be able to reduce emissions of CO_2. The following potential annual CO_2 emission savings number is a broad estimate based on the estimated costs associated with the data center suggested improved and is not meant to reflect actual realized savings at your data center.

Potential Annual CO_2 Savings

574815900 lbs

Suggested Next Steps

Category		
Air Management	Place supply devices in cold aisles only	Perforated floor tiles or over-head supply diffusers should only be placed in the cold aisles to match the “consumption” of air by the electronic equipment. Too little or too much supply air results in poor overall thermal and/or energy conditions. Note that the hot aisles are supposed to be hot, and supplies should not be placed in those areas.
Air Management	Implement a tile/diffuser location program	A program should be in place to maintain the alternating hot and cold aisle configuration of perforated tiles or over-head diffusers. There should be no reason to place tiles or diffusers in the hot equipment aisles.
Air Management	Seal floor leaks (including cable cutouts)	A large fraction of the air from the air-handler may be lost through leaks in the raised floor. The leaks are often hidden under the equipment racks and not visible during a casual walk-through audit. Such leakage often causes by-pass air that does not contribute to cooling the electronic equipment. There are a number of commercial products that can be used to seal the raised floor.
Air Management	Use supplemental cooling (for example, high density areas)	Equipment areas with high heat densities and/or significantly higher heat densities than the average density (>8) may be prime candidates for supplemental cooling, including liquid-cooled solutions. Supplemental cooling solutions are generally best suited for controlling occasional point loads rather than a large number of racks.
Air Management	Use adequate ratio system flow to rack flow (target 1.0 or RTI=100%)	Generally, the supply airflow should closely match the equipment airflow. The Return Temperature Index (RTI) is a measure of the level of by-pass air or recirculation air in the equipment room. Both effects are detrimental to the thermal and energy performance of the data center. The target is 100% whereas >100% implies recirculation air and <100% implies by-pass air.
Air Management	Balance the air-distribution system (diffusers/tiles)	Over-head ducted systems can be adequately balanced using conventional methods whereas raised-floor systems are balanced by using “enough” perforated tiles. The latter often becomes more an art rather than science, especially since the pressure difference across the floor is small.
Air Management	Shut off CRAC/H units	If it is determined that a lower airflow volume is desired and the CRAC/CRAH units do not have variable speed fans, adjustment is limited to shutting off individual units. This is not a precise way of controlling the air volume, but it can still yield acceptable results. Some experimentation may be required to determine which units can be shut off without compromising adequate cooling of the IT equipment.

Air Management	Implement an air-balancing program	Generally, the supply flow should closely match the equipment flow. The Return Temperature Index (RTI) is a measure of by-pass air or recirculation air. Both are detrimental to the performance of the data center. The target is 100% whereas >100% implies recirculation air and <100% implies by-pass air.
Air Management	Control all fans in parallel. Add pressure sensor (under floor or in duct) for control of fans. Consider fan reset by demand.	If all the supply fans serving a given space are identical and equipped with variable speed drives, fan energy is minimized by running all the fans (including redundant units) at the same speed.
Air Management	Consider adding either an air or waterside economizer to the existing CRAH/AHU(s)	If the data center is served by cooling units that can be practically served with outside air, and there is a feasible exhaust air path, consider implementing airside economizing. In economizing mode, 100% outside air is drawn in to the data center and returned to the outdoors after one pass. This scheme will offset or even eliminate cooling compressor energy whenever the energy content of the outside air is less than the energy content of the return air. The higher the nominal return air temperature, the more viable economizing hours there will be. To ensure that summer peak electric demand is not increased due to fan energy, design for a low pressure drop intake and exhaust paths. Off-the-shelf air handlers and AC units can often be ordered with an economizer option direct from the manufacturer.
Air Management	If the existing economizer(s) have never been commissioned or have not been retrocommissioned in the past 2 years, retrocommission them	While airside economizers can offer large energy savings (particularly in milder climates), they need regular service to operate properly. The outside air sensors that control when the economizer opens and closes must be kept calibrated. The actuators and linkages that control the economizer louvers must be kept lubricated and in adjustment. The entire economizer system should be tested at least once a year to ensure it operates as intended.
Air Management	Remove abandoned cable and other obstructions from underfloor and over-head.	Under-floor and over-head obstructions often interfere with the distribution of cooling air. Such interferences can significantly reduce the air handlers' airflow as well as negatively affect the air distribution. The cooling capacity of a raised floor depends on its effective height, which can be increased by removing obstructions that are not in use.
Air Management	Implement alternating hot aisle/cold aisles	This is generally the first step towards separating hot and cold air, which is key to air management. Cold air is supplied into the cold front aisles, the electronic gear moves the air from the front to the rear and/or front to the top, and the hot exhaust air is returned to the air handler from the hot rear aisles. Some data centers are not suitable for hot/cold aisles, including those with non-optimal gear (not moving air from front to rear/top).

Air Management	Provide physical separation of hot and cold air: Provide semi-enclosed aisles (e.g., aisle end doors) Provide flexible strip curtains to enclose aisles Provide rigid enclosures to enclose aisles Use in-rack ducted exhaust	Physical barriers can successfully be used to avoid mixing the hot and cold air, allowing reduction in airflow and fan energy as well as increase in supply/return temperaturses and chiller efficiency. There are four principal ways of providing physical separation:
Air Management	Convert to VFD fans that allow variation of airflow to meet cooling demand.	This action allows variation of airflow to meet cooling demand. Traditionally, few CRAC units have the capability to vary the airflow in real time, and adjusting the supply temperature is the only option. With variable speed drives, the capacity control can be modified to improve the cooling effectiveness of the electronic equipment as well as save fan and cooling energy.
Cooling	Add VSDs to cooling tower fans	Cooling towers are typically equipped with a single-speed or a two-speed fan motor. The motor cycles on and off to maintain the desired condenser water temperature. Adding a variable speed drive (VSD) to the motor offers several advantages. It saves energy by operating continuously at a lower speed rather than cycling between a higher speed and off. It saves the wear and tear that occurs with cyclic operation, and is less noisy. And it allows more precise control of the condenser water temperature.
Cooling	Add integrated waterside economizer to plant	This action requires a water-cooled chilled water plant; i.e., a plant that includes cooling towers. During periods of low wetbulb temperature (often at night), the cooling towers can produce water temperatures low enough to precool the chilled water returning from the facility, effectively removing a portion of the load from the energy-intensive chillers. During the lowest wetbulb periods, the towers may be able to cool the chilled water return all the way down to the chilled water supply temperature setpoint, allowing the chillers to be shut off entirely. The air handlers see the same chilled water supply temperature at all times, allowing them to maintain the required temperature and humidity requirements. Free cooling also offers an additional level of redundancy by providing a non-compressor cooling solution for portions of the year.
Cooling	Recalibrate CHWS temperature sensors.	A chiller's efficiency is directly affected by the temperature of the chilled water (CHW) it is required to produce. A colder CHW supply temperature typically results in lower chiller efficiency, all other factors held equal. An out-of-calibration CHW supply temperature sensor can cause a chiller plant to produce an unnecessarily cold CHW temperature and waste energy. In addition, a too-cold CHW temperature can cause undesired dehumidification at the cooling coils. This places an extra load on the cooling system and additional energy use.

Cooling	Recalibrate CWS temperature sensors.	A water-cooled chiller's efficiency is directly affected by the temperature of the condenser water (CW) entering the condenser. A higher CW supply temperature typically results in lower chiller efficiency, all other factors held equal. An out-of-calibration CW supply temperature sensor can cause the cooling towers to produce a warmer than desired CW temperature and in turn cause the chiller plant to work unnecessarily hard.
Cooling	If the existing chillers are in poor condition or over 5 years old, evaluate them for replacement	Chillers are typically the greatest energy-using components in the cooling system. Recent advances in chiller technology, especially variable-speed compressors, offer more efficient operation. For these reasons, it is often worthwhile to examine the cost-effectiveness of replacing existing chillers if they are more than 5 years old or are in poor condition.
Cooling	Convert all 3 way valves to 2 way and close off all bypasses. Add VSD to pumps. Control pump speed to pressure. Consider reset of pressure setpoint by demand.	Older chilled water distribution systems are designed with 3-way valves at the cooling coils. A constant flow of chilled water is delivered to each coil location. Each coil is equipped with a bypass leg, and each 3-way valve modulates to divert as much water through the coil as is currently needed for cooling purposes. The remaining water bypasses the coil. This method is energy intensive. With the advent of inexpensive, reliable variable speed drives for pump motors, the preferred method is eliminate the bypasses and replace the 3-way valves with 2-way valves. The 2-way valves modulate as needed to serve the cooling load, and the pump motor speed varies in response to the demand (by maintaining a constant pressure at the far end of the distribution loop). In facilities that experience a varying load, it may be cost effective to go one step further and program the control system to vary the pressure setpoint in response to the position of the most-open 2-way valve.
Environmental Conditions	Consider increasing the supply temperature	A low supply temperature makes the chiller system less efficient and limits the utilization of economizers. Enclosed architectures allow the highest supply temperatures (near the upper end of the recommended intake temperature range) since mixing of hot and cold air is minimized. In contrast, the supply temperature in open architectures is often dictated by the hottest intake temperature.
Environmental Conditions	Place temperature/humidity sensors so they mimic the IT equipment intake conditions	IT equipment manufacturers design their products to operate reliably within a given range of intake temperature and humidity. The temperature and humidity limits imposed on the cooling system that serves the data center are intended to match or exceed the IT equipment specifications. However, the temperature and humidity sensors are often integral to the cooling equipment and are not located at the IT equipment intakes. The condition of the air supplied by the cooling system is often significantly different by the time it reaches the IT equipment intakes. It is usually not practical to provide sensors at the intake of

		every piece of IT equipment, but a few representative locations can be selected. Adjusting the cooling system sensor location in order to provide the air condition that is needed at the IT equipment intake often results in more efficient operation.
Environmental Conditions	Network the CRAC/CRAH controls	CRAC/CRAH units are typically self-contained, complete with an on-board control system and air temperature and humidity sensors. The sensors may not be calibrated to begin with, or they may drift out of adjustment over time. In a data center with many CRACs/CRAHs it is not unusual to find some units humidifying while others are simultaneously dehumidifying. There may also be significant differences in supply air temperatures. Both of these situations waste energy. Controlling all the CRACs/CRAHs from a common set of sensors avoids this.
Environmental Conditions	Add personnel and cable grounding to allow lower IT equipment intake humidities	The lower humidity limit in data centers is often set relatively high (40% RH at the IT equipment intake is common) to guard against damage to the equipment due to electrostatic discharge (ESD). Maintaining this level of humidity is energy intensive if the humidifiers use electricity to make steam (this is the most common type). Energy can be saved if the allowed lower humidity limit can be lowered, particularly if the cooling system has an airside economizer. ESD can be kept in check by conductive flooring materials, good cable grounding methods, and providing grounded wrist straps for technicians to use while working on equipment.
Environmental Conditions	Consider disabling or eliminating humidification controls or reducing the humidification setpoint	Tightly controlled humidity can be very costly in data centers since humidification and dehumidification are involved. A wider humidity range allows significant utilization of free cooling in most climate zones by utilizing effective air-side economizers. In addition, open-water systems are high-maintenance items.
Environmental Conditions	Consider disabling or eliminating dehumidification controls or increasing the dehumidification setpoint	Most modern IT equipment is designed to operate reliably when the intake air humidity is between 20% and 80% RH. However, 55% RH is a typical upper humidity level in many existing data centers. Maintaining this relatively low upper limit comes at an energy cost. Raising the limit can save energy, particularly if the cooling system has an airside economizer. In some climates it is possible to maintain an acceptable upper limit without ever needed to actively dehumidify. In this case, consider disabling or removing the dehumidification controls entirely.
Global	Consider upgrading all cooling supply fan, pump, and cooling tower fan motors to premium efficiency.	Premium efficiency motors are generally a few percent more efficient than their baseline counterparts. The efficiency gains are modest, but the incremental first cost tends to be low as well, especially when replacing existing motors that have reached the end of their service life. Specifying a premium efficiency motor is almost always cost effective for applications with long or continuous runtimes.

