

5-2012

Thermally Aware, Energy-Based Techniques for Improving Data Center Energy Efficiency

Dustin W. Demetriou
Syracuse University

Follow this and additional works at: http://surface.syr.edu/mae_etd



Part of the [Mechanical Engineering Commons](#)

Recommended Citation

Demetriou, Dustin W., "Thermally Aware, Energy-Based Techniques for Improving Data Center Energy Efficiency" (2012).
Mechanical and Aerospace Engineering - Dissertations. Paper 65.

This Dissertation is brought to you for free and open access by the College of Engineering and Computer Science at SURFACE. It has been accepted for inclusion in Mechanical and Aerospace Engineering - Dissertations by an authorized administrator of SURFACE. For more information, please contact surface@syr.edu.

Abstract

This work investigates the practical implementation of so-called thermally aware, energy optimized load placement in air-cooled, raised floor data centers to reduce the overall energy consumption, while maintaining the reliability of the IT equipment. The work takes a systematic approach to modeling the data center's airflow, thermodynamic and heat transfer characteristics – beginning with simplified, physics-inspired models and eventually developing a high-fidelity, experimentally validated thermo-hydraulic model of the data center's cooling and power infrastructure. The simplified analysis was able to highlight the importance of considering the trade-off between low air supply temperature and increased airflow rate, as well as the deleterious effect of temperature non-uniformity at the inlet of the racks on the data center's cooling infrastructure power consumption. The analysis enabled the development of a novel approach to reducing the energy consumption in enclosed aisle data centers using bypass recirculation. The development and experimental validation of a high-fidelity thermo-hydraulic model proceeded using the insights gained from the simple analysis. Using these tools, the study of optimum load placement is undertaken using computational fluid dynamics as the primary tool for analyzing the complex airflow and temperature patterns in the data center and is used to develop a rich dataset for the development of a reduced order model using proper orthogonal decomposition. The outcome of this work is the development of a robust set of rules that facilitate the energy efficient placement of the IT load amongst the operating servers in the data center and operation of the cooling infrastructure. The approach uses real-time temperature measurements at the inlet of the racks to remove IT load from the servers with the warmest inlet temperature (or add load to the servers with the coldest

inlet temperature). These strategies are compared to conventional load placement techniques and show superior performance by considering the holistic optimization of the data center and cooling infrastructure for a range of data center IT utilization levels, operating strategies and ambient conditions.

**Thermally aware, energy-based techniques for improving
data center energy efficiency**

By

Dustin W. Demetriou

B.S. Manhattan College, 2006

M.S. Syracuse University, 2008

Submitted in partial fulfillment of the requirements for the degree of doctor
of philosophy in Mechanical and Aerospace Engineering

Syracuse University
Syracuse, New York, U.S.A.
May 2012

Copyright © Dustin W. Demetriou, 2012
All Rights Reserved

Acknowledgements

I would like to extend my deepest gratitude to my advisor Dr. H. Ezzat Khalifa for his support and invaluable insight during the course of this research as well as during my Master's degree. Through our frequent interactions and his invaluable guidance, I have obtained a great deal of knowledge about not only the topics in this dissertation but also many topics deemed necessary for a successful career as a mechanical engineer. My appreciation for his investment in my academic and professional careers cannot be overstated. I would also like to acknowledge the efforts of Dr. John F. Dannenhoffer, who has provided constant guidance and feedback in every aspect of this dissertation, especially in the areas of computational and numerical modeling. His weekly efforts have shaped my understanding of how successful research is done. This dissertation would not have been successful without the insights provided by Drs. Roger Schmidt and Madhusudan Iyengar of IBM, whose in depth knowledge of the data center industry provided immeasurable guidance to a field that was relatively foreign to me at the start. I would like to thank the remaining members of my committee, Drs. Thong Q. Dang, Can Isik and Mark Glauser, who have all provided beneficial guidance that is reflected throughout this work.

I have also had the opportunity to work with several other faculty and researchers who have all provided invaluable help. Dr. Basman Elhadidi of Cairo University was vital in the development of the proper orthogonal decomposition codes. I have collaborated closely with Mr. Dan Rice of Syracuse University on the development of many software

tools, including the thermo-hydraulic model. The validation data was collected with the generous help of Joe Caricari, John Palmer, Sal Rosato, and Gerald Beattie from Grubb & Ellis, IBM Poughkeepsie Site. I would also like to thank the many other students and staff at Syracuse University who I have had the privilege of working with.

The work reported in this dissertation was performed with financial sponsorship from the IBM Corporation and the Syracuse Center of Excellence in Environmental and Energy Systems.

Table of Contents

Acknowledgements.....	v
List of Figures.....	xi
List of Tables.....	xvi
Nomenclature.....	xviii
1 Introduction	1-1
1.1 Background and Problem Definition	1-2
1.1.1 Trends in Data Center Energy Use	1-2
1.1.2 Data Center Systems	1-11
1.1.3 Data Center Thermal Environment.....	1-21
1.1.4 Metric Used for Evaluating Data Centers	1-26
1.2 Previous Work	1-32
1.2.1 Thermal Profiling.....	1-33
1.2.2 Computational Modeling of Airflow and Temperature Fields.....	1-34
1.2.3 Cooling and Power Infrastructure Modeling.....	1-41
1.2.4 Enclosed Aisle Data Centers.....	1-48
1.2.5 Reduced-Order Modeling	1-49
1.2.6 Load Placement.....	1-53
1.3 Best Practices for Data Center Design.....	1-58
1.4 Research Gaps in Data Centers.....	1-59
1.5 Objectives & Scope.....	1-61
1.6 Importance of Work.....	1-63
2 Modeling Considerations in Data Centers	2-65
2.1 Simple, Physics-based Model of a Data Center	2-66
2.1.1 Rack Inlet Temperature Non-Uniformity.....	2-66
2.1.2 Thermal Analysis	2-68
2.1.3 Cooling Infrastructure Power Consumption	2-72
2.1.4 Summary of Simple Model Assumptions	2-79
2.2 Thermo-hydraulic Model of a Data Center’s Cooling Infrastructure	2-80
2.2.1 Motor-driven Centrifugal Chiller.....	2-81

2.2.2	Effectiveness-NTU Model of a Wet Cooling Tower	2-82
2.2.3	Hydraulic Network.....	2-86
2.2.4	Chilled Water and Cooling Water Pumps.....	2-89
2.2.5	Computer Room Air Handler.....	2-89
2.2.6	Uninterruptable Power Supply (UPS).....	2-91
2.2.7	Creating a Coupled Simulation Environment	2-91
2.3	Computational Fluid Dynamics for Investigating the Data Center’s Airflow and Temperature distribution.....	2-95
2.3.1	Conservation of Mass	2-95
2.3.2	Conservation of Momentum	2-96
2.3.3	Conservation of Energy	2-100
2.3.4	Turbulence Modeling.....	2-100
3	Investigation into the Optimization of a Data Center’s Infrastructure using the Simple Model.....	3-103
3.1	Optimization of Enclosed Aisle Data Centers	3-103
3.1.1	Constant Server Flow	3-104
3.1.2	Effect of Server Temperature Rise.....	3-108
3.2	Optimization of Open Aisle Data Centers	3-110
3.2.1	Constant Server Flow.....	3-111
3.2.2	Effect of Server Temperature Rise.....	3-116
3.3	Chapter Conclusions	3-119
4	Verifying Assumptions of the Simple Model.....	4-121
4.1	Computational Domain and Setup.....	4-121
4.2	Practical Implementation of Bypass Recirculation.....	4-124
4.3	Additional Parameters for Data Center Airflow Analysis	4-129
4.4	Non-Uniformity Parameter, θ	4-137
4.5	Temperature Linearity and a Data Center Specific Archimedes Number	4-140
4.6	Chapter Conclusions	4-150
5	Experimental Validation of the Thermo-Hydraulic Model.....	5-152
5.1	Buildings 710 and 027 Data Collection	5-152
5.2	Post-Processing of Measurements	5-165

5.3	Thermo-hydraulic Model Validation	5-167
5.3.1	Validation Results	5-168
5.3.2	Discussion of Validation	5-174
5.4	Application of the Thermo-Hydraulic Model	5-178
5.4.1	Effect of Ambient Conditions	5-178
5.4.2	Impact of Chilled Water Set Point	5-183
5.5	Chapter Conclusions	5-184
6	Systematic Investigation of Thermally Aware, Energy-based Load Placement in Open Aisle Data Centers	6-186
6.1	Data Center Used in Load Placement Study	6-187
6.2	Workload Placement in Open Aisle Data Centers	6-189
6.2.1	Proposed IT Load Placement Scenarios	6-190
6.2.2	Optimization Procedure	6-197
6.2.3	Computational Fluid Dynamics Simulations	6-202
6.3	Application of Proper Orthogonal Decomposition to Data Centers	6-216
6.3.1	Proper Orthogonal Decomposition Theory	6-217
6.3.2	Application of POD to Thermally Aware, Energy-based Load Placement	6-220
6.3.3	Section Conclusions	6-239
6.4	Workload Placement Optimization Results	6-240
6.4.1	Baseline Scenarios	6-241
6.4.2	100% Useful IT Power	6-245
6.4.3	75% Useful IT Power	6-251
6.4.4	50% Useful IT Power	6-256
6.4.5	0% Useful IT Power	6-259
6.4.6	Verification of a Linear and One-to-One Temperature Field	6-262
6.4.7	Section Conclusions	6-266
7	Further Analysis of Thermally Aware, Energy-based Load Placement	7-270
7.1	Further Assessment of Optimum IT Load Placement	7-270
7.1.1	Optimum Load Placement for Reducing IT Load by Turning off the Hottest Chassis	7-271
7.1.2	Effect of Ambient Conditions	7-273

7.1.3	Turning off Idle Chassis' Power and Airflow.....	7-275
7.1.4	Optimum Workload Placement for Increasing IT Load by Turning on the Coldest Chassis	7-278
7.1.5	Dynamic Analysis of IT Load Placement	7-280
7.2	Optimum Workload Placement in Enclosed Aisle Data Centers.....	7-285
7.2.1	Effect of Ambient Conditions.....	7-290
7.2.2	Effect of Server Temperature Rise.....	7-292
7.2.3	Section Conclusions.....	7-294
7.3	Developing a Control Methodology for the Implementation of Thermally Aware, Energy-based Load Placement.....	7-295
7.4	Practical Implementation Issues.....	7-301
8	Conclusions.....	8-305
8.1	Summary of Results.....	8-305
8.2	Suggestions for Future Work.....	8-311
	Appendix A: Input Data Used for Thermo-Hydraulic Model Validation.....	8-314
	Appendix B: Relating the Maximum Chassis' Inlet Temperature to the Archimedes Number	8-332
	Appendix C: Using Design of Experiments to Investigate Optimum Load Placement.....	8-339
9	References.....	9-350

List of Figures

Figure 1.1 - U.S. Electricity Consumption by Sector (U.S. DOE, 2009; Koomey, 2011).....	1-4
Figure 1.2 - EPA Projections of Data Center Energy Use (EPA, 2007).....	1-4
Figure 1.3 - ASHRAE Power Trends for IT Equipment (ASHRAE, 2008b).....	1-6
Figure 1.4 - Breakdown of Energy Use in 40 U.S. Data Centers for 2007/2008 and 2010.....	1-10
Figure 1.5 - Standard Hot Aisle/Cold Aisle Arrangement in Raised Floor Data Centers.....	1-13
Figure 1.6 - Hot Air Recirculation from the Exhaust of Racks Back to the Cold Aisle.....	1-14
Figure 1.7 - Alternative Cooling Configurations (Schmidt, Cruz, Iyengar, 2005).....	1-15
Figure 1.8 - Schematic of a Data Center with a Central Chilled Water Plant.....	1-17
Figure 1.9 - AC Data Center Power Conversion Stages (Intel, 2007).....	1-18
Figure 1.10 - ASHRAE Environmental Guidelines for IT Equipment (ASHRAE, 2011).....	1-24
Figure 2.1 - Simple Model of an Air-Cooled Data Center.....	2-68
Figure 2.2 - Typical Isentropic Efficiency of a Single-Screw Compressor (ASHRAE, 2000)...	2-77
Figure 2.3 - Computed COP Ratio for a Vapor-Compression System Using R134a.....	2-78
Figure 2.4 - Regression Coefficients for ε_r Map.....	2-79
Figure 2.5 - Part Load Efficiency for Various UPS Topologies (adapted from LBNL, 2005)...	2-91
Figure 2.6 - Flow Chart of Thermo-Hydraulic Model.....	2-94
Figure 3.1 - Simple Model Optimization Results: Enclosed Aisle.....	3-106
Figure 3.2 - Preferred Configuration for Enclosed Aisle Data Centers with Bypass Recirculation.....	3-107
Figure 3.3 – Component-by-Component Energy Breakdown for $\theta = 0.0$	3-109
Figure 3.4 - Enclosed Aisle: Effect of Server Temperature Rise.....	3-110
Figure 3.5 - Simple Model Optimization Results: Open Aisle.....	3-113
Figure 3.6 – Component-by-Component Energy Breakdown for $\theta = 4.0$	3-115
Figure 3.7 - Refrigeration System COP for $\theta = 0.0$ and 4.0.....	3-115
Figure 3.8 - Effect of Server Temperature Rise on Optimum Cooling Power for Open Aisle Data Centers.....	3-117
Figure 3.9 - Effect of Server Temperature Rise on Optimum \bar{v}	3-118
Figure 3.10 - Effect of Server Temperature Rise on Optimum Supply Air Temperature.....	3-119
Figure 4.1 - Data Center Geometries Used in CFD Studies.....	4-122
Figure 4.2 - Layouts for Bypass Recirculation.....	4-125

Figure 4.3 - Measured Turbulence Intensity at CRAH Discharge.....	4-127
Figure 4.4 – Temperature (in K) Contours 8" Below the Raised Floor	4-128
Figure 4.5 - Relationship between $\bar{\psi}$ and ψ_c from CFD	4-133
Figure 4.6 - Relationship between F_o and ψ_c from CFD.....	4-133
Figure 4.7 - Relationship between λ_b and ψ_c from CFD.....	4-136
Figure 4.8 - Normalized Energy Results for Different Configurations and Leakage	4-136
Figure 4.9 - Comparison of Methods to Compute θ	4-139
Figure 4.10 - Effect of Ar on Temperature Linearity	4-146
Figure 4.11 - Inlet Temperature Contours for $\psi_T = 1.0$	4-147
Figure 4.12 - Inlet Temperature Contours for $\psi_T = 0.8$	4-148
Figure 4.13 - Inlet Temperature Contours Comparing Buoyancy for $\psi_T = 0.8 \Delta T_m = 20^\circ\text{C}$	4-149
Figure 5.1 - Weather Data for a Winter Day in Poughkeepsie, NY.....	5-161
Figure 5.2 - Comparison of Load Measurements	5-167
Figure 5.3 - Comparison of Cooling Tower Heat Rejection.....	5-169
Figure 5.4 - Comparison of Chiller Power.....	5-169
Figure 5.5 - Comparison of Chiller COP	5-170
Figure 5.6 - Comparison of Condenser Return Temperature.....	5-170
Figure 5.7 - Comparison of Condenser Supply Temperature	5-171
Figure 5.8 - Comparison of CRAH Heat Removal (Individual).....	5-172
Figure 5.9 - Comparison of CRAH Supply Air Temperature (Individual).....	5-172
Figure 5.10 - Comparison of CRAH Heat Removal (floor-by-floor)	5-173
Figure 5.11 - Predicted System Hydraulic Network Characteristics	5-174
Figure 5.12 - CRAH-by-CRAH Cooling Load Measurements.....	5-176
Figure 5.13 - Weather Data for a Summer Day in Poughkeepsie, NY	5-179
Figure 5.14 – Simulation Results: Data Center Energy for a Summer Day in Poughkeepsie, NY.....	5-181
Figure 5.15 – Simulation Results: Data Center Energy for a Winter Day in Poughkeepsie, NY.....	5-182
Figure 5.16 - Simulation Results: Changing Chilled Water Temperature	5-184
Figure 6.1 - Relationship between Useful IT and Data Center Heat Load	6-188
Figure 6.2 - Proposed Load Placement Scenarios for 75% Useful IT	6-195
Figure 6.3 - Proposed Load Placement Scenarios for 50% Useful IT	6-196

Figure 6.4 - Flow Chart for Infrastructure Optimization	6-201
Figure 6.5 - Energy Content of POD Modes	6-222
Figure 6.6 - Reconstruction Error for Useful IT of 100%.....	6-224
Figure 6.7 - Reconstruction Error for Useful IT of 75%.....	6-225
Figure 6.8 - Reconstruction Error for Useful IT of 50%.....	6-226
Figure 6.9 - Reconstruction Error for Useful IT of 25%.....	6-227
Figure 6.10 - Reconstruction Error for Useful IT of 0%.....	6-228
Figure 6.11 - POD Reconstruction Comparison with CFD Data.....	6-230
Figure 6.12 – Mean of the Rack’s Inlet Temperature Distribution for POD Data.....	6-231
Figure 6.13 - Temperature Reconstruction Contours as the Number of Modes in the Expansion is increased for 100% Useful IT	6-233
Figure 6.14 - Temperature Reconstruction Contours as the Number of Modes in the Expansion is increased for 50% Useful IT	6-234
Figure 6.15 - Amplitude Coefficient Distributions.....	6-236
Figure 6.16 - Error Plots for Interpolated Datasets.....	6-238
Figure 6.17 - POD Interpolation Comparison with CFD Data	6-239
Figure 6.18 - Energy Optimization Details of the Baseline Scenarios for Useful IT	6-244
Figure 6.19 - Component-by-Component Breakdown for Baseline Scenarios.....	6-245
Figure 6.20 - Energy Optimization Details of the 100% Useful IT	6-247
Figure 6.21 - Variation in Optimum Cooling Power with Chilled Water Flow Rate for 100% Useful IT Scenario	6-248
Figure 6.22 - Optimization Results for 100% Useful IT.....	6-250
Figure 6.23 - Component-by-Component Breakdown of Cooling Power for 100% Useful IT	6-251
Figure 6.24 - Optimization Results for Different Scenarios at 75% Useful IT.....	6-253
Figure 6.25 - Component-by-Component Breakdown of Cooling Power for 75% Useful IT ..	6-255
Figure 6.26 - Optimization Results for Different Scenarios at 50% Useful IT.....	6-257
Figure 6.27 - Component-by-component Breakdown of Cooling Power for 50% Useful IT...6-259	
Figure 6.28 - Optimization Results for 0% Useful IT.....	6-261
Figure 6.29 - Component-by-Component Breakdown of Cooling Power for 0% Useful IT	6-262
Figure 6.30 - Expected Error if the δT Assumption was not Valid	6-266
Figure 6.31 - Summary of Load Placement Scenario Analysis	6-269
Figure 7.1 - Normalized Cooling Power for Thermally Aware, Energy-based Load Placement.....	7-272

Figure 7.2 - Data Center Operation Map for Thermally Aware, Energy-based Load Placement.....	7-273
Figure 7.3 - Normalized Cooling Power at Different Ambient Temperatures for Thermally Aware, Energy-based Load Placement	7-274
Figure 7.4 - Optimum ψ_T at Different Ambient Temperatures for Thermally Aware, Energy-based Load Placement.....	7-275
Figure 7.5 – Idle vs. Shut Off Control at 75% Useful IT.....	7-277
Figure 7.6 - Idle vs. Shut Off Control at 50% Useful IT	7-278
Figure 7.7 - Comparing Scenarios 7 and 8 for 75% Useful IT	7-279
Figure 7.8 - Comparing Scenarios 7 and 8 for 50% Useful IT	7-280
Figure 7.9 - Comparing the Energy Consumption of Dynamic vs. Static Load Placement.....	7-283
Figure 7.10 - Comparison of Load Placement Arrangement for Dynamic vs. Static Operation	7-284
Figure 7.11 - Normalized Cooling Power Consumption with $\Delta T_m = 18.2^\circ\text{C}$ and $T_{wb} = 21.9^\circ\text{C}$	7-287
Figure 7.12 - Reduction in Cooling Power Consumption from a Conventional Enclosed Aisle with $\Delta T_m = 18.2^\circ\text{C}$ and $T_{wb} = 21.9^\circ\text{C}$	7-288
Figure 7.13 - CRAH Supply Air Temperature Requirements with $\Delta T_m = 18.2^\circ\text{C}$ and $T_{wb} = 21.9^\circ\text{C}$	7-290
Figure 7.14 - Effect of Ambient Temperature on the Optimum Cooling Power Consumption with $\Delta T_m = 18.2^\circ\text{C}$	7-291
Figure 7.15 - Effect of Ambient Temperature on the Optimum CRAH Airflow Fraction with $\Delta T_m = 18.2^\circ\text{C}$	7-292
Figure 7.16 - Effect of Chassis Temperature Rise on the Optimum Cooling Power Consumption with $T_{wb} = 21.9^\circ\text{C}$	7-293
Figure 7.17 - Effect of Chassis Temperature Rise on the Optimum CRAH Airflow Fraction with $T_{wb} = 21.9^\circ\text{C}$	7-294
Figure 7.18 - Proposed Control System for Implementing Thermally Aware, Energy-based Load Placement.....	7-300
Figure 8.1 - Inherent Problem with Air-Cooled Data Center.....	8-313
Figure A.1 - B027 Utility Plant Schematic	8-315
Figure A.2 - B710 Hydraulic Schematic (Pipe Lengths).....	8-316
Figure A.3 - B710 Hydraulic Schematic (Pipe Diameters).....	8-317

Figure A.4 – Chilled Water Valve Control.....	8-320
Figure A.5 – Computed Chilled Water Valve Discharge Coefficient C_D vs. Valve Opening ..	8-321
Figure A.6 - Computed Chilled Water Valve Correction Factor β vs. Valve Opening	8-321
Figure A.7 - B027 Infrastructure Design Point Summary	8-323
Figure A.8 - B027 Chilled Water Pump Characteristics.....	8-324
Figure A.9 - B027 Condenser Water Pump Characteristics	8-325
Figure A.10 – B027 Cooling Tower Performance Data	8-326
Figure A.11 – B027 Cooling Tower Fan Control	8-327
Figure A.12 - 1000 Ton Chiller Performance Data	8-329
Figure A.13 - 1200 Ton Chiller Performance Data	8-331
Figure B.1 – Chassis’ Inlet Temperature for $Ar = 0.54$	8-334
Figure B.2 – Chassis’ Inlet Temperature for $Ar = 1.17$	8-334
Figure B.3 - Slope of Temperature Curve.....	8-335
Figure B.4 - Intercept of Temperature Function	8-335
Figure B.5 – Functional Relationship for $(T_{max} - T_a)^*$ vs. Ar	8-336
Figure B.6 - Validation of Temperature versus Ar Function.....	8-337
Figure C.1 - Experimental Cube for a 2^3 Factorial Design	8-340
Figure C.2 - Normalized Energy Consumption Results for all DOE Cases	8-349
Figure C.3 - Comparison of Load Placement between Scenario 7 and Best DOE Case	8-349

List of Tables

Table 1.1 - 2011 ASHRAE Thermal Guidelines (Adapted from ASHRAE, 2011).....	1-25
Table 3.1 - Parameters Used in Simple Analysis.....	3-104
Table 4.1 - Boundary Conditions Used in Computational Domain.....	4-123
Table 4.2 - Rack Inlet Temperature for Various Bypass Geometries.....	4-129
Table 5.1 - Inventory of Equipment in IBM B710 and B027.....	5-153
Table 5.2 - Refrigeration Unit Measurements.....	5-155
Table 5.3 - Evaporator and Condenser Temperature Measurements.....	5-156
Table 5.4 - Cooling Tower Fan Measurements.....	5-156
Table 5.5 - CRAH Temperature Measurements: Lower Level.....	5-157
Table 5.6 - CRAH Temperature Measurements: First Floor.....	5-158
Table 5.7 - CRAH Temperature Measurements: Second Floor E.....	5-159
Table 5.8 - CRAH Temperature Measurements: Second Floor.....	5-160
Table 5.9 - Raised Floor Electrical Measurements: Lower Level.....	5-162
Table 5.10 - Raised Floor Electrical Measurements: First Floor.....	5-163
Table 5.11 - Raised Floor Electrical Measurements: Second Floor.....	5-164
Table 6.1 – Cooling Equipment Design Points for Load Placement Simulations.....	6-189
Table 6.2 - CFD Results for 100% Useful IT Load.....	6-202
Table 6.3 - CFD Results for Scenario 1 with 75% Useful IT Load.....	6-203
Table 6.4 - CFD Results for Scenario 2 with 75% Useful IT Load.....	6-203
Table 6.5 - CFD Results for Scenario 3 with 75% Useful IT Load.....	6-204
Table 6.6 - CFD Results for Scenario 4 with 75% Useful IT Load.....	6-204
Table 6.7 - CFD Results for Scenario 5 with 75% Useful IT Load.....	6-205
Table 6.8 - CFD Results for Scenario 6 with 75% Useful IT Load.....	6-205
Table 6.9 - CFD Results for Scenario 7 with 75% Useful IT Load.....	6-206
Table 6.10 - CFD Results for Scenario 1 with 50% Useful IT Load.....	6-207
Table 6.11 - CFD Results for Scenario 2 with 50% Useful IT Load.....	6-208
Table 6.12 - CFD Results for Scenario 3 with 50% Useful IT Load.....	6-209
Table 6.13 - CFD Results for Scenario 4 with 50% Useful IT Load.....	6-210
Table 6.14 - CFD Results for Scenario 5 with 50% Useful IT Load.....	6-211
Table 6.15 - CFD Results for Scenario 6 with 50% Useful IT Load.....	6-212

Table 6.16 - CFD Results for Scenario 7 with 50% Useful IT Load.....	6-213
Table 6.17 - CFD Results for Scenario 7 with 25% Useful IT Load.....	6-214
Table 6.18 - CFD Results with 0% Useful IT Load.....	6-215
Table 6.19 - CFD Results for Baseline Scenarios.....	6-242
Table 6.20 - Verification of One-to-One Temperature Field for 100% Useful IT.....	6-264
Table 6.21 - Verification of One-to-One Temperature Field for 50% and 75% Useful IT	6-264
Table A.1 - Minor Losses used in Thermo-Hydraulic Model.....	8-318
Table A.2 - CRAH Unit Design Characteristics	8-318
Table C.1 – Factorial Design Effect Coefficient Table	8-342
Table C.2 - 2^{8-4} Fractional Factorial Design Effect Coefficient Table.....	8-343
Table C.3 - Experimental Design for a 2^{32-26} Fractional Factorial Design.....	8-346

Nomenclature

A	area (m ²)
AA	coefficient
Ar	Archimedes number
c _p	specific heat (kJ/kg-K)
C	heat capacity (kW/K); mass fraction (kg/kg)
C _D	discharge coefficient
COP	coefficient of performance
CRAC	computer room air conditioner
CRAH	computer room air handler
D	diameter (m); mass diffusivity
DCiE	data center infrastructure efficiency
f	friction factor (m)
F	momentum body force
g	gravity (m/s ²)
Gr	Grashof number
h	specific enthalpy (kJ/kg)
I	turbulence intensity (%); current (amps)
IT	information technology
L	length (m)
Le	Lewis number
k	loss coefficient

K	pressure coefficient (Pa/m ⁶ /s ²)
\dot{m}	mass flow rate (kg/s)
M	molar mass (kg/kmol)
n	cooling tower exponent
N	number
NTU	number of transfer units
p	pressure (Pa)
P	power (kW)
PDU	power distribution unit
\dot{Q}	heat rate (kW)
Re	Reynolds number
\bar{R}	universal gas constant (kJ/kmol-K)
t	time (s)
T	temperature (°C or K)
TCO	total cost of ownership
u	velocity component (m/s)
UA	overall conductance (kW/K)
UPS	uninterruptable power supply
V	volume (m ³); Voltage
\dot{V}	volume flow rate (m ³ /s)
VFD	variable frequency drive
x	coordinate direction (m)

Greek Symbols

χ	condenser parameter
ψ	capture index
ψ_T	tile cold air supply fraction
λ	leakage parameter
Λ	leakage parameter
θ	recirculation non-uniformity
τ	temperature non-uniformity
η	efficiency
ε	effectiveness; turbulent dissipation
ε_r	Carnot efficiency ratio
σ	cooling tower coefficient; porosity
α	thermal diffusivity; opening ratio
ρ	density
Ω	rotational speed
β	thermal expansion coefficient; CRAH valve parameter
μ	viscosity
ω	humidity ratio
ζ	roughness (m)
γ	temperature ratio

Subscripts

m	mixed
---	-------

x exhaust
ax heat exchanger exhaust
a CRAC exit
c condenser
C cooling; CRAC; CRAH
e evaporator
F fan
eq equivalent
ov overall
s saturation
sh shaft
w water

Superscripts

d design
eff effective
in inlet
max maximum
min minimum
out exit
* redline

1 Introduction

The United States Environmental Protection Agency (EPA) defines a data center as a facility whose primary function is to house the data processing, data storage, networking and communication infrastructure that has become essential to the daily operation of almost every sector of the economy (EPA, 2007). The main objective of data center design and operation is to provide an acceptable thermal environment for the information technology (IT) equipment in order to maintain a high level of reliability. In the early 1990s, the total cost of ownership of a data center was driven primarily by the high initial cost of purchasing the IT hardware. Even as recently as 1996, the cost of purchasing new hardware was estimated to be 87% higher than the annual cost of providing cooling and power to the equipment (ASHRAE, 2008b). IT equipment reliability was maintained without a concern for the energy consumption of the data center. In recent years, the industry has undergone a shift in paradigm related to how the total cost of ownership (TCO) of the data center is partitioned. A number of trends in the IT industry have led to a significant increase in data center's energy consumption: reliance on the Internet, increased use of electronic transactions, national security, and the demand for high performance scientific computing, to name a few. The focus of this research is on developing tools and techniques to improve the energy efficiency of data centers, while maintaining the reliability of the IT equipment.

1.1 Background and Problem Definition

The purpose of this research is to evaluate the potential energy savings possible in data centers by optimizing their design and operation, without compromising the reliability of the IT equipment. Specifically, this work uses a systematic modeling approach - from simple physics-based models to higher fidelity, experimentally validated models – to optimize the operation of the data center’s cooling infrastructure and develop so-called thermally aware, energy-optimized load placement guidelines. To demonstrate these guidelines, computational fluid dynamics (CFD) simulations are used to obtain and understand the complex airflow and temperature patterns inherent in data centers. This work will demonstrate guidelines that use knowledge of the thermal environment of the data center from real-time sensor measurements to provide near-optimum load placement and cooling infrastructure operation solutions.

1.1.1 Trends in Data Center Energy Use

As a response to Congress’s Public Law 109-431 request, the United States Environmental Protection Agency (EPA), along with the United States Department of Energy (DOE), prepared a “Report to Congress on Server and Data Center Energy Efficiency (EPA, 2007),” which assessed the current energy use trends in data centers and provided guidelines and recommendations for pursuing energy efficiency. In PL 109-431, it was estimated that in 2006, the nation’s data centers consumed 61 billion kWh of electricity, enough to support 7.1 million average sized four-person homes (Lui, 2010). It was anticipated that with current trends in the IT industry, the energy consumption of

data centers would double every 5 years. Koomey (2011) provided up to date statistics on data center's electricity use. He showed that while still significant, U.S. data center's electricity use increased by only 36% between 2005 and 2010, instead of doubling as expected. He attributed this to a lower server installed base than predicted by EPA and not necessarily to efficiency improvements. The lower installed base was partially attributed to the current economic recession and it is not clear if this trend will continue in the future. Figure 1.1 provides historical and projected electricity use of different sectors in the United States. In 2010, U.S. data centers accounted for ~2% of the total electricity consumption of the United States. Koomey also estimated that worldwide electricity use by data centers increased by about 56% between 2005 and 2010, accounting for ~1.5% of the total world electricity consumption. It is anticipated that data center electricity use will continue to grow at a much faster rate than any other sector if no measures are taken to improve data center efficiency. In its report, the EPA outlined possibilities for significant energy improvement in data centers through the use and research of innovative technologies and best practices. Figure 1.2 gives projections by the EPA for significant reductions in electricity use by implementing these state-of-the-art technologies and best practices, highlighting opportunities to reduce electricity consumption by as much as 74 billion kWh.

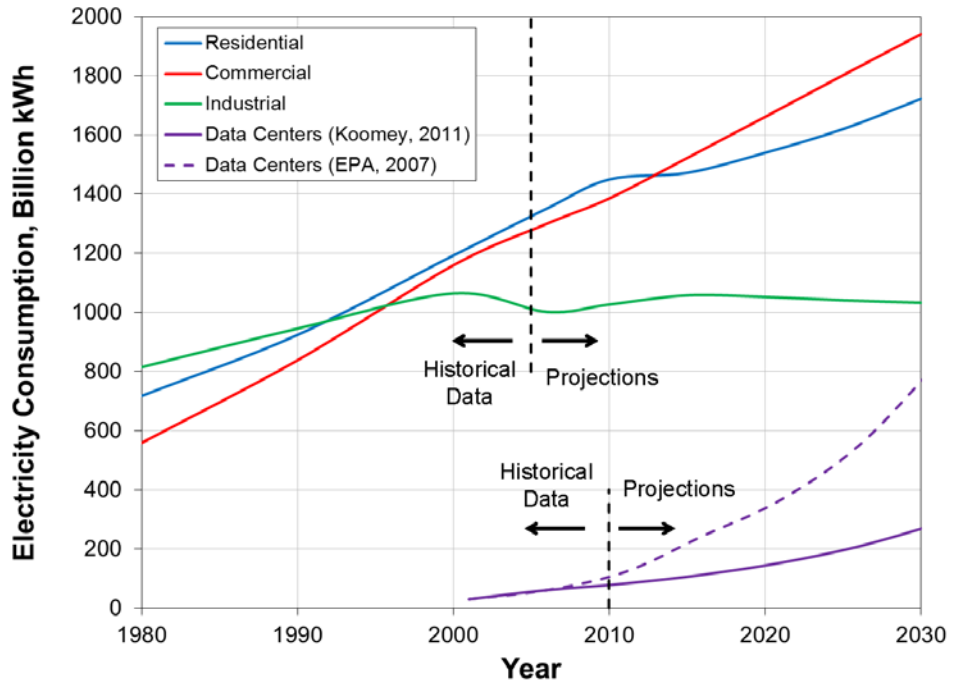


Figure 1.1 - U.S. Electricity Consumption by Sector (U.S. DOE, 2009; Kooimey, 2011)

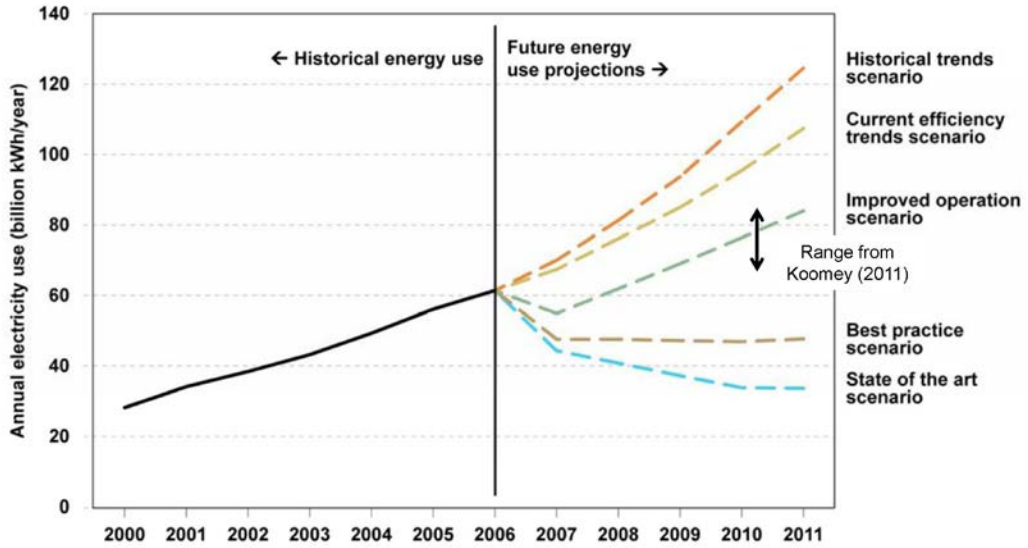


Figure 1.2 - EPA Projections of Data Center Energy Use (EPA, 2007)

The shift in thinking related to data center TCO – from one based solely on the IT equipment capital cost to one that accounts for both the capital cost and the energy cost – has been driven by a number of factors. First, between 1999 and 2005, the electricity consumption of data centers rose more than 39% (Salim and Tozer, 2010), while worldwide IT hardware costs stayed almost constant. Statistics for 2010 show that for an estimated installed base of 45 million additional servers, the hardware cost would be around \$60 billion, whereas the cooling and power costs were estimated to be over \$40 billion. To contrast this, in 1996, for an estimated additional installed base of 5 million servers, the hardware cost was \$65 billion and the power and cooling cost was around \$10 billion (ASHRAE, 2008b). Several studies have even shown that the cooling and power costs are greater than the cost of the IT equipment it supports (ASHRAE, 2008b). Secondly, the average retail price of electricity in the United States increased nearly 50% between 1995 and 2009, from 4.66 cents/kWh to 6.95 cents/kWh, in the industrial sector (EIA, 2010). Lastly, the increased reliance on computing infrastructure has led to the miniaturization of microprocessors and a sharp increase in power (and subsequently heat) density, which necessitates a higher cooling demand. The power and cooling infrastructure that supports the IT equipment consumes upwards of 50% of the total data center electricity (Salim and Tozer, 2010). A major challenge in data center design comes from a lifecycle mismatch between the IT equipment and the cooling infrastructure. The lifecycle of the IT equipment is only a few years, whereas the lifecycle of the cooling infrastructure is typically 15 – 25 years (Rambo and Joshi, 2007). Figure 1.3 gives typical rack heat loads per rack footprint as a function of product announcement year. In 1990, a typical rack's heat dissipation was no more than 1 kW; whereas, the same rack populated

with today's servers can easily exceed 30 kW. Clearly, the same cooling infrastructure designed in 1990 would not be suitable today. The high cost of owning and operating a data center has forced data center operators to maximize the computing output by packing many servers into smaller spaces. Large Internet data centers can house 50,000 servers (Rolia et al., 2000), with an electricity demand on the order of 10s of MW.

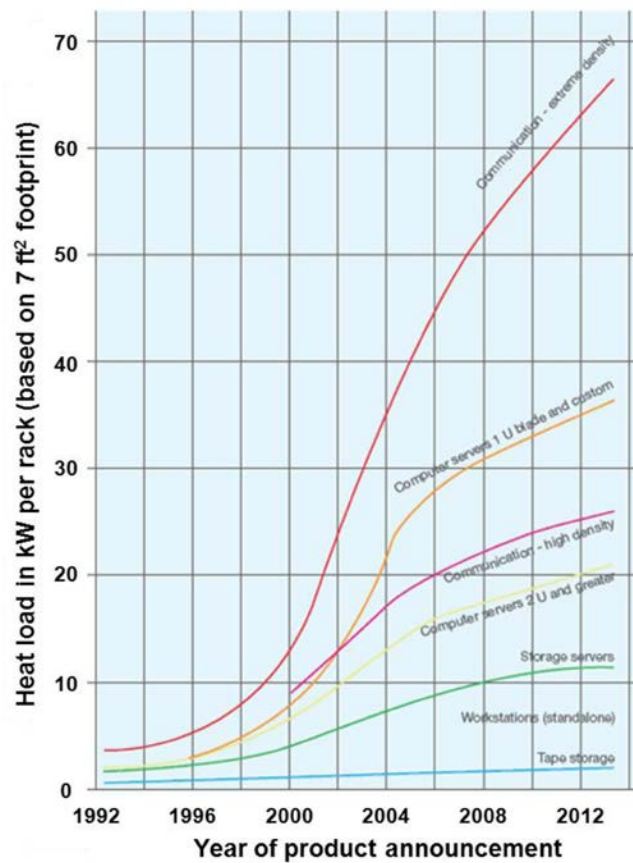


Figure 1.3 - ASHRAE Power Trends for IT Equipment (ASHRAE, 2008b)

Aside from the EPA, DOE, and the IT manufacturers, several other organizations have recognized the need for improvements in data center's energy efficiency. In 2003, the American Society of Heating, Refrigeration, and Air-Conditioning Engineers (ASHRAE) along with key technical experts from the IT industry created ASHRAE Technical Committee (TC) 9.9 (ASHRAE, 2011b), "Mission Critical Facilities" to produce and publish information and provide training related to HVAC for data center facilities. Since its inception, TC 9.9 has published a series of books that provide trends and best practices for data center operation (ASHRAE 2008b; ASHRAE 2008c). The Green Grid (2011) is a global consortium of IT companies and professionals seeking to achieve optimal energy efficiency in data centers. A key contribution of the Green Grid has been the development of the widely adopted data center metric Power Usage Effectiveness (PUE) (The Green Grid, 2007, 2008a), which is defined as,

$$PUE = \frac{\textit{Total Facility Power}}{\textit{IT Equipment Power}}. \quad (1.1)$$

The numerator of PUE contains the total facility power, which is defined as the power measured at the utility meter that is dedicated solely to the data center. This includes the power distributed to the uninterruptible power supplies (UPS), switchgear, generators, power distribution units (PDU), batteries, distribution losses outside the IT equipment, chillers, CRAC/CRAH blowers, pumps, cooling tower fans, lighting and IT equipment power. The denominator of PUE is the total power distributed to the IT equipment, which is defined as "the power used to manage, process, store or route data within the data

center.” Included in this are compute, storage, network equipment, keyboard, video, and mouse (KVM) switches, monitors and workstations/laptops used to control the data center. A PUE of 1.0 is ideal, meaning all the power that goes into the data center facility is utilized in the IT equipment and no power is used for cooling or any infrastructure beyond IT equipment. Several other organizations have been founded to extend the knowledge base in data centers, including The Uptime Institute (2011) and DatacenterDynamics (2011).

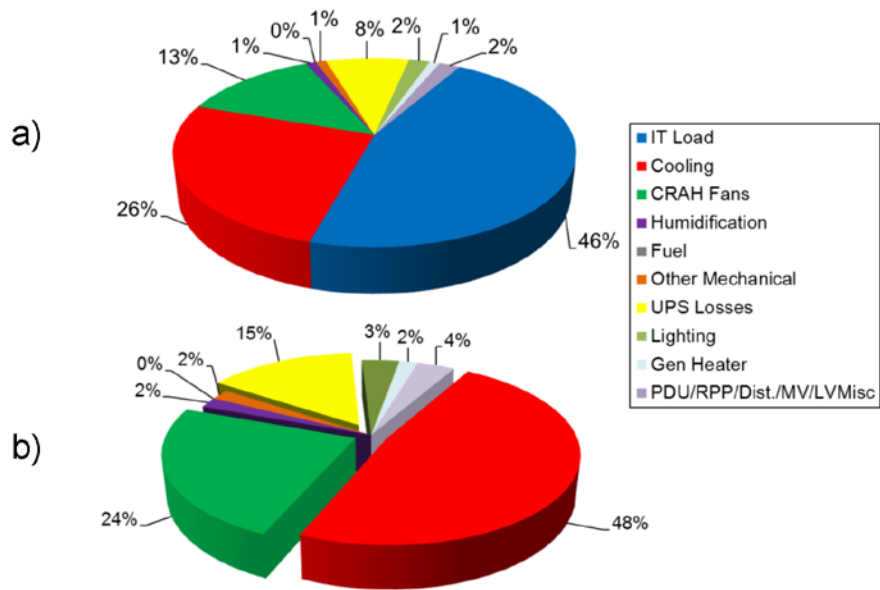
An increased focus on data center’s energy consumption has resulted in numerous benchmarking studies, which looked at the breakdown of the energy among the IT equipment and the supporting infrastructure. Lawrence Berkeley National Laboratory (LBNL) (2007) performed one of the first benchmarking studies on a collection of 12 data centers. They found that the average PUE of the 12 data centers was 2.17, concluding that on average more than half of the power delivered to the data center goes to the supporting infrastructure and not to the IT equipment itself. A more recent benchmarking study published by Salim and Tozer (2010) presented a detailed annual energy assessment of 40 data centers. The results, which were obtained during 2007 and 2008, show that the average annual PUE ranged from 1.69 to 3.57. The overall average of the 40 data centers was 2.19, meaning that for every kW of IT power another 1.19 kW of additional power is used for cooling and electrical power conditioning and distribution. Figure 1.4 presents the breakdown of the energy consumption of the 40 data centers. The study found a strong correlation between data center size and PUE. Smaller data centers (less than 10,000 ft²) typically had a higher annual PUE than large data centers (greater

than 30,000 ft²). Large-sized data centers typically implemented various methods of free cooling, had larger, higher efficiency equipment and had more cooperation between IT and facilities departments. An important concept in data center design is the tier level, as defined by the Uptime Institute (2009). The tier level describes the requirement for the data center's infrastructure to maintain uptime and reliability needs. The tier levels range from Tier 1, which provide basic site infrastructure to maintain 99.671% availability, to Tier 4, which provide fault-tolerant infrastructure to maintain 99.995% availability.

While the tier level certainly has an impact on data center's reliability, the study by Salim and Tozer (2010) found no correlation between the infrastructure's power consumption and the tier level.

A more recent dataset obtain from Salim and Tozer of the same 40 data centers showed a drop in average PUE to 1.89, due to awareness and implementation of various technologies since the initial benchmarking study. Google (2011) has been one of the leaders in providing benchmarking data for their data centers. Insights into the energy efficiency of several of their large-sized data centers in Mountain View, California showed monitored average PUE values between 1.15 - 1.20. However, in many cases, low PUE values are a function of climate with the use of free cooling and the results would not be realized in data centers located in other regions. In addition, internet data centers do not require the same level of reliability as mission critical facilities (i.e., banks, government data centers, etc.), which is important when direct free-cooling methods are used.

Data Center Energy Breakdown 2007/08 - Average PUE = 2.19



Data Center Energy Breakdown 2010 - Average PUE = 1.89

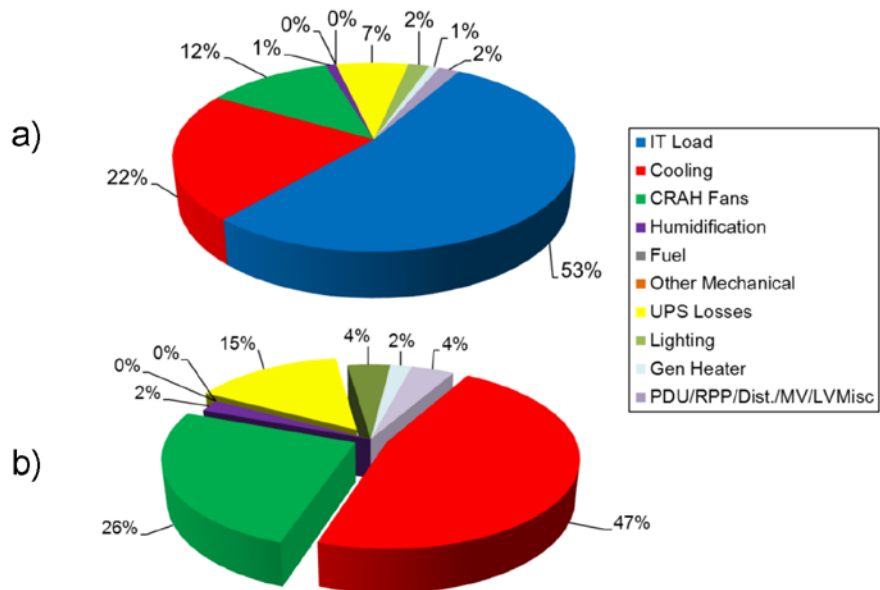


Figure 1.4 - Breakdown of Energy Use in 40 U.S. Data Centers for 2007/2008 and 2010

a) Includes IT Power b) Cooling and Power Infrastructure

(Data was Obtained from Dr. Munther Salim and used with Permission)

1.1.1.1 Guidelines for Improving Energy Efficiency

To improve beyond the current state-of-the-art, The Green Grid (2008c) outlined seven strategies and directions of research that should lead to improved energy efficiency.

These include:

1. Develop an effective air management strategy by use of,
 - a) Enclosed aisle
 - b) Infrastructure optimization
 - c) Optimized load placement
2. Moving the cooling system closer to the load, i.e.,
 - a) Liquid cooling
 - b) In-row coolers
 - c) Rack heat exchangers
3. Operate the data center at a higher ΔT
4. Design equipment that can handle higher component temperatures.
5. Install economizers to provide free-cooling when available.
6. Use higher efficiency equipment.
7. Use of dynamic controls, i.e.,
 - a) CRAC fan speed control
 - b) Compressor variable frequency drives

This dissertation studies several of these guidelines in detail, specifically, the use of effective air management strategies and the use of dynamic controls.

1.1.2 Data Center Systems

This work will focus on the design and operation of raised-floor, air-cooled data center (RF/AC) configurations – although alternative configurations will be discussed as

needed. A typical RF/AC data center can be divided into two subsystems: the raised-floor space and the cooling infrastructure. This section provides a description of each of these subsystems.

1.1.2.1 Raised-Floor Space

The raised-floor space of the data center houses the IT equipment, PDU and the CRAH units. The management of airflow and temperature patterns within the data center is paramount to providing an acceptable thermal environment to the IT equipment. The most common arrangement is to house the computer servers (i.e., the module that contains the CPU, hard drive, RAM, etc.) in storage units known as racks. Industry standard rack sizes have a width of either 19” or 24”, depth of 42” and a height of 78” or 45U (1U = 1.75”). Typical horizontal server sizes are 1U and 2U tall, while so-called blade server systems are typically 7U tall, but house 10+ vertical servers. For cooling purposes, racks employ front-to-back cooling, where cold air is ingested into the front of the rack and is exhausted from the back of the rack, after being heated by the IT equipment. The racks are typically arranged in a hot aisle, cold aisle configuration, as pictured in Figure 1.7, where the front of the racks face the cold aisle and the rear of the racks face the hot aisle. The most common layout is the raised-floor configuration, where the chilled air is delivered from a pressurized under-floor plenum through perforated tiles located in the cold aisle. This configuration allows for some separation of the cold air supply from the warm exhaust. However, after the air leaves the racks, a fraction returns to the raised floor plenum through the CRAH units, where it is cooled back to the

necessary temperature and a fraction is recirculated back to the cold aisle where it is mixed with the cold supply air.

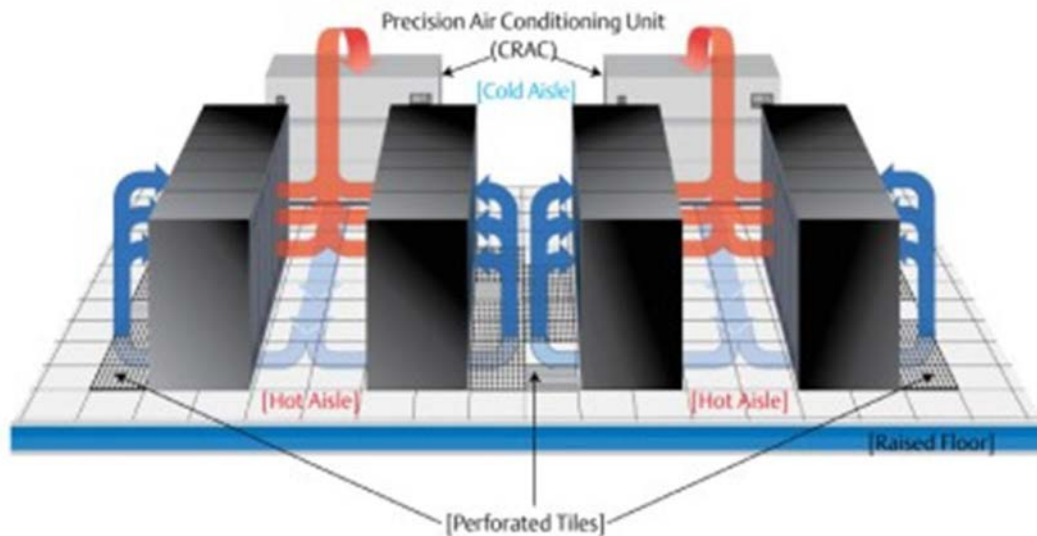


Figure 1.5 - Standard Hot Aisle/Cold Aisle Arrangement in Raised Floor Data Centers

The phenomena of hot air recirculation as shown in Figure 1.6, where hot exhaust air is entrained into the chilled air in the cold aisle before it enters the IT equipment, is one of the key drivers for increased energy consumption of the cooling infrastructure. This problem causes increased inlet temperatures to the equipment, which could pose reliability concerns if not properly controlled. This is prevalent in data centers for two reasons: 1) in order to save air-moving energy, data center operators typically supply less air through the perforated tiles than is required by the equipment and 2) the placement of CRAH units in the proximity of the cold aisle may cause air to short-circuit directly to the inlet of the CRAH units.

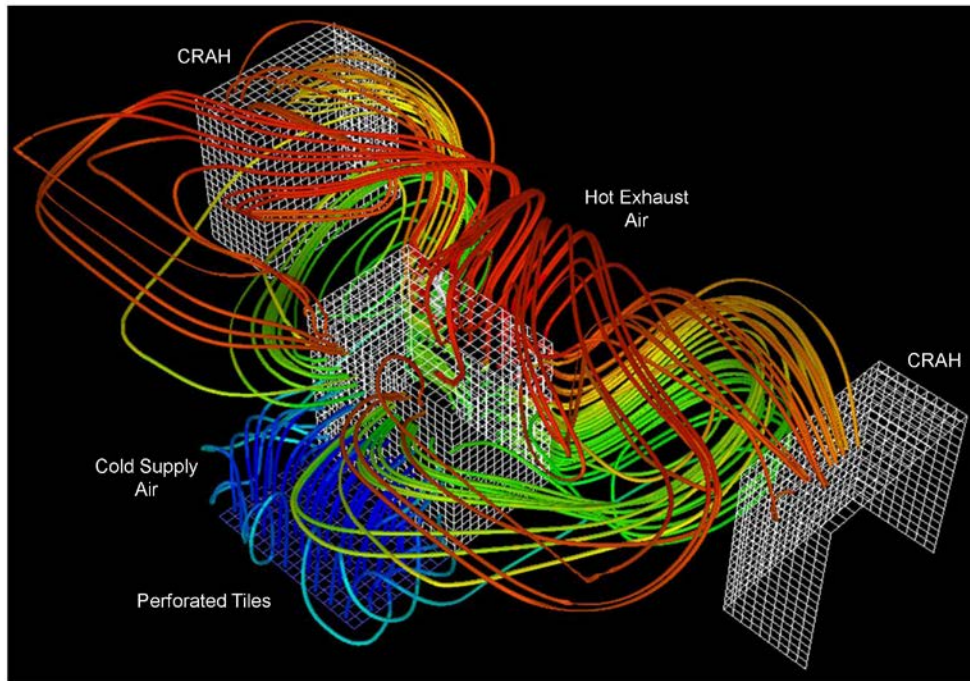


Figure 1.6 - Hot Air Recirculation from the Exhaust of Racks Back to the Cold Aisle

Several other airflow configurations have been investigated. Schmidt, Cruz and Iyengar (2005) describe several configurations, which are pictured in Figure 1.7. Stahl and Belady (2001) introduced an overhead cooling system and showed that in combination with a raised floor could sustain 30% higher heat loads.

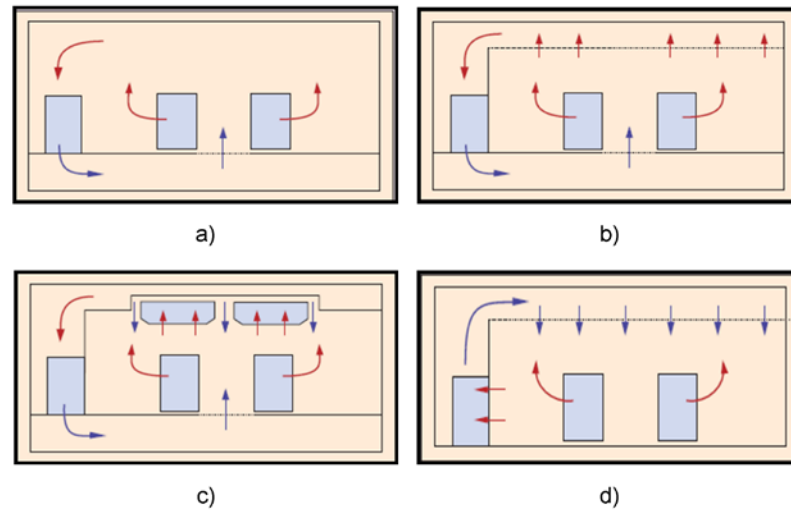


Figure 1.7 - Alternative Cooling Configurations (Schmidt, Cruz, Iyengar, 2005)

a) under-floor air distribution with CRAH room return, b) under-floor air distribution with ceiling plenum return, c) under-floor air distribution with local ceiling cooling units and d) overhead air distribution with CRAH room return

A growing trend in data center design is to enclose the cold aisle or hot aisle using side barriers and a roof. This strategy has the clear advantage that none of the hot exhaust air from the racks enters the IT equipment, eliminating any temperature gradient at the servers because of recirculation. This strategy would guarantee that the servers receive air directly from the perforated tiles at the supply air temperature. However, the flow rate through the tiles would need to be equal to the flow rate through the servers, which typically is not the case in open-aisle data centers. The increase in airflow through the perforated tiles would increase the energy consumption of the air moving devices. Google (2011) implements a form of cold aisle containment by placing the servers in containers, similar to tractor-trailers, to eliminate the recirculation of hot air.

1.1.2.2 Cooling and Power Infrastructure

Most high-density enterprise data centers use centralized chilled water plants to provide chilled water to computer room air handling units (CRAH) housed in the data center raised floor space. Figure 1.8 shows the typical equipment found in a RF/AC data center with a central chilled water plant. A data center chilled water plant is identical to a plant found in a conventional building; however, an office building produces heat fluxes on the average of $5 - 10 \text{ W/ft}^2$, whereas data centers have been documented to produce heat fluxes as high as $150 - 750 \text{ W/ft}^2$ (Lui, 2010). The chilled water plant contains of three loops: 1) *heat rejection loop* (cooling tower, condenser cooling water pump and the condenser) 2) *refrigeration loop* (compressor, condenser, throttling device and evaporator) and 3) *chilled water loop* (evaporator, chilled water pumps and the CRAH units). Each of these is needed in order to provide the necessary cooling to the data center raised floor. Although required for the data center to operate, the power consumption of these components is considered non-IT and is placed in the numerator of the PUE. In addition, a network of pipes that have their own hydraulic and thermal characteristics connects these subsystems.

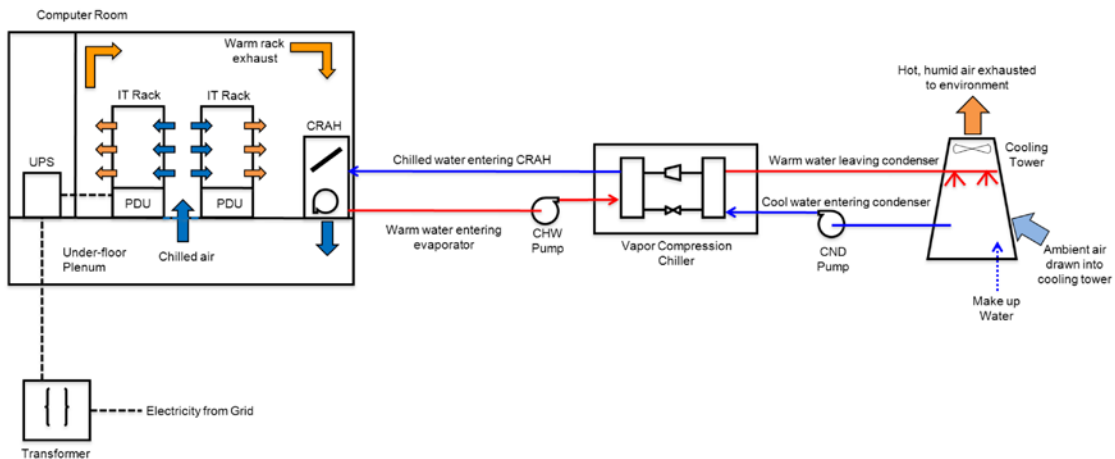


Figure 1.8 - Schematic of a Data Center with a Central Chilled Water Plant

Two types of air conditioning units are common in data center applications. A Computer Room Air Handler (CRAH) is a chilled water unit, which uses chilled water from a central chiller plant as the coolant to cool the data center's cooling air. Computer room air conditioners (CRAC) are direct expansion (DX) vapor compression units, which cool the air using a boiling refrigerant. Typically, CRAC units are used in smaller data centers due to their reduced efficiency compared to a central chilled water plant (Salim and Tozer, 2010). Both of these systems provide chilled air to an under-floor plenum with one or more fans housed in the unit. These fans must overcome the pressure losses due to the CRAC/H heat exchanger and filters, as well as the pressure loss as the air travels in the data center.

Power distribution and conditioning equipment for the IT equipment, which is typically housed on the raised floor space, include power distribution units (PDU), power supply

units (PSU) and other necessary power conversion equipment. Figure 1.9 shows the typical AC data center's power conversion stages and associated efficiencies. Any inefficiency in power conversion shows up as heat dissipation, which must be removed via the cooling infrastructure.

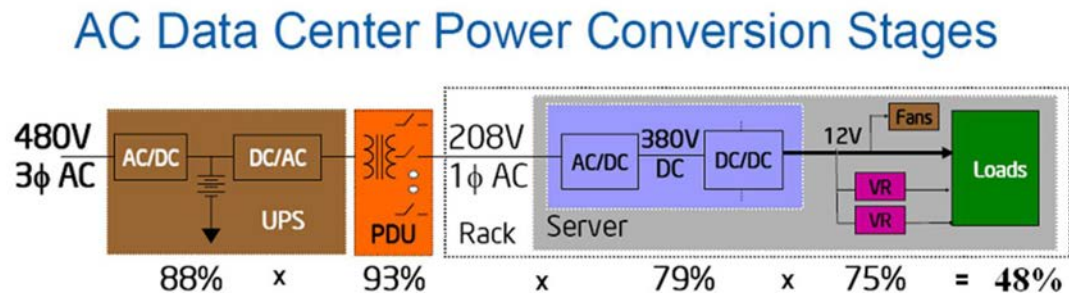


Figure 1.9 - AC Data Center Power Conversion Stages (Intel, 2007)

Figure 1.4 showed the typical energy breakdown for the cooling and power infrastructure. The LBNL benchmarking studies showed that the infrastructure consumed 54% of the total power delivered to the data center. Of the components considered infrastructure, three in particular consume a substantial portion of this power – the cooling (or compressor power) at 48%, the CRAH fans at 24% and the UPS/PDU Losses at 15%. The high cooling power can be attributed to two causes. First, many small data center utilize air-cooled direct expansion CRAC units which have a relatively low coefficient of performance (COP), the ratio of the cooling obtained to the required work input, compared to centralized chilled water systems. The COP of a typical air-cooled DX unit is only around 2.5, compared to even a legacy water-cooled chiller, which can have a typical COP of ~6.0. A water-cooled chiller produced today is far more efficient and can

easily have a design COP in the 6.5 – 7.5 range. However, in many instances, data center cooling equipment operates off-design because of variations in IT load and ambient conditions, which can significantly degrade its performance. The second largest consumers of power in the cooling infrastructure are the CRAH fans. The CRAH/CRAC units typically have a large pressure drop, relative to the rest of the air-distribution path, due to the presence of filters and the heat exchanger. With the large volume of air being moved in data center, the CRAH/CRAC blowers consume a significant amount of power to move this air across the large pressure resistance. In addition, to assure reliability, data centers typically operate far more CRAH units than required. Many of these units move air throughout the data center, hence consuming fan power, but do not contribute to cooling. In many cases, legacy CRAH units are installed that typically have constant speed fans and lower efficiency motors. Typical CRAH units today will have the option of installing a variable frequency drive (VFD) in order to reduce the fan speed thus consuming a fraction of the power. The third most significant component of the infrastructure was the UPS/PDU losses. Generally speaking, UPS/PDU losses are reduced by operating the UPS/PDU at a higher load factor.

1.1.2.2.1 Economizers

While not essential to this research, it is worth looking at two common methods to reduce the significant power consumption of compressor-based cooling options. Both of these methods rely on favorable ambient conditions in order to provide “free-cooling” to the data center.

1.1.2.2.1.1 Air-side Economizers

The use of outdoor air (OA) is required in traditional office buildings by ASHRAE Standard 62.1 (2004a), “Ventilation of Acceptable Indoor Air Quality,” which requires a minimum amount of OA to provide acceptable indoor air quality, and ASHRAE Standard 90.1 (2004b), “Energy Standard for Buildings,” which requires the use of airside economizers for office buildings. Currently, these standards do not apply to data centers because of the “mission critical” label, but ASHRAE TC 9.9 has recently begun work with the Standard 90.1 committee to develop an energy standard for data centers. The environmental guidelines of ASHRAE aim at increasing the number of hours available for bringing in ambient air directly to the data center by increasing the upper limit on dry bulb temperature and lowering the moisture limit. However, acceptance of direct airside economizer use in data centers has been limited for several reasons. Contamination of the equipment from particles and pollutant gases in the OA is a concern. Particles settling on the circuit boards can cause electrical shorting and corrode circuit boards. Chlorine and sulfur-bearing salts can be a significant problem for printed circuit boards (Lui, 2010). The European Union directive 2002/95/EC “on the Restriction of the use of certain Hazardous Substances in electrical and electronic equipment,” eliminated the use of lead in electronic products. Research has shown that printed circuit boards made using lead-free materials are more susceptible to corrosion, which has been noticeable in hard disk drives (Muller, 2010). Moisture levels of the air are a matter of concern for OA economizer use in data centers. Studies have shown that dryer air results in a greater risk of electro-static discharge (ESD). A strong correlation has been seen between the air dew point and charge creation, which lead ASHRAE to base a low moisture limit on dew point. It is anticipated that in certain climates the need for humidification could be a

significant energy drawback (Lui, 2010). It has also been shown that conductive anodic filament growth is strongly correlated to relative humidity (ASHRAE, 2008a). However, Tschudi (2007) argued that with only modest improvements in filtration, particle concentrations could be negated.

1.1.2.2.1.2 Water-side Economizers

Waterside economizers take advantage of cool condenser water from the cooling tower in order to provide either full or partial cooling of the chilled water. Two main types of systems exist. In a direct system, the condenser water is brought directly to the CRAH units for cooling; however, fouling in the cooling coils is a concern in these systems (ASHRAE, 2008b). In an indirect system, a heat exchanger is used to separate the condenser water stream from the chilled water stream. Indirect systems come in two forms: parallel and series. A parallel configuration only allows operation when the economizer can meet the full load. Series operation allows the economizer to pre-cool the chilled water in partial mode or take the entire load, depending on the condition of the condenser water. Series mode allows for an increased number of operational hours for economizer use (ASHRAE, 2008b).

1.1.3 Data Center Thermal Environment

The primary function of the systems described in Section 1.1.2 is to provide an acceptable thermal environment in order to maintain the reliability of the IT equipment. In 2004, ASHRAE TC 9.9 compiled a set of guidelines for the acceptable condition of the air entering the IT equipment. In 2008, these guidelines were expanded to allow

greater flexibility in facility operation (ASHRAE, 2008a). In 2011, a substantial update to the guidelines was issued by ASHRAE TC 9.9 (ASHRAE, 2011). This document went beyond providing only thermal environment recommendations for end-users and outlined a procedure for using the guidelines to minimize the TCO of a data center considering both energy and reliability. The guidelines delineate data centers into one of 6 classes: A1 – A4, B and C. Each class has a distinct set of environmental criteria depending on the type of equipment, level of control and overall reliability needs – this must be selected based on the end user’s business priorities. This research will consider data centers that fall into classes A1 and A2, which include enterprise and volume servers and storage along with personal computers and workstations that are tightly controlled. Classes A3 and A4 also pertain to the same types of IT equipment, but were included in the guidelines to illustrate where the industry would like new equipment to be designed. Currently, IT manufacturers do not sell equipment that can operate reliably in classes A3 and A4.

Two specifications are given in the guidelines that relate to the thermal and moisture conditions of the air entering the IT equipment. The *recommended envelope* “defines the limits under which IT equipment would operate the most reliably while still achieving reasonably energy efficient data center operation”. The *allowable envelope* gives a maximum limit under which equipment can be operated for short periods of time and still maintain functionality. The distinction becomes very important when considering free cooling methods that want to maximize the number of hours one can operate against the allowable envelope and still maintain IT equipment reliability. Figure 1.10 provides a

psychrometric chart showing the recommended and allowable envelopes for classes A1 and A2 data centers. Table 1.1 provides details on the full scope of the 2011 ASHRAE thermal guidelines. The environmental conditions given in Figure 1.10 and Table 1.1 are for the air entering the IT equipment and should be measured 2” in front of the air inlet.

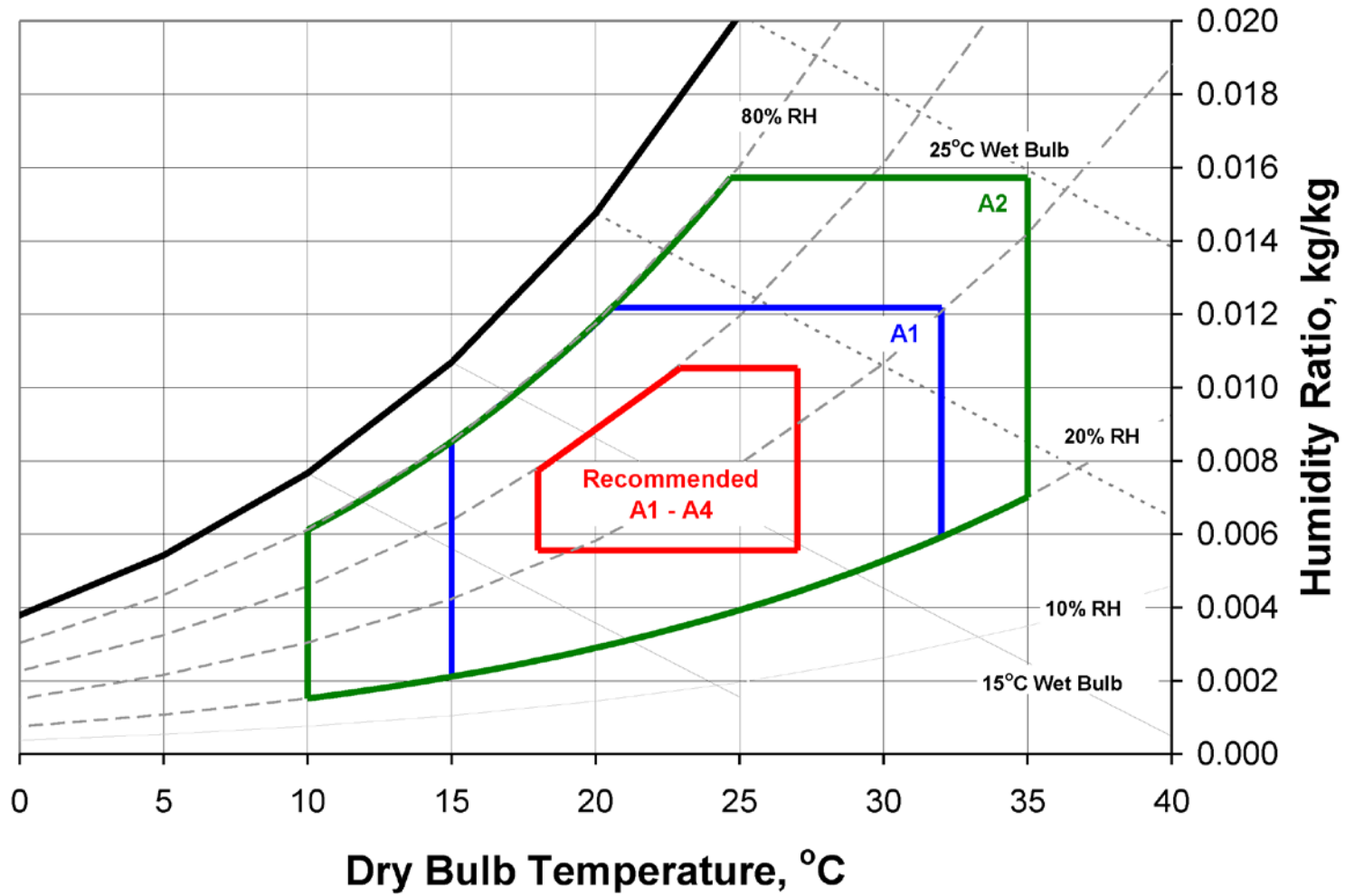


Figure 1.10 - ASHRAE Environmental Guidelines for IT Equipment (ASHRAE, 2011)

Table 1.1 - 2011 ASHRAE Thermal Guidelines (Adapted from ASHRAE, 2011)

Classes	Operational Products					Powered Off Products		
	Dry-bulb Temperature °C	Humidity Range	Max Dew Point °C	Max Elevation m	Max Rate of Change °C/h	Dry-Bulb Temperature °C	Relative Humidity %	Max Dew Point °C
A1 to A4	18 to 27	5.5°C DP to 60% RH and 15°C DP	Recommended Envelope (Applies to all A classes)					
Allowable Envelope								
A1	15 to 32	20% to 80% RH	17	3050	5/20	5 to 45	8 to 80	27
A2	10 to 35	20% to 80% RH	21	3050	5/20	5 to 45	8 to 80	27
A3	5 to 40	-12°C DP & 8% to 85% RH	24	3050	5/20	5 to 45	8 to 85	27
A4	5 to 45	-12°C DP & 8% to 90% RH	24	3050	5/20	5 to 45	8 to 90	27
B	5 to 35	8% to 80% RH	28	3050	NA	5 to 45	8 to 80	29
C	5 to 40	8% to 80% RH	28	3050	NA	5 to 45	8 to 80	29

1.1.4 Metric Used for Evaluating Data Centers

A number of metrics have been developed to give data center operators and designers a better understanding of the operation of their data center. Throughout the course of this dissertation, many of these metrics are considered and alternative metric are developed as needed. The following section outlines the most widely used metrics in use today.

1.1.4.1 Data Center Infrastructure Efficiency, DCiE

The Green Grid (2007) defines the reciprocal of its PUE as the data center's infrastructure efficiency (DCiE). In the white paper "The Green Grid Metrics: Data Center Infrastructure Efficiency Detailed Analysis," (2008a) The Green Grid outlines the necessary steps to derive the DCiE for a data center. Three levels of detail are described:

1. Level 1 – collect power measurements monthly from the UPS and at the main distribution panel feeding all mechanical equipment.
2. Level 2 – collect power measurements on a daily basis from the PDUs supplying the racks and the distribution system feeding each piece of mechanical equipment separately.
3. Level 3 – collect data on a continuous basis from each piece of equipment individually.

Clearly, each level adds more complexity to data collection. In order to make the DCiE a beneficial metric for the data center, data must be measured for a period that is longer

than the cyclic variations in efficiency, which are yearly in most facilities (The Green Grid, 2008a). Due to these variations, the Green Grid does not recommend estimates of DCiE from manufacturer rated values.

There are two major deficiencies in both PUE and DCiE when considering non-traditional energy savings techniques. First, no benefit is available for generating on-site power or for the use of waste heat, which become important when considering applications such as on-site cogeneration. Secondly, when considering optimized load placement and operation of the servers, a reduction in IT load does not necessarily reduce the PUE and in many instances increases it. The industry recognized these deficiencies and is working on the development of the Energy Reuse Effectiveness (Green Grid, 2010), Compute Power Efficiency (Malone and Belady, 2008) and Data Center Energy Productivity (The Green Grid, 2008b) to overcome these deficiencies. The Energy Reuse Effectiveness gives credit for the use of waste heat that is used for either providing heating or cooling to the data center; however, this metric does not distinguish between heat and work provided by electricity and therefore is thermodynamically incorrect. Both the Data Center Energy Productivity and Compute Power Efficiency look to quantify the actual useful work obtained (in terms of CPU utilization) compared to the power required to perform the work. Since the formulation includes the actual CPU utilization, data center operators can begin to address both efficient IT power utilization and cooling and power infrastructure. Shah, Bash, Kumari et al. (2011) define a “coefficient of performance grand” (COP_G) for the data center as the ratio of useful IT power to the total power required to run the data center. In their definition, the useful IT removes the

contribution to the typical definition of IT (such as in the definition of PUE and DCiE) from chip leakage, server fans and any idle processor power. They present several case studies looking at how server fan algorithms and workload consolidation can lead to misleading results when the conventional DciE and PUE are used to compare data centers that have continually changing IT and facility power.

1.1.4.2 Capture Index, ψ

VanGilder and Shrivastava (2007) proposed the use of the capture index (ψ), which is based on the airflow patterns in the data center, in order to determine the amount of air that is recirculated from the exhaust of the racks to the inlet of the racks. The capture index is obtained by releasing passive contaminants at the perforated tiles that “track” the airflow patterns. The cold aisle capture index is defined as,

$$\psi = \frac{C_i}{C_{cooling}} , \quad (1.2)$$

where, C_i is the mass fraction of the contaminant at the inlet of server i and $C_{cooling}$ is the mass fraction at the cooling source (the capture index cannot be greater than 1 or less than 0 and can be computed at the server, rack or aisle level). The capture index is useful in determining how much of the air that originates from the cooling source (i.e., perforated tiles) is ingested by the racks. Clearly, a high value of ψ indicates less mixing between the warm exhaust air and the cold supply air. The capture index requires that the

tracer be filtered out of the exit of the servers and can be easily computed using CFD but cannot be measured in actual data centers.

1.1.4.3 Rack Cooling Effectiveness

Herrlin (2005) proposed the use of the rack cooling effectiveness to determine if there is sufficient cooling of the racks, based on an industry standard for rack inlet temperatures. The rack-cooling index high (RCI_{HI}) is a measure of the absence of server over-temperatures, where 100% means no over-temperatures exist and 0% means all racks are over-temperature. RCI_{HI} is defined as:

$$RCI_{HI} = 1 - \frac{\sum_{i=1}^N (T_i^{in} - T_{\max-rec})_{T_i > T_{\max-rec}}}{N(T_{\max-all} - T_{\max-rec})} \quad (1.3)$$

Where, T_i^{in} is the temperature at server inlet i , N is the total number of servers, $T_{\max-rec}$ is the maximum recommended inlet temperature per some guideline and $T_{\max-all}$ is the maximum allowable inlet temperature per some guideline.

1.1.4.4 Return Temperature Index

Herrlin (2008) proposed the use of the return temperature index (RTI) as a measure of the energy performance of the data center's air-management system. The RTI is defined as,

$$RTI = \frac{\Delta T_C}{\Delta T_m}, \quad (1.4)$$

where, ΔT_C is the difference between the return and supply air temperature of the CRAH units and ΔT_m is the temperature rise of the IT equipment. Although intended to be energy metric, the fact that it is based solely on temperature makes it a measure of the level of by-pass or recirculated air. A value of 1.0 indicates no mixing between the hot and cold air streams – such as in the case of an enclosed cold aisle. A value above 1.0 suggests mainly recirculation of the hot air to the cold aisle, whereas a value below 1.0 suggests air is emanating from the perforated tiles but bypasses the rack.

1.1.4.5 Supply and Return Heat Indices

Sharma, Patel and Bash (2002) introduced the supply heat index (SHI) and return heat index (RHI). The supply heat index is a measure of the heat that infiltrates the cold aisle by recirculation. SHI is defined as,

$$SHI = \frac{\sum_{i=1}^N (T_i^{in} - T_{tile})}{\sum_{i=1}^N (T_i^{out} - T_{tile})}, \quad (1.5)$$

where, T_i^{in} , T_i^{out} are the inlet and exit temperatures of server i , respectively, and T_{tile} is the temperature of the supply air in the cold aisle. Assuming a constant specific heat and identical server flow rates, the numerator represents the sensible heat gained by the air in the cold aisle before it enters the rack and the denominator represents the total sensible heat gained by the air leaving the rack exhausts.