IT Equipment	Evaluate the potential savings from upgrading to newer equipment.	IT technology evolves rapidly, and improvements in energy performance are often provided in newer equipment. A cost-benefit analysis will reveal when it makes economic sense to replace existing equipment.
IT Equipment	Consider consolidating to network-attached (NAS or SAN) storage and using diskless servers.	Servers typically have on-board mechanical disk drives. These drives are responsible for a significant percentage of the server's total energy use, but they often have a low utilization rate. Converting to solid-state memory at the servers, or consolidating to a network-attached (NAS or SAN) data storage device may be a path to an effective energy performance improvement.
IT Equipment	Assess storage usage and move less performance-sensitive data to higher capacity, more efficient media.	It is not uncommon to have more storage allocated to processing tasks than is needed, and to have the storage accessed infrequently. This can result in poor energy performance, as storage devices draw energy whether they are in active use or not. Investigating data storage utilization patterns can reveal opportunities, such as moving less performance-sensitive data to higher capacity, more efficient media.
IT Equipment Power Chain	If existing UPS is older than 10 years, retrofit UPS topologies for more efficient ones	UPS technology continues to evolve. If the existing UPS is scheduled for replacement, be sure to specify a high-efficiency UPS topology. If the existing UPS more than 10 years old it may be cost-effective to replace it with a new system right away.
IT Equipment Power Chain	Standby Generator block heater / heater water jacket(s) (HWJ) operate with thermostat control	In many areas of the country the engine blocks of the emergency backup generators are kept warm with electric resistance heat to help promote rapid, reliable starting. Often these heaters are very simple devices that provide continuous heat without any thermostat control. Adding a thermostat will help minimize the electric use of the heater.
IT Equipment Power Chain	Change UPS DC capacitors if older than 5 years	The DC capacitors in typical UPS systems tend to lose effectiveness over time. This can result in the inverter failing to operate under load, and increased ripple current in the batteries. Not only does this result in less efficient operation, it becomes a safety issue as well. The DC capacitors usually have the same design lifetime as the batteries; approximately 5 years. The capacitors should be checked regularly.
IT Equipment Power Chain	Shut Down UPS Modules, Stand-by Generators, PDUs when Redundancy Level is High Enough	In some facilities, the array of UPS modules and/or PDUs has more than enough capacity to serve the load. It may be possible to shut down some modules and still retain the required level of redundancy. This will allow the remaining units to operate at a higher load factor, which usually translates to higher efficiency.
Lighting	Install Occupancy Sensors to Control Lights	Many data centers are unoccupied for long periods of time. Controlling the data center lights with occupancy sensors

directly saves lighting energy. This also reduces the heat load, saving cooling system energy.

Revisions

7/11/2011

- Future cabinet layout updated with supply grille locations
- Future cabinets updated with specific loads including blade equipment
- Blanking plates replace doors on future cabinets with zero load

CFD Modeling Objectives

- Improve data center efficiency and effectiveness
- Develop optimized airflow and control strategies

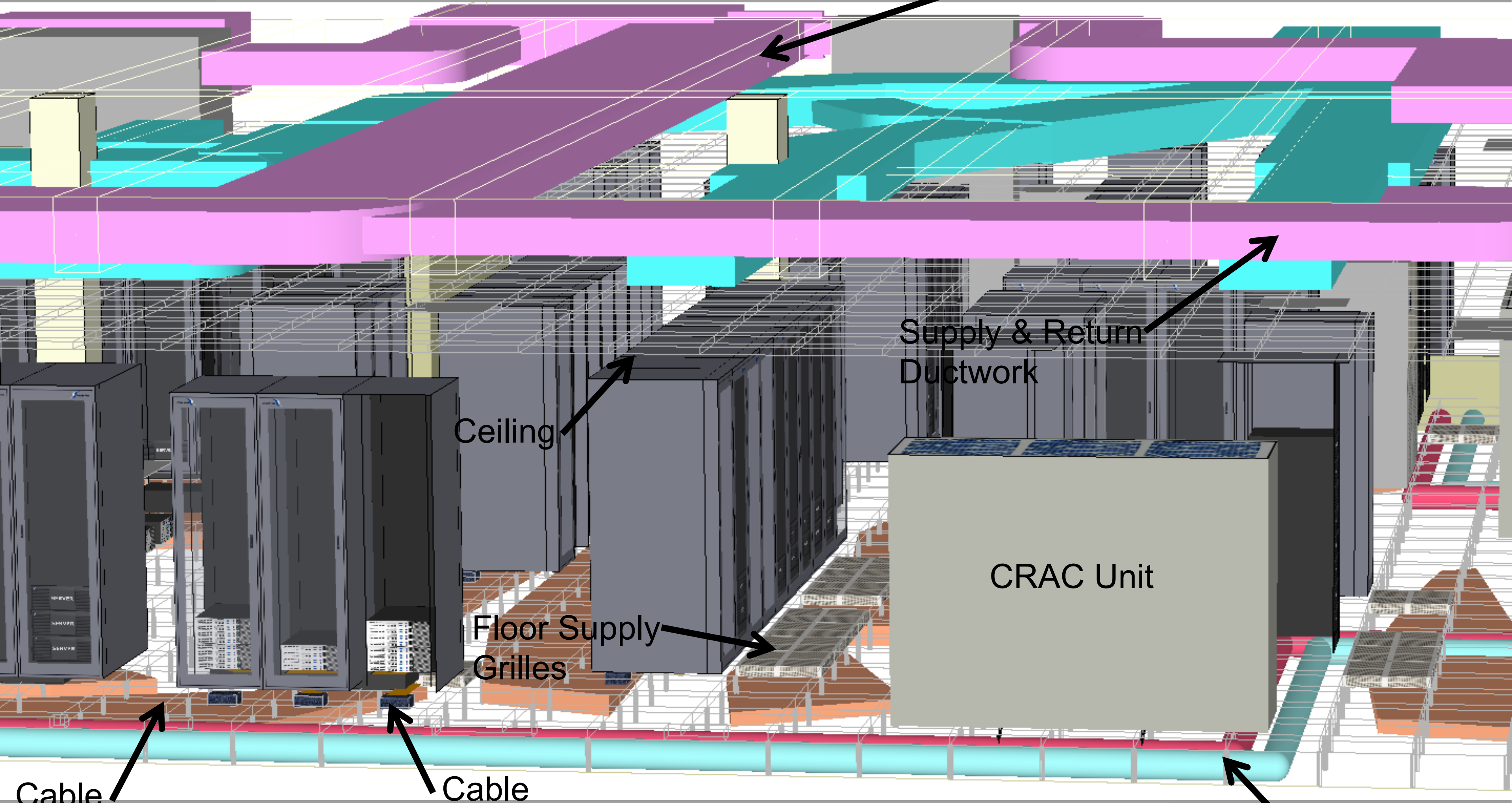
CFD Modeling Assumptions

- All CRAC units operate independently to maintain underfloor air pressure
- IT loads based upon existing readings
- All iterations assume existing floor holes and cable penetrations have been sealed and cabinets are fully constructed
- All iterations use blanking plates in areas where columns conflict with aisles
- Iteration IT load based upon spreadsheet provided by Columbia
- Cold aisle containment constructed using vinyl partitions

Modeling Scenarios

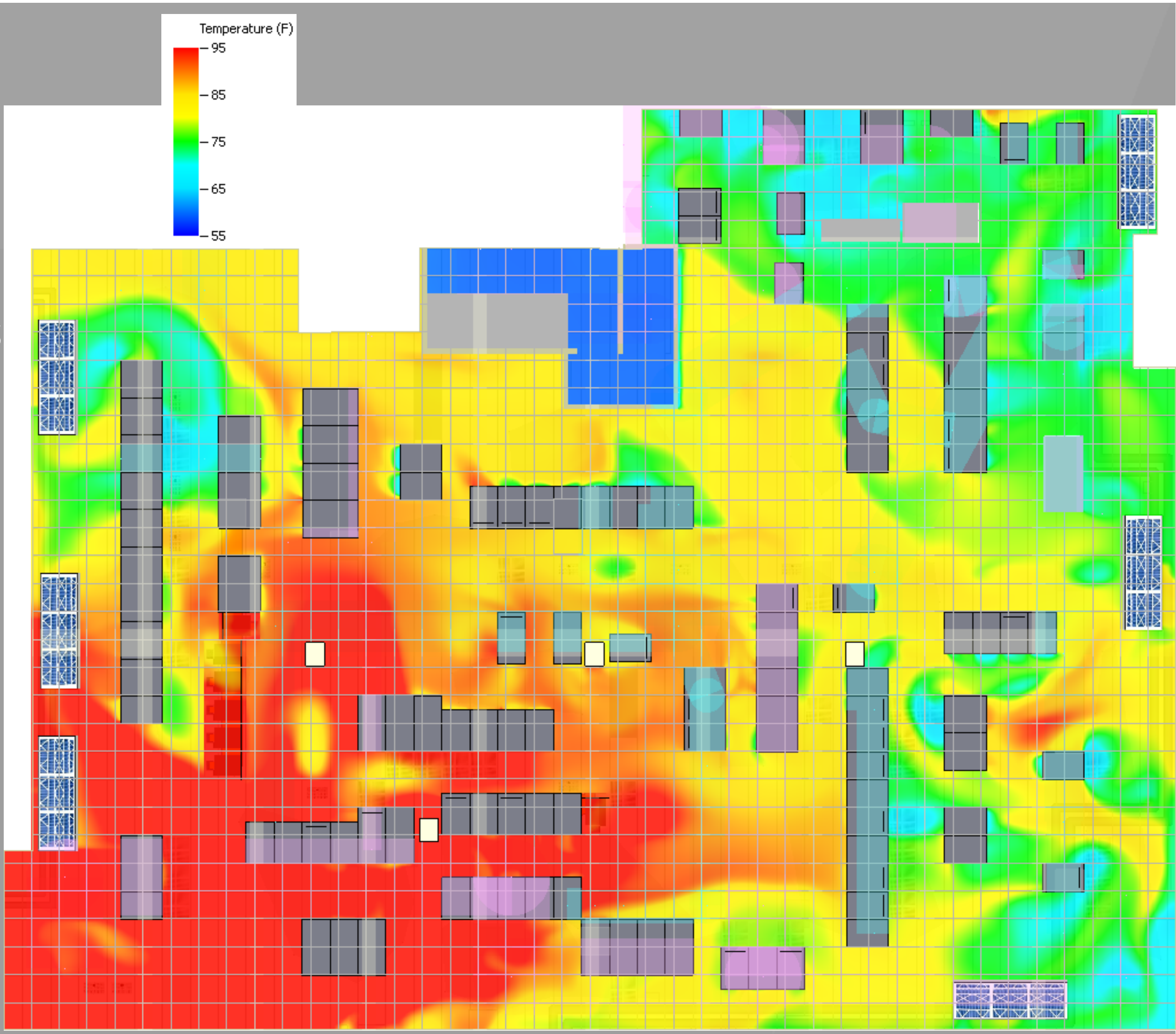
- Baseline - Existing data center layout
- Iteration #1 - Future layout with reconfigured hot and cold aisles
- Iteration #2 - Future layout and ducted CRAC units to ceiling plenum
- Iteration #3 - Future layout, ducted CRACs, and cold aisle containment

Baseline Model Geometry



Results – Baseline

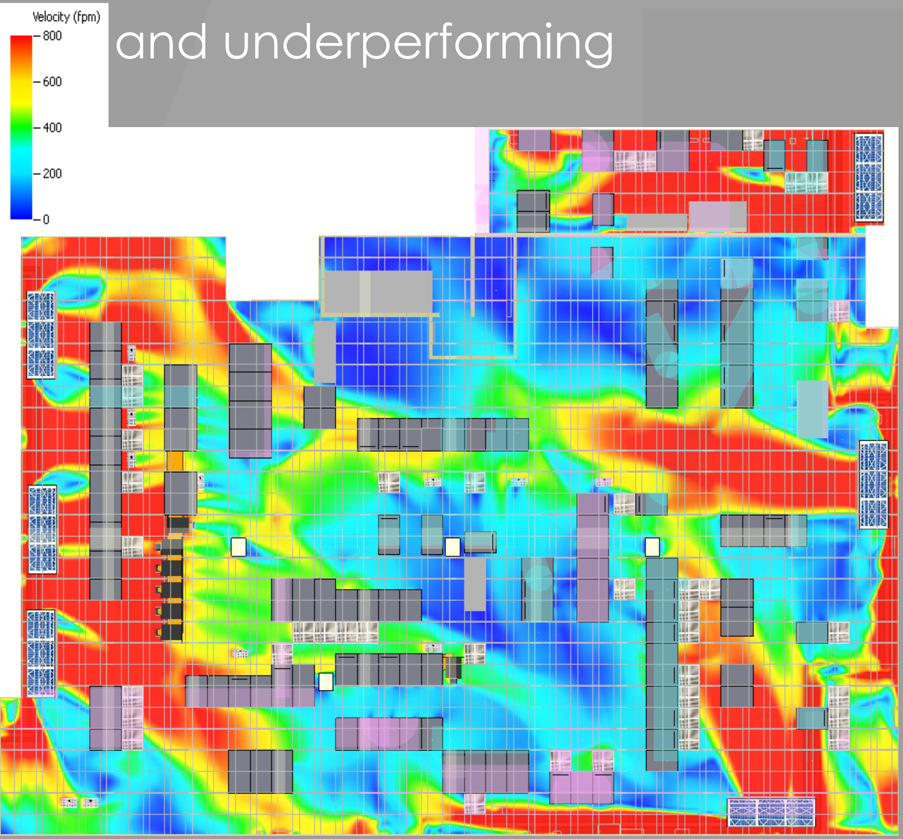
- Temperature profile at 6'-0" shows heat load returning to CRAC units through other equipment.
- Cold spots and hot spots define problem areas.