The return heat index was developed in order to investigate the degree of mixing the rack exhaust air undergoes before it returns to the CRAH units. RHI is defined as,

$$RHI = \frac{\sum_{k=1}^{N_c} \dot{m}_k c_p (T_k^{in} - T_{ref})}{\sum_{i=1}^N \dot{m}_i c_p (T_i^{out} - T_{ref})} = 1 - SHI \quad . \quad (1.6)$$

where, k is a summation over the N_c CRAH units, i is a summation over the N servers, \dot{m} is the mass flow rate and c_p is the specific heat. The numerator of RHI represents the total heat dissipation of the data center. Clearly, an increase in the server inlet air temperature would reduce RHI, concluding that the air undergoes a higher degree of mixing before entering the IT equipment. Higher values of RHI would indicate a better aisle design with less mixing. To satisfy the overall energy balance of the data center, the addition of RHI and SHI must equal unity.

1.1.4.6 Metrics Summary

This section outlined many of the common metrics used in the design and analysis of data centers within the IT industry. Herrlin and Compiano (2010) have developed a high-level software tool for displaying these metrics to the data center operator. Other metrics have also been proposed, including several by Tozer, Kurkjian, and Salim (2009) to understand by-pass flow, negative pressure and the balance of server and CRAH airflow. While these are useful metrics, they are impractical to obtain in day-to-day operation.

Herrlin and Compiano (2010) state that good insight into the operation of a data center

can be obtained by monitoring only the DCiE, RCI and RTI, all of which are obtainable under real-time operation. Sisk, Khalell, et al. (2009) detail the instrumentation, data acquisition and software systems for a mixed-use data center at Pacific Northwest Laboratory (PNNL) in order to evaluate real-time DCiE. The project was part of a larger research program focused on highly instrumented water and air-cooled data centers. Even in this research facility, they were unable to measure all the necessary DCiE data and missing data had to be supplemented with manufacturer specifications, highlighting the tremendous effort needed in collecting real-time information in data centers.

1.2 Previous Work

In recent years, as data center energy use has come under more scrutiny from the IT industry and government agencies, increased work has been done on understanding, assessing, modeling and optimizing data center design and operation in order to improve energy efficiency. Prior to the start of this work, many authors have studied several of the topics covered in this dissertation. Therefore, a review of existing literature has been incorporated under the following topics: thermal profiling, computational modeling of the airflow and temperature fields, cooling infrastructure modeling, enclosed aisle data centers, reduced-order modeling and load placement.

1.2.1 Thermal Profiling

Thermal profiling refers to the process of experimentally characterizing the “thermal map” of a data center. This process is extremely useful in understanding the thermal environment around the IT equipment. Schmidt (2004) was one of the first to measure the thermal profile of a high-density data center. He developed a methodology for collecting airflow, temperature and power data in operational data centers. The measurements showed that discrepancies in the CRAH airflow could be as large as 20% between the manufacturers’ published data and the actual flow rate, due to cable obstructions under-floor, dirty filters and turning vane installation in CRAH units. He also concluded that less than half of the airflow introduced in the data center space was coming from the perforated tiles themselves. Over 1/3 of the airflow was coming from cable cutouts and leakage. Temperature measurements throughout the data center showed that the air entering the racks was at a significantly higher temperature than the supply air. In many cases, the air was 15°C – 20°C warmer than the supply air, indicating a significant fraction of recirculated air. One section of the data centers showed elevated rack inlet temperatures due to short-circuiting of the cold supply air from the perforated tiles to nearby CRAH units. Lastly, it was concluded that although a significant fraction of the cold supply air was coming through leakage holes and cable cut-outs, this air was not effective at providing cooling to the racks. Following the guidelines of Schmidt (2004), Schmidt, Iyengar, Beaty and Shrivastava (2005) profiled another high-density data center. The conclusions drawn in their study were similar to those by Schmidt (2004). In both cases, it was also seen that the total supply flow (tile + leakage) rate was only 2/3 of the required rack flow rate; hence, at least 1/3 of the required server flow was coming

from recirculation of hot air from the rack exhaust. A collection of similar case studies on a number of high-density data centers was published by ASHRAE (2008c).

Karlsson and Moshfegh (2005) studied the temperature profile at the inlet of racks by using infrared cameras, to visualize the temperature distribution, and a traversing grid of thermocouples. The results showed that a temperature gradient existed along the racks, where higher inlet temperatures were seen at the top of the rack. This result is common in air-cooled data centers and has been well documented. Servers located in the bottom of racks in the center of the cold aisle receive 100% cold supply air. Servers toward the top of a rack and toward the end of an aisle, receive a mixture of cold supply air and warm recirculated air. Typically, the server located in the top of the rack located on the end of an aisle receives the most recirculated air and will therefore be the hottest (Figure 1.6).

1.2.2 Computational Modeling of Airflow and Temperature Fields

The development of easy-to-use computational fluid dynamics (CFD) tools has allowed numerous researchers and designers the ability to understand the impact of different design considerations. Shrivastava (2008) studied the impact of seven different airflow configurations, which were a combination of floor and ceiling supply diffusers and floor and ceiling returns. He concluded that the floor supply ceiling return and the floor supply room return configurations were the most effective at lowering the inlet temperature to the racks. Shrivastava also used a statistical variance based analysis to study the effects of ceiling height, tile flow rate and the location of return vents, concluding that the tile flow

rate had the largest impact on reducing rack inlet temperatures. Shrivastava (2008) concluded that increasing the ceiling height from 8 to 12 ft. had the effect of cooling the server's inlet temperatures. In a similar study, Sharma, Bash, and Patel (2002) performed several parametric studies, for a floor supply room return data center configuration in order to understand the effect that the cold aisle width, hot aisle width, ceiling height and spacing between the racks and the wall had on the inlet temperature to the servers. They concluded that the spacing between the racks and the wall had no effect on the inlet rack temperatures, increasing the cold aisle width reduced the heat infiltration by recirculation to the cold aisle and a decrease in the hot aisle width reduced the mixing occurring in the data center and improved inlet air temperatures. Schmidt, Cruz and Iyengar (2005) attempted to reduce the temperature of the exhaust air before it entered the racks by splitting the cold air distribution between the cold aisle and the hot aisle; however, this did not help in reducing the rack inlet temperatures.

1.2.2.1 Validation of Computational Models

The collection of data in operating data centers is a complex task. Several authors have discussed the need for more detail measurements in data centers. They conclude that intrusive measurements in data centers are near impossible because of reliability and uptime constraints. This fact forces the investigation of a data center's thermal map to be done using primarily computational fluid dynamics (CFD). Recognizing the need for this in the industry, several commercially available CFD packages have been developed to address data center design (Innovative Research, 2011; Applied Math Modeling, 2011).

Shrivastava, Iyengar et al. (2006) compared the experimental measurements obtained by Schmidt, Iyengar et al. (2005) to a CFD model using a commercially available CFD package. The model used the standard k- ϵ turbulence model and included the effects of buoyancy using the Boussinesq approximation. The model was of a 76' x 98' x 11' data center and used ~1.5M cells (for an estimated uniform grid size of ~ 4"). Detailed modeling of the racks and the under-floor plenum were omitted. The results compared the CFD to the measured inlet temperatures to the racks. Overall, 83% of the rack's inlet temperature predictions agreed to within 7°C. Iyengar, Schmidt, Hammann, and VanGilder (2007) performed a numerical/experimental comparison on a small test cell. They collected a dense set of temperature data at all locations in the data center raised floor space. The standard k- ϵ turbulence model was used. No server or under-floor plenum details were modeled. The results showed average temperature differences of 3.1°C, 3.2°C and 2.7°C, at heights of 0.5 ft, 4.5 ft and 8.5 ft, respectively. The results showed that the CFD under-predicted the mixing of the hot and cold jets before entering the racks. Several other comparison studies have been performed. Abdelmaksoud, Dang, Khalifa et al. (2010) improved upon the results of Iyengar, Schmidt, Hammann, and VanGilder (2007) by including several features that were not previously included, such as perforated tile flow and rack exhaust modeling to conserve both mass and momentum, inclusion of buoyancy and improved thermal boundary conditions on the floor. Noh, Song, and Chun (1998) showed on average, 2.4°C errors between experimental and numerical predictions for a 38 W/ft² data center. Patel, Bash, and Belady (2001) showed a 14% average discrepancy between CFD and experimental temperatures for a high-density data center. The largest errors were obtained at the top of the racks, where the

recirculation and mixing of the air streams is largest. The study by Patel et al. is one of the only studies that compared point-wise velocity measurements in the data center. They found errors of greater than 10% in all velocity measurements.

1.2.2.2 Using CFD in Data Centers

Validation studies of computational models in data centers have been quite limited. Full-field measurement techniques, such as particle image velocimetry or laser-induced fluorescence could be used to provide this data but are challenging to use in data centers because of the environmental constraints (Rambo and Joshi, 2007). Substantial work has been done to understand simplifications and assumptions in CFD modeling, including under-floor plenum modeling, perforated tile modeling, turbulence modeling, rack modeling and grid requirements.

1.2.2.2.1 Under-floor Plenum Modeling

In many CFD studies of the raised-floor space, the under-floor plenum is omitted from the computational domain and a uniform velocity boundary condition is used for all perforated tiles. Several studies have shown that the flow distribution through the tiles is not uniform, in which case omitting the under-floor plenum could result in a significant error with a uniform velocity boundary condition. Schmidt, Karki et al. (2001) experimentally measured the flow rate through individual tiles using a flow hood measuring device. Comparisons between CFD and experimental measurements of the plenum tile flow rate distribution were done in (Schmidt, Karki et al., 2001; Schmidt, Karki et al., 2004; VanGilder and Schmidt, 2005), showing that all CFD predictions had

errors greater than 10%, with certain areas showing greater than 100% error. These studies only compared the flow rate through the tiles and did not compare the actual temperature or velocity distribution within the plenum or the affect that the non-uniformity had on the inlet temperature of the racks. In all of these studies, the CRAH unit is modeled as black box, heat extraction device, with an assumed flow rate. Details about the pressure drop or exhaust geometry were omitted. Rambo and Joshi (2007) show that, based on manufacturer specifications, because of the pressure-flow characteristics of the blowers in commercially available CRAH units, large changes in blower flow rate are possible for relatively small changes in pressure resistance at a given fan speed. They concluded that CRAH units should be modeled by taking into account the blower characteristics.

1.2.2.2.2 Perforated Tile Modeling

Taking into account the resistance of the perforated tiles could have a significant effect on the distribution of flow in the under-floor plenum. Typical perforated tiles found in data centers range from 25% to 85% open area. Without significant obstructions under-floor, the perforated tiles are the largest pressure resistance external to the CRAH (or CRAC), which will dominate how uniform the flow rate is. When modeling the under-floor plenum, the tiles are typically modeled as a lumped resistance using the relationship $\Delta p = kV^2 + cV$, where the coefficients k and c can be found experimentally. While it has been shown that a porous jump condition is sufficient for modeling the pressure resistance of the tile, the perforations in the tile greatly affect the momentum of the jets leaving. To match the correct momentum, one could model the perforations in the tile

one-by-one; however, numerically this is not a feasible since the perforations are approximately an mm in diameter and such a fine computational grid would greatly increase the computational time. Abdelmaksoud, Dang et al. (2010) showed good agreement between experimental and numerical results for an isolated perforated tile by including an additional momentum body force term in the Navier-Stokes equations. This body force term corrects for the momentum deficit due to the perforations by either entraining air into the “side” of the momentum source volume or adjusting the pressure field around the volume.

1.2.2.2.3 Turbulence Modeling

Indoor flows are generally characterized by transitional or turbulent mixed convection. Traditionally in data center CFD modeling, the standard k- ϵ turbulence model has been used. The lack of validation data has forced CFD users to use the model without confidence in its accuracy. While it may not be the only reason for poor agreement, several studies (Shrivastava, Iyengar et al., 2006; Iyengar, Schmidt, Hamann, and VanGilder, 2007; Patel, Bash, Belady, 2001; Rambo and Joshi, 2007) have shown large discrepancies using the k- ϵ model. Cruz, Joshi, Iyengar, and Schmidt (2009) tested several turbulence models along with a laminar flow model on the test cell built by Iyengar, Schmidt, Hamann, and VanGilder (2007). It was concluded that the zero equation and Spalart-Allmaras turbulence models produced the lowest errors in temperature comparisons. Zhang, VanGilder, Iyengar and Schmidt (2008) found no large difference when they studied several different turbulence models. To date, turbulence

modeling in data centers is still an open research question, which can only be answered with higher-fidelity experimental data.

1.2.2.2.4 Rack and Server Modeling

Zhang, VanGilder, Iyengar and Schmidt (2008) have studied the level of detail needed in rack and server modeling. Three levels of detail were studied a) racks as a black box heat-addition device, b) server as a black box heat addition device but details of the rack door and frame are modeled, and c) the racks and server simulators are modeled exactly. Their computational model used some grid clustering but had typical grid sizes of approximately 6". They concluded that there was no difference in the results based on these three modeling techniques; therefore, the black box technique is recommended since it is the most computationally efficient. Rambo and Joshi (2007) showed that using a pressure-flow relationship for the fans in the rack could have a 10% variation in the net rack flow rate predicted by CFD compared to assuming a constant flow model. Khankari (2009) studied the recirculation patterns within racks and showed significant back-to-front recirculation when an opening is left inside the rack. They recommend the use of blanking panels to eliminate this.

1.2.2.2.5 Grid Requirements

VanGilder and Zhang (2008) studied the effect of performing coarse grid CFD for 10 different data center layouts. They concluded that the best return for computational time and accuracy was to use a 10" grid size, which resulted in an average computational time of 4 minutes and 87% of the rack's capture indices were predicted within 80% of a

baseline case that used a 2” grid size. Moving to a 6” grid size resulted in 92% accuracy in capture index but the computational time increased to 25 minutes. The study considered about 500 iterations to be a reasonable number to reach “sufficient convergence”. All computations were performed on a Pentium 4 3.0 GHz personal computer with 1 GB of RAM. They further commented that localized grid clustering has the potential of reducing solution time relative to uniform grid spacing, but do not give any recommendations about zones for refinement. These conclusions should only be used as recommendations and when performing an analysis for a different configuration, the necessary grid resolution will depend on a number of factors, including grid stretching, accuracy, turbulence model and y^+ , to name a few.

1.2.3 Cooling and Power Infrastructure Modeling

Reliability concerns in data centers forces the cooling infrastructure to be over-designed, which typically leads to lower efficiency operation of equipment. Even though off-design operation could have a significant energy impact, little work has been done understanding its impact on data center’s energy efficiency. Lui (2010) contributes this to the limitations in currently available simulations models in accounting for the dynamic behavior of the equipment. In a substantial review of data center modeling, Rambo and Joshi (2007) did not identify any work in this area.

Many legacy data centers employ a thermal management strategy that consists of a single sensor feedback signal, which acts as an indication of the heat being dissipated in the

room and controls the temperature of the CRAH's discharge air. This sensor is placed in the CRAH's return air stream (Bash, Patel, and Sharma, 2006). Several researchers have studied the dynamic optimization of CRAC/CRAH operation. Boucher et al. (2004) studied how variable control and actuation in the data center affects the state of the data center space. In their study, they looked at changing the CRAC supply temperature, changing the CRAC airflow rate using variable speed fans and variable open plenum tiles (electronically actuated floor tiles developed by the authors). They suggest the use of a feedback controller based on rack inlet temperature and CRAH supply temperatures. One of the goals was to control rack recirculation by varying CRAH fan speeds; however, the authors conclude that without advanced non-linear control methods (fuzzy logic, neural networks) this control strategy would not be ideal. They also comment that variable floor tiles could have a significant impact on local control of rack inlet temperature. Hayama, Enai et al. (2003) and Furihata, Hayama, and Enai (2003) looked to develop an air conditioning methodology that reduces the volume of supply air while maintaining proper cooling of IT equipment. They introduced a strategy that made tile damper adjustments that track with the heat load of the racks.

Bash, Patel, and Sharma (2006) experimentally tested a data center environmental control system that uses a distributed sensor network to manipulate conventional CRAC units. A cascaded control algorithm was used to evaluate the data from a sensor network and manipulate supply temperature and flow rate from individual CRACs to ensure thermal management with reduced operational expense. The "Dynamic Smart Cooling (DSC)" controller makes decisions based on predefined regions of influence for each CRAC unit

and the current rack inlet temperatures. Estimates of the power consumption were made based on a model developed by Patel, Sharma et al. (2002) of a simple direct expansion CRAC system relating the coefficient of performance (COP) to the supply air temperature. The DSC controller showed a reduction in energy consumption by as much as 58% compared to a conventional control strategy.

Iyengar and Schmidt (2007, 2009) developed an analytical design point model to predict the thermal performance and energy efficiency of the data center cooling loop. The model required inputs of the flow rates and pressure losses in each component. The model was applied to a 5.88 MW data center. The study concluded that the cooling infrastructure consumed 2.61 MW, with the chiller consuming 41% of the infrastructure's total (i.e., of the 2.61 MW). Several parametric studies were performed showing that a change in chilled water set point of 10°C resulted in an 8% decrease in total plant energy. Pelley et al. (2009) recognized that the development of a model that incorporates important dynamics was essential to understanding data center's energy consumption. As a starting point, they developed a simplified system level model of the data center subsystems. This model allowed for back-of-the-envelope computations that replace interactions with simple parametric models of the components. They showed that substantial savings were possible between a data center located in Ann Arbor, MI and Austin, TX because of the effect of ambient condition on chiller power.

Breen, Walsh et al. (2010) developed a simple thermodynamic and heat transfer based model to evaluate various operating strategies in data centers from the cooling tower to

the chip. They concluded that improved energy efficiency can be achieved by increasing the inlet air temperature to the server, because of the reduced need for refrigeration. In a second paper, Walsh, Breen et al. (2010) also studied the effect of various server fan algorithms on the energy consumption of the data center. They showed that increasing the inlet temperature to the servers is the recommended operating strategy only when the chip temperature can vary linearly with rack inlet temperature. Furthermore, they found that if a constant chip temperature is maintained, the trade-off between refrigeration and fan power must be considered in order to realize optimal energy-efficient operation. Breen, Walsh, et al. (2011) also studied the effect of chip leakage power on the overall COP_G of the data center and showed that increasing operating temperatures to the thermal limit of the IT equipment does not necessarily improve energy efficiency when the chips have a high leakage gradient. Iyengar, Schmidt and Caricari (2010) used CFD to study the energy savings from reducing the flow rate through CRAH units using motor speed control. It was found that reducing CRAH flow showed significantly more savings than changing the chiller set point temperature (12.6% vs. 3.6% of IT load). They also describe a systematic methodology for shutting down CRAH units.

Braun, Mitchell, and Klein (1987) addressed performance and control characteristics of a large scale cooling system at the Dallas/Fort Worth airport, by comparing measurements and several computational models. A variable-speed controlled chiller shows significant performance advantages compared to a fixed-speed vane controlled chiller.

Measurements showed energy savings as large as 40% during fall weather, especially with a small exit temperature difference between the evaporator and condenser and at low

load. He addressed the optimization of the cooling tower and condenser water and concluded that the region around optimal condenser water flow and cooling tower airflow is flat; therefore, the optimum operating point is relatively insensitive to either. The optimal operation corresponds to a cooling tower thermal capacity ratio of around 1.0, attributing to an energy savings up to 10%.

The works of Bejan and Ledezma (1996) and Bejan (1982, 1995) have highlighted the importance of considering the fan power requirement when optimizing the design of combined thermal/fluid/heat transfer systems, in order to minimize entropy generation. Much of this work has focused on using simple physics-based models that capture the most important characteristics of the system to uncover fundamental design trade-offs.

Hellmer (2010) used a design point (with no off-design performance) model to study the energy impact of a number of cooling systems using hour-by-hour weather data for various cities. He considered four systems: refrigeration only, refrigeration with dry coolers, refrigeration with water economizers and refrigeration with air economizers. Three humidification methods were also investigated: steam, ultrasonic and evaporative. A single operating point of 22°C supply air temperature, 26.7°C return temperature, and 40% RH was assumed. Based on the analysis, it was concluded that air economizers with evaporative humidifiers are the most energy efficient of the systems studied, although the savings are quite small for hot and humid climates. Regardless of the humidifier type, air economizers with humidifiers consumed less water than water side economizer or refrigeration systems. He concluded that typically, water economizers are less effective at

improving energy efficiency because they have a shorter operational season. The dry cooler model did not show any savings versus the refrigeration only model. This was due mainly to the added fan power required in the air handler.

Sorell (2007) provided one of the first looks at airside economizer use in data centers by looking at the range of enthalpy of outdoor air over a year for several cities. Use temperature bins, he concluded that in cities such as San Francisco, CA and London, England, OA economizer use is possible for more than 8000 hours/year. Even in hot and humid climates, such as Dallas, TX, it is anticipated that OA economizers could theoretically be used for more than half the year. Patterson, Atwood, and Miner (2009) conducted a ten-month test between two data centers at Intel's facility in New Mexico. One of the facilities was configured to be a standard data center with CRAC units supplying air at 20°C. The other data center implemented extensive use of airside economizers, which kept the supply air between 18°C and 32°C and only used the CRAC units when the air was above 32°C. Results of the study showed that for greater than 90% of the year, the energy savings were as high as 74% by using an airside economizer. In this study, neither the contaminant level nor the moisture of the OA was controlled. After the experiment, servers from both sides were sent back to the manufacturer for testing. The servers in the airside economizer were easily distinguishable because of the substantial dust build up both inside and outside; however, extensive testing by the manufacturer concluded that there was no corrosion and no cause for reliability concerns. Patterson et al. also argue that the dew point was frequently below the 5.5°C limit and there were no reliability concerns due to ESD since the servers were always grounded.

Both Wilman (2007) and Patterson et al. (2009) have showed an added benefit of air economizers in data centers. Even when the ambient temperature is high, it is still lower than the return air from the data center; therefore, even at high ambient conditions, mechanical refrigeration savings are possible by bringing in the OA. Davidson (2009) studied several dehumidification processes and their energy impact on airside economizers using a yearly temperature-RH bin method in Los Angeles, CA. The first approach looked at mechanical refrigeration to dehumidify, but concluded that because of the necessary reheat this process is not energy efficient. Second, the use of desiccant dehumidification was studied and showed that unless a source of waste heat with a temperature greater than 100°F was available, the desiccant dehumidification process was also energy inefficient.

ASHRAE (2008b) provides estimates of waterside economizer use in various cities based on wet bulb temperature. They estimate that for a 9°C chilled water set point, the number of available hours can range from 3% in Los Angeles, CA to 51% in Denver, CO. Increasing the chilled water set point temperature to 19°C significantly increased the number of hours to 68% and 93% for Los Angeles and Denver, respectively. Stein (2009) evaluated waterside economizers for several cities using DOE2.2 DesignDay simulations. He showed that using waterside economizers had the potential of reducing HVAC energy by 30%, even in warm climates. Lui (2010) discusses several design considerations for waterside economizers, including control strategies for waterside economizer operation in partial and full cooling modes.

1.2.4 Enclosed Aisle Data Centers

Enclosed aisle data center configurations clearly have the advantage of providing complete separation of the hot and cold air streams; hence, providing thermally uniform conditions to the inlet of the racks. However, the isolation of the cold aisle allows no recirculated air to be used in the cooling process and therefore 100% of the required rack flow must be provided through the perforated tiles in the enclosed aisle. In recent years, a number of white papers have been written showing significant energy savings using aisle containment (Moss, 2009a; Moss, 2009b; Neimann, 2008; Fink, 2008)

Schmidt, Vallury, and Iyengar (2011) discuss three different forms of aisle containment: cold aisle containment, hot aisle containment, and rack exhaust chimney containment. Using computational fluid dynamics, a case study is performed on an 8944 ft² high-density data center to compare the energy savings potential of implementing cold aisle containment to a conventional open aisle design – showing a potential of saving 59% of the energy required for CRAC units. The savings arise because in the open aisle configuration, the data center must be “flooded” with cold air in order to meet the rack’s inlet temperature constraint. In their example, the CRAC units must provide 2.5 times the required rack flow in the open aisle configuration to meet the rack’s inlet temperature constraint.

Gondipalli et al. (2009) recognized that in typical open aisle data centers, the CRAH units do not supply 100% of the required rack air and some recirculation is necessary. As a means of providing controlled recirculation in enclosed aisle configurations, they

investigated various types of openings placed on the roof and side barriers of the aisle enclosure. Simulations were performed and the configurations were compared based on the ability to provide acceptable inlet rack temperatures, added pressure drop across the servers and velocity in the enclosed aisle. It was concluded that placing slots at the bottom of the enclosure, porous roofs and doors, and solid roof and meshed doors all violated one or more constraint. They found that solid doors and a slit on the roof was an acceptable design because of its simple implementation and ability to meet all criteria. Villa (2010) experimentally investigated cold aisle containment for high-density applications. He concluded that the enclosed aisle approach allowed for higher-density loads than the typical open aisle.

1.2.5 Reduced-Order Modeling

Computational fluid dynamics has certainly been the tool of choice when evaluating the data center's airflow and temperature distribution. While CFD provides detailed descriptions of the data center air and temperature fields, its use in optimization or real-time control is impractical because of its requirement of large computational resources and lengthy execution time. Reduced order models (ROM) allow for statistics-based representations of the physics describing complex systems. These ROMs can typically be run in real or near-real time. Several attempts have been made at describing the nonlinear air and temperature fields of a data center using reduced order models.

VanGilder, Zhang and Shrivastava (2007) developed a CFD-based partially-decoupled aisle method (CFD-PDA), where a cold aisle or hot aisle is analyzed as if it were decoupled from the rest of the room. Within the aisle, a first-order CFD solver is used assuming the “top” boundary condition is “open” and along the ends, a prescribed inflow or outflow boundary is assigned. The end airflow is modeled by linear regression using the model developed by Shrivastava, VanGilder, and Sammakia (2006, 2007), where the airflow at the rack boundaries is correlated to rack power and airflow distribution, supply air temperature, ambient reference temperature and the length of the cluster. Several applications of the model are presented in which the capture index at each of the racks in a typical hot aisle/cold aisle data center was computed. The CFD-PDA method shows errors greater than 10% for all comparisons compared to the full room CFD. The sources of error were mainly due to the assignment of boundary conditions. VanGilder and Shrivastava (2006) also developed a more simplified model, which uses the same end airflow boundary condition, but assumes the flow within the cold aisle can be computed using potential flow theory and the superposition of solutions. Shrivastava (2008) used the CFD-PDA model to develop a database of solutions that was used to train and develop a neural network, which was used to compute the capture index for a single cluster hot aisle. The neural network was capable of predicting the CFD-PDA results to within 10% over 95% of the time; however, by using the CFD-PDA method, greater than 10% error in the capture index prediction was already introduced.

Healey, VanGilder, Sheffer, and Zhang (2011) used potential flow modeling as a fast technique for obtaining a data center’s airflow and temperature field. A 3-dimensional

potential-flow numerical model was used to study 8 different data center layouts with rack heat loads in the range of 2 kW to 12 kW. The results were compared to CFD solutions and showed that errors as large as 5°C were seen in many of the rack's inlet temperature predictions using potential flow because of the inability of potential flow to capture turbulent mixing and buoyancy. Yarlanki, Das, Hamann et al. (2011) used potential flow modeling to predict thermal zones in the data center by tracing streamlines from the exit of the CRAH units to the rack inlets. Thermal zones are the regions of influence each CRAH unit has on the raised floor space. Similarly, Lopez and Hamann (2011) formulate the problem of finding thermal zones as a boundary value problem for convective transport. This technique leads to easy identification of thermal zones in post-processing. In order to derive thermal zones, Li and Hamann (2011) use a statistics-based approach to correlate temperatures measured at sensors located throughout the data center to the supply temperature of the CRAHs.

Tang, Mukherjee, Gupta, and Crayton (2006) developed an abstract heat flow model, where the recirculation to server, i , is expressed as a coefficient that represents the percentage of server i 's flow that originates from every other server. A cross interference matrix of these coefficients is determined using a database of CFD solutions for a given data center geometry. The abstract heat model was compared to full CFD and showed significant errors in predicting the actual temperature; however, the trends in inlet rack temperature were captured. Inherent in the development of abstract heat flow model is the assumption of temperature linearity (i.e., the temperature field changes linearly with

supply air temperature), which is an appropriate assumption only if buoyancy and radiation effects are negligible.

Using an artificial neural network, Moore, Chase, and Ranganathan (2006b) proposed a methodology to infer a detailed thermal map of the data center from a stream of instrumentation data that includes workload distribution, CRAC airflow rate, CRAC supply temperature and CRAC and server location. Computational fluid dynamics simulations were used to test the applicability of the neural network for different workload placement scenarios. Seventy-five out of three hundred and sixty simulations were used for training the neural network and the remaining were used for validation. The learning process showed that over 75% of the predictions were within 0.5°C and 92% were within 1°C. Moore, Chase, and Ranganathan (2006a) also attempted to derive thermal maps using a combination of CPU workload and internal component temperatures using a neural network, taking into account the time dependence of the temperature by keeping previous measurements in the machine learning function. The training of the network required the collection of external and internal temperatures and CPU workload data for each server in the data center. An application of the neural network showed predictions of the inlet temperature to within 1°C for over 80% of the predictions and within 1.5°C for over 90% of the predictions. However, because of differences in servers, a commissioning phase would be needed for each data center. Song, Murray, and Sammakia (2011) used artificial neural networks as a predictive tool for obtaining individual tile flow rates and rack inlet temperatures as a function of plenum height, tile perforation, and air leakage.

Elhadidi and Khalifa (2005) and Khalifa, Elhadidi and Dannenhoffer (2007) were the first to use proper orthogonal decomposition (POD) to model the indoor environment. Specifically, they used POD to compute the distribution of contaminants in a large open space, where the assumption of a “well mixed” condition is not appropriate, and the spatial gradients of velocity or concentration are of practical significance. Further, they developed a methodology to couple the POD solution to a lumped-parameter flow network zonal model that relies on the “well-mixed” assumption. This approach was applied to a typical office building connected to a large, non-uniform atrium space in order to compute the flow and contaminant fields throughout the office and atrium in near-real time. Samadiani, Joshi et al. (2009) used POD to derive the thermal map of a data center as a function of CRAH flow rate. The POD ROM was developed from a series of full field temperature measurements for seven different CRAH flow rates. Two other data sets were obtained to test the interpolation ability of the reduced order model. Mean temperature errors of 0.60°C and 0.75°C were seen for these two cases; however, there were maximum errors in the domain of 6.1°C and 8.2°C .

1.2.6 Load Placement

A growing trend in data center operation is to reduce the amount of operational IT equipment by either moving the workload amongst the operational servers (this is often referred to as virtualization) or consolidate the workload to a smaller number of machines in order to reduce the IT energy consumption. The difference is that typically the servers

are completely shut off during consolidation. However, many factors influence why a data center operator would choose not to use a consolidation strategy - including the time necessary to restart a computer after it has been shut down and the effect on the airflow and temperature distribution in the data center. A common strategy for load placement is to use software-based virtual machines. A virtual machine is essentially a “server within a server”. The host server operating system creates a virtual execution environment in which a guest operating system can execute. This allows any computational task to be executed on any server and be moved efficiently amongst all servers in a data center. Studies have shown that a data center’s workload intensity could change by a factor of 3 to 7 depending on time-of-day (Rolia, Singhal and Friedrich, 2000). The operational question to answer becomes, how does one place the necessary computational tasks amongst the servers in the data center, so that the energy consumption of the data center is reduced without violating the inlet temperature constraint of the hottest server?

Sharma, Bash et al. (2005) were the first to address this problem. They proposed the idea of dynamic thermal management, where power consumption is adjusted based on sensor readings in order to find workload placement scenarios that promoted a uniform temperature distribution in the data center. They scaled the power in each rack based on the difference between its exhaust temperature and the cold aisle temperature. Using CFD, they showed a reduction in hot spot temperatures of up to 3°C.

Moore, Chase, Ranganathan, and Sharma (2005) developed several algorithms for workload placement, which reduce the energy consumed by the cooling infrastructure.

Their workload placement algorithms had two goals: 1) prevent server inlet temperature from crossing a pre-defined “safe” threshold and 2) maximize the temperature of the air the CRAC’s supply air. Similarly, Tang, Gupta, and Varsamopoulos (2007) investigated several job placement strategies, including those proposed by Moore et al. (2005), and compared them to an optimal job placement strategy that was determined using a genetic optimization algorithm, which was developed by Tang (2009). In both pieces of work, the power consumption of the cooling equipment was computed by a coefficient of performance that was a monotonically increasing function of the CRAC’s supply air temperature only.

The works of Moore et al. (2005), Tang (2009), Tang, Gupta, Stanzione and Crayton (2006), and Tang, Gupta, and Varsamopoulou (2007, 2008) proposed several workload placement algorithms. The proposed algorithms were:

- *Uniform Workload* – the total workload is distributed evenly over all servers.
- *Coolest Inlets* – assign the workload to the fewest number of servers possible starting with the servers that have the coldest inlet temperature.
- *One Pass Analog* – scale the workload based on the outlet temperature in an attempt to create a uniform outlet temperature for all servers.
- *MinHR* – assign the workload starting with the servers that recirculate the least heat to the cold aisle.

Moore (2005) applied these algorithms to a data center with four rows of seven racks showing that the *MinHR* algorithm outperformed all other algorithms. However, in order to use *MinHR*, the heat recirculation fractions had to be obtained using computational fluid dynamics simulations. This calibration phase took 56 hours to complete for a relatively small data center and needs to be re-done anytime a configuration change is made. If servers not in use could be turned off completely, *MinHR* could cut cooling costs by 33%. However, the heat recirculation fractions were obtained from a data set that did not consider turning off server completely and no update to the heat recirculation matrix was done when servers were shut off.

Tang, Gupta, and Varsamopoulos (2008) and Tang (2009) also studied optimum load placement by using the reduced order abstract heat flow model for a small-scale data center (9.6m x 8.4m x 3.6m) with two rows of five racks. Tang, Gupta, and Varsamopoulos (2008) used a genetic optimization algorithm and showed 24% to 35% reduction in energy consumption compared to *Uniform Workload* and *One Pass Analog*. The genetic optimization algorithm of Tang et al. typically placed the workload beginning with servers that were in the upper portion of racks.

Moore et al. (2006b) further studied job placement using a neural network reduced order model to find the optimum workload placement for a given thermal map using a coordinate-space search algorithm and compared the results to the above described workload placement algorithms. Depending on IT utilization level, this optimization algorithm resulted in a 13% - 25% reduction in cooling energy compared to a *Uniform*

Workload. However, the results of the simpler *MINHR* algorithm were comparable to the full optimization, both showing savings depending on utilization level.

Mukherjee, Banerjee, Varsamopoulos, and Gupta (2009) investigated the challenges and benefits of extending the previous work, by Tang (2009) on spatial load placement, to include temporal variations. In his work, he was given a number of jobs with start time, duration and finish time. The goal was to determine how to order the jobs spatially and temporally to minimize energy consumption. The method uses the slack between the user-estimated execution time and the actual execution time of the job to determine how much delay could be introduced to improve energy efficiency without violating the user's expectations. Mukherjee's work highlighted that workload placement solutions that rely on detailed optimization algorithms to determine resource allocation are too computationally intensive to be practical in operational data centers. In his work, computing the workload placement using the genetic optimization algorithm took ~2.5 hours for a small data center.

Shrivastava (2008) studied the optimization of arranging racks and in-row coolers in the cold aisle using the CFD-PDA tool. The goal was to distribute the rack heat loads in order to maintain all CI values above 80%.

1.3 Best Practices for Data Center Design

Schmidt and Iyengar (2006) compiled a substantial review of design recommendations, which was later compiled and expanded into an ASHRAE publication (ASHRAE, 2008b). Rasmussen and Torell (2007) developed a systematic procedure for establishing a floor plan in raised floor data centers to provide effective air distribution to the IT equipment. A summary of these best practices is given below.

1. Ventilation Design

- a. The best traditional ventilation scheme is a raised floor for chilled air and a ceiling return for exhaust.
- b. A layout that allows hot air unobstructed access to the return of the CRAC helps at lowering the server inlet temperature.
- c. Airflow through the perforated tiles becomes more uniform when all CRAC units discharge air in the same direction.

2. Raised Floor Plenums

- a. Increased plenum height leads to a more uniform airflow distribution.
- b. Low flow or back-flow can occur when tiles and IT equipment are placed near CRAC units due to the low static pressure under-floor in these locations.

3. CRAC unit placement

- a. Turning vanes and baffles could reduce CRAC flow by as much as 15%.

- b. Dampers should not be used in perforated tiles because of the increased pressure resistance.

4. Rack Placement

- a. Place high-power IT equipment at floor locations that have high static pressure under-floor. This allows the highest possible airflow in the cold aisle. Typically, this is furthest away from the CRAC units.
- b. To get the expected hot-aisle, cold-aisle behavior, at least seven racks should be placed together (Rambo and Joshi, 2007).

1.4 Research Gaps in Data Centers

The review of existing literature related to energy efficient data centers revealed several key areas in which further research is merited.

1. A limited number of studies have been done to understand the optimization space of air-cooled data centers because the design space is so large that a formal optimization using CFD is computationally very prohibitive. To limit the range of options and parameters to be explored in the resource-intensive more rigorous analyses, simple physics-based models that capture the important thermal-fluid characteristics of data centers should be used in the early stages of the conceptual design process to define energy-saving approaches and near-optimum design and operating parameters

2. Recent industry guidelines for improving the energy efficiency of data centers have focused on increasing the allowable inlet air temperature for IT equipment. These guidelines make sense only when the refrigeration power dominates the facility energy consumption. However, benchmarking studies have shown that the fans and blowers in computer room air conditioning units consume a significant amount of power. Therefore, the optimal design and operation of data center systems must consider all energy-consuming components.
3. Several authors have acknowledged the need for energy simulation tools that can account for the dynamic operation of data center's cooling infrastructure. The need for higher fidelity, experimentally validated models that consider the off-design performance of the cooling and power distribution equipment are essential for studying energy efficient solutions to data center design and operation.
4. Several researchers have developed load placement algorithms that reduce the infrastructure energy consumption; however, many of these algorithms had the disadvantage of requiring a long computational time in order to find the optimum placement of IT loads amongst the operational servers in the data center. Furthermore, none of these studies considered the optimization of job placement and the cooling infrastructure. The development of physics-based, heuristic guidelines that use knowledge of the thermal environment through real-time measurements are imperative for the energy efficient operation of data centers.

1.5 Objectives & Scope

The research described herein has two main objectives, whose aim is to improve the overall energy efficiency of data centers.

1. Develop physics-based thermodynamic models to enable the holistic understanding, prediction and optimization of the energy consumption and the heat transfer phenomenon in a data center.
2. Formulate control methodologies that enable so-called thermally aware, energy optimized load placement.

The development of thermodynamic models is necessary in order to facilitate the study of optimized load placement. However, much insight can be gained through model development and application. Therefore, several intermediate studies are done to understand data center systems:

- 1.a. Develop a simplified physics-based model of the data center's airflow, temperature and energy characteristics for both enclosed and open aisles.
- 1.b. Verify several of the assumptions in the simplified model using computational fluid dynamics including, the practical implementation of bypass recirculation, recirculation non-uniformity, leakage, and buoyancy.
- 1.c. Develop a high-fidelity holistic thermo-hydraulic model of the data center's energy and power infrastructure – from the rack to the cooling tower.

- 1.d. Experimentally validate the thermo-hydraulic model against an operating 6MW data center in Poughkeepsie, NY.
- 1.e. Use the thermo-hydraulic model to investigate the effect of ambient conditions and chilled water set point temperature on the data center's energy efficiency.

With a validated modeling methodology in place, the development of thermally aware, energy optimized load placement strategies proceeds as follows:

- 2.a. Study several practical options for implementing load placement in open aisle data centers based on feature-based heuristics or real-time sensor measurements.
- 2.b. Perform a detailed analysis of the energy optimization of load placement in homogenous enclosed aisle data centers, including the effect of ambient temperature and rack temperature rise.
- 2.c. Develop a reduced order model of the data center's temperature field in order to facilitate a detailed analysis of optimized load placement in open aisle data centers.
- 2.d. Perform a detailed analysis of the work in 2.a. using the reduced order model in 2.c. to develop rules for implementing optimized load placement in operating data centers at a range of IT utilization and ambient conditions.
- 2.e. Study the effect of server control on the optimum operation of the cooling infrastructure
- 2.f. Develop optimized strategies for increasing IT load demand.

2.g. Compare the dynamic operation of the load placement rules to the steady-state studies performed thus far.

1.6 Importance of Work

With the rapid increase in worldwide data center electricity use, at nearly 54% in the last 5 years, it is becoming imperative to improve their energy efficiency. To complicate this, changing the mindset of data center operators is becoming increasingly difficult due to the mission critical nature of the application. Recently, data center best practices have focused on increasing the temperature of the thermal environment around the IT equipment in order to save energy on refrigeration. The hypothesis of this work is that there are further efficiency improvements possible by considering the combined optimization of the entire data center and cooling infrastructure systems. This work proposes a systematic approach to modeling the data center's systems – from the rack to the cooling tower – by employing simplified physics-inspired models, experimentally validated lumped parameter models and high fidelity computational fluid dynamics models. This research produced significant results, including, proving that the optimum operation of the data center must consider both the power required for refrigeration and the power to move the air throughout the data center, a novel technique for reducing the energy consumption of enclosed aisle data centers, the development of experimentally validated models for estimating the energy consumption of the entire data center system, reduced order methodologies for predicting the rack's inlet temperature distribution caused by the complex airflow and temperature patterns in the data center, and strategies

that use real-time sensor measurements to implement efficient load placement techniques in operating data centers.

2 Modeling Considerations in Data Centers

Modeling the energy consumption of a data center requires consideration of both the cooling infrastructure and the raised floor space. The cooling infrastructure modeling necessitates the characterization of the off-design performance of the power and cooling equipment, from the CRAH to the cooling tower, which is subject to fluctuations in ambient conditions, IT load and operation. The raised floor space modeling entails complex airflow and temperature patterns that emerge because of the interaction between forced convection (CRAH and rack fans) and natural convection (buoyancy driven flow arising from temperature gradients). This section focuses on the development of models to evaluate the energy consumption of a data center. As a starting point, a simple, physics-based model is developed, which is extremely useful in the early stages of the conceptual design process to define energy saving approaches and near-optimum design and operating parameters. Second, a higher fidelity, coupled thermodynamic-hydraulic model of the data center's cooling infrastructure is presented, which characterizes the off-design performance characteristics of all the cooling, power, and hydraulic systems in the data center. Finally, the fundamentals of computational fluid dynamics are presented in order to describe the turbulent airflow and temperature fields in the data center.

2.1 Simple, Physics-based Model of a Data Center

This section focuses on the development of a simplified physics-based model of an air-cooled data center. The model explicitly incorporates the effect of recirculation non-uniformity – a measure that characterizes the inlet temperature non-uniformity over an aisle of server racks. Furthermore, a simplified thermodynamic/heat transfer-based model of data center air recirculation and thermal balance and their effects on cooling infrastructure power consumption are given. Further details are provided in Khalifa and Demetriou (2010).

2.1.1 Rack Inlet Temperature Non-Uniformity

Consider a cluster of identical servers arranged in racks and organized in a cold aisle/hot aisle configuration. The air entering each server will be a mixture of cold air discharged by the CRAHs and issuing from the perforated tiles and recirculated warm air entrained from the rack exhaust. The fraction of server total flow that is recirculated exhaust air can be expressed by the symbol ϕ . We expect that, with open aisles, the servers close to the bottom and center of the cold aisles will be receiving cold tile air and essentially zero recirculation. On the other hand, servers close to the top and edges of the aisle will be receiving a considerable fraction of recirculated warm exhaust air. For a given constant server flow rate, \dot{V}_s , we will assume that the recirculation fraction of server j in rack i is $0 \leq \phi_{ij} \leq 1$, or the fraction of cold supply air entering server (i, j) is ψ_{ij} , which is identical to the so-called capture index defined by VanGilder and Shrivastava (2007).

Since the server inlet temperature is a resultant of the mixing of cold supply air issuing from the perforated tiles and recirculated hot air from the exhaust of the servers, we will base this analysis on the recirculation fraction, φ_{ij} . It is expected that φ will be a function of the location of a server in the aisle, server power, server flow rate and cold air supply rate. A detailed knowledge of this function can only be obtained from exhaustive measurements or CFD simulations of the data center's flow and temperature fields. For this simple analysis, only one important characteristic of the recirculation fraction is considered, a simple measure of its non-uniformity over the inlet face of the racks in an aisle, defined here as the ratio, θ , of the maximum recirculation fraction, φ_{max} , to its average value over the inlet face of the racks in an aisle.

$$\theta = \frac{\varphi_{max} - \bar{\varphi}}{\bar{\varphi}} = \frac{\bar{\psi} - \psi_{min}}{1 - \bar{\psi}} \quad (2.1)$$

A value of $\theta = 0.0$ corresponds to a uniform temperature over the racks inlet face, which may be achievable with enclosed aisles. It is expected that θ will depend on rack and tile layout, the cold air supply flow rate, and server temperature rise.

It should be emphasized here that $\bar{\psi}$ is a measure of the capture index and should not be interpreted as the ratio of the aggregate tile flow to the aggregate server flow. This ratio is typically higher than $\bar{\psi}$ because some of the air emanating from the tiles bypasses the

servers altogether and blends with the server exhaust. In this analysis, such bypass flow is treated as leakage (i.e., part of λ).

2.1.2 Thermal Analysis

A simple model for a server, rack or a group of servers/racks in an aisle is depicted in Figure 2.1. In this model, all the servers have been aggregated into a single “super” rack operating at an IT power P_m .

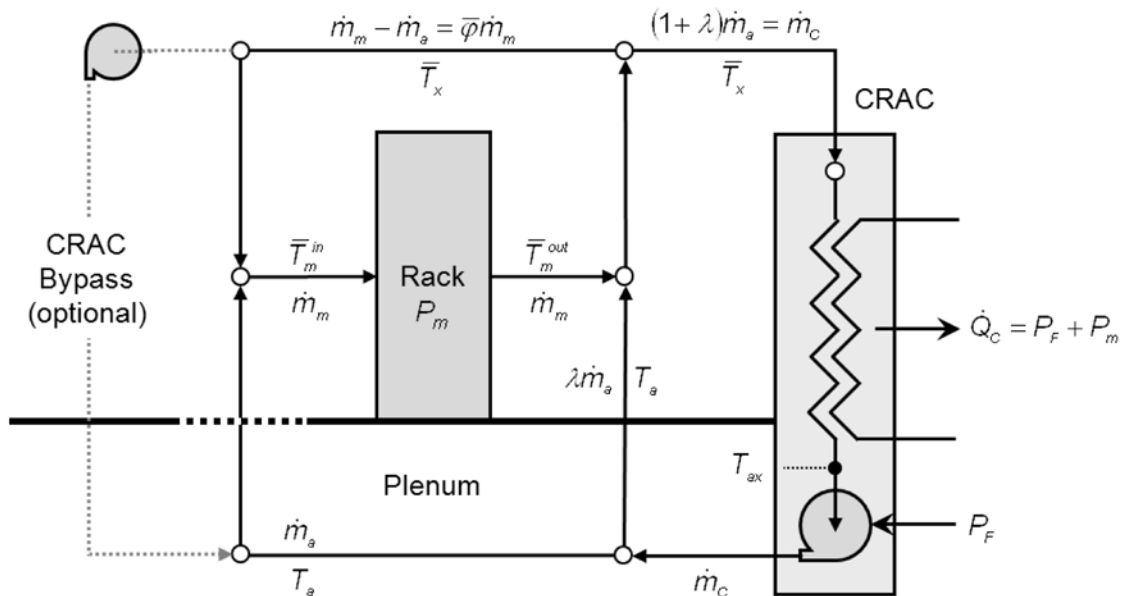


Figure 2.1 - Simple Model of an Air-Cooled Data Center

The rack is cooled by a stream of mixed air m_m entering the rack at an average mixed temperature T_m^{in} . From an energy balance, the average rack exit temperature is given by,

$$\bar{T}_m^{out} = \bar{T}_m^{in} + P_m / (\dot{m}_m c_p) = \bar{T}_m^{in} + \Delta T_m \quad (2.2)$$

The cold air at T_a from the under-floor plenum will be divided into an active cooling component, \dot{m}_a that eventually enters the racks after mixing with recirculated air, and a leakage component, $\lambda \dot{m}_a$, that bypasses the racks and blends homogeneously with the data center air. This leakage includes all the cold air that is not captured by the racks (i.e., leakage through cable cutouts, flow that “escapes” from the cold aisle and blends with the air in the data center space, etc.). Therefore, the data center’s exhaust air temperature, which is also the CRAC inlet temperature, will be given by,

$$\bar{T}_x = T_a + \frac{P_m}{(1 + \lambda)\dot{m}_a c_p} \quad (2.3)$$

The exhaust temperature must also satisfy an energy balance at the rack exhaust. Given by,

$$\bar{T}_x = \frac{\dot{m}_m \bar{T}_m^{out} + \lambda \dot{m}_a T_a}{\dot{m}_m + \lambda \dot{m}_a} = \frac{\bar{T}_m^{out} + \lambda(1 - \bar{\varphi})T_a}{1 + \lambda(1 - \bar{\varphi})}. \quad (2.4)$$

If we assume that all the recirculated air entering the servers is at the average exhaust temperature \bar{T}_x , then the average rack inlet temperature must obey,

$$\bar{T}_m^{in} = \bar{\varphi}\bar{T}_x + (1 - \bar{\varphi})T_a. \quad (2.5)$$

The average recirculation fraction, $\bar{\varphi}$, in Equation 2.5 is defined as,

$$\bar{\varphi} = 1 - \frac{\dot{m}_a}{\dot{m}_m} = 1 - \bar{\psi} \quad (2.6)$$

Combining Equations 2.3 and 2.5 together with the definition of ΔT_m from Equation 2.2 yields,

$$\bar{T}_m^{in} = T_a + \frac{\bar{\varphi}\Delta T_m}{(1 + \lambda)(1 - \bar{\varphi})} \quad (2.7a)$$

$$\frac{\bar{T}_m^{in} - T_a}{\Delta T_m} = \frac{1}{(1 + \lambda)(1/\bar{\varphi} - 1)} \equiv \Delta\tau_m^{in} \quad (2.7b)$$

For the server experiencing the highest recirculation (i.e., the highest inlet temperature),

$$T_{\max}^{in} = T_a + \frac{\varphi_{\max}\Delta T_m}{(1 + \lambda)(1 - \bar{\varphi})} \quad (2.8a)$$

$$\frac{\bar{T}_{\max}^{in} - T_a}{\Delta T_m} = \frac{(1 + \theta)}{(1 + \lambda)(1/\bar{\varphi} - 1)} \equiv \Delta\tau_{\max}^{in} \quad (2.8b)$$

For a given cold air supply rate and rack flow rate (i.e., for a given $\bar{\varphi}$ there is a maximum cold air supply temperature T_a^* beyond which the redline constraint T^* (i.e., that temperature defined by the ASHRAE guidelines to maintain reliability) will be violated. This temperature is given by,

$$T_a^* = T^* - \frac{(1 + \theta)\Delta T_m}{(1 + \lambda)(1/\bar{\varphi} - 1)} \quad (2.9a)$$

$$\frac{T^* - T_a^*}{\Delta T_m} = \frac{(1 + \theta)}{(1 + \lambda)(1/\bar{\varphi} - 1)} \equiv \Delta \tau^* \quad (2.9b)$$

The left-hand-side of Eq. (2.9b) is the normalized temperature range at the inlet of the racks. As a measure of rack inlet temperature non-uniformity, the normalized temperature range is higher the higher the value of θ and the higher the average recirculation fraction (i.e., lower $\bar{\varphi}$).

The relationship between rack inlet temperature non-uniformity and θ can be further elucidated by dividing Equation 2.8b by Equation 2.7b to yield the following simple relationship between the recirculation non-uniformity parameter, θ , and the corresponding temperature non-uniformity at the inlet face of the racks, expressed here as a temperature rise above the cold air supply temperature, T_a .

$$\theta = \frac{T_{\max}^{in} - \bar{T}_m^{in}}{\bar{T}_m^{in} - T_a} \quad (2.10)$$

A limiting ideal case arises if by some means it was possible to supply each and every operating server with air at the same inlet temperature (i.e., if the flow entering the servers was thermally uniform, and composed of a fraction $\bar{\psi}$ of cold air supply and a controlled fraction $\bar{\phi} = 1 - \bar{\psi}$ of hot air recirculation). We note that under these conditions, $\theta = 0.0$ and Equation 2.9a dictates a definite relationship between $\bar{\phi}$ and its conjugate cold air supply temperature, T_a^* . We note here the special case corresponding to $\bar{\phi} = 0$, (i.e., the case of no recirculation at all). This is the case of the enclosed aisle in which $T_a^* = T^*$ (i.e., the air is supplied at the maximum possible temperature – the redline temperature).

2.1.3 Cooling Infrastructure Power Consumption

The CRAC power consumption P_C comprises two major components that account for the dominant part of the cooling infrastructure power consumption. A refrigeration component, P_R , that is expected to be higher the lower the CRAC exit temperature, T_a , and the higher the outdoor temperature, T_{oa} , and an air moving (fan/blower) component, P_F , that is expected to vary with the CRAC air flow rate, \dot{V}_C . The variation in fan power is given by,

$$P_F \approx \frac{K_C N_C \dot{V}_C^3}{\eta_F}, \quad (2.11)$$

in which η_F is the overall efficiency of the CRAC fans, assumed here to be driven by controllable variable-speed motors. Therefore, the overall cooling power consumption is given by,

$$P_C \approx \frac{K_C N_C \dot{V}_C^3}{\eta_F} + P_R(T_a, T_{OA}, \dot{Q}_C), \quad (2.12)$$

where, the CRAC cooling capacity, \dot{Q}_C is given by,

$$\dot{Q}_C = P_m + P_F + P_P. \quad (2.13)$$

P_P is the sum of other parasitic cooling loads due to lights, wall heat transfer, etc., which we will neglect in this analysis. The power consumption of the refrigeration system is given by,

$$P_R = \frac{\dot{Q}_C}{COP(T_e, T_c, \eta_C)}, \quad (2.14)$$

where, COP is the refrigeration system's overall coefficient of performance. For a vapor-compression refrigeration system, the COP can be represented reasonably well as a function of evaporator temperature T_e , condenser temperature T_c and the compressor overall efficiency η_c . It is convenient to express the real COP as a fraction, ε_r , of the ideal Carnot COP at the same T_e and T_c ,

$$COP(T_e, T_c, \eta_c) = \varepsilon_R(T_e, T_c, \eta_c) \frac{T_e}{T_c - T_e} \quad (2.15)$$

in which Carnot efficiency $\varepsilon_r \leq \eta_c$ accounts for all irreversibility in the real vapor-compression cycle.

The saturated evaporator temperature, T_e must be lower than the cold air supply temperature and the saturated condenser temperature must be higher than the outdoor ambient temperature (i.e., dry bulb or wet bulb temperature, depending on the type of heat rejection system). The temperature difference between T_a and T_e is expected to depend on the evaporator or CRAC heat exchanger design (flow arrangement, UA, etc.), airflow rate and evaporator load. Similarly, the difference between T_c and T_{oa} is expected to depend on condenser design (flow arrangement, UA, etc.), coolant (air or water) flow rate, and the condenser load, given by,

$$\dot{Q}_o = \dot{Q}_c \left[1 + \frac{1}{COP} \right]. \quad (2.16)$$

Further, because the blowers in typical CRACs/CRAHs are placed downstream of the heat exchanger, the CRAH exit temperature, T_a , will be higher than the air temperature exiting the CRAC/CRAH heat exchanger (evaporator for a CRAC) owing to the CRAC/CRAH fan/motor heating effect, which can be significant in air-cooled data centers. Therefore, the CRAC heat exchanger exit temperature is given by,

$$T_{ax} = T_a - \frac{P_F}{(1 + \lambda)\dot{m}_a c_p}. \quad (2.17)$$

It is this temperature, not T_a , that must be used in determining the appropriate evaporator temperature to be used in *COP* calculations. This is an important distinction when the fan/motor power dissipation is significant, as it is in a data center application.

To compute T_e and T_c , simplified effectiveness-NTU relationships for heat exchangers with one of the two fluids changing phase (i.e., boiling or condensing) are used (Incopera and DeWitt, 2005):

$$T_e = \bar{T}_x - \frac{T_x - T_{ax}}{1 - \exp\left[-\frac{UA_e}{(1 + \lambda)\dot{m}_a c_p}\right]} \quad (2.18)$$

$$T_c = T_{OA} - \frac{\dot{Q}_o / \dot{m}_o c_{po}}{1 - \exp\left[-\frac{UA_c}{\dot{m}_o c_{po}}\right]} \quad (2.19)$$

Here c_{po} and \dot{m}_o are the specific heat and mass flow rate of the condenser coolant, and T_{oa} is the outdoor ambient temperature. Equation 2.19 can be further simplified by observing that for a constant \dot{m}_o and UA_c , $(T_c - T_{OA})$ is directly proportional to \dot{Q}_o .

$$T_c \cong \frac{T_{OA} + \chi \dot{Q}_c (1 - 1/\varepsilon_R)}{1 - \frac{\chi \dot{Q}_c}{\varepsilon_R T_e}} \quad (2.20)$$

Lastly, a relationship is needed to determine the ratio of real *COP* to Carnot *COP*, ε_r , which was expressed as a function of the saturation evaporator and condenser temperatures and compressor efficiency. For this analysis, ε_r was obtained by computing the performance of a vapor-compression refrigeration cycle, with 5°C subcooling and 0°C of superheat that uses R-134a as a working fluid. The vapor compression cycle consists of four processes:

- Process 1-2: compression of the refrigerant to the condenser pressure
- Process 2-3: heat transfer from the refrigerant as it flows at constant pressure through the condenser
- Process 3-4: throttling to a two-phase liquid-vapor mixture

- Process 4-1: heat transfer to the refrigerant as it flows at constant pressure through the evaporator

Irreversibilities in the compression process were considered by using a map of the isentropic efficiency of a typical single-screw compressor with a variable volume ratio (ASHRAE Systems and Equipment, 2000), as shown in Figure 2.2.

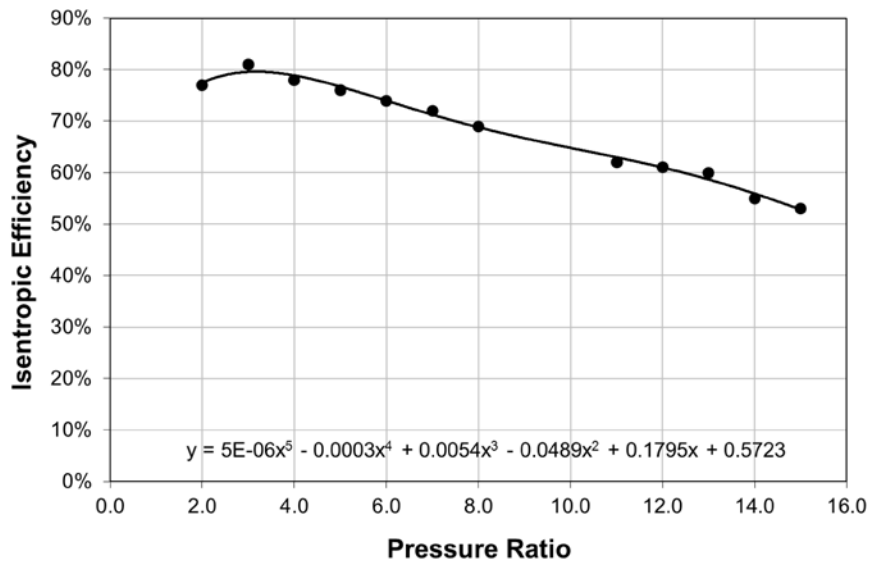


Figure 2.2 - Typical Isentropic Efficiency of a Single-Screw Compressor (ASHRAE, 2000)

A constant motor efficiency of 90% was assumed. Figure 2.3 gives the computed ε_r of the system as a function of saturated evaporator and condenser temperature. The COP ratio was expressed as,

$$\varepsilon_R = AA(T_c) \cdot T_e^2 + BB(T_c) \cdot T_e + CC(T_c) \quad (2.21)$$

where, the coefficients AA, BB and CC were obtained from the regression shown in Figure 2.4.

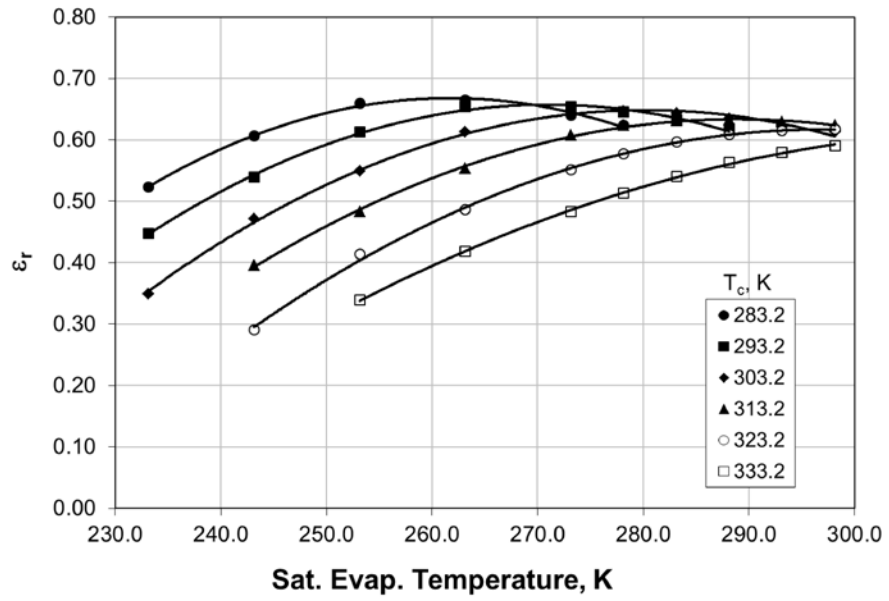


Figure 2.3 - Computed COP Ratio for a Vapor-Compression System Using R134a

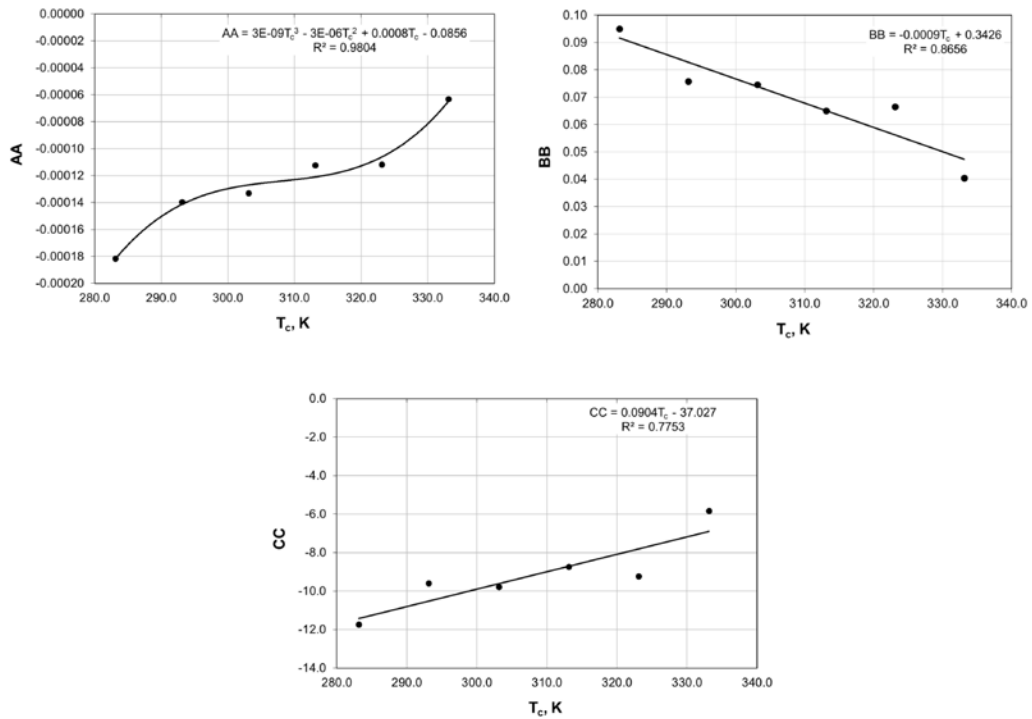


Figure 2.4 - Regression Coefficients for ϵ_r Map

2.1.4 Summary of Simple Model Assumptions

The simplified thermodynamic model presented herein is subject to a number of simplifying assumptions, such as,

- All of the servers/chassis are identical (power and airflow)
- The model does not differentiate between different data center configurations (e.g., number of racks in an aisle). However, the non-uniformity of temperature expected in an aisle is captured through the recirculation non-uniformity parameter, θ .
- The leakage parameter, λ , accounts for unintentional leakage through cracks and cable cut-outs, as well as cold air supplied to the cold aisle but not directly

ingested by the racks (i.e., it "spills" out of the aisle and short-circuits to a CRAC/CRAH).

- Leakage flow mixes with rack exhaust, resulting in an average mixed exhaust temperature \bar{T}_x .
- Recirculated air is at the average mixed exhaust temperature \bar{T}_x and therefore the server with the most recirculation, ϕ_{max} , has the highest inlet temperature.
- The efficiencies of the vapor compression cooling system and the CRAC/CRAH fan are independent of their respective pressure rise.

2.2 Thermo-hydraulic Model of a Data Center's Cooling Infrastructure

Many data center analyses focus on design point operation, which in reality is rarely realized due to the reliability needs and the constantly changing operating states of data centers. This section focuses on the development of a model that will allow data center operators and designers the ability to evaluate the energy use of a data center. The framework of this modeling environment is generic enough as to allow for easy integration and assessment of new, innovative technologies to the data center. The modeling methodology couples a thermodynamic model of the cooling and power equipment and a hydraulic pipe network. Inherent in this model is the ability to capture the off-design operating conditions of the data center's infrastructure caused by changes in ambient conditions and fluctuations in required IT load. In the ensuing sections, the

details of the modeling methodology of each component in the data center cooling loop are discussed followed by details of the software development. Further details of the model development can be found in Demetriou, Khalifa, Iyengar and Schmidt (2011a; 2011b).

2.2.1 Motor-driven Centrifugal Chiller

Braun (1987) showed that the dimensionless shaft power consumption, P_{sh} , of fixed speed vane-controlled or variable speed centrifugal chillers could be computed using the correlation given by Equation 2.22, where the coefficients ($a_0 - a_5$) can be determined with linear least squares fit applied to a chiller performance map consisting of evaporator load \dot{Q}_e , condenser exit temperature T_{cd}^{out} and evaporator exit temperature T_{ev}^{out} . (Note: all quantities are normalized by a design condition noted by a superscript d).

$$\frac{P_{sh}}{P^d} = a_0 + a_1\xi + a_2\xi^2 + a_3\gamma + a_4\gamma^2 + a_5\xi\gamma$$

$$\xi = \left(\frac{\dot{Q}_e}{\dot{Q}_e^d} \right) \tag{2.22}$$

$$\gamma = \left(\frac{T_c^{out} - T_e^{out}}{T_c^{out,d} - T_e^{out,d}} \right)$$

For a given evaporator exit temperature (commonly referred to as chilled water temperature), the evaporator load can be computed by a steady state energy balance.

$$\dot{Q}_e = \dot{m}_e c_p (T_e^{in} - T_e^{out}) \quad (2.23)$$

The load removed by the condenser is the sum of the load removed in the evaporator and the work input to the compressor, given by Equation 2.24, with the only unknown being the dimensionless quantity γ .

$$\dot{m}_{cd} c_p \left[\gamma (T_c^{out,d} - T_c^{in,d}) - (T_c^{in} - T_e^{out}) \right] = \dot{Q}_e + P_s^d (a_0 + a_1 \xi + a_2 \xi^2 + a_3 \gamma + a_4 \gamma^2 + a_5 \xi \gamma) \quad (2.24)$$

The actual electric power input of the chiller is computed by dividing the power computed by Equation 2.22 by the motor efficiency. The motor efficiency is computed based on a curve fit to data describing the motor efficiency as a percentage of part loads.

2.2.2 Effectiveness-NTU Model of a Wet Cooling Tower

The performance model of a counter flow or cross flow cooling tower and sump, where a hot water stream is in direct contact with an air stream and cooled as a result of sensible heat transfer due to temperature differences with the air and mass transfer resulting from evaporation to the air was developed by Braun et al. (1989). The water loss due to evaporation to the tower air stream is replaced with make-up water to the tower sump. Braun et al. (1989) has shown an effectiveness-NTU model for a cooling tower can be derived from the Merkel analysis (Merkel, 1925). The airside effectiveness,

$$\varepsilon_{air} = \frac{1 - \exp[-NTU(1 - C_r)]}{1 - C_r \exp[-NTU(1 - C_r)]}, \quad (2.25)$$

is defined as the ratio of the actual heat transfer to the maximum possible airside heat transfer that would occur if the exit air stream was saturated at the temperature of the incoming water. The heat capacity ratio C_r is defined as,

$$C_r = \frac{\dot{m}_{air} c_s}{\dot{m}_w c_w} = \frac{C_s}{C_w}, \quad (2.26)$$

where, C_w is the heat capacity rate of water and c_s is defined as the saturation specific heat. The saturation specific heat c_s is estimated by the average slope of the air saturation enthalpy versus temperature curve by,

$$c_s = \frac{h_s^{in} \Big|_{T_w^{in}} - h_s^{out} \Big|_{T_w^{out}}}{T_w^{in} - T_w^{out}}. \quad (2.27)$$

The actual heat transfer in terms of the airside effectiveness is computed by,

$$\dot{Q} = \varepsilon_{air} \dot{m}_{air} \left(h_s^{in} \Big|_{T_w^{in}} - h_{air}^{in} \right) \quad (2.28)$$

The exit enthalpy of the air and exit temperature of the water can be determined from an energy balance on each stream respectively.

$$h_{air}^{out} = h_{air}^{in} + \varepsilon_{air} \left(h_s^{in} \Big|_{T_w^{in}} - h_{air}^{in} \right) \quad (2.29)$$

$$T_w^{out} = T_{ref} + \frac{\dot{m}_w^{in} c_{p,w} (T_w^{in} - T_{ref}) - \dot{m}_{air} (h_a^{out} - h_a^{in})}{\dot{m}_w^{out} c_{p,w}} \quad (2.30)$$

The solution of Equations 2.25 – 2.30 can be obtained iteratively. The exit water flow rate of the cooling tower is computed by mass balance,

$$\dot{m}_w^{out} = \dot{m}_w^{in} - \dot{m}_{air} (\omega_{air}^{out} - \omega_{air}^{in}) \quad (2.31)$$

With the assumption of a Lewis number¹ equal to unity and an effective saturation humidity ratio for the entire cooling tower volume, Braun et al. (1989) showed that the air exit humidity ratio can be expressed as,

$$\omega_{air}^{out} = \omega_s^{eff} \Big|_{T_w^{in}} + \left(\omega_{air}^{in} - \omega_s^{eff} \Big|_{T_w^{in}} \right) e^{-NTU} \quad (2.32)$$

¹ The Lewis number is defined as, $Le = \alpha/D$, where α is the thermal diffusivity and D is the mass diffusivity. It can also be defined as the ratio of the Schmidt number to the Prandtl number.

The effective saturation humidity ratio can be obtained by psychrometric data using an effective saturation enthalpy for the cooling tower defined by Braun et al. (1989) as,

$$h_s^{eff} \Big|_{T_w^{in}} = h_{air}^{in} + \frac{h_{air}^{out} - h_{air}^{in}}{1 - e^{-NTU}} \quad (2.33)$$

Equation 2.34 provides a correlation to determine the number of transfer unites (NTU) based on performance data for a specific cooling tower. The NTU can be correlated to performance data of a specific cooling tower, which includes air flow rate, air dry bulb temperature, air wet bulb temperature, water flow rate, water inlet temperature and water outlet temperature at a number of operating conditions as,

$$NTU = \sigma \left(\frac{\dot{m}_w}{\dot{m}_{air}} \right)^{1+n}, \quad (2.34)$$

where, σ and n are empirical constants that can be obtained by fitting a straight line to a log-log plot of NTU versus flow rate ratio. Simpson and Sherwood (1946) provided typical values of σ (0.5 to 5.0) and n (-0.35 to -1.1).

The sump temperature is computed from Equation 2.35, which provides an energy balance of the cooling tower sump under the assumptions that the sump volume remains

constant, the volume is fully mixed and the flow rate of makeup water is equal to the rate at which water was evaporated.

$$V\rho \frac{dT_{sump}}{dt} = \dot{m}_w^{out} (T_w^{out} - T_{sump}) + \left[(\dot{m}_w^{in} - \dot{m}_w^{out}) (T_{makeup} - T_{sump}) \right] \quad (2.35)$$

Finally, the power consumption of the cooling tower fan is assumed to obey fan laws, where the total power consumption can be written as,

$$P_{F,CT} = \Omega^3 P_{F,max} \quad , \quad (2.36)$$

where, Ω is the relative fan speed and $P_{F,max}$ is the maximum fan power consumption.

2.2.3 Hydraulic Network

The hydraulic model allows for the calculation of the flow distribution in a pipe network based on the pressure distribution. The pipe network is represented by a collection of flow resistances, which could be representative of valves, elbows, tee pieces, length of pipe, etc. The major flow distribution change being captured in the hydraulic model is related to the variations in CRAH operation. Predicting the water flow rate to each CRAH will allow for the calculation of the heat removal in each device. Stand-alone, the hydraulic model does not incorporate a thermodynamic-based model for the heat transfer occurring in the various data center components.

Each CRAH unit has a controllable flow valve associated with it, which along with the heat exchanger can be represented as a flow resistance. In this model, it is assumed that the fractional open area, $\alpha = A/A^*$, of the valve is determined based on an open loop controller, where the current state is only a function of the return air temperature to the CRAH. The characteristics of this control are obtained from the manufacturer for a specific CRAH unit. The flow resistance of the CRAH unit is modeled as an orifice, where the pressure-flow relationship is represented by,

$$\dot{V} = C_D A \sqrt{\frac{2\Delta p}{\rho}}, \quad (2.37)$$

where, C_D , is the discharge coefficient and A is the flow area. However, the discharge coefficient is a function of valve opening ratio, α . To incorporate this characteristic, the valve discharge coefficient is corrected by a multiplier β , which is a function of α . The discharge coefficient can then be written as,

$$C_D(\alpha) = C_D^* \beta(\alpha), \quad (2.38)$$

where C_D^* is assumed constant and β is the correction factor. The orifice equation can then be re-written as,

$$\dot{V} = C_D^* \beta(\alpha) \alpha A^* \sqrt{\frac{2\Delta p}{\rho}}. \quad (2.39)$$

This modified equation allows for variations in both C_D and A . The resistance of the chiller's evaporator is modeled in a similar manner to the CRAH unit; however, in this instance, the discharge coefficient is variable and provided as a function of Reynolds number assuming a constant flow area. Lastly, the pressure loss in the pipes is modeled using the Darcy equation,

$$\Delta p = f \frac{(L + L_{eq})}{D_H} \frac{\rho}{2A^2} \dot{V}^2, \quad (2.40)$$

where the friction factor of the pipe, f , is computed using the Haaland approximation (Mathworks, 2008),

$$\frac{1}{\sqrt{f}} = -1.8 \log \left[\left(\frac{\zeta/D}{3.7} \right)^{1.11} + \frac{6.9}{Re_D} \right]. \quad (2.41)$$

The minor losses in the pipe section are considered by computing an equivalent pipe length, which is then added to the actual pipe length. The equivalent pipe length is related to the loss coefficient, k , which is typically provided in manufacturer's handbooks for valves, tee-pieces, elbows, etc., by,

$$L_{eq} = \frac{D_H \sum k}{f} \quad (2.42)$$

2.2.4 Chilled Water and Cooling Water Pumps

Modeling of both the chilled water and condenser water pumps is carried out using a model whose flow rate is based upon the intersection between a series of pump head curves at different pump speeds, and the system pressure drop, which is computed using the hydraulic model. The power consumption of the pump can be computed from Equation 2.43, where the overall efficiency, η_{ov} , is defined as the motor efficiency times the pumping efficiency,

$$P_p = \frac{\dot{m}\Delta p}{\rho\eta_{ov}} \quad (2.43)$$

2.2.5 Computer Room Air Handler

At this point, the chilled water flow rate to each CRAH unit is known from the hydraulic model and the remainder of the heat exchanger quantities can be determined from an NTU-effectiveness model of the CRAH heat exchanger. The CRAH heat exchanger is assumed to be cross flow where both fluids are unmixed. The effectiveness of the heat exchanger is given by Incropera and Dewitt (2002) as,

$$\varepsilon = 1 - \exp\left[\left(\frac{1}{C_r}\right)NTU^{0.22}\left\{\exp\left[1 - C_rNTU^{0.78}\right] - 1\right\}\right], \quad (2.44)$$

where, C_r is the ratio of minimum to maximum heat capacity rates of the air and water streams and NTU is defined as,

$$NTU = \frac{UA}{C_{min}}, \quad (2.45)$$

where, UA is the overall conductance for the CRAH heat exchanger. Typical values based on manufacturer data range from 10,000 – 25,000 W/K. Finally, the heat transfer rate, exit temperatures of the water and exit temperature of the air streams can be computed by energy balances of each stream.

A number of other control strategies are commercially available for CRAH units, including water bypass for controlling the exit air temperature and closed loop feedback-based controllers. All of these control methods utilize the equations presented above and have been implemented into the simulation engine. Each CRAH unit also has one or more centrifugal blowers, which move a substantial volume of air against a relatively large pressure drop; therefore, they consume a significant amount of power. The characteristic of the blowers are modeled in a similar manner to the pumps, which were described above. It should be noted that work input to the blowers also shows up as heat dissipation, which results in a temperature rise of the air before it exits the CRAH.

2.2.6 Uninterruptible Power Supply (UPS)

The uninterruptible power supply (UPS) is modeled using a performance map look up table of four different UPS topologies: average double conversion, high efficiency double conversion, high efficiency delta conversion and high efficiency flywheel. The look-up table data was obtained from a Lawrence Berkeley National Labs study (2005), where the efficiency is given as a function of the normalized load. Figure 2.5 shows the data for each of the topologies.

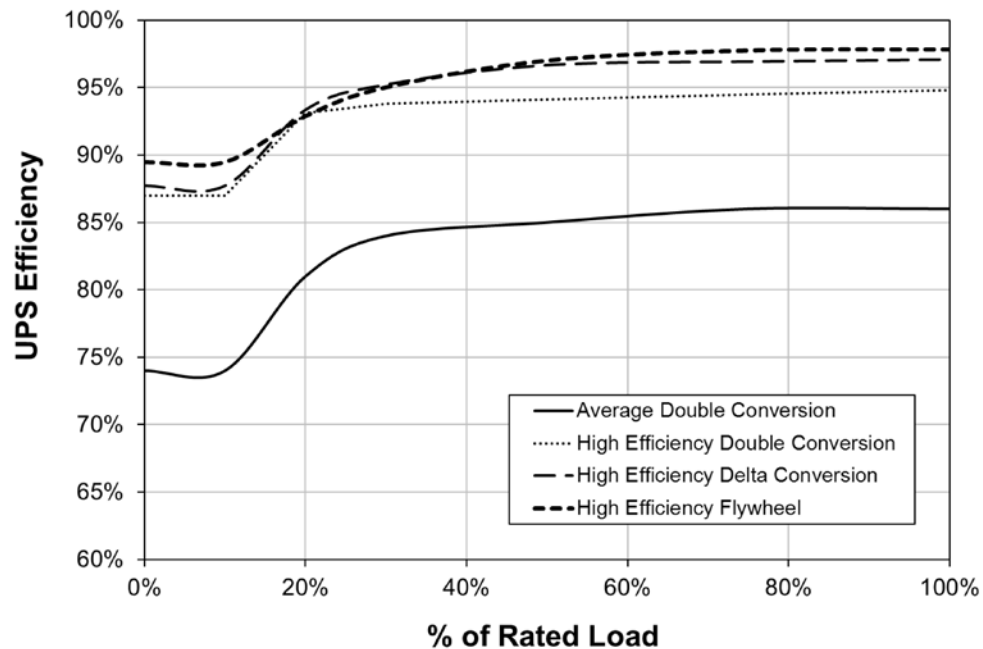


Figure 2.5 - Part Load Efficiency for Various UPS Topologies (adapted from LBNL, 2005)

2.2.7 Creating a Coupled Simulation Environment

The coupled thermo-hydraulic model uses TRNSYS (Klein, 2002) to perform the thermodynamic calculations and Simulink to perform the hydraulic network simulations.

TRNSYS (TRaNsient System Simulation Program) is a commercially available code, which originally was developed at the University of Wisconsin-Madison for use in solar thermal applications. However, due to its modular approach to modeling and easy to use graphical interface has since been adopted for energy simulations for many types of systems. TRNSYS represents a component, such as a pump, cooling tower, etc., as a subroutine that lays out the mathematical description of the component. Any model can be easily integrated to a simulation through a modular dynamic link library (DLL). The TRNSYS kernel is then used as the simulation engine for providing communication between components and iteratively solving all components as a coupled system.

Simulink (Mathworks, 2008) is a simulation environment for modeling dynamic and embedded systems, under a variety of time-dependent forcing functions. Similar to TRNSYS, Simulink represents the system as a collection of coupled mathematical models. Simhydraulics is a library of blocks in Simulink that allows modeling of hydraulic and hydro mechanical systems. The library includes more than 45 hydraulic and mechanical components, including pumps, valves, accumulators and pipelines. New components can easily be added to Simulink through Matlab code or application programming interfaces (API).

The simulation environment developed here uses TRNSYS as the solution engine for coupling Matlab/Simulink to TRNSYS. The Matlab coupling to TRNSYS is done through a component model interface, which opens an instance of Matlab as a separate process. The Matlab component requires an *m-file*, which is run as a script in the Matlab

workspace. The structure of the *m-file* script is almost identical to the Fortran-based code required of a typical TRNSYS component. The coupling of Matlab to Simulink is done through a call to the Matlab function *sim*, which runs a dynamic simulation in Simulink. The *sim* function requires as input a Simulink model file as well as any data require of that specific model. A flow chart of the software communication is given in Figure 2.6.

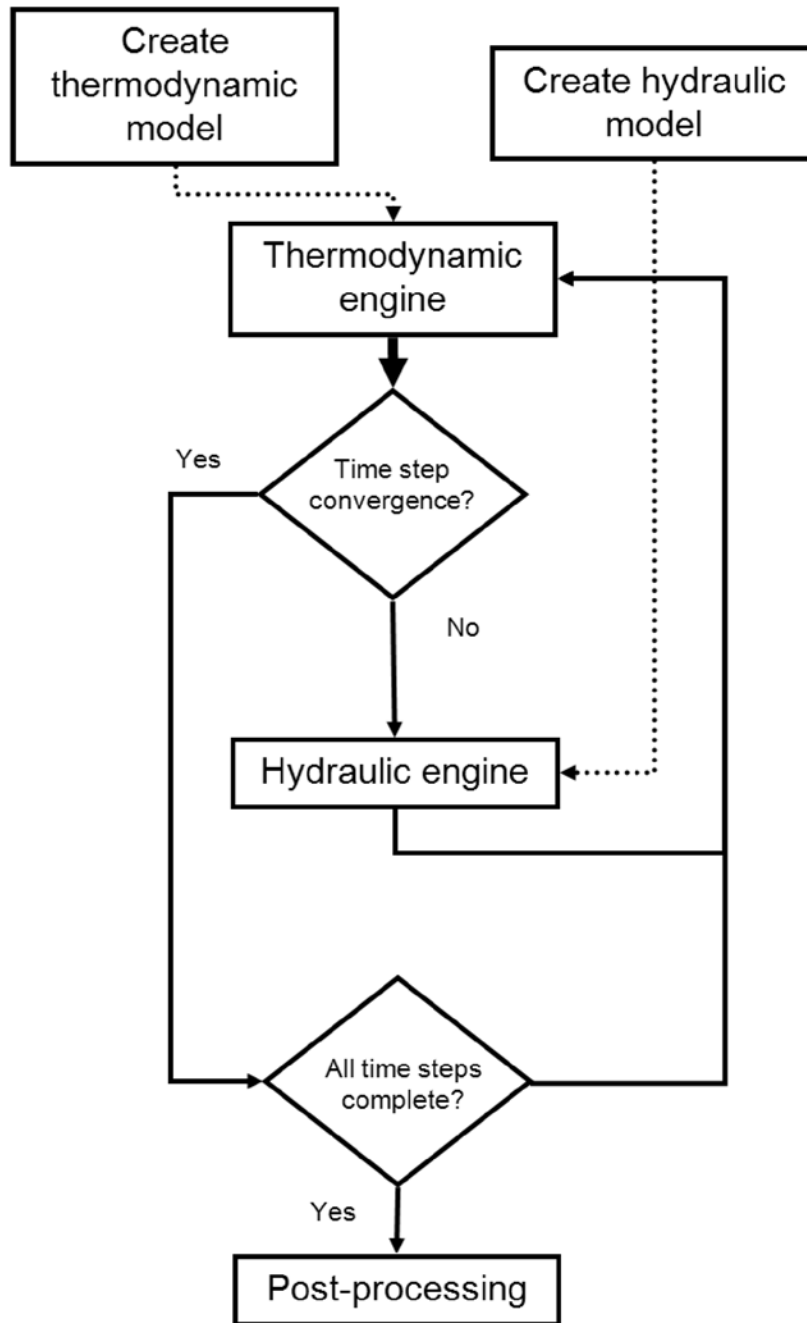


Figure 2.6 - Flow Chart of Thermo-Hydraulic Model

2.3 *Computational Fluid Dynamics for Investigating the Data Center's Airflow and Temperature distribution*

A consequence of an air-cooled, raised floor data center is the emergence of complex airflow and temperature fields within the space. Predicting the thermal environment is paramount to designing effective strategies for maintaining the reliability of the IT equipment. Furthermore, intrusive measurements are typically prohibited in data centers due to reliability considerations; therefore, computational fluid dynamics becomes the main tool for investigating the airflow and temperature distribution of a data center. This section will describe the fundamental equations for conservation of mass, momentum and energy as well as necessary turbulence models. For this work, the commercial CFD package Fluent is used.

2.3.1 Conservation of Mass

Mass conservation leads to the continuity equation for a fluid, given as,

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_i}(\rho u_i) = 0 \quad (2.46)$$

where, ρ is the fluid density and u is the velocity (note: Einstein notation is assumed). In data centers, incompressibility can be assumed since the Mach number is much less than 0.3. The assumption of incompressibility leads to the following,

$$\frac{\partial}{\partial x_i}(u_i) = 0 \quad (2.47)$$

2.3.2 Conservation of Momentum

With the assumption of incompressibility, conservation of momentum is given in each of the three coordinate directions as,

$$\frac{\partial}{\partial t}(\rho u_i) + \frac{\partial}{\partial x_j}(\rho u_i u_j) = -\frac{\partial p}{\partial x_i} + \frac{\partial}{\partial x_i} \left[\mu \frac{\partial u_i}{\partial x_j} \right] + g_i, \quad (2.48)$$

where, μ is the fluid viscosity, p is the hydrostatic pressure and g_i is the body force in direction i .

In data centers with low flow and large heat dissipation (i.e., high server temperature rise), flows driven by natural convection, due to density difference caused by temperature differences, may be important. To determine if buoyancy driven flows are important, the Archimedes number (Ar) is considered. Archimedes number is defined as the ratio of the Grashoff number (Gr) to the square of the Reynolds number (Re).

$$Ar = \frac{Gr}{Re^2} = \frac{g\beta\Delta TL}{u^2} \quad (2.49)$$

β is the volumetric thermal expansion coefficient ($\beta = 1/T$ for a perfect gas), ΔT , L and u are characteristic temperature difference, length scale and velocity, in the flow, respectively. Archimedes number represents the ratio of buoyancy forces to inertial forces. Therefore, an Archimedes number close to or greater than unity shows that thermal buoyancy forces may be just as important as inertial forces. In future sections, an Ar in terms of data center characteristics will be developed and several studies will be presented that show the importance of considering buoyancy in high-density data center applications.

Density differences in the flow field are considered using the so-called Incompressible Ideal Gas Law, based on recommendations by Dygert et al. (2010), which showed many advantages of the Incompressible Ideal Gas model over the commonly used Boussinesq approximation. The density of the fluid in the flow field is computed by,

$$\rho = \frac{p_{op}}{\frac{R_u}{M} T}, \quad (2.50)$$

where, p_{op} is an operating pressure, R_u is the universal gas constant and M is the molecular weight of the gas. In this formulation, density is a function of local temperature only.

Many flow features of the data centers (i.e., perforations on tiles and racks) make them impractical for explicitly modeling due to the difference in length scale compared to the overall room. For this reason, Abdelmaksoud et al. (2010) recommended the use of a momentum source for modeling perforated surfaces in the data center. The momentum source method corrects for the momentum deficit in the jet emanating from the entire surface by adding a body-force field in the computational volume immediately adjacent to the perforated surface. Consider the example of a perforated tile in the data center. In CFD, a perforated tile is typically modeled as 100% open. Mass conservation will be assured by setting the normal velocity through the tile as (\dot{V} / A_{tile}) , where A_{tile} is the physical face area of the tile. Clearly, the correct jet momentum is obtained only if the perforation of the tile is 100%. If the tile is not fully open, the correct velocity through a pore is $\sim(\dot{V} / \sigma A_{tile})$, corresponding to the total tile momentum flux of $\rho\dot{V}(\dot{V} / \sigma A_{tile})$. To correct for this momentum deficit, a body-force field, given as,

$$F_i = \frac{I}{V} \rho \dot{V} \left(\frac{\dot{V}}{\sigma A_{tile}} - \frac{\dot{V}}{A_{tile}} \right), \quad (2.51)$$

is added to the momentum equation. In this formulation, F_i represents the body force per unit volume in direction i , V represents the volume of the finite region adjacent to the perforated surface where the body-force is applied and σ is the fraction of perforated area.

The perforated tile airflow distribution (from the plenum to the raised floor) could have a significant impact on the amount of cooling provided to the racks. The main feature for determining the uniformity of the airflow distribution to the perforated tiles is the tile resistance. Computationally, the pressure loss in the tiles is modeled using the *porous media* model, in which a finite cell volume is selected for which the porous media is applied. The pressure loss through the porous media is given by,

$$\Delta p_i = - \left(\sum_{j=1}^3 D_{ij} \mu u_j + \sum_{j=1}^3 C_{ij} \frac{1}{2} \rho |u| u_j \right), \quad (2.52)$$

where, D_{ij} and C_{ij} are coefficient matrices specifying the model constants in each of the three coordinate directions. The first term on the RHS is the Darcy's Law term where the pressure loss is proportional the velocity, this term is most appropriate for laminar flows and is typically ignored in data center applications. The second term on the RHS is the inertial loss term in the porous media. In Fluent, heat transfer in the porous media is handled by specifying an effective thermal conductivity, k_{eff} , given by,

$$k_{eff} = \gamma k_f + (1 - \gamma) k_s. \quad (2.53)$$

k_{eff} is a volume average of the fluid k_f and solid k_s thermal conductivities. Turbulence is handled in a porous region by solving the transport equations for turbulence, described in Section 2.3.4; however, it is assumed that the solid region has no effect on turbulence

generation or dissipation. An alternative approach is to specify the region as laminar, in which case the turbulence quantities at the inlet of the porous zone are simply advected to the exit of the porous zone.

2.3.3 Conservation of Energy

With the assumption of incompressibility, the energy equation can be written in terms of static enthalpy, h , as,

$$\frac{\partial}{\partial t}(\rho h) + \frac{\partial}{\partial x_i}(\rho u_i h) = \frac{\partial}{\partial x_i} \left[\frac{\mu}{Pr} \frac{\partial h}{\partial x_i} \right], \quad (2.54)$$

where for an ideal gas the static enthalpy is defined as $h = c_p (T - T_{ref})$ with $T_{ref} = 298.15$

K and Pr is the fluid Prandtl number $Pr = \frac{c_p \mu}{k}$. When using the pressure-based solver in

Fluent, viscous dissipation, kinetic energy and pressure work are ignored due to the assumptions of incompressibility and low speed flow.

2.3.4 Turbulence Modeling

Turbulent flow is a regime of fluid flow that is characterized by chaotic and stochastic property changes. The diffusivity of turbulence causes rapid mixing and increased rates of momentum and heat transfer. By nature, turbulence is rotational and three-dimensional and is characterized by high levels of fluctuating vorticity. To study turbulence, a

Reynolds decomposition of the Navier-Stokes equations is performed, in which the instantaneous quantity is decomposed into its time-average and fluctuating quantities. This results in the well-known Reynolds Averaged Navier-Stokes equations (RANS). A consequence of this decomposition is the emergence of the non-linear Reynolds stress term $\left(-\overline{\rho u_i' u_j'}\right)$ in the governing equations, where the prime designates a fluctuating velocity component. This term leads to the closure problem of turbulence and further modeling is required.

Over the years, a wide range of turbulence models has been developed in order to handle different aspects of turbulent flow. In the area of data centers, many researcher have studied the performance of different turbulence models for predicting the airflow and temperature fields in a data center; however, limited experimental data has made many of these studies inconclusive. For this work, the Realizable $k-\varepsilon$ turbulence model with standard wall functions was adopted, based on studies of other indoor airflows (Shih et al., 1995; Russo, 2010). This model was found superior for predicting flows that include planar and round jets. The transport equation for k is given as,

$$\frac{\partial}{\partial x_i}(\rho k u_i) = \frac{\partial}{\partial x_i} \left[\left(\mu + \frac{\mu_t}{Pr_k} \right) \frac{\partial k}{\partial x_j} \right] - \overline{\rho u_i' u_j'} \frac{\partial u_i}{\partial x_j} + \beta g_i \frac{\mu_t}{Pr_t} \frac{\partial T}{\partial x_i} - \rho \varepsilon \quad (2.55)$$

where, k is the turbulent kinetic energy, ε is the turbulent dissipation, μ_t is the turbulent viscosity, and Pr_t and Pr_k are the Prandtl numbers for turbulence and kinetic energy, respectively.

The Realizable k - ε model provides improvements over the other models in the k - ε family of turbulence models by improving the model constants needed in the dissipation equation, which is given by (Shih et al., 1995),

$$\begin{aligned} \frac{\partial}{\partial x_i}(\rho \varepsilon u_i) = & \frac{\partial}{\partial x_i} \left[\left(\mu + \frac{\mu_t}{Pr_\varepsilon} \right) \frac{\partial \varepsilon}{\partial x_j} \right] - \rho C_1 S_\varepsilon + \rho C_2 \frac{\varepsilon^2}{k + \sqrt{\nu \varepsilon}} \\ & + C_{1\varepsilon} \frac{\varepsilon}{k} C_{3\varepsilon} \beta g_i \frac{\mu_t}{Pr_t} \frac{\partial T}{\partial x_i} \end{aligned} \quad (2.56)$$

where, $C_1 = \max \left[0.43, \frac{l}{l+5} \right]$, $l = S \frac{k}{\varepsilon}$, $S = \sqrt{2 S_{ij} S_{ij}}$, C_2 , $C_{1\varepsilon}$ and $C_{3\varepsilon}$ are model constants, l

is the turbulence length scale and S_{ij} is the mean strain rate tensor.

3 Investigation into the Optimization of a Data Center's Infrastructure using the Simple Model

This chapter uses the simplified thermodynamic model to explore optimization possibilities in air-cooled data centers. The results of this analysis will highlight the important features that need to be considered when optimizing the operation of air-cooled data centers. The use of the simple model offers the ability to limit the range of options and parameters to be explored in what would otherwise be a resource-intensive optimization analysis.

3.1 Optimization of Enclosed Aisle Data Centers

A clear advantage of an enclosed aisle data center is that no recirculated air from the exhaust of the racks can be entrained into the supply air stream in the cold aisle and therefore, thermally uniform conditions are obtained at the inlet to the racks. The simple model was used to investigate the optimization of an enclosed aisle, air-cooled data center with 1024 kW of IT load (i.e., a data center in which $\theta = 0.0$). Typical values were assumed for model parameters and system characteristics, as given in Table 3.1.

Table 3.1 - Parameters Used in Simple Analysis

Operating IT Power [kW]	1024.0
Average Rack Temperature Rise [°C]	10.0
Leakage as a % of Tile Flow, λ [%]	20.0
CRAC Fan Pressure Coefficient, K_C [Pa/(m ⁶ /s ²)]	11.1
Bypass Fan Pressure Coefficient, K_C [Pa/(m ⁶ /s ²)]	1.11
Server Redline Temperature [°C]	27.0
Outdoor Ambient Temperature [°C]	30.0
CRAC Fan Overall Efficiency, η_F	0.65
Compressor Motor Efficiency, η_m	0.90
Default ε_R (relative to Carnot)	0.55
Evaporator Conductance, UA_e [kW/K/CRAC]	14.0
Condenser Thermal Resistance, χ [K/kW]	0.08
Number of CRACs, N_C	10

3.1.1 Constant Server Flow

Figure 3.1 presents the results for a case with constant server flow (i.e., a specified server temperature rise of 10°C). The figure shows the cooling infrastructure power consumption and the conjugate CRAC exit temperature for a range of CRAC air supply fraction, (i.e., the fraction of air that passes through the CRAC and is cooled).

There exists an optimal value $\bar{\psi} = 0.548$ and its conjugate cold air supply temperature $T_a = 20.13^\circ C$ that minimize the overall power consumption, subject to the redline

constraint. It is evident that a higher CRAC exit temperature does not always lead to lower power consumption in air-cooled data centers, where CRAC fan power consumption is a significant contributor to the overall cooling infrastructure power. Ignoring the CRAC fan power may be justified in situations in which the refrigeration power consumption is dominant, as it would be when inefficient refrigeration systems are employed in very warm climates. In most other situations, exclusive focus on reducing refrigeration power consumption without regard to the power consumed in moving the cooling air would lead to sub-optimum, possibly misleading results.

Enclosing the cold aisle is one practical arrangement for achieving a thermally uniform inlet temperature. With an enclosed cold aisle, no recirculated air can be entrained into the supply air stream, and the perforated tiles within the enclosure must provide as much air as the servers need. This air will enter the servers at a uniform temperature as high as the redline temperature T^* . This is the point corresponding to the extreme right ($\bar{\psi} = 1.0$) of Figure 3.1. Yet, paradoxically, the power consumption at this point is much higher than that at the minimum of the power consumption curve, which corresponds to a significant degree of recirculation, ($\bar{\psi} = 0.548$). This apparent paradox stems from the fact that if all the flow issuing from the enclosed perforated tiles emanates from the CRACs, considerable fan power must be consumed to overcome the relatively large pressure drop across the CRAC heat exchangers and filters. This is both wasteful and unnecessary in view of the relatively high temperature at which air enters the servers. Based on this example, the implementation of a CRAC bypass branch in conjunction with

aisle containment shows the potential of reducing the power consumption by 43% (173kW vs. 306kW) compared to a conventional enclosed aisle.

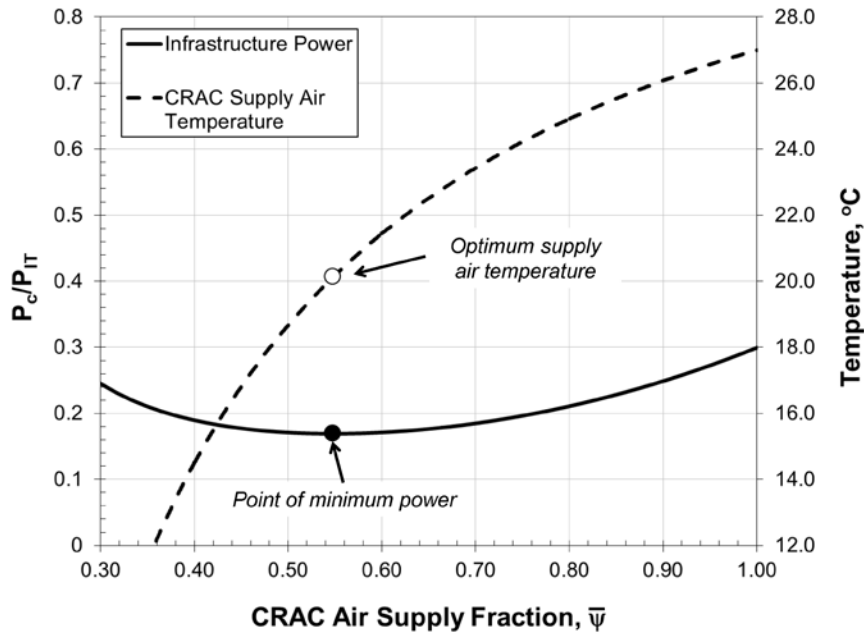


Figure 3.1 - Simple Model Optimization Results: Enclosed Aisle

In order to realize the potential energy savings of an enclosed aisle with bypass recirculation, a method for controlling the amount of recirculated air, while providing thermally uniform inlet conditions to the racks, is needed. One potential solution resides in the provision of a CRAC bypass recirculation branch like that shown schematically in Figure 2.1. Because this recirculated air must be thoroughly mixed with the CRAC cold air supply to ensure uniformity, it may be necessary to introduce this recirculated air into the under floor plenum using a low-lift fan to overcome the modest Δp in the plenum (~10% of that in the CRAC/CRAH for the same flow rate), for example. The power

consumption of a bypass fan has been accounted for in the results presented in Figure 3.1. Figure 3.2 provides a schematic of this preferred configuration (Khalifa and Demetriou, 2011). It may also be possible to introduce recirculated air directly into the cold aisle by means of induction louvers and low-lift fans, along with mixing fans to “homogenize” the air inside the enclosure. The bypass branch and its controls may also be integrated into the CRACs themselves.

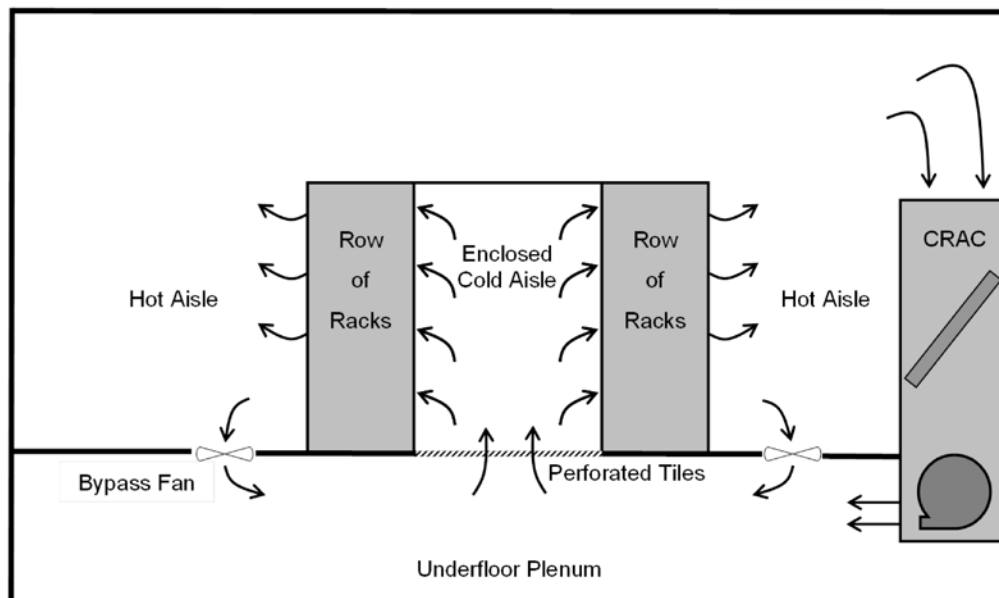


Figure 3.2 - Preferred Configuration for Enclosed Aisle Data Centers with Bypass Recirculation

Enclosing the hot aisle would produce similar benefits, but may also provide some practical advantages. For example, recirculation bypass can be provided directly by installing a number of actively dampered tiles in the floor of the hot aisle. These tiles may be equipped with low-lift axial fans to overcome the additional few mm-of-water pressure difference in the under floor plenum. In such a case, the hot aisle must be ducted

to the inlets of the CRACs. For now, it is surmised that the bypass air will thoroughly mix with the cold air from the CRAC inside the plenum before issuing from the cold aisle perforated tiles to enter the racks (this is verified in Section 4.2). Figures 3.3 illustrate the breakdown of power consumption for the refrigeration system and the fans (CRAC + bypass). For the special case of a conventional enclosed aisle without CRAC bypass ($\bar{\psi} = 1.0$), the refrigeration power is relatively low (~13% of IT power) owing to the relatively high supply air temperature ($T_a = 27.0^\circ C$), which improves the refrigeration system's *COP* considerably. In this case, the entire flow required by the racks plus the leakage passes through the CRAC, resulting in a high CRAC fan power use (~17% of IT power). It should be noted, however, that many data centers are typically supplied with air at much lower temperatures, in the 12°C - 18°C range, and consequently have higher refrigeration power consumption.

3.1.2 Effect of Server Temperature Rise

The results depicted in Figure 3.1 were for servers with a 10°C temperature rise. The effect of the server temperature rise on optimized cooling power consumption and its conjugate CRAC air supply fraction are shown in Figure 3.4 (Note: Figure 3.4 shows the minimum power consumption point for each value of server temperature rise). As the server temperature rise increases (server flow decreases for the same server power), the cooling power consumption decreases. However, the advantage of bypass recirculation diminishes as the server temperature rise increases and disappears entirely at a server temperature rise around 20°C, in this example case. However, enclosing the aisles

together with bypass recirculation (when beneficial) also reduces the dependence of the optimum cooling power on server temperature rise as evidenced by the relatively small savings in power consumption (<10%) as the server temperature rise is increased from 8°C to 26°C in Figure 3.4.

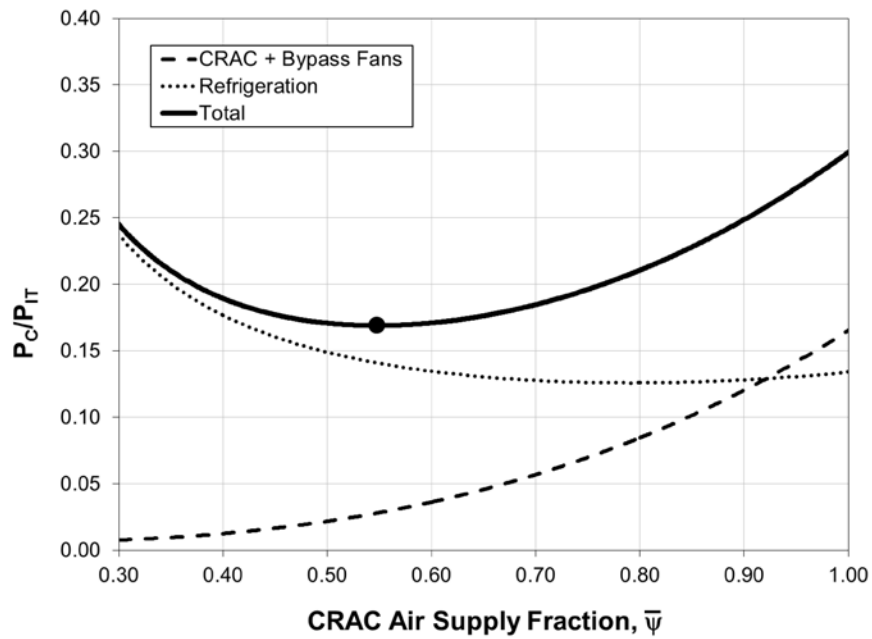


Figure 3.3 – Component-by-Component Energy Breakdown for $\theta = 0.0$

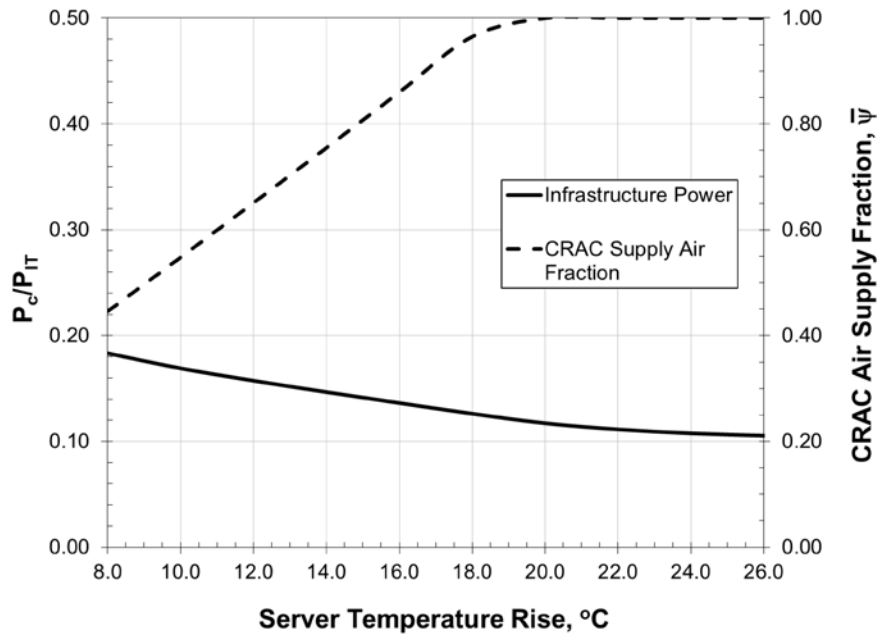


Figure 3.4 - Enclosed Aisle: Effect of Server Temperature Rise

3.2 Optimization of Open Aisle Data Centers

The development of the simple model explicitly incorporated the effect of recirculation non-uniformity through the parameter θ . This measure characterized the inlet temperature non-uniformity over an aisle of server racks. The simple model was used to investigate the optimization of an open aisle, air-cooled data center with 1024kW of IT load, for a range of non-uniformity, $0.0 \leq \theta \leq 7.0$. This analysis will highlight the important features that need to be considered when optimizing the operation of air-cooled data centers and elucidate the deleterious effect of temperature non-uniformity at the inlet of the racks on the data center's cooling infrastructure power consumption. Typical values were assumed for model parameters and system characteristics as given in Table 3.1.

3.2.1 Constant Server Flow

Figure 3.5 presents sample results for a range of recirculation flow non-uniformity parameter θ (Equation 2.1) from uniform ($\theta = 0.0$) to highly non-uniform ($\theta = 7.0$). For each value of θ , Figure 3.5 shows the variation of the normalized cooling power consumption with the average cold air supply fraction $\bar{\psi}$.

It can be seen that there are optimum values of $\bar{\psi}$ and its conjugate cold air supply temperature T_a^* that minimize the power consumption, subject to the redline constraint. Figure 3.5 also displays the loci of the minimum cooling power consumption (black circles) for various θ and the corresponding optimum cold air supply temperature, required to satisfy the redline constraint (black squares). The minima of the power consumption curves shift upward and toward a higher $\bar{\psi}$ as the degree of recirculation non-uniformity increases. Therefore, rack, perforated tile, CRAC placement, and flow control must aim to reduce or eliminate server-to-server recirculation non-uniformity as much as practicable. It should be recalled that θ , while a direct measure of recirculation fraction non-uniformity, is only an indirect measure of rack inlet temperature non-uniformity. Furthermore, as $\bar{\psi} \rightarrow 1.0$ very little of the rack exhaust will be recirculated to the rack inlet. Therefore, even if this small amount of recirculation is non-uniform, the effect on rack inlet temperature will also be very small, leading to modest rack inlet temperature non-uniformity, even with high values of θ , as implied by Equation 2.9b.

It is also evident that a higher CRAC exit temperature does not always lead to lower power consumption in air-cooled data centers, where CRAC fan power consumption is a significant contributor to the overall cooling infrastructure power. Again, exclusive focus on reducing refrigeration power consumption without regard to the power consumed in moving the cooling air would lead to sub-optimum, possibly misleading results.

From Figure 3.5, it can be seen that the lowest power consumption is achieved when the temperature at the inlet of each server is uniform (i.e., $\theta = 0$), which corresponds to the case of an enclosing the cold aisle. This case was discussed in detail in Section 3.1.

Comparing the case of a typical enclosed aisle ($\theta = 0.0$, $\bar{\psi} = 1.0$) with a case of an open-aisle with some degree of non-uniformity ($\theta = 4.0$), the result show that even with some non-uniformity, a potential of 10% energy reduction (275 kW vs. 306 kW) is possible compared to the conventional enclosed aisle, where all the flow is forced through the CRAC (i.e., $\bar{\psi} = 1.0$).

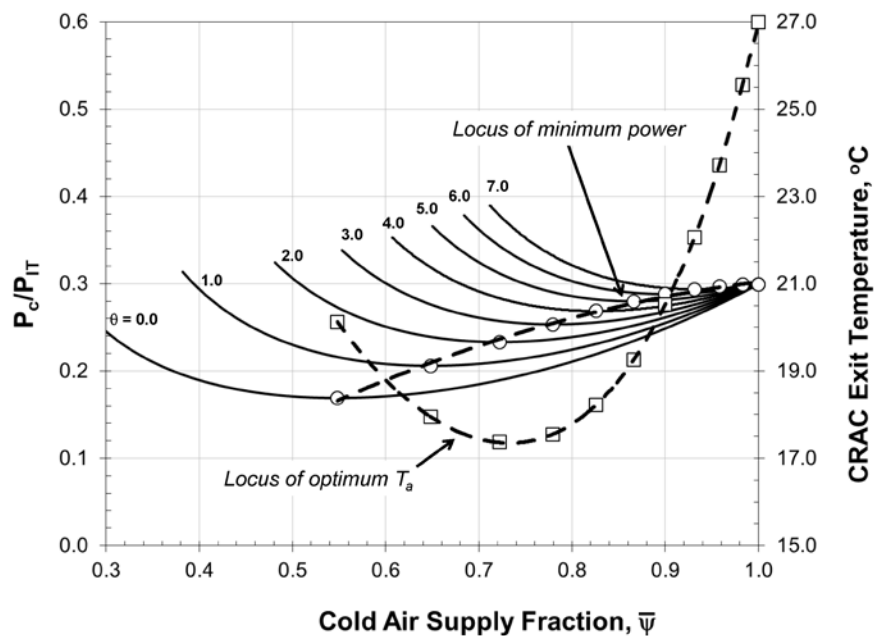


Figure 3.5 - Simple Model Optimization Results: Open Aisle

However, the benefits of using bypass recirculation in conjunction with an enclosed aisle were previously shown. Using the CRAC bypass branch ($\theta = 0.0$, $\bar{\psi} = 0.484$) shows a possible reduction in energy consumption by 43% (173 kW vs. 306 kW) compared to the typical enclosed aisle configuration ($\theta = 0.0$, $\bar{\psi} = 1.0$). Alternatively, being able to provide thermally-uniform conditions with the addition of the CRAC bypass shows savings of 37% (173 kW vs. 275 kW) compared to the case of an optimized open-aisle data center with a typical degree of recirculation non-uniformity ($\theta = 4.0$).

Figures 3.3 and 3.6 illustrate the breakdown of power consumption for two cases: a data center with an enclosed aisle ($\theta = 0.0$) and an open aisle data center with some degree of

recirculation non-uniformity ($\theta = 4.0$). The results show that as θ increases, the minimum power point shifts toward a larger cold air supply fraction, $\bar{\psi}$, and is associated with a higher power consumption. For the case of a conventional open aisle data center ($\theta = 4.0$), depicted in Figure 3.6, the refrigeration power at the minimum power point is ~17% of the IT power, whereas the CRAH fan power is ~9%, which fall within the ranges reported in by Salim and Tozer (2010).

Figure 3.7 shows the COP of the vapor-compression refrigeration system for four different supply air temperatures (6 °C, 12 °C, 18 °C and 24 °C), for a non-uniformity of $\theta = 0.0$ and $\theta = 4.0$. The results show the degradation of the COP as the supply air temperature is lowered and as the degree of non-uniformity increases. We note that higher non-uniformity necessitates a higher value of $\bar{\psi}$ (i.e., a higher supply air flow rate) and hence higher CRAC fan power at a given supply air temperature. This, in turn, requires that the air exits the CRAC heat exchanger at a lower temperature (i.e., for a larger θ we have a lower T_{ax}), and consequently a lower evaporator saturation temperature, and a lower refrigeration COP.

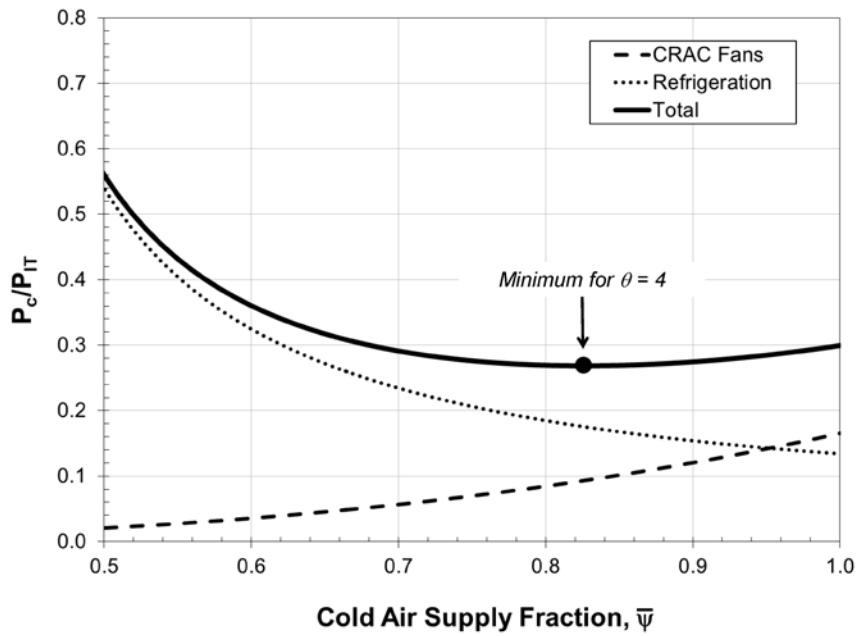


Figure 3.6 – Component-by-Component Energy Breakdown for $\theta = 4.0$

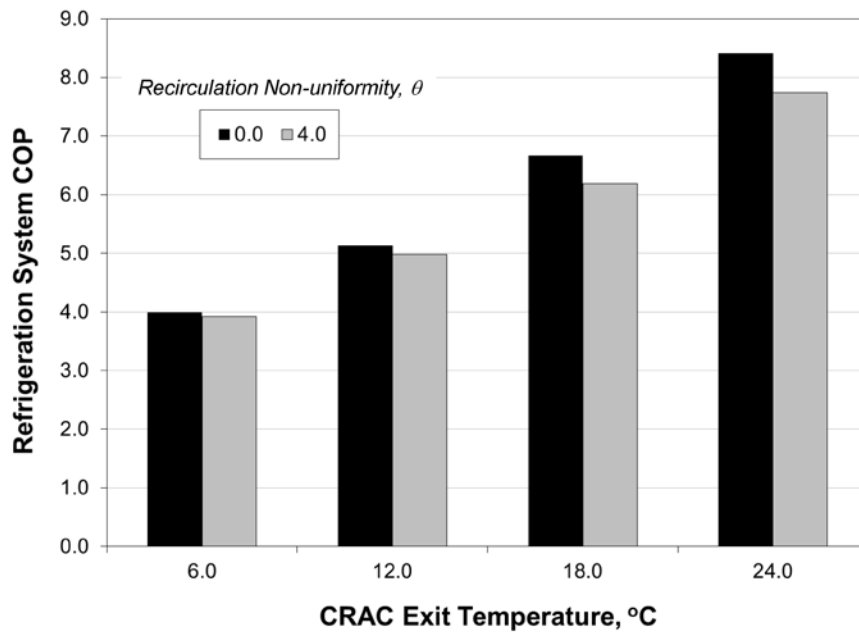


Figure 3.7 - Refrigeration System COP for $\theta = 0.0$ and 4.0

3.2.2 Effect of Server Temperature Rise

While the results depicted in Figure 3.5 are for servers with a 10°C temperature rise, the effect of the server temperature rise on optimized cooling power consumption is shown as a function of the rack recirculation non-uniformity parameter, θ , in Figure 3.8. As the server temperature rise increases (server flow decreases for the same server power), the cooling power consumption decreases and exhibits a flatter dependence on θ , especially at higher θ . At server temperature rise above $\sim 10^\circ\text{C}$, the curves are terminated when the minimum power consumption occurs at $\bar{\psi} = 1.0$, with $T_a = T^*$. Any value of θ beyond the terminal point has no practical relevance. This is because non-uniformity can arise only from recirculation (i.e., $\bar{\varphi} > 0.0$) and a value of $\bar{\varphi} = 0.0$ implies the absence of a mechanism for the rack inlet temperature to be non-uniform.

The optimum cold air supply fraction and the optimum conjugate CRAC exit temperature vary with the server temperature rise as shown in Figures 3.9 and 3.10. As can be seen in Figure 3.9, the optimum cold air supply fraction increases as the server temperature rise increases, terminating at unity as $\theta \rightarrow \theta^*$. For thermally uniform ($\theta = 0$) the benefits of bypass recirculation diminish markedly for server temperature rise of $\sim 16^\circ\text{C}$ or higher.

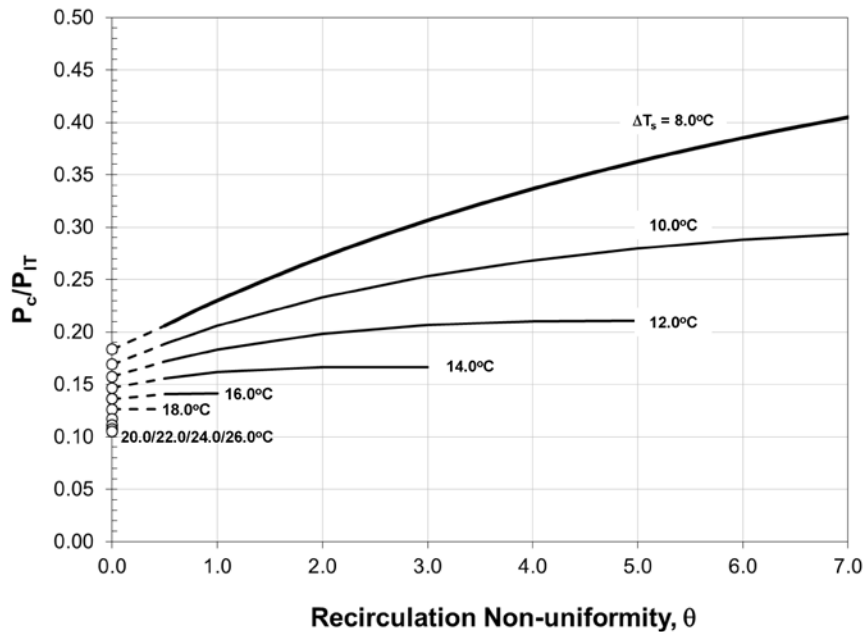


Figure 3.8 - Effect of Server Temperature Rise on Optimum Cooling Power for Open Aisle Data Centers

The effect of server temperature rise (or server flow rate) on the conjugate optimum CRAC air supply temperature is shown in Figure 3.10, which also highlights the restricted range of possible CRAC air supply temperatures at higher server temperature rise. Any additional flow beyond the maximum total server flow is equivalent to increased leakage flow, which would increase the cooling power consumption. Obviously, it is possible to lower the supply air temperature below the values indicated in Figure 3.10, resulting in a maximum server inlet temperature below the redline. However, this will also lead to an increase in the refrigeration system's power consumption and, consequently, the total cooling infrastructure's power consumption, unless the supply temperature reduction is also accompanied by a commensurate reduction in server flow.

Nevertheless, many air-cooled data centers with open aisles use servers with modest temperature rise ($\sim 10^{\circ}\text{C}$), and are operated with 10-20% recirculation, which could lead to a recirculation non-uniformity parameter, θ , in excess of 6, along with excessive cooling power consumption. Significant reductions in cooling power consumption can only be achieved through reduction in recirculation, accompanied by controlled bypass recirculation, redistribution of perforated tiles and their flow rates, or controlling rack flow to achieve similar outcomes. Enclosing the cold or hot aisles is an especially effective energy-saving solution for servers with high temperature rise, although bypass recirculation would be of little or no value for such servers.

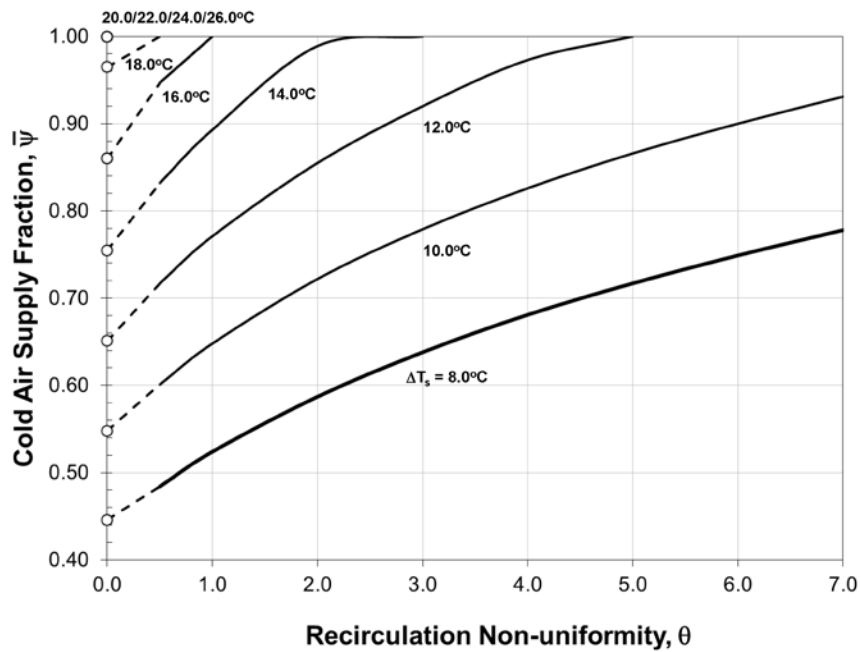


Figure 3.9 - Effect of Server Temperature Rise on Optimum $\bar{\psi}$

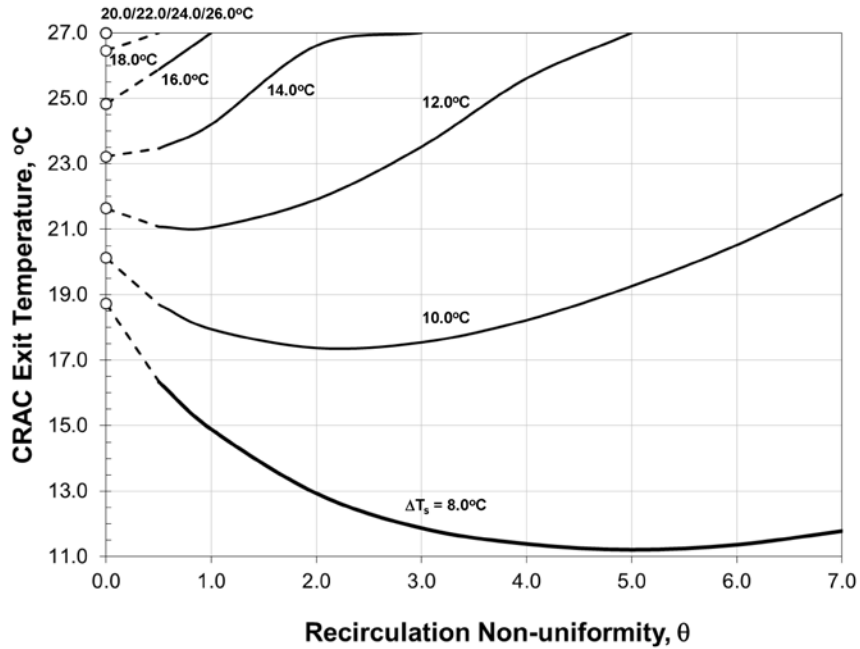


Figure 3.10 - Effect of Server Temperature Rise on Optimum Supply Air Temperature

3.3 Chapter Conclusions

The simple analysis presented in this section provides a flexible and fast tool for exploring optimization possibilities in air-cooled data centers. It can be used for identifying optimal, energy-efficient designs and operating scenarios. The methodology embodied in this simple analysis can be used in the early stages of the conceptual design process to define energy saving approaches and near-optimum design and operating parameters such as flow rates and air supply temperatures, as well as to carry-out tradeoff investigations of cooling infrastructure sizing and performance characteristics. While this simple model is not a substitute for detailed, higher fidelity analysis and optimization

studies, its most useful benefits stem from its ability to limit the range of options and parameters to be explored in the resource-intensive more rigorous optimization analyses.

Through the simple model, the following was revealed:

1. the importance of the trade-off of low air supply temperature vs. increased air flow rate
2. the energy-saving potential of bypass recirculation in enclosed aisle configurations
3. the effect of server temperature rise (flow rate) on energy optimization
4. the deleterious effect of flow non-uniformity at the inlet of the racks on the data center's cooling infrastructure power consumption.

A more detailed account of the latter factor requires detailed CFD analyses or extensive temperature mapping of the data center air space and is the subject of the next chapter.

4 Verifying Assumptions of the Simple Model

The simple model of an air-cooled data center, developed in Section 2.1, was most useful in limiting the range of options and parameters to be explored in the resource-intensive more rigorous optimization analyses. However, a number of assumptions were necessary to allow for a simplified analysis. This section focuses on verifying many of these assumptions using higher fidelity computational fluid dynamics simulations.

4.1 Computational Domain and Setup

Throughout this work, different data center geometries are created to investigate several different aspects of data center design. Figures 4.1a and 4.1b provide detailed schematics of two such geometries, specifically, a data center with a 16 rack aisle (Figure 4.1a) and a data center with a 10 rack aisle (Figure 4.1b). Both of these geometries conform to a standard hot-aisle/cold-aisle arrangement, with a 36" deep under-floor plenum.

Guidelines for determining cold-aisle spacing, hot-aisle spacing and CRAH placement were obtained from Rasmussen and Torell (2007). Symmetry boundary conditions are applied at the right and front walls; therefore, only $\frac{1}{4}$ of the data center is modeled. The racks and CRAH dimensions are obtained using industry standard equipment. The racks are divided into 4 "chassis" in which different boundary conditions could be applied. A chassis is $\frac{1}{4}$ the height of the rack, which conforms to typical 8U blade server configurations. Three pathways are provided for air to flow from the under-floor plenum to the raised-floor space, through perforated tiles (2' x 2'), through cable cut-outs behind

the racks (8" x 8") and distributed leakage over the entire raised floor (representative of leakage through the seams of tiles). Each of these paths is modeled as a porous media, where the percentage of flow emanating from each of the three paths is determined by modifying the coefficients in Equation 2.52. The CRAH units discharge air horizontally into the under-floor plenum, as if a turning vane were installed. Table 4.1 provides details of all the boundary conditions used in the simulations.

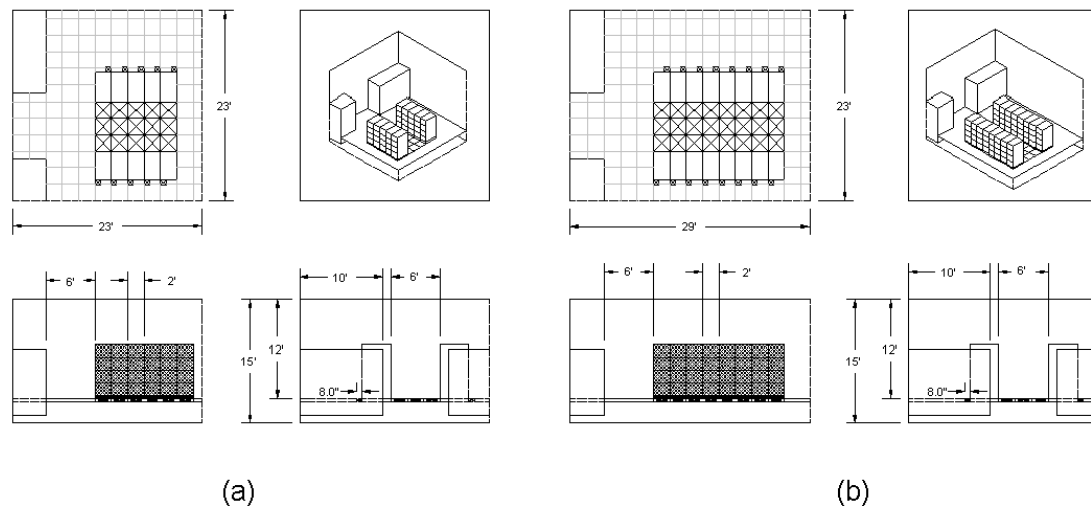


Figure 4.1 - Data Center Geometries Used in CFD Studies

a) 10 rack aisle and b) 16 rack aisle

The incompressible Navier-Stokes equations were solved using the commercial software package, Fluent. The Realizable $k-\epsilon$ turbulence model was used along with the standard wall treatment option as discussed in Section 2.3. Standard wall treatment is appropriate so long as the wall y^+ is greater than 30, which was verified for these simulations.

Second-order accurate upwind schemes were employed to solve the momentum and

energy equations, and a second-order accurate scheme was used for the pressure interpolation. The SIMPLEC algorithm was used for pressure-velocity coupling. The computational geometries grids were developed using the commercial software, Gambit. A structured grid was created where the cell size was (2" x 2" x 2"). It is worth noting that the cell size used in this dissertation is much finer than what is typically used in industry. VanGilder and Zhang (2008) recommended a cell size of (6" x 6" x 6"). However, prior to performing this work, a grid sensitivity study was conducted on a small data center test cell. Grid sizes of (2" x 2" x 2"), (4" x 4" x 4") and (6" x 6" x 6") were studied. The 4" and 6" grid size both exhibited high numerical diffusion compared to the 2" cell size. This added numerical diffusion had a significant effect on under-predicting the mixing and recirculation of hot air into the cold aisle, which is the key driver for increased temperatures at the inlet to the IT equipment. To this end, the (2" x 2" x 2") cell size was used, which resulted in $\sim 1.9M$ cells in the 16-rack geometry shown in Figure 4.1a.

Table 4.1 - Boundary Conditions Used in Computational Domain

Feature	Boundary Condition Type	Specification
Racks	<i>Recirculation</i>	\dot{m} , P , I (leaving), l (leaving)
CRAH Supply	<i>Velocity</i>	V , T , I , l
CRAH Return	<i>Outflow</i>	% of flow at each outflow
Walls	<i>Wall</i>	<i>No-slip, adiabatic</i>
Perforated Tiles, Cut-Outs, Raised-Floor	<i>Porous Media</i>	C_{ij}

4.2 Practical Implementation of Bypass Recirculation

One of the main advantages of enclosed aisle data centers is the ability to provide thermally uniform conditions at the inlet of the IT equipment. Section 3.1 highlighted that the energy savings potential of the bypass recirculation branch in conjunction with an enclosed cold aisle relied on the ability to mix thoroughly the bypassed recirculated air with the chilled air provided by the CRAH to the under-floor plenum before providing this air through the perforated tiles, as depicted in Figure 3.2. This section uses computational fluid dynamics to investigate two practical configurations for implementing a bypass recirculation branch in air-cooled, raised-floor data centers, to assess the ability of providing thermally uniform conditions to the IT equipment.

Figure 4.2 provides the plan view for two possible configurations that allow for bypassing a fraction of the air into the under-floor plenum. Both configurations assume that floor tiles, equipped with low-lift fans, are used to push the bypassed air into the under-floor plenum (these tiles have a star pattern). The fans would be designed to overcome the modest pressure differential between the under-floor plenum and raised-floor space. The data center geometry given in Figure 4.1a is modified to include an enclosed aisle with a non-porous roof and side barriers, which forces the tile airflow to be equal to the required server airflow. The under-floor plenum feeds cooling air to the raised floor space through perforated tiles (~85% of CRAH flow), cable cut-outs behind the racks (~14% of CRAH flow) and distributed leakage over the raised floor (~2% of CRAH flow).

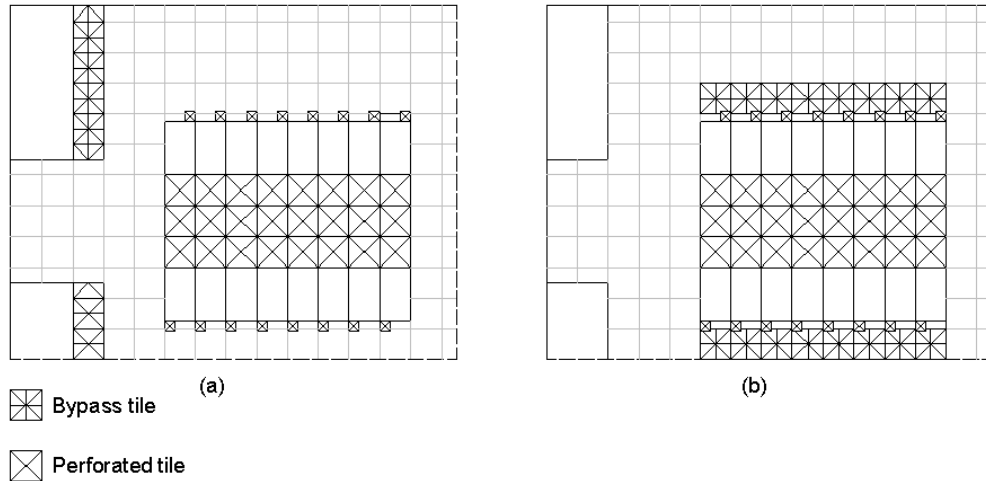


Figure 4.2 - Layouts for Bypass Recirculation

a) Bypass near CRAH and b) Bypass in Hot Aisle

Figure 4.2a shows *layout 1* in which the bypass tiles are placed in front of the CRAH units. The rationale for this configuration is that the close proximity of the bypass tiles to the CRAH units will promote mixing before the air arrives near the perforated tiles.

Figure 4.2b shows *layout 2* in which the bypass tiles are placed in the hot-aisle in an attempt to entrain the hot air into the cold air before it exits through the perforated tiles.

Additionally, *layout 1* is used to investigate the difference in under-floor air distribution for two different CRAH flow arrangements, specifically, a CRAH with an installed turning vane that discharges the flow horizontally in the plenum and a CRAH without the turning vane that discharges the flow vertically into the plenum.

As an example, consider an enclosed cold aisle data center populated with typical high volume servers, which have a temperature rise of 16°C. From the results in Figure 3.4, the optimum operation of this data center would be when 12% of the required rack flow

was bypassed around the CRAH and 86% of the flow was provided by the CRAH and cooled to a temperature of $\sim 24.3^{\circ}\text{C}$. In this analysis, the bypassed flow is distributed equally among the bypass tiles.

Since the temperature uniformity of the air provided to the cold aisle is a consequence of mixing in the under-floor plenum, it is important to prescribe appropriate turbulence boundary conditions at the discharge of the CRAH units. To this end, measurements were performed at the discharge of a typical CRAH unit with an installed turning vane using an omni-directional hot-wire anemometer. Figures 4.3a and 4.3b give the results of the measurements at the discharge of the turning vane at two locations: a) in-line with the exit of the blower and b) in between the two blowers. The results shows that while the velocity was relatively uniform, between 7.0 and 8.0 m/s over the discharge, the turbulence intensity (I) varied significantly from 10 - 70%. Since details of the CRAH geometry are not modeled in the CFD simulations performed throughout this dissertation, only an average turbulence intensity of 50% at 7.0 m/s is used (resulting in an average turbulent kinetic energy, $\bar{k} = 18.4 \text{ m}^2/\text{s}^2$). Turbulence length scale values (l) could not be measured and instead the recommendation for fully developed channel flow of $0.07D_H$ is used (Fluent, 2011)², even though the flow is unlikely to be fully developed at the discharge of the unit.

² It is important to remember that governing equations for k and ε are solved and not I and l . These

quantities are related by, $k = \frac{3}{2}(u_{avg} I)^2$ and $\varepsilon = C_{\mu}^{0.75} \frac{k^{1.5}}{l}$. Therefore, turbulent intensity

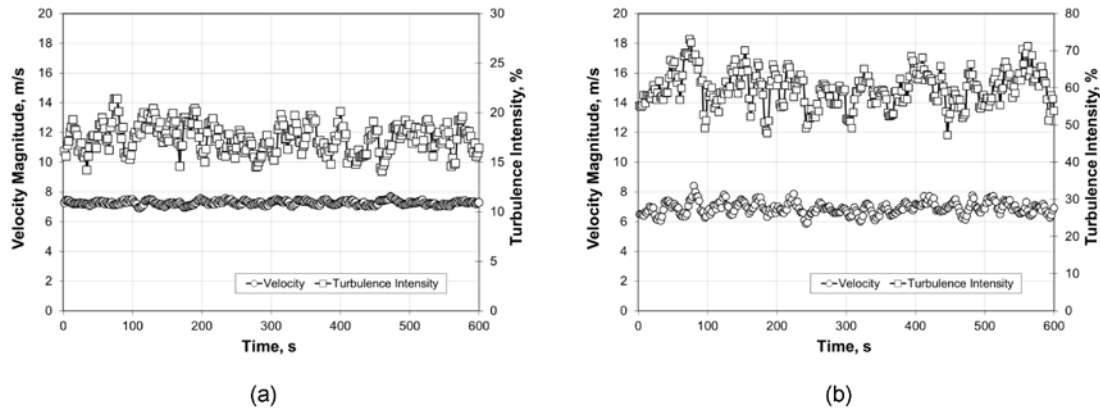


Figure 4.3 - Measured Turbulence Intensity at CRAH Discharge

a) in-line with blower and b) between blowers

Figure 4.4 shows temperature contours 8" below the raised floor in the under-floor plenum for the three cases studied. For the case of *layout 1* with a horizontal discharge (Figure 4.4a), a large amount of mixing occurs because the bypass flow and CRAH flow are discharged in a cross-flow arrangement, which promotes turbulent mixing far from the perforated tiles. In the case of the down-flow CRAH (Figure 4.4b), the two jets are discharged parallel to one another and result in two distinct airflow paths with limited mixing at the interface between jets. This results in a more non-uniform plenum temperature distribution. Figure 4.4c shows the results for *layout 2* with a horizontal flow CRAH. An arrangement with the bypass tiles placed in the hot aisle produces limited mixing between the hot and cold air streams. This is mostly due to the close proximity of the bypass tiles to the perforated tiles.

measurements must always be specified with the velocity in which they were measured in order to match the correct turbulent kinetic energy of the flow.

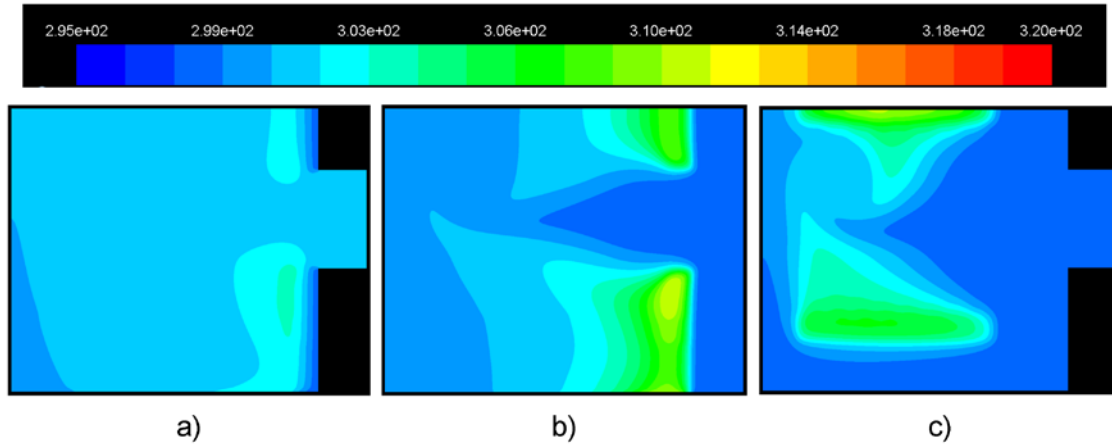


Figure 4.4 – Temperature (in K) Contours 8'' Below the Raised Floor

a) Layout 1 with Horizontal Flow CRAH, b) Layout 1 with Down Flow CRAH and c) Layout 2 with Horizontal Flow CRAH

While the under-floor temperature distribution helps to understand the resulting flow pattern, the real judge of the effectiveness of any of the proposed arrangements is in how uniform the temperature entering the IT equipment is. Table 4.2 gives the range of temperatures experienced at the inlet of the IT equipment in the enclosed cold aisle. *Layout 1*, where the bypass tiles are placed near the CRAH with an installed turning vane, does a reasonable job of mixing the air in the under-floor plenum. This configuration results in only a 1.1°C range of inlet temperature at the IT equipment. Based on Equation 2.10, this arrangement produces a recirculation non-uniformity parameter of $\theta = 0.6$.

Table 4.2 - Rack Inlet Temperature for Various Bypass Geometries

DF = down-flow configuration, HF = horizontal-flow configuration

Flow Arrangement	Layout 1		Layout 2
	HF	DF	HF
$T_{\max}, ^\circ\text{C}$	28.0	27.6	29.3
$T_{\min}, ^\circ\text{C}$	26.9	25.3	24.9
Range, $^\circ\text{C}$	1.1	2.3	4.5

4.3 Additional Parameters for Data Center Airflow Analysis

In the simple model developed in Section 2.1, optimal values of the supply air fraction of rack flow, $\bar{\psi}$, and the cold air supply temperature, T_a , that minimize the cooling infrastructure energy consumption were obtained for given recirculation non-uniformity parameter, θ , server temperature rise ΔT_m and leakage parameter λ . However, neither $\bar{\psi}$ or θ are explicitly known or are easily measured. Other relationships must be developed that define optimum operating conditions in terms of measurable or controllable parameters. The two obvious such parameters are the cooling air supply temperature, T_a , and the total CRAH air supply relative to the rack flow, ψ_c , defined as,

$$\psi_c = \frac{\dot{m}_c}{\dot{m}_m}, \quad (4.1)$$

where, \dot{m}_c and \dot{m}_m are the CRAH and rack flow rates, respectively. To develop such formulations, several key flow rates are defined,

\dot{m}_a = mass flow rate of air emanating from the perforated tiles and captured by the racks

\dot{m}_b = mass flow rate of air emanating from the tiles that bypasses the racks

\dot{m}_l = cold air flow leaking from the plenum outside the cold aisle

\dot{m}_T = cold air flow emanating from the perforated tiles in the cold aisle

These mass flow rates must obey the following relations and give rise to the following definitions,

$$\bar{\psi} = \frac{\dot{m}_a}{\dot{m}_m} \quad (4.2)$$

$$\psi_T = \frac{\dot{m}_T}{\dot{m}_m} \quad (4.3)$$

$$\dot{m}_T = \dot{m}_a + \dot{m}_b \quad (4.4)$$

$$\dot{m}_b = \lambda_b \bar{\psi} \dot{m}_m \quad (4.5)$$

$$\psi_T = \bar{\psi} (1 + \lambda_b) \quad (4.6)$$

$$\dot{m}_c = \dot{m}_a + \dot{m}_b + \dot{m}_l = \dot{m}_T + \dot{m}_l \quad (4.7)$$

$$\dot{m}_l = \lambda_o \bar{\psi} \dot{m}_m \quad (4.8)$$

$$\psi_c = \bar{\psi} (1 + \lambda_b + \lambda_o) = \psi_T \frac{(1 + \lambda_b + \lambda_o)}{(1 + \lambda_b)} \quad (4.9)$$

In this formula, $\lambda_b + \lambda_o$ is equivalent to λ , the leakage fraction used previously in the simple model, in which it was stated that λ represented both the actual leakage through cable cutouts and tile perimeter, as well as any flow that bypasses the racks and is blended with the rack's exhaust. This variable was treated as a given parameter, such that,

$$\bar{\psi} = \frac{\psi_c}{(1 + \lambda)} \quad (4.10)$$

It is expected that λ_b will be a function of ψ_c and that $\bar{\psi} \rightarrow 0$ and $\lambda_b \rightarrow 0$ as $\psi_c \rightarrow 0$ and $\bar{\psi} \rightarrow 1$ as $\psi_c \rightarrow \infty$.

It is expected that a relationship between $\bar{\psi}$ and ψ_c can be obtained using computational fluid dynamics, for typical hot aisle/cold aisle arrangements. This relationship can be elucidated by defining a function F_o , which expresses this relationship in the absence of leakage (i.e., when $\lambda_o = 0$),

$$F_o(\psi_c) = \frac{1}{1 + \lambda_b(\psi_c)}. \quad (4.11)$$

The functions $F_o(\psi_c)$ or $\lambda_b(\psi_c)$ were obtained using CFD for the data center geometries shown in Figures 4.1a and 4.1b, for server temperature rise of 10°C, 15°C and 20°C and

for a range of tile cold air fraction, ψ_T , from 0.50 to 1.25 and with $\lambda_o = 0$. Figure 4.5 shows the results of the CFD study along with one possible curve fit, which can be used to approximate this relation between $\bar{\psi}$ and ψ_c in the simple model. For this analysis, the cumulative normal distribution function was used with $\mu = 0$ and $\sigma = 0.76$. Figure 4.5 also shows the qualitative effect of intentional leakage, λ_o , which was included through the application of Equation 4.9 to the fit of the CFD data. Figure 4.6 shows the relationship between F_o and ψ_c from the CFD data as well as the curve fit using the cumulative normal distribution function. In light of this discussion, expressing the leakage as a fraction of the tile flow that is captured by the racks is not a practical definition for obtaining during operation. Instead, the leakage parameter Λ_o is introduced that expresses the leakage as the ratio of the leakage flow to the CRAH flow,

$$\Lambda_o = \frac{\dot{m}_l}{\dot{m}_c}. \quad (4.12)$$

It can also be shown by mass balance considerations that λ_o and Λ_o are related by,

$$\lambda_o = \frac{\Lambda_o(I + \lambda_b)}{(I - \Lambda_o)} = \frac{\Lambda_o}{F_o(I - \Lambda_o)}. \quad (4.13)$$

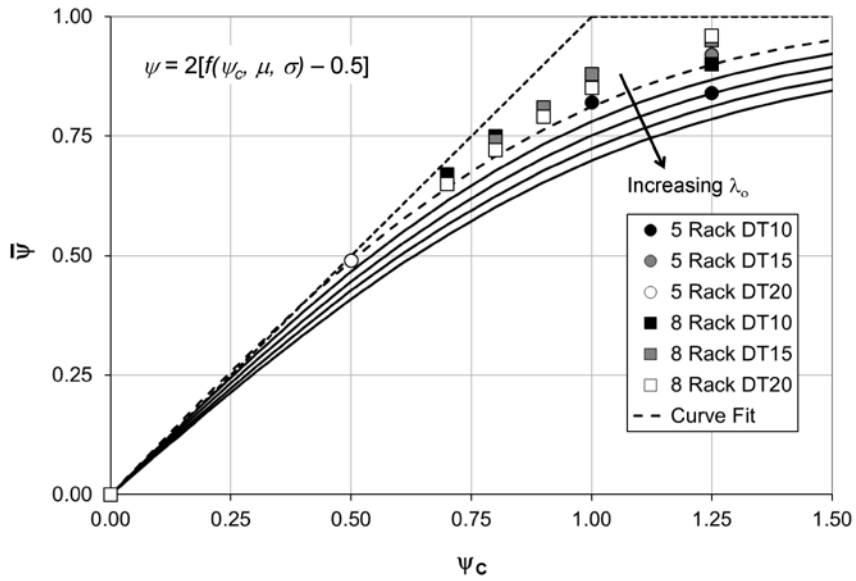


Figure 4.5 - Relationship between $\bar{\psi}$ and ψ_c from CFD

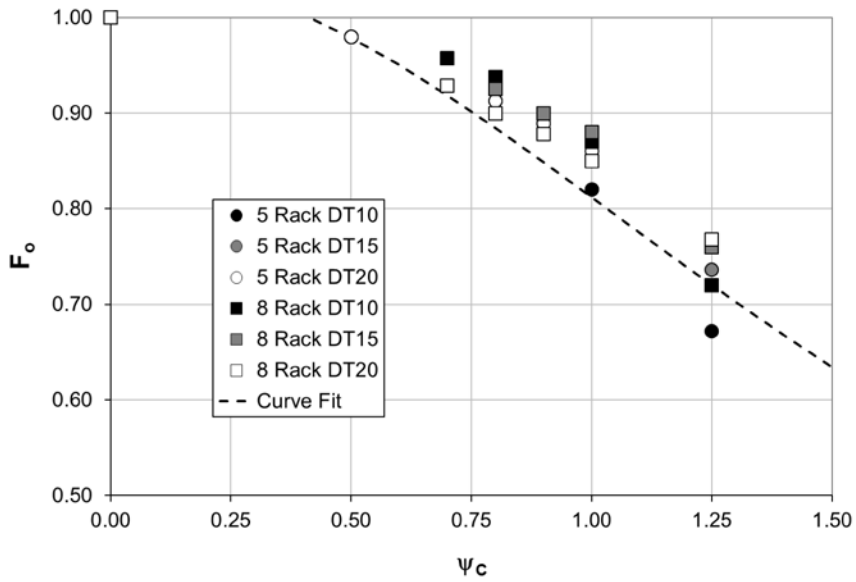


Figure 4.6 - Relationship between F_o and ψ_c from CFD

The leakage parameter used in the simple model, λ , is a composite leakage that includes the effect of both intentional leakage, λ_o (i.e., that through cable cut-outs, perimeter leakage or intentionally open tiles outside the cold aisle), and a “pseudo” leakage, λ_b , which represents the fraction of the air that is introduced to the cold aisle through the perforated tiles but bypasses the racks and either blends with the rack exhaust or is short-circuited to the CRAH units. The pseudo leakage, λ_b can be computed by,

$$\lambda_b = \frac{\psi_T - \bar{\psi}}{\bar{\psi}}. \quad (4.14)$$

The value of the leakage parameter needed in the simple model is the summation of the intentional and “pseudo” leakages. To this end, the results from the CFD study were used to determine a relationship between λ_b and ψ_c for the case of $\lambda_o = 0$, using Equation 4.11. This relationship is shown in Figure 4.7.

The energy consumption results of Figure 3.5 are revisited, this time considering realistic leakage, λ , as a function of CRAH flow rate ratio, ψ_c . Figure 4.8 plots four normalized power consumption curves as a function of cold air supply fraction. The curves labeled “EA (20%)” and “OA (20%)” are the previously report results for the enclosed aisle ($\theta = 0.0$) and open aisle ($\theta = 4.0$) with $\lambda = 0.20$, respectively. The results showed that an optimized open aisle configuration with some non-uniformity ($\theta = 4.0$) used 10% less energy than a conventional enclosed aisle configuration ($\theta = 0$, $\bar{\psi} = 1.0$). However,

while $\lambda = 0.20$ is a reasonable leakage for an enclosed aisle data center, this may not be the case for an open aisle configuration in light of the previous discussion, where it was shown that the leakage λ is actually the summation of the intentional and unintentional leakage. The curve labeled “OA (50%)” in Figure 4.8 shows an example of an open aisle data center with $\theta = 4.0$ but $\lambda = 0.50$. As expected, the optimum point shift towards more recirculation and higher power consumption. In this case, the open aisle data center consumes 16% more power than the conventional enclosed aisle (367 kW vs. 307 kW) and 53% more power than the optimized enclosed aisle configuration with bypass recirculation (367 kW vs. 173 kW). The curve labeled “OA (Variable)” in Figure 4.8 shows the results if the relationship between λ_b and ψ_c given in Figure 4.7 is applied, assuming $\lambda_o = 0.20$. As expected, the power consumptions falls between the $\lambda = 0.20$ and $\lambda = 0.50$ cases, showing negligible savings in power compared to the conventional enclosed aisle case.

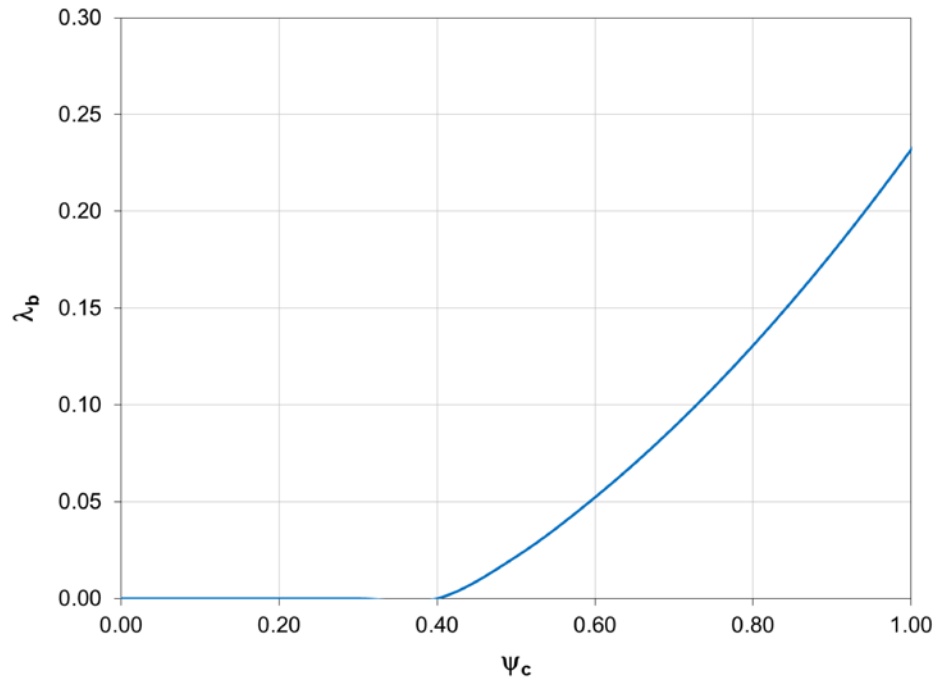


Figure 4.7 - Relationship between λ_b and ψ_c from CFD

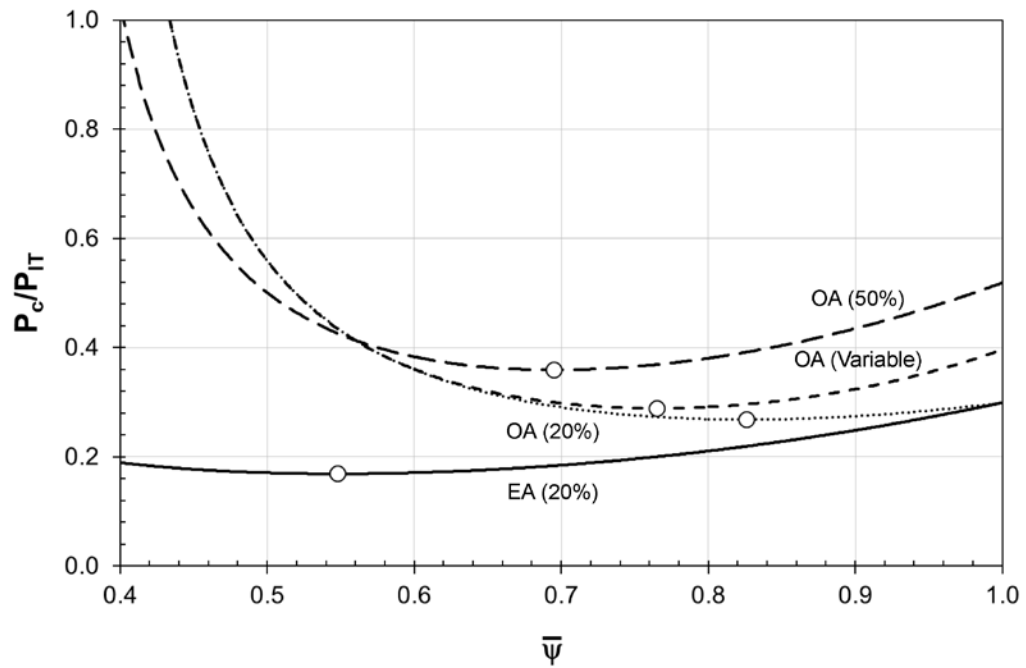


Figure 4.8 - Normalized Energy Results for Different Configurations and Leakage

4.4 Non-Uniformity Parameter, θ

The simple model developed in Section 2.1 introduced a recirculation non-uniformity parameter, θ . This parameter was used parametrically to highlight the deleterious effect that non-uniformity has on the energy consumption of an open-aisle, air-cooled data center. Using several simplifying assumptions, the recirculation non-uniformity was related to the temperature non-uniformity at the inlet to a row of IT racks (see Equation 2.10). The key assumption that led to the development of Equation 2.10 was that all of the air is recirculated at a mixed average exhaust temperature. A consequence of this is that the server with the most recirculation, ϕ_{max} , also has the highest inlet temperature. This assumption plays an important role in computing θ , which was directly related to the overall infrastructure energy consumption.