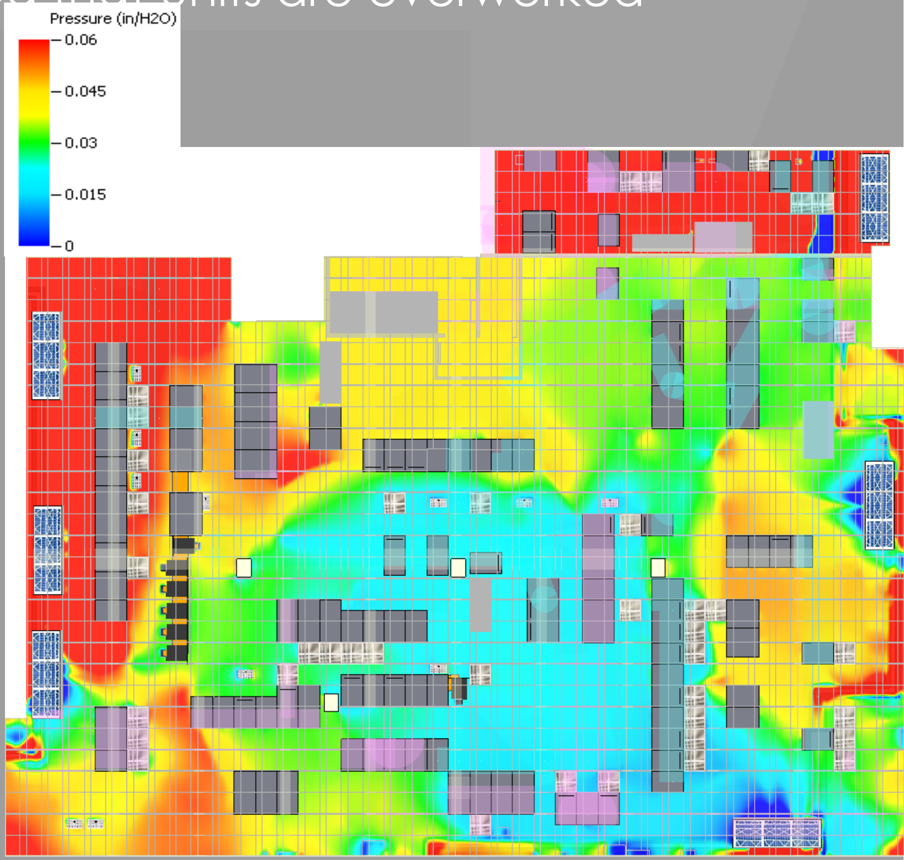
Temperature Profile @ 6'-0"

Results – Baseline

- Underfloor cabling chokes off airflow to the middle of the room
- Pressure highest at CRAC units indicated that units are overworked and underperforming



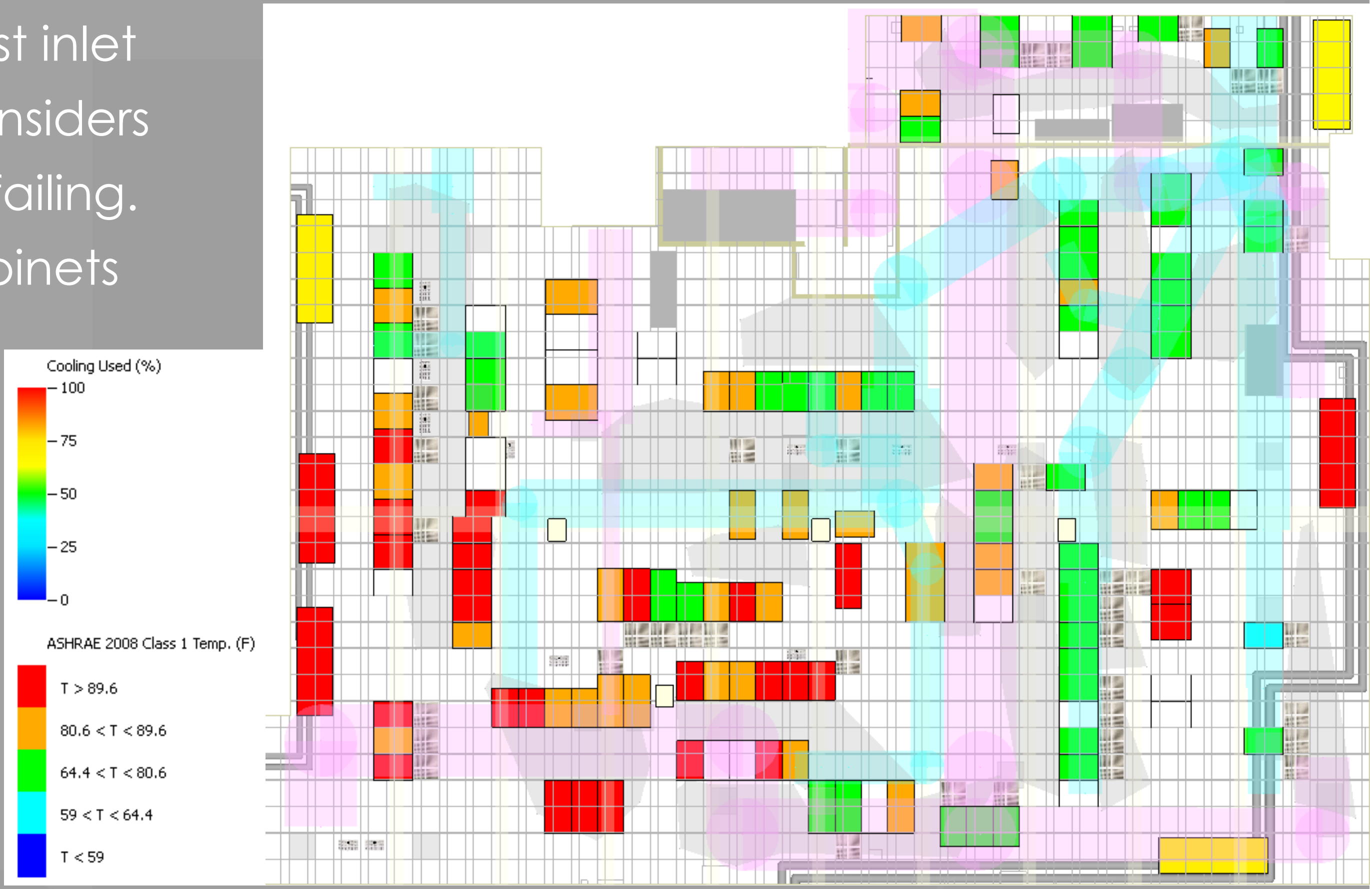
Velocity Profile @ -4”



Pressure Profile @ -4”

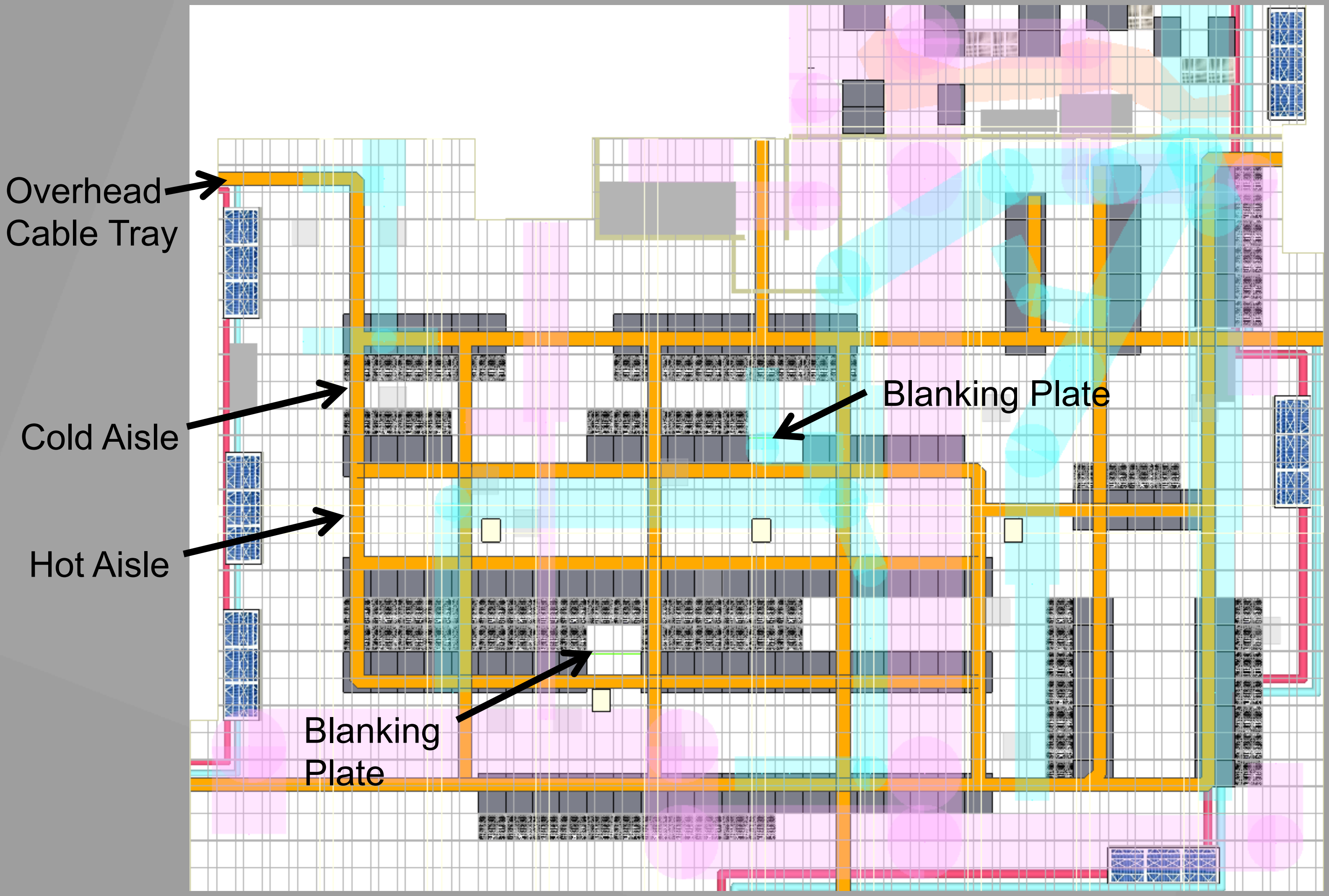
Results – Baseline

- ASHRAE Cabinet Compliance determines the highest inlet temperatures and considers anything over 80F as failing. Approximately 56 cabinets fail.
- Cabinets in areas of hot air recirculation tend to fail ASHRAE compliance.
- Cooling percentage is highest with warmest return air temperatures.



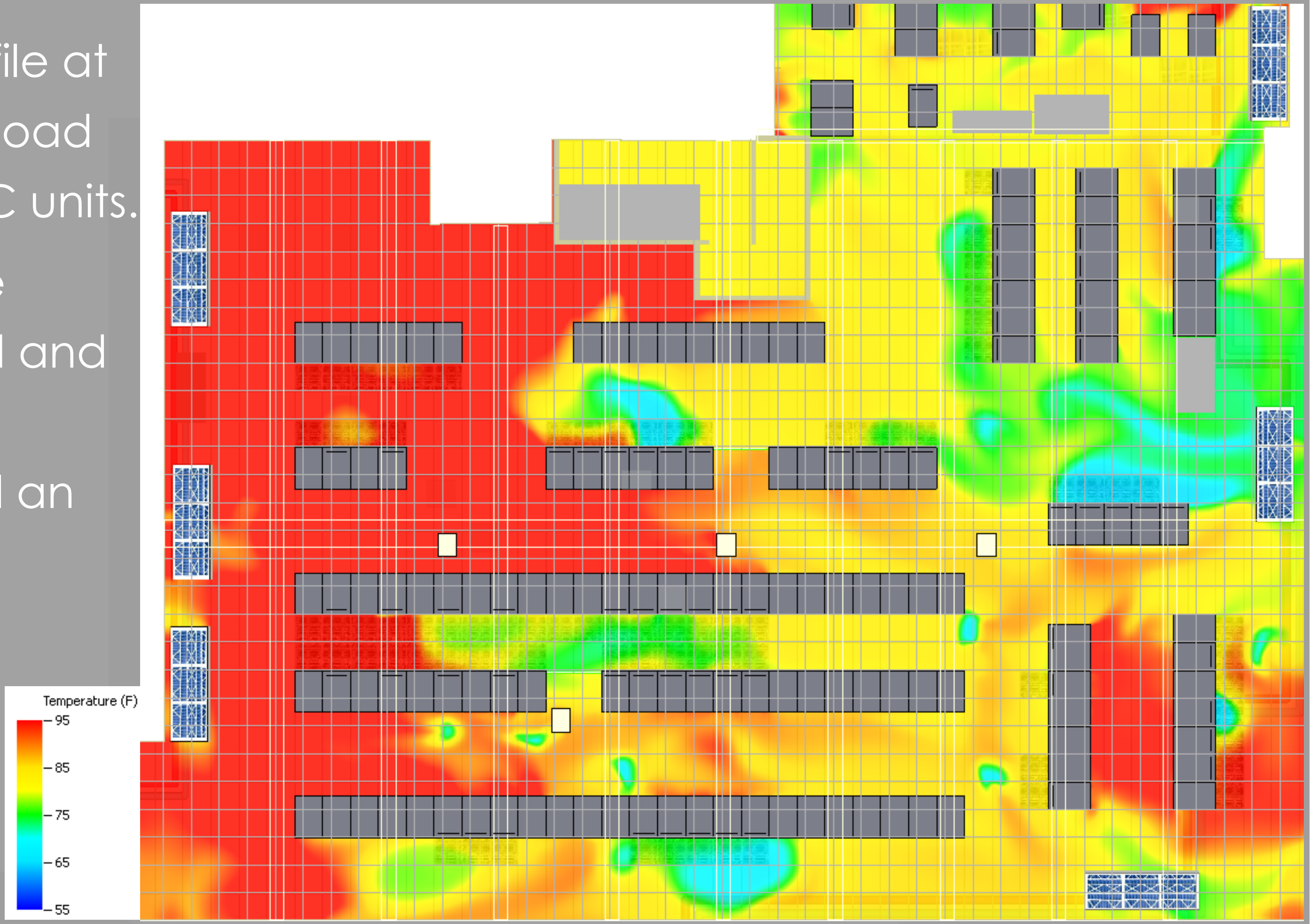
CRAC Cooling % / ASHRAE Cabinet Compliance