To assess the validity of this assumption, a number of CFD studies are performed on the data center geometries described in Section 4.1 and the recirculation non-uniformity parameter was computed in two way:

1. its basic definition (Equation 2.1), which is denoted $\bar{\phi}_{act}$
2. its simplified definition (Equation 2.10), which is denoted $\bar{\phi}_{model}$

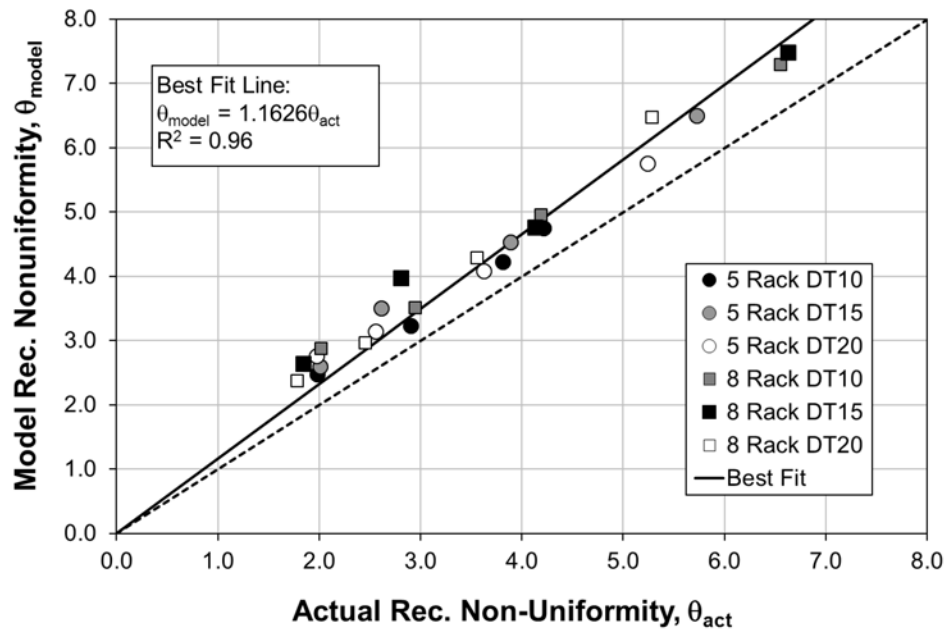
The study was performed for three different rack temperature rises: 10 °C, 15 °C and 20 °C, and 4 CRAH air supply fractions: $\psi_C = 0.7, 0.8, 0.9$ and 1.0. No “intentional” leakage, λ_o , is considered in this study. Equation 2.10 requires the computation of the

actual cold air that enters the IT equipment, which is equivalent to the capture index, $\bar{\psi}$.

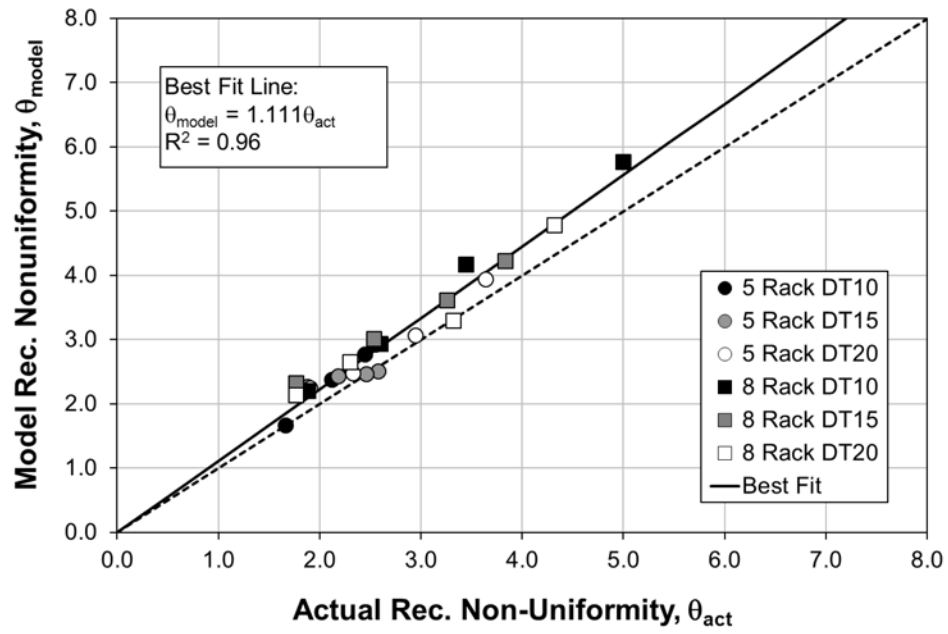
The capture index is computed following the tracer gas approach employed by VanGilder and Shrivastava (2008).

Clearly, the method by which the maximum temperature in the row of IT equipment is computed affects the computed value of θ . For example, a single point measurement at the hottest point would always result in a higher value of θ than an average of multiple sensors. In a real application, an area representative of the most vulnerable server (i.e., the inlet of the hottest server) would have to be used. For this study, as a means of providing a bound to the results, $\bar{\varphi}_{model}$ was computed by two methods. In *method 1*, $\bar{\varphi}_{model}$ is computed based on the maximum temperature in any 2" x 2" grid cell on the face of the IT racks. In *method 2*, $\bar{\varphi}_{model}$ is computed based on the mass-weighted average over the face of a chassis (i.e., 1/4 of a rack).

Figures 4.9a and 4.9b provide a comparison between the recirculation non-uniformity parameter based on the simple model derivation $\bar{\varphi}_{model}$ and the original definition, $\bar{\varphi}_{act}$. If there was no consequence of the assumptions in the simple model and the two methods agreed perfectly, all points would fall on the dashed 45° line. However, the simple model result in an over-prediction of the recirculation non-uniformity, due to the assumption of an average mixed exhaust temperature. As expected, taking a average over a larger area to compute the maximum temperature tends to improve the agreement between $\bar{\varphi}_{model}$ and $\bar{\varphi}_{act}$, as can be seen by comparing Figures 4.9a and 4.9b.



a)



b)

Figure 4.9 - Comparison of Methods to Compute θ

Using: a) 2" x 2" Cell and b) Chassis Average

Figures 4.9a and 4.9b also show best-fit lines to the data, which gives the reader a sense of the error introduced in computing θ using the simple model derivation. A standard deviation of 11% is seen when the maximum temperature is computed based on the 2"x2" grid cell (Figure 4.9a). In the example of the chassis average, a standard deviation of only 8% is seen between the actual recirculation non-uniformity and the estimated non-uniformity. Even with all of the simplifying assumptions used to develop the simple model, the accuracy of the results are fairly good, with an expectation that the recirculation non-uniformity parameter can be estimated directly to within a 10-15% uncertainty by knowing only the temperature distribution at the inlet of the IT equipment.

4.5 Temperature Linearity and a Data Center Specific Archimedes

Number

The dependence of the thermal state on the air supply temperature is expected to be linear if buoyancy and radiation effects are negligible. This stems from the fact that, under these conditions, the momentum equation is decoupled from the energy equation (the velocity field is independent of the temperature field), and the energy equation for low-speed flow is linear in temperature, i.e.,

$$u_i \frac{\partial T}{\partial x_i} = \alpha_o \frac{\partial^2 T}{\partial x_i \partial x_i} \quad (4.15)$$

in which u_i is the i^{th} component of the velocity vector and α_o is the overall (molecular plus turbulent) thermal diffusivity. In high-density datacenters, radiation heat transfer is of negligible effect but buoyancy may be significant at low flow conditions and/or high power conditions. On the other hand, buoyancy effects may be small at high flow conditions, and the temperature field's dependence on the air supply temperature may remain approximately linear, (i.e., a δT change in the supply temperature would cause a δT change of the air temperature everywhere, for adiabatic walls and fixed heat generation rate. For simplicity, this is referred to as having a one-to-one temperature field).

In most low-density data centers, the flow is characterized by forced convection. However, current trends show an increase in data center's heat flux and concurrently the use of CRAH variable frequency drives to reduce the airflow. The combined effect of these trends points towards a flow in the data center that may begin to be affected by buoyancy. As a means of characterizing the effect of buoyancy within high-density data centers, an Archimedes number (Ar) that is tailored for data center applications is introduced. The Archimedes number, which is the ratio of the buoyancy force to inertial force, is defined as,

$$Ar = \frac{g\beta L\Delta T}{u^2}. \quad (4.16)$$

A convenient length scale, L , for data center applications is the rack height, H_r , and temperature scale, ΔT , is the rack temperature rise, ΔT_m . An appropriate velocity scale is considered the fully open perforated tile velocity. The conditions at the perforated tile are related to the rack airflow using the definition of the tile cold air supply fraction, ψ_T . Using these parameters, the Archimedes number can be written as,

$$Ar = \frac{g\beta H_r \Delta T_m (n_t A_t)^2}{\dot{V}_m^2 \psi_T^2} \quad (4.17)$$

where, A_t is the perforated tile area, n_t is the number of perforated tiles per rack and \dot{V}_m is the rack volumetric flow rate. The Archimedes number is a dimensionless quantity that represents the ratio of the buoyancy force to the inertial force in the flow. An $Ar \gg 1.0$ indicates a flow that is dominated by natural convection; whereas, an $Ar \ll 1.0$ indicates a flow dominated by forced convection. Since many of the geometries in the data centers are based on industry standard dimensions, Equation 4.17 can be simplified to,

$$Ar = 0.009 \frac{\Delta T_m n_t^2}{\dot{V}_m^2 \psi_T^2}, \quad (4.18a)$$

where, $H_r = 1.98 \text{ m}$, $A_t = 0.3716 \text{ m}^2$, $g = 9.81 \text{ m/s}^2$ and $\beta = (1/293) \text{ K}^{-1}$. Alternatively, Equation 4.17 could be written in terms of rack power, P_m , as,

$$Ar = 0.014 \frac{\Delta T_m^3 n_t^2}{P_m^2 \psi_T^2}, \quad (4.18b)$$

where, the rack P_m is related to \dot{V}_m using an energy balance of the racks (i.e.

$P_m = \rho c_p \dot{V}_m \Delta T_m$). The choice of parameter used in this development of Ar were done

based for convenience; however, other parameter could certainly be used that might be more easily measured in data centers, such as relating ΔT to the supply and return temperature difference of the CRAH unit and u to the CRAH airflow rate.

As Ar becomes greater than unity, it is expected that the assumption that a δT change in the supply air temperature would result in a δT change in the temperature everywhere would become weaker. To assess the importance of buoyancy, CFD simulations are run at three different supply air temperatures (14 °C, 18 °C and 22 °C) for two cases, $Ar = 0.22$ and $Ar = 1.80$, for the 10 rack geometry shown in Figure 4.1. Figures 4.10a and 4.10b compare the mass-weighted average inlet temperature for several of the chassis in the upper portion of the IT racks. The CFD simulations are represented as solid lines, while the dashed lines represent the inlet temperatures if the temperature field remained perfectly one-to-one. A comparison of Figures 4.10a and 4.10b show that while the temperature response does remain linear, the assumption of one-to-one is weaker for the case of $Ar = 1.80$.

Buoyancy also plays a role in affecting the temperature distribution at the inlet to a row of IT racks. At higher Ar , the temperature distribution becomes more stratified - a consequence of the cold air remaining closer to floor and the hot air rising in the presence of increased buoyancy force. By comparison, for low Ar cases, the flow is dominated by forced convection and is more jet-like. Figures 4.11 and 4.12 compare the temperature distribution at the inlet to a row of IT racks in the cold aisle (the figures show the inlet plane of the cold aisle as if you were standing in the cold aisle facing the 5 racks and each rack has 4 chassis). Figure 4.11 illustrates the difference in temperature patterns as Ar increases for a given tile air supply fraction. The top picture shows the case of a low Ar where the flow is dominated by forced convection, which results in almost all the recirculation being entrained by the outer racks. The bottom picture show the case of a high Ar , where the flow is dominated by buoyancy. The temperature pattern exhibits pronounced stratification, where the servers at the upper portion of all racks receive a significant portion of recirculated air, not just those racks at the end of the aisle. Similarly, Figure 4.12 shows another example of the various flow regimes in high-density data centers with a lower air supply fraction ($\psi_T = 0.8$). In both cases, the asymmetry of the temperature patterns is attributed to the location of the CRAH units, which are located closer to the left-most rack. It is worth noting that as the momentum of the tile flow decreases (reduced ψ_T , which has the effect of increasing Ar), the effect of CRAH suction plays a more important role in distorting the temperature pattern. For example whereas the top picture in Figure 4.11, which has relatively high momentum ($V_a = 1.74 \text{ m/s}$), shows a jet-like profile, which is only slightly distorted, the bottom picture in Figure 4.12

shows a highly distorted, stratified temperature profile because of the low momentum of the perforated tile flow ($V_a = 0.69 \text{ m/s}$).

To ascertain that the stratification is indeed an effect of buoyancy and not just a consequence of a lower momentum jet, a simulation was done of $Ar = 2.81$ ($\Delta T_m = 20^\circ\text{C}$ and $\psi_T = 0.8$) but the effect of buoyancy was excluded. Figure 4.13 compares the temperature contours with and without buoyancy. It can be seen that the temperature stratification is indeed a consequence of the buoyancy-dominated flow for higher Ar cases. Even for a case where the momentum of the perforated tiles is relatively low ($\psi_T = 0.8$), without buoyancy a jet-like flow is produced since the buoyancy force is not keeping the cold air towards the floor.

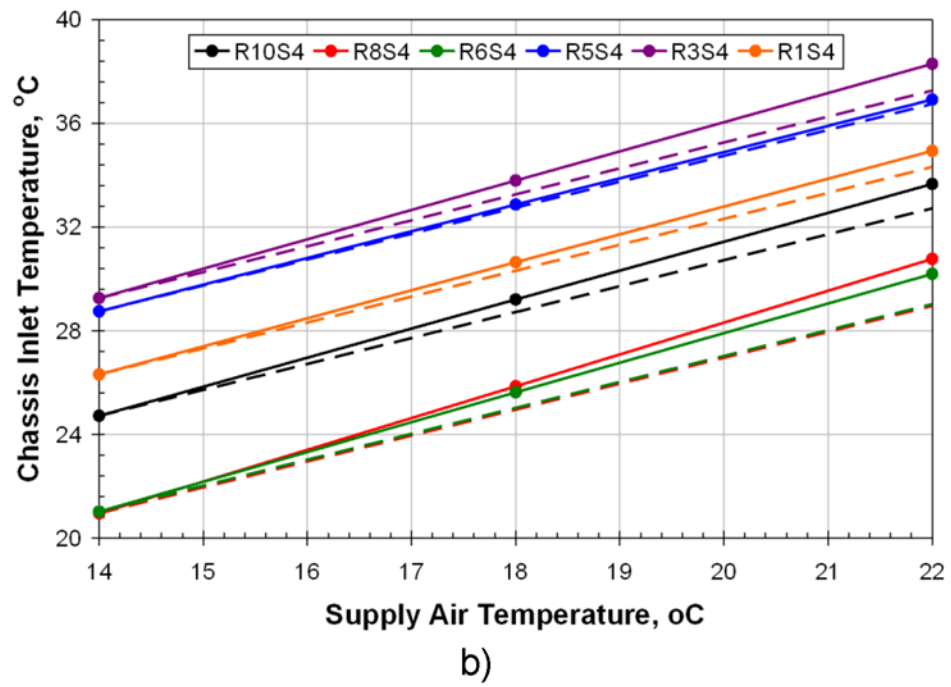
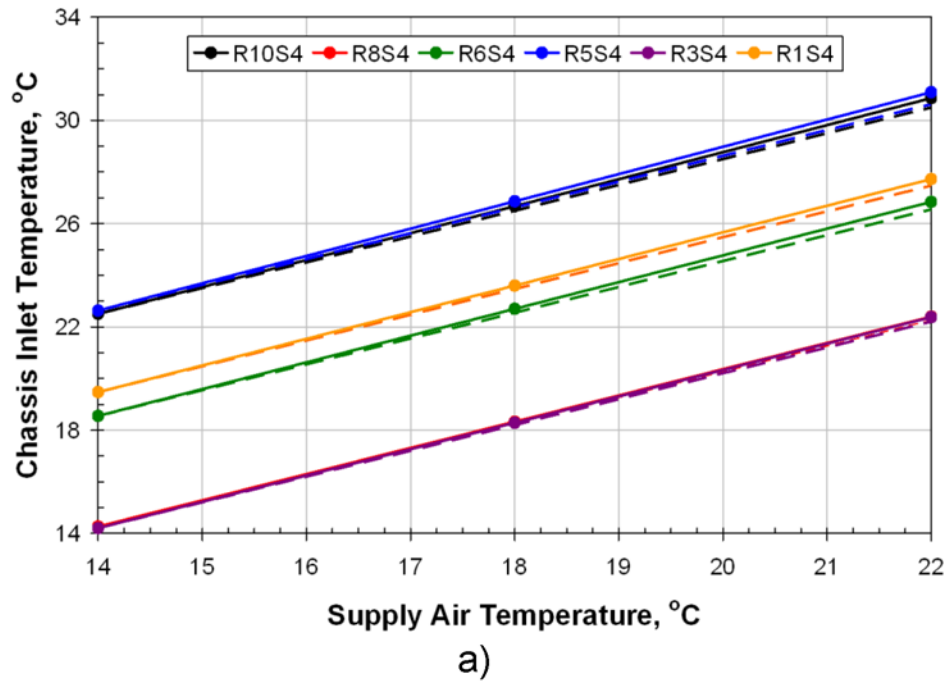


Figure 4.10 - Effect of Ar on Temperature Linearity

a) Ar = 0.22 and b) Ar = 1.80

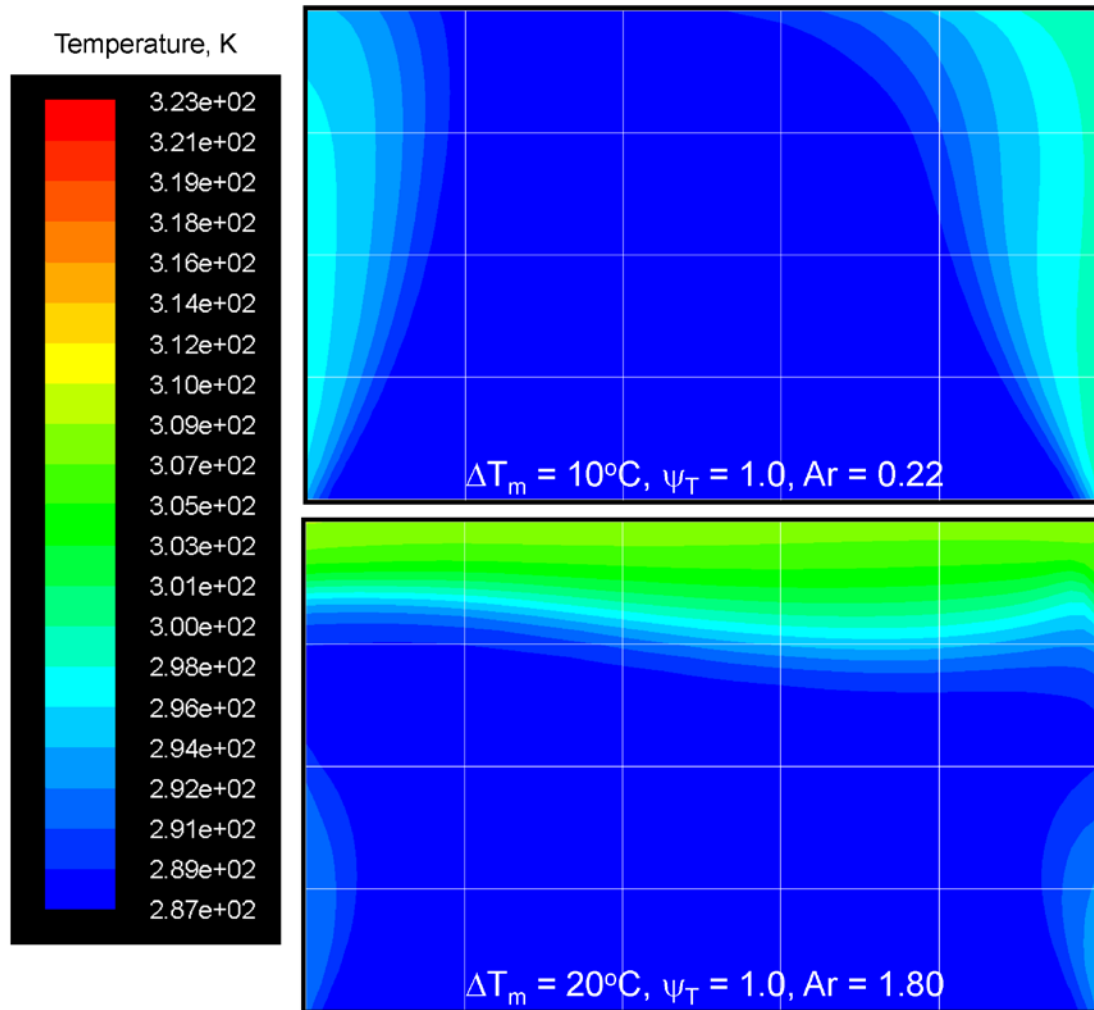


Figure 4.11 - Inlet Temperature Contours for $\psi_T = 1.0$

Top: $\Delta T_m = 10^\circ\text{C}, Ar = 0.22$ Bottom: $\Delta T_m = 20^\circ\text{C}, Ar = 1.80$

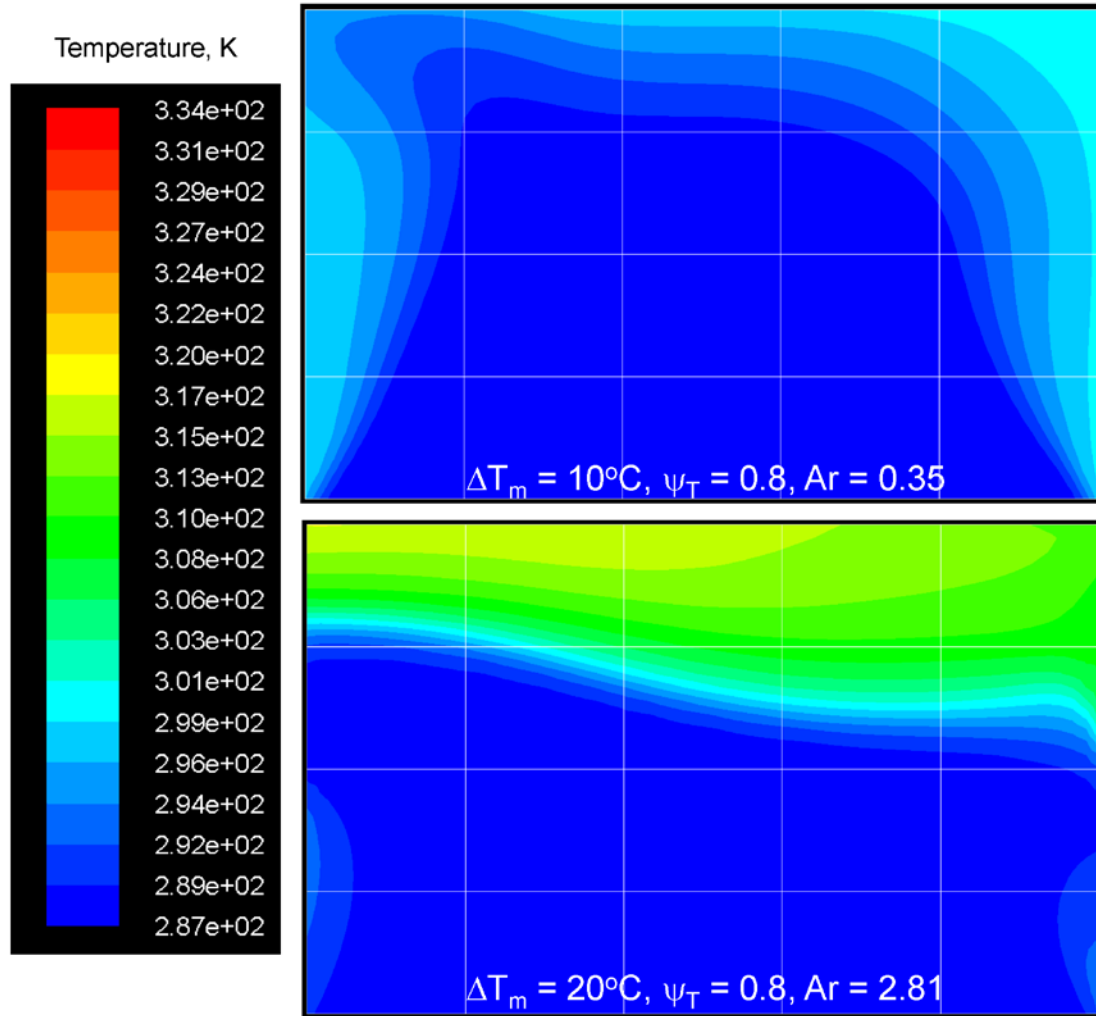


Figure 4.12 - Inlet Temperature Contours for $\psi_T = 0.8$

Top: $\Delta T_m = 10^\circ\text{C}, Ar = 0.35$; Bottom: $\Delta T_m = 20^\circ\text{C}, Ar = 2.81$

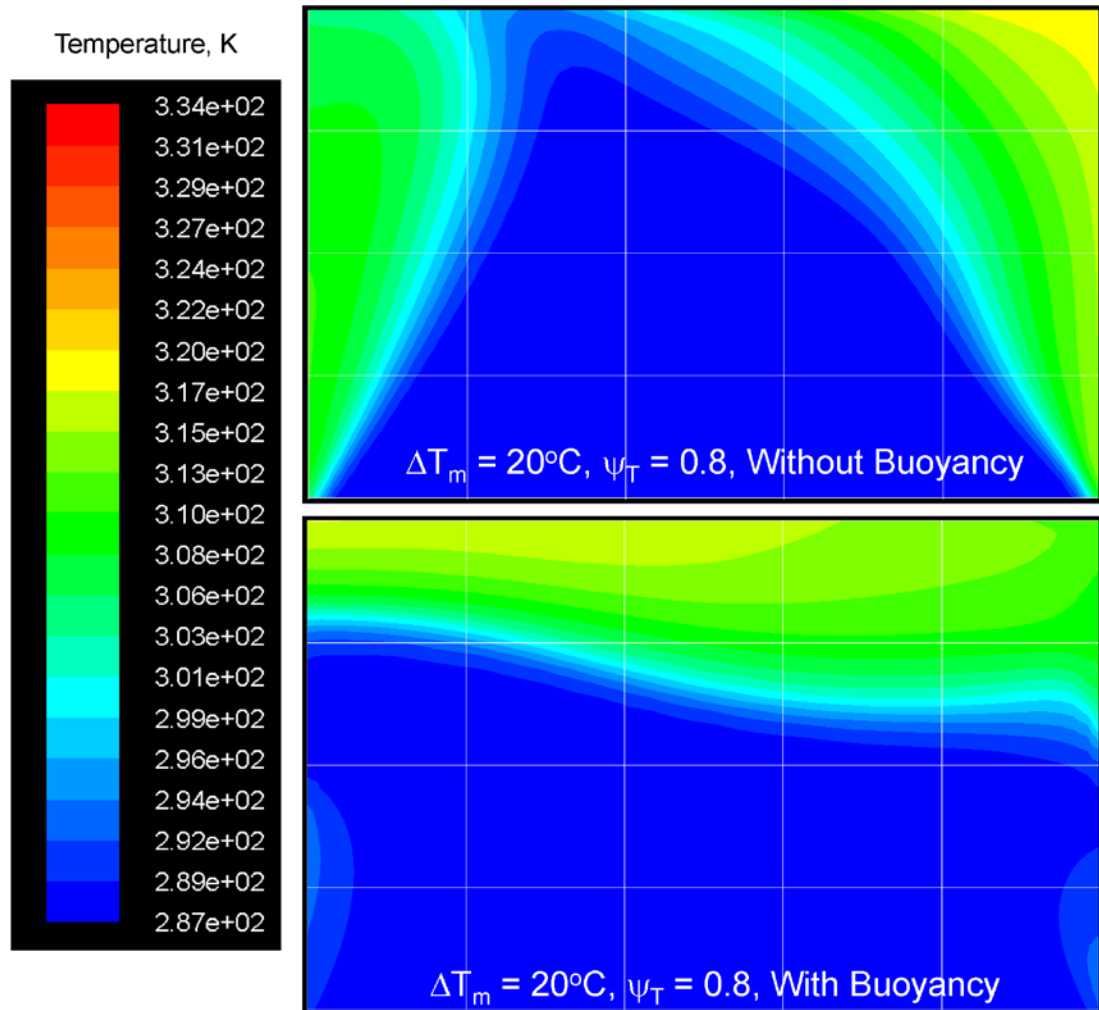


Figure 4.13 - Inlet Temperature Contours Comparing Buoyancy for $\psi_T = 0.8$ $\Delta T_m = 20^\circ\text{C}$

Top: without buoyancy; Bottom: with buoyancy

4.6 Chapter Conclusions

This chapter used computational fluid dynamics to verify several of the assumptions made during the development of the simple model of a data center and its cooling infrastructure presented in Section 2.1. Several additional parameters were introduced that quantify air distribution in air-cooled data centers. Specifically, these parameters looked at various forms of leakage flow present in data centers and its effect on the overall infrastructure energy consumption was studied. The results showed that in open-aisle data centers, the leakage parameter λ , introduced in the simple model, could be significantly higher than in an enclosed aisle configuration. When realistic leakage was considered, the energy savings of the optimized open aisle data center vanished compared to a conventional enclosed aisle configuration. In addition, the energy savings of the bypass recirculation branch in an enclosed aisle data center are greater when compared to the open aisle data center with realistic leakage.

Next, it was shown that a reasonably accurate estimate of the recirculation non-uniformity parameter, θ , could be obtained from the rack inlet temperature distribution, even with the inherent assumption of a single mixed exhaust temperature.

The effect of buoyancy in high-density data centers was studied and provided insight into the drastic effect buoyancy can have on changing the temperature patterns in the data center for cases where the Archimedes number is greater than unity. The analysis shows the difference in exhaust recirculation patterns to the inlet of the racks in an aisle in the

presence of a strong buoyancy force, which is likely in many high-density applications. It also shows that the assumption of a linear, one-to-one change in the temperature field with a change in the supply air temperature is a weak assumption for high Ar data centers.

5 Experimental Validation of the Thermo-Hydraulic Model

In Section 2.2, a computational tool was presented that allows data center operators and designers the ability to evaluate the energy consumption of their data center configuration. Inherent in this model was the ability to capture off-design operating conditions of the data center's cooling infrastructure caused by changes in ambient conditions and fluctuations in required IT load. In this chapter, the thermo-hydraulic model is experimentally validated based on an existing IT data center and chiller plant located in Poughkeepsie, New York. The model is used to study the data center under several climatic and operating conditions – namely, winter and summer. The resulting performance in terms of the PUE was found to be 1.57 and 1.73, respectively, for the two climatic conditions.

5.1 Buildings 710 and 027 Data Collection

The data center and chilled water plant used for data collection was a “tier 3” data center (B710) located on the IBM campus in Poughkeepsie, New York. As outlined by the Uptime Institute (2010), a “tier 3” data center is “concurrently maintainable,” providing 99.982% reliability. The data center consists of multiple power and cooling distribution paths; however, only one of them is active, resulting in (N+1) redundant components. B710 houses three data centers for a total raised-floor area of 12,913 m² (39,000 ft²). The total design heat load of the three data centers was 6245 kW. Other auxiliary and HVAC loads of the data center (UPS, lighting, etc.) totaled 1459 kW. The data center is served

by a near-by chiller plant (B027) with a redundant refrigeration capacity of 11,957 kW (3400 tons). Variable frequency drive (VFD) capabilities are provided on the chilled water pumps and cooling tower fans. Table 5.1 provides an inventory of the equipment in both the data center and the chilled water plant along with the design rating as per manufacturer specification. Detailed schematics and performance data is given in Appendix A.

Table 5.1 - Inventory of Equipment in IBM B710 and B027

Item	Quantity	Design Rating
Raised Floor Area	-	Lower Level: 2787 m ² (30,000 ft ²) First Floor: 3066 m ² (33,000 ft ²) Second Floor: 7060 m ² (76,000 ft ²)
Design IT Heat Flux	-	Lower Level: 269 W/m ² First Floor: 431 W/m ² Second Floor: 592 W/m ²
Auxiliary and HVAC Heat Loads (i.e., UPS, Lighting, etc.)	-	~1460 kW
Centrifugal Chillers	2	4220 kW (1200 ton), 0.56 kW/ton
R-134a, Vane Controlled	1	3517 kW (1000 ton), 0.59 kW/ton
Cooling Towers	3	462,471 m ³ /h, 44.7 kW
CRAH	134	105 kW (30 ton), 21,068 m ³ /h, 5.56 kW
Chilled Water Pump	3	75 kW, 636 m ³ /h, 30.5 m
Cooling Water Pump	3	75 kW, 681 m ³ /h, 26 m

A set of data during the actual operation of the data center was obtained over a 24-hour period in February 2009, where temperature, flow and power measurements (as appropriate) were made on all of the components listed in Table 5.1. During the day of data collection, the data center was operated in waterside economizer mode between 00:00 - 13:00 h and chiller mode from 14:00 – 23:00 h. For validation purposes, the data collected during chiller operation was used. Based on the collected data, the PUE of the data center was 1.57 during the operation of the chiller. This includes the power consumed in all of the components in the data center and the cooling infrastructure. The total measured evaporator load during data collection was ~4 MW. Tables 5.2 – 5.4 provide the measurements done at the chillers and cooling tower. During the collection period, only a single chiller (1200-ton model) and a single cooling tower were operated.

Data was collected at each of the CRAH units inside the data center over a 4-h period. Because of the complexity in acquiring these measurements, only a single dataset of all 134 CRAH units was collected. The measurements were obtained by walking through the data center, measuring, and recording the operation of each of the 134 CRAH units. This was done between the hours of 13:00 - 17:00. Due to time constraints, the airflow rate through each of the CRAH units could not be measured and instead was estimated as an average value based on previous measurements to be 16,990 m³/h (10,000 CFM). In the previous work, the flow rate through the CRAH units was measured using a calibrated flow hood balometer with an accuracy of $\pm 3\% \pm 12 \text{ m}^3/\text{h}$ ($\pm 3\% \pm 7 \text{ CFM}$). Even though an average value was used in this study, CRAH-to-CRAH airflow differences are expected due to difference in control strategy, under-floor pressure, obstructions at the CRAH

discharge, and dirty filters. Because of the operational nature of this data center, intervention was not possible to investigate these effects or obtain new airflow measurements. Supply and return air temperatures were measured in each CRAH unit using a handheld thermocouple. The thermocouple accuracy was estimated based on manufacturer specifications to be $\pm 1^{\circ}\text{C}$ ($\pm 1.8^{\circ}\text{F}$). For each of the 134 CRAH units, the temperature was measured at one point at the intake and one point at each of the two fan exhausts. Tables 5.5 – 5.8 gives the result of these measurements. The numerical average of the two fan discharge measurements is given in the tables. Highlighted entries are CRAH units that were moving air but not removing any load. The increase in temperature of the air is due to the fan heating effect.

Table 5.2 - Refrigeration Unit Measurements

Time	Evaporator Load, kW	Input Current, amps
14:00	4721	186
15:00	3841	134
16:00	4021	138
17:00	4039	142
18:00	4116	138
19:00	4166	140
20:00	4060	136
21:00	3876	133
22:00	3979	134
23:00	4003	132

Table 5.3 - Evaporator and Condenser Temperature Measurements

Time	Chilled Water, °C	Condenser Water, °C	
		Return	Supply
14:00	7.9	25.0	18.3
15:00	7.2	19.4	13.3
16:00	7.2	18.3	12.8
17:00	6.9	19.4	13.3
18:00	7.2	19.4	12.8
19:00	7.3	19.4	12.8
20:00	7.2	18.9	12.8
21:00	7.1	18.9	12.8
22:00	7.1	18.9	12.8
23:00	7.2	18.3	12.2

Table 5.4 - Cooling Tower Fan Measurements

Time	Fan Speed, Hz (Tower #1)
14:00	51
15:00	60
16:00	60
17:00	60
18:00	60
19:00	56
20:00	54
21:00	53
22:00	52
23:00	51

Table 5.5 - CRAH Temperature Measurements: Lower Level

Unit	Return, °C	Supply, °C	Unit	Return, °C	Supply, °C
1	OFF		21	22.6	19.6
2	OFF		22	23.2	24.1
3	23.0	15.1	23	21.9	20.7
4	25.4	18.1	24	22.7	23.7
5	23.3	23.7	25	22.6	20.0
6	25.4	24.8	26	22.3	22.8
7	23.7	18.3	27	22.0	15.0
8	27.2	16.9	28	22.0	15.0
9	22.8	10.0	29	22.0	15.0
10	22.3	23.7	30	22.0	15.0
11	22.6	23.3	31	22.0	15.0
12	23.3	22.6			
13	23.5	21.9			
14	OFF				
15	OFF				
16	OFF				
17	21.8	11.6			
18	OFF				
19	OFF				
20	22.5	21.1			

Table 5.6 - CRAH Temperature Measurements: First Floor

Unit	Return, °C	Supply, °C	Unit	Return, °C	Supply, °C
1	23.6	24.5	21	22.6	15.1
2	23.6	20.1	22	19.5	16.1
3	23.0	20.9	23	OFF	
4	12.7	13.9	24	OFF	
5	23.6	18.0	25	OFF	
6	23.9	12.4	26	OFF	
7	22.2	22.6	27	OFF	
8	22.6	11.3	28	OFF	
9	22.4	20.7	29	OFF	
10	24.1	11.6	30	OFF	
11	21.2	10.8	31	OFF	
12	22.1	22.8			
13	21.6	12.1			
14	21.5	22.4			
15	21.1	22.2			
16	21.8	21.4			
17	22.9	17.2			
18	23.0	23.7			
19	24.4	18.0			
20	22.5	20.0			

Table 5.7 - CRAH Temperature Measurements: Second Floor E

Unit	Return, °C	Supply, °C	Unit	Return, °C	Supply, °C
1	26.7	19.5	21	OFF	
2	26.0	16.3	22	23.4	19.4
3	24.8	17.4	23	26.0	12.3
4	25.1	17.0	24	25.1	13.3
5	23.6	11.4	25	24.2	15.9
6	OFF		26	23.5	12.6
7	24.7	12.5	27	OFF	
8	22.8	12.9	28	OFF	
9	OFF		29	OFF	
10	23.8	12.2	30	24.8	12.7
11	22.4	17.7	31	24.4	12.9
12	24.1	13.4	32	26.2	13.4
13	24.6	18.3	33	25.9	13.2
14	23.9	13.1			
15	23.3	16.4			
16	OFF		36	24.7	12.7
17	25.9	20.1	37	26.4	12.7
18	OFF				
19	OFF		39	27.6	12.6
20	25.4	20.9			

Table 5.8 - CRAH Temperature Measurements: Second Floor

Unit	Return, °C	Supply, °C	Unit	Return, °C	Supply, °C
34	OFF		54	25.7	13.4
35	26.7	19.5	55	26.4	28.1
36			56	25.3	10.7
37			57	24.8	19.8
38			58	23.7	18.6
39	OFF		59	22.7	13.1
40	25.3	16.3	60		
41	27.6	10.2	61	24.4	11.9
42	23.9	14.7	62	22.7	12.7
43	24.8	16.8	63	22.6	19.5
44	24.4	9.9	64	23.9	20.7
45	24.0	20.5	65	25.4	22.1
46	24.0	20.0	66	22.9	14.6
47	25.7	22.6	67	OFF	
48	OFF		68	22.7	23.7
49	24.6	11.5	69	OFF	
50	25.3	10.5	70	21.9	17.3
51	27.5	12.1	71	24.1	19.2
52	26.1	12.0	72	21.3	14.8
53	OFF				

During the CRAH measurement period, measurements of the IT-connected UPS loads and the CRAH unit fan input power were obtained. These measurements are given in Tables 5.9 – 5.11. It was estimated that in addition to the loads given in Tables 5.9 and 5.10, there was 108 kW of unconnected UPS IT load on the lower level and 180 kW of unconnected UPS IT load on the first floor. The difference between these measurements and the evaporator load measurements given in Table 5.2 is considered non-IT related auxiliary load, which was unable to be measured during data collection. Ambient weather data was obtained from the Poughkeepsie, New York airport weather station (Weather Underground, 2009) for the day that the validation data was collected. This data is provided in Figure 5.1.

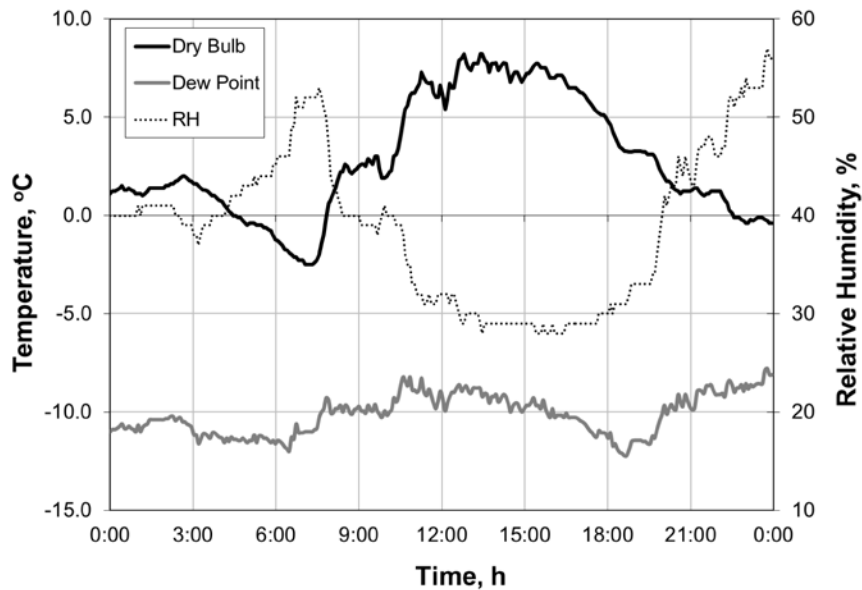


Figure 5.1 - Weather Data for a Winter Day in Poughkeepsie, NY

Table 5.9 - Raised Floor Electrical Measurements: Lower Level

Time	CRAH Fan Power, kW	UPS-connected IT Load, kW
13:00	96	257
13:15	96	258
13:30	96	258
13:45	96	258
14:00	96	257
14:15	96	258
14:30	96	257
14:45	96	258
15:00	96	257
15:15	96	259
15:30	96	259
15:45	96	258
16:00	96	259
16:15	96	259
16:30	96	259
16:45	96	259

Table 5.10 - Raised Floor Electrical Measurements: First Floor

Time	CRAH Fan Power, kW	UPS-connected IT Load, kW
13:00	162	264
13:15	159	264
13:30	161	264
13:45	159	264
14:00	161	264
14:15	160	264
14:30	159	264
14:45	161	264
15:00	159	264
15:15	161	264
15:30	157	264
15:45	154	264
16:00	143	264
16:15	140	264
16:30	142	264
16:45	142	264

Table 5.11 - Raised Floor Electrical Measurements: Second Floor

Time	CRAH Fan Power, kW	UPS-connected IT Load, kW
13:00	290	2492
13:15	289	2517
13:30	290	2507
13:45	288	2531
14:00	290	2519
14:15	290	2497
14:30	289	2518
14:45	289	2506
15:00	290	2499
15:15	289	2531
15:30	290	2526
15:45	290	2525
16:00	289	2529
16:15	289	2519
16:30	289	2531
16:45	289	2538

5.2 Post-Processing of Measurements

With the measured data, several quantities were computed, which will be used later for validation of the thermo-hydraulic model. The electrical input power to the chiller was computed from the amperage measurements by,

$$P_R = \sqrt{3} \times V \times I \times PF, \quad (5.1)$$

where, V is the input voltage, I is the measured current and PF is the power factor for a three-phase electric motor. The power factor is defined as the ratio of the active power to the apparent power. For this study, the power factor was computed from full load amperage (FLA) data of the chiller by,

$$PF = \frac{P_{FLA}}{\sqrt{3}V_{FLA}I_{FLA}}, \quad (5.2)$$

At full load, the chiller required 675 kW of electrical power and had an input voltage and current of 2300 V and 190 amp, resulting in a power factor of $PF = 0.89$.

With the input power to the chiller known, the cooling water flow rate was computed based on a steady-state energy balance of the cooling tower using the supply and return water temperature measurements as,

$$\dot{m}_{CND} = \frac{\dot{Q}_{CT}}{c_p (T_{CT,w}^{return} - T_{CT,w}^{supply})} \quad (5.3)$$

where, the total cooling tower heat removal \dot{Q}_{CT} is the sum of the computed chiller input power and the measured evaporator load.

The heat removal of each CRAH unit \dot{Q}_C in the raised floor space was computed by a steady-state energy balance on the air-side as,

$$\dot{Q}_C = \dot{m}_C c_p (T_{C,a}^{return} - T_{C,a}^{supply}), \quad (5.4)$$

where, \dot{m}_C is the estimated 5.78 kg/s (10,000 CFM @ $\rho = 1.225 \text{ kg/m}^3$) and T_C^{return} , T_C^{supply} are the measured return and supply air temperatures given in Tables 5.5 – 5.8.

Three methods were used to obtain the data center's heat load: the evaporator measurements at the chilled water plant, thermal measurements of the CRAH units, and the electrical measurements of the UPS load and CRAH fans. Figure 5.2 compares these measurement techniques to assure some consistency in the measurements. Excellent agreement was obtained between the three measurement techniques, therefore, providing a good foundation for performing a validation of the thermo-hydraulic model.

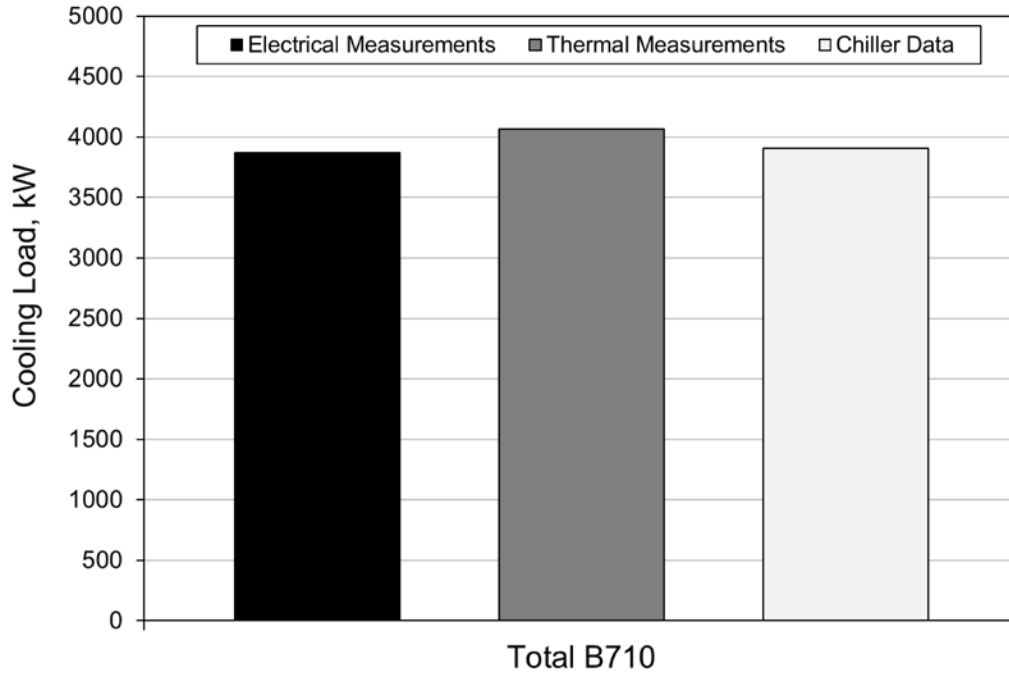


Figure 5.2 - Comparison of Load Measurements

5.3 Thermo-hydraulic Model Validation

Using the thermo-hydraulic model that was developed in Section 2.2, a comprehensive model of B710 and B027 was created. The necessary input data came from manufacturer specifications for each piece of equipment. These data sheets can be found in Appendix A. This section details the validation of the thermo-hydraulic model using the experimental data described in Sections 5.1 and 5.2.

5.3.1 Validation Results

Because there was no time synchronization between the CRAH thermal measurements and the data obtained at the chiller-plant, it would be unfair to expect the thermo-hydraulic model to perform well over the entire 7-hour period. Therefore, the validation was performed in two steps. In the first step, the measured chiller evaporator load was imposed on the chilled water loop and the prediction of the chiller power, condenser supply and return temperatures and the cooling tower heat removal was performed. Figure 5.3 gives the percent difference between the thermo-hydraulic model's predicted and measured cooling tower heat removal. Figures 5.4 and 5.5 show the percent difference between the thermo-hydraulic model's predicted and the measured chiller power and chiller's coefficient of performance, $COP = \dot{Q}_e / P_R$. Based on the measurements, the chiller was operating in the range of $0.54 - 0.56 \text{ kW/ton}$. In all cases, agreement within 5% was obtained.

Figures 5.6 and 5.7 plot the computed cooling tower supply and return temperatures against the measured cooling tower supply and return temperatures. On these figures, a perfect agreement would result in all points falling on the 45° line. The square of the correlation coefficient (R^2) value between the measurements and predictions is given on the figures.

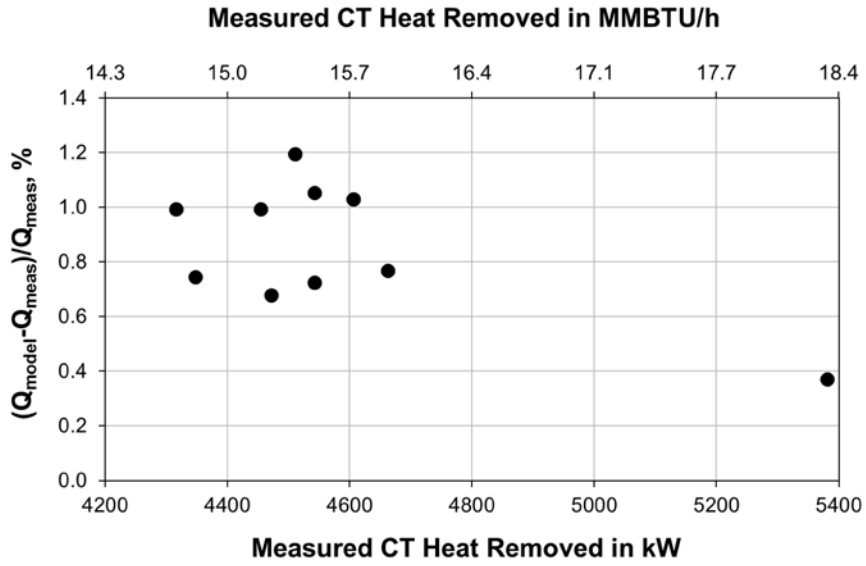


Figure 5.3 - Comparison of Cooling Tower Heat Rejection

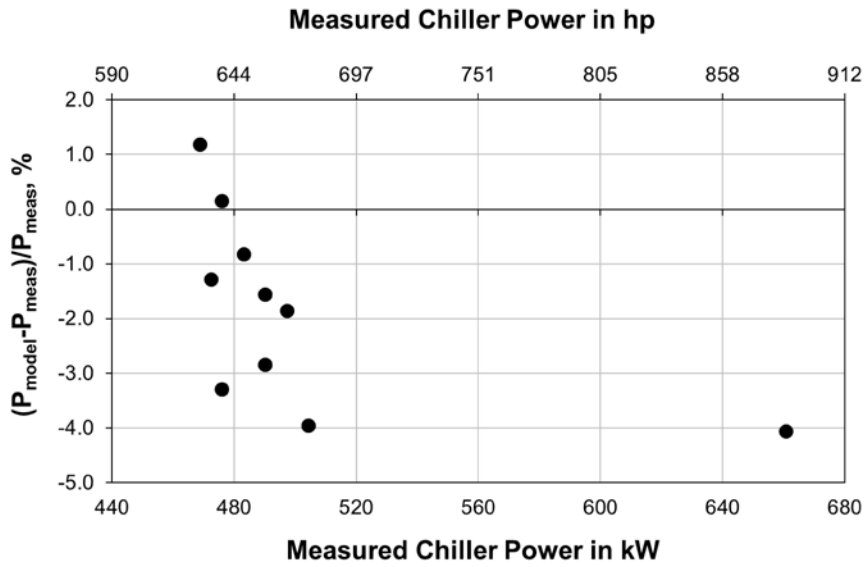


Figure 5.4 - Comparison of Chiller Power

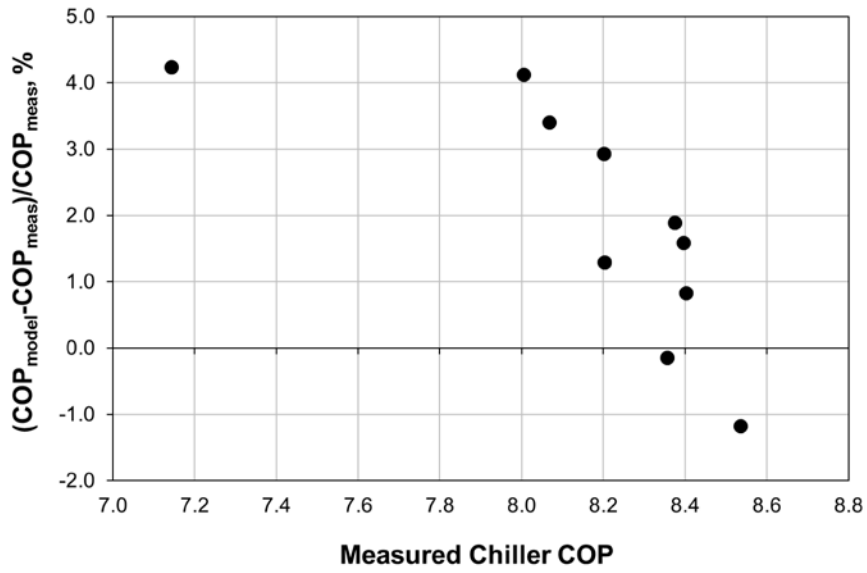


Figure 5.5 - Comparison of Chiller COP

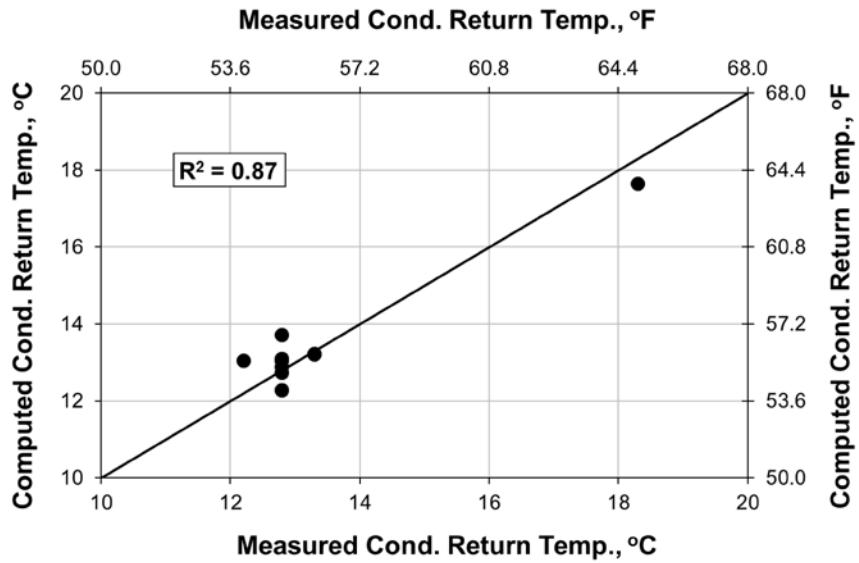


Figure 5.6 - Comparison of Condenser Return Temperature

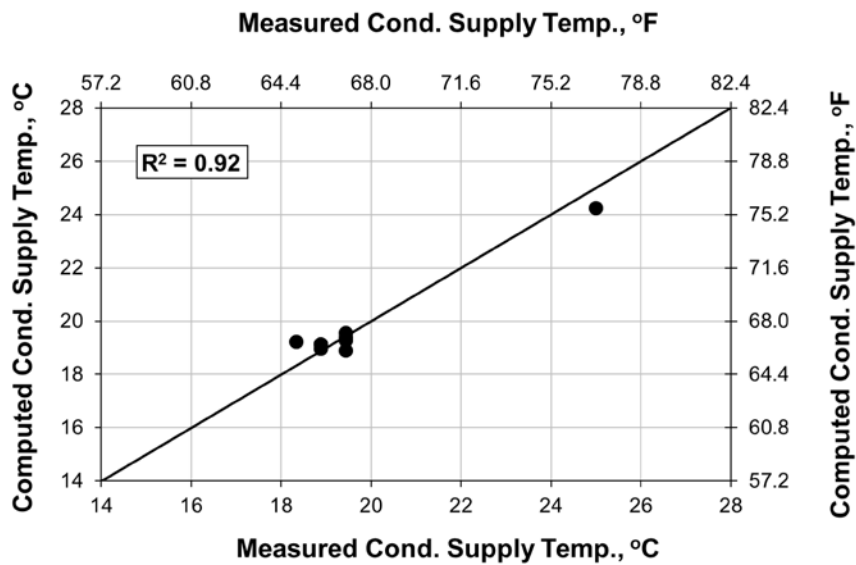


Figure 5.7 - Comparison of Condenser Supply Temperature

In step two of the validation procedure, the measured CRAH return temperatures were used as inputs to the CRAHs in the thermo-hydraulic model. This allows for the prediction of the CRAH-by-CRAH heat removal and air exit temperature. Figures 5.8 and 5.9 plot the computed heat removal and air exit temperature for each of the CRAH against the thermal measurements. Figure 5.10 plots the cumulative heat load for each of the floors in the data center along with the total predicted heat load of B710. The percent difference between the thermo-hydraulic model and the thermal measurements is also given.

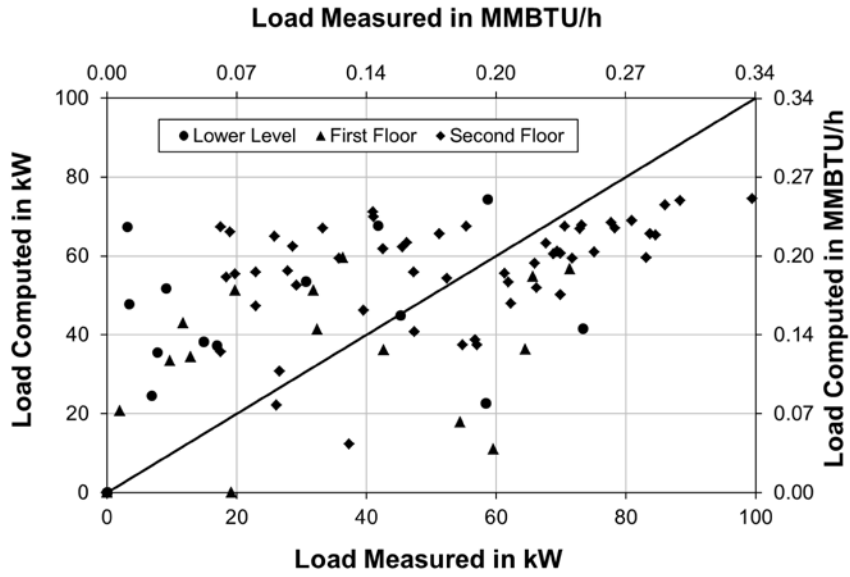


Figure 5.8 - Comparison of CRAH Heat Removal (Individual)

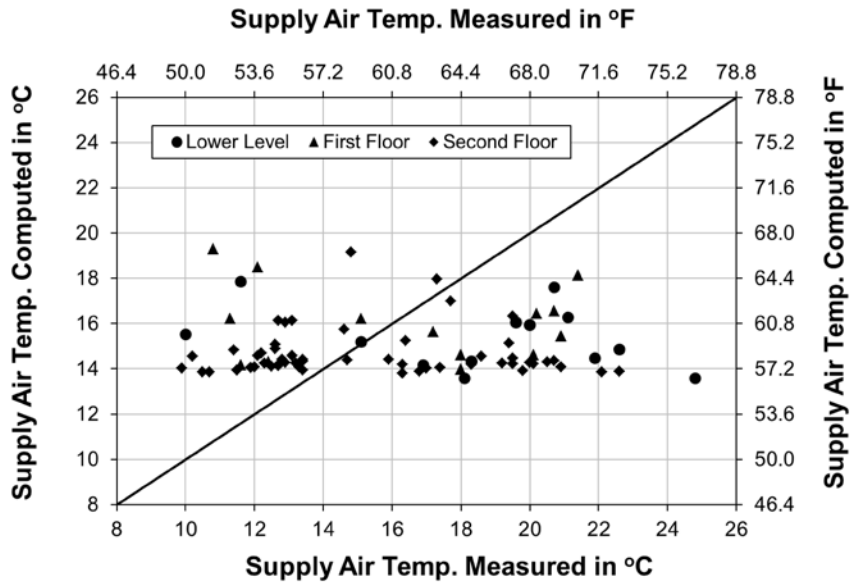


Figure 5.9 - Comparison of CRAH Supply Air Temperature (Individual)

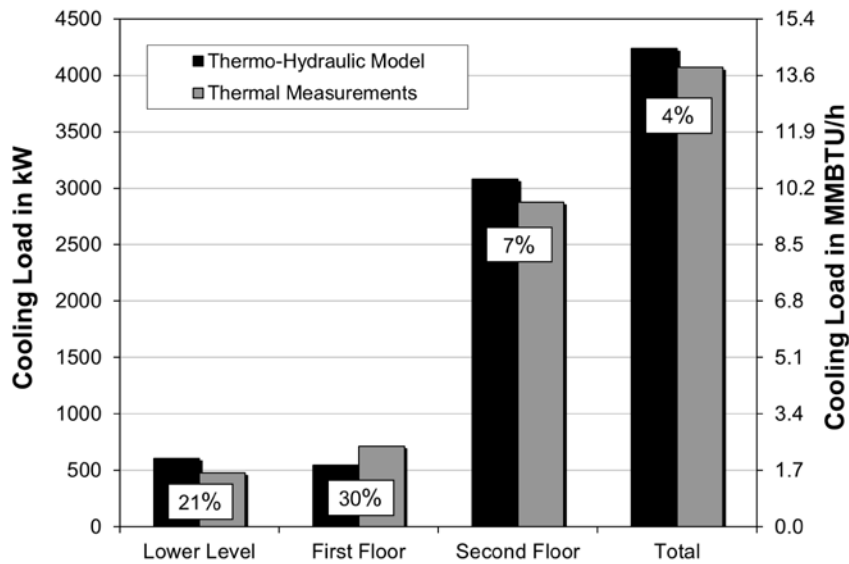


Figure 5.10 - Comparison of CRAH Heat Removal (floor-by-floor)

Figure 5.11 plots the predicted system curve based on the hydraulic model, along with the chilled water pump curve. The first system curve is the overall system curve for the entire data center and chilled water plant facility. Only a single measurement of the chilled water flow rate was available. This measurement showed the pump operating at 2595 GPM of chilled water. The second system curve is the system curve for the data center building only (B710), which does not include the chiller evaporator and the infrastructure connecting B710 and B027. At the predicted operating point of 2051 GPM, the predicted pressure drop was 101.8 kPa (14.8 PSI_G). Again, a single measurement of the pressure drop of the B710 facility was available and showed the B710 pressure drop as 158.6 kPa (23.0 PSI_G). In the context of this study, this is considered good agreement, considering a more accurate estimate of the pressure drop would rely on the accuracy and details of all

the system's minor and major losses, which were not fully available during this study. The system's pipe diameters, lengths and estimates of the minor losses are detailed in Appendix A.

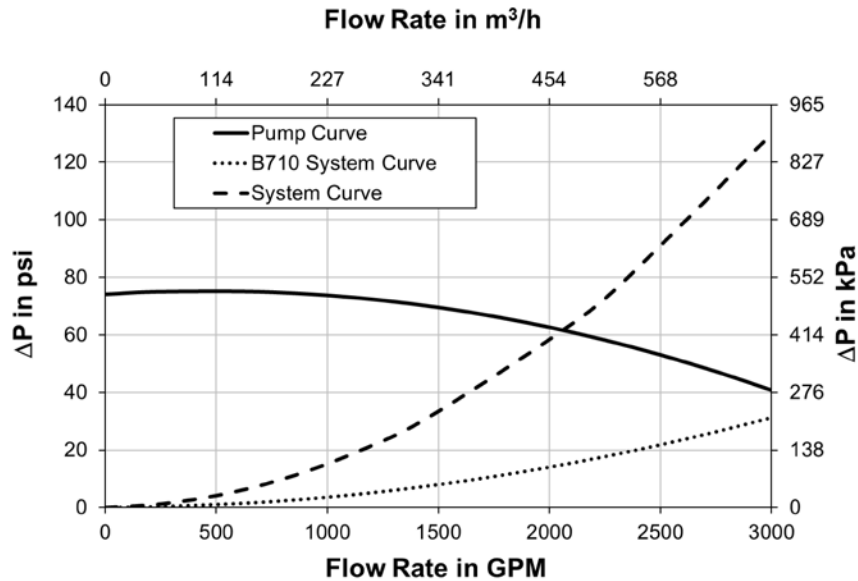


Figure 5.11 - Predicted System Hydraulic Network Characteristics

5.3.2 Discussion of Validation

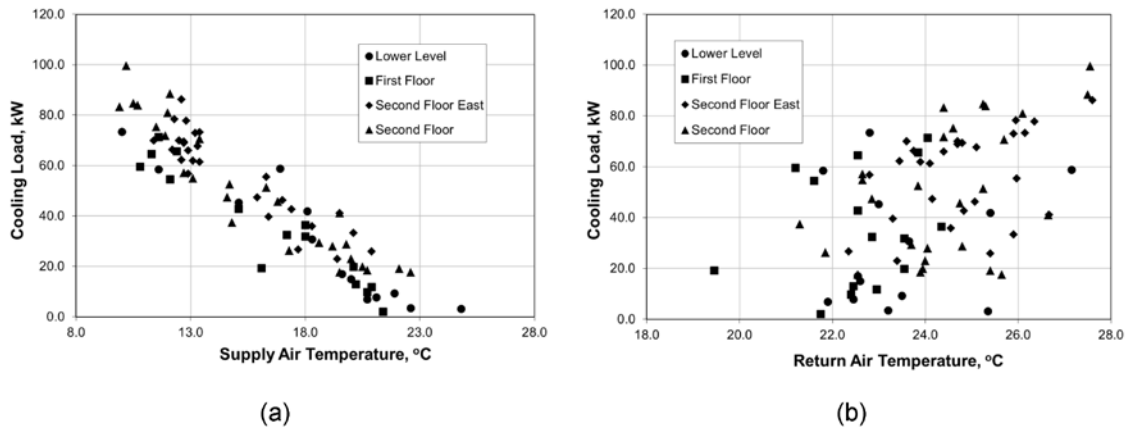
It was shown that good agreement could be obtained between the thermo-hydraulic model and the experimental data in the aggregate; however, slightly inferior agreement was seen when comparing the CRAH model predictions floor-by-floor. There are several explanations for this. First, the measurements done in B710 were extremely complicated due to the data center being operational. Therefore, no intervention was possible and the state of the data center was changing frequently, which may have led to transients that are

not captured in the thermo-hydraulic model. Secondly, the accuracy of the measurements done could be suspect. The airflow rate through the CRAH units was estimated based on previous measurements, but many factors could affect the actual flow, including dirty filters and under-floor plenum pressure. When measuring the temperature of the CRAH units, it is highly likely that temperature gradients existed at both the supply and return. Since the temperature measurements were only done at a limited number of points, this gradient could not be captured.

The accuracy of the temperature measurements, based on the handheld thermocouple, was estimated to be $\pm 1^{\circ}\text{C}$ ($\pm 1.8^{\circ}\text{F}$). Within the scope of this study, these errors are quite large since the temperature difference between CRAH supply and return is approximately 10°C . In addition, a 1°C error could have a significant effect on the prediction of the valve opening in the CRAH model, which effects the predicted heat removal of the CRAH (the valve characteristics are given in Appendix A).

One of the big unknowns during data collection was the exact operation of the CRAH valves. Based on manufacturer input, it was ascertained that the valves were controlled by the return air temperature, which linearly modulated the chilled water valve. However, it could not be determined if the installed units were operated per manufacturer specification. Because of the nature of this open-loop control algorithm, a linear trend was expected between the return air temperature and the cooling load. Figures 5.12a and 5.12b show the cooling load data plotted against supply and return temperature, respectively. A noticeable trend was found in the cooling load vs. supply air temperature

but this was not the case for the cooling load vs. return air temperature. Figures 5.12a and 5.12b further point to incomplete mixing at the return of the CRAH units. It is surmised that a uniform temperature is obtained at the CRAH supply, due to the presence of the filters and heat exchanger pressure drop, which would tend to spread the flow more evenly at the supply.



**Figure 5.12 - CRAH-by-CRAH Cooling Load Measurements
vs. a) Supply Air Temperature and b) Return Air Temperature**

Figures 5.4 and 5.5 also indicate that larger errors are produced for larger power consumption (or lower COP). The trend can be partly attributed to the method by which the overall power consumption was computed. Because the chiller input power included the motor losses, the calculation of the chiller power consumption accounted for changes in motor efficiency as a function of load, with the inclusion of a curve of the motor efficiency vs. motor part-load ratio. However, during data collection no information was available related to the motor specifications and therefore, a generic motor performance

curve was used, in which the peak efficiency of the motor occurred at 85% of the maximum load. Furthermore, the data used to obtain the coefficients in Equation 2.22 was obtained from a manufacturer specification and not based on data collected for the actual chiller in operation. A combination of the above would affect the predicted power consumption, but without more detailed measurements and motor-specific data, the approach adopted was judged adequate. A more important aspect of this work is that the trends are captured correctly, and the model results agreed well with the measured data based on relatively easy to obtain manufacturer data. A number of different generic models for the performance of the chiller, cooling tower, etc. were considered before selecting the ones presented here. All required about the same types of inputs, and the model selected was favored because its input data were readily available.

This is the first study reported in literature to develop and validate a dynamic model that looks at the data center's power consumption from a holistic viewpoint. The inherent difficulty in experimentally validating a model of this nature stems from the fact that data center operation is extremely sensitive to reliability and uptime constraints, which limits any intervention in installing new measurement equipment and conducting controlled experiments. The industry has certainly recognized the need for more experimental measurements in operating data centers. In the future, it is anticipated that higher fidelity data will be available for validating dynamic system simulation models, such as the one developed here. As a starting point, it would be desirable to validate the model over a larger range of operating conditions. However, due to the shear effort required for collecting detailed data in an operating data center, the only data that could be obtained

for validation was collected over a small window during the winter. One of the intentions of this study was to guide development of experimental plans to collect the appropriate data for future validation studies.

5.4 Application of the Thermo-Hydraulic Model

The efficiency of the data center cooling infrastructure is dependent on a number of external conditions, including the IT load, ambient weather conditions, off-design performance of the cooling infrastructure, control strategy, etc. The most widely used metric in the data center industry to evaluate energy performance is the power usage effectiveness (PUE). To this end, the validated thermo-hydraulic model can now be used with reasonable confidence to understand the impact of these external factors.

5.4.1 Effect of Ambient Conditions

To evaluate the effect of ambient conditions, ambient weather data was obtained from the Poughkeepsie, New York airport weather station (Weather Underground, 2009) for the winter day (previously shown in Figure 5.1) in which the validation data in Section 5.1 was collected and for a typical summer day. Figure 5.13 shows the weather data for the typical summer day in Poughkeepsie, NY.

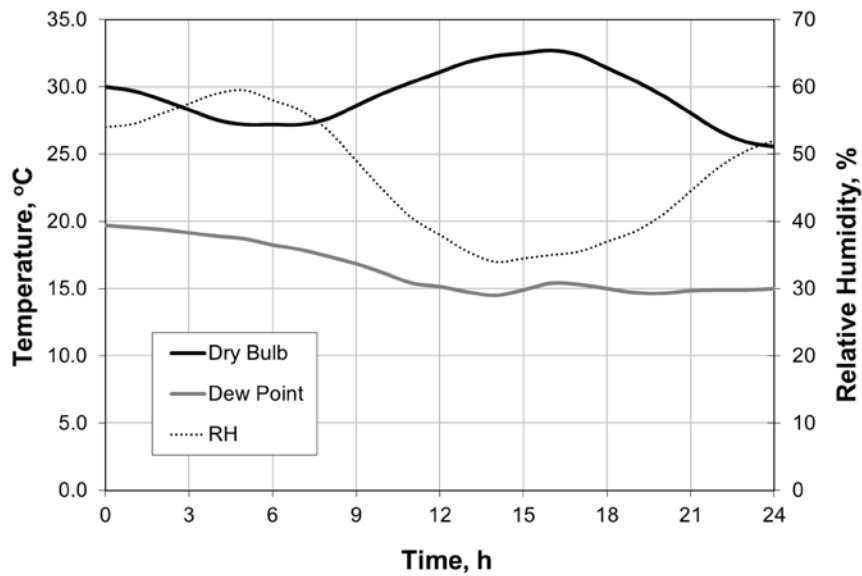


Figure 5.13 - Weather Data for a Summer Day in Poughkeepsie, NY

From the collected data in Section 5.1, a sample operating strategy was extracted, including chilled water set point temperature, cooling tower blower speed and number of pumps, chillers and CRAH units operating. A 1-day simulation was performed using the thermo-hydraulic model to understand the effect of ambient weather conditions. For these simulations, the IT load was based on the thermal measurements reported in Section 5.1. This load was assumed constant over the 24-hour period being simulated. Two hour-by-hour simulations were performed based on the typical winter and summer days shown in Figures 5.1 and 5.13, respectively. The results in Figures 5.14 and 5.15 represent the daily averaged energy usage, for summer and winter, respectively, where the average is computed from the 24 hours of the hour-by-hour simulation. In each figure, the top chart includes the IT load as a slice of the pie and therefore directly represents the DCiE of the

data center. The bottom chart represents just the cooling infrastructure. The results show that a significant difference in the energy consumption of the data center is possible depending on the season of operation. Winter operation of the data center resulted in a PUE of 1.57, whereas for the same IT load, the summer operation PUE was 1.73. This difference was mainly due to the degraded COP of the chiller with higher condenser temperatures. Even in summer operation, B710 is a fairly energy efficient data center when compared to typical U.S. data centers. Much of this efficiency comes from energy efficient chillers, pump variable frequency drives and cooling tower blower variable frequency drives.