Iteration #1 Room Geometry



Results – Iteration #1

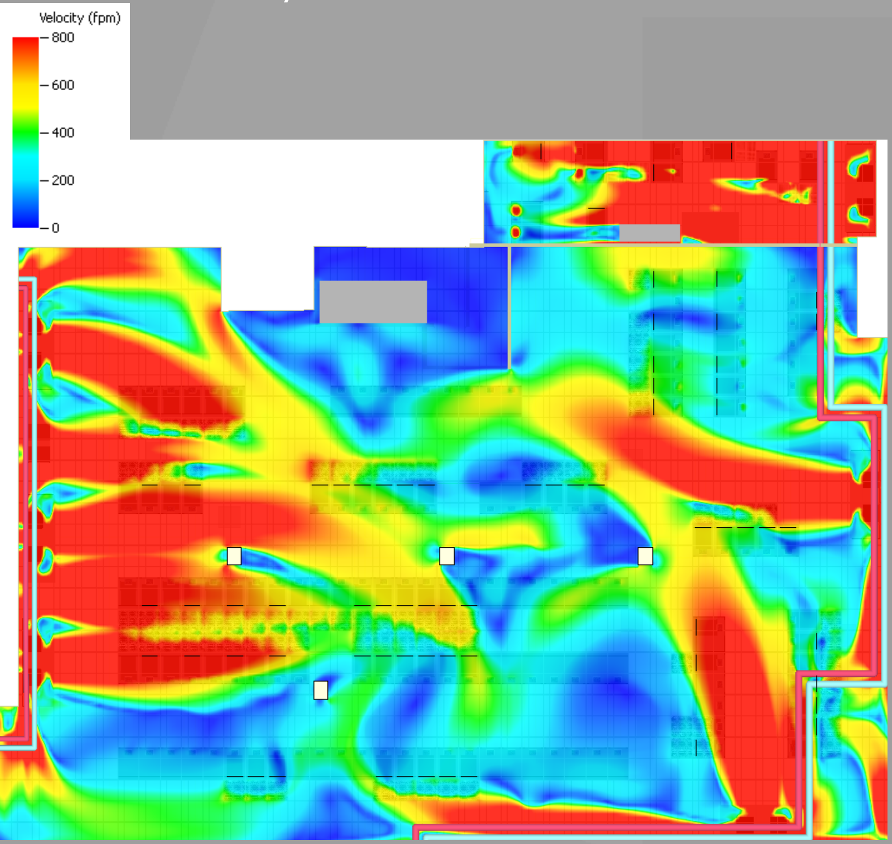
- Temperature profile at 6'-0" shows heat load returning to CRAC units.
- Heat load is more evenly distributed and controlled but recirculation is still an issue.



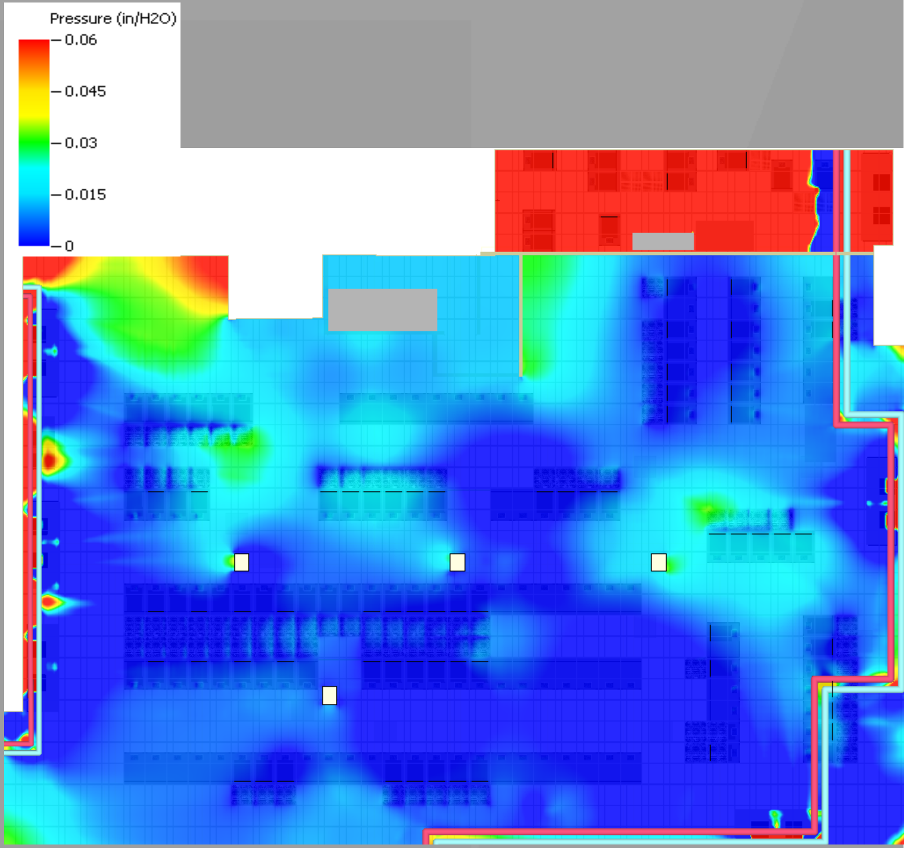
Temperature Profile @ 6'-0"

Results – Iteration #1

- Removal of underfloor cabling has improved both velocity and pressure
- Velocity streams are well-defined and pressure is consistent.



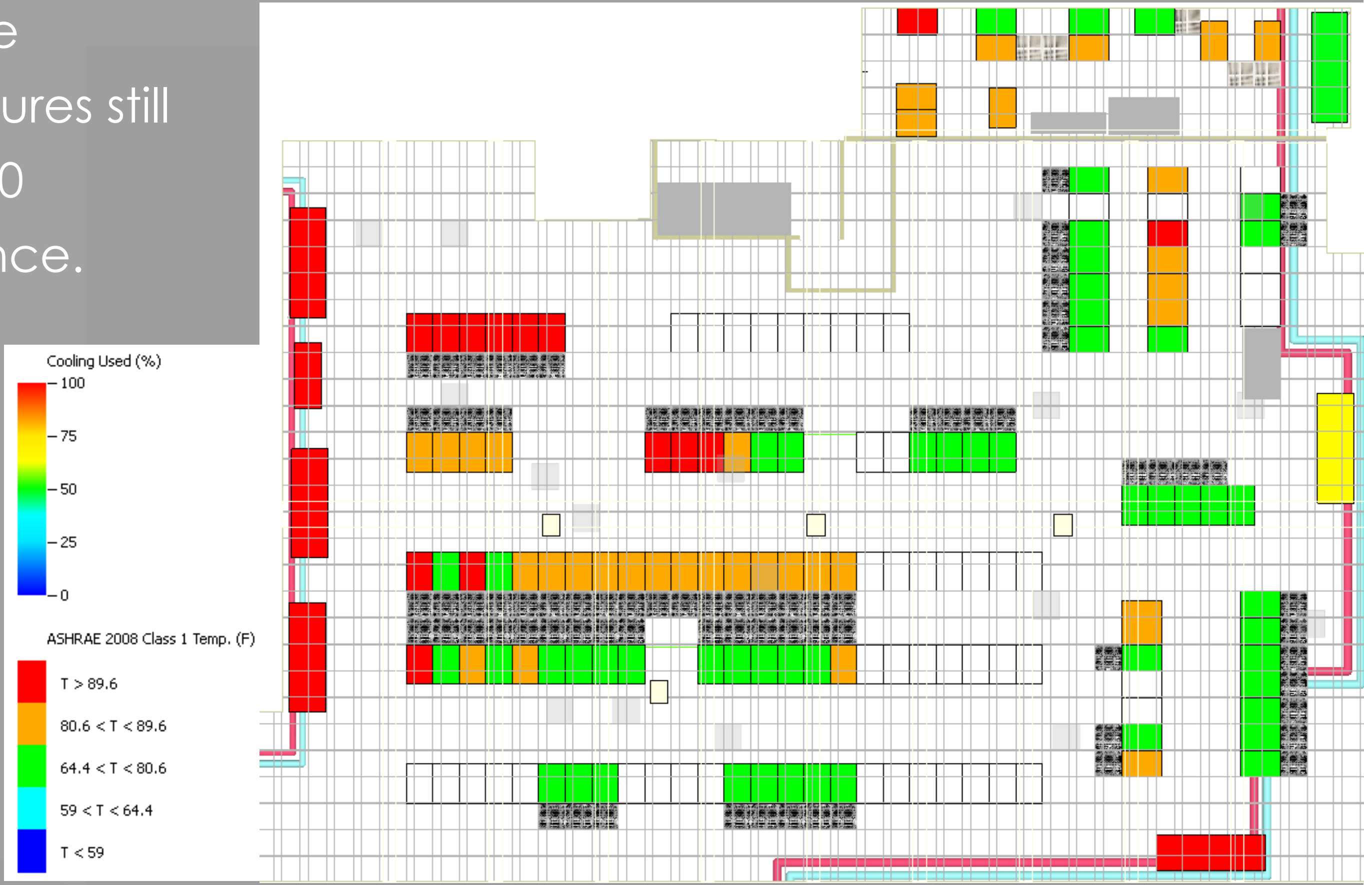
Velocity Profile @ -4”



Pressure Profile @ -4”

Results – Iteration #1

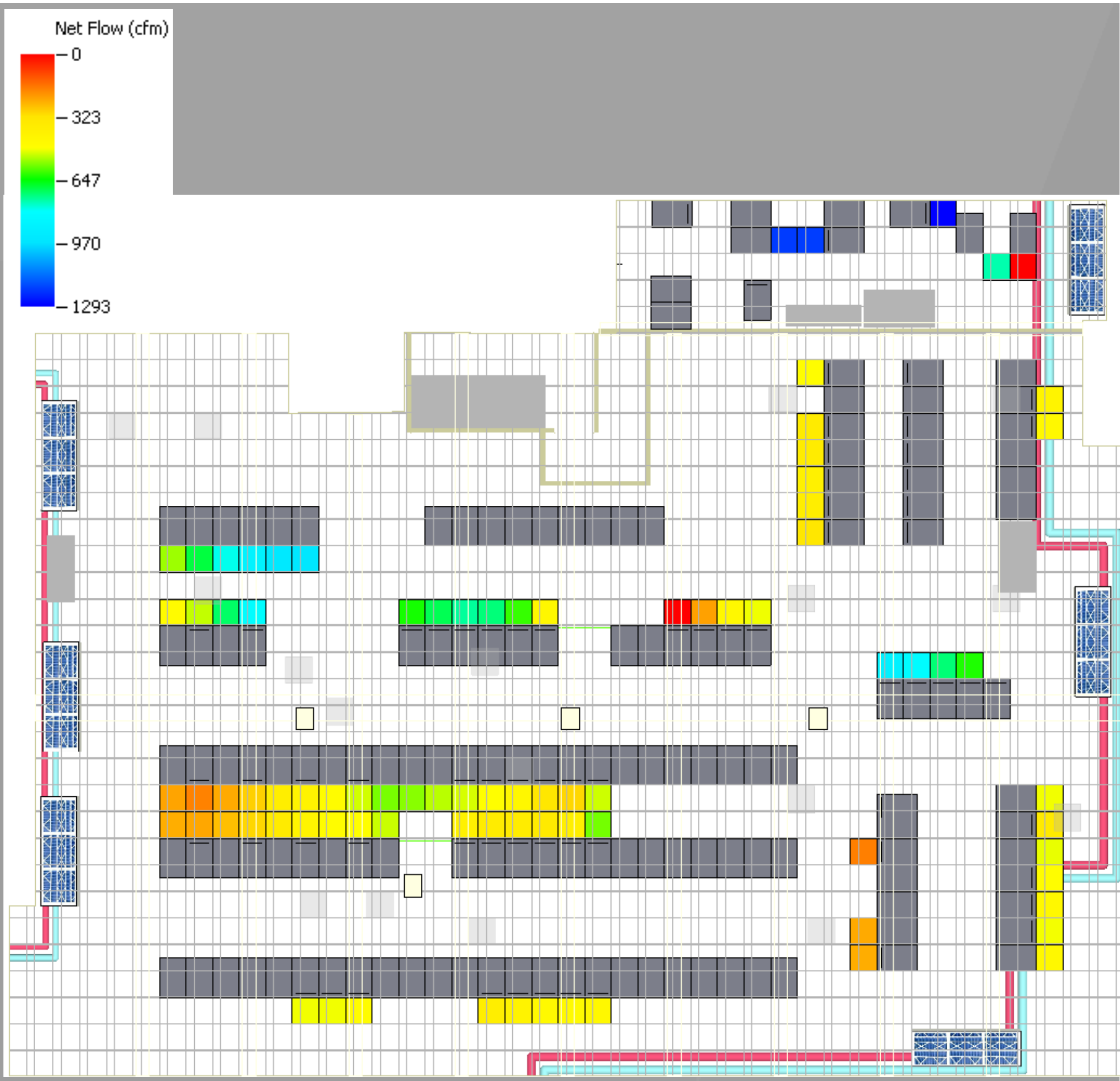
- ASHRAE Cabinet Compliance has improved from the Baseline but many failures still exist. Approximately 40 cabinets fail compliance.
- Cabinets without dedicated supply air grilles are most likely to fail compliance.