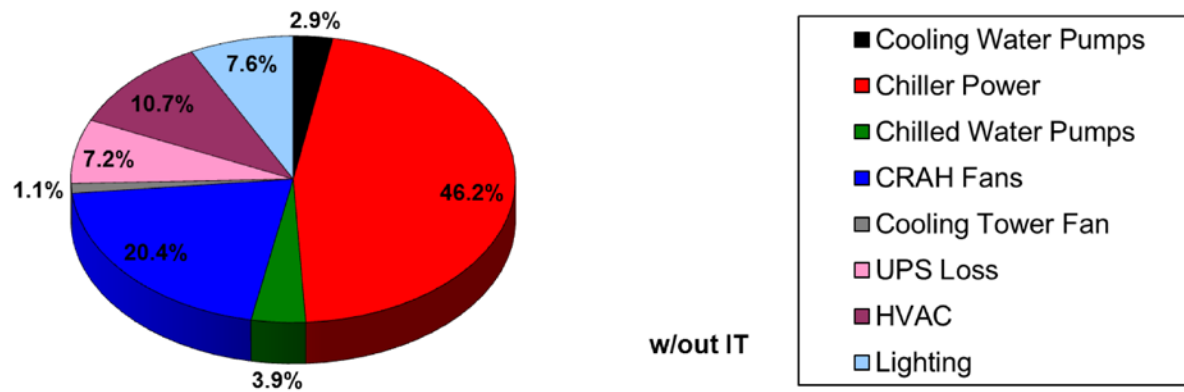
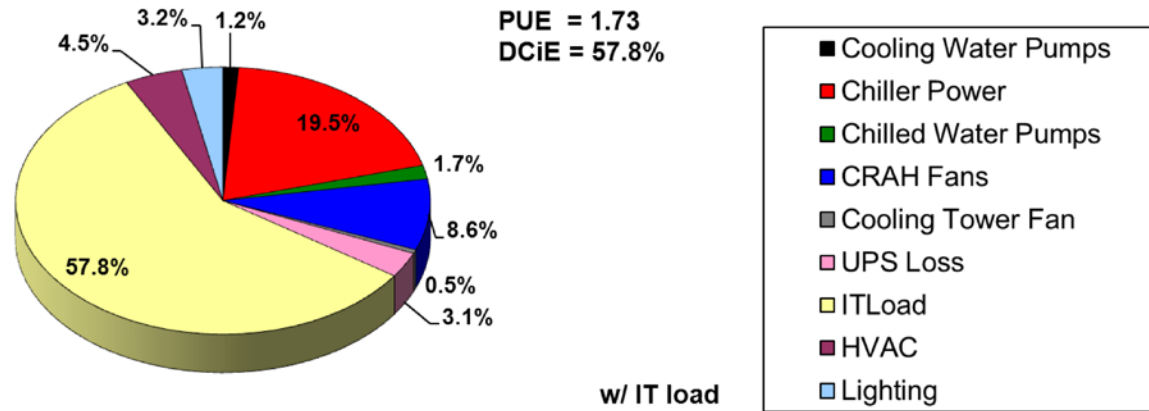


Figure 5.14 – Simulation Results: Data Center Energy for a Summer Day in Poughkeepsie, NY

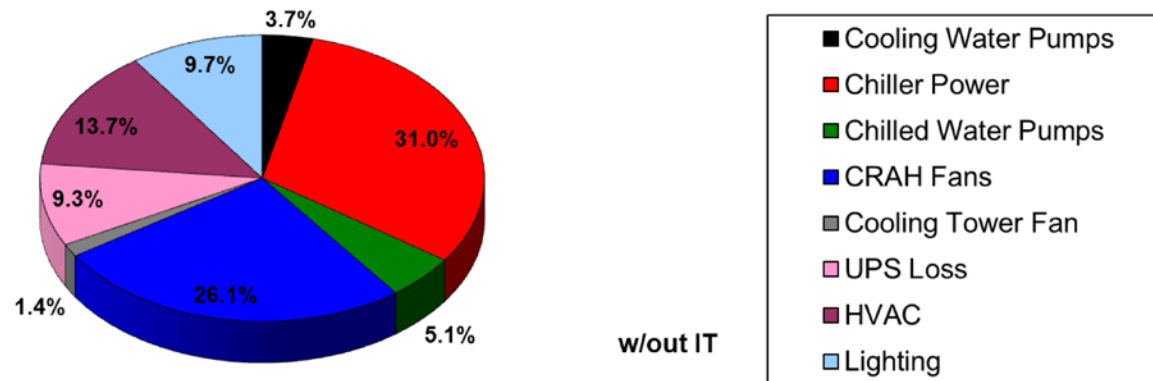
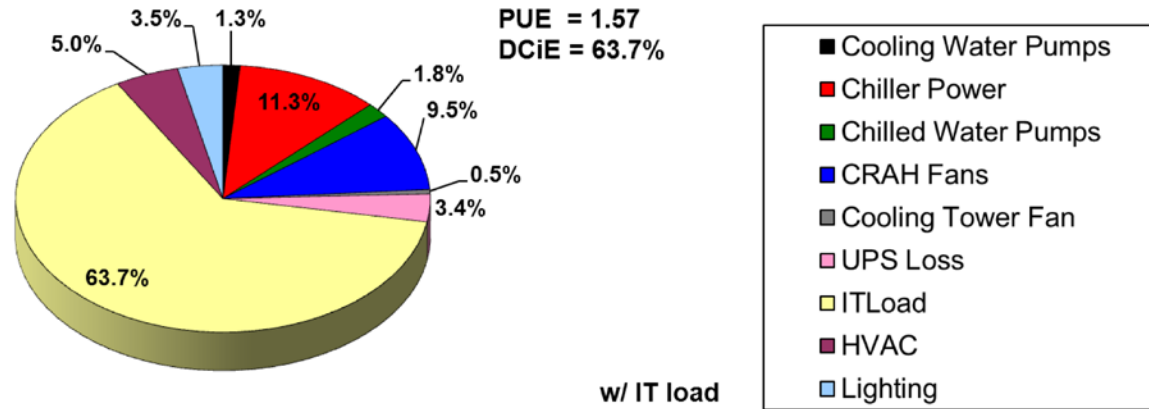


Figure 5.15 – Simulation Results: Data Center Energy for a Winter Day in Poughkeepsie, NY

5.4.2 Impact of Chilled Water Set Point

The ASHRAE Environmental Guidelines for Datacom Equipment expanded the recommended environmental envelope in order to promote energy efficiency in data centers. The new guideline allows for rack inlet air temperatures as high as 27°C (80.6°F). Clearly, with constant speed CRAH units in the data centers, higher air temperatures are realized by increasing the temperature of the chilled water supplied; therefore, saving energy in the chiller by improving its COP.

Simulations were conducted for a range of chilled water temperatures and the effect on the zone's air supply temperature and the facility PUE was investigated. Figure 5.16 show the effect of the chilled water set point on the facility PUE for winter and summer. It can be seen that significant energy savings are possible during the summer months by implementing the ASHRAE recommendations. In this facility, a reduction in PUE from 1.69 to 1.59 is possible by increasing the chilled water set point from the typical 7.2°C (45°F) to 16.0°C (60.8°F). Based on this CRAH heat exchanger model, a 16°C chilled water set point would result in 27°C (80.6°F) supply air from the CRAH units. However, supplying 27°C air to the data center and maintaining inlet server temperature at no greater than 27°C is typically not possible due to hot air recirculation, unless an enclosed aisle is implemented. Although the new environmental guidelines increased the inlet temperature maximum to 27°C, which indeed saves refrigeration energy as just shown, this may not be the optimum strategy for improving the energy efficiency of air-cooled data centers in light of the results obtained in Chapter 3, due to the relatively high power

consumption of the CRAH blowers. However, because the CRAH units in B710 were constant speed, energy savings from modulating the CRAH airflow were not possible.

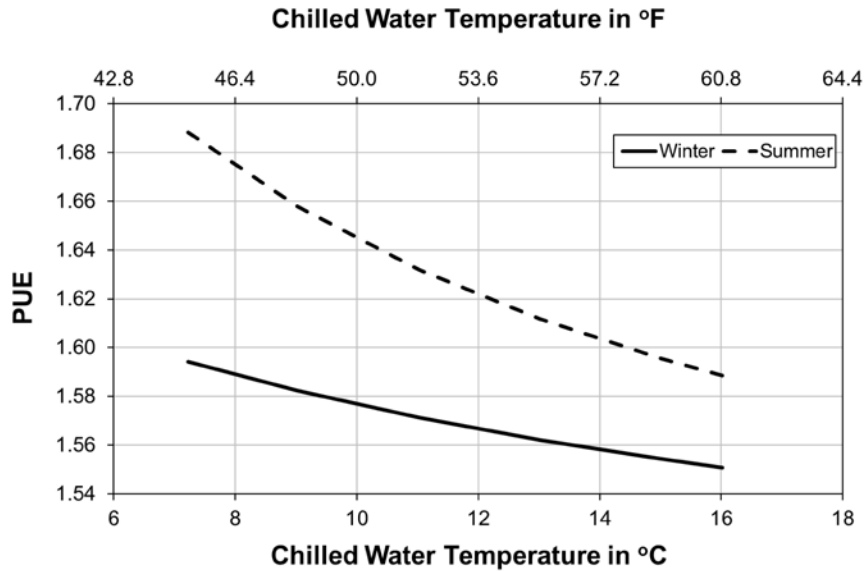


Figure 5.16 - Simulation Results: Changing Chilled Water Temperature

5.5 Chapter Conclusions

A coupled thermodynamic and hydraulic simulation environment was developed, which gives data center designers and engineers the ability to evaluate the energy consumption of various data center configurations at the system and data center levels. As a case study and validation platform, the model was applied to an operating data center located in Poughkeepsie, NY. Exceptional agreement was obtained between the model and the collected data in the aggregate. The models developed here are extremely flexible and

versatile. In their current state, these models are useable and accurate, with input data that is easily obtainable from equipment manufacturer's specifications.

The validated model is further extended to perform several studies that evaluate the data center's energy efficiency at off-design conditions. As designed, B710/B027 it is a fairly energy efficient data center, with a PUE in the range of 1.5; however, based on the dynamic simulations done here, which included the effects of ambient conditions, operating strategy and equipment off-design performance, it was seen that the PUE could be as high as 1.7 during the warmer summer months. Without the dynamic capabilities offered by this model, these changes would go unnoticed with design point simulations. Studies using the B710 data center also showed that further energy efficiency is possible by adopting the new ASHRAE environmental guidelines and increasing the air temperature supplied by the CRAH units. This mode of operation allows for increased chilled water temperature, which both decreases the energy consumption of the chiller and allows for a larger number of hours of economizer use. The annual benefits of an economizer could be easily captured with the thermo-hydraulic model by running yearly simulations and changing the switch over temperature between the chillers and the economizers in the control strategy.

6 Systematic Investigation of Thermally Aware, Energy-based Load Placement in Open Aisle Data Centers

Consider an air-cooled data center consisting of N_R identical racks, each containing n_s identical servers, for a total of $N = N_R \times n_s$ servers. The data center is cooled by a set of N_C identical CRAHs, whose flow rate and supply temperature can be modulated. Each server can exist only in one of two states, either *idle* or *on*. The thermal state (thermal map) of the data center can be represented by a matrix of temperatures $T_{i,j}^{in}$, describing the inlet temperature of each server (or chassis) i in rack j . An operating mode m of the data center is one in which the data center is required to meet a given level of virtualized IT processing load, corresponding to P_m of IT (server or chassis) power usage, which is met by operating m servers at full power such that $P_m = mP_{full} + (N - m)P_{idle}$, in which $(N - m)$ is the number of idle servers. The CRAHs can be controlled to provide \dot{V}_C of cooling air at a temperature T_a such that no server shall receive air at a temperature higher than T^* , the maximum safe server inlet temperature (the redline). There is a large number, K_m , of possible combinations of m operating servers out of the total N servers in the data center, namely, $K_m = N!/m!(N - m)!$. For example, $K_m > 5 \times 10^{20}$ even for a modest $N = 100$ and $m = 80$.

For each combination K of the m operating servers, we would like to find the conjugate pair of CRAH airflow rate and supply air temperature that will minimize the total data

center's cooling infrastructure power consumption. Globally, we would like to find which of the K_m combinations yields the lowest value of the infrastructure power consumption. Thus we seek optimum IT placement scenarios from the K_m possible arrangements, with the pair of conjugate values of CRAH air flow rate and air supply temperature yielding the minimum cooling infrastructure power consumption, without violating the redline temperature constraint anywhere in the data center. Intuitively we would expect that this state would correspond to the highest possible average server inlet temperature compatible with the redline temperature constraint.

However, the sheer effort involved in performing a formal analysis, using computational fluid dynamics or detailed experimental measurements, such as that described above is so prohibitive, rendering it impracticable. Clearly, other methods that are much more computationally efficient must be pursued to derive rules to guide server placement for thermally constrained, energy-optimized IT load distribution.

6.1 Data Center Used in Load Placement Study

The approach undertaken in this dissertation to study thermally aware, energy-based load placement uses a representative hot aisle/cold aisle data center to derive physics-based heuristics that can guide the practical implementation of IT load placement in operating data centers. The layout of the data center used was given in Figure 4.1b. The figure shows $\frac{1}{4}$ of the data center layout with symmetry boundary conditions applied. All of the analysis presented was done for the entire data center, which consists of 64 racks. Based

on a survey of typical high-volume servers available on the market, racks were selected that consume 16 kW at full power. The racks are divided into 4 chassis, which are representative of blade server configurations. Since this work is concerned with load placement, the chassis can be operated in one of two states, an *on* state or an *idle* state. In the *on* state, a chassis consumes 4 kW of IT power and requires 637 m³/h (375 CFM) of airflow. In the *idle* state, a chassis consumes 1 kW of IT power and requires 510 m³/h (300 CFM) of airflow. At this point, a distinction between the useful IT load and the heat load, P_{IT} , of the data center is needed. Figure 6.1 gives a schematic of the distinction. Even when the data center is producing zero useful IT (i.e., all servers are *idle*), electricity is needed in the servers in order to run the fans, hard drives, memory, etc. This residual power is still dissipated as heat in the data center and must be removed by the cooling infrastructure.

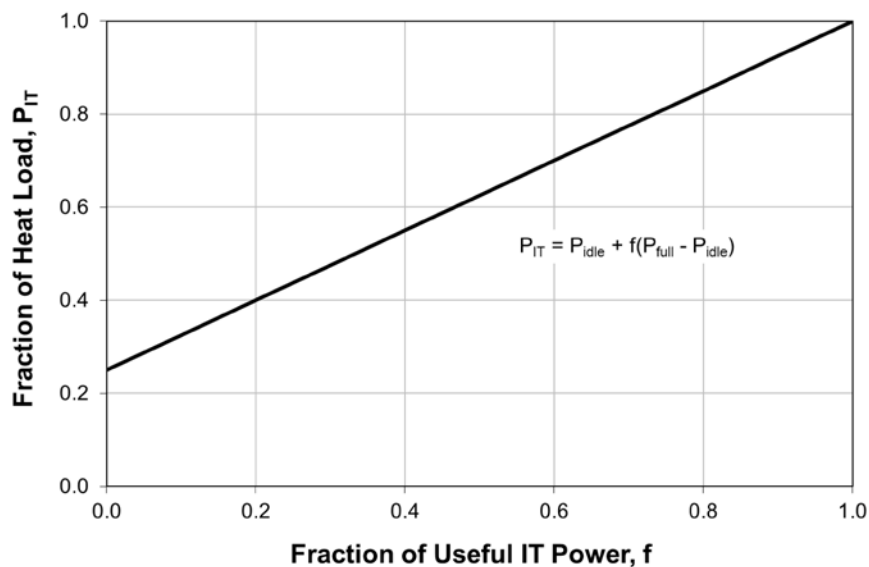


Figure 6.1 - Relationship between Useful IT and Data Center Heat Load

For the remainder of this dissertation, the experimentally validated thermo-hydraulic model is used to evaluate the off-design energy performance of the cooling infrastructure. For this analysis, the infrastructure must be designed to remove the full IT load of 1024 kW plus any heat dissipated by the CRAH fans. The design points of the cooling equipment are given in Table 6.1.

Table 6.1 – Cooling Equipment Design Points for Load Placement Simulations

Item	Quantity	Design Rating
Chillers	1	1758 kW (500 ton), 0.59 kW/ton
Cooling Towers	1	231235 m ³ /h (136100 CFM), 14.7 kW (20 hp) Variable Speed
CRAH	6	40776 m ³ /h (24,000 CFM), 11.5 kW (15.6 hp) Variable speed
Chilled Water Pump	2	341 m ³ /h (1500 GPM), 186 kPa (27 psi), 24.2 kW (33 hp), BEP = 85% Variable speed
Cooling Water Pump	1	454 m ³ /h (2000 GPM), 69 kPa (10 psi), 14.0 kW (19 hp), $\eta = 65\%$

6.2 Workload Placement in Open Aisle Data Centers

Recently, workload placement in data centers has received increased attention in the IT industry. A number of studies have hinted at the potential savings possible by optimizing the placement of the workload amongst the servers in the data center. However, these workload placement algorithms were primarily developed using resource intensive

optimization algorithms that have the disadvantage of requiring a long computational time in order to find the optimum placement of the IT load amongst the operational servers. Furthermore, none of these studies considered the optimization of both the job placement and the cooling infrastructure. In this section, the combined energy optimization of the IT load placement and the cooling infrastructure is considered. The objective of the study is to develop strategies that can reduce the cooling infrastructure's power consumption, while maintaining the reliability of the IT assets. To begin, several load placement scenarios are introduced that are easily implemented in operating data centers. These heuristics are based on either geometrical traits of the data center, a prior physics-based knowledge of the airflow and temperature patterns or measurements that are easily obtainable during operation. For this work, computational fluid dynamics is used to evaluate the airflow and temperature distribution for each of the proposed load placement scenarios. Even though CFD is used in the preliminary development of the heuristics, the intention is not to use CFD and develop new rules for each-and-every data center. Instead, the rules-of-thumb are based on characteristics that are common to most data centers.

6.2.1 Proposed IT Load Placement Scenarios

This section focuses on developing and assessing different IT load placement options in open-aisle, air-cooled data centers. The proposed scenarios are based on either geometrical traits of the data center, a prior physics-based knowledge of the airflow and

temperature patterns or measurements that are easily obtainable during operation. This assessment phase studies four useful IT load levels: 100%, 75%, 50% and 0%.

6.2.1.1 100% Useful IT Power

In the case of 100% useful IT, every chassis must be operated in the *on* state and there is no latitude for moving IT load throughout the data center. Clearly, the 100% useful IT case has the highest heat dissipation; however, it should provide the best match for the cooling infrastructure, which had to be designed to remove the maximum possible heat load.

6.2.1.2 75% Useful IT Power

Figure 6.2 shows the layout for each of the seven proposed scenarios for a useful IT load of 75% (192/256 chassis *on* and 64/256 chassis *idle*). An *x* denotes a chassis that is placed in the *idle* state and a % indicates a chassis where the useful IT has been turned down by 75%. The rationale for choosing the scenarios is as follows:

- Scenario 1: this scenario turns the chassis that are expected to have the coldest inlet temperature to *idle*. It is expected that these chassis are located towards the center of the cold aisle and near the bottom of the rack. There are two underlying reasons why this is a reasonable strategy to consider. First, since these chassis typically have the lowest inlet temperature, they are furthest from reaching their redline constraint, which would be higher if they were *idle*. Secondly, these chassis receive only cold air directly from the perforated tiles. When these chassis

are placed in an *idle* state and consequently their flow rate is reduced, more cold air becomes available for the chassis located in the upper portion of the racks.

- Scenario 2: this scenario turns the chassis in the upper portion of the racks to *idle*. Based on knowledge of the airflow and temperature patterns, hot air recirculation over the top of the aisle causes elevated temperature at the inlet of the IT equipment located in the upper portion of racks. Therefore, since *idle* servers have a higher redline temperature threshold, it is anticipated that a higher CRAH supply air temperature could be used.
- Scenario 3: this scenario turns off the racks located at the end of the cold aisle. Based on knowledge of the airflow and temperature patterns, it is expected that hot air will recirculate around the corner of the aisle, from the hot aisle to the cold aisle, and be entrained into the cold air emanating from the perforated tiles. The hot air mixing that occurs near the end of the aisle causes elevated inlet temperatures to the racks located at the end of the aisle. Therefore, since *idle* servers have a higher redline temperature threshold, it is anticipated that a higher CRAH supply air temperature could be used.
- Scenario 4: this scenario is a combination of Scenarios 2 and 3. Like those scenarios, the intent of Scenario 4 is to increase the temperature of the supply air by switching to *idle* the chassis that are expected to suffer most from hot air recirculation, which are the upper corner chassis and outer racks.
- Scenario 5: this scenario randomly distributes the IT load amongst the servers. The servers placed in the *idle* state were selected based on a random number generator.

- Scenario 6: this scenario distributed the required IT load evenly amongst all chassis, such that, $P_{IT} = P_{IDLE} + f(P_{ON} - P_{IDLE})$. The chassis' airflow rate is also varied linearly between the flow at the *on* state and the flow at the *idle* state.
- Scenario 7: this scenario uses temperature measurements at the inlet of each chassis to turn the servers with the highest inlet temperature to *idle*. Most IT equipment available on the market today has one or more built in thermistors, which can communicate with central data acquisition and building management systems. The proposed algorithm shuts down chassis, starting with the hottest, until the required IT load is reached. Again, the hypothesis being that since *idle* servers have a higher redline temperature threshold, it is anticipated that a higher CRAH supply air temperature could be used. In practice, the workload at a future instance in time would be determined from the current state. Therefore, for this analysis, the chassis that are shut to *idle* are determined from the optimized 100% useful IT load scenario, which is described in Section 6.2.3.2.

6.2.1.3 50% Useful IT Power

Figure 6.3 shows the layout for each of the seven proposed scenarios for a useful IT load of 50% (128/256 chassis *on* and 128/256 chassis *idle*). An *x* denotes a chassis that is placed in the *idle* state and a % indicates a chassis where the useful IT has been turned down by 50%. The rationale for choosing the scenarios follows that given for the 75% useful IT load case.

6.2.1.4 0% Useful IT Power

In the 0% useful IT load case, no useful computations are being performed in the data center; however, all chassis are operated in the *idle* state. Therefore, the cooling infrastructure must still remove the idle power heat dissipation that is necessary to operate the server fans, hard disk drives, memory, etc. Even though this may not be a practical scenario for operational data centers, it is still of academic interest insomuch as it results the most off-design operation of the cooling infrastructure.

	1	2	3	4	5	6	7	8
4								
3			x	x	x	x		
2		x	x	x	x	x	x	
1		x	x	x	x	x	x	

	1	2	3	4	5	6	7	8
4	x	x	x	x	x	x	x	x
3	x	x	x	x	x	x	x	x
2								
1								

	9	10	11	12	13	14	15	16
1		x	x	x	x	x	x	
2		x	x	x	x	x	x	
3			x	x	x	x		
4								

	9	10	11	12	13	14	15	16
1								
2								
3	x	x	x	x	x	x	x	x
4	x	x	x	x	x	x	x	x

Scenario 1

Scenario 2

	1	2	3	4	5	6	7	8
4	x	x					x	x
3	x	x					x	x
2	x	x					x	x
1	x	x					x	x

	1	2	3	4	5	6	7	8
4	x	x	x	x	x	x	x	x
3	x	x					x	x
2	x							x
1	x							x

	9	10	11	12	13	14	15	16
1	x	x					x	x
2	x	x					x	x
3	x	x					x	x
4	x	x					x	x

	9	10	11	12	13	14	15	16
1	x							x
2	x							x
3	x	x					x	x
4	x	x	x	x	x	x	x	x

Scenario 3

Scenario 4

	1	2	3	4	5	6	7	8
4	x	x			x	x		
3	x	x		x			x	
2		x				x	x	
1		x	x		x	x		x

	1	2	3	4	5	6	7	8
4	%	%	%	%	%	%	%	%
3	%	%	%	%	%	%	%	%
2	%	%	%	%	%	%	%	%
1	%	%	%	%	%	%	%	%

	9	10	11	12	13	14	15	16	
1		x	x			x	x		x
2		x	x			x	x		x
3	x	x		x			x		
4	x	x			x	x			

	9	10	11	12	13	14	15	16
1	%	%	%	%	%	%	%	%
2	%	%	%	%	%	%	%	%
3	%	%	%	%	%	%	%	%
4	%	%	%	%	%	%	%	%

Scenario 5

Scenario 6

	1	2	3	4	5	6	7	8
4	x	x	x	x	x	x	x	x
3	x				x	x	x	x
2	x							x
1	x							x

	9	10	11	12	13	14	15	16
1	x							x
2	x							x
3	x						x	x
4	x	x	x	x	x	x	x	x

Scenario 7

Figure 6.2 - Proposed Load Placement Scenarios for 75% Useful IT

	1	2	3	4	5	6	7	8
4								
3								
2			X	X	X	X		
1			X	X	X	X		

	1	2	3	4	5	6	7	8
4	X	X	X	X	X	X	X	X
3								
2								
1								

	9	10	11	12	13	14	15	16
1			X	X	X	X		
2			X	X	X	X		
3								
4								

	9	10	11	12	13	14	15	16
1								
2								
3								
4	X	X	X	X	X	X	X	X

Scenario 1

Scenario 2

	1	2	3	4	5	6	7	8
4	X							X
3	X							X
2	X							X
1	X							X

	1	2	3	4	5	6	7	8
4	X	X					X	X
3	X	X					X	X
2								
1								

	9	10	11	12	13	14	15	16
1	X							X
2	X							X
3	X							X
4	X							X

	9	10	11	12	13	14	15	16
1								
2								
3	X	X					X	X
4	X	X					X	X

Scenario 3

Scenario 4

	1	2	3	4	5	6	7	8
4	X							X
3	X		X					X
2	X							
1						X		X

	1	2	3	4	5	6	7	8
4	%	%	%	%	%	%	%	%
3	%	%	%	%	%	%	%	%
2	%	%	%	%	%	%	%	%
1	%	%	%	%	%	%	%	%

	9	10	11	12	13	14	15	16
1						X		X
2	X							
3	X		X				X	
4	X						X	

	9	10	11	12	13	14	15	16
1	%	%	%	%	%	%	%	%
2	%	%	%	%	%	%	%	%
3	%	%	%	%	%	%	%	%
4	%	%	%	%	%	%	%	%

Scenario 5

Scenario 6

	1	2	3	4	5	6	7	8
4	X	X	X	X	X	X	X	X
3								X
2								X
1								

	9	10	11	12	13	14	15	16
1								
2								X
3	X							X
4	X					X		X

Scenario 7

Figure 6.3 - Proposed Load Placement Scenarios for 50% Useful IT

6.2.2 Optimization Procedure

Each of the proposed scenarios was tested for a range of tile airflow fraction, ψ_T , and for a leakage fraction, $\Lambda = 0.25$. The procedure for performing the optimization is shown in Figure 6.4. For each IT load scenario, IT power level, and ψ_T , a computational fluid dynamics simulation is run to determine the airflow and temperature fields in the data center. Two changes in the data center raised floor lead to reductions in the cooling infrastructure's power consumption, namely, changes in the CRAH airflow (i.e., via changing ψ_T) and changes in the supply air temperature (i.e., via changes in the required chilled water temperature). However, regardless of what changes are made, the reliability of the IT equipment cannot be sacrificed. Therefore, the following constraints are applied to each case,

$$\begin{aligned} T_{i,j}^{in,on} &\leq 27^\circ C \\ T_{i,j}^{in,idle} &\leq 32^\circ C \end{aligned} \tag{6.1}$$

A simplifying assumption is necessary in order to keep the number of necessary CFD simulations reasonable. The results of Section 4.5 indicated that for data centers with a low Ar , a δT change in the supply air temperature would result in a δT change in the temperature field everywhere in the data center. To this end, all CFD simulations are performed at a supply air temperature of $T_a = 14.0^\circ C$. However, for a given ψ_T , there is no guarantee that a $14^\circ C$ supply temperature will satisfy the constraints given by

Equation 6.1. Therefore, the supply air temperature (and as a consequence all temperatures in the data center) is adjusted by,

$$\delta T_a = \min \left[\max \left(T_{i,j}^{in,on} - 27^\circ C \right), \max \left(T_{i,j}^{in,idle} - 32^\circ C \right) \right]. \quad (6.2)$$

This step is done in post-processing and therefore a new CFD simulation is not performed for each δT adjustment. Again, this simplification is only justified for data centers with a “small” Ar and will be verified in Section 6.4.2. Once the required supply air temperature is computed, the focus moves to optimizing the cooling infrastructure with the thermo-hydraulic model.

The optimum operation of the cooling infrastructure must focus on finding the combination of CRAH airflow rate, chilled water temperature and chilled water flow rate that meets the IT demands of the data center, at the minimum power consumption. A schematic of the data center’s cooling and power infrastructure was given in Figure 1.5. The link between the data center’s raised floor space and the chilled water plant is the CRAH heat exchanger. The optimization of the infrastructure is governed by the well-known NTU- ε heat exchanger relations (Incropera and DeWitt, 2005). For a CRAH unit with chilled water flowing through the heat exchanger tubes and warm air passing over the heat exchanger coil, these equations are given as,

$$\dot{Q}_C = C_{\min} \varepsilon \left(T_{C,a}^{return} - T_{C,w}^{supply} \right) \quad (6.3)$$

$$\varepsilon = f\left(NTU = \frac{UA}{C_{\min}}, C_r\right) \quad (6.4)$$

in which ε is the heat exchanger effectiveness, $T_{C,a}^{return}$ and $T_{C,w}^{sup ply}$ are the air and water inlet temperatures, respectively, UA is the heat exchanger overall conductance, and C is the heat capacity rate defined as the mass flow rate times the specific heat. The heat capacity ratio C_r is defined as,

$$C_r = \frac{C_{\min}}{C_{\max}} = \frac{\min(c_{p,w}\dot{m}_{CHW}, c_{p,a}\dot{m}_{C,a})}{\max(c_{p,w}\dot{m}_{CHW}, c_{p,a}\dot{m}_{C,a})}. \quad (6.5)$$

In addition to the above equations, the first law of thermodynamics provides the following relationships,

$$T_{C,a}^{sup ply} = T_{C,a}^{return} - \frac{\dot{Q}_C}{\dot{m}_{C,a}c_{p,a}} \quad (6.6)$$

$$T_{C,w}^{return} = T_{C,w}^{sup ply} + \frac{\dot{Q}_C}{\dot{m}_{CHW}c_{p,w}}. \quad (6.7)$$

For this analysis, the CRAH airflow rate is used parametrically, assuming the CRAH units have variable speed drives that control the fan RPM (i.e., flow rate). Similarly, the chilled water flow rate is obtained by selecting a pump speed and having the hydraulic

model compute the flow distribution. Therefore, for a given IT power, the remainder of the heat exchanger quantities can be uniquely determined by choosing one other parameter to evaluate parametrically. In this instance, the chilled water temperature is used and therefore, the CRAH air exit temperature is computed. The heat exchanger conductance was obtained from manufacturer specifications, with typical values of UA ranging from 14.0 kW/K – 25.0 kW/K.

The preceding analysis showed that without an appropriate control algorithm, the exit air from the CRAH heat exchanger remains uncontrolled, and governed by the heat exchanger NTU- ϵ equations. However, many commercially available CRAH units control the exit air temperature by bypassing a fraction of the chilled water around the heat exchanger coil using a three-way valve. A similar algorithm was implemented into the CRAH heat exchanger model. Therefore, for a given air discharge temperature, the model determines the amount of water that must be bypassed around the coil to meet the discharge air temperature constraint. If there is not a sufficient water flow rate at the given chilled water temperature, the model computes the minimum possible temperature, but provides an error since this is not a feasible operating state. The bypassed water is mixed with the fraction of the water that goes through the coil before exiting the CRAH unit. This model requires the iterative solution of Equations 6.3 – 6.7. This model of the CRAH heat exchanger was coupled into the simulation environment discussed in Section 2.2, which already has the off-design performance models of the entire cooling infrastructure. The total power consumption of the cooling infrastructure is then given by,

$$P_C = P_{F,CT} + P_{P,CND} + P_{P,CHW} + P_{F,C} + P_R, \quad (6.8)$$

where, $P_{F,CT}$ is the power consumption of the cooling tower fan (variable speed based on ambient wet-bulb temperature), $P_{P,CND}$ is the power consumption of the condenser pump (constant speed), $P_{P,CHW}$ is the power consumption of the chiller water pump (variable speed), $P_{F,C}$ is the power consumption of CRAH fans (variable speed), and P_R is the chiller power consumption.

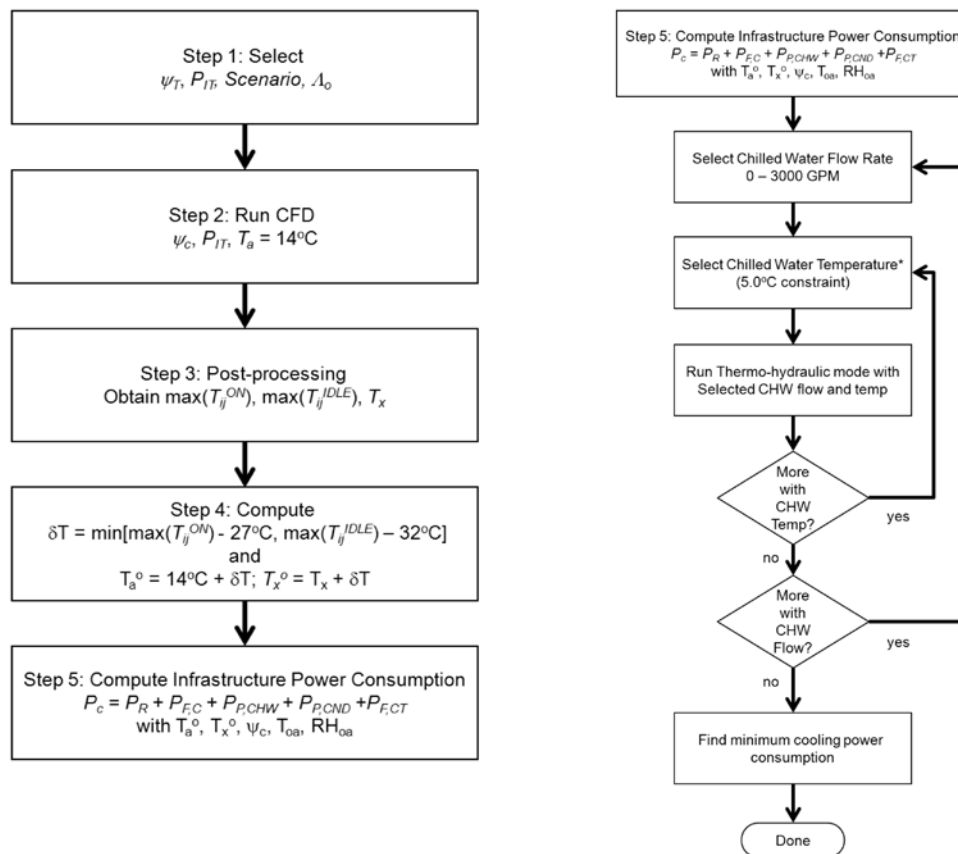


Figure 6.4 - Flow Chart for Infrastructure Optimization

6.2.3 Computational Fluid Dynamics Simulations

With an optimization methodology in place, the next step is to perform the analysis to determine the effectiveness of each of the proposed IT load placement scenarios. For the analysis, 141 computational fluid dynamics simulations were done, for a range of useful IT, load placement scenario and ψ_T . The results of these analyses are given in Tables 6.2 – 6.18, along with the computed supply air δT and the optimum supply air temperature, T_a^* , that meets the chassis' inlet temperature constraint given by Equation 6.1.

Table 6.2 - CFD Results for 100% Useful IT Load

ψ_T	Ar	\dot{m}_C , kg/s	$\max(T_{i,j}^{in,on})$, °C	$\max(T_{i,j}^{in,idle})$, °C	δT , °C	T_a^* , °C
1.0	0.75	11.60	27.81		-0.81	13.19
0.95	0.83	11.02	29.72		-2.72	11.28
0.90	0.93	10.44	31.58		-4.58	9.42
0.85	1.04	9.86	33.09		-6.09	7.91
0.80	1.17	9.28	34.05		-7.05	6.95
0.75	1.33	8.70	34.12		-7.12	6.88
0.70	1.53	8.12	38.54		-11.54	2.46

Table 6.3 - CFD Results for Scenario 1 with 75% Useful IT Load

ψ_T	Ar	\dot{m}_C , kg/s	$\max(T_{i,j}^{in,on})$, °C	$\max(T_{i,j}^{in,idle})$, °C	δT , °C	T_a^* , °C
1.00	0.65	11.02	26.05	14.00	0.95	14.95
0.95	0.72	10.47	27.21	14.00	-0.21	13.79
0.90	0.80	9.92	28.19	14.00	-1.19	12.81
0.85	0.89	9.37	28.93	14.00	-1.93	12.07
0.80	1.01	8.82	29.73	14.01	-2.73	11.27
0.75	1.15	8.27	30.60	14.06	-3.60	10.40
0.70	1.32	7.72	31.48	14.21	-4.48	9.52

Table 6.4 - CFD Results for Scenario 2 with 75% Useful IT Load

ψ_T	Ar	\dot{m}_C , kg/s	$\max(T_{i,j}^{in,on})$, °C	$\max(T_{i,j}^{in,idle})$, °C	δT , °C	T_a^* , °C
1.00	0.65	11.02	26.90	22.23	4.77	18.77
0.95	0.72	10.47	23.29	28.51	3.49	17.49
0.90	0.80	9.92	24.45	30.18	1.82	15.82
0.85	0.89	9.37	25.30	31.62	0.38	14.38
0.80	1.01	8.82	26.14	32.56	-0.56	13.44
0.75	1.15	8.27	27.62	33.45	-1.45	12.55
0.70	1.32	7.72	28.91	33.64	-1.91	12.09

Table 6.5 - CFD Results for Scenario 3 with 75% Useful IT Load

ψ_T	Ar	\dot{m}_C , kg/s	$\max(T_{i,j}^{in,on})$, °C	$\max(T_{i,j}^{in,idle})$, °C	δT , °C	T_a^* , °C
1.00	0.65	11.02	17.98	23.85	8.15	22.15
0.95	0.72	10.47	20.78	25.85	6.15	20.15
0.90	0.80	9.92	23.17	27.40	3.83	17.83
0.85	0.89	9.37	25.34	28.76	1.66	15.66
0.80	1.01	8.82	27.23	29.99	-0.23	13.77
0.75	1.15	8.27	29.86	30.99	-2.86	11.14
0.70	1.32	7.72	32.02	32.13	-5.02	8.98

Table 6.6 - CFD Results for Scenario 4 with 75% Useful IT Load

ψ_T	Ar	\dot{m}_C , kg/s	$\max(T_{i,j}^{in,on})$, °C	$\max(T_{i,j}^{in,idle})$, °C	δT , °C	T_a^* , °C
1.00	0.65	11.02	19.62	24.41	7.38	21.38
0.95	0.72	10.47	20.78	25.85	6.15	20.15
0.90	0.80	9.92	21.30	27.19	4.81	18.81
0.85	0.89	9.37	23.96	28.46	3.04	17.04
0.80	1.01	8.82	26.06	29.48	0.94	14.94
0.75	1.15	8.27	27.92	30.53	-0.92	13.08
0.70	1.32	7.72	29.76	32.00	-2.76	11.24

Table 6.7 - CFD Results for Scenario 5 with 75% Useful IT Load

ψ_T	Ar	\dot{m}_C , kg/s	$\max(T_{i,j}^{in,on})$, °C	$\max(T_{i,j}^{in,idle})$, °C	δT , °C	T_a^* , °C
1.00	0.65	11.02	25.67	19.32	1.33	15.33
0.95	0.72	10.47	27.20	21.66	-0.20	13.80
0.90	0.80	9.92	28.97	24.48	-1.97	12.03
0.85	0.89	9.37	30.49	26.91	-3.49	10.51
0.80	1.01	8.82	31.57	28.84	-4.57	9.43
0.75	1.15	8.27	32.78	30.65	-5.78	8.22
0.70	1.32	7.72	34.01	32.34	-7.01	6.99

Table 6.8 - CFD Results for Scenario 6 with 75% Useful IT Load

ψ_T	Ar	\dot{m}_C , kg/s	$\max(T_{i,j}^{in,on})$, °C	$\max(T_{i,j}^{in,idle})$, °C	δT , °C	T_a^* , °C
1.00	0.65	11.02	25.88		1.12	15.12
0.95	0.72	10.47	27.45		-0.45	13.55
0.90	0.80	9.92	29.04		-2.04	11.96
0.85	0.89	9.37	30.40		-3.40	10.60
0.80	1.01	8.82	31.30		-4.30	9.70
0.75	1.15	8.27	31.36		-4.36	9.64
0.70	1.32	7.72	32.06		-5.06	8.94

Table 6.9 - CFD Results for Scenario 7 with 75% Useful IT Load

ψ_T	Ar	\dot{m}_C , kg/s	$\max(T_{i,j}^{in,on})$, °C	$\max(T_{i,j}^{in,idle})$, °C	δT , °C	T_a^* , °C
1.00	0.65	11.02	17.49	24.05	7.95	21.95
0.95	0.72	10.47	17.98	25.58	6.42	20.42
0.90	0.80	9.92	18.88	26.80	5.20	19.20
0.85	0.89	9.37	19.70	27.97	4.03	18.03
0.80	1.01	8.82	20.67	29.09	2.91	16.91
0.75	1.15	8.27	23.20	30.24	1.76	15.76
0.70	1.32	7.72	26.12	31.37	0.63	14.63

Table 6.10 - CFD Results for Scenario 1 with 50% Useful IT Load

ψ_T	Ar	\dot{m}_C , kg/s	$\max(T_{i,j}^{in,on})$, °C	$\max(T_{i,j}^{in,idle})$, °C	δT , °C	T_a^* , °C
1.00	0.54	10.44	23.67	14.08	3.33	17.33
0.95	0.60	9.92	24.42	14.08	2.58	16.58
0.90	0.67	9.40	24.99	14.15	2.01	16.01
0.85	0.75	8.88	25.60	14.60	1.40	15.40
0.80	0.85	8.35	26.29	15.82	0.71	14.71
0.75	0.97	7.83	26.88	17.80	0.12	14.12
0.70	1.11	7.31	28.51	21.01	-1.51	12.49
0.60	1.51	6.27	34.63	26.39	-7.63	6.37
0.50	2.17	5.22	37.84	30.09	-10.84	3.16

Table 6.11 - CFD Results for Scenario 2 with 50% Useful IT Load

ψ_T	Ar	\dot{m}_C , kg/s	$\max(T_{i,j}^{in,on})$, °C	$\max(T_{i,j}^{in,idle})$, °C	δT , °C	T_a^* , °C
1.00	0.54	10.44	17.60	24.90	7.10	21.10
0.95	0.60	9.92	17.34	25.99	6.01	20.01
0.90	0.67	9.40	17.24	27.00	5.00	19.00
0.85	0.75	8.88	17.28	27.95	4.05	18.05
0.80	0.85	8.35	17.59	28.88	3.12	17.12
0.75	0.97	7.83	18.17	29.86	2.14	16.14
0.70	1.11	7.31	18.93	30.97	1.03	15.03
0.60	1.51	6.27	21.13	33.63	-1.63	12.37
0.50	2.17	5.22	26.28	37.37	-5.37	8.63

Table 6.12 - CFD Results for Scenario 3 with 50% Useful IT Load

ψ_T	Ar	\dot{m}_C , kg/s	$\max(T_{i,j}^{in,on})$, °C	$\max(T_{i,j}^{in,idle})$, °C	δT , °C	T_a^* , °C
1.00	0.54	10.44	16.77	21.25	10.23	24.23
0.95	0.60	9.92	18.27	21.99	8.73	22.73
0.90	0.67	9.40	20.27	22.71	6.73	20.73
0.85	0.75	8.88	22.56	23.45	4.44	18.44
0.80	0.85	8.35	24.80	25.78	2.20	16.20
0.75	0.97	7.83	27.37	30.64	-0.37	13.63
0.70	1.11	7.31	33.25	34.66	-6.25	7.75
0.60	1.51	6.27	35.01	35.05	-8.01	5.99
0.50	2.17	5.22	39.79	39.36	-12.79	1.21

Table 6.13 - CFD Results for Scenario 4 with 50% Useful IT Load

ψ_T	Ar	\dot{m}_C , kg/s	$\max(T_{i,j}^{in,on})$, °C	$\max(T_{i,j}^{in,idle})$, °C	δT , °C	T_a^* , °C
1.00	0.54	10.44	14.17	22.25	9.75	23.75
0.95	0.60	9.92	14.24	23.72	8.28	22.28
0.90	0.67	9.40	14.97	25.40	6.60	20.60
0.85	0.75	8.88	16.17	26.47	5.53	19.53
0.80	0.85	8.35	17.62	27.36	4.64	18.64
0.75	0.97	7.83	19.36	28.23	3.77	17.77
0.70	1.11	7.31	21.04	29.15	2.85	16.85
0.60	1.51	6.27	26.44	32.44	-0.44	13.56
0.50	2.17	5.22	35.61	38.62	-8.61	5.39

Table 6.14 - CFD Results for Scenario 5 with 50% Useful IT Load

ψ_T	Ar	\dot{m}_C , kg/s	$\max(T_{i,j}^{in,on})$, °C	$\max(T_{i,j}^{in,idle})$, °C	δT , °C	T_a^* , °C
1.00	0.54	10.44	24.73	17.20	2.27	16.27
0.95	0.60	9.92	26.05	20.09	0.95	14.95
0.90	0.67	9.40	27.25	22.84	-0.25	13.75
0.85	0.75	8.88	28.31	25.23	-1.31	12.69
0.80	0.85	8.35	28.80	26.97	-1.80	12.20
0.75	0.97	7.83	29.43	28.57	-2.43	11.57
0.70	1.11	7.31	30.86	30.20	-3.86	10.14
0.60	1.51	6.27	34.32	33.89	-7.32	6.68
0.50	2.17	5.22	39.07	38.34	-12.07	1.93

Table 6.15 - CFD Results for Scenario 6 with 50% Useful IT Load

ψ_T	Ar	\dot{m}_C , kg/s	$\max(T_{i,j}^{in,on})$, °C	$\max(T_{i,j}^{in,idle})$, °C	δT , °C	T_a^* , °C
1.00	0.54	10.44	23.72		3.28	17.28
0.95	0.60	9.92	24.88		2.12	16.12
0.90	0.67	9.40	26.19		0.81	14.81
0.85	0.75	8.88	27.37		-0.37	13.63
0.80	0.85	8.35	28.29		-1.29	12.71
0.75	0.97	7.83	28.47		-1.47	12.53
0.70	1.11	7.31	28.86		-1.86	12.14
0.60	1.51	6.27	31.15		-4.15	9.85
0.50	2.17	5.22	37.59		-10.59	3.41

Table 6.16 - CFD Results for Scenario 7 with 50% Useful IT Load

ψ_T	Ar	\dot{m}_C , kg/s	$\max(T_{i,j}^{in,on})$, °C	$\max(T_{i,j}^{in,idle})$, °C	δT , °C	T_a^* , °C
1.00	0.54	10.44	14.24	23.00	9.00	23.00
0.95	0.60	9.92	14.32	24.04	7.96	21.96
0.90	0.67	9.40	14.47	24.97	7.03	21.03
0.85	0.75	8.88	14.71	26.00	6.00	20.00
0.80	0.85	8.35	15.21	27.09	4.91	18.91
0.75	0.97	7.83	16.44	28.13	3.87	17.87
0.70	1.11	7.31	18.16	29.09	2.91	16.91
0.60	1.51	6.27	23.12	31.65	0.35	14.35
0.50	2.17	5.22	31.70	36.58	-4.70	9.30

Table 6.17 - CFD Results for Scenario 7 with 25% Useful IT Load

Ψ_T	Ar	\dot{m}_C , kg/s	$\max(T_{i,j}^{in,on})$, °C	$\max(T_{i,j}^{in,idle})$, °C	δT , °C	T_a^* , °C
1.00		9.86	14.00	20.67	11.33	25.33
0.90		8.88	14.00	22.00	10.00	24.00
0.80		7.89	14.00	23.66	8.34	22.34
0.70		6.90	14.04	25.70	6.30	20.30
0.60		5.92	14.38	28.21	3.79	17.79
0.50		4.93	16.04	31.37	0.63	14.63
0.40		3.95	20.68	34.85	-2.85	11.15

Table 6.18 - CFD Results with 0% Useful IT Load

ψ_T	Ar	\dot{m}_C , kg/s	$\max(T_{i,j}^{in,on})$, °C	$\max(T_{i,j}^{in,idle})$, °C	δT , °C	T_a^* , °C
1.00	0.37	9.28		18.48	13.52	27.52
0.95	0.41	8.82		18.87	13.13	27.13
0.90	0.45	8.35		19.29	12.71	26.71
0.85	0.51	7.89		19.78	12.22	26.22
0.80	0.57	7.43		20.32	11.68	25.68
0.75	0.65	6.96		20.93	11.07	25.07
0.70	0.75	6.50		21.41	10.59	24.59
0.65	0.87	6.03		21.61	10.39	24.39
0.60	1.02	5.57		22.06	9.94	23.94
0.55	1.21	5.11		22.64	9.36	23.36
0.50	1.47	4.64		23.55	8.45	22.45
0.40	2.29	3.71		27.63	4.37	18.37

6.3 Application of Proper Orthogonal Decomposition to Data Centers

Computational fluid dynamics provided a rich dataset for studying optimum thermally aware, energy-based load placement. However, expanding on this analysis using computational fluid dynamic or detailed experimental measurements is quite prohibitive. Therefore, the focus changes to developing a statistics-based reduced-order model using proper orthogonal decomposition (POD), as described by Holmes et al. (1996). POD has been successfully applied to many applications, including indoor airflows (Elhadidi and Khalifa, 2005; Khalifa, Elhadidi and Dannenhoffer, 2007; Rambo and Joshi, 2005), turbulent flows (Lumley, 1981; Arndt et al., 1997), development of simplified flow control techniques (Efe and Ozbay, 2003; Ly and Tran, 2001; Podvin and Lumley, 1998) and for character/face recognition (Everson and Sirovich, 1995). POD represents the solution domain in terms of the most “energetic” characteristics. In the case of data centers, these characteristics may include CRAH airflow rate, CRAH supply air temperature, IT power level or IT load placement. To develop the POD tool, the CFD solutions are used to derive “empirical” eigenmodes for each placement scenario, which can be computed and stored. These eigenmodes can then be used to evaluate the chassis’ inlet temperature distribution at a range of parameters within the design space, with a considerable reduction in computational time compared to a detailed CFD analysis.

6.3.1 Proper Orthogonal Decomposition Theory

To construct the POD eigenmodes, an ensemble of snapshots, $\mathbf{T}_i(\mathbf{x})$, $i = 1, 2, \dots, N_s$, where \mathbf{T}_i represents a temperature solution set, \mathbf{x} is the spatial coordinate and N_s is the number of snapshots, is considered. In this study, each snapshot is obtained from a CFD solution containing the temperature at each of the N_p computational cells at the inlet to the racks in the cold aisle, for different operating conditions. The goal of POD is to represent any of the snapshots by,

$$\mathbf{T}_i(\mathbf{x}) = \bar{\mathbf{T}}(\mathbf{x}) + \boldsymbol{\tau}(\mathbf{x}) \approx \bar{\mathbf{T}}(\mathbf{x}) + \sum_{k=1}^{N_s} c_k \boldsymbol{\varphi}_k(\mathbf{x}) \quad (6.9)$$

where, $\bar{\mathbf{T}}(\mathbf{x})$ is the ensemble average of the snapshots at each spatial point and $\boldsymbol{\tau}(\mathbf{x})$ is the deviation of the spatial point from the average, which can be expanded in terms of the eigenmodes $\boldsymbol{\varphi}_k$ with c_k representing the amplitude of each mode. The amplitudes depend on the variables used to develop the different operating scenarios. It is expected that the eigenmodes should reconstruct a snapshot, with yet-to-be determined accuracy as N_s is reduced.

The solution to the problem seeks the orthonormal set of eigenmodes $\boldsymbol{\varphi}_k$ that minimize the average squared error of the reconstructed field and the original snapshot, i.e.,

$$\varepsilon = \left\| \boldsymbol{\tau} - \sum_{k=1}^{N_s} c_k \boldsymbol{\Phi}_k \right\|_{\mathbf{x}}^2 = \min \quad (6.10)$$

It is expected that the first eigenmode would have the highest variance and account for as much of the variability in the data as possible. Each succeeding eigenmode would then have the highest variance possible under the constraint that it be orthogonal to the preceding eigenmodes. The eigenmodes satisfying this criterion are the solution of the classical Fredholm eigenvalue problem (Holmes et al., 1996). The solution of this eigenvalue problem can be solved by one of two methods, the direct method or the method of snapshots (Sirovich, 1987). The direct method computes a two-point correlation matrix between each variable and spatial point in the ensemble. The size of this matrix would be $(N_s N_p)^2$, which for typical problems is quite large and renders the solution of the eigenvalue problem computationally intensive. On the other hand, the method of snapshots uses a correlation matrix \mathbf{R} with size N_s^2 . Therefore, when $N_s \ll N_p$ the method of snapshots provides a significant computational advantage. The method of snapshots solves the following eigenvalue problem,

$$\mathbf{R}\mathbf{v} = \lambda\mathbf{v}, \quad (6.11)$$

where, \mathbf{v} and λ are the eigenvectors and eigenvalues of the correlation matrix, whose elements are given by,

$$R_{ij} = \langle \boldsymbol{\tau}_i(\mathbf{x}) \boldsymbol{\tau}_j(\mathbf{x}) \rangle = \frac{1}{N_s} \int \boldsymbol{\tau}_i(\mathbf{x}) \boldsymbol{\tau}_j^T(\mathbf{x}) d\mathbf{x} \approx \frac{1}{N_s} \sum_{k=1}^{N_s} \boldsymbol{\tau}_i(\mathbf{x}) \boldsymbol{\tau}_j^T(\mathbf{x}) \quad (6.12)$$

The correlation matrix \mathbf{R} is symmetric and positive definite and therefore, all eigenvalues will be real and the eigenmodes are orthogonal, as expected. The eigenmodes represent a complete set of all solutions within the ensemble space; therefore, given the values of the amplitude coefficients, c_k , can be used to optimally reconstruct any solution by a linear combination of the eigenmodes (i.e., interpolation). The eigenmodes $\boldsymbol{\phi}_k(\mathbf{x})$ can be computed from,

$$\boldsymbol{\phi}_k = \sum_{k=1}^{N_s} \mathbf{v}_k \boldsymbol{\tau}_k(\mathbf{x}). \quad (6.13)$$

The amplitude coefficients c_k are obtained by applying the inner product of $\boldsymbol{\tau}_i(\mathbf{x})$ in Equation 6.9,

$$c_k = \frac{\langle \boldsymbol{\tau}_i \boldsymbol{\phi}_k \rangle}{\langle \boldsymbol{\phi}_k \boldsymbol{\phi}_k \rangle}. \quad (6.14)$$

The amplitude coefficients are also orthogonal and therefore uncorrelated. The coefficients will depend on the choice of design variables and once known can be used for interpolating within the ensemble space.

A useful measure of the expected accuracy by keeping N_n of the N_s eigenmodes in the expansion given by Equation 6.9 is the relative energy, E_n , that is captured by the N_n eigenmodes,

$$E_n = \frac{\sum_{i=1}^{N_n} \lambda_i}{\sum_{i=1}^{N_s} \lambda_i} . \quad (6.15)$$

The relative energy can give an indication of how well a snapshot in the original ensemble can be reconstructed; however, it gives no indication about additional data sets that will be interpolated using the amplitude coefficients, c_k . Typically, one would choose the N_n eigenmodes to keep based on some convergence criterion based on the desired level of accuracy needed for the application (Sirovich and Everson, 1992).

6.3.2 Application of POD to Thermally Aware, Energy-based Load Placement

The CFD analysis for developing thermally aware, energy-based load placement techniques was done over a limited range of conditions, owing to the fact that performing detailed simulations for all possible scenarios was impractical. Therefore, the proper orthogonal decomposition theory is applied to the CFD data sets obtained in Section 6.2.3. The results presented here are for IT placement Scenario 7, in which the IT load was removed beginning with those chassis with the highest inlet temperature.

The analysis considers the datasets already computed, as given in Tables 6.2, 6.9, 6.16, 6.17 and 6.18. The complete ensemble of snapshots contains solutions for useful IT levels of 0%, 25%, 50%, 75% and 100% and $\psi_T = 0.40 - 1.00$, in increments of 0.10. The solutions that have already been obtained at other values of ψ_T , in the above tables, will be saved and used for validation of the POD model. Therefore, the ensemble consists of 35 snapshots, with the choice of design variables being the useful IT level and ψ_T . The interest in using the POD is to be able to reproduce the rack's inlet temperature distribution efficiently using limited computational resources. Therefore, for each snapshot, the temperature at each of the $N_p = 6912$ computational grid points along the rack's inlet are stored. However, the prior work on thermally aware, energy-based load placement focused on chassis level temperatures. Therefore, after the POD is used to generate the 6912 temperatures at each of the computational cells, the data is reduced by averaging the 108 computational cells on each chassis to obtain the 64 chassis inlet temperatures. This approach was used to avoid any smearing of the temperature field by performing the averaging first and then developing the POD model. Another benefit of this approach is that the analysis can easily be extended to look at 1U, 2U or other chassis and server sizes, without performing additional work, if required.

Figure 6.5 shows the energy content of the POD eigenmodes. The modes converge rapidly with over 93% of the variance captured in the first eigenmode. With six eigenmodes, over 99% of the relative energy is captured, indicating the POD method can reduce the information contained in the original 35 datasets, down to a few empirical eigenmodes.

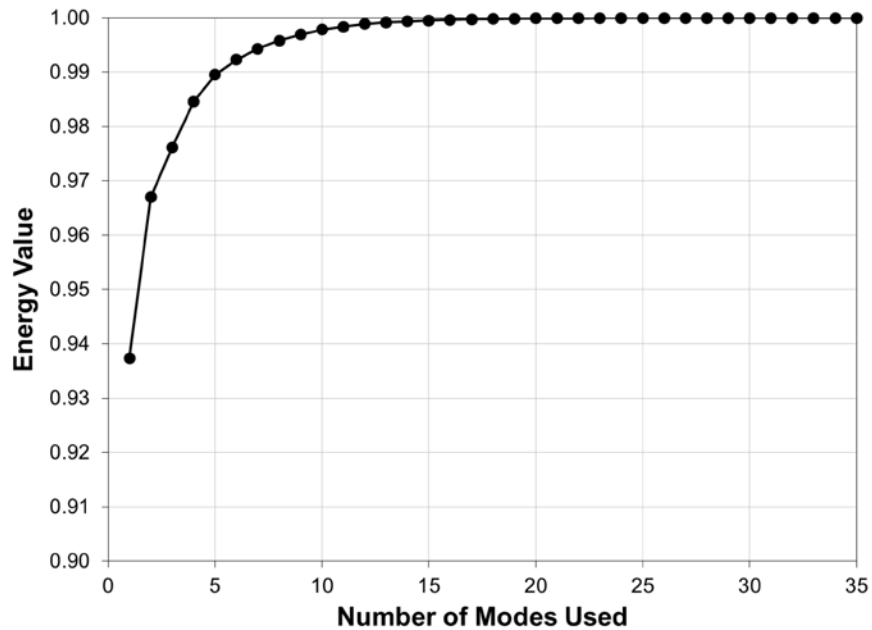
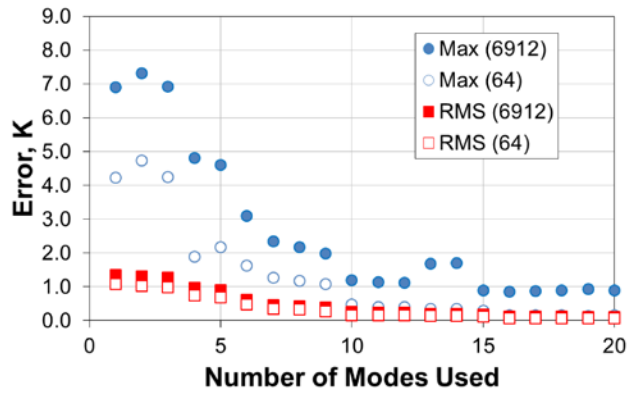


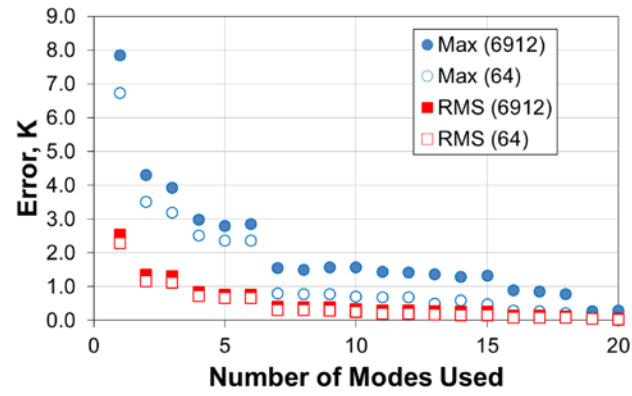
Figure 6.5 - Energy Content of POD Modes

While the energy content of the eigenmodes provides a quick metric for evaluating the accuracy of the POD approach, the first real test of the effectiveness of the POD is in how well it can reconstruct the original datasets. This is evaluated by considering the root mean squared (RMS) and maximum temperature errors obtained as additional modes are kept in the expansion given by Equation 6.5. Figures 6.6 – 6.10 show the errors for useful IT loads of 100%, 75%, 50%, 25% and 0%, respectively, for ψ_T of 1.00, 0.80, 0.60 and 0.40, as modes 1 – 20 are retained in the expansion. The figures show both the results of the full $N_p = 6012$ as well as the reduced dataset of 64 chassis temperatures. Overall, the error converges rapidly for all the snapshot in the ensemble, with RMS and maximum errors of less than 0.2°C and 1.0°C for the 64 chassis using about 12 modes. Obviously,

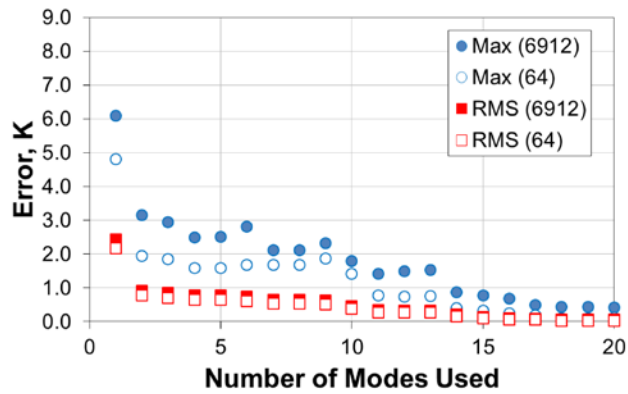
larger errors are obtained if all 6012 points are being predicted; however, with about 15 modes, even these can be reconstructed to within a 1°C maximum error.



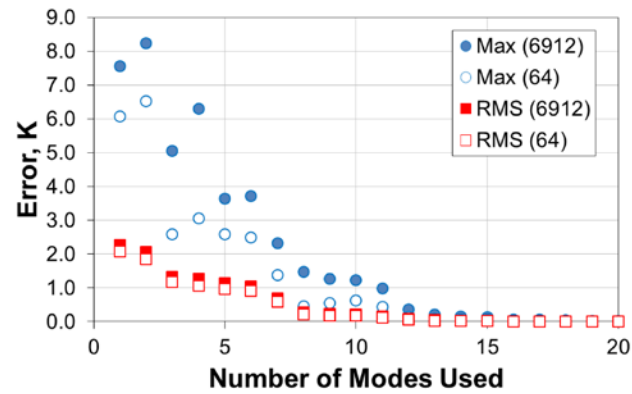
$\psi_T = 1.0$



$\psi_T = 0.8$

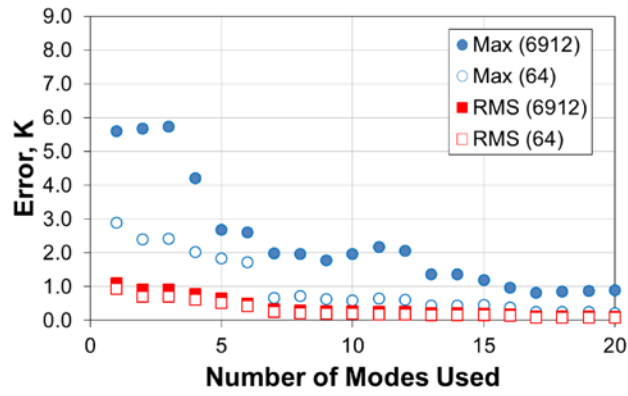


$\psi_T = 0.6$

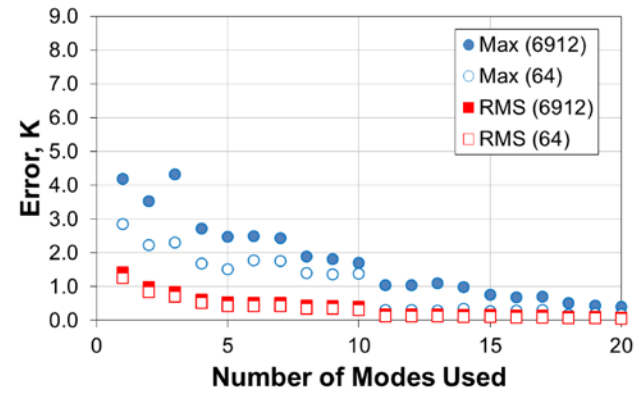


$\psi_T = 0.4$

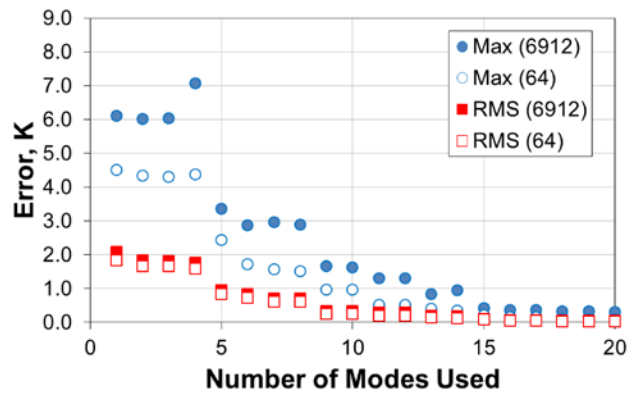
Figure 6.6 - Reconstruction Error for Useful IT of 100%



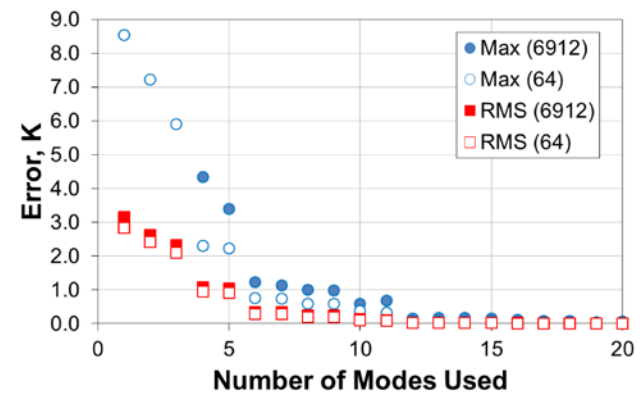
$\psi_T = 1.0$



$\psi_T = 0.8$

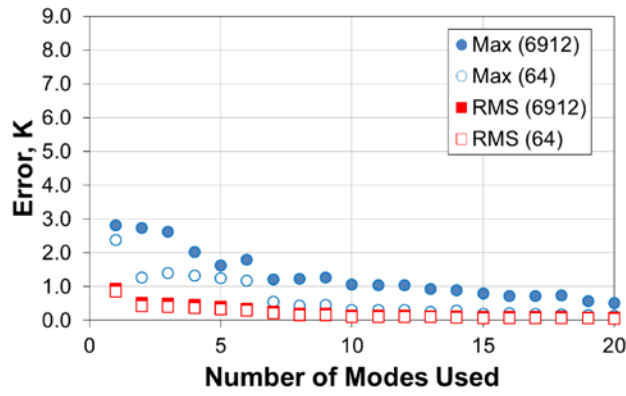


$\psi_T = 0.6$

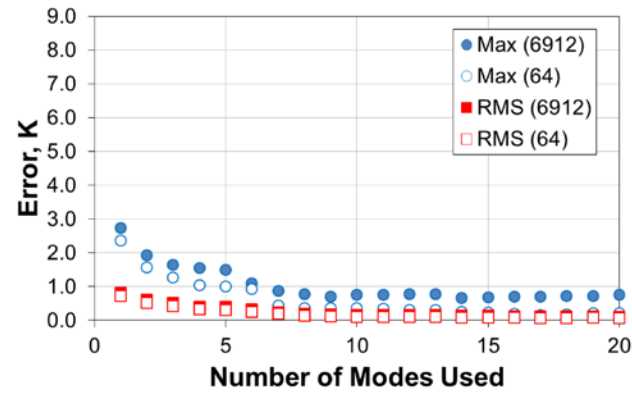


$\psi_T = 0.4$

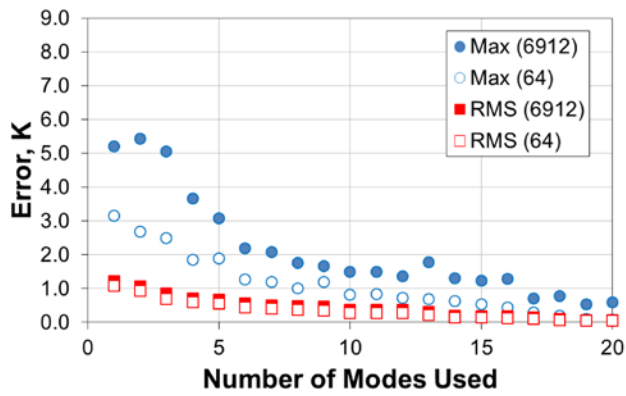
Figure 6.7 - Reconstruction Error for Useful IT of 75%



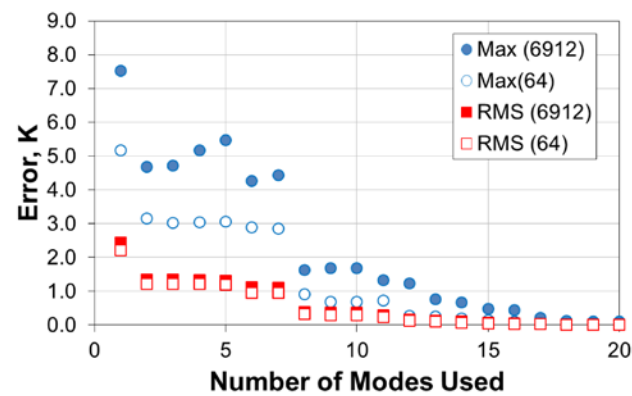
$\psi_T = 1.0$



$\psi_T = 0.8$

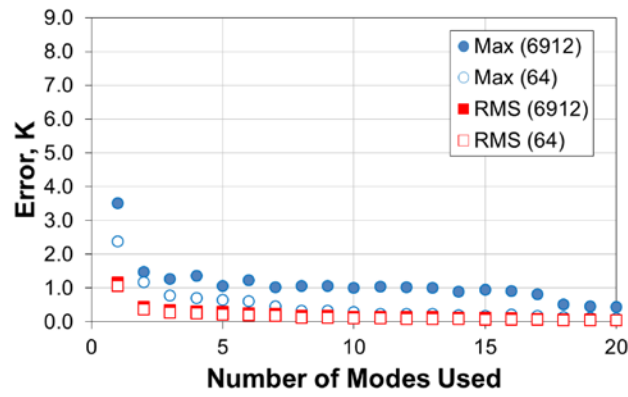


$\psi_T = 0.6$

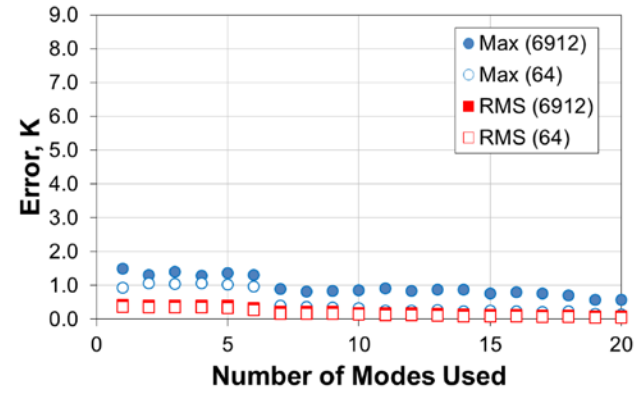


$\psi_T = 0.4$

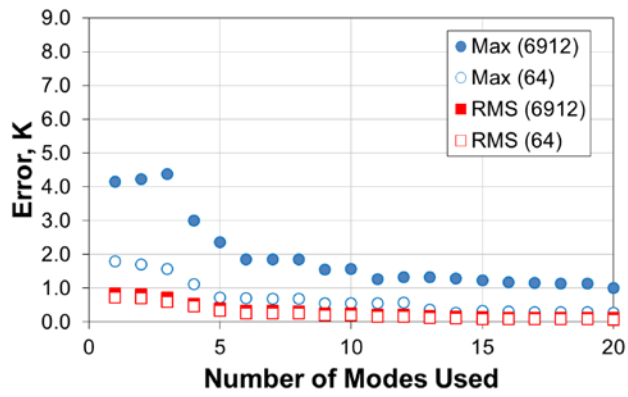
Figure 6.8 - Reconstruction Error for Useful IT of 50%



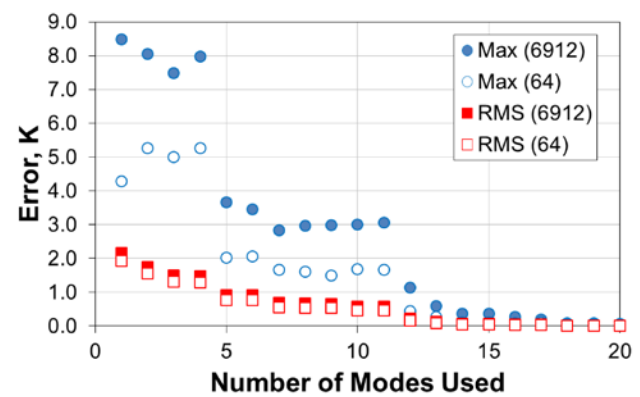
$\psi_T = 1.0$



$\psi_T = 0.8$

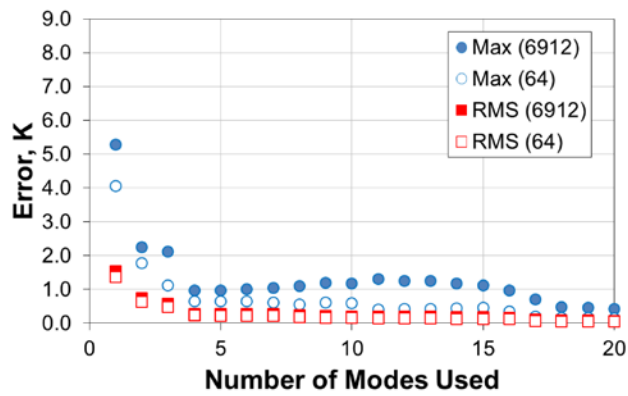


$\psi_T = 0.6$

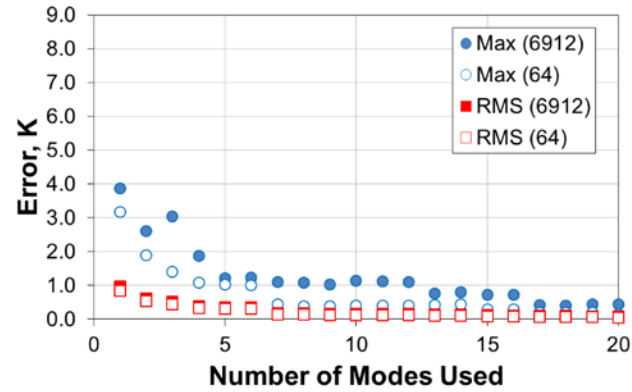


$\psi_T = 0.4$

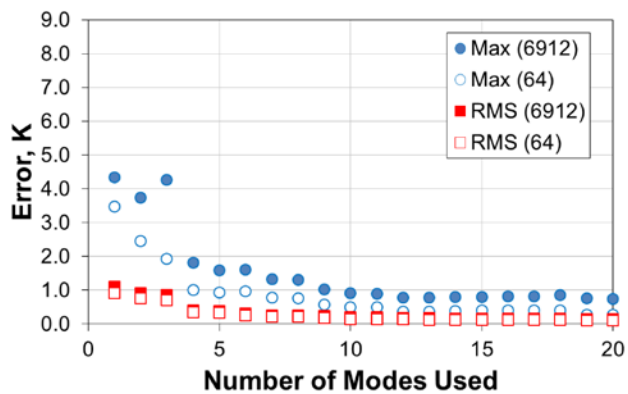
Figure 6.9 - Reconstruction Error for Useful IT of 25%



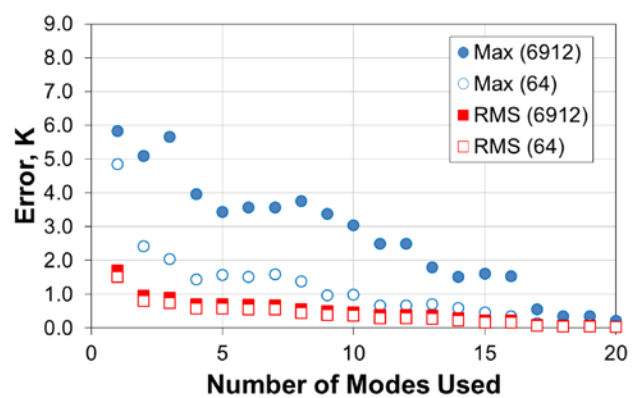
$\psi_T = 1.0$



$\psi_T = 0.8$



$\psi_T = 0.6$

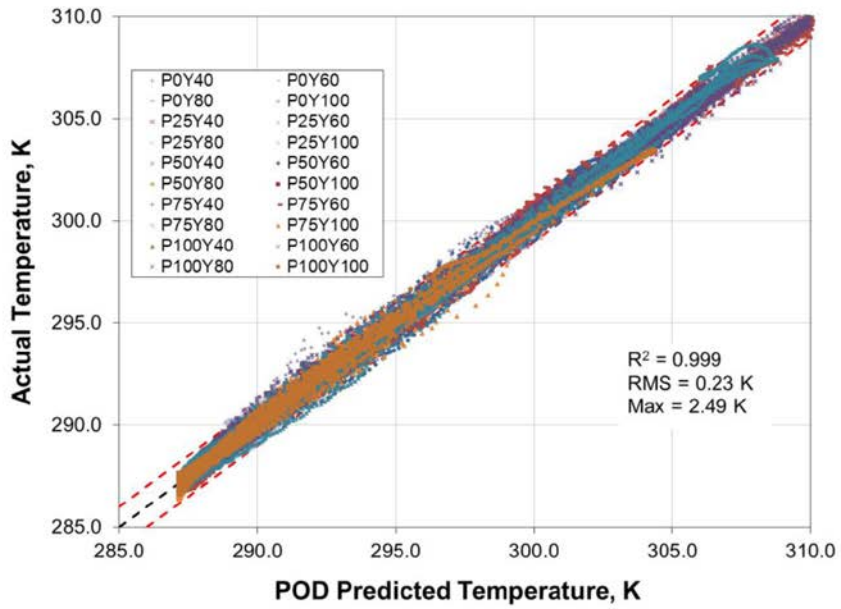


$\psi_T = 0.4$

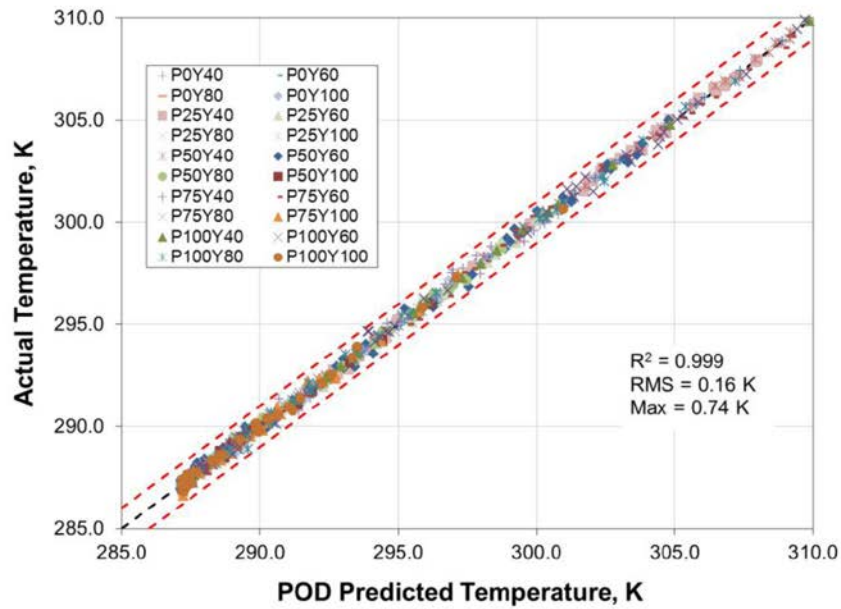
Figure 6.10 - Reconstruction Error for Useful IT of 0%

With these results in mind, the analysis proceeds with using 12 eigenmodes. Figure 6.11 compares the reconstruction capabilities of the POD method using 12 eigenmodes with the actual temperatures from the CFD results, for 20 of the realistic snapshots³. Figure 6.11a shows the results for the entire POD dataset, while Figure 6.11b shows the results of the reduced set of 64 chassis for each snapshot. The figures also provide $\pm 1^\circ\text{C}$ error bars. By retaining 12 modes in the POD expansion, reconstruction results in a RMS error of 0.16°C and a max error of 0.74°C , for the 1280 chassis temperatures in the 20 snapshots. Theoretically, retaining twelve modes as opposed to the six predicted by the eigenvalue energy content would result in additional computational time, obviously at some cost in accuracy; however, on a standard personal computer the reconstruction of a dataset is completed in a fraction of a second and therefore, there is really no penalty for retaining the twelve modes as opposed to six. In addition, we can reconstruct an infinite number of solutions at hardly any computational cost, compared to running a CFD analysis, which for this work takes nearly seven hours to complete a single solution using 32 nodes on a high performance computer cluster.

³ While included in the development of the POD, several of the snapshots have been omitted from reconstruction because they would be unrealistic scenarios in operating data centers. For example, a case of 100% useful IT and an $\psi_T = 0.4$, is highly unrealistic since the temperature of the chilled air that would be required to meet the inlet temperature constraint is far below what could be provided using chilled water CRAH units.



a)



b)

Figure 6.11 - POD Reconstruction Comparison with CFD Data

a) all 6912 temperatures and b) 64 chassis temperatures

In addition to being a set of orthogonal basis functions to satisfy the decomposition in Equation 6.5, the eigenmodes also contain information related to the physics of the temperature distribution. As was presented by Sirovich and Everson (1992) for the solution of the so-called “Rogue Gallery” problem, in which POD was used as a tool for face recognition, the different eigenmodes represent different features of the solution. In that problem, Sirovich and Everson (1992) showed that from the first eigenmode emerged an oval representing the outline of the face. Further modes provided details of the eyes, mouth, nose, etc. We take a similar approach for describing the physics that leads to the rack’s inlet temperature distribution. Figure 6.12 shows the mean of the temperature distribution snapshots at the inlet of one of the rows of racks in the cold aisle (the view is as if you were standing in the cold aisle facing the IT equipment). The mean temperature distribution highlights several of the expected airflow and temperature features, including, the chassis towards the bottom of the racks and center of the aisle receiving only cold air and hot air recirculation at the top of the racks and side of the aisle.

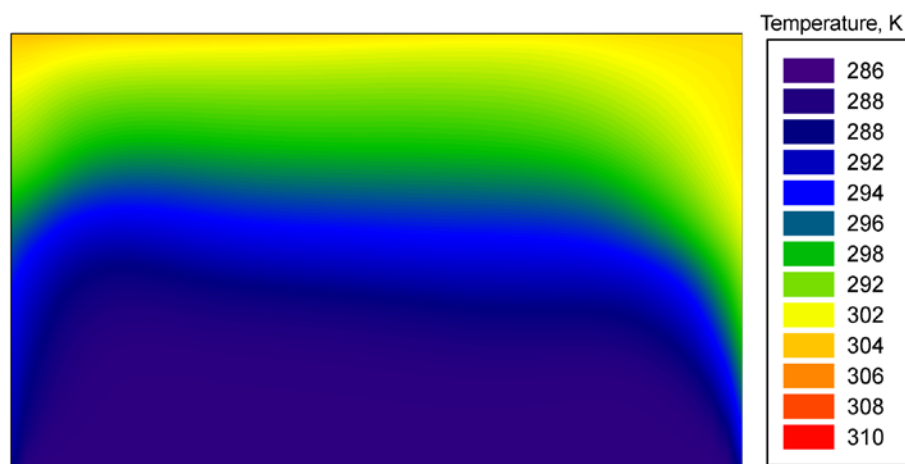


Figure 6.12 – Mean of the Rack’s Inlet Temperature Distribution for POD Data

Figures 6.13 and 6.14 show several reconstructed snapshots, while keeping 1, 4, 8, and 12 eigenmodes, along with the actual temperature distribution. Clearly, mode 1, which by design contains the most variance, reconstructs the general shape of the temperature distribution; however, it significantly under-predicts the magnitude of hot air recirculation. By mode 4, the precise details of the recirculation pattern is apparent and by mode 8, most of the details of the temperature distribution are noticeable, including the asymmetry due to the CRAH units only being on the left side. Interestingly, for the cases with low Ar , the temperature pattern is predicted accurately by mode 8; whereas, the higher Ar cases required more modes and even in some cases, the finer details of the solution are not captured with 12 modes. One explanation for this can be attributed to the change in temperature pattern as Ar becomes greater than unity – the pattern changes from a jet-like profile to a stratified pattern as buoyancy becomes important. The lower momentum flow typically has more flow features since they are greatly affected by CRAH suction and less by the jet momentum. These patterns are highly non-uniform and unsymmetrical, with subtle features not in the low Ar cases. In addition, the original dataset contained fewer snapshots with a stratified profile and therefore, it is expected that these cases would require more information to capture accurately.

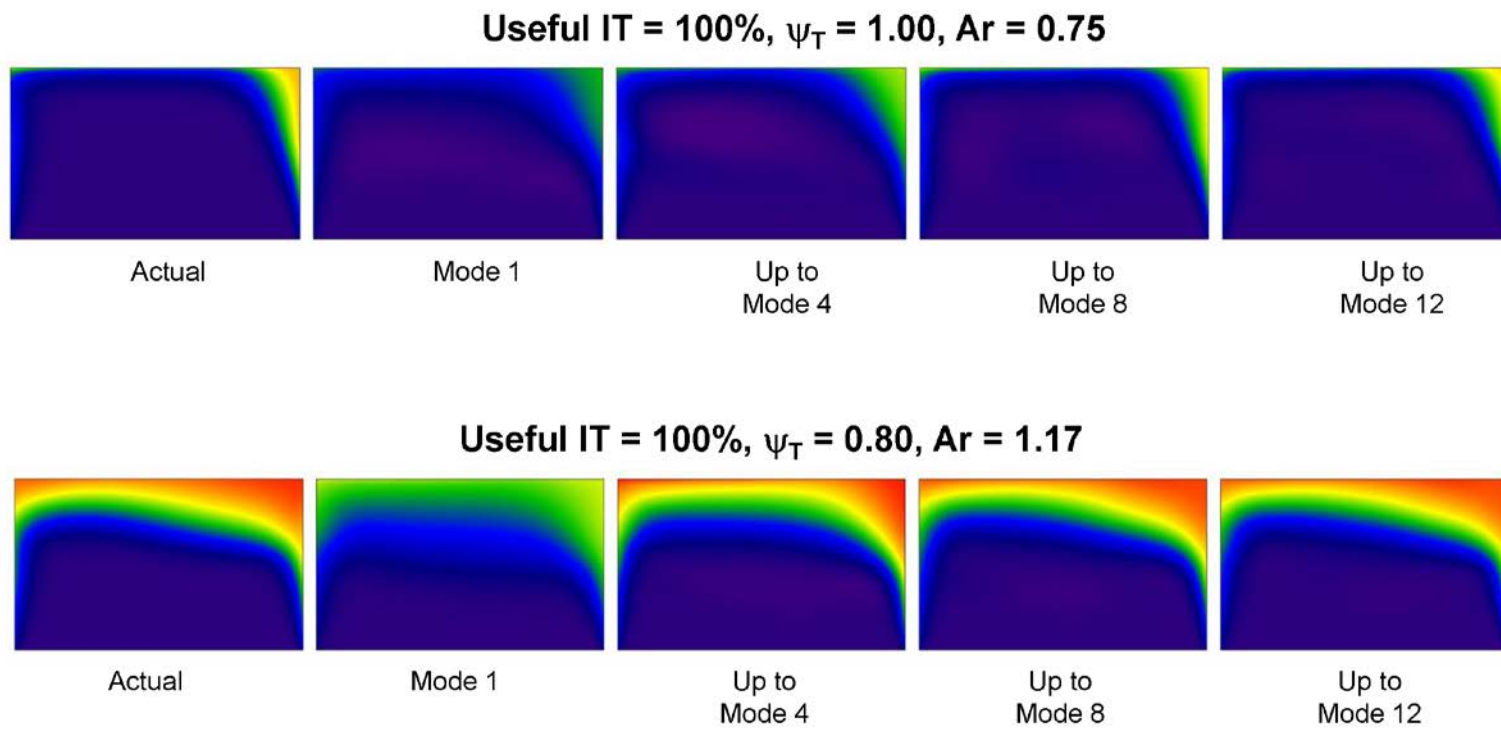


Figure 6.13 - Temperature Reconstruction Contours as the Number of Modes in the Expansion is increased for 100% Useful IT

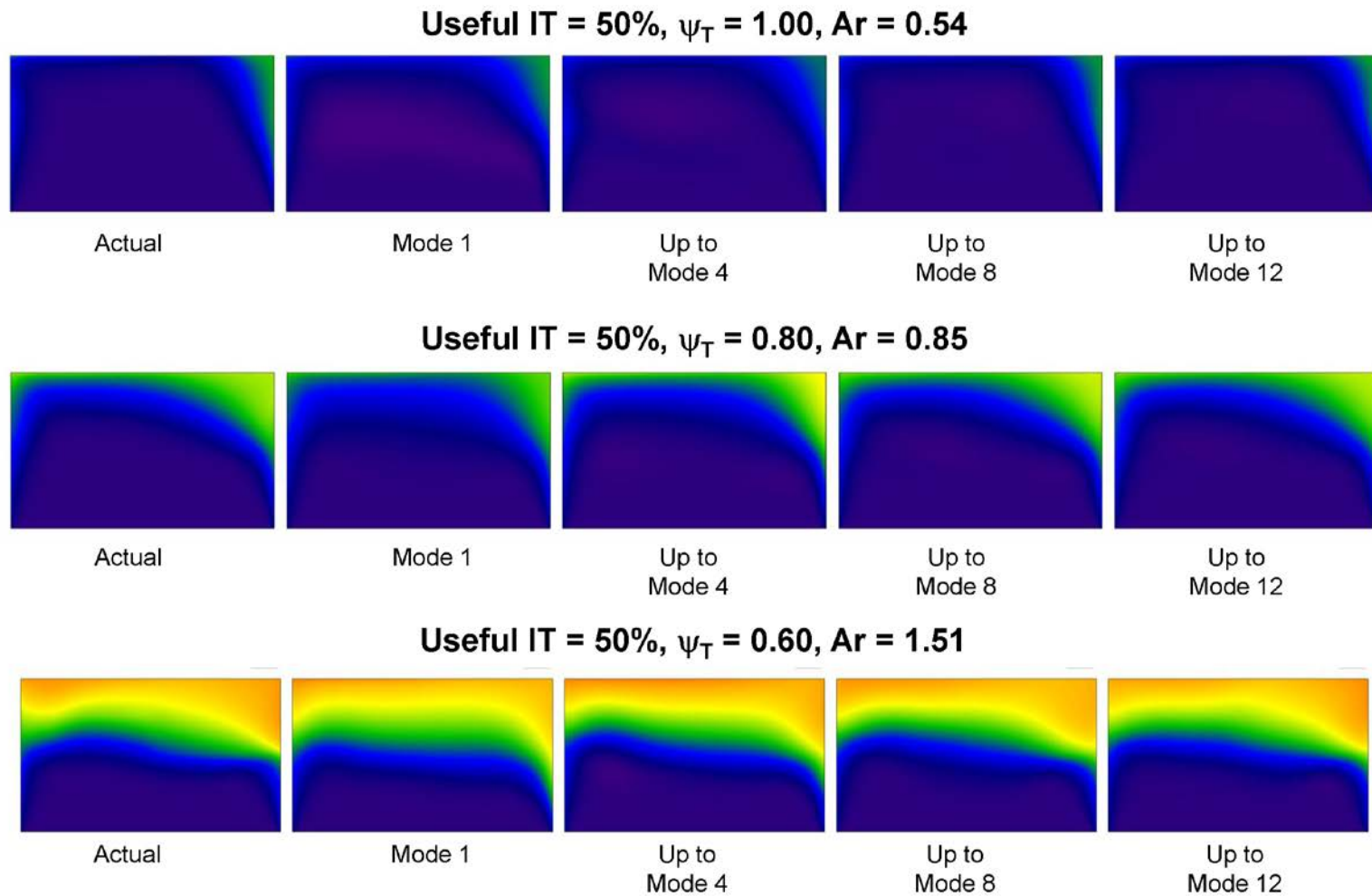
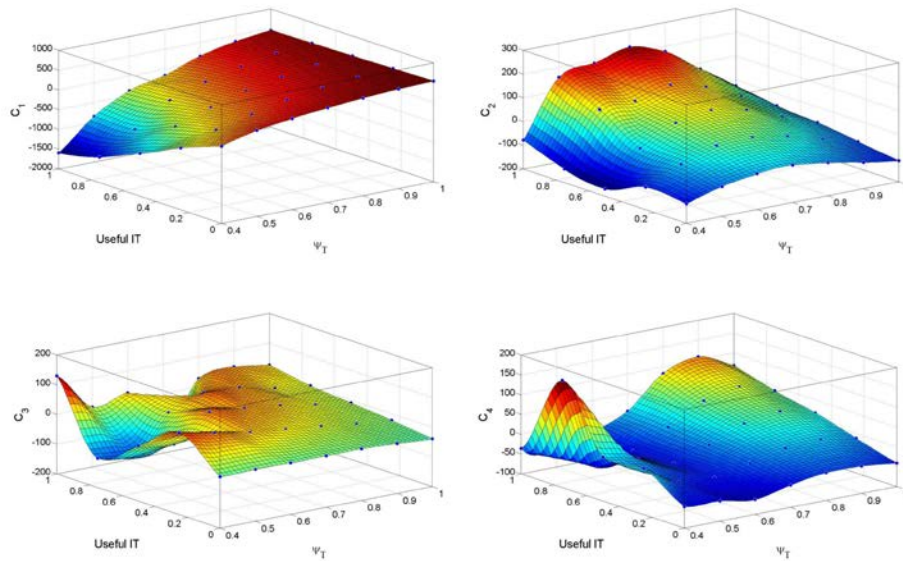


Figure 6.14 - Temperature Reconstruction Contours as the Number of Modes in the Expansion is increased for 50% Useful IT

The real strength of developing a POD model is in its ability to be used as an interpolation tool. To do this, the amplitude coefficients, c_k , must be given as a function of the design variables, in this case, the useful IT power and ψ_T . Figure 6.15 shows the relationship of the first twelve expansion coefficients. From the figures, it can be seen that the first mode has the highest amplitude and smoothest coefficient distribution. Indeed, higher order modes have waiver coefficient distributions, which should be expected. For instance in a Fourier decomposition, the lower order modes have the highest amplitude and least number of zero crossings. The coefficient distributions have been fit using cubic splines in order to estimate the higher order modes more accurately, as suggested by Bui-Thanah et al. (2003).



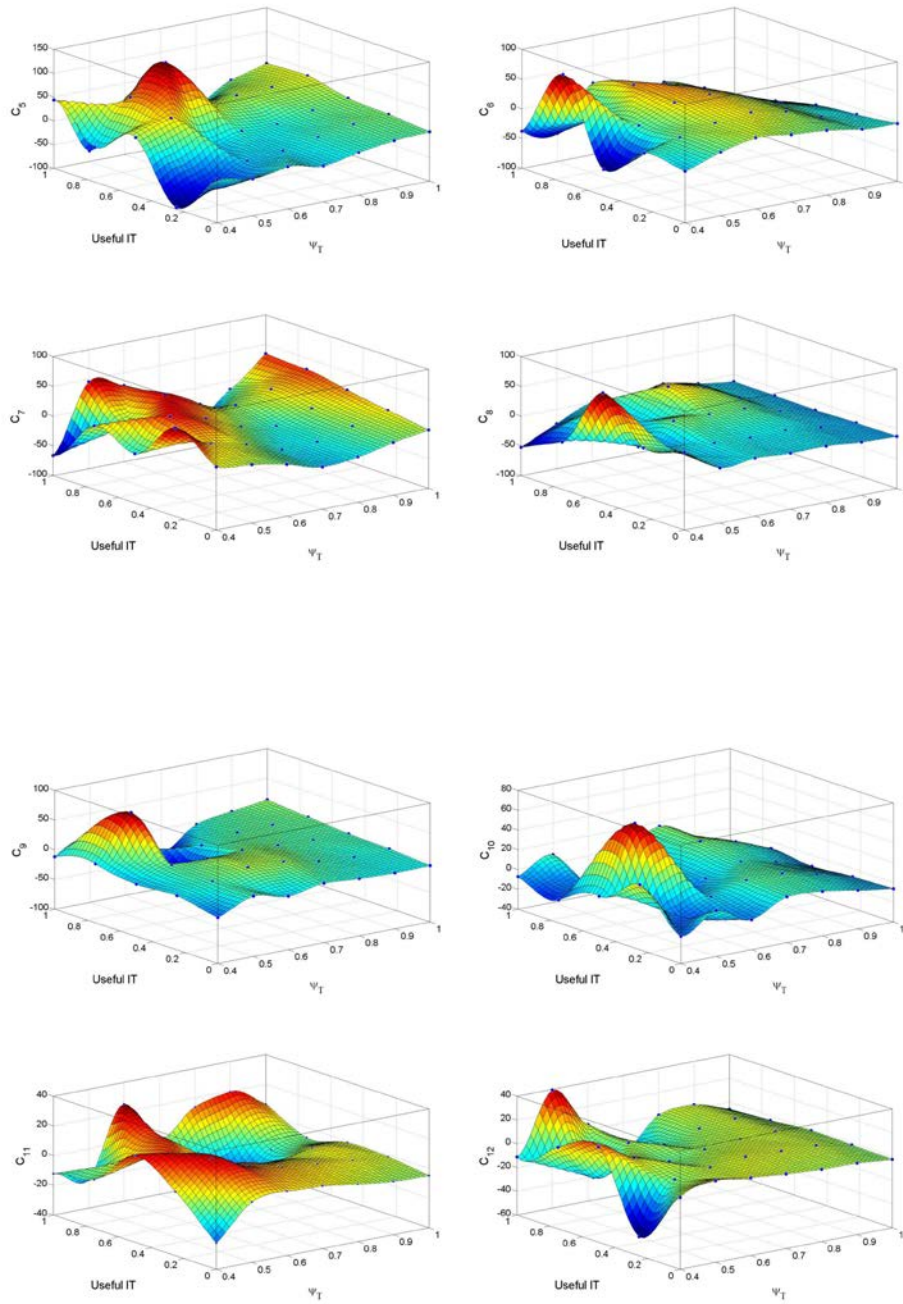
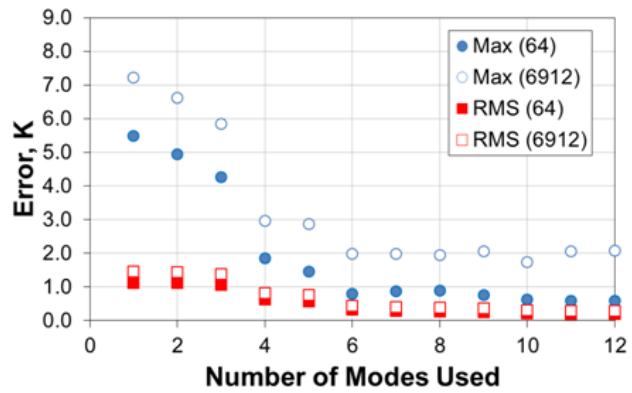
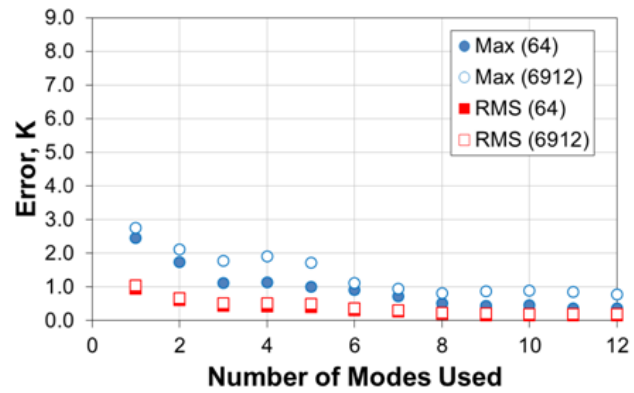


Figure 6.15 - Amplitude Coefficient Distributions

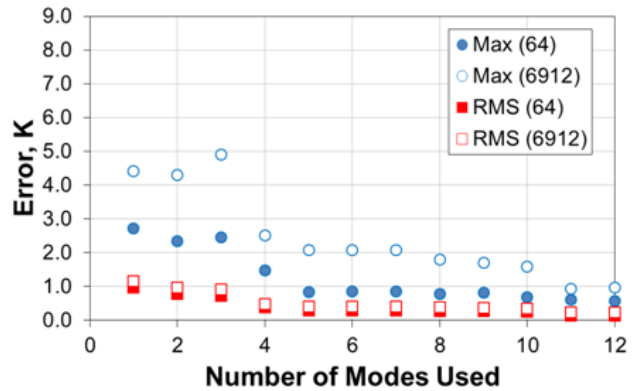
Using the amplitude coefficient distributions, the POD model is used to interpolate for datasets that were not included in the original ensemble. We consider the CFD results presented in Tables 6.2, 6.9, 6.16 and 6.18 that were not used in the development of the POD for validation. Theoretically, it can be shown that the error in reconstruction will go to zero as the number of modes in the expansion is increased. However, this is not necessarily the case for the interpolated datasets since there is error introduced by estimating the coefficients from the distributions in Figure 6.15. For example, Figure 6.16 shows the error in interpolation for four sample cases. The results for the 100% useful IT and $\psi_T = 0.95$ show that the error decrease gradually until mode 6 and the increases slightly from mode 6 to mode 12, because of the error introduced by interpolating the higher order coefficients. In general, this may suggest keeping fewer than 12 modes in the expansion to avoid this effect. However, the interest of this study is in determining the 64 chassis' inlet temperature and the results show that the averaging will smooth out much of the error introduced by the interpolation.



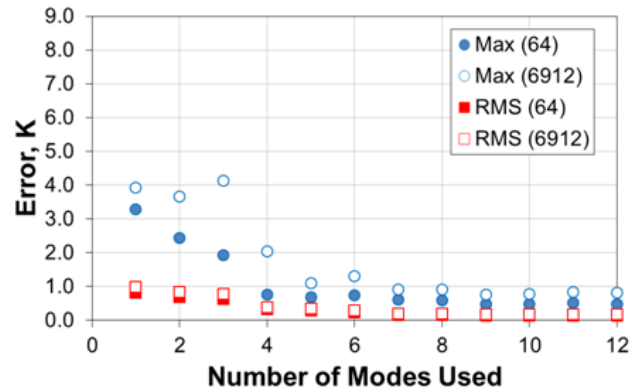
Useful IT = 100%, $\psi_T = 0.95$



Useful IT = 50%, $\psi_T = 0.75$



Useful IT = 75%, $\psi_T = 0.85$



Useful IT = 0%, $\psi_T = 0.65$

Figure 6.16 - Error Plots for Interpolated Datasets

Figure 6.17 gives a comparison of the POD predicted temperature and the actual temperature for all the validation cases. The maximum temperature error of the 832 chassis' temperatures in the 13 validation cases was 0.95°C with an RMS error of 0.16°C. The validation results provide confidence that the POD model can be used for performing the work on thermally aware, energy-based load placement.

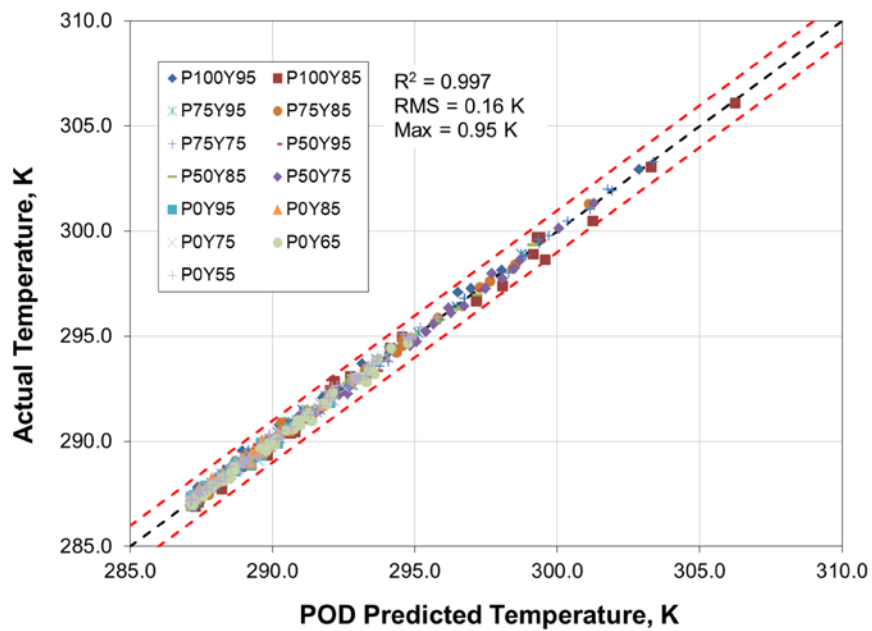


Figure 6.17 - POD Interpolation Comparison with CFD Data

6.3.3 Section Conclusions

This section introduced a methodology for developing a reduced order model, using proper orthogonal decomposition, to predict the rack's inlet temperature distribution. The method uses a limited set of computational fluid dynamics data at different useful IT

levels and tile airflow fractions, ψ_T . The model was able to reconstruct these datasets to within 0.16°C RMS error. The model was also used to interpolate successfully for alternative configurations that were not included in the original dataset. Therefore, the POD model can be used to assess optimum thermally aware, energy-based load placement strategies at an infinite combination of useful IT and ψ_T . The POD model can generate a new design alternative in a fraction of a second on a standard personal computer. In addition, the model was able to interpolate configurations that were not included in the original ensemble to within a maximum error of 1°C, using a fraction of the information. The number of modes needed to generate a dataset within the 1°C error was reduced from the full set of 35 empirical eigenfunctions to twelve eigenmodes using POD. It is anticipated that these models can be used as predictive tools in operating data centers to assess the outcomes of load placement adjustments to the rack's inlet temperature distribution before changes are made that could affect the IT equipment's reliability.

6.4 Workload Placement Optimization Results

With a fast and efficient method for predicting the rack's inlet temperature distribution at an infinite number of design scenarios, we proceed with evaluating the load placement scenarios proposed in Section 6.2.1.

6.4.1 Baseline Scenarios

First, as a means of comparison, a baseline for each of the proposed scenarios is introduced based on what would typically be done in data centers today. In many instances, there is inadequate cooperation between facility operators and the IT specialists. Without this communication, sub-optimal operation of the data center is likely possible, as each organization would optimize based on its own need. The baseline cases used in this work specifically address the situation where IT load placement is considered without any changes made to the cooling infrastructure. The IT placement considered is the uniform IT placement described by Scenario 6, as given in Section 6.2.1. It is assumed that no changes are made to the CRAH unit airflow or supply temperature settings from the 100% useful IT load case with $\psi_T = 1.0$ (i.e., those given in Table 6.2 for $\psi_T = 1.00$). The CRAH's supply air temperature is also fixed at the temperature that satisfies the chassis' inlet temperature constraint for the case of 100% useful IT and $\psi_T = 1.0$. The consequence of these baseline scenarios is that at a reduced IT load, significant over-cooling of the IT equipment is done, both in terms of the amount of air provided to the data center and the temperature of the air. Clearly, if the supply air temperature identically meets the chassis' inlet temperature constraint at 100% useful IT, if the flow rate remains fixed and the IT load is reduced, it is expected that the inlet temperatures would tend to lower values. Therefore, since the chassis' inlet temperature constraint is not met identically in the reduced IT load scenarios, there would be room for further increasing the supply air temperature until the constraint is met; however, this is not done in the baseline scenarios.

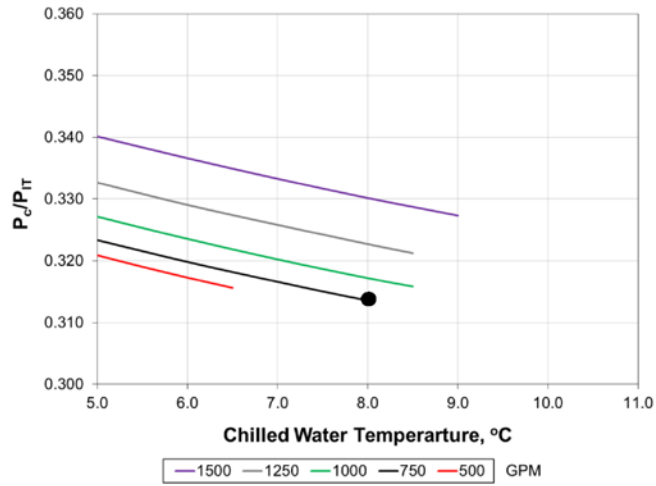
Table 6.19 provides the results of the CFD analysis for the baseline scenarios. To begin the analysis, the optimization of the baseline cases is described in detail for a given ambient wet bulb temperature of 21.9°C (30°C and 50% RH)⁴. Figure 6.18 plots the normalized cooling power versus the chilled water temperature for a range of chilled water flow rates, for each of the four baseline scenarios. As a reminder, each of the baseline cases has the same CRAH airflow rate and supply air temperature. For a given chilled water flow rate, the chilled water temperature can be increased until an infeasible solution is obtained, as governed by the NTU- ϵ heat exchanger model described in Section 6.2.2. As the chilled water temperature is increased at a given chilled water flow rate, the only quantity that changes is the chiller's refrigeration power consumption.

Table 6.19 - CFD Results for Baseline Scenarios

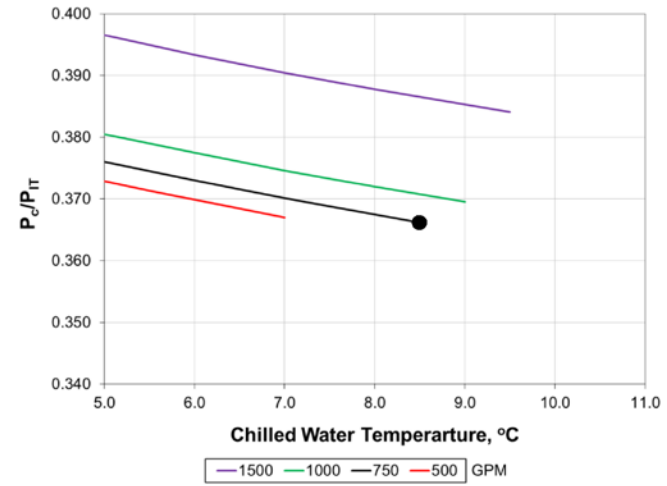
Useful IT	ψ_T	Ar	\dot{m}_C , kg/s	$\max(T_{i,j}^{in,on})$, °C	T_a , °C
0	1.30	0.22	11.64	16.68	13.19
50	1.16	0.40	11.64	20.66	13.19
75	1.10	0.53	11.64	23.64	13.19
100	1.00	0.75	11.64	27.00	13.19

⁴ In this section, all of the analysis will be done for an ambient wet-bulb temperature of 21.9°C.

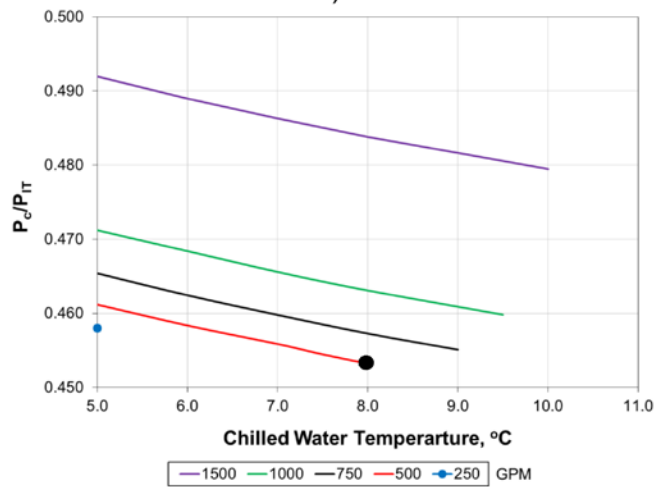
The results show the monotonically decreasing behavior of the chiller's power consumption as the chilled water temperature is increased, at a given IT load. Therefore, the lowest energy operating point, for a given chilled water flow rate, corresponds to the maximum achievable chilled water temperature. This procedure is repeated for a range of chilled water flow rates. Interestingly, the lowest chilled water flow rate does not always lead to the lowest cooling power consumption; in light of the fact that even though the pumping power is reduced, higher refrigeration power is needed since a lower chilled water temperature is necessary to satisfy the NTU- ϵ model. Therefore, the optimum operating point corresponds to the lowest cooling power consumption in the chilled water flow rate vs. chilled water temperature space, as given by the black circle on each of the plots in Figure 6.18. Figure 6.19 provides a detailed breakdown of the power consumption of each component. In Figure 6.19, the actual power consumption in kW is plotted to stress the point that the fan power consumption remains the same in each of the baseline scenarios. Changes in the chiller power consumption occur because the IT load is different in each case and not because the chilled water temperature is being adjusted.



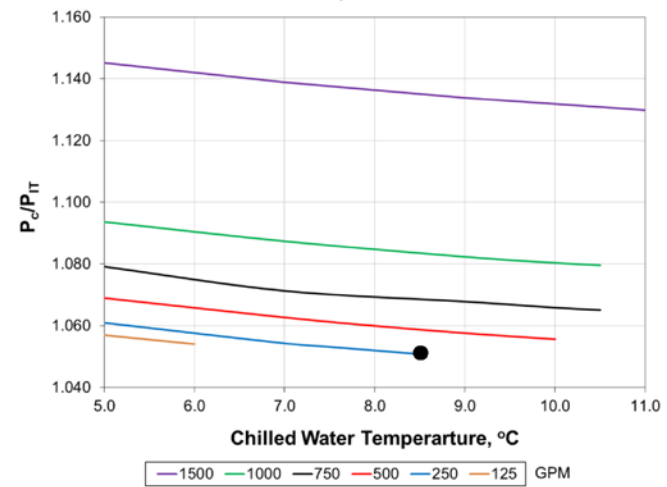
a)



b)



c)



d)

Figure 6.18 - Energy Optimization Details of the Baseline Scenarios for Useful IT

a)100%, b) 75%, c) 50% and d) 0%

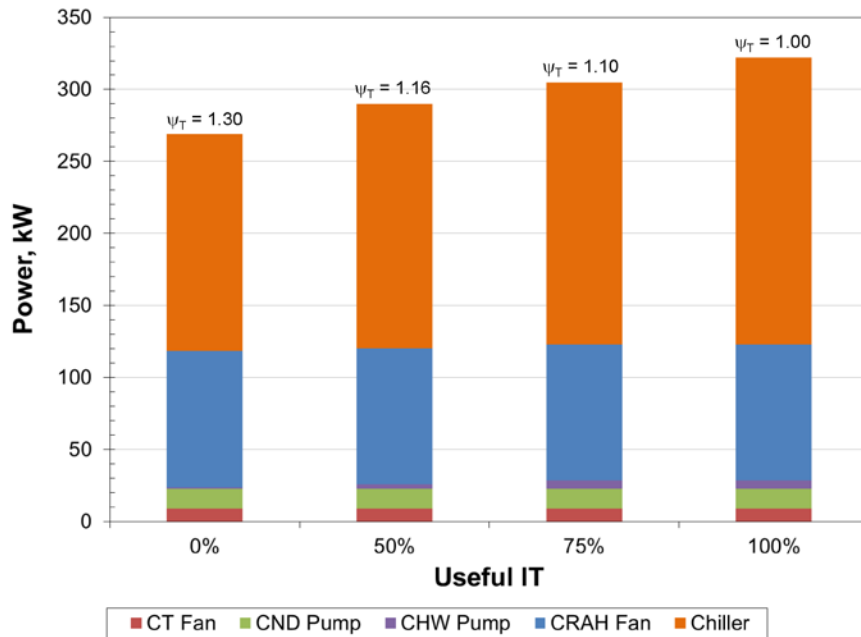


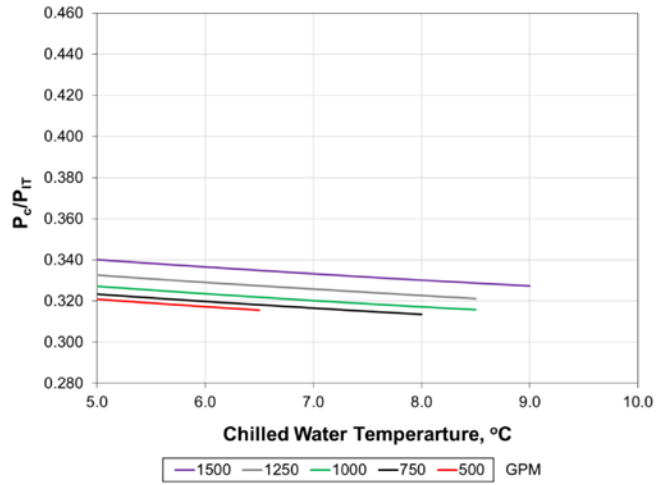
Figure 6.19 - Component-by-Component Breakdown for Baseline Scenarios

6.4.2 100% Useful IT Power

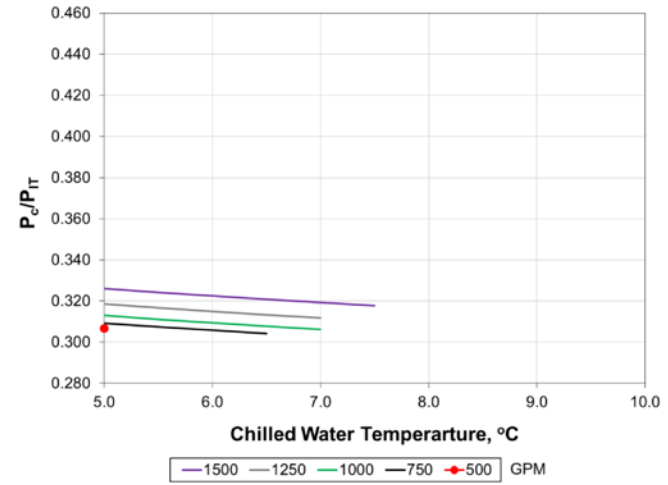
With the baseline results established, the next step is to optimize the data center at each of the IT load levels. Figure 6.20 plots the normalized cooling power versus chilled water temperature for the case of 100% useful IT, over a range of chilled water flow rates.

Unlike in the baseline scenarios, where the CRAH's airflow was fixed, the CRAH's airflow rate is now used as an additional variable in the optimization (i.e., by changing ψ_T). Each of the charts in Figure 6.20 shows the optimization for a different value of ψ_T , from 1.00 – 0.85. Again, the plots show the monotonically decreasing nature of the chiller's power consumption with increasing chilled water temperature. In a given plot,

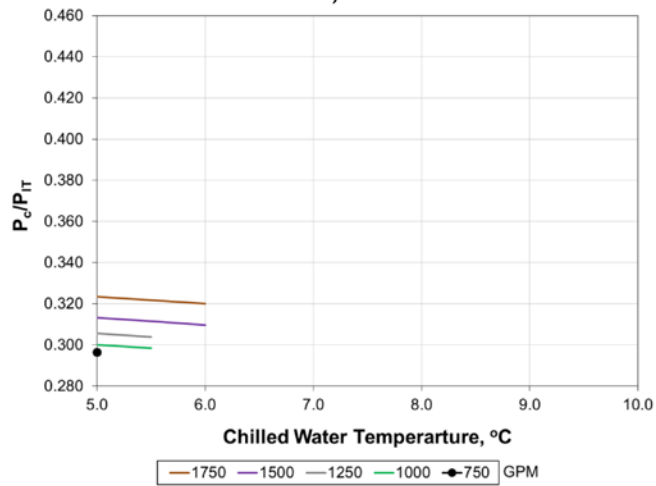
for example Figure 6.20a, since ψ_T is fixed, the CRAH's fan power consumption is also fixed. Figure 6.21 shows the minimum points from each of the plots on Figure 6.20, as a function of the chilled water flow rate for the range of ψ_T considered. The curves for $\psi_T = 1.00$ and 0.95 exhibit a mathematical minimum of power consumption, whereas, the curves for $\psi_T = 0.90$ and 0.85 are terminated at the low flow rate end. The termination of the curves occurs when the lowest achievable chilled water temperature reaches a self-imposed constraint of 5°C (41°F). While this constraint is self-imposed, it is based on physical understanding of the system. Obviously, since these are chilled water based CRAH units, it is impossible to provide water lower than its freezing point of 0°C (32°F). In addition, some temperature differential will exist in the evaporator heat exchanger, between the chilled water and the boiling refrigerant, which is typically a few degrees Celsius. In these instances, the chiller's evaporator would be operated at a saturation temperature near 1.7°C (35°F) to avoid spots of freezing on the heat exchanger coil.



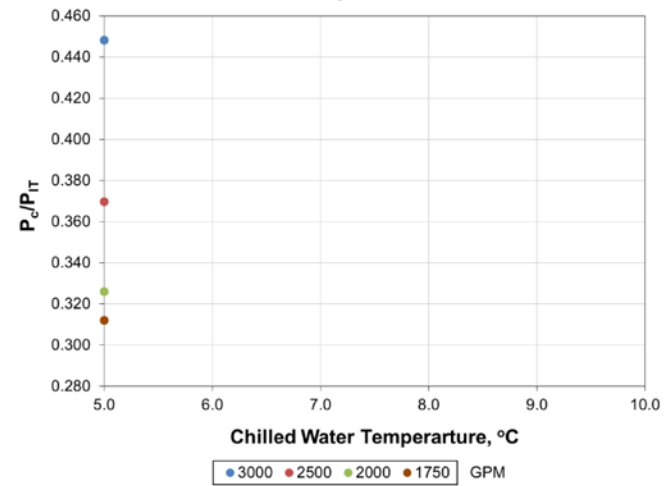
a)



b)



c)



d)

Figure 6.20 - Energy Optimization Details of the 100% Useful IT

a) $\psi_T = 1.00$, b) $\psi_T = 0.95$, c) $\psi_T = 0.90$ and d) $\psi_T = 0.85$

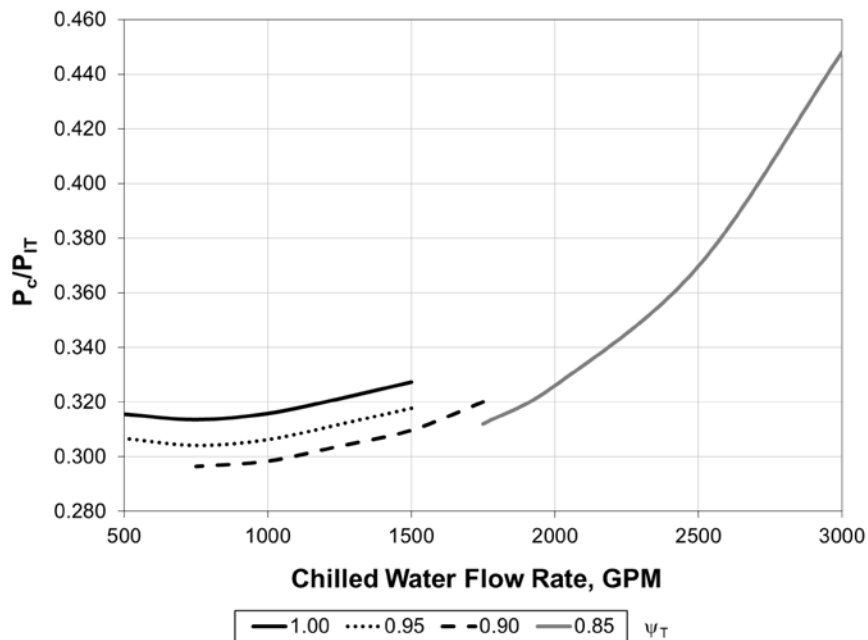
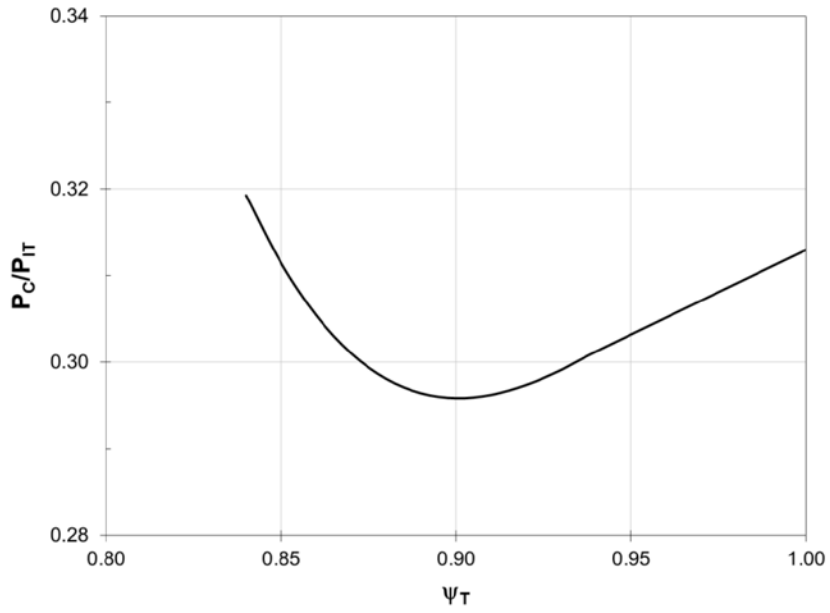


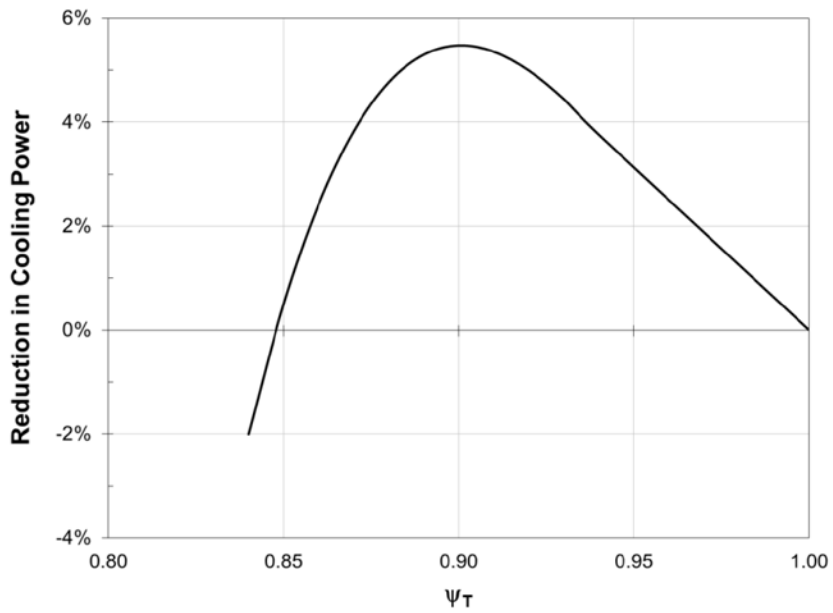
Figure 6.21 - Variation in Optimum Cooling Power with Chilled Water Flow Rate for 100% Useful IT Scenario

Figure 6.22 shows the results of the optimization for the 100% useful IT load scenario. The results are presented in two ways. Figure 6.22a plots the normalized cooling power consumption versus ψ_T , whereas Figure 6.22b plots the reduction in cooling power from the baseline case presented in Section 6.4.1. Operating the data center at a tile flow rate ratio near $\psi_T = 0.9$ results in the minimum cooling power consumption, with about a 6% savings in the cooling power. Figure 6. shows the breakdown of the power consumption of each component at each value of ψ_T . The minimum point occurs at $\psi_T = 0.90$ for several reasons. Although the optimum chilled water temperature at $\psi_T = 0.90$ is 5.0°C and at $\psi_T = 1.00$ is 8.0°C , this difference results in only a 10 kW (5% of chiller power) reduction in chiller power, whereas the CRAH fans' power is reduced by 25 kW (26% of

the CRAH fan power). The sharp increase in power consumption going from $\psi_T = 0.9$ to $\psi_T = 0.85$ is attributed to the increase in chilled water pumping power required to remove the given load, even at the minimum chilled water temperature of 5°C. To pump the required 1750 GPM of chilled water at $\psi_T = 0.85$ required an increase in pumping power of 27.4 kW. This increase is only partially offset by a decrease of 10.5 kW in the CRAH fans' power and a 0.9 kW decrease in refrigeration power as the flow is decreased from $\psi_T = 0.90$ to 0.85. It should be noted that the decrease in refrigeration power is attributed to the fact the total heat load of the data center is reduced when the CRAH's fans are operated at $\psi_T = 0.85$ instead of 0.90, and not because the chilled water temperature was reduced, since the optimum chilled water temperature in both cases was 5.0°C.



a)



b)

Figure 6.22 - Optimization Results for 100% Useful IT

a) Normalized Cooling Power and b) Reduction in Cooling Power from Baseline

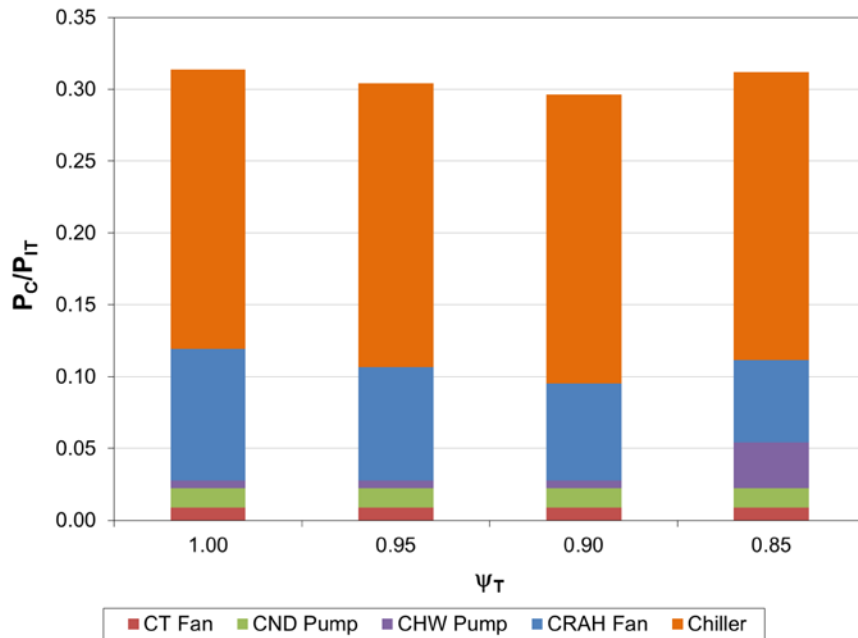
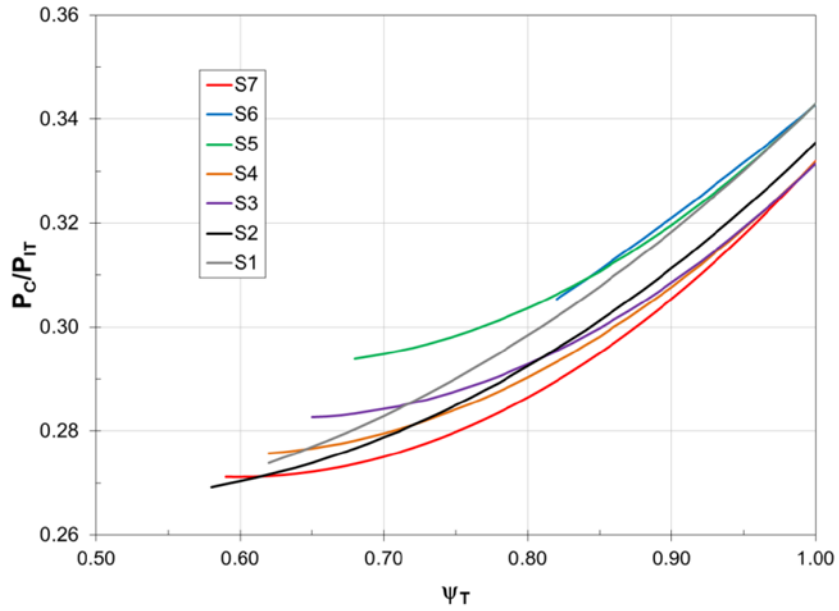


Figure 6.23 - Component-by-Component Breakdown of Cooling Power for 100% Useful IT

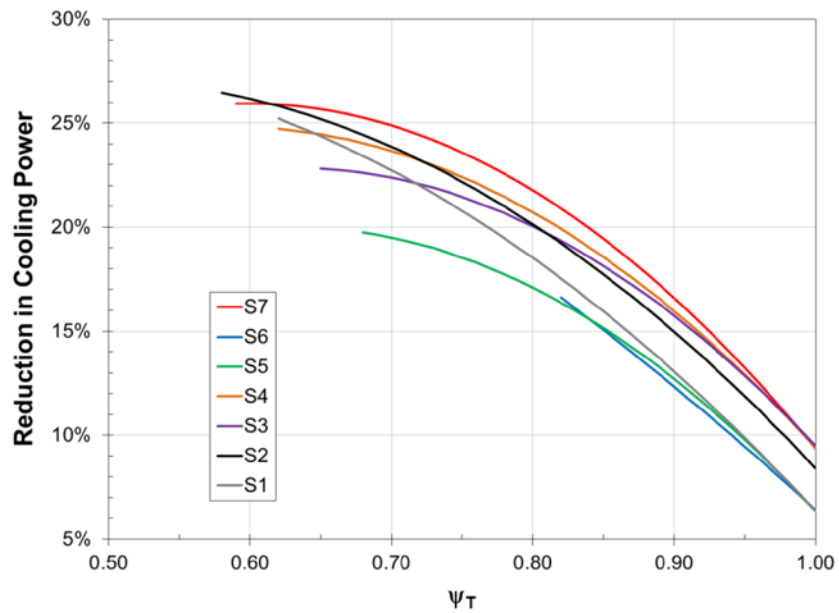
6.4.3 75% Useful IT Power

The energy optimization for each of the load placement scenarios described in Section 6.2.1.2 is considered here. The optimization procedure is identical to that done for the 100% useful IT load case; however, for brevity, the intermediate steps will be omitted and only the final optimization results are presented. Figure 6.24a and 6.24b provide the normalized power consumption and reduction in cooling power consumption results, respectively, as a function of ψ_T , for each of the IT load scenarios (S1 – S7). Clearly, the results in Tables 6.3 – 6.9, show that the IT load placement in the data center has some effect on the changing the airflow and temperature field, as evident by the differences in maximum chassis' inlet temperature for each of the placement scenarios. A consequence

of this is that there is no guarantee that a given scenario will be able to operate at all values of ψ_T , since this could result in a supply air temperature that is lower than the self-imposed 5.0°C chilled water temperature constraint. This is evident by the difference in terminal points of each of the scenarios. For example, Scenario 6 (uniform placement) reaches a terminal point at $\psi_T = 0.82$ with a savings in cooling power of 16%, compared to the 75% useful IT baseline; whereas, Scenario 2 (turn off the upper chassis) extends to a $\psi_T = 0.58$ with a savings in cooling power of 26%. Clearly, there is a reduction in fan power as ψ_T is reduced from 0.82 to 0.58. This possible reduction in ψ_T between scenarios 6 and 2 is attributed to a more favorable temperature field and airflow recirculation pattern because of a more thermally aware load placement strategy. Scenario 6 also proves to be an inferior load placement option because it does not take advantage of the increased inlet air temperature possible by placing chassis in *idle* operation, since all chassis are operated at part load.



a)



b)

Figure 6.24 - Optimization Results for Different Scenarios at 75% Useful IT
a) Normalized Cooling Power and b) Reduction in Cooling Power from Baseline

If we consider all the load placement scenarios, the two that perform the poorest are the uniform workload (Scenario 6) and the random placement (Scenario 5), providing confidence that there is a benefit to distributing the workload using knowledge of the thermal environment, as was done in the other scenarios. At their optimum points, none of the five thermally aware scenarios (S1-S4, S7) provides a significant advantage, with only a 3.5% reduction in cooling power difference between Scenarios 3 and 7. However, Scenario 7 appears to perform the best across the spectrum of ψ_T ⁵. Scenario 7 also provides several practical implementation benefits for operating data centers. To recall, Scenario 7 turned to *idle* the chassis that had the highest inlet temperature, until the required IT load was reached. In operational data centers, this is relatively easy to implement since most IT equipment available on the market has one or more built in thermistors, which can be connected to the data center's management and control system. In addition, shutting to *idle* the hottest servers requires no a priori knowledge of the data center layout, and is relatively independent of choices made about where the equipment is placed in the data center.

Figure 6.25 provides the breakdown of the power at a near-optimum point for each scenario. The optimum chilled water temperature and flow rate are 6.5/6.0/6.0/8.0/5.0/6.0/5.5°C and 500/500/500/500/1000/500/500 GPM, for each of the

⁵ At higher ψ_T , Scenario 2 performs slightly better than Scenario 7; however, the less than 1% difference in cooling power is well within the range of modeling error. Therefore, the results are considered essentially the same.

seven scenarios, respectively. Interestingly, for all scenarios, the optimum chilled water temperature is low because there are more saving possible by operating at a lower ψ_T than operating at a higher chilled water temperature. It needs to be emphasized that although the new ASHRAE environmental guidelines called for higher rack inlet temperatures to improve energy efficiency, the path to realizing this is not through reducing the chiller power alone. More emphasis needs to be placed on reducing the equally important CRAH fan's power consumption in order to arrive at the minimum energy consumption.

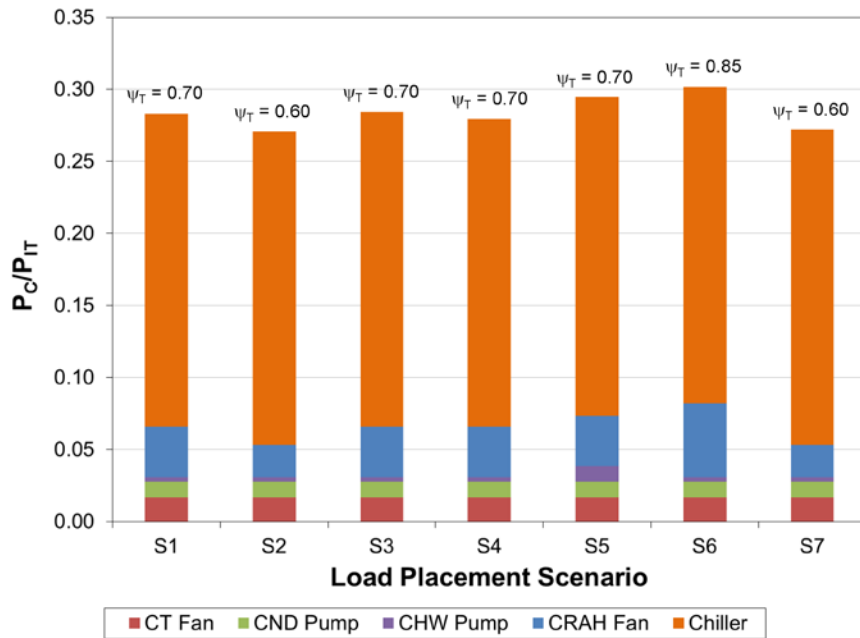
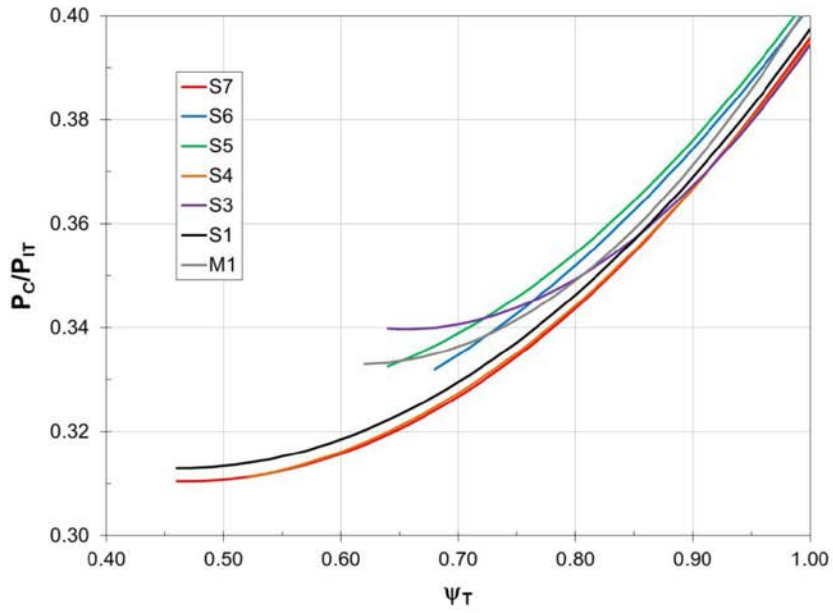


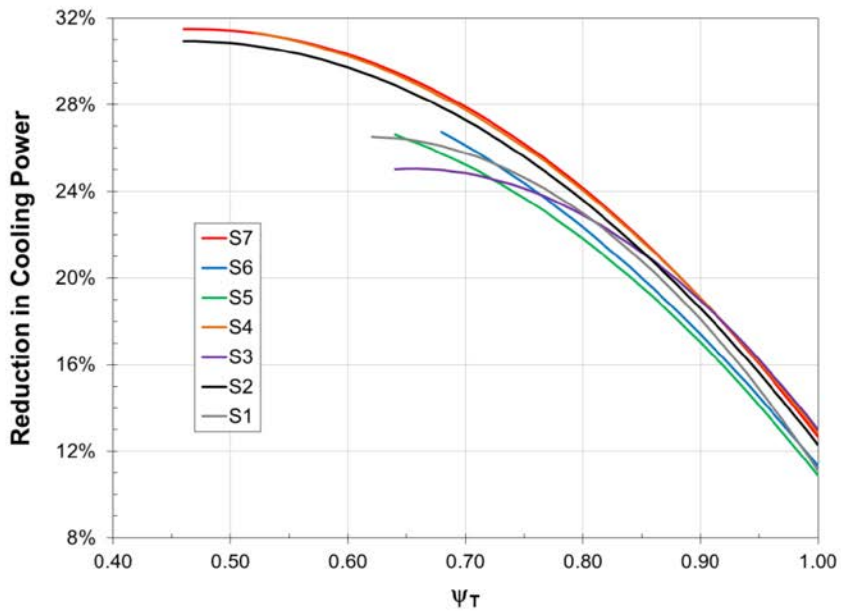
Figure 6.25 - Component-by-Component Breakdown of Cooling Power for 75% Useful IT

6.4.4 50% Useful IT Power

In a similar manner as the 75% useful IT scenarios, this section presents the results for the 50% useful IT load scenarios. A robust load placement strategy should show significant savings at all IT load levels. Figure 6.26 provides the normalized energy consumption results and reduction in cooling power results versus ψ_T , for each of the load placement scenarios. The results show again, that Scenario 7 (turn to *idle* the hottest chassis) out-performs the other IT load placement scenarios in terms of reducing the cooling infrastructure's power consumption. At $\psi_T = 0.50$, Scenario 7 shows the potential for a 32% reduction in cooling power consumption. In addition, Scenario 7 exhibits larger savings at all IT utilization levels compared to the other six scenarios, as was the case at 75% useful IT. The potential savings of Scenario 7 follow identical rationale as was given for the 75% useful IT scenarios. Placing the hottest chassis in an *idle* state allows for a significantly reduced airflow because those chassis are allowed to experience a higher inlet air temperature. Reducing ψ_T from 1.00 to 0.50 results in a 59 kW reduction in CRAH fan power, whereas, there is only a 16 kW reduction in refrigeration power as the chilled water temperature is reduced from 18.5°C at $\psi_T = 1.00$ to 7.0°C at $\psi_T = 0.50$.



a)



b)

Figure 6.26 - Optimization Results for Different Scenarios at 50% Useful IT
a) Normalized Cooling Power and b) Reduction in Cooling Power from Baseline

Figure 6.27 provides the component-by-component breakdown at the near-optimum point for each load placement scenario. The optimum chilled water temperature and chilled water flow rate are 10.0/7.5/5.5/10.0/8.0/10.0/7.0°C and 500/500/500/250/500/500/250 GPM, for each of the seven scenarios. Overall, the 50% useful IT scenarios have higher optimum chilled water temperatures than the 75% useful IT scenarios. The simple reason for this is that even though the non-dimensional ψ_T is used to relate to the CRAH's airflow rate, the magnitude of the airflow rate is greater for 75% useful IT than 50% useful IT, at a given value of ψ_T . Therefore, the magnitude of the energy savings as ψ_T is reduced is not as large as it was in the 75% useful IT scenarios and slightly more emphasis needs to be placed on using higher temperature chilled water. The results of Figure 6.27 shows that even though the optimum chilled water temperatures range from 5.5°C – 10.0°C, depending on scenario, the chiller power is almost identical across the board. Again, stressing that the reduction in fan power at low ψ_T dominates the energy optimization and should be the focus during actual operation.

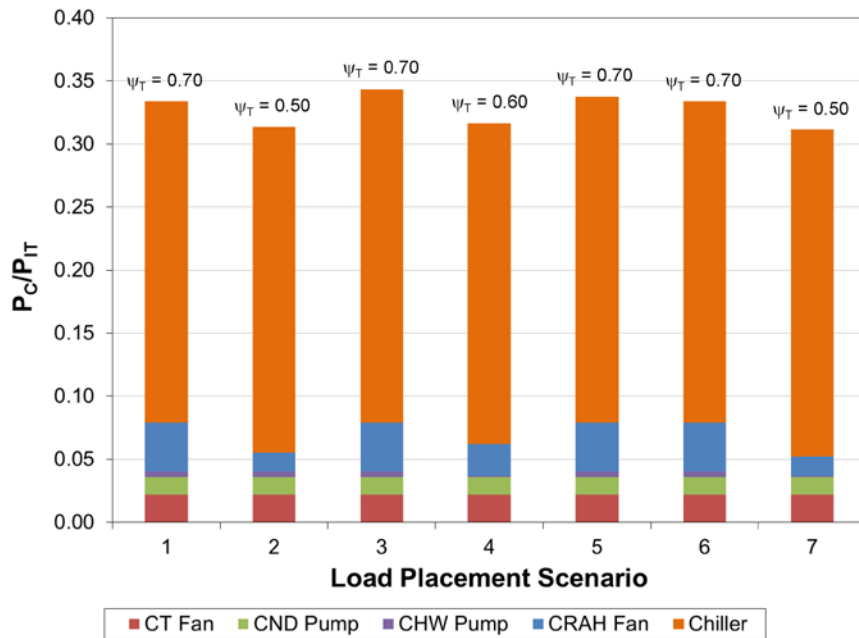


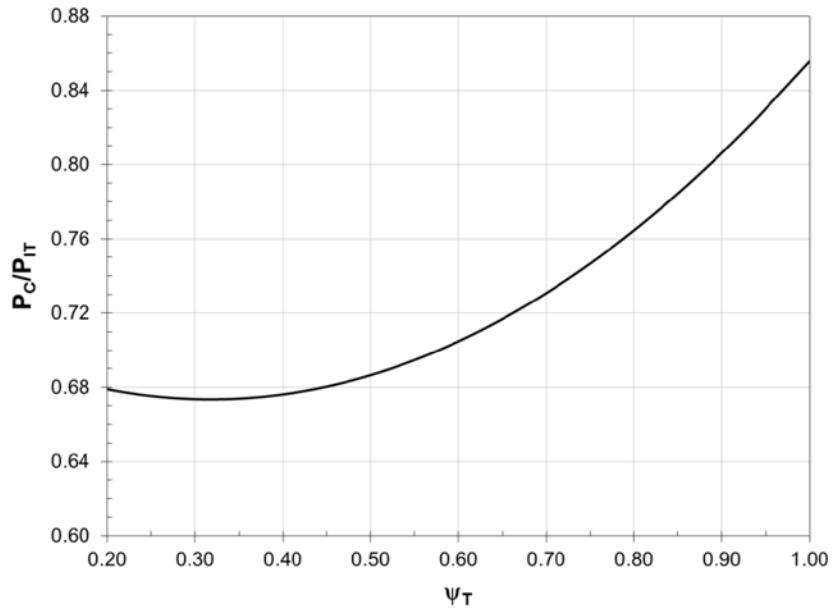
Figure 6.27 - Component-by-component Breakdown of Cooling Power for 50% Useful IT

6.4.5 0% Useful IT Power

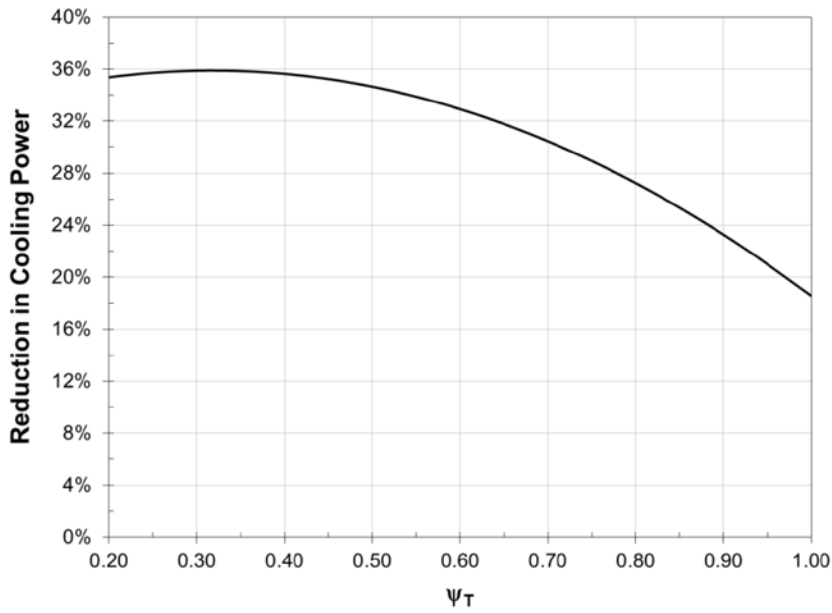
While not a practical operating scenario from the viewpoint of IT, the 0% useful IT case provides a terminal point for the operation of the cooling infrastructure. The 0% useful IT scenario provides the smallest heat removal for the cooling infrastructure and therefore, is the worst mismatch for the performance of the cooling equipment. However, at the significantly reduced heat load, it should be anticipated that the CRAH airflow could be substantially reduced. This point is shown in Figure 6.28, which plots the cooling infrastructure's power consumption versus ψ_T . The optimum operation of the cooling infrastructure in this case occurs around $\psi_T = 0.40$. Unlike many of the previous scenarios, which were terminated because of the self-imposed chilled water temperature

constraint, the 0% useful IT displays a mathematical minimum in the power consumption, since the reduction in refrigeration power consumption is not offset by the small reduction in fan power consumption at low ψ_T .

Figure 6.29 provides the component-by-component breakdown in power consumption as ψ_T is reduced from 1.00 to 0.40. The optimum chilled water temperature and flow rate at these four points are 19.0/19.0/15.0/13.5°C and 125/125/75/75 GPM. As ψ_T is reduced from 0.40, the combination of the increase in chilled water pumping power and increased refrigeration power, results in the increase in overall power consumption, which is not offset by the decrease in CRAH fan power at low ψ_T .



a)



b)

Figure 6.28 - Optimization Results for 0% Useful IT

a) Normalized Cooling Power and b) Reduction in Cooling Power from Baseline

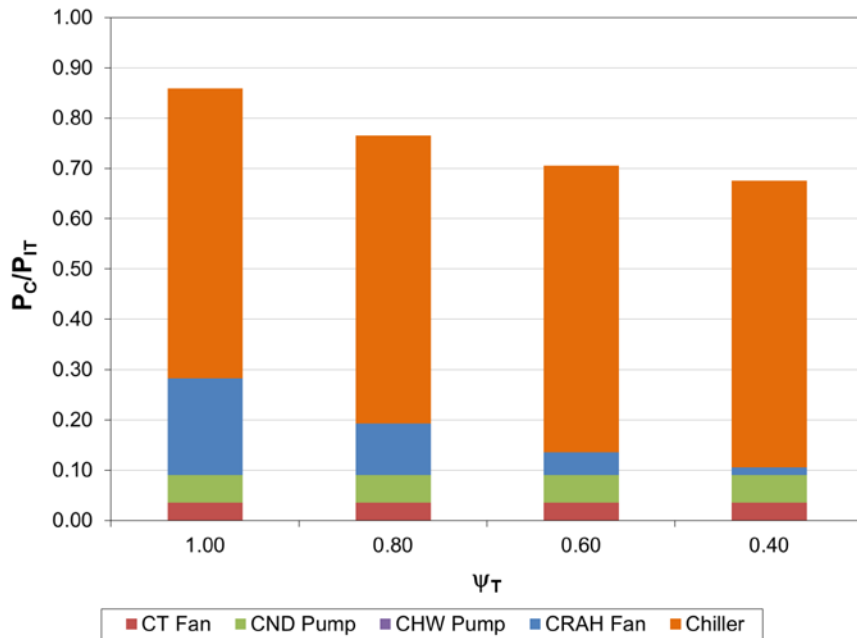


Figure 6.29 - Component-by-Component Breakdown of Cooling Power for 0% Useful IT

6.4.6 Verification of a Linear and One-to-One Temperature Field

The analysis presented thus far has assumed that once a solution to the temperature field of the data center was known, at a given supply air temperature, the solution at any other supply air temperature could be obtained by applying a δT_a change to the temperature field everywhere in the data center. In Section 4.5, it was shown that for data centers with a small Ar (i.e., lower buoyancy flows), this held true, at least for the temperatures at the most vulnerable chassis. However, for data centers with larger Ar , the assumption did not hold as well, and could result in a $2^\circ\text{C} - 3^\circ\text{C}$ temperature error at the most vulnerable chassis. In this section, the applicability of the δT_a change is considered for the load placement simulations presented thus far in Chapter 6.

The analysis is performed by first using the computed δT_a change in supply air temperature, as given in Tables 6.2 – 6.18, and performing an additional CFD simulation at that supply air temperature to determine the actual temperature field in the data center. If the constraints given by Equation 6.1 are not met identically, an additional CFD case is performed at the new δT_a , computed based on the new CFD case. This process is repeated iteratively until the constraints are identically met. This process was done for several of the scenarios performed during the load placement study. First, the 100% useful IT load scenario was studied, since it is expected, based on Equation 4.18b, that for racks with a higher temperature rise (caused by the airflow control algorithm at higher rack power), a higher data center Archimedes number would be obtained – especially at low airflow conditions. Table 6.20 provides the results of the study for the 100% useful IT load case for the ψ_T that result in reasonable operating conditions in the data center. The table also provides the resulting temperature at the CRAH units as a reference for a point outside of the cold aisle.

The results show that at the lowest airflow condition of $\psi_T = 0.85$, the largest error in the required supply air temperature is obtained, as expected because of the higher Ar . However, an error of only 0.2°C would result in no change in the cooling infrastructure's energy consumption because of the discrete nature of the control and optimization. The largest Ar obtainable in the 100% useful IT case was 1.04 at $\psi_T = 0.85$. The possibility of obtaining higher Ar is more prevalent at reduced load because of the ability to reduce the airflow rate significantly. Therefore, several of the optimum cases at 50% and 75% useful

IT are considered. Similar to the previous results, Table 6.21 provides the results of this analysis for these scenarios.

Table 6.20 - Verification of One-to-One Temperature Field for 100% Useful IT

ψ_T	Ar	T_a based on δT_a , °C	T_a based on CFD, °C	T_{x1} , °C		T_{x2} , °C	
				δT_a	CFD	δT_a	CFD
1.00	0.75	13.2	13.2	29.2	29.2	25.0	25.0
0.95	0.83	11.3	11.3	27.9	27.9	24.2	24.2
0.90	0.93	9.4	9.4	26.8	26.9	23.3	23.3
0.85	1.04	7.9	8.1	26.4	26.3	23.0	22.8

Table 6.21 - Verification of One-to-One Temperature Field for 50% and 75% Useful IT

Useful IT	ψ_T	Ar	T_a based on δT_a , °C	T_a based on CFD, °C	T_{x1} , °C		T_{x2} , °C	
					δT_a	CFD	δT_a	CFD
50	0.50	2.17	9.3	9.3	29.3	29.4	30.2	30.1
50	0.60	1.51	14.4	14.4	31.3	31.3	31.2	31.2
75	0.70	1.31	14.6	14.6	33.2	33.2	31.0	31.0

The preceding analysis showed that the assumption of a δT change in supply air temperature is an acceptable approach, at least for the scenarios considered throughout this dissertation that are used for developing load placement options in raised-floor, air-cooled data centers.

It remains beneficial to determine the error in the cooling infrastructure's power consumption if the δT assumption was not applicable. For this task, Scenario 7 with a useful IT load of 50% and $\psi_T = 0.60$ is considered. The error in the cooling infrastructure's power consumption is shown in Figure 6.30 for errors in the supply air temperature of $\pm 4.0^\circ\text{C}$. Positive values indicate conservative errors that predict a lower supply air temperature than required; therefore, resulting in lower energy savings. The figure shows that over the entire $\pm 4.0^\circ\text{C}$ range of supply air temperatures considered, only about a 3% change in the power consumption is realized. However, errors of 4°C are needed in the δT estimate of the supply air temperature in order to produce these errors in power consumption and errors this large were not seen in the CFD results. The analysis presented here confirms that the simple assumption of a δT change in the supply air temperature results in a δT change in the temperature field everywhere in the data center is reasonable for the magnitude of airflow and power levels considered throughout this dissertation.

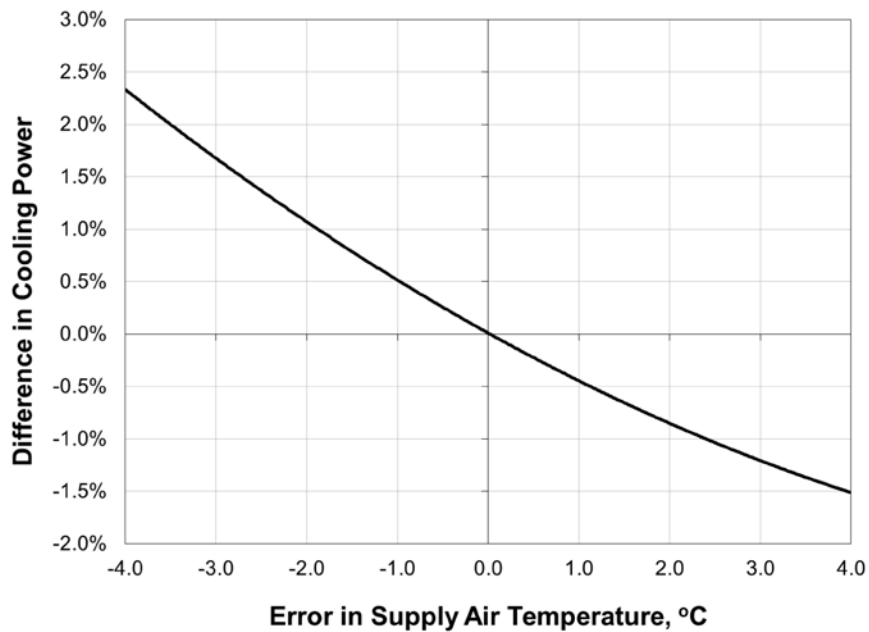


Figure 6.30 - Expected Error if the δT Assumption was not Valid

6.4.7 Section Conclusions

The work presented in this section showed the potential for reducing the cooling infrastructure's energy consumption through so-called thermally aware, energy optimized load placement. These strategies use prior or measured knowledge of the thermal environment to place the required workload efficiently amongst the servers in the data center. The notion of optimizing load placement based on the thermal environment is currently at the forefront of data center research. When studying this optimization problem, it is essential to consider the power consumed by all equipment in the cooling infrastructure. It is worth noting here that several studies on data center load placement have considered only the refrigeration portion of the total power consumption (Sharma,

Bash, et. al, 2005; Moore, Chase, et. al, 2005; Tang et. al, 2008) ignoring the equally important CRAH fan's power consumption. This may be justified in situations in which the refrigeration power consumption is dominant, as it would be when inefficient refrigeration systems are employed in very warm climates. In most other situations, exclusive focus on reducing refrigeration power consumption without regard to the power consumed in moving the cooling air would lead to sub-optimum, possibly misleading results, unless the CRAH flow rate (fan speed) is constant. In light of the results of this work, the load placement techniques previously developed would certainly be sub-optimal in terms of minimizing the total cooling power consumption.