CRAC Cooling % / ASHRAE Cabinet Compliance

Results – Iteration #1

- Supply grille net flow is a good indication of the quantity of air delivered to cabinets in different areas of the space.
- High density cabinets should be located in peak airflow areas.

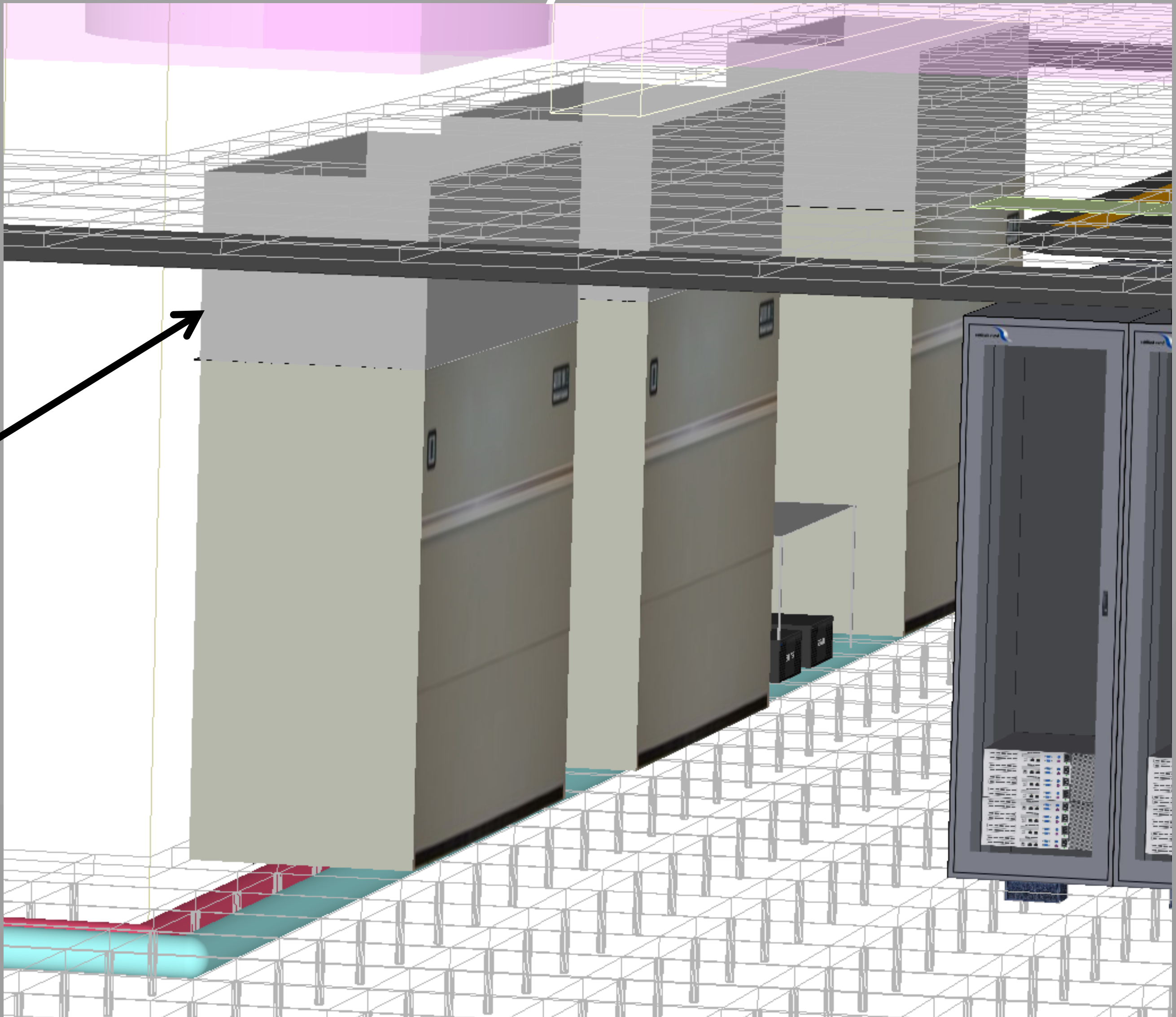


Supply Grille Net Flow

Conclusions – Iteration #1

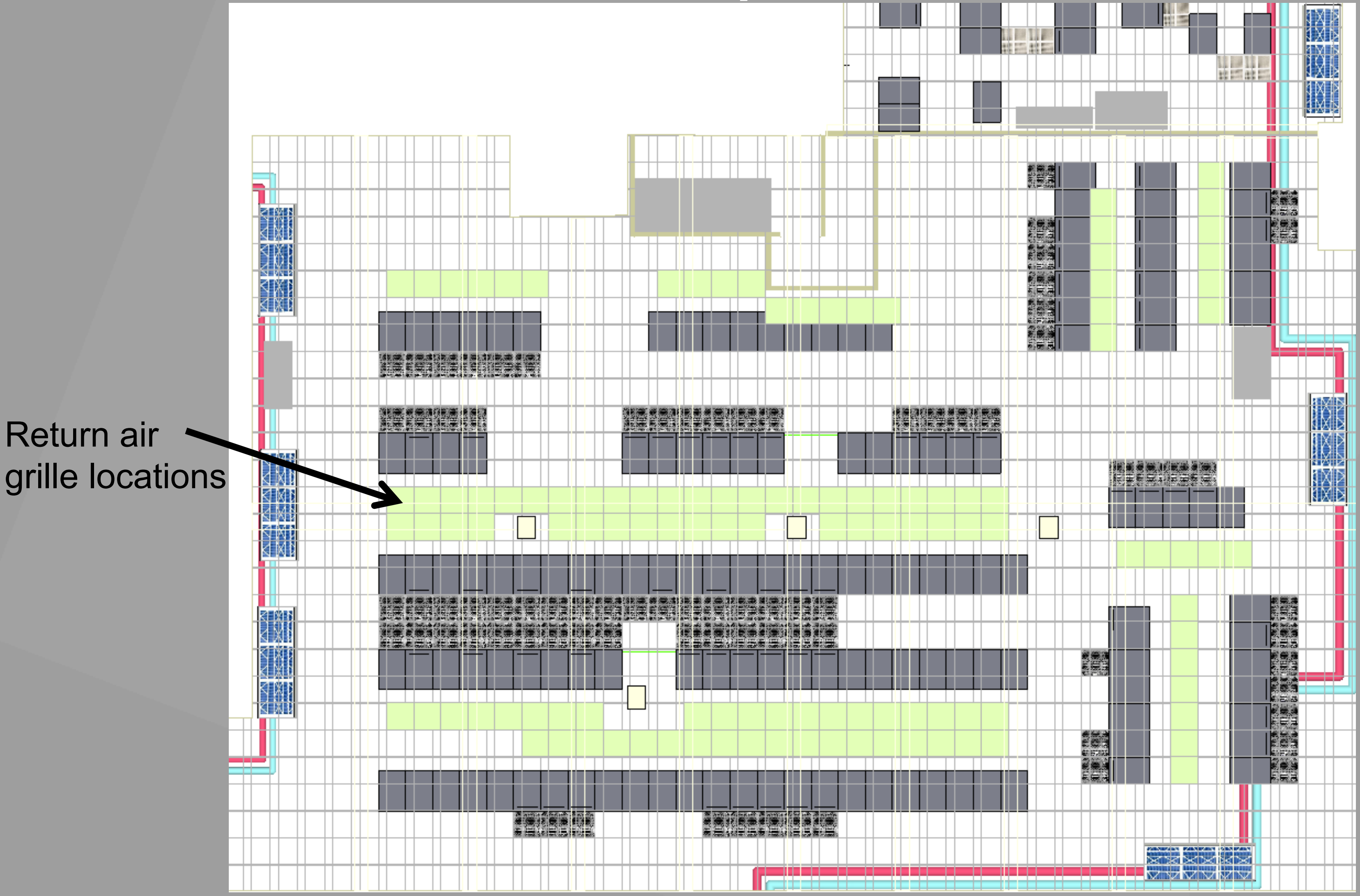
- Clearing the underfloor void of all cabling demonstrates the most dramatic improvement to airflow. Underfloor air pressure does not peak in pockets which in turn delivers air more evenly to all supply air grilles.
- Hot and cold aisle row configuration improves upon the return air temperatures at the CRAC units but does not solve all recirculation issues.

Iteration #2 Room Geometry



Ducted return air to ceiling void

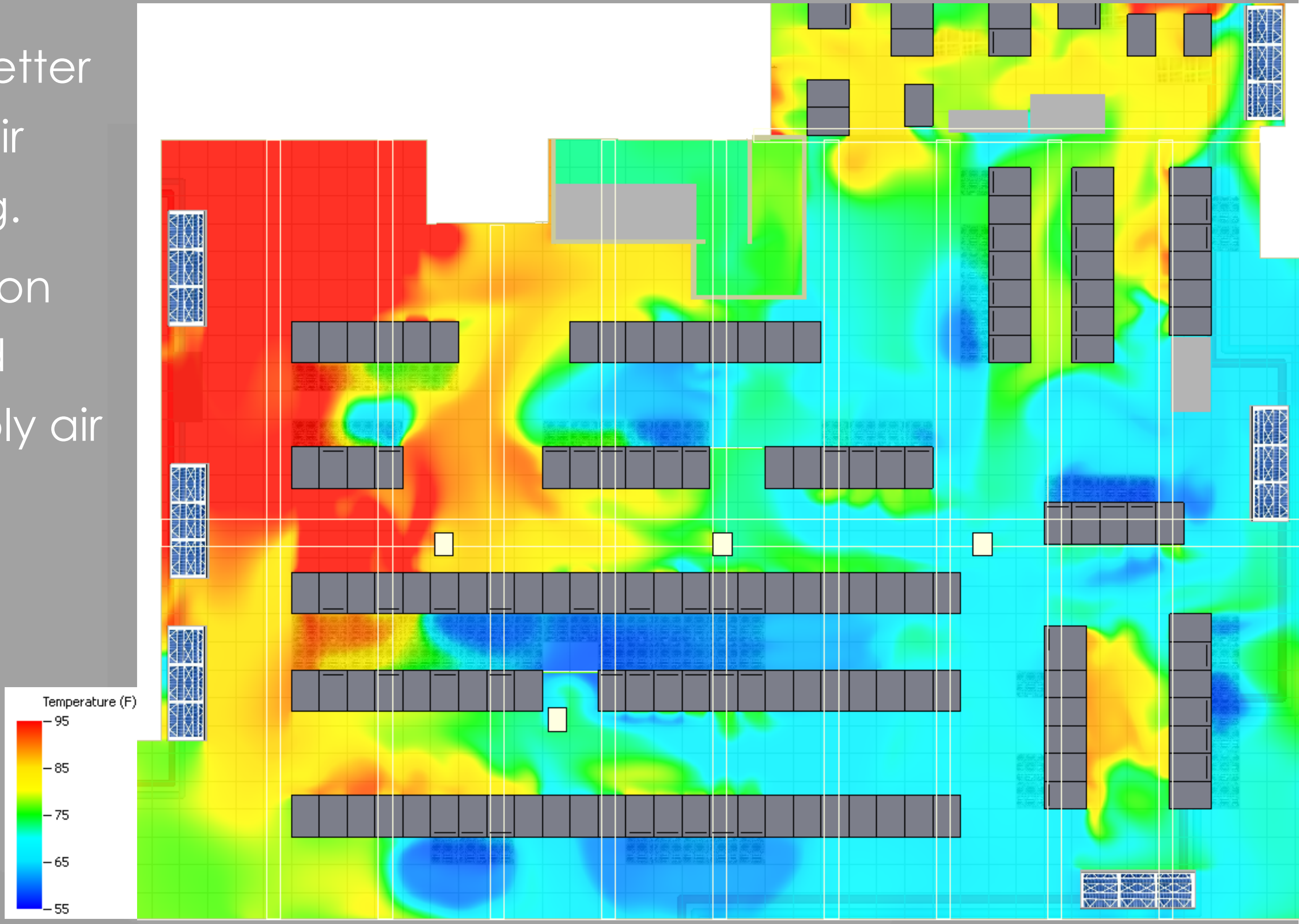
Iteration #2 Room Geometry



Return air grille locations

Results – Iteration #2

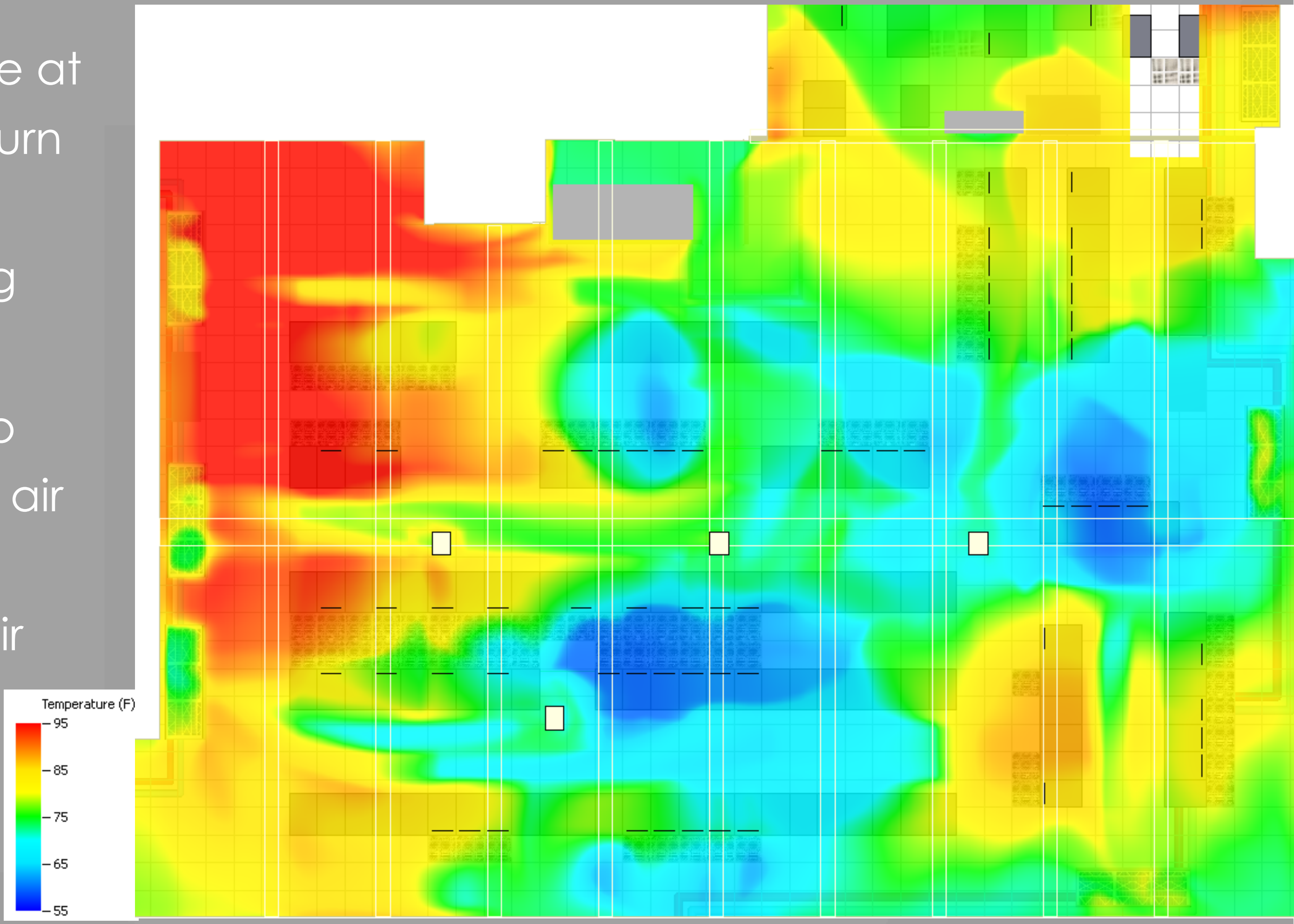
- Hot aisle does a better job returning the air through the ceiling.
- Minimal recirculation only occurs in cold aisles without supply air grilles.



Temperature Profile @ 6'-0"

Results – Iteration #2

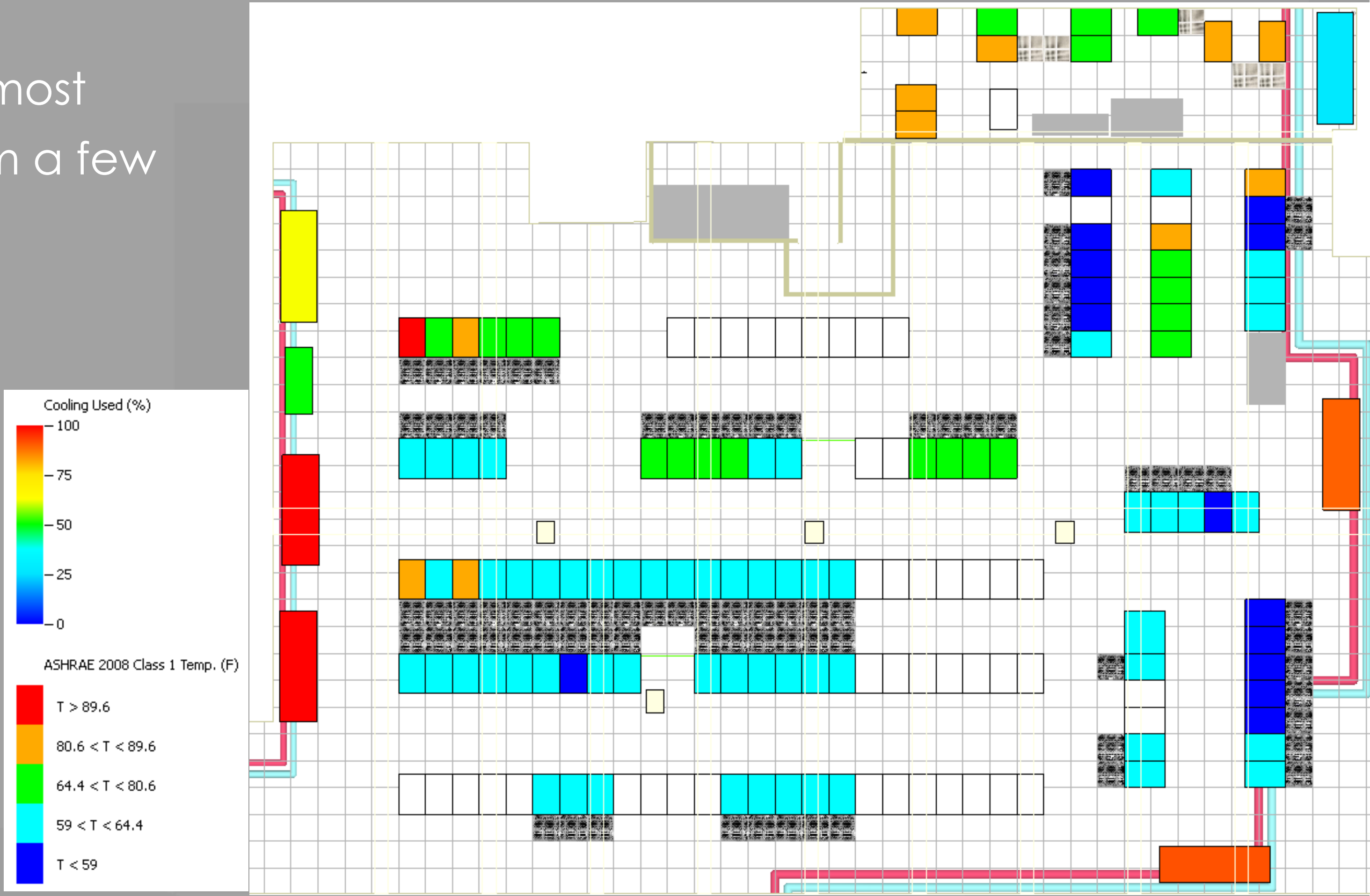
- Temperature profile at 8’-6” shows the return air temperature entering the ceiling void.
- This can be used to improve the return air temperature by relocating return air grilles for a best fit.



Temperature Profile @ 8’-6”

Results – Iteration #2

- ASHRAE Cabinet Compliance is almost perfect aside from a few cabinets.
- 6 cabinets fail compliance.

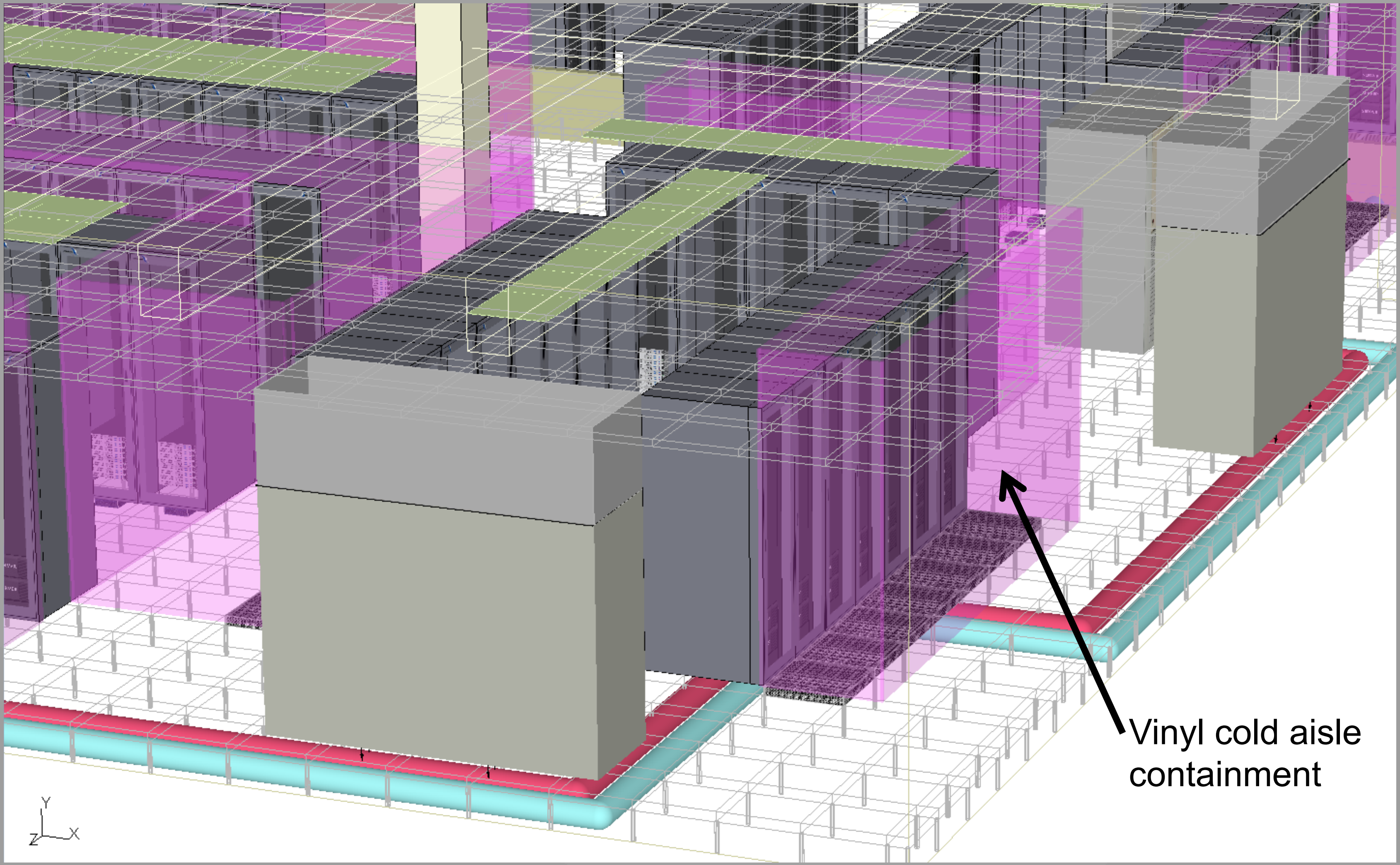


CRAC Cooling % / ASHRAE Cabinet Compliance

Conclusions – Iteration #2

- Ducting the return air of the CRAC units to the ceiling plenum will improve return air temperatures at the unit if consideration is taken for the placement of all return air grilles.
- This design route decreases recirculation at the cabinet level more than any other benefits. This is important to ensure the safety of all equipment.