Figure 6.31 provides a summary of the results obtained in this section. The figure plots the data center's total power consumption (cooling + IT) versus the data center's heat load fraction; all quantities are normalized by the maximum IT load of 1024 kW. The figure compares the minimum power consumption point for each IT load scenarios and shows relatively little advantage of using one load placement scenario over the others, when considering reductions in the total power consumption of the data center, as long as consideration is given to optimizing the cooling infrastructure. However, this analysis showed that the most practical and energy efficient scenario for implementation in operating data centers was to turn to *idle* the chassis starting with those that have the highest inlet temperature. This scenario took advantage of the expanded environmental guidelines for IT equipment, which allows for higher inlet temperatures for equipment that is *idle*. A key finding of this work is that significantly more savings in the cooling infrastructure's power consumption is seen by reducing the CRAH's airflow rate than by

providing higher chilled water temperatures from the refrigeration units. Therefore, the path to realizing the industry's goal of higher IT equipment inlet temperatures should be through a reduction in airflow and not necessarily through higher CRAH supply air temperatures.

Compared to the baseline cases, which do not consider the optimization of the cooling infrastructure, the results illustrated the potential for a 1.3% - 18.3% reduction in the total data center power and 5.5% - 35.7% reduction in the cooling power, depending on IT utilization level, by considering the holistic optimization of the data center. Figure 6.31 also shows the line of constant $(P_C + P_{IT})/P_{IT}^o = 1.34$, which is the value obtained at the 100% useful IT load case. This line is significant because it highlights the degradation of the cooling equipment's performance at part-load. For example, consider the situation where the 1024 kW data center is made up of four identical 256 kW data centers, each with a $(P_C + P_{IT})/P_{IT}^o = 1.34$. If only 50% of the useful IT was needed, two of the four data centers would be turned off, which would have no effect on the performance of the remaining two data centers. Therefore, $(P_C + P_{IT})/P_{IT}^o$ would remain 1.34. However, this is not that case when the cooling equipment has off-design performance that is different from its design rating. The figure shows that the performance of the equipment improves slightly between 75% - 90% useful IT and degrades significantly between 0% - 20% useful IT.

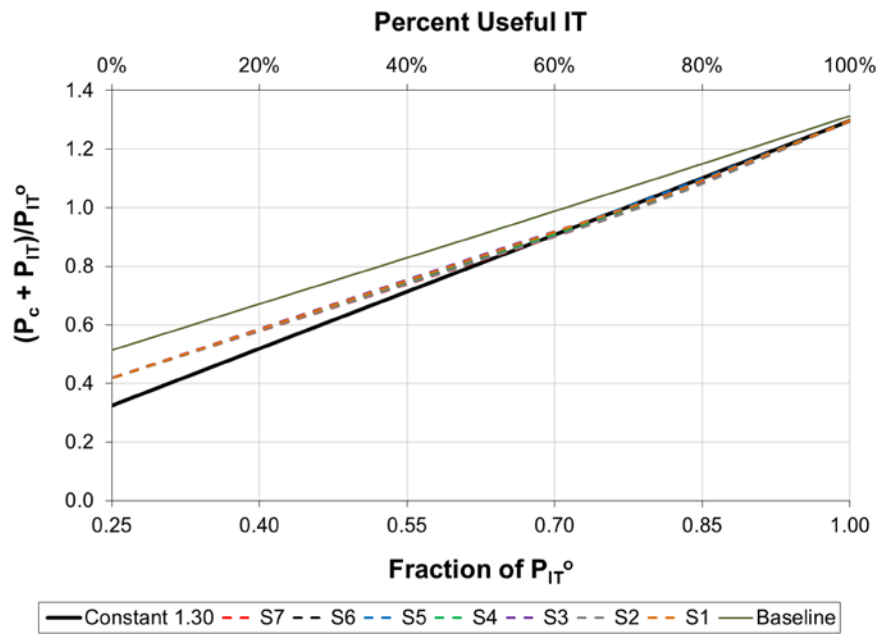


Figure 6.31 - Summary of Load Placement Scenario Analysis

7 Further Analysis of Thermally Aware, Energy-based Load Placement

The preceding chapter studied so-called thermally aware, energy-optimized load placement in order to develop physics-based heuristic guidelines for facilitating efficient load placement in data centers. Several realistic load placement options were studied to surmise methods for reducing the overall power consumption of the data center and cooling infrastructure, while maintaining the reliability of the IT equipment. The work highlighted the benefits of a robust load placement technique that uses real-time measurements of the chassis' inlet temperature to remove load from those chassis with the highest inlet temperature. This chapter extends that work to look at other IT load utilization levels in order to develop rules for applying optimized load placement in operating data centers. Additionally, a more detailed analysis of workload placement is completed, which looks at the effect of ambient conditions, placement scenarios for increasing IT load, chassis operation, and dynamic vs. static load placement. Lastly, thermally aware, energy optimized load placement techniques in enclosed aisle data centers are considered.

7.1 Further Assessment of Optimum IT Load Placement

The POD model developed in Section 6.3 allowed us to generate an infinite number of design configurations for the optimum load placement scenario described in Chapter 6.

The results presented in Section 6.4 are expanded to assess the performance of the IT

placement scenario, where the chassis are shut to *idle* beginning with the hottest (Scenario 7).

7.1.1 Optimum Load Placement for Reducing IT Load by Turning off the Hottest Chassis

The POD reduced order model was used to generate temperature data at four additional IT utilization levels: 87.5%, 62.5%, 37.5% and 12.5%, for a range of ψ_T . Figure 7.1 plots the normalized cooling power consumption for each of these IT levels along with isotherms of the CRAH's supply air ranging from 10°C to 20°C. At lower IT utilization levels, there is only a slight savings in cooling power if the data center was operated at 20°C supply air temperature instead of 10°C. For an IT utilization level of 25%, these savings are only 7% of the cooling power. However, at higher utilization levels, the increase in energy consumption while operating at 20°C can be large – greater than 20% in the 75% IT utilization case, for example. Therefore, at lower utilization there is more freedom in operation, since the energy consumption around the optimum is relatively flat. At higher utilization, more care must be taken to assure the data center is operating at its optimum efficiency. This result is certainly counterintuitive because the expectation is that a higher supply air temperature would result in lower power consumption; however, the large savings in fan power consumption at lower ψ_T trump the savings from higher supply air temperature. The conclusion of a substantial portion of this dissertation is the need for a holistic accounting of the infrastructure's energy consumption. The results of any optimization of the data center's cooling infrastructure would certainly be sub-

optimal without such a consideration. Any optimization based solely on reducing the refrigeration power consumption is flawed if the results are to be applied to data centers with central vapor-compression chilled water plants, such as those considered in this work. In these cases, the precise control of the chilled water flow and temperature is secondary to the need for variable speed control on the CRAH units. Only in the case of inefficient refrigeration systems, such as in those data centers that employ air-cooled direct-expansion CRAC units, would a refrigeration only analysis be justified.

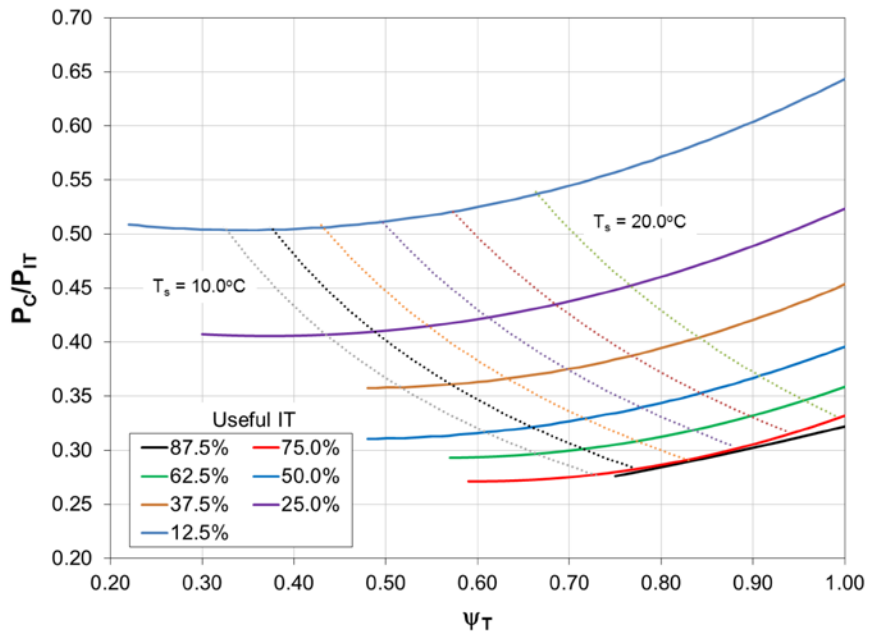


Figure 7.1 - Normalized Cooling Power for Thermally Aware, Energy-based Load Placement

Similarly, Figure 7.2 plots the results as a three-dimensional surface of useful IT, ψ_T and the CRAH's supply air temperature. These results are useful for selecting a near-optimum operating point for the infrastructure at a needed IT utilization level. This result will

prove useful in the context of the practical implementation of thermally aware, energy-based load placement, which is discussed in Section 7.3.

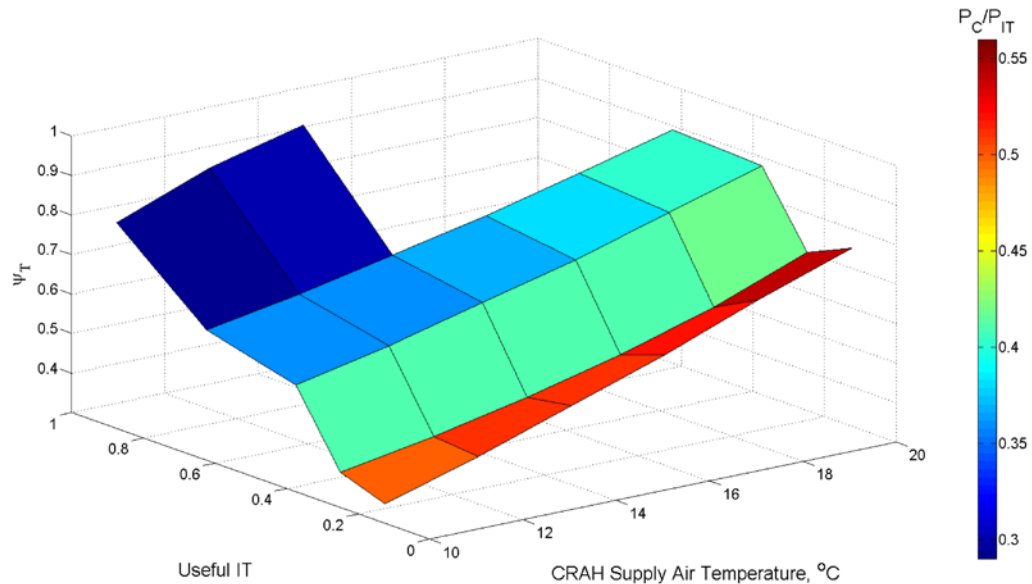


Figure 7.2 - Data Center Operation Map for Thermally Aware, Energy-based Load Placement

7.1.2 Effect of Ambient Conditions

The results presented thus far were for an ambient wet-bulb temperature of 21.9°C.

Figure 7.3 presents the normalized energy consumption results for three different ambient wet-bulb temperatures: 13.7°C, 21.9°C and 30.2°C. As expected, higher ambient temperatures result in larger power consumption because of the degraded performance of the chiller and the required increase in cooling tower fan power. The minimum point in the power consumption curves around 80% useful IT reflects the off-design nature of the cooling infrastructure's power consumption. In this case, the chiller has a rated COP of

6.2 at a part load ratio around 0.80. This characteristic is common in the design of the infrastructure because of the large amount of time it is operated off-design.

Figure 7.4 provides the optimum ψ_T at each of the ambient wet-bulb temperatures for the full range of IT utilization. At higher ambient temperatures, the optimum operation shifts to a higher value of ψ_T . For example, at a 50% IT utilization level, the optimum CRAH operation is at a $\psi_T = 0.51$ and $\psi_T = 0.63$ for wet-bulb temperatures of 13.7°C and 30.2°C, respectively. This result reflects the increased focus on reducing the chiller power consumption at higher ambient temperatures, in addition to reducing the CRAH fan power. Although more focus is placed on the refrigeration power at higher ambient temperatures, the savings in CRAH fan power are still significant and cannot be ignored.

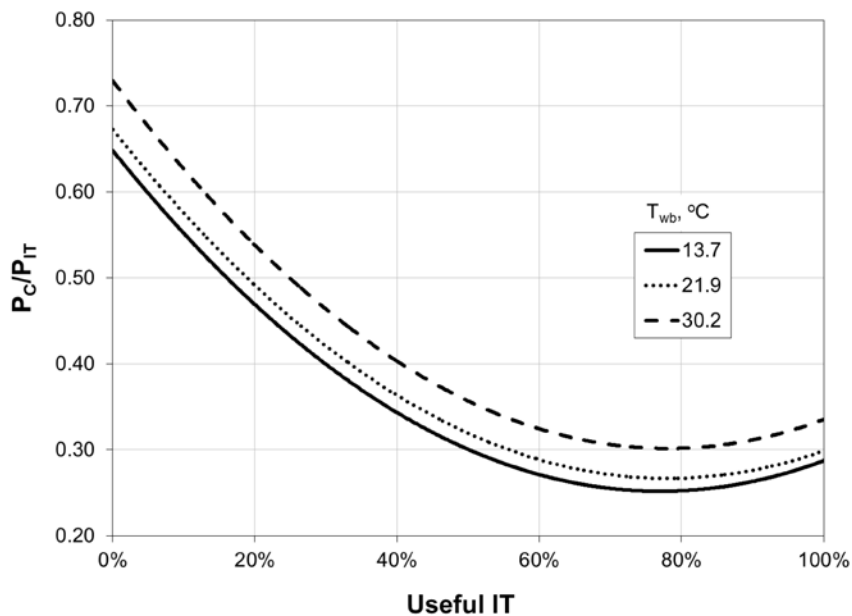


Figure 7.3 - Normalized Cooling Power at Different Ambient Temperatures for Thermally Aware, Energy-based Load Placement

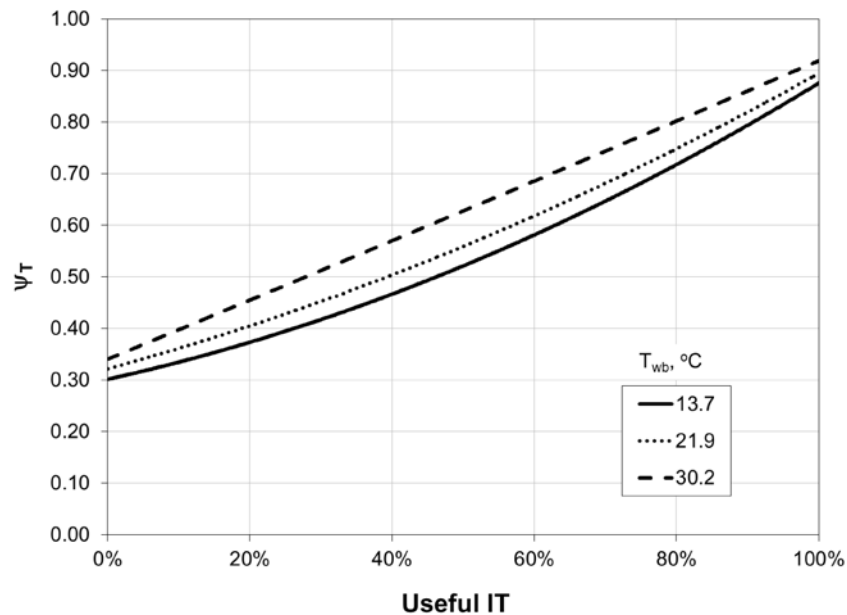


Figure 7.4 - Optimum ψ_T at Different Ambient Temperatures for Thermally Aware, Energy-based Load Placement

7.1.3 Turning off Idle Chassis' Power and Airflow

All of the analysis so far has focused on load placement where the appropriate chassis are placed in an *idle* mode, in which the power and airflow are reduced. In cases where the instant availability of the IT equipment is not a concern, it might be anticipated that the IT equipment could be completely shut *off*, in terms of both the airflow and power. In this case, the data center would not experience the residual heat dissipation that would have to be removed in the *idle* state. In addition, the data center would require a reduced airflow rate, which would save fan power compared to a case of *idle* servers at the same ψ_T . We consider this scenario for the optimum load placement (Scenario 7). However, in this case the chassis' heat load and airflow are set to zero for the virtualized chassis. Because of the effect that the changes in airflow and heat load (through buoyancy) have on the

data center's temperature distribution, additional CFD cases, that completely *shut off* the appropriate chassis, are performed for 75% and 50% useful IT.

Figures 7.5 and 7.6 compare the case where the chassis are placed in an *idle* state (S7) with the case where the chassis are completely turned *off* (S7 – Shut Off), for 75% and 50% useful IT, respectively, for a range of ψ_T . The data is normalized by the useful IT load, which is 512 kW for 50% useful IT and 768 kW for 75% useful IT. As anticipated, the optimum point shifts towards higher ψ_T . However, the optimum points correspond to the same fan operating point. For example, in the 50% useful IT load case, the near-optimum points of $\psi_T = 0.50$ (for S7) and $\psi_T = 0.90$ (for S7- Shut Off) correspond to the same CRAH airflow rate of 15291 m³/h (9000 CFM). The savings in the infrastructure's power consumption between these cases comes from the 6°C increase in chilled water temperature made possible by the combination of a 20% reduction in heat load and a higher fraction of required chassis airflow being provided to the data center, when the chassis are turned *off* completely.

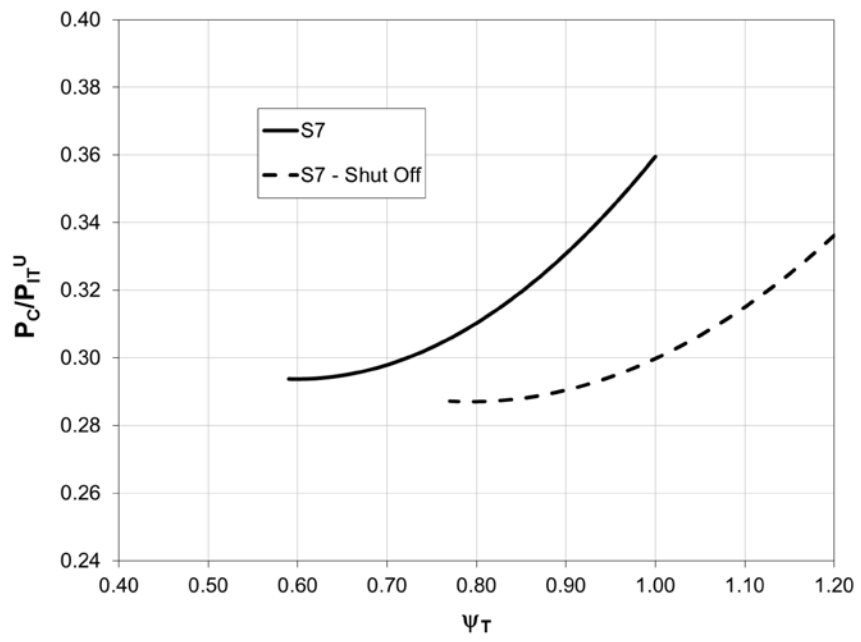


Figure 7.5 – Idle vs. Shut Off Control at 75% Useful IT

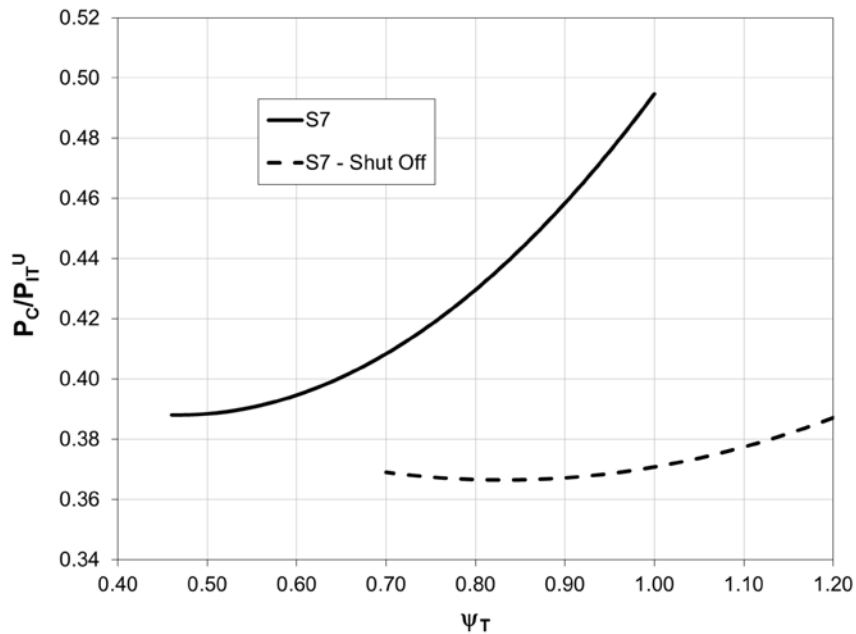


Figure 7.6 - Idle vs. Shut Off Control at 50% Useful IT

7.1.4 Optimum Workload Placement for Increasing IT Load by Turning on the Coldest Chassis

Each of the IT load placement scenarios considered so far were for when the data center's IT load was being reduced. The analysis found that an efficient method for reducing the IT load was to remove the load starting with the chassis that have the hottest inlet temperature. It is just as likely that the load in the data center would need to be increased and strategies for performing this efficiently need to be considered. Therefore, the complementary case (Scenario 8) is considered where the chassis are turned *on* beginning with those that have the coldest inlet temperature. Figures 7.7 and 7.8 compare Scenarios 7 and 8, for 75% and 50% useful IT, respectively, by turning *on* the 128 and 192 coldest servers in the optimized *idle* scenario. The results show that Scenario 7 and Scenario 8

compare well in terms of the optimum infrastructure cooling power over the entire range of ψ_T . Therefore, with the same ease that *turning off the hottest chassis* could be implemented in operating data centers using the chassis' internal temperature sensor, *turning on the coldest chassis* could be implemented, assuming the chassis remained in an *idle* state and access to the inlet temperature was maintained.

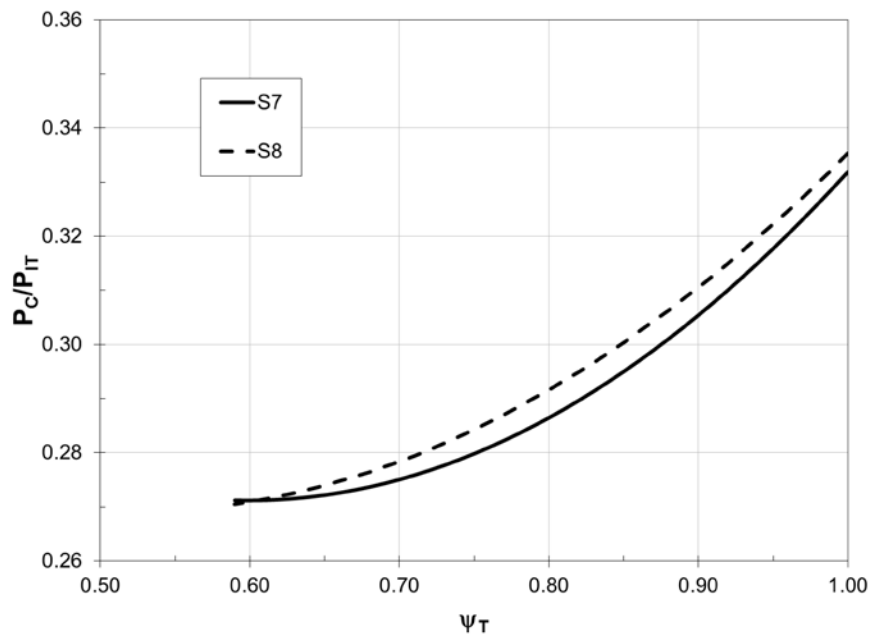


Figure 7.7 - Comparing Scenarios 7 and 8 for 75% Useful IT

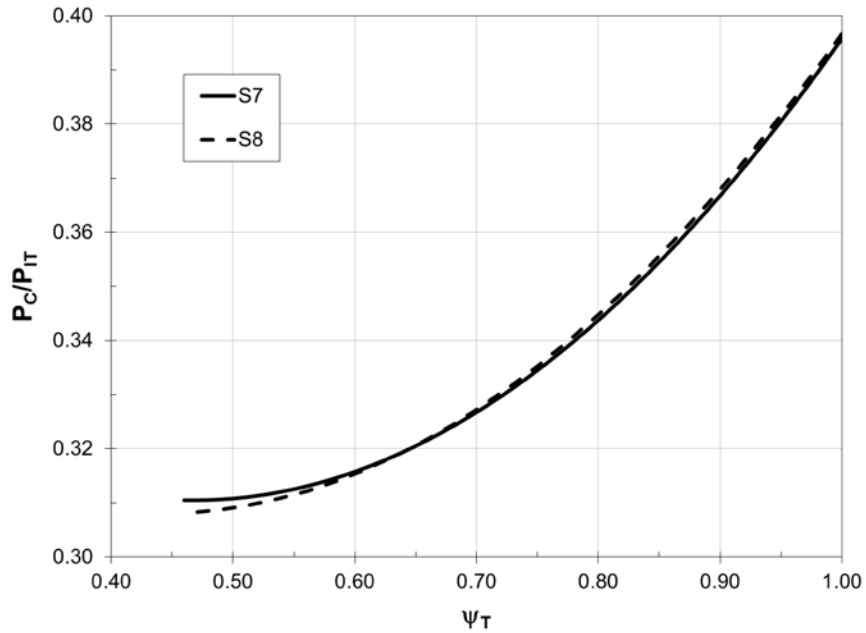


Figure 7.8 - Comparing Scenarios 7 and 8 for 50% Useful IT

7.1.5 Dynamic Analysis of IT Load Placement

Until now, IT load placement was done by turning to *idle* the chassis that were the hottest in the optimized 100% IT load scenario (or turning *on* the chassis with the coldest inlet temperature in the 0% useful IT load scenario). However, it may not always be the case that the data center will experience an increase or reduction in IT load from operation at 0% or 100% useful IT. For example, the data center might experience a succession of IT load reductions from 100% to 75% and then to 25%. From a case such as this arises an interesting problem if real-time temperature sensors are used to place the necessary IT load. If the data center needs to be operated at 50% useful IT, there is no guarantee that the 50% of the chassis that would be turned to *idle* from the 100% useful IT scenario would identically match the additional 25% of the chassis that would need to be turned to

idle had the data center been operating at 75% useful IT. Scenarios such as this give rise to two options for implementing load placement.

1. *Static*: develop an IT load placement algorithm during commissioning when the data center is operated at both 0% and 100% useful IT load. During this period, a queue of the hottest and coldest chassis in the data center is developed. When implementing load placement in the operating data center, the next chassis on the queue is selected until the required IT load is reached. This is identical to what has been done previously in this dissertation.
2. *Dynamic*: use the real-time sensor measurements and choose the hottest or coldest chassis based on the current operating state of the data center.

Figure 7.9 compares the normalized energy consumption for both the *static* and *dynamic* modes of operation for both the scenario of increasing (Scenario 8) and decreasing IT load (Scenario 7). For the *dynamic* analysis, the load was increased or decreased in steps of 25%. For example, for the scenario of *shutting idle the hottest chassis*, the first 25% were turned to *idle* based on the operation at 100% useful IT. Once a CFD solution was obtained, an additional 25% (of the chassis that were not already in an *idle* state) of the chassis are turned to *idle* based on the thermal state of the data center at 75% useful IT, therefore, resulting in 50% of the chassis being placed in an *idle* state. Lastly, to get to 25% useful IT, an additional 25% of the chassis were turned to *idle* from the 50% useful IT thermal state. The results show that there is little difference between the optimum power consumption for both the *static* and *dynamic* modes at the three useful IT load

levels: 25%, 50% and 75%. Although there is a negligible difference in the energy consumption for the two modes, the placement of the IT load is slightly different between them. Figure 7.10 compares the placement for each of the three IT levels for both the *static* and *dynamic* modes of operation. Obviously, the 75% IT load case is identical for the *static* and *dynamic* modes, since they are both derived from the 100% useful IT load case. However, for the case of 50% and 25% useful IT, it can be seen that the *dynamic* algorithm places the IT load on the chassis in the upper portion of the racks, as opposed to the *static* algorithm, which places the load on the lower chassis in the outer racks. The reason for this has to do with the very different airflow patterns of the two cases. For the *static* algorithm, the queue for load placement was developed with a $\psi_T = 0.90$; therefore, there was a significant amount of the required chassis' flow available to the upper chassis. However, in the *dynamic* mode, the data center had been energy optimized resulting in an operating point of $\psi_T = 0.70$ and therefore, a smaller fraction of the airflow is available to the upper chassis, which increases their temperature.

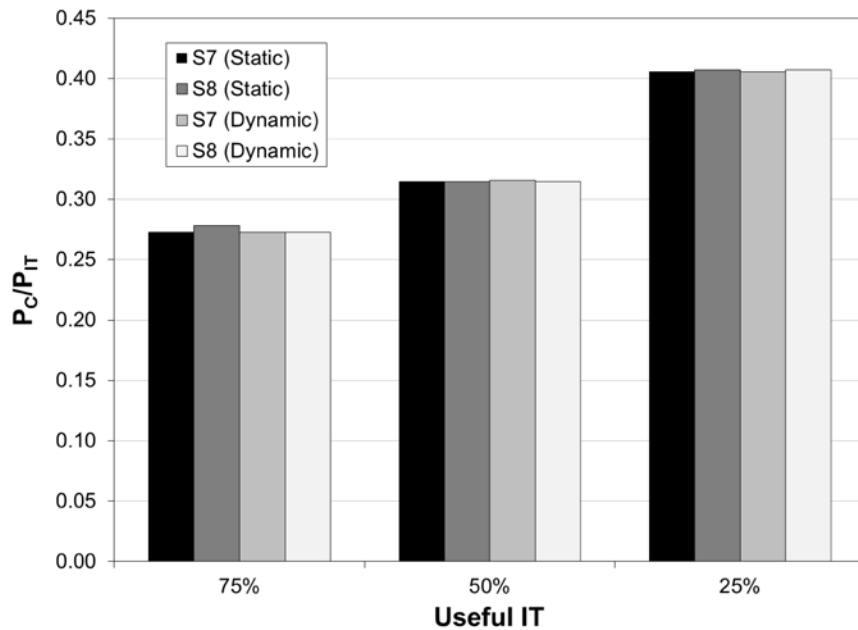


Figure 7.9 - Comparing the Energy Consumption of Dynamic vs. Static Load Placement

Although the difference in energy consumption of the different modes was negligible after optimizing the cooling infrastructure, there are a number of practical advantages for using the *dynamic* algorithm. Foremost, it is able to adapt to the changing thermal environment of the data center and eliminate the load from the actual chassis that are the most vulnerable to reliability issues from over-temperature. Secondly, if chassis with different airflow and/or power characteristics are installed after the original commissioning phase, there is no guarantee that the original placement queue would still correctly identify the hottest or coldest chassis. Therefore, after new equipment is installed, a new commissioning phase would be needed, which is likely impossible once the data center is operating.

	1	2	3	4	5	6	7	8
4	X	X	X	X	X	X	X	X
3								X
2								X
1								

	1	2	3	4	5	6	7	8
4	X	X	X	X	X	X	X	X
3								X
2								X
1								

1								
2								X
3	X							X
4	X						X	X
	9	10	11	12	13	14	15	16

1								
2								X
3	X							X
4	X						X	X
	9	10	11	12	13	14	15	16

75% - Scenario 7 - Static

75% - Scenario 7 - Dynamic

	1	2	3	4	5	6	7	8
4	X	X	X	X	X	X	X	X
3	X				X	X	X	X
2	X							X
1	X							X

	1	2	3	4	5	6	7	8
4	X	X	X	X	X	X	X	X
3	X	X	X	X	X	X	X	X
2	X							X
1								

1	X							X
2	X							X
3	X						X	X
4	X	X	X	X	X	X	X	X
	9	10	11	12	13	14	15	16

1								
2								X
3	X	X	X				X	X
4	X	X	X	X	X	X	X	X
	9	10	11	12	13	14	15	16

50% - Scenario 7 - Static

50% - Scenario 7 - Dynamic

	1	2	3	4	5	6	7	8
4	X	X	X	X	X	X	X	X
3	X	X	X	X	X	X	X	X
2	X	X				X	X	X
1	X	X						X

	1	2	3	4	5	6	7	8
4	X	X	X	X	X	X	X	X
3	X	X	X	X	X	X	X	X
2	X	X	X	X	X	X	X	X
1	X							X

1	X	X					X	X
2	X	X					X	X
3	X	X	X	X	X	X	X	X
4	X	X	X	X	X	X	X	X
	9	10	11	12	13	14	15	16

1	X							X
2	X					X	X	X
3	X	X	X	X	X	X	X	X
4	X	X	X	X	X	X	X	X
	9	10	11	12	13	14	15	16

25% - Scenario 7 - Static

25% - Scenario 7 - Dynamic

Figure 7.10 - Comparison of Load Placement Arrangement for Dynamic vs. Static Operation

7.2 Optimum Workload Placement in Enclosed Aisle Data Centers

The preceding section focused on the development of feature-based heuristics to facilitate optimized load placement in open aisle data centers. The complex airflow and temperature distribution in open aisle data centers provided a challenging problem for load placement. Enclosing the cold aisle provides a solution to the problem of a non-uniform temperature distribution at the rack's inlet. In addition, an enclosed cold aisle data center was shown to reduce the overall energy consumption compared to its open aisle counterpart. Therefore, it becomes logical to ask whether enclosed aisle data centers provide advantages in thermally aware, energy-based load placement.

This section focuses on developing load placement strategies in enclosed aisle data centers. The formidable problem of optimal workload placement, which has received increased interest by the IT industry and is the subject of numerous investigations, becomes much simpler, almost trivial, when the aisles are enclosed. This is because there is no preference among identical chassis in an enclosed aisle based on the thermal environment, which will be uniform in this case, making for a wide range of equivalent load placement possibilities within the data center, or at least those racks in the data center that share an enclosed cold aisle. To this end, developing optimal load placement scenarios for enclosed aisle data centers becomes a problem in optimizing the operation of the cooling infrastructure only. The importance of considering both the fan and refrigeration power consumption in enclosed aisle data centers has been shown previously and lead to the introduction of the CRAH bypass recirculation branch in

Section 3.1. The analysis of the enclosed aisle configuration, follows the methodology presented in Section 2.1, where a single mixed rack exhaust temperature is computed based on an energy balance of the enclosed cold aisle. A typical leakage fraction of $\Lambda_o = 0.25$ is assumed. The optimization results presented here use the thermo-hydraulic model and follow the procedure outlined in Section 6.2.2.

Figure 7.11 shows the computed infrastructure power consumption normalized by the IT power as a function of CRAH airflow fraction, as the useful IT load is varied from 0% to 100%, for an ambient dry-bulb temperature of 30°C and relative humidity of 50% ($T_{wb} = 21.9^\circ\text{C}$). Alternatively, Figure 7.12 provides the results as a reduction in cooling power from the conventional enclosed aisle, where all the flow is provided by the CRAH units and cooled to the redline temperature. The use of an optimized bypass recirculation configuration shows the potential of a 21% – 25% reduction in the cooling infrastructure’s power consumption, compared to the conventional enclosed aisle, depending on IT utilization. At all IT power levels, the optimum operating point corresponds to one with a significant degree of bypass recirculation, which is highlighted by the loci of minimum power points shown on Figure 7.12. The optimum operation corresponds to CRAH air supply fractions of 0.28 at 0% useful IT and increasing to 0.54 at 100% useful IT. However, the optimum bypass fraction appears to be relatively insensitive to deviations, as highlighted by the flatness of the power consumption curve around the optimum. Therefore, precisely locating the optimum CRAH airflow fraction becomes inconsequential and provides data center operators with some flexibility in selecting CRAH airflow and temperature settings. Figure 7.11 provides isotherms of the

CRAH's supply air ranging from 12°C to 20°C, which illustrates the wide range of possible operating scenarios in which significant energy savings are realized, compared to a conventional enclosed aisle.

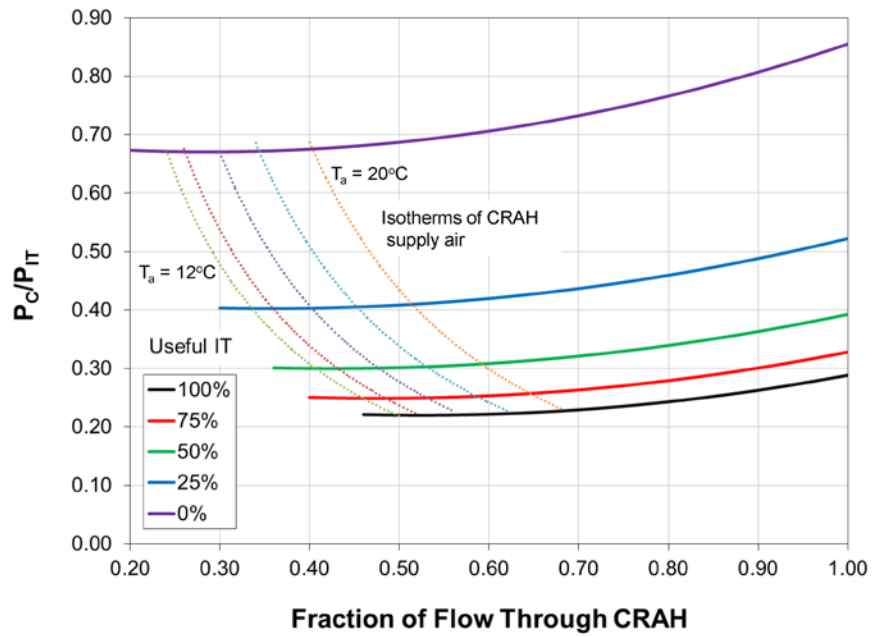


Figure 7.11 - Normalized Cooling Power Consumption with $\Delta T_m = 18.2^\circ\text{C}$ and $T_{wb} = 21.9^\circ\text{C}$

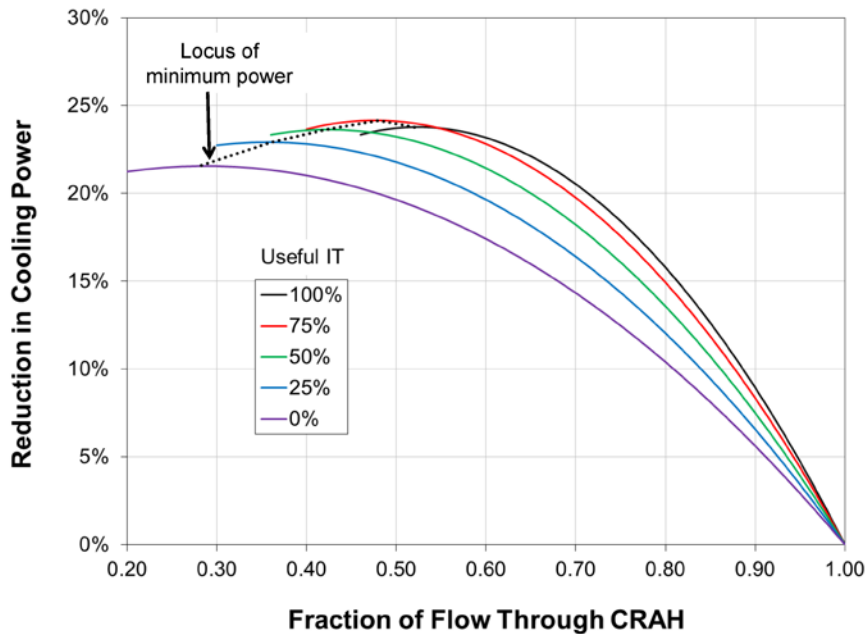


Figure 7.12 - Reduction in Cooling Power Consumption from a Conventional Enclosed Aisle with $\Delta T_m = 18.2^\circ\text{C}$ and $T_{wb} = 21.9^\circ\text{C}$

The increase in CRAH's air supply fraction, as the IT load increases, can be attributed to two facts. First, as the IT power is reduced, less power is needed for refrigeration, which places more emphasis on reducing the CRAH's fan power; however, this point alone does not fully account for the trend in light of the degraded performance of the refrigeration equipment at reduced load. Secondly, the chassis used in this analysis have a significantly lower temperature rise when *idle* than when *on*. Therefore, while the power consumption is reduced significantly (from 4 kW to 1 kW) the airflow rate at *idle* power is only 127.4 m³/h (75 CFM) less when *on*, again, indicating solutions that reduce the power required by the CRAH fans. Obviously, the latter is a characteristic of the IT equipment used in the data center and would vary based on manufacturer. However, the equipment selected

here is representative of typical high-volume equipment available from many manufacturers.

Despite the insensitivity of the CRAH's operation to the minimum power consumption, it is still useful to study the hypothetical optimum operation as given by the locus of minimum power consumption on Figure 7.12. Figure 7.13 provides the required CRAH supply air temperature, as a function of the CRAH airflow fraction, along with the locus of optimum CRAH's supply air temperature, at each IT load level. Clearly, when there is no bypass, the air is provided at the redline temperature of 27°C. As more bypass air is introduced, the CRAH must cool the fraction that passes over the CRAH's heat exchanger to a lower temperature before it is mixed with the bypass air. However, the optimum supply air temperature is significantly lower than the redline temperature, owing to the fact that the minimum power consumption occurs at a significant fraction of bypass air. These results indicated that the savings in fan power, due to the bypass air, outweighs the penalty in providing lower temperature chilled water.

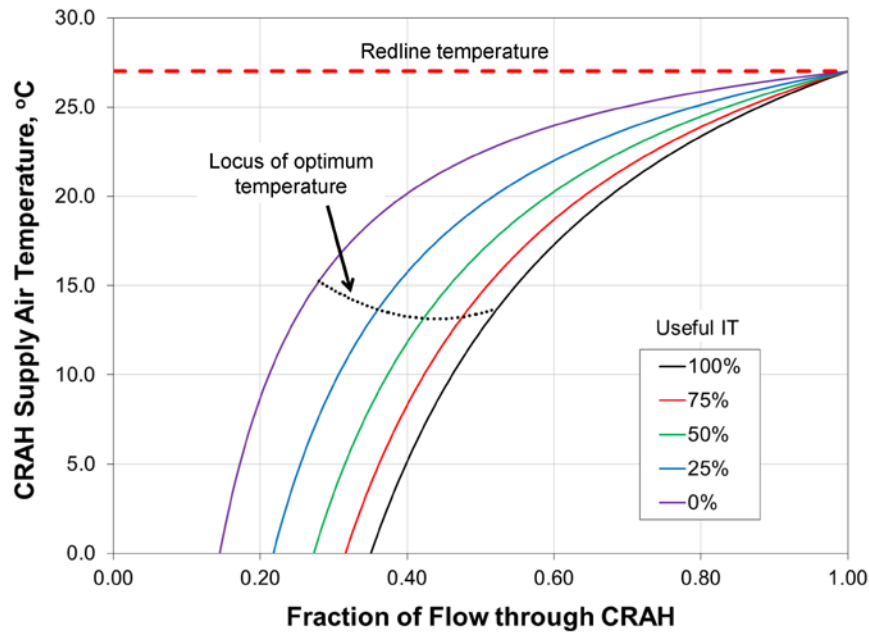


Figure 7.13 - CRAH Supply Air Temperature Requirements with $\Delta T_m = 18.2^\circ\text{C}$ and $T_{wb} = 21.9^\circ\text{C}$

7.2.1 Effect of Ambient Conditions

While the results in Figures 7.11 – 7.13 were for an ambient wet-bulb temperature of 21.9°C , the effect of ambient conditions on the infrastructure is given in Figure 7.14 for three different ambient wet-bulb temperatures (Note: Figures 7.14 – 7.17 provide the optimum points at each IT level). Similarly, Figure 7.15 shows the optimum CRAH's air supply fraction at each IT level for the three ambient temperatures. At higher ambient temperatures, the optimum amount of bypass air is reduced, as more emphasis is placed on reducing the refrigeration load, due to the reduced chiller COP. However, for all IT levels and ambient temperatures, savings are realized by implementing a bypass recirculation branch in an enclosed aisle data center, compared to the conventional

enclosed aisle configuration. Even for high ambient wet-bulb temperatures, greater than a 20% reduction in cooling power is possible, compared to a conventional enclosed aisle.

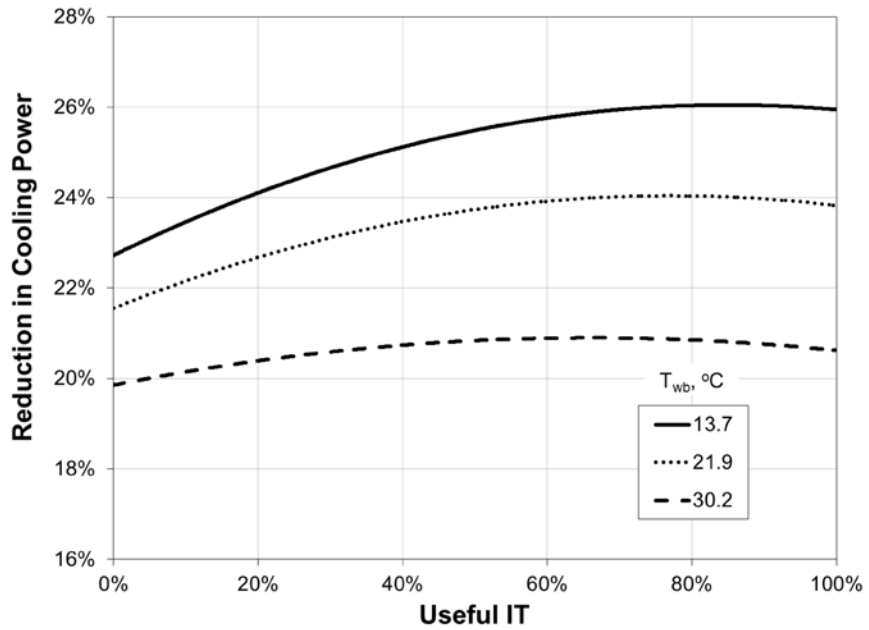


Figure 7.14 - Effect of Ambient Temperature on the Optimum Cooling Power Consumption with

$$\Delta T_m = 18.2^\circ\text{C}$$

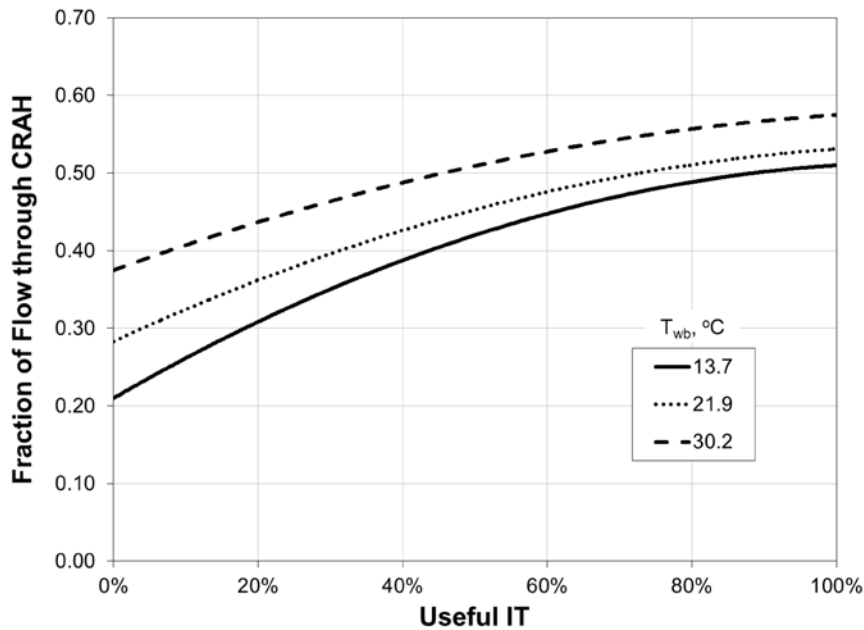


Figure 7.15 - Effect of Ambient Temperature on the Optimum CRAH Airflow Fraction with $\Delta T_m = 18.2^\circ\text{C}$

7.2.2 Effect of Server Temperature Rise

Figures 7.16 and 7.17 show the effect of the chassis' temperature rise on the infrastructure optimization for an ambient wet-bulb temperature of a 21.9°C . The chassis temperature rise is changed by reducing the chassis airflow rate at the *on* state ($\Delta T_m = 18.2^\circ\text{C}$, 20.2°C and 22.7°C at rack airflow rates of 1500 CFM, 1350 CFM and 1200 CFM, respectively). The *idle* state airflow rate remained unchanged in this analysis. Clearly, servers with a higher temperature rise (i.e., lower airflow rate) reduce the CRAH's fan power consumption and therefore, since the refrigeration load (at a given IT power) is nearly the same, lower savings are awarded by the use of bypass recirculation.

If the servers have a higher temperature rise than the infrastructure was designed for, the optimum operating point shifts towards more airflow through the CRAH.

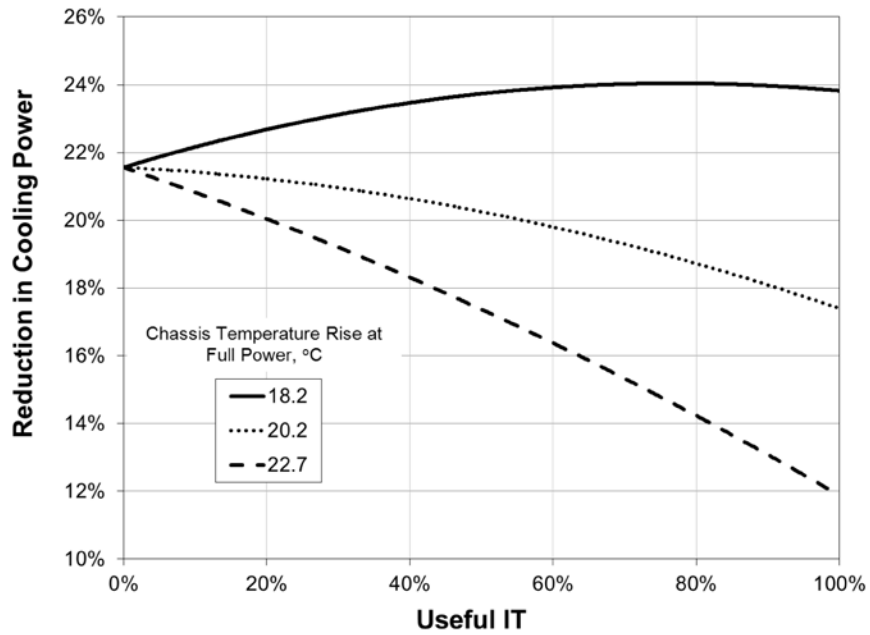


Figure 7.16 - Effect of Chassis Temperature Rise on the Optimum Cooling Power Consumption with

$$T_{wb} = 21.9^{\circ}\text{C}$$

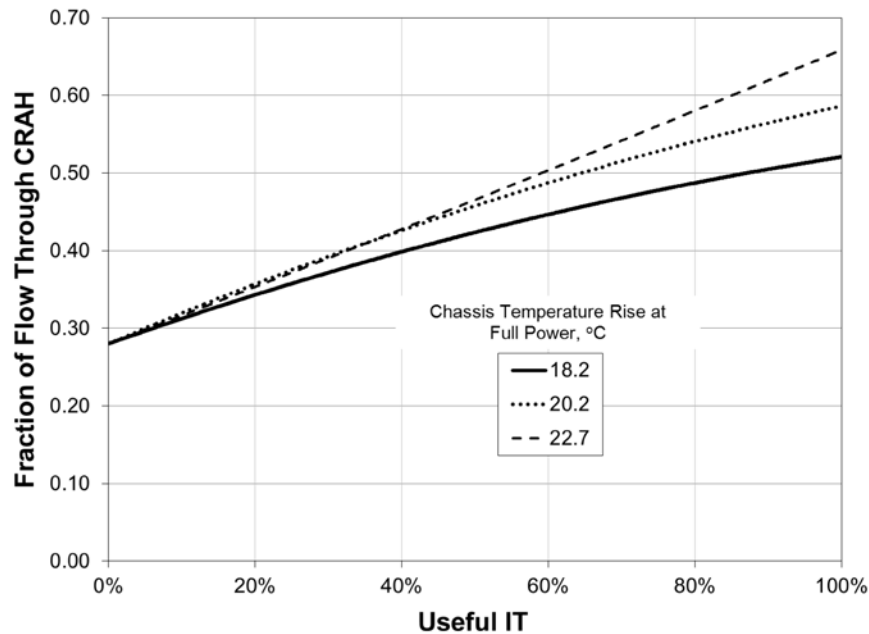


Figure 7.17 - Effect of Chassis Temperature Rise on the Optimum CRAH Airflow Fraction with $T_{wb} = 21.9^{\circ}\text{C}$

7.2.3 Section Conclusions

The notion of optimizing load placement based on the thermal environment is currently at the forefront of research on data centers. When studying this optimization problem, it is essential to consider the power consumed by all equipment in the cooling infrastructure. Throughout this dissertation, it has become evident that a higher CRAH exit temperature does not always lead to lower power consumption in air-cooled data centers, where CRAH fan power consumption is a significant contributor to the overall cooling infrastructure power. The potential energy benefits of enclosing the cold aisle have been well documented. However, in addition to their energy benefit, enclosed aisle data centers facilitate so-called thermally aware load placement, because there is no

preference based on the thermal environment, which is uniform in this case. For heterogeneous data centers, it is anticipated that the metric for determining which servers to operate would be based more on computing performance than the thermal environment. However, there is some indication that more savings would be possible by selecting equipment whose power and airflow requirements are a better match to the designed infrastructure.

The results presented in this section highlighted the added benefits of bypass recirculation in enclosed aisle data centers, especially those that are operated at a reduced IT load, since the optimum amount of bypass recirculation increases mightily as the IT load in the data center is reduced, owing to the off-design performance characteristics of the cooling infrastructure.

7.3 Developing a Control Methodology for the Implementation of Thermally Aware, Energy-based Load Placement

With an abundant number of simulation results in place, the purpose of this section is to propose potential solutions for implementing thermally aware, energy-based load placement in operating data centers and provide preliminary thoughts regarding control implementation and logic to use in designing future experimental studies.

The results of this dissertation showed that the near-optimum operation of the data center could be found using relatively easy to implement sensors and actuators - many of which

are already in place, even in legacy data centers. At the data center level, only temperature sensors at the inlet to the IT equipment are necessary. These sensors are readily available in almost all IT equipment currently on the market. The output of these sensors can easily be communicated to a central building management system or a data center specific application, such as Tivoli (IBM, 2011).

The computer room air-handling units would require variable speed drives, which would allow for control of the delivered airflow. The installation of VFD control on CRAH units is a popular option on today's models. The one point that needs to be addressed when controlling the airflow through the CRAHs is that the heat exchanger can deliver the required capacity at the reduced flow solutions being proposed here. Many manufacturers, such as Emerson (2011), have already considered this in the design of the CRAH's built-in controls. While specific knowledge of the algorithms used in their units is proprietary, the solution relies on the use of a three-way valve to control the water bypassed around the heat exchanger coil, similar to the proposed solution given in Section 6.2.2. Since the work of this dissertation showed that the largest energy savings came from controlling the airflow, a study of the data center raised floor would be needed to determine the region of each CRAH's influence in relationship to perforated tile locations. This can be done computationally using the tools developed by Lopez and Hamann (2011). Once this is known, the delivered CRAH airflow (i.e., the VFD setting), can be related to ψ_T through the simple analysis presented in Section 4.3.

At the chilled water plant, the results of this work relied upon the ability to provide capacity and chilled water temperature control of the refrigeration units as well as speed control of the chilled water pumps – both of which are common today. For centrifugal chillers, capacity control is provided by either adjusting the inlet guide vanes or compressor variable speed control. Control of the chilled water temperature is a standard feature based on the evaporator heat exchanger design.

The proposed approach for implementing the thermally aware, energy-based load placement methodologies developed here is summarized below for the case where the IT load needs to be reduced:

1. Determine the required reduction in IT load based on the current workload of the data center.
2. Remove the required IT load from the hottest servers, which is determined via the IT equipment's inlet temperature sensors (in the case of increasing load, the load is added via virtual machines or other means to the servers with the coldest inlet temperature).
3. Determine the appropriate CRAH airflow rate for the given IT utilization level based on a map similar to the one given in Figure 7.2.
4. Use a feedback-based control loop to adjust the CRAH's supply air temperature to that which identically meets the IT equipment's inlet temperature constraint given by Equation 6.1. This would be done by knowing the region of each CRAH's influence.

The proposed approach uses a simple proportional-integral-derivative (PID) controller based on the IT equipment's inlet temperature measurements. The CRAH's supply air temperature is adjusted to meet the inlet temperature constraint of the hottest server. The PID controller is suggested because the derivative control will tend to reduce the overshoot of the set point in the controller, hence, limiting reliability concerns of the equipment due to elevated inlet temperatures. Even with the simple controller proposed here, a number of research questions remain that need to be considered in future work.

While the flow physics governed by the Navier-Stokes equations is inherently non-linear, the PID controller itself is a linear control. Therefore, its performance in a non-linear system is suspect. However, several of the results in this dissertation showed that the response of the temperature field to changes in the supply air temperature remains nearly linear for reasonable operating scenarios. Another important area of research needed throughout the IT industry is related to the transient response of the data center. Several studies have looked at the transient time for changes such as turning off CRAH units, changing power levels in servers, etc. (Sharma et al., 2005; Gondipalli et al., 2010; Ibrahim et al., 2010); however, none of these studies considered the thermal capacitance of the equipment. Detailed experimental studies should be conducted to understand the transient characteristics of the data center. The transient characteristics could play a role in the selection of the derivative control, as it tends to slow the transient response of the system, which could be important if the transient response of the data center leads to unwanted behavior – such as elevated temperatures at the inlet to the IT equipment.

If the performance of the PID controller proves unsatisfactory, it can be improved by incorporating a model-based (or feed forward) control with such methods as the technique of proper orthogonal decomposition, which was discussed in Section 6.3. The feed forward control can provide a good portion of the controller output and the feedback control can correct for any remaining error. This type of control can improve the response and stability of the system. This can also be done using other reduced order models such as neural networks or fuzzy logic as discussed by Moore et al. (2006a). Moore et al. (2006b) also described a neural network based model that predicted the rack's inlet temperatures in real-time and was self-improving using measurements collected during operation of the data center. A conceptual schematic of the proposed control system is shown in Figure 7.18, in which the CRAH units are controlled via programmable logic controllers (PLC) that implement the PID control. The PLC is being supplied with the temperature measurements being collected by the building management system.

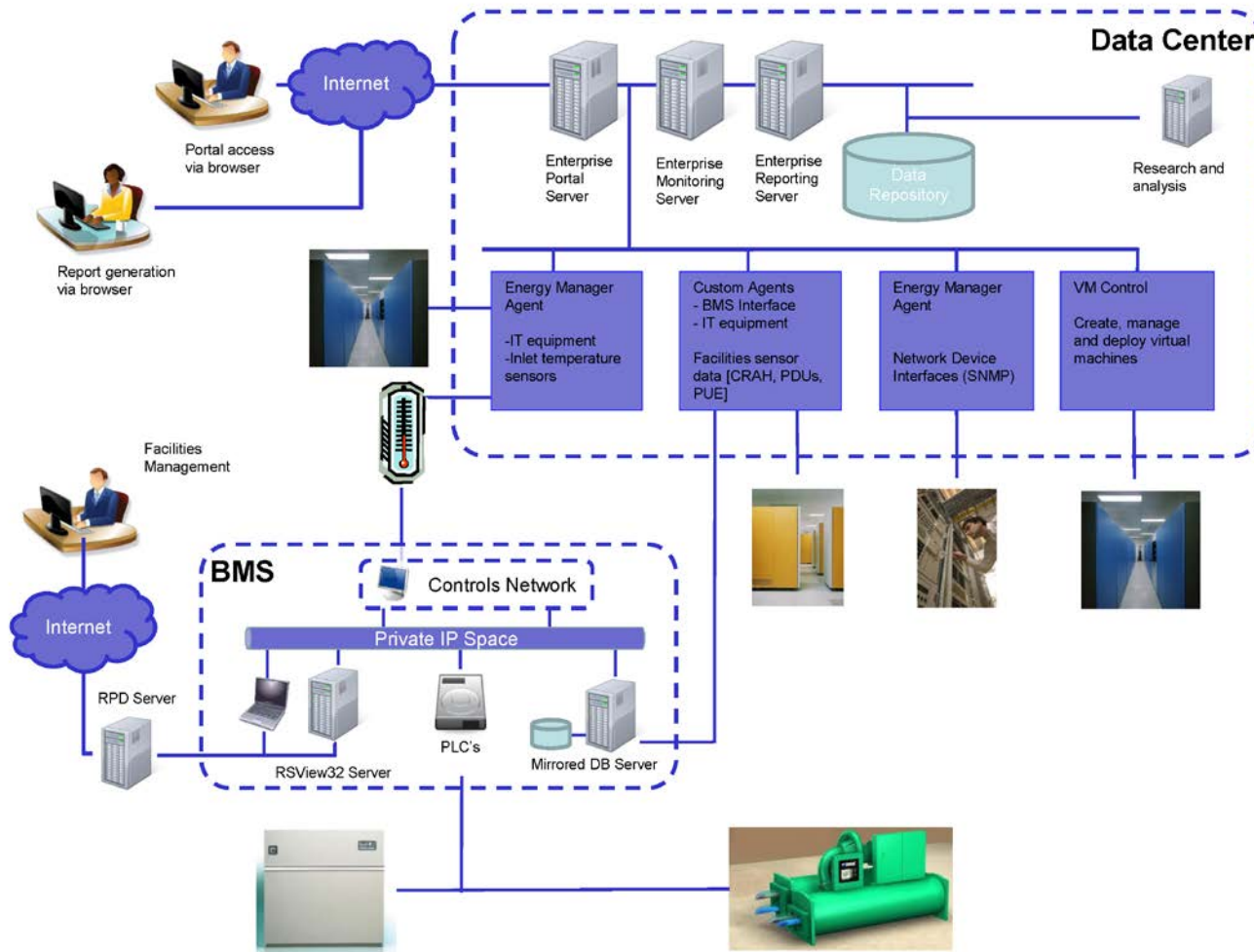


Figure 7.18 - Proposed Control System for Implementing Thermally Aware, Energy-based Load Placement

7.4 Practical Implementation Issues

Data centers have many disadvantages when relying on real-time measurements for prediction and control, due to the mission critical nature of the application. In most cases, a very limited set of sensor data is available for use. Throughout this dissertation, many tools, techniques and parameters were adopted that relied on measurements to facilitate the optimization of air-cooled raised-floor data centers. This section will review these methods and discuss their practical application in operating data centers.

The heuristic rules proposed in Chapter 6 relied on easy-to-obtain information – either based on geometrical properties of the data center or real-time temperature measurements at the inlet to the IT equipment – to facilitate load placement. Implementing these rules into existing virtualization software is a relatively straightforward task. The two main justifications for significant energy savings were:

1. using load placement strategies that took advantage of the IT equipment's higher inlet temperature constraint when *idle* and
2. optimizing the tradeoff between CRAC/H supply air temperature and flow rate.

The first of these is a manufacturer design constraint in order to maintain acceptable component temperatures. Currently, most manufacturers design equipment to operate in classes A1 or A2 (referring to Table 1.1); however, the updated ASHRAE guidelines (2011) proposed new classes (A3 and A4) in which the industry would like to see IT

equipment designed. In the 2011 standard, ASHRAE added to the allowable guidelines an extended thermal environment for equipment that is powered *off* or *idle* – this was leveraged in this dissertation to formulate the IT equipment’s inlet temperature constraint (Equation 6.1). Based on this guideline, from a thermal environment perspective, there is little benefit to operating a chassis at partial power – since it would still be required to conform to the *fully on* constraint. Undeniably, this was the reason that Scenario 6 (uniform workload) was inferior to all of the other proposed workload placement algorithms.

Secondly, one needs a method for controlling the airflow delivered to the perforated tiles near the IT equipment. Controlling the airflow of a CRAC/H unit is relatively straightforward with an easy-to-install variable speed drive; however, measuring or predicting the airflow distribution in the under-floor plenum is not as straightforward in real-time. Several researchers have proposed methods for determining the regions of CRAC/H influence. These regions are primarily a function of the plenum geometry and therefore can be obtained using either computational fluid dynamics⁶ or the reduced order methods proposed by Hamann, Lopez and Stepanchuk (2011), which uses real-time measurements to compute the CRAC/H thermal zones.

⁶ The commercially available CFD tool TileFlow (Innovative Research 2012) provides an easy-to-use interface for determining the tile airflow distribution in raised-floor data centers.

The technique of proper orthogonal decomposition was proposed as a tool for predicting the rack's inlet temperature distribution – since this is the key driver for both increased energy consumption and reduced equipment reliability. Other researchers have considered the real-time prediction of the rack's thermal profile using reduced-order models (Moore et. al, 2006b; Samandiani et al., 2009). However, only one study considered changes in CRAC/H airflow rate (Samandiani et al., 2009) even though adjusting the flow rate was shown to have a significant effect on both the rack's thermal profile as well as the data center's energy consumption. The work in this dissertation highlighted the fact that for a given data center configuration, the thermal profile is effected most by the airflow rate (i.e., ψ_T) and the load placement technique – hence was the motivation for developing the POD in terms of these two parameters. The work presented in Appendix B attempted to determine the hottest chassis as a function of the Archimedes number only; however, while the results showed promise, they ultimately displayed a necessary dependence on the useful IT and/or ψ_T .

While measuring ψ_T in real-time is almost impossible in operating data centers, Samandiani et al. (2009) suggested using the CRAC/H VFD setting as a surrogate airflow parameter since this is readily available during operation. Formulating a reduced order model using POD in terms of the VFD setting is undoubtedly possible. The key to using the CRAC/H VFD setting as a surrogate parameter is in identifying the regions of influence of the various CRAC/H units in the data center. Although the CRAC/H VFD setting is easy to obtain, it can only be used as an estimate of the actual flow rate since

the actual flow rate is affected by dirty filters in the units, under floor obstructions, etc., which cannot be considered unless the regions are being computed in real-time.

Lastly, the Archimedes number derived in Section 4.5 was convenient for this work because it was written in terms of the tile-to-rack cold air fraction, ψ_T , and the rack thermal characteristics, ΔT_m and \dot{V}_m or P_m . However, the Archimedes number can be written in any form as long as an appropriate velocity and temperature scale are used. In data center applications, another possible temperature scale is the overall temperature rise (or power consumption) of the data center, which could be measured at the CRAC/H units for example. One of the hardest parameters to obtain however is an appropriate velocity scale (or flow rate). In reality, the non-dimensional parameter ψ_T is impossible to measure in real-time. In fact, most airflow related quantities are not realistic for real-time measurement and in some cases at all. Again, the best hope for obtaining an airflow quantity is by relating the CRAC/H VFD setting to the flow rate. Whichever method is adopted, it is important to remain consistent in the definition of Ar used (similar to how the Reynolds number for pipe flows is always based on the hydraulic diameter of the pipe) to compare the airflow characteristics of data centers. While the physical interpretation of the number will not change, the numerical value will change depending on the definition used.

8 Conclusions

The main objectives of this work were to develop tools and techniques for studying the holistic energy consumption of an IT data center and to define optimum operating strategies for efficient IT load placement, without compromising the reliability of the IT equipment. The approach undertaken was a systematic modeling approach – beginning with simple physics-inspired models then moving towards higher-fidelity models as the optimization space was reduced through the simpler techniques.

8.1 Summary of Results

The simple analysis presented in Chapter 2 provided a flexible and fast tool for exploring optimization possibilities in air-cooled data centers. It proved very useful in identifying optimal, energy-efficient designs and operating scenarios. The methodology embodied in this simple analysis can be used in the early stages of the conceptual design process to define energy saving approaches and near-optimum design and operating parameters such as flow rates and air supply temperatures, as well as to carryout tradeoff investigations of the cooling infrastructure's sizing and performance. While this simple model was not a substitute for detailed, higher fidelity analysis and optimization studies, its most useful benefits stem from its ability to limit the range of options and parameters to be explored in the resource-intensive more rigorous optimization analyses. The simple analysis presented in Chapter 3 was able to highlight the importance of the trade-off of low air supply temperature vs. increased airflow rate, while identifying the energy-saving

potential of providing thermally uniform conditions to the inlet of the racks. A simple metric of recirculation non-uniformity was defined. Using this metric, the deleterious effect of flow non-uniformity at the inlet of the racks on the data center's cooling infrastructure power consumption was shown.

The simple model also proved highly useful in formulating the novel energy saving technique of recirculation bypass in enclosed aisle data centers. A practical bypass tile configuration was proposed in which tiles with low-lift fans are placed near the CRAC units on the raised floor to inject the bypass flow into the under-floor plenum. This configuration did a reasonable job at promoting mixing of the bypass air and CRAC air in the under-floor plenum. It is expected that in the future, other configurations could be proposed to improve the uniformity of the inlet temperature. It is anticipated that installing a bypass branch within the CRAC itself would be a preferred approach.

A more detailed account of the recirculation non-uniformity metric along with a verification of several of the assumptions necessary in the development of the simple model was carried out using detailed computational fluid dynamics analyses in Chapter 4. It was shown that the simple model provides a reasonably accurate estimate of the recirculation non-uniformity parameter, which was related to the overall cooling infrastructure's power consumption, even with the inherent assumption of a single mixed exhaust temperature. This simplification made it possible to compute the recirculation non-uniformity metric using only temperature measurements at the inlet of the IT equipment, which are readily available in most operating data centers.

Next, using an Archimedes number tailored for data center applications, the effect of buoyancy in high-density data centers was studied to provide insight into the drastic effect it could have on altering the temperature patterns at the inlet of the IT equipment. The analysis shows the difference in exhaust recirculation patterns to the inlet of the racks in the presence of a strong buoyancy force. It also showed that the assumption of a linear, one-to-one change in the temperature field with a change in the supply air temperature is a weaker assumption for high Ar data centers. Lastly, several parameters were introduced to account for the effect of leakage on the data center's overall energy consumption.

Chapter 5 presented a detailed experimental validation of a coupled thermodynamic and hydraulic simulation environment, which gives data center designers and engineers the ability to evaluate the energy consumption of various data center configurations at the system and data center levels. The development of this tool was quite monumental in the industry as it was the first study to validate experimentally a holistic model of the data center and cooling infrastructure's power consumption. The inherent difficulty in experimentally validating a model of this nature stems from the fact that the data center operation is extremely sensitive to reliability and uptime constraints, which limits any intervention in installing new measurement equipment and conducting controlled experiments.

In the future, it is anticipated that higher fidelity data will be available for validating dynamic system simulation models like the one developed here. As a starting point, it would be desirable to validate the model over a larger range of operating conditions. One of the intentions of this study was to guide development of experimental plans to collect the appropriate data for future validation studies. The thermo-hydraulic model has already been extended to include waterside economizers, airside economizers, evaporative cooling, and co-generation systems.

With experimentally validated tools in place, the remainder of the dissertation focused on the development of control methodologies for implementing efficient IT load placement. Chapter 6 introduced several practical IT load placement scenarios. The approach used in this dissertation differed from previous research in that it focused on developing rules based on feature or physics-based characteristics of the data center or measurements that are obtainable during real-time operation. The previous work by other researchers all focused on developing rules through intricate optimization algorithms, which had to be computed ahead of time due to the need for a long computational time to develop the placement of the IT load. During the exploratory phase of this work, seven load placement scenarios were considered using computational fluid dynamics to compute the airflow and temperature distribution in a representative hot aisle/cold aisle data center.

With these simulations, a reduced order model, using proper orthogonal decomposition, was developed. The model was useful in expanding the analysis of the optimum load placement, by providing a simple and fast method for computing the temperature

distribution at the inlet of the racks at an infinite number of design conditions, with little computational effort. The reduced order model was capable of generating a temperature field, with an accuracy of better than 0.2°C RMS, in less than a second on a standard PC, compared to a CFD simulation that needed 7 hours to complete using 32 processors on a high-performance computer.

The optimization analysis showed that the most effective scenario was one that removed IT load from the equipment that had the highest inlet air temperature. These servers can easily be found in practice using built-in temperature sensors, which most IT equipment currently have. While this proved a robust and energy efficient strategy for load placement, there was less than a 10% difference in the cooling infrastructure's power consumption between the different scenarios considered. However, potential savings of greater than 20% in the cooling infrastructure's power consumption were realized by simultaneously considering the optimization of the cooling infrastructure and load placement. A substantial portion of the savings came from allowing the inlet temperature of *idle* servers to be 5°C higher than servers that were *on*, which is consistent with the newest ASHRAE environmental guidelines (ASHRAE, 2011). The results showed that scenarios that allowed for significant reductions in the supply air flow proved superior to those that relied on higher supply air temperatures to meet the IT equipment's inlet air temperature constraint. This result may be received with criticism from the industry, since the current best practice is to raise the supply air temperature as much as possible in order to reduce the refrigeration power consumption. However, only when studying the problem from a holistic perspective would it be realized that significantly more savings

are possible through reduced airflow that by increasing the chilled water temperature set point. Ignoring the CRAH's fan power may be justified in situations in which the refrigeration power consumption is dominant, as it would be when inefficient refrigeration systems are employed in very warm climates. However, exclusive focus on reducing refrigeration power consumption without regard to the power consumed in moving the cooling air would certainly lead to sub-optimum, possibly misleading results.

Chapter 7 extended the work on efficient load placement by discussing control rules for implementing thermally aware, energy optimized load placement in data centers, in which equal consideration was given to optimizing the load placement and the cooling infrastructure. The load placement work was extended to study other conditions that may be prevalent in data centers. If the servers were completely shut *off*, the optimum operation of the cooling infrastructure tended towards a higher server-to-CRAH airflow fraction compared to the scenario where the servers are placed in an *idle* mode. In addition, it was shown that a complementary strategy for increasing the IT load was to add load to the servers with the coldest inlet temperature. The results for both increasing and decreasing the IT load were shown to hold both when the data center was operated in a *dynamic* fashion as well as if the placement algorithm was generated from *steady* conditions at the *full* and *idle* power levels. The benefit of a dynamic operation is that the algorithm is capable of choosing the servers that are actually the hottest and can adjust more easily to changes in the data center's equipment configuration and operation.

Lastly, efficient load placement in homogenous enclosed aisle data centers was considered. Because there is no preference for load placement when the thermal environment is uniform, the problem of load placement focuses entirely on the optimum operation of the cooling infrastructure, which results in the implementation of the novel bypass recirculation branch, as discussed in Chapter 3.

8.2 Suggestions for Future Work

The work in this dissertation highlighted the benefits of using a systematic modeling approach for understanding the potential benefits of optimizing the design and operation of air-cooled data centers; however, continued study would be beneficial in several areas. First, a detailed experimental validation of the computational fluid dynamics simulations is necessary, not only for this work but throughout the IT industry. While the simulations done here included all the essential physics of the flow, validation against experimental data is ideal. Secondly, experimental implementation of the load placement rules would be useful for determining how robust the rules are in environments that are non-uniform and highly variable.

While all the analysis presented in this work focused on active cooling infrastructure utilizing vapor-compression refrigeration (direct expansion or chiller), it could be easily extended to include economizer operation. When economizers are used in series or in parallel with, or in place of vapor-compression refrigeration, the trade-off between refrigeration and air moving power consumption is expected to shift toward a greater

emphasis on reducing air moving power usage since the power for active refrigeration in this case will be greatly reduced, or may be eliminated altogether under suitable climatic conditions.

All of the analysis in this dissertation focused on raised-floor, air-cooled data centers.

While these are certainly a proven technology and the most obvious choice for implementation in the near future, they will not be the sole solution for cooling as the data center's heat flux continues to increase. Figure 8.1 shows the problem with air-cooled data centers as the rack heat load increases. With current technologies, a typical limit on the per tile airflow rate is around 1,100 CFM. From the figure, it can be seen that this limit is quickly surpassed as the rack's heat load increases beyond about 20 kW.

Therefore, technologies that present better heat transfer characteristics than air must be pursued. Currently, many technologies available either improve the effectiveness of typical CRAH-based data centers or eliminate the need for air-cooling altogether. Moving the cooling source closer to the heat load is certainly an effective way of eliminating hot spots in the data center and ultimately improves the effectiveness of air-cooling.

Technologies such as in-row coolers, which place the CRAC/CRAH in-line with the IT rack, and rear-door heat exchangers, which remove a fraction of the heat load via water-cooling at the rack, are aimed at improving the efficiency of air-cooling. Water-cooling at the chip level provides a better solution from a heat transfer perspective, since water has a thermal capacitance that is more than four times that of air, and eliminates the need for air-cooling completely. IT equipment that accommodates chip-level water-cooling is becoming more prevalent on the market; however, infiltration into the industry is not

widespread due to the fear of water leakage, which could destroy the IT equipment. No matter which solution is pursued, they all have research questions that remain unanswered. Many of these questions could be answered using the tools and techniques developed throughout this dissertation.

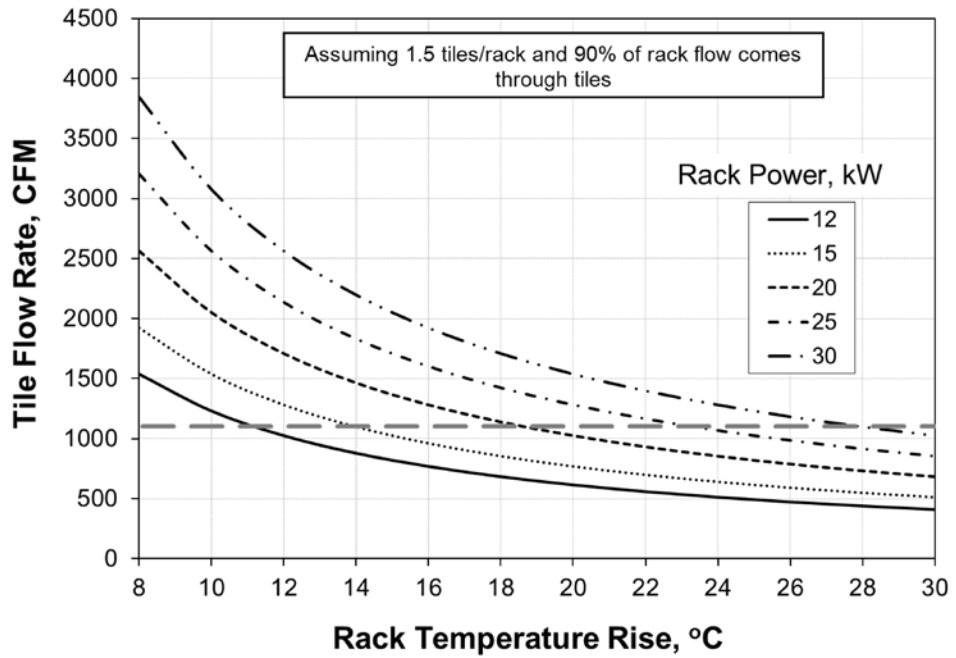


Figure 8.1 - Inherent Problem with Air-Cooled Data Center

Appendix A: Input Data Used for Thermo-Hydraulic Model Validation

This appendix contains the input data used in the validation of the thermo-hydraulic model presented in Chapter 5. All of the information was obtained from both site measurements of the data center (B710) and chilled water plant (B027) or from manufacturer catalogues.

A required input of the hydraulic model is an accounting for the major and minor pressure losses in the hydraulic network. The major losses are easily computed by knowing the length and material of each pipe in the network. This information was obtained from Figures A.1 – A.3. However, limited information was available related to the minor losses in the system (i.e., valves, fittings, 90° elbows, etc.); Therefore, these were estimated based on the hydraulic schematics. The values used for the loss coefficient, k , in the model were obtained from Janna (1998) and are given in Table A.1.