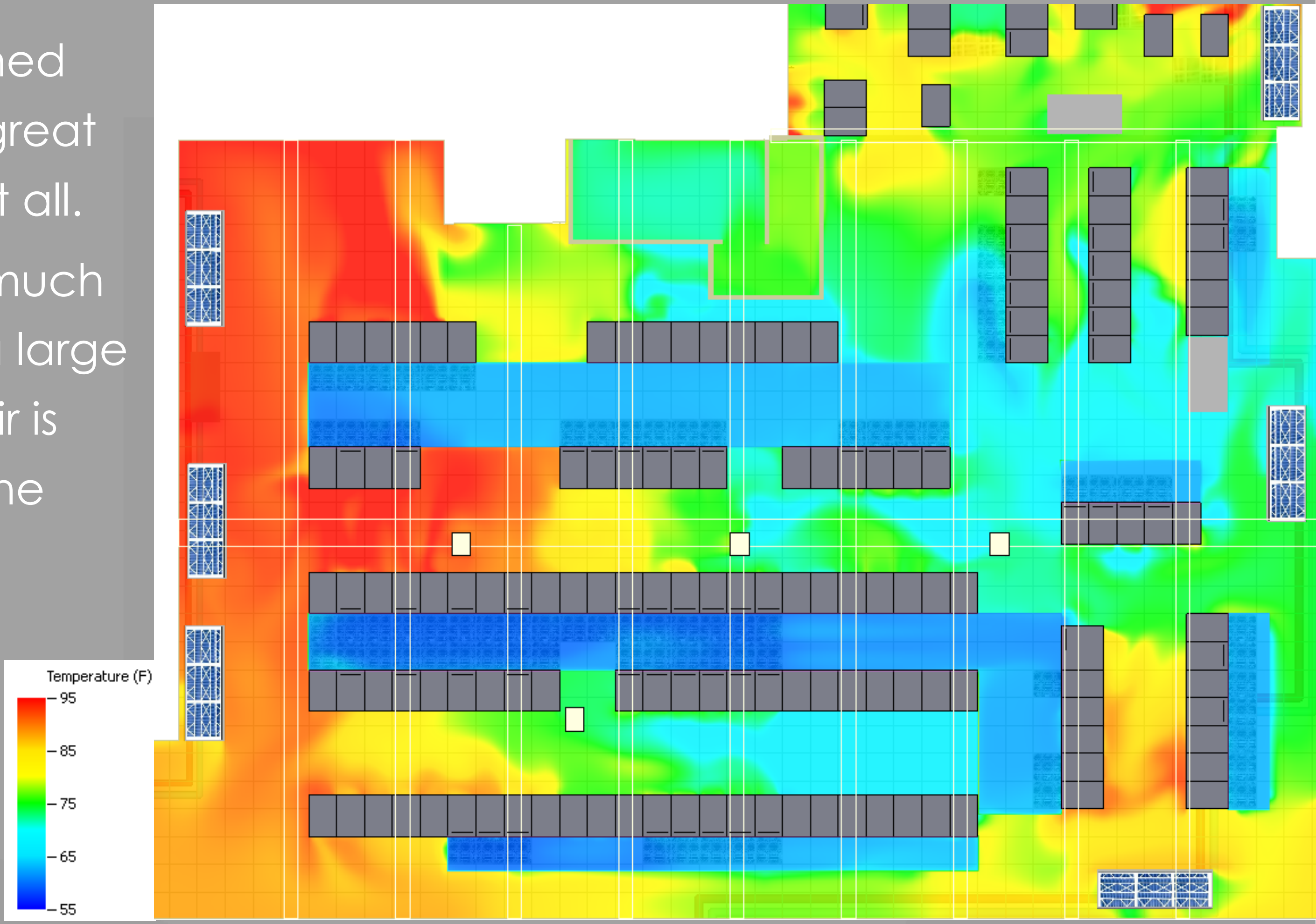
Iteration #3 Room Geometry



Vinyl cold aisle containment

Results – Iteration #3

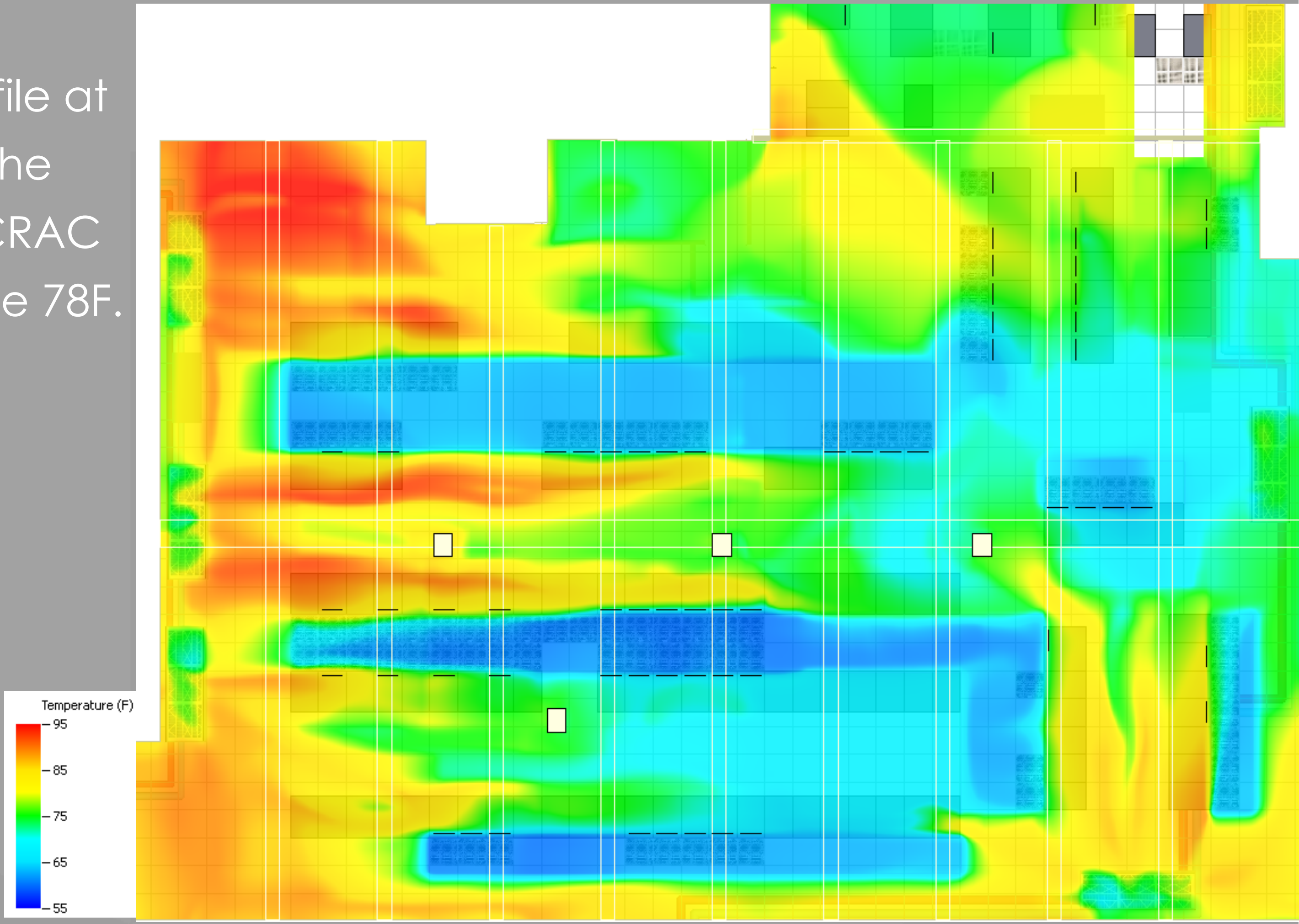
- Cold air is contained but IT load is not great enough to utilize it all.
- Return aisles are much cooler because a large amount of cold air is passing through the cabinets.



Temperature Profile @ 6'-0"

Results – Iteration #3

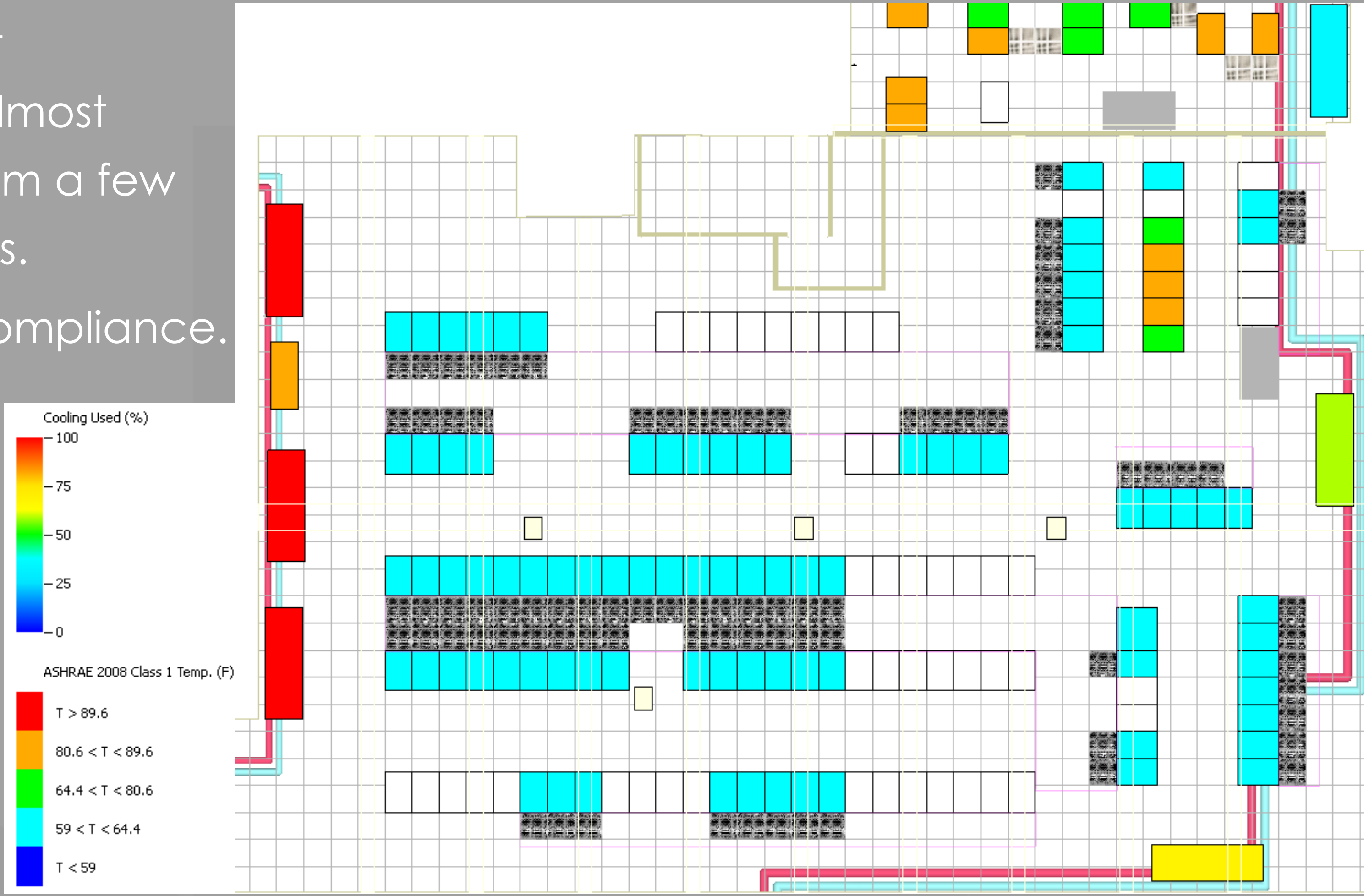
- Temperature profile at 8’-6” shows that the return air to the CRAC units is on average 78F.



Temperature Profile @ 8’-6”

Results – Iteration #3

- ASHRAE Cabinet Compliance is almost perfect aside from a few network cabinets.
- 3 cabinets fail compliance.



CRAC Cooling % / ASHRAE Cabinet Compliance

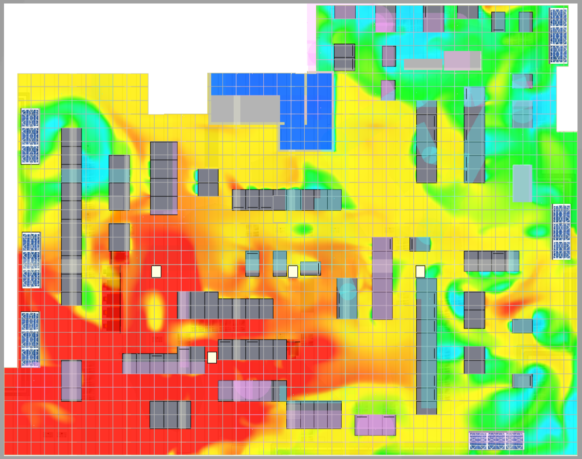
Conclusions – Iteration #3

- Cold aisle containment ensures delivery of cold air to the equipment without the chance of hot air recirculation.
- Containment is the ultimate in air efficiency.

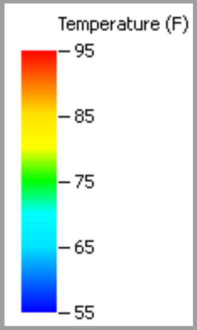
Overall Conclusions / Recommendations

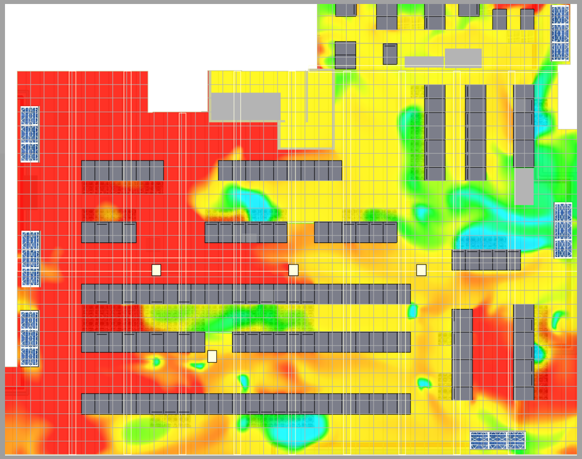
- Ducted CRAC units without containment work best for this data center.
- Containment will ensure optimal performance.
- Above ceiling ductwork will add to the static pressure on the CRAC fans when the return is ducted. Removing as much unnecessary ductwork as possible will save fan energy.
- Removing data and power cabling below the raised floor improves the delivery of the cold air the best.
- Hot and cold aisle reconfiguration is in the best interest of the equipment.

Overall Conclusions / Recommendations

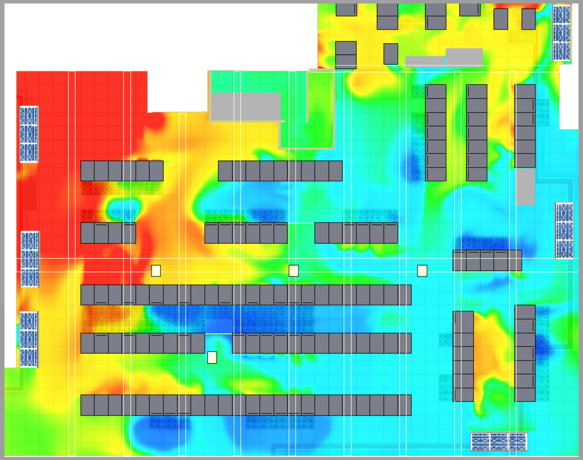


Baseline – Temperature Profile @ 6’-0”

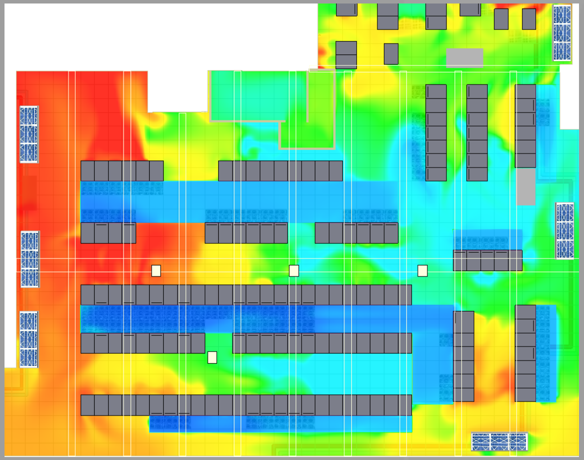




Iteration #1 – Temperature Profile @ 6’-0”

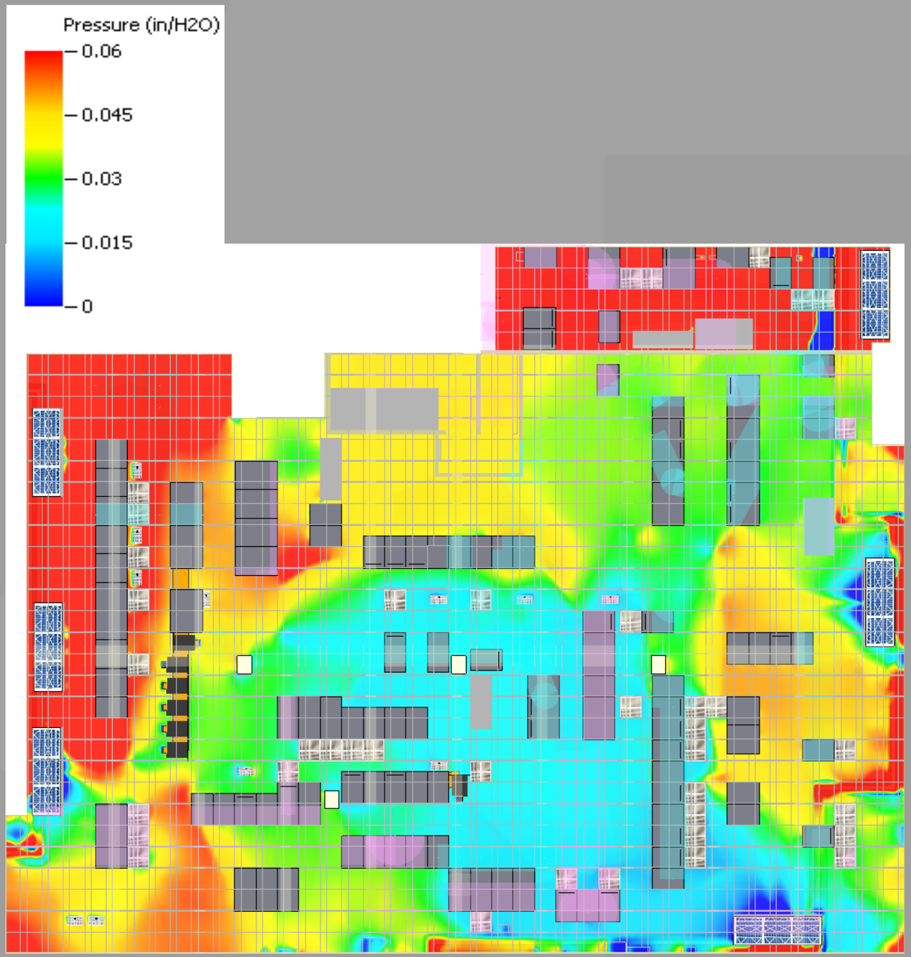


Iteration #2 – Temperature Profile @ 6’-0”

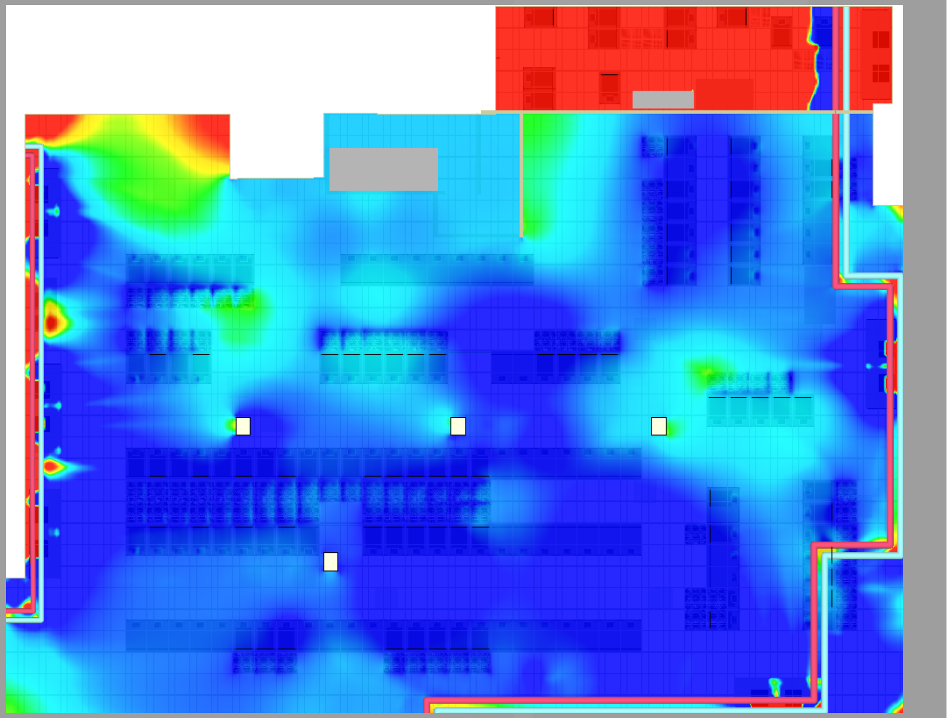


Iteration #3 – Temperature Profile @ 6’-0”

Overall Conclusions / Recommendations

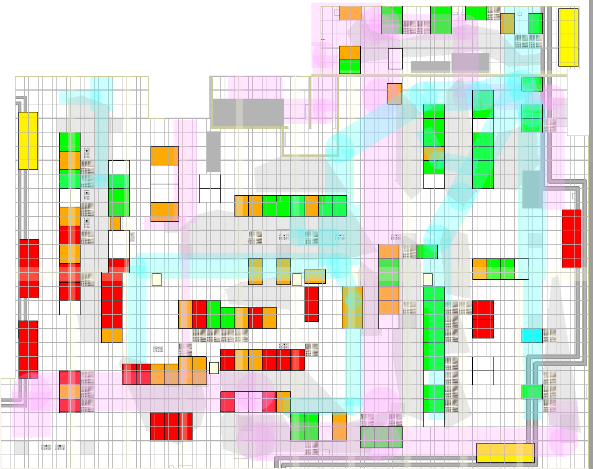


Baseline – Pressure Profile @ -4”

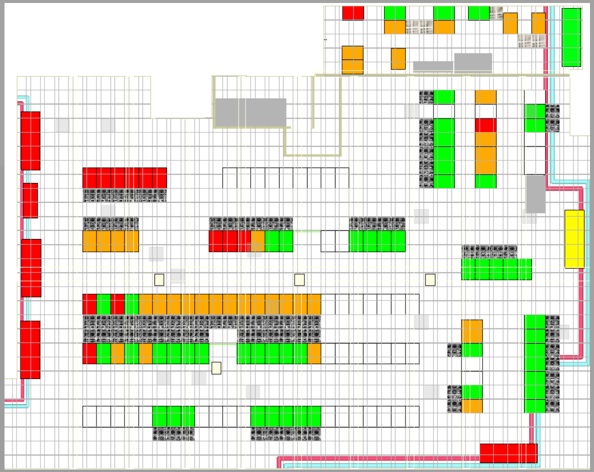


Iteration #1 – Pressure Profile @ -4”

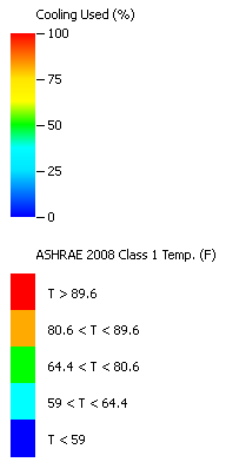
Overall Conclusions / Recommendations

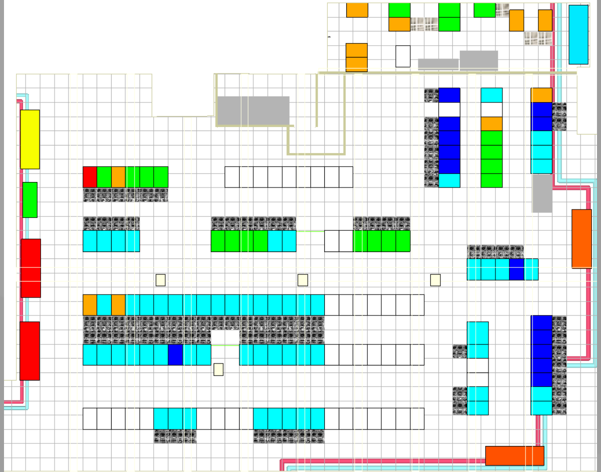


Baseline – CRAC Cooling % / ASHRAE Cabinet Compliance

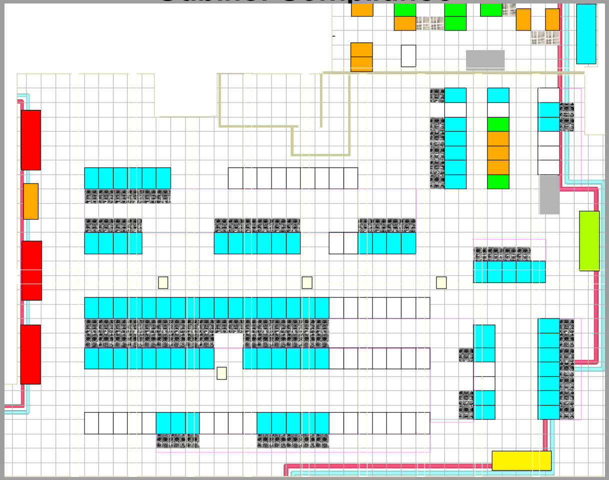


Iteration #1 – CRAC Cooling % / ASHRAE Cabinet Compliance





Iteration #2 – CRAC Cooling % / ASHRAE Cabinet Compliance



Iteration #3 – CRAC Cooling % / ASHRAE Cabinet Compliance

Questions?