B/027 AUX UTILITY PLANT

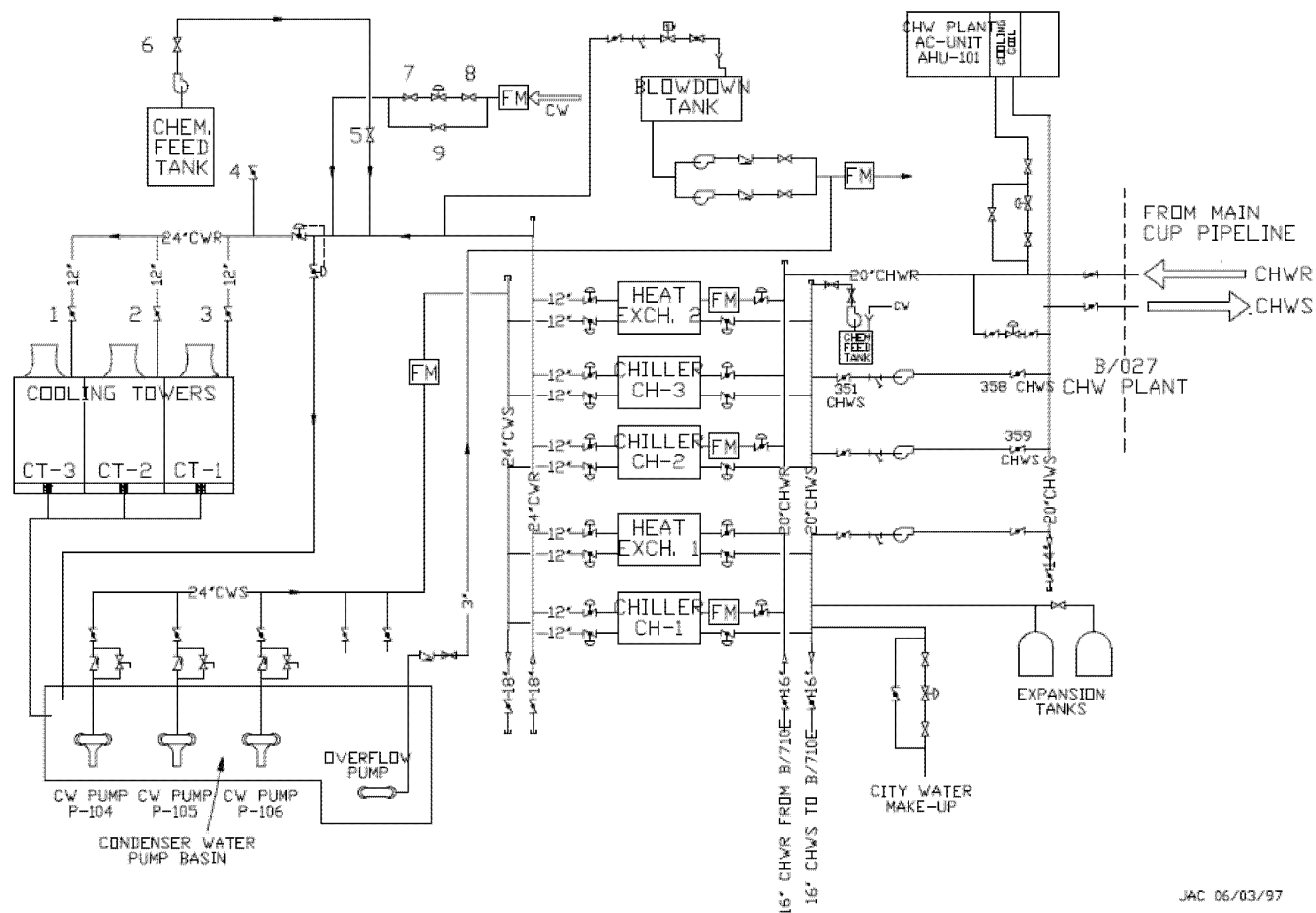


Figure A.1 - B027 Utility Plant Schematic

B/710, B/710E AND B/027 CHW PIPING

B/027 PEAK LOAD HISTORY

2000 - 230 Tons
 2001 - 370 Tons
 2002 - 90 Tons
 2003 - 85 Tons
 2004 - 270 Tons on 5/22 @ 2:40 AM
 2005 - 230 Tons on 1/28 @ 9:00 AM
 2006 - 200 Tons on 1/22 @ 3:00 PM

* SEE INDICATED OR HARBORER AREAS CURRENTLY BEING FED BY B/027
 * CHW INDICATES DUCTIVE PIPING NOT IN USE WITH CURRENT VALVE ALIGNMENT
B/027 REDUNDANT CAPACITY - 2200 TONS

- NOTES
1. EP Transmitter on B/710E-2nd Floor controls B/710E Pump 03 VFD, currently inactive
 2. EP Transmitter on ACL controls B/710E Pump 1 VFD
 3. EP Transmitter on B/710E-2nd Floor controls B/710E EP Bypass in Basement
 4. Automatic Control Valves in B/027 for the section into B/710E

B/027 UPS TARGET 2006 DESIGN SPECS

2006 UPS - 20 TONS
 PANE 1 SWITCH FE - 40 TONS
 LEAD CONTROL ROOM (ACH) - 40 TONS
 SHEP TEST LAB (MAY 04) - 5 TONS
 ACS - 140 TONS
 ACS - 40 TONS
 PANE 0 LINE CONTROL ROOM - 40 TONS
TOTAL LOAD 400 TONS

LOADING LOSS

PAN FLOOR - 700 sq ft @ 2.50 W/SGY = 1750 TONS
 1ST FLOOR - 500 sq ft @ 4.00 W/SGY = 2000 TONS
 LEVEL LEVEL - 300 sq ft @ 5.00 W/SGY = 1500 TONS
TOTAL - 5250 TONS
PLANT LOAD - 200 TONS

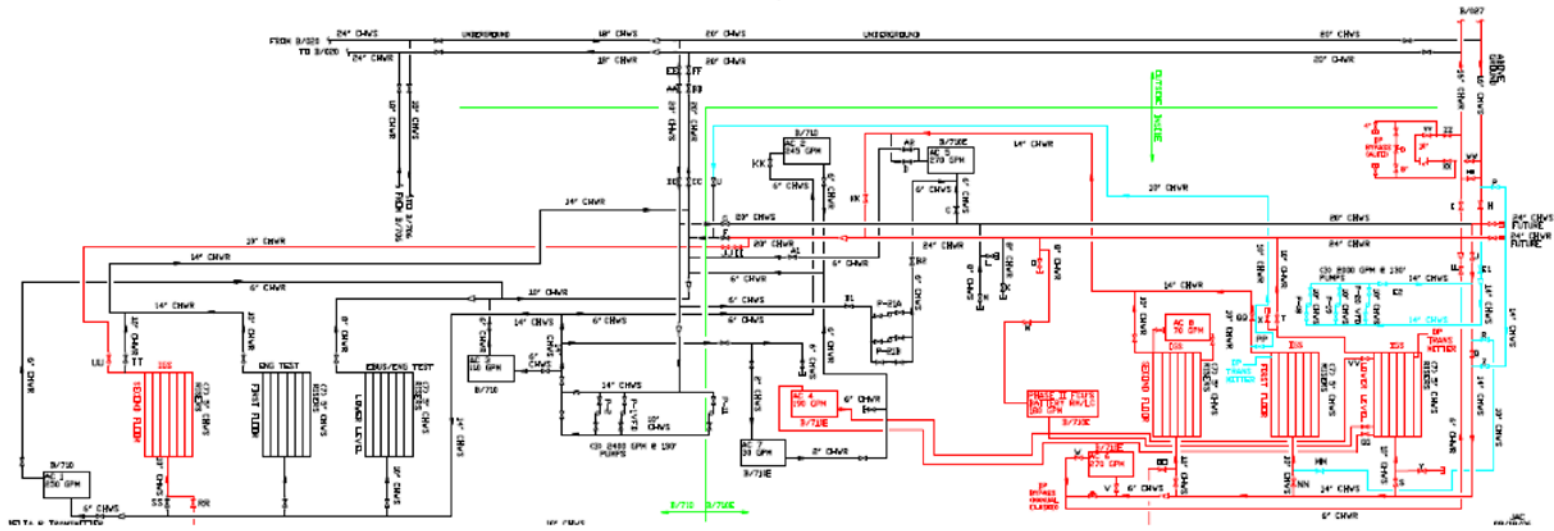


Figure A.2 - B710 Hydraulic Schematic (Pipe Lengths)

Table A.1 - Minor Losses used in Thermo-Hydraulic Model

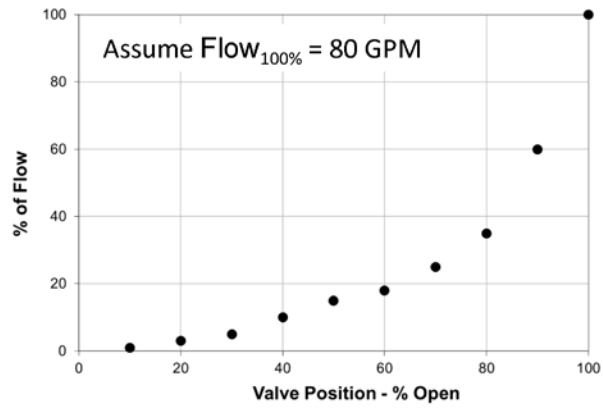
Fitting Type	Quantity	$k/\text{fitting}$	Σk	L_{eq}, m
90° Elbow	35	0.31	10.85	245.0
Tee	19	0.14	2.66	60.1
Butterfly Valve	24	12.80	307.20	6935.9

The computer room air handlers in B710 were Liebert Model FH529C. These units have constant speed fans, with a rated airflow rate of 12,000 CFM. The actual airflow through a unit was estimated to be 10,000 CFM because of the presence of dirty filters, under-floor obstructions, etc. This was based on previous measurements done in B710. The thermal characteristic of the units (overall conductance, UA) needed for Equation 2.44 was computed based on manufacturer's design point data, as given in Table A.2.

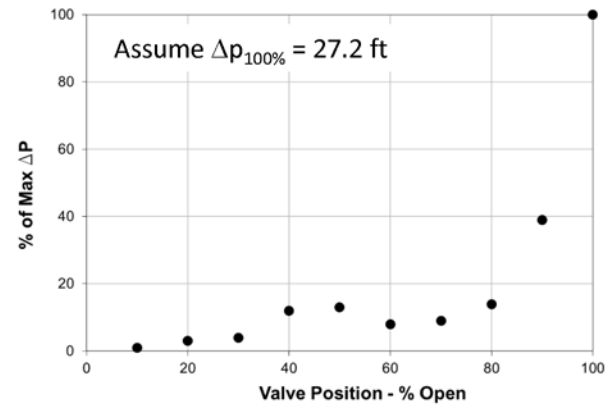
Table A.2 - CRAH Unit Design Characteristics

Model FH529C	Liebert Design Point				
	1	2	3	4	5
$T_w^{in}, ^\circ C$	7.2	7.2	7.2	7.2	7.2
$T_a^{in}, ^\circ C$	26.7	23.9	23.9	22.2	22.2
\dot{Q}_s, kW	105.2	89.5	90.9	80.4	81.8
\dot{V}_a, CFM	12000	12000	12000	12000	12000
\dot{V}_w, GPM	107.2	80.1	72.7	65.6	59.6
$UA, kW/k$	12.8	13.5	14.6	14.5	16.1

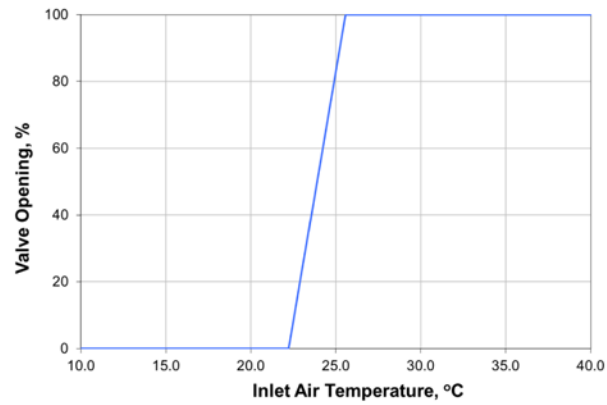
Another feature of the thermo-hydraulic model is the use of realistic CRAH chilled water valve control. The algorithm used in the B710 units adjusted the opening of the chilled water valve based on the unit's return air temperature. The valve characteristics, as obtained from the manufacturer, are given in Figures A.4. From this data, the discharge coefficient and correction factor inputs needed in Equations 2.37 – 2.39 can be computed, and are given in Figures A.5 and A.6.



(a)



(b)



(c)

Figure A.4 – Chilled Water Valve Control

a) water flow fraction vs. valve opening, b) pressure loss vs. valve opening, and c) valve opening vs. CRAH return air temperature

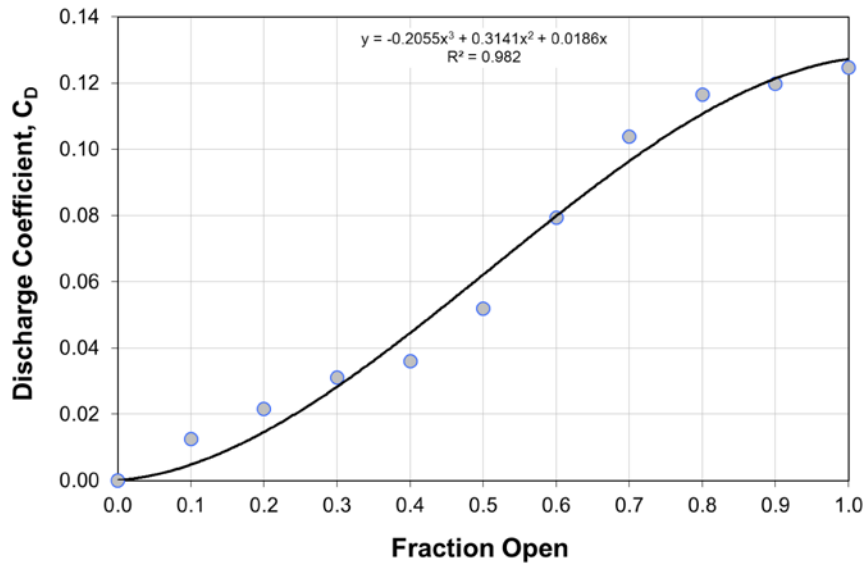


Figure A.5 – Computed Chilled Water Valve Discharge Coefficient C_D vs. Valve Opening

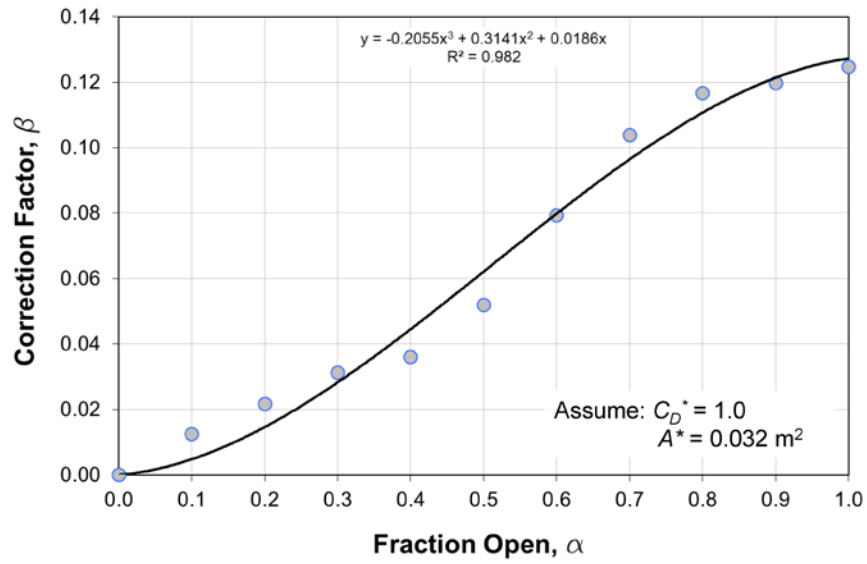


Figure A.6 - Computed Chilled Water Valve Correction Factor β vs. Valve Opening

With this information, the input to the CRAH model would be a measurement of the of return air temperature, from which the valve opening can be computed by Figure A.4c.

With the valve open percentage, the CRAH's hydraulic losses can be fully characterized by computing β from Figure A.6.

A summary of the cooling equipment's design point performance is given in Figure A.7.

A detailed description of each piece of equipment is given in successive figures, starting with Figure A.8. These figures provide a detailed performance map of each piece of equipment over a range of operating conditions. These details allow for the off-deign performance of the cooling infrastructure to be captured at a range of ambient conditions, IT utilization level, control strategy, etc.

B/027 Auxiliary Utility Plant Equipment

Chillers - (1) 1000 Ton York R134A Unit

- Evaporator GPM - 2400
- Evap. EWT/LWT - 55/45
- Evap. Pressure Drop in Ft - 18.3
- Condenser GPM - 3000
- Condenser EWT/LWT - 85/94
- Condenser Pressure Drop in Ft - 13.5

Chillers - (2) 1200 Ton York R134A Units

- Evaporator GPM - 2400
- Evap. EWT/LWT - 46/58
- Evap. Pressure Drop in Ft - 13.4
- Condenser GPM - 3000
- Condenser EWT/LWT - 85/96.1
- Condenser Pressure Drop in Ft - 10.6

Cooling Towers - (3) 1200 Marley Units

- Summer Flow GPM - 3000
- Summer EWT/LWT - 96/85
- Summer OA WB - 78
- Winter Flow GPM - 3000
- Winter EWT/LWT - 52.4/46
- Winter OA WB - 38
- Max Motor Hp - 60

Pumps

- Chilled Water - 2800 GPM at 100 Ft of head, 100 HP 1750 RPM, Weinman 8L1
VFD's installed on CHW pumps, DP transmitter in B/027 and B/710E Basement
- Condenser Water - 3000 GPM at 86 Ft of head, 100 HP 1750 RPM, Weinman 10L3

Heat Exchangers

- Primary GPM - 3000
- Primary EWT/LWT - 46/52.4
- Primary Pressure Drop in Ft - 22.8
- Secondary GPM - 2000
- Secondary EWT/LWT - 57.6/48
- Secondary Pressure Drop in Ft - 10.4
- Model - BAC EC9-237-2M

Figure A.7 - B027 Infrastructure Design Point Summary

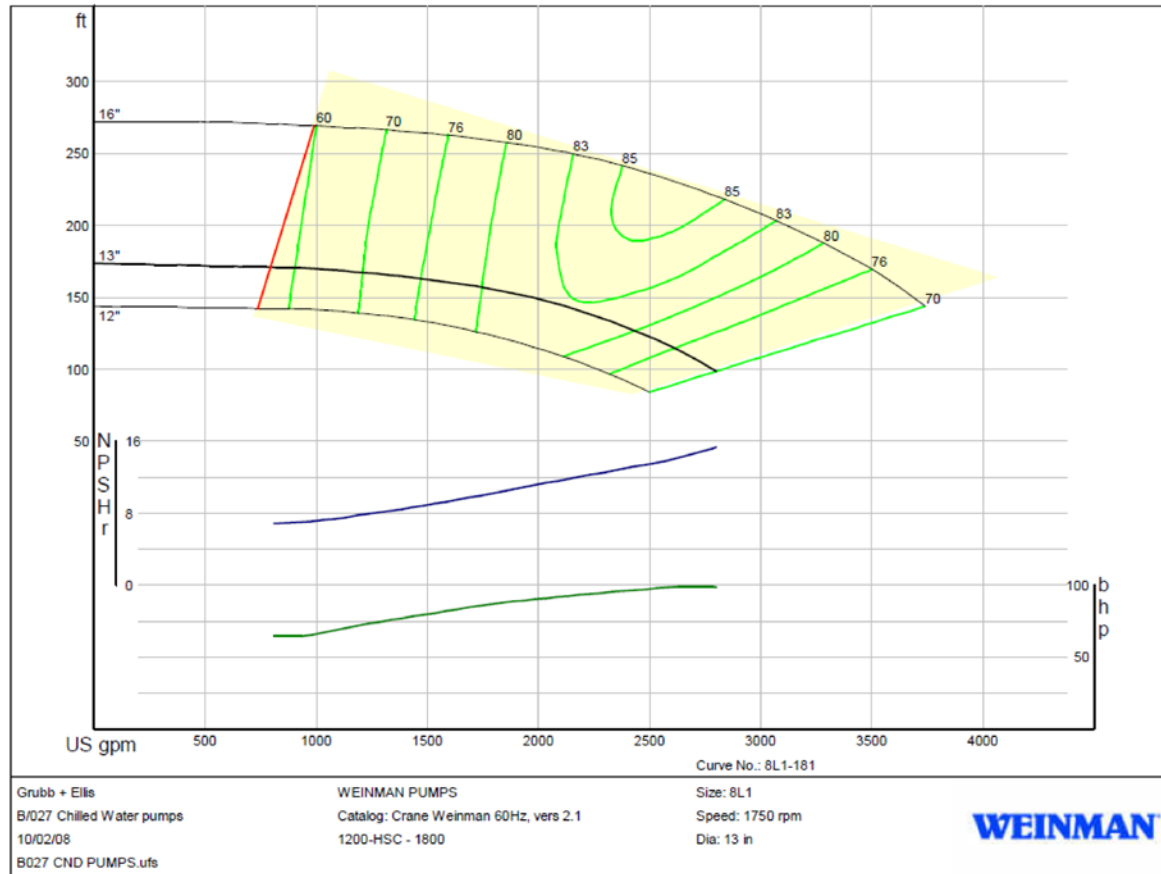


Figure A.8 - B027 Chilled Water Pump Characteristics

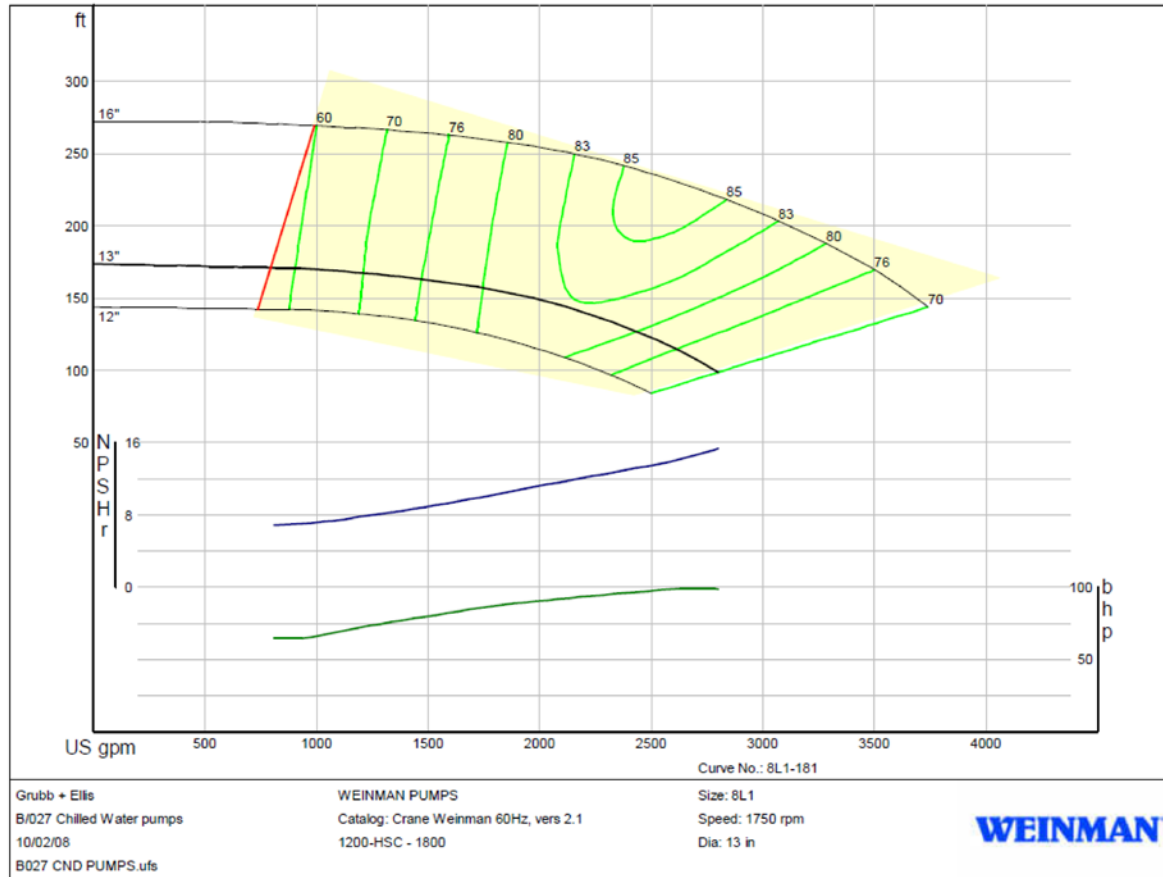


Figure A.9 - B027 Condenser Water Pump Characteristics

F.R. FOOOTE CO., INC.
 32 Fuller Road • P.O. Box 5129 • Albany, NY 12205
 Telephone (518) 482-7251 • Fax (518) 482-0250
 Vermont and Massachusetts (888) 482-7251
 Visit our Website @ www.frfoote.com

MANUFACTURER'S REPRESENTATIVE
 100 Beethoven Blvd., Yonkers, New York, NY 10590

SCHEDULE

TAG	MODEL	GPM	SUMMER			MHP	ELECTRIC
			EWI	LWT	WB		
CT-1	NC8312H	3000	96°	85°	78°	60	480/3/60
CT-2	NC8312H	3000	96°	85°	78°	60	480/3/60
CT-3	NC8312H	3000	96°	85°	78°	60	480/3/60

- Tower is Galvanized Steel construction with 304 SS cold water basin and hot water basin.
- Cooling towers to be provided with (1) ladder/ safety gage and (1) Common handrail around all three towers.
- Plenum Walkway and Interior Mechanical Equipment Access Platform included for each tower.
- Towers are connected with a fan deck walkway.
- Motors for use with VFD.
- 2 HC Hot Water Balancing Valves (10") provided with each tower.
- Extended lube lines included for gearreducer.
- Basin heaters consist of (2) 15KW, 460V/3/60 heaters.
- Velocity Recovery Stacks Included
- Metrix DPDT Vibration Switch
- Rosemount Model 3144PDIANAB4 Universal Temperature Transmitter with mounting bracket for remote mounting and Model 0068R11C30N120 RTD sensor, 12" long with 3" nipple union extension to be provided per tower shipped loose to be field mounted.
- Drack Model PTX 7211 Industrial Pressure Transmitter provided loose per tower.

INCORPORATED SINCE 1961

Cooling Towers • Pumps • Dehumidifiers • Boiler Breaching and Stack • Duct Work • Air and Water Filtration • Infrared Heating

Vehicle Exhaust Systems • Kitchen Ventilation • Fans • Corrosion Resistant Ventilation Systems

IBM UPS Upgrade
 Pizzagalli Construction
 Poughkeepsie, NY

04-Feb-03
 11:32:17 AM
 Todd Strong

Marley UPDATE Version 3.57
 Unitary Product Data And Thermal Evaluation
 Copyright © 2003, The Marley Cooling Tower Company

Product Line:	NC Class	
Model	NC8312H1	
No. of Cells	1	
Motor	60 HP; 1800 RPM	
Motor Output	60.0 BHP	
Tower Flow Rate	3000 GPM	
Hot Water Temp.	96.00 °F	
Cold Water Temp.	85.00 °F	
Wet-Bulb Temp.	78.00 °F	
## CTI Certified ##		
Sound Level	80 dBA/Cell, 5.00 ft from Air Inlet Face *	
Fan	12.00 ft Dia., 6 Blades	
Fan Speed	289 RPM, 10895 ft/min	
Air Flow	272200 CFM Per Cell	
	272200 CFM Total	
Weights:		
	Per Cell **	Total **
Shipping	19970 lb	19970 lb
Max. Operating	42610 lb	42610 lb
Dimensions:		
	Per Cell	Total
Width	22.42 ft	22.42 ft
Length	13.90 ft	13.90 ft
Height	19.81 ft	19.81 ft
Static Lift	19.21 ft	

For CAD layouts refer to DXF file NC8312

Minimum Clearance for Enclosures:
 Clearance required on air inlet sides of tower without altering performance. Assumes no air from below tower.

Solid Wall	9.67 ft
50% Open Wall	7.23 ft

Collection Basin Heater Sizing:
 Minimum ambient temperature to maintain water at 40.00 °F

Heater kW/Cell	30.0	24.0	18.0	15.0	12.0	9.0	7.5
Ambient Temp °F	-20.40	-7.45	5.50	11.98	18.45	24.93	28.17

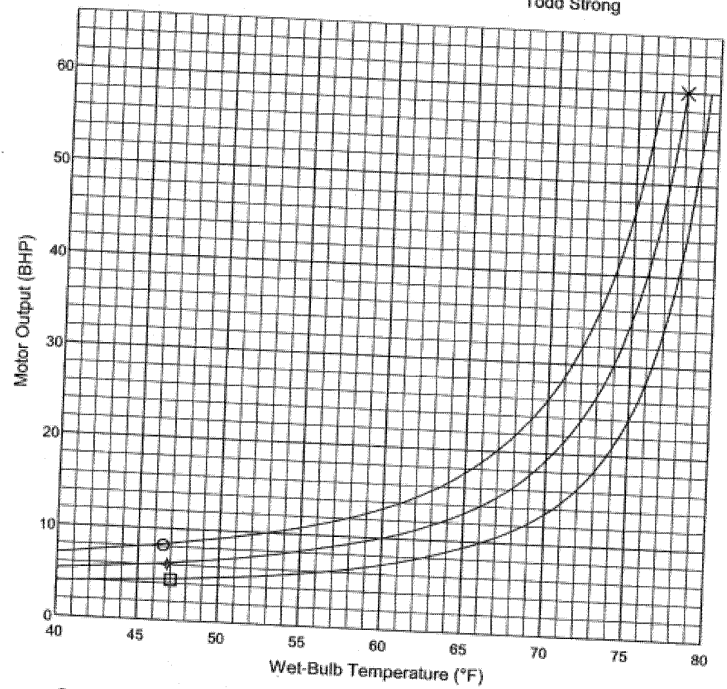
NOTE: ** Weights do not include options. Refer to sales drawing.
 NOTE: * See sound report for additional information.

Figure A.10 – B027 Cooling Tower Performance Data

IBM UPS Upgrade
 Pizzagalli Construction
 Poughkeepsie, NY

04-Feb-03
 11:26:13 AM
 Todd Strong

Marley UPDATE Version 3.57
 NC Class
 Model NC8312H1
 Flow Rate 3000 GPM
 Motor Output 60.0 BHP
 Motor RPM 1800
 Fan RPM 289
 (VFD Operation)
 Number of Cells 1
 CW Set Point 85.00 °F



Design Conditions:
 Flow Rate 3000 GPM
 Hot Water 96.00 °F
 Cold Water 85.00 °F
 Wet-Bulb 78.00 °F

○ 12.00 °F Range
 ◇ 10.00 °F Range
 □ 8.00 °F Range
 X Design Point

Copyright © 2003, The Marley Cooling Tower Company

Figure A.11 – B027 Cooling Tower Fan Control

PROJECT - YORK ORDER NO. 00-105349-01
 YORK SERIAL NO. G8KM25056S

SALES ENGINEER - RCN
 CUSTOMER - IBM - POUGHKEEPSIE

DATE: 02/18/05
 PAGE: 1 OF 3

MODEL YKHHBJ1-CZE (MOTOR SPECIFIED BY USER)
 REFRIGERANT 134A GEAR CODE EK(SPEC)
 RATED CAPACITY (TONS) 1000 SPECIFIED CAPACITY (TONS) 1000
 INPUT POWER (KW) 592 MAX MOTOR LOAD(KW) 617
 VOLTAGE/Hz 2300/60
 DIFFUSER FIXED
 FLA 169 LRA 1100
 INRUSH (AMPS) 1100
 FULL LOAD (KW/TON) 0.592 NPLV
 STARTER TYPE (7) ACROSS-THE-LINE (100%INRUSH) - 3 LEAD 0.519

	EVAPORATOR	CONDENSER
FLUID	WATER	WATER
TUBE	181*	230*
PASSES	2*	2*
FOUL FACTOR	0.00025*	0.00100*
FLUID ENT TEMP (F)	55.00	85.00*
FLUID LEV TEMP (F)	45.00*	94.29
FLUID FLOW (GPM)	2400.0*	3000.0*
FLUID PRDROF (FT)	18.7	13.4

(*) DESIGNATES SPECIFIED INPUT

ESTIMATED PART LOAD PERFORMANCE:

LAST DATA LINE IS PART LOAD OPERATING LIMIT WITH PRE-ROTATION VANES

% LOAD	CAP (TONS)	% POWER (KW)	INPUT (F)	ERPT (F)	ELFT (F)	CEFT (F)	CLFT (F)	PERF (KW/TON)
100.0	1000.0	100.0	592	55.00	45.00	85.00	94.29	0.592
90.0	900.0	87.7	519	54.00	45.00	85.00	93.34	0.577
80.0	800.0	77.2	457	53.00	45.00	85.00	92.40	0.571
72.1	721.2	69.9	414	52.21	45.00	85.00	91.67	0.574

% LOAD	CAP (TONS)	% POWER (KW)	INPUT (F)	ERPT (F)	ELFT (F)	CEFT (F)	CLFT (F)	PERF (KW/TON)
100.0	1000.0	100.0	540	55.00	45.00	80.00	89.17	0.540
90.0	900.0	88.5	478	54.00	45.00	80.00	88.25	0.531
80.0	800.0	78.5	424	53.00	45.00	80.00	87.32	0.530
70.0	700.0	69.6	376	52.00	45.00	80.00	86.42	0.537
60.0	600.0	61.5	332	51.00	45.00	80.00	85.52	0.553
50.0	500.0	53.9	291	50.00	45.00	80.00	84.62	0.582
40.0	400.0	46.7	252	49.00	45.00	80.00	83.74	0.630
30.0	300.0	39.6	214	48.00	45.00	80.00	82.85	0.713
20.0	200.0	33.0	178	47.00	45.00	80.00	81.97	0.890
17.3	172.9	31.1	168	46.73	45.00	80.00	81.73	0.972

PROJECT - YORK ORDER NO. 00-105349-01
 YORK SERIAL NO. G8KM25056S

SALES ENGINEER - RCN
 CUSTOMER - IBM - POUGHKEEPSIE

DATE: 02/18/05
 PAGE: 2 OF 3

ESTIMATED PART LOAD PERFORMANCE:

LAST DATA LINE IS PART LOAD OPERATING LIMIT WITH PRE-ROTATION VANES

% LOAD	CAP (TONS)	% POWER (KW)	INPUT (F)	ERPT (F)	ELFT (F)	CEFT (F)	CLFT (F)	PERF (KW/TON)
100.0	1000.0	100.0	501	55.00	45.00	75.00	84.09	0.501
90.0	900.0	88.8	445	54.00	45.00	75.00	83.17	0.494
80.0	800.0	78.2	397	53.00	45.00	75.00	82.26	0.496
70.0	700.0	70.5	353	52.00	45.00	75.00	81.36	0.504
60.0	600.0	62.5	313	51.00	45.00	75.00	80.47	0.522
50.0	500.0	55.1	276	50.00	45.00	75.00	79.59	0.552
40.0	400.0	47.9	240	49.00	45.00	75.00	78.71	0.600
30.0	300.0	40.9	205	48.00	45.00	75.00	77.83	0.683
20.0	200.0	33.9	170	47.00	45.00	75.00	76.95	0.850
15.0	150.4	30.7	154	46.50	45.00	75.00	76.52	1.024

% LOAD	CAP (TONS)	% POWER (KW)	INPUT (F)	ERPT (F)	ELFT (F)	CEFT (F)	CLFT (F)	PERF (KW/TON)
100.0	1000.0	100.0	467	55.00	45.00	70.00	79.01	0.467
90.0	900.0	89.5	418	54.00	45.00	70.00	78.10	0.464
80.0	800.0	80.1	374	53.00	45.00	70.00	77.20	0.468
70.0	700.0	71.5	334	52.00	45.00	70.00	76.32	0.477
60.0	600.0	63.8	298	51.00	45.00	70.00	75.43	0.497
50.0	500.0	56.3	263	50.00	45.00	70.00	74.55	0.526
40.0	400.0	49.0	229	49.00	45.00	70.00	73.68	0.572
30.0	300.0	42.0	196	48.00	45.00	70.00	72.81	0.653
20.0	200.0	34.9	163	47.00	45.00	70.00	71.93	0.815
15.0	150.4	31.5	147	46.50	45.00	70.00	71.50	0.978

% LOAD	CAP (TONS)	% POWER (KW)	INPUT (F)	ERPT (F)	ELFT (F)	CEFT (F)	CLFT (F)	PERF (KW/TON)
100.0	1000.0	100.0	440	55.00	45.00	65.00	73.95	0.440
90.0	900.0	90.0	396	54.00	45.00	65.00	73.05	0.440
80.0	800.0	80.7	355	53.00	45.00	65.00	72.16	0.444
70.0	700.0	72.5	319	52.00	45.00	65.00	71.28	0.456
60.0	600.0	64.8	285	51.00	45.00	65.00	70.40	0.475
50.0	500.0	57.3	252	50.00	45.00	65.00	69.53	0.504
40.0	400.0	50.0	220	49.00	45.00	65.00	68.66	0.550
30.0	300.0	42.7	188	48.00	45.00	65.00	67.79	0.627
20.0	200.0	35.5	156	47.00	45.00	65.00	66.92	0.780
15.0	150.4	32.0	141	46.50	45.00	65.00	66.48	0.938

PROJECT - YORK ORDER NO. 00-105349-01
 YORK SERIAL NO. GHEM280565

SALES ENGINEER - RCN
 CUSTOMER - IBM - POOGHEEPSIE

DATE: 02/18/05
 PAGE: 3 OF 3

ESTIMATED PART LOAD PERFORMANCE:

LAST DATA LINE IS PART LOAD OPERATING LIMIT WITH PRE-ROTATION VANES

% LOAD	CAP (TONS)	% POWER	INPUT (KW)	EEFT (F)	ELFT (F)	CEFT (F)	CLFT (F)	PERF (KW/TON)
100.0	1000.0	100.0	417	55.00	45.00	60.00	68.89	0.417
90.0	900.0	90.4	377	54.00	45.00	60.00	68.00	0.419
80.0	800.0	81.5	340	53.00	45.00	60.00	67.12	0.425
70.0	700.0	73.4	306	52.00	45.00	60.00	66.25	0.437
60.0	600.0	65.5	273	51.00	45.00	60.00	65.37	0.455
50.0	500.0	58.0	242	50.00	45.00	60.00	64.50	0.484
40.0	400.0	50.8	212	49.00	45.00	60.00	63.63	0.530
30.0	300.0	43.4	181	48.00	45.00	60.00	62.77	0.603
20.0	200.0	35.7	149	47.00	45.00	60.00	61.90	0.745
15.0	150.4	32.1	134	46.50	45.00	60.00	61.47	0.891

% LOAD	CAP (TONS)	% POWER	INPUT (KW)	EEFT (F)	ELFT (F)	CEFT (F)	CLFT (F)	PERF (KW/TON)
100.0	1000.0	100.0	399	55.00	45.00	55.00	63.84	0.399
90.0	900.0	90.5	361	54.00	45.00	55.00	62.96	0.401
80.0	800.0	82.0	327	53.00	45.00	55.00	62.09	0.409
70.0	700.0	73.7	294	52.00	45.00	55.00	61.22	0.420
60.0	600.0	65.9	263	51.00	45.00	55.00	60.35	0.438
50.0	500.0	58.4	233	50.00	45.00	55.00	59.48	0.466
40.0	400.0	50.9	203	49.00	45.00	55.00	58.62	0.507
30.0	300.0	43.4	173	48.00	45.00	55.00	57.75	0.577
20.0	200.0	35.6	142	47.00	45.00	55.00	56.88	0.710
15.0	150.4	31.3	125	46.50	45.00	55.00	56.45	0.831

Figure A.12 - 1000 Ton Chiller Performance Data

B/027 #1 #2

PROJECT - YORK ORDER NO. 99-105265-03
YORK SERIAL NO. SGHM595540

SALES ENGINEER - RCN
CUSTOMER - IBM - POUGHKEEPSIE

DATE: 02/11/05
PAGE: 1 OF 3

MODEL	YKGTBJ2-CBE	(MOTOR SPECIFIED BY USER)
REFRIGERANT	134A	GEAR CODE
RATED CAPACITY (TONS)	1200	EF(SPEC)
INPUT POWER (KW)	675	SPECIFIED CAPACITY (TONS)
VOLTAGE/Hz	2300/60	MAX MOTOR LOAD(KW)
DIFFUSER	FIXED	
FLA	190	LEA
INRUSH (AMPS)	1230	
FULL LOAD (KW/TON)	0.563	NPLV
STARTER TYPE (7) ACROSS-THE-LINE	(100%INRUSH) - 3 LEAD	

	EVAPORATOR	CONDENSER
FLUID	WATER	WATER
TUBE	181*	230*
PASSES	2*	2*
FOUL FACTOR	0.00025*	0.00025*
FLUID ENT TEMP(F)	58.00	85.00*
FLUID LEV TEMP(F)	46.00*	96.08
FLUID FLOW(GPM)	2400.0*	3000.0*
FLUID PRDROP (FT)	13.7	10.8

ESTIMATED PART LOAD PERFORMANCE:

LAST DATA LINE IS PART LOAD OPERATING LIMIT WITH PRE-ROTATION VANES

% LOAD	CAP (TONS)	% POWER (KW)	INPUT (F)	EEFT (F)	ELFT (F)	CEFT (F)	CLFT (F)	PERF (KW/TON)
100.0	1200.0	100.0	675	58.00	46.00	85.00	96.08	0.562
90.0	1080.0	88.4	597	56.80	46.00	85.00	94.95	0.553
80.0	960.0	77.6	524	55.60	46.00	85.00	93.83	0.546
70.0	840.0	68.3	461	54.40	46.00	85.00	92.73	0.549
60.0	720.0	60.0	405	53.20	46.00	85.00	91.64	0.562
53.6	643.3	55.0	371	52.43	46.00	85.00	90.95	0.577

PROJECT - YORK ORDER NO. 99-105265-03
YORK SERIAL NO. SGHM595540

SALES ENGINEER - RCN
CUSTOMER - IBM - POUGHKEEPSIE

DATE: 02/11/05
PAGE: 2 OF 3

ESTIMATED PART LOAD PERFORMANCE:

LAST DATA LINE IS PART LOAD OPERATING LIMIT WITH PRE-ROTATION VANES

% LOAD	CAP (TONS)	% POWER (KW)	INPUT (F)	EEFT (F)	ELFT (F)	CEFT (F)	CLFT (F)	PERF (KW/TON)
100.0	1200.0	100.0	623	58.00	46.00	80.00	90.97	0.519
90.0	1080.0	88.4	551	56.80	46.00	80.00	89.85	0.510
80.0	960.0	78.2	487	55.60	46.00	80.00	88.74	0.507
70.0	840.0	69.2	431	54.40	46.00	80.00	87.66	0.513
60.0	720.0	60.8	379	53.20	46.00	80.00	86.58	0.526
50.0	600.0	53.3	332	52.00	46.00	80.00	85.51	0.553
40.0	480.0	45.9	286	50.80	46.00	80.00	84.45	0.596
30.0	360.0	38.8	242	49.60	46.00	80.00	83.39	0.672
20.0	240.0	31.8	198	48.40	46.00	80.00	82.33	0.825
15.0	180.5	28.4	177	47.80	46.00	80.00	81.81	0.981

% LOAD	CAP (TONS)	% POWER (KW)	INPUT (F)	EEFT (F)	ELFT (F)	CEFT (F)	CLFT (F)	PERF (KW/TON)
100.0	1200.0	100.0	577	58.00	46.00	75.00	85.86	0.481
90.0	1080.0	88.9	513	56.80	46.00	75.00	84.76	0.475
80.0	960.0	79.0	456	55.60	46.00	75.00	83.67	0.475
70.0	840.0	70.2	405	54.40	46.00	75.00	82.60	0.482
60.0	720.0	62.0	358	53.20	46.00	75.00	81.53	0.497
50.0	600.0	54.4	314	52.00	46.00	75.00	80.47	0.523
40.0	480.0	47.1	272	50.80	46.00	75.00	79.41	0.567
30.0	360.0	39.9	230	49.60	46.00	75.00	78.36	0.639
20.0	240.0	32.8	189	48.40	46.00	75.00	77.31	0.788
15.0	180.5	29.3	169	47.80	46.00	75.00	76.79	0.937

% LOAD	CAP (TONS)	% POWER (KW)	INPUT (F)	EEFT (F)	ELFT (F)	CEFT (F)	CLFT (F)	PERF (KW/TON)
100.0	1200.0	100.0	539	58.00	46.00	70.00	80.77	0.449
90.0	1080.0	89.4	482	56.80	46.00	70.00	79.68	0.446
80.0	960.0	79.8	430	55.60	46.00	70.00	78.61	0.448
70.0	840.0	71.1	383	54.40	46.00	70.00	77.54	0.456
60.0	720.0	63.1	340	53.20	46.00	70.00	76.49	0.472
50.0	600.0	55.5	299	52.00	46.00	70.00	75.43	0.498
40.0	480.0	48.2	260	50.80	46.00	70.00	74.38	0.542
30.0	360.0	40.8	220	49.60	46.00	70.00	73.34	0.621
20.0	240.0	33.6	181	48.40	46.00	70.00	72.29	0.754
15.0	180.5	30.1	162	47.80	46.00	70.00	71.77	0.898

PROJECT - YORK ORDER NO. 99-105265-03
 YORK SERIAL NO. SGRMS95540

SALES ENGINEER - RCN
 CUSTOMER - IBM - POUKKEEPSIE

DATE: 02/11/05
 PAGE: 3 OF 3

ESTIMATED PART LOAD PERFORMANCE:

LAST DATA LINE IS PART LOAD OPERATING LIMIT WITH PRE-ROTATION VANES

% LOAD	CAP (TONS)	% POWER	INPUT (KW)	EEFT (F)	ELFT (F)	CEFT (F)	CLFT (F)	PERF (KW/TON)
100.0	1200.0	100.0	508	58.00	46.00	65.00	75.69	0.423
90.0	1080.0	89.8	456	56.80	46.00	65.00	74.62	0.422
80.0	960.0	80.5	409	55.60	46.00	65.00	73.56	0.426
70.0	840.0	71.9	365	54.40	46.00	65.00	72.50	0.435
60.0	720.0	64.0	325	53.20	46.00	65.00	71.45	0.451
50.0	600.0	56.3	286	52.00	46.00	65.00	70.40	0.477
40.0	480.0	48.8	248	50.80	46.00	65.00	69.36	0.517
30.0	360.0	41.3	210	49.60	46.00	65.00	68.31	0.583
20.0	240.0	33.9	172	48.40	46.00	65.00	67.27	0.717
15.2	182.3	30.3	154	47.82	46.00	65.00	66.76	0.845

% LOAD	CAP (TONS)	% POWER	INPUT (KW)	EEFT (F)	ELFT (F)	CEFT (F)	CLFT (F)	PERF (KW/TON)
100.0	1200.0	100.0	482	58.00	46.00	60.00	70.63	0.402
90.0	1080.0	90.0	434	56.80	46.00	60.00	69.56	0.402
80.0	960.0	80.9	390	55.60	46.00	60.00	68.51	0.406
70.0	840.0	72.4	349	54.40	46.00	60.00	67.46	0.415
60.0	720.0	64.5	311	53.20	46.00	60.00	66.41	0.432
50.0	600.0	56.8	274	52.00	46.00	60.00	65.37	0.457
40.0	480.0	49.2	237	50.80	46.00	60.00	64.33	0.494
30.0	360.0	41.5	200	49.60	46.00	60.00	63.29	0.556
20.0	240.0	33.8	163	48.40	46.00	60.00	62.25	0.679
15.5	185.9	30.3	146	47.86	46.00	60.00	61.77	0.785

% LOAD	CAP (TONS)	% POWER	INPUT (KW)	EEFT (F)	ELFT (F)	CEFT (F)	CLFT (F)	PERF (KW/TON)
100.0	1200.0	100.0	459	58.00	46.00	55.00	65.58	0.382
90.0	1080.0	90.2	414	56.80	46.00	55.00	64.52	0.383
80.0	960.0	81.3	373	55.60	46.00	55.00	63.47	0.389
70.0	840.0	73.0	335	54.40	46.00	55.00	62.42	0.399
60.0	720.0	64.9	298	53.20	46.00	55.00	61.38	0.414
50.0	600.0	57.1	262	52.00	46.00	55.00	60.34	0.437
40.0	480.0	49.5	227	50.80	46.00	55.00	59.30	0.473
30.0	360.0	41.4	190	49.60	46.00	55.00	58.27	0.528
20.0	240.0	33.3	153	48.40	46.00	55.00	57.22	0.637
16.0	191.6	29.8	137	47.92	46.00	55.00	56.80	0.715

Figure A.13 - 1200 Ton Chiller Performance Data

Appendix B: Relating the Maximum Chassis' Inlet Temperature to the Archimedes Number

The CFD results presented in Section 4.5 suggested a linear relationship between the rack's maximum inlet temperature and the Archimedes' number, i.e.,

$$\tau_{\max} = \tau_{\max}^*(Ar) + \alpha(Ar)(T_a - T_a^*), \quad (\text{B.1})$$

where, $\tau_{\max} = \max(T_{i,j}^{in}) - T_a$ and “*” corresponds to a reference value. It is expected that $\alpha \rightarrow 1$ as $Ar \rightarrow 0$ (i.e., a δT change in T_a will cause a δT change in $\max(T_{i,j}^{in})$). In this Appendix, CFD is used to determine the functional dependence of α and τ_{\max}^* on Ar .

To demonstrate the idea, we revisit the CFD results presented in Chapter 6 that were used for developing load placement rules. Seven CFD cases are selected at random, with Ar ranging from 0.54 to 1.17. Each case was run at four supply air temperatures (10 °C, 14 °C, 18 °C and 20 °C) and the temperature distribution at the inlet to the racks is recorded. Figures B.1 and B.2 plot the mass-weighted average inlet temperature at several of the chassis in the upper section of the racks for the example cases of $Ar = 0.54$ and 1.17, respectively. The solid lines represent the CFD results and the dashed lines represent the expected behavior if a δT change in the supply air temperature resulted in a δT change everywhere. As before, at the higher Ar , there is more deviation from the δT behavior;

however, the temperature response does remain linear. Therefore, we proceed with trying to develop a functional relationship such as the one in Equation B.1.

To start, we focus on developing an expression for the slope from the seven selected cases. Figure C.3 plots $\max(T_{i,j}^{in}) - \max(T_{i,j}^{in*})$ versus $T_a - T_a^*$ for each of the seven cases. The figure shows that the slope of all the cases is nearly 45° ; therefore, concluding that $\alpha = 1.0$ in Equation B.1, and is not a function of Ar .

Since the slope of Equation B.1 was not a function of Ar , the focus becomes finding a functional relationship for the intercept of Equation B.1. Figure B.4 plots

$\max(T_{i,j}^{in}) - T_a^*$ versus $T_a - T_a^*$. The figure shows that $T_{\max} - T_a^*$ is a monotonically increasing function of Ar , for the cases considered here. Additionally, two cases were randomly selected that had an $Ar = 0.75$. Both of these cases result in the same intercept, providing confidence that the intercept is a function of Ar only.

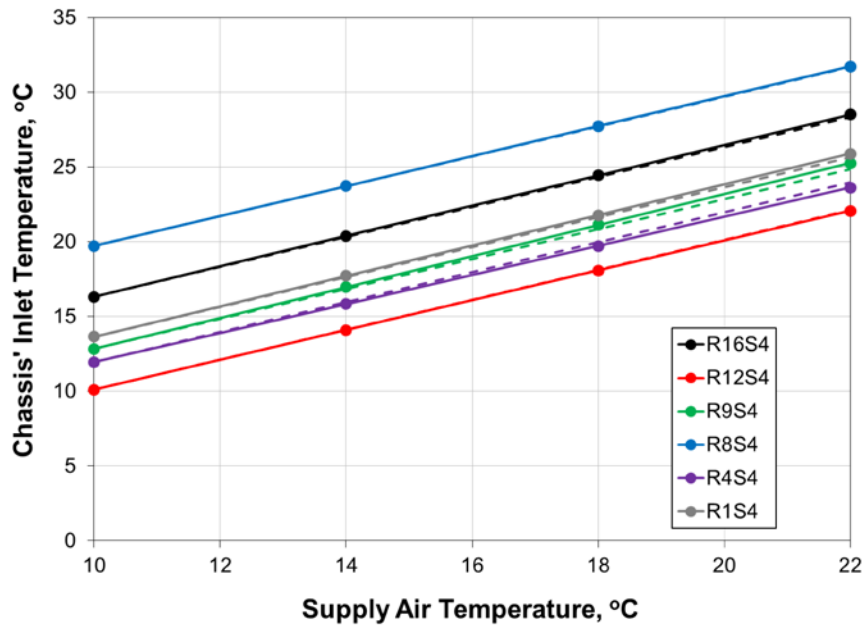


Figure B.1 – Chassis' Inlet Temperature for Ar = 0.54

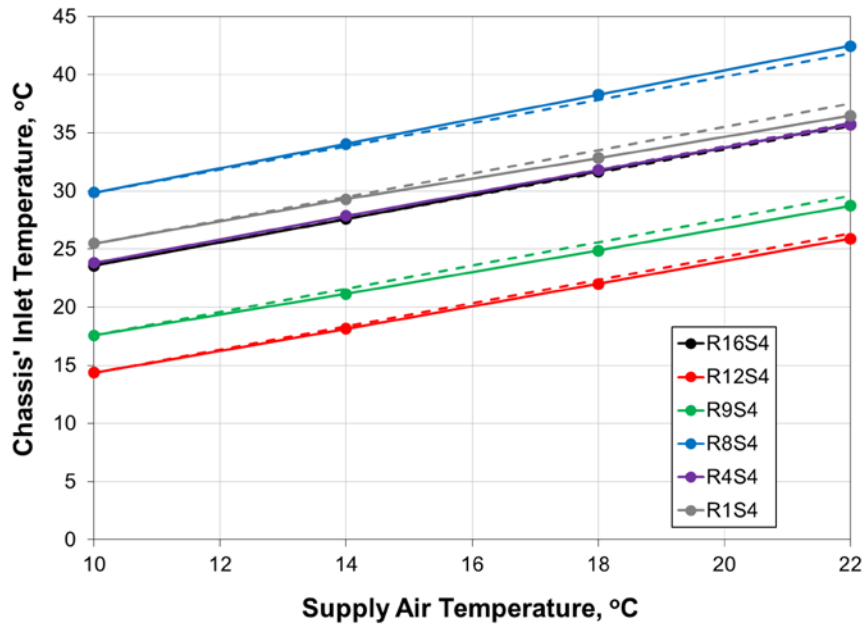


Figure B.2 – Chassis' Inlet Temperature for Ar = 1.17

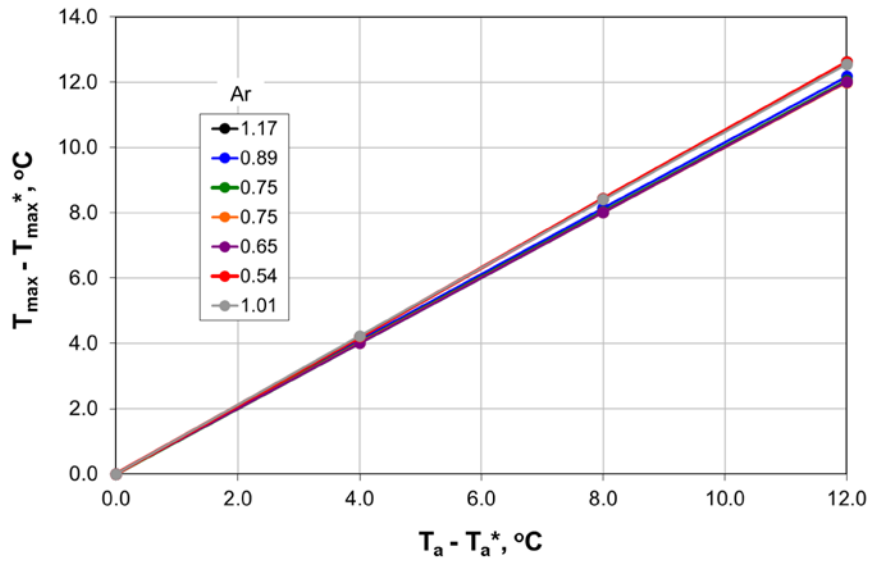


Figure B.3 - Slope of Temperature Curve

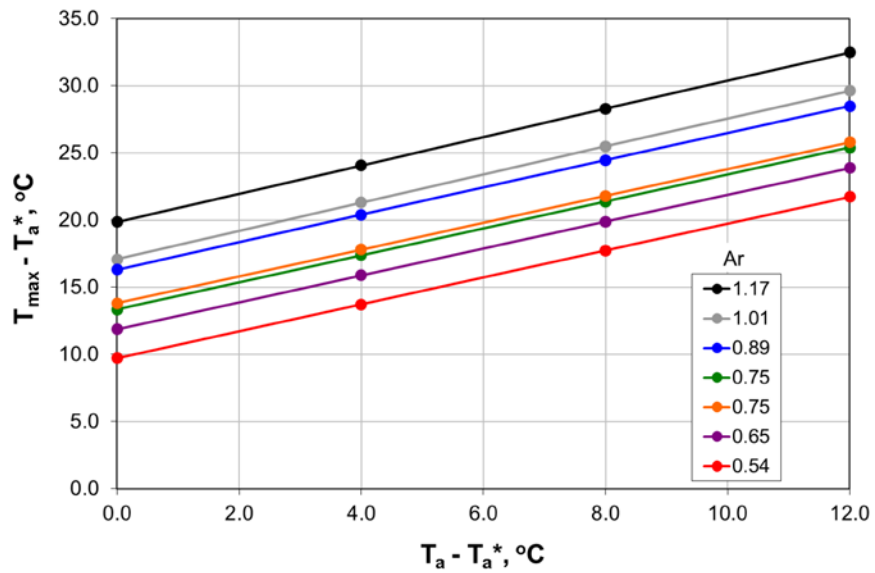


Figure B.4 - Intercept of Temperature Function

Based on the analysis just presented, it is anticipated that the maximum inlet temperature, $\max(T_{i,j}^{in})$ for any case can be found as a function of Ar by,

$$T_{\max} = \tau_{\max}^*(Ar) + T_a \quad (B.2)$$

Figure B.5 plots τ_{\max}^* as a function of Ar for the seven randomly selected cases. A second-order polynomial was selected to fit the data. This function provides the expected behavior that $\tau_{\max}^* \rightarrow 0$ as $Ar \rightarrow 0$.

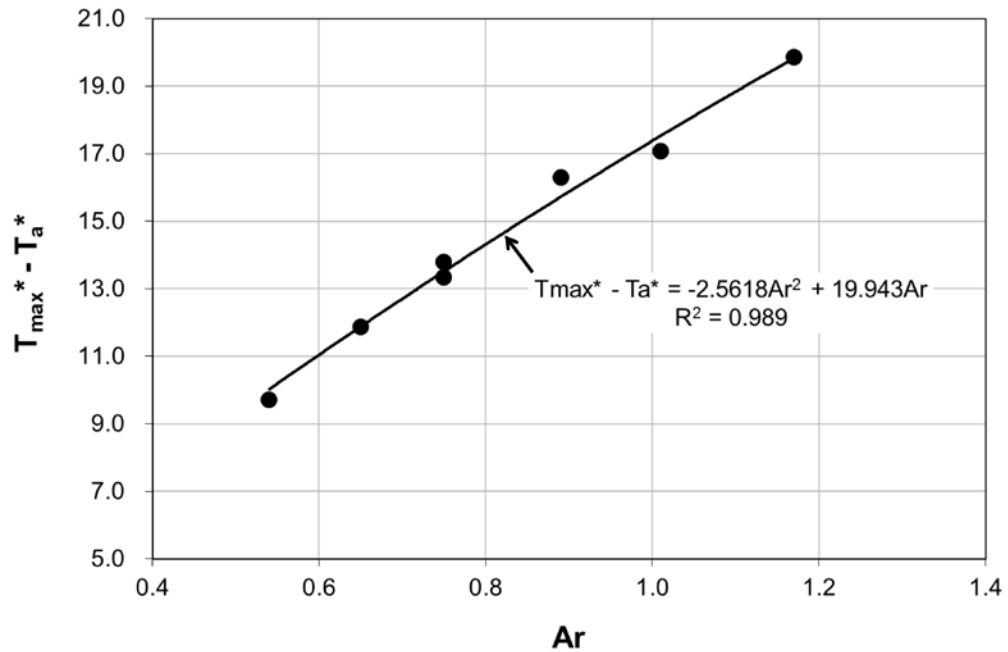


Figure B.5 – Functional Relationship for $(T_{\max} - T_a)^*$ vs. Ar

With a relationship in place, we test the predictability of the method with the remainder of the CFD cases done in Chapter 6. Figure B.6 plots the predicted maximum inlet temperature versus the actual maximum inlet temperature obtained from CFD. If the agreement was perfect, all the points would fall on the 45° dashed line. The simple method proposed here does a reasonable job of predicting the maximum inlet temperature as only a function of the data centers Archimedes number. The maximum temperature errors of the 18 cases considered was ~3.5°C.

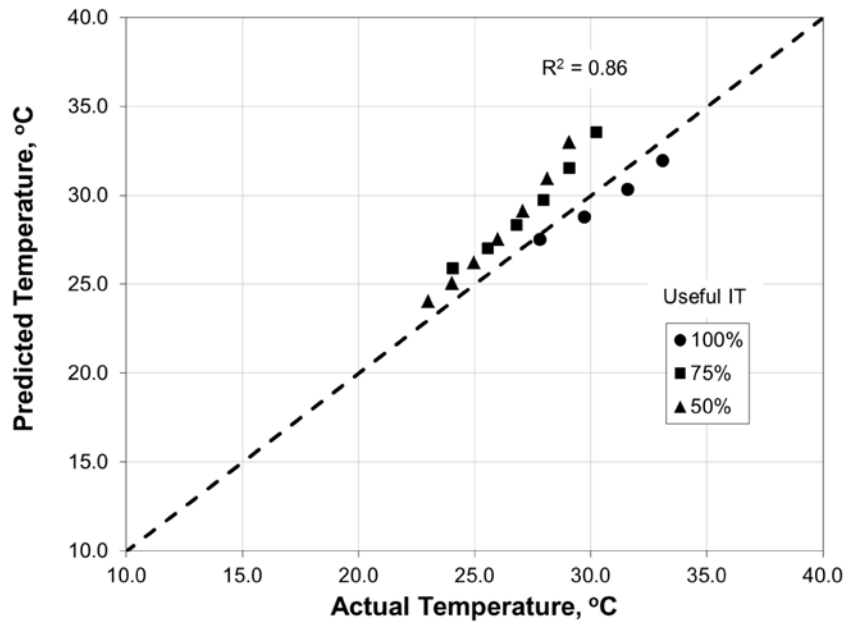


Figure B.6 - Validation of Temperature versus Ar Function

The proposed method does not provide the most accurate estimate of the maximum inlet temperature and hence was the reason the method of proper orthogonal decomposition was pursued in Chapter 6. However, it certainly provides a very simple tool for predicting

the maximum inlet temperature. The results of Figure B.6 clearly show two distinct slopes, one for the 100% useful IT case and another for the reduced IT load scenarios. This behavior suggests that including variables besides the Archimedes number that may improve the regression results. For example, the change in recirculation patterns caused by turning to *idle* chassis seems to affect the results, as seen by the two distinct sets of points in Figure B.6. This method provides a starting point for the future development of simple tools to predict the maximum inlet temperature. The development of fast, accurate and simple tools such as these may facilitate predictive methods for load placement in the instance where temperature sensors at the inlet of the IT equipment are not available.

Appendix C: Using Design of Experiments to Investigate Optimum Load Placement

If we consider a data center that has forty racks, each containing four chassis, the data center operator has the choice of operating any of these chassis, to produce the required IT load. In order to find the effect each chassis has on the thermal environment, one would need to perform experiments that varied each chassis one at a time to cover all possible combinations. Assuming that each variable can take one of two states (e.g., *idle* and *on*) the number of experiments or computations needed would be 2^{246} , which is prohibitively expensive, even computationally. In the literature, this technique is referred to as a full factorial experimental design.

The design and analysis of experiments (D.O.E.) (Montgomery, 2001) is a well-developed technique for reducing a large number of experimental scenarios to a manageable number, while still obtaining the same conclusions. This is known as a fractional experimental design. The use of fractional D.O.E. assumes that higher order interactions have a much smaller effect on the response than low order interactions. A consequence of reducing the number of experiments is that higher-order interactions get aliased together and the magnitude of these interactions cannot be uniquely determined, only the combined effect can. As an example, consider an experiment where a variable T is being predicted as a function of three factors (A, B, C) each of which has two levels (high and low), for a total of $2^3 = 8$ total experiments. Figure C.1 shows a schematic of

the experimental design, where each corner of the cube represents one experiment that will be conducted.

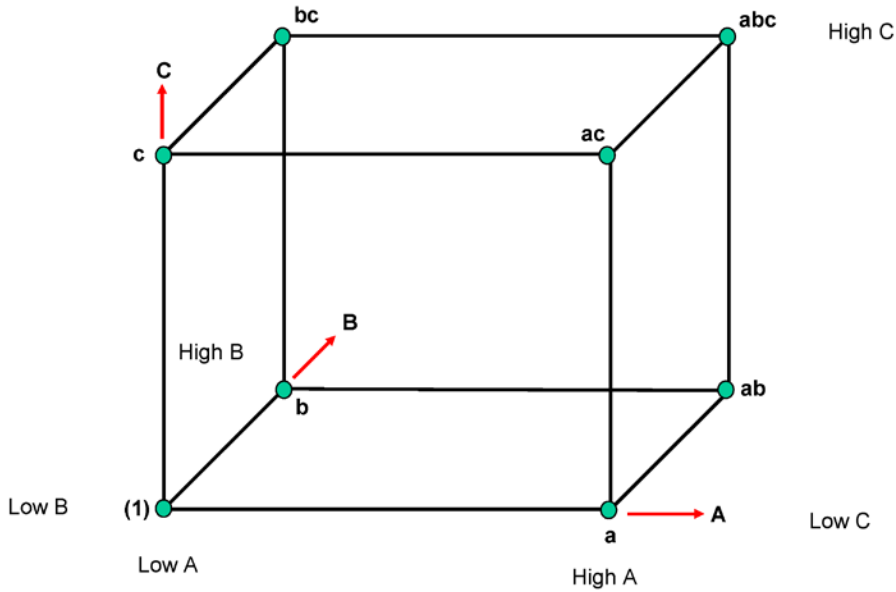


Figure C.1 - Experimental Cube for a 2^3 Factorial Design

The naming scheme in Figure C.1 is as follows: a lowercase letter represents a variable at its high level (i.e. experiment “a” has factor A at its high level and factors B and C at their low level; experiment “ab” has factors A and B at their high levels and factor C at its low level; experiment (1) represents all factors at their low level). The following expansion for the variable T can be written,

$$\begin{aligned}
 T = T_o + \Delta A \frac{\partial T}{\partial A} + \Delta B \frac{\partial T}{\partial B} + \Delta C \frac{\partial T}{\partial C} + \frac{2\Delta A \Delta B}{2!} \frac{\partial^2 T}{\partial A \partial B} + \frac{2\Delta A \Delta C}{2!} \frac{\partial^2 T}{\partial A \partial C} + \\
 \frac{2\Delta B \Delta C}{2!} \frac{\partial^2 T}{\partial B \partial C} + \frac{3\Delta A \Delta B \Delta C}{3!} \frac{\partial^3 T}{\partial A \partial B \partial C}
 \end{aligned}
 \tag{C.1}$$

where, T_o is the mean value of T and the effect coefficients are the partial derivatives. The purpose of experimental design is to determine the magnitude of each of the effect coefficients. The value of the coefficient is determined by running a set of appropriate experiments. For example, estimating $\frac{\partial T}{\partial A}$ by finite difference results in,

$$\frac{\partial T}{\partial A} = \frac{T|_{high} - T|_{low}}{2\Delta A} . \quad (C.2)$$

In this case, since this is a partial derivative with respect to A , $T|_{high}$ represents the average value of all the experiments where A is at its high level and $T|_{low}$ represents the average value of all the experiments where A is at its low level. Relating this to the naming scheme from Figure C.1, the value of the effect coefficient can be written as,

$$\frac{\partial T}{\partial A} = \frac{1}{4} [-(1) + a - b + ab - c + ac - bc + abc] \quad (C.3)$$

A similar procedure can be done for each effect coefficient. These results are summarized in Table C.1. To determine the value of an effect coefficient from the table, the column of the effect coefficient is multiplied by its corresponding treatment combination. These are then summed together to obtain the effect value. For example, the value of effect coefficient AC is “+(1)-a+b-ab-c+ac-bc+abc”. Column I is the average of all eight experiments.

Table C.1 – Factorial Design Effect Coefficient Table

Treatment Combination	Effect							
	I	A	B	AB	C	AC	BC	ABC
(1)	+	-	-	+	-	+	+	-
a	+	+	-	-	-	-	+	+
b	+	-	+	-	-	+	-	+
ab	+	+	+	+	-	-	-	-
c	+	-	-	+	+	-	-	+
ac	+	+	-	-	+	+	-	-
bc	+	-	+	-	+	-	+	-
abc	+	+	+	+	+	+	+	+

Table C.1 shows that all of the columns are independent (i.e., none are identical); therefore, all of the effects are independent and can be uniquely determined by running all eight experiments. However, in many case running the full factorial set of experiments is too cost or time prohibitive, especially for experiments where a large number of variables need to be considered. The fractional factorial design of experiments specifically addresses the situation in which only a subset of the experiments is possible. For example, consider the case where only four of the eight experiments in Figure C.1 can be performed. The goal of the design of experiments methodology (Montgomery, 2001) is to find the subset of experiments to run that will reduce aliasing. One of the key assumptions needed is that higher-order effects (i.e., ABC) have less influence on the response than low-order effects (i.e., A). Therefore, the higher-order effects are aliased

with the lower order effects and the outcome is a high order effect plus a small residual effect due to the aliasing with a higher order term. For an experimental design where four of eight experiments can be done, the D.O.E. methodology chooses to select only those experiments where ABC is at its high level (+) in Table C.1. If this is done, the set of four experiments in Table C.2 is obtained.

Table C.2 - 2^{8-4} Fractional Factorial Design Effect Coefficient Table

Treatment Combination	Factorial Effect							
	I	A	B	AB	C	AC	BC	ABC
a	+	+	-	-	-	-	+	+
b	+	-	+	-	-	+	-	+
c	+	-	-	+	+	-	-	+
abc	+	+	+	+	+	+	+	+

An examination of Table C.2 shows that not all of the columns are independent. For instance, column A is identical to column BC; therefore, when the effect of A is estimated, the result is actually the combined effect of A and BC. This is referred to as aliasing. The magnitude of aliasing in any design is determined by computing its resolution and luckily can be determined before running the actual experiments. The most common resolutions of an experimental design are defined as:

- Resolution III – no main effect is aliased with any other main effect, but main effects are aliased with two-factor interactions. Two-factor interactions can be aliased with each other.
- Resolution IV – no main effect is aliased with any other main effect or two-factor interaction. Two-factor interactions are aliased with each other.
- Resolution V – no main effect or two-factor interaction is aliased with any other main effect or two-factor interaction. Two-factor interactions are aliased with three-factor interactions.

The above example was a resolution III fractional factorial design of experiments since the main effect A was aliased with the two-factor interaction BC.

The design of experiments methodology provides a means of reducing a large set of experimental runs into a reasonable number, while preserving the response. DOE was used to investigate thermally aware, energy-based load placement using CFD to surmise whether better load placement options exist than the ones considered in Section 6.2. For this exercise, the same data center configuration as given in Section 6.1 was considered, where the chassis can be operated at an *idle* or *on* state. However, for this analysis each rack was divided into 2 sections – an upper and lower section – where each section had the power and flow characteristics of two chassis. This results in the option to place the load on any of the 32 variables (or sections) in the 16 racks. If the full factorial design was to be completed, $2^{32} = 4.29 \times 10^9$ experiments would be needed, which is impossible even using a high-performance computer. Therefore, a resolution IV experimental design

was selected, which results in 64 CFD cases to run in order to uniquely determine the main effects. These 64 CFD cases were run for a useful IT load of 50% and a $\psi_T = 0.80$. Table C.3 provides a detail of each of the CFD runs necessary, where a (-1) means that section is placed in an *idle* state and a (+1) means that section is placed in an *on* state. For each of the 64 CFD cases, the chassis' inlet temperature distribution and CRAH's inlet temperature was recorded. The simple model detailed in Section 2.1 was used to evaluate the energy consumption of each of the cases. For each case, the δT_a transformation shown in Equation 6.2 was applied to determine the supply air temperature that identically meets the constraint given by Equation 6.1.

Table C.3 - Experimental Design for a 2^{32-26} Fractional Factorial Design

Run ID	Variable																																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32		
1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1		
2	-1	-1	-1	-1	-1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	
3	-1	-1	-1	-1	1	-1	1	1	1	1	1	1	1	1	-1	-1	-1	-1	-1	-1	-1	1	1	1	1	1	1	1	1	-1	-1	-1	-1	
4	-1	-1	-1	-1	1	1	-1	-1	-1	-1	-1	-1	-1	-1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	-1	-1	-1	-1	
5	-1	-1	-1	1	-1	-1	1	1	1	1	1	-1	-1	-1	1	1	1	1	-1	-1	-1	1	1	1	1	1	1	-1	-1	1	1	1	-1	
6	-1	-1	-1	1	-1	1	-1	-1	-1	-1	1	1	1	1	-1	-1	-1	-1	1	1	1	1	1	1	1	1	-1	-1	-1	1	1	1	-1	
7	-1	-1	-1	1	1	-1	-1	-1	-1	-1	1	1	1	1	1	1	1	1	-1	-1	-1	-1	-1	-1	-1	-1	1	1	1	1	1	1	-1	
8	-1	-1	-1	1	1	1	1	1	1	1	-1	-1	-1	-1	-1	-1	-1	-1	1	1	1	-1	-1	-1	-1	-1	1	1	1	1	1	1	-1	
9	-1	-1	1	-1	-1	-1	1	1	-1	-1	1	1	-1	-1	1	1	-1	-1	1	1	-1	1	1	1	-1	-1	1	1	-1	1	1	1	1	
10	-1	-1	1	1	-1	1	-1	-1	1	1	-1	-1	1	1	-1	-1	1	1	-1	-1	1	1	1	1	-1	-1	1	1	-1	1	1	1	-1	1
11	-1	-1	1	-1	1	-1	-1	-1	1	1	-1	-1	1	1	1	1	-1	-1	1	1	-1	-1	-1	1	1	-1	-1	1	1	1	1	-1	1	
12	-1	-1	1	-1	1	1	1	1	-1	-1	1	1	-1	-1	-1	-1	1	1	-1	-1	1	-1	-1	1	1	-1	-1	1	1	1	1	-1	1	
13	-1	-1	1	1	-1	-1	-1	-1	1	1	1	1	-1	-1	-1	-1	1	1	1	1	1	-1	-1	-1	1	1	1	1	-1	-1	-1	1	1	
14	-1	-1	1	1	-1	1	1	1	-1	-1	-1	-1	1	1	1	1	-1	-1	-1	-1	-1	1	-1	-1	-1	1	1	1	1	-1	-1	-1	1	1
15	-1	-1	1	1	1	-1	1	1	-1	-1	-1	-1	1	1	-1	-1	1	1	1	1	1	-1	1	1	-1	-1	-1	-1	1	-1	-1	-1	1	1
16	-1	-1	1	1	1	1	-1	-1	1	1	1	1	-1	-1	1	1	-1	-1	-1	-1	1	1	1	1	-1	-1	-1	-1	1	-1	-1	1	1	
17	-1	1	-1	-1	-1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	1	-1	1	-1	1	-1	1	1	-1	1	1	1	
18	-1	1	-1	-1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	1	-1	1	1
19	-1	1	-1	-1	1	-1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1
20	-1	1	-1	-1	1	1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	1
21	-1	1	-1	1	-1	-1	-1	1	-1	1	1	-1	1	-1	-1	1	-1	1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	1
22	-1	1	-1	1	-1	1	1	-1	1	-1	-1	1	-1	1	1	-1	1	-1	-1	1	-1	1	1	-1	1	-1	1	-1	1	-1	1	-1	1	1
23	-1	1	-1	1	1	-1	1	-1	1	-1	-1	1	-1	1	-1	1	-1	1	1	-1	1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1
24	-1	1	-1	1	1	1	-1	1	-1	1	1	-1	1	-1	1	-1	1	-1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1
25	-1	1	1	-1	-1	-1	-1	1	1	-1	-1	1	1	-1	-1	1	1	-1	-1	1	1	-1	1	1	-1	-1	1	-1	1	-1	1	-1	1	-1
26	-1	1	1	-1	-1	1	1	-1	-1	1	1	-1	-1	1	1	-1	-1	1	1	-1	-1	-1	1	1	-1	-1	1	1	-1	1	1	-1	1	-1
27	-1	1	1	-1	1	-1	1	-1	-1	1	1	-1	-1	1	-1	1	1	-1	-1	1	1	1	-1	-1	1	1	-1	-1	1	1	-1	1	-1	1
28	-1	1	1	-1	1	1	-1	1	1	-1	-1	1	1	-1	1	-1	-1	1	1	-1	1	-1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1
29	-1	1	1	1	-1	-1	1	-1	-1	1	-1	1	1	-1	1	-1	-1	1	-1	1	1	1	-1	-1	1	-1	1	-1	1	1	-1	1	-1	1
30	-1	1	1	1	-1	1	-1	1	1	-1	1	-1	-1	1	-1	1	1	-1	1	-1	-1	1	-1	-1	1	-1	1	1	1	-1	-1	-1	1	-1
31	-1	1	1	1	1	-1	-1	1	1	-1	1	-1	-1	1	1	-1	-1	1	-1	1	1	-1	1	1	-1	1	-1	-1	1	-1	-1	1	-1	-1
32	-1	1	1	1	1	1	1	-1	-1	1	-1	1	1	-1	-1	1	1	-1	1	-1	-1	-1	1	1	-1	1	-1	-1	1	-1	-1	1	-1	-1

Figure C.2 plots the normalized cooling energy consumption of each of the 64 cases considered, along with a line for Scenario 7, which removed the load from the hottest chassis. The results show that most of the cases considered are inferior to Scenario 7, which was considered in Chapters 6 and 7. The results show that only one of the sixty-four DOE placement scenarios resulted in lower power consumption than Scenario 7 and even this scenario resulted in less than a 1% reduction in the cooling power consumption. Figure C.3 shows the load placement distributions for both Scenario 7 and the best option for the D.O.E. work. The placement distributions both remove the load from those servers that are expected to be the hottest – those in the upper portion of the rack. However, the best D.O.E. case tends to remove load from all of the upper chassis and not the side chassis – as was done in Scenario 7. Even though the load placement is different, it turns out that in many cases, the chassis that were selected in the D.O.E. case 47 were less than 1°C away from being selected by the “turning off the hottest” algorithm. The rigorous approach to selecting a representative subset of experiments to run using D.O.E. provides confidence that the selection of the “turning off the hottest” algorithm is a robust solution and that a far-superior solution would not be found using a more detailed optimization study, for example using a genetic algorithm as proposed by Tang et al. (2008). The “turning off the hottest” algorithm also provides implementation benefits as discussed compared to the D.O.E. case, which really has no geometrical characteristics to develop an easy-to-implement heuristic.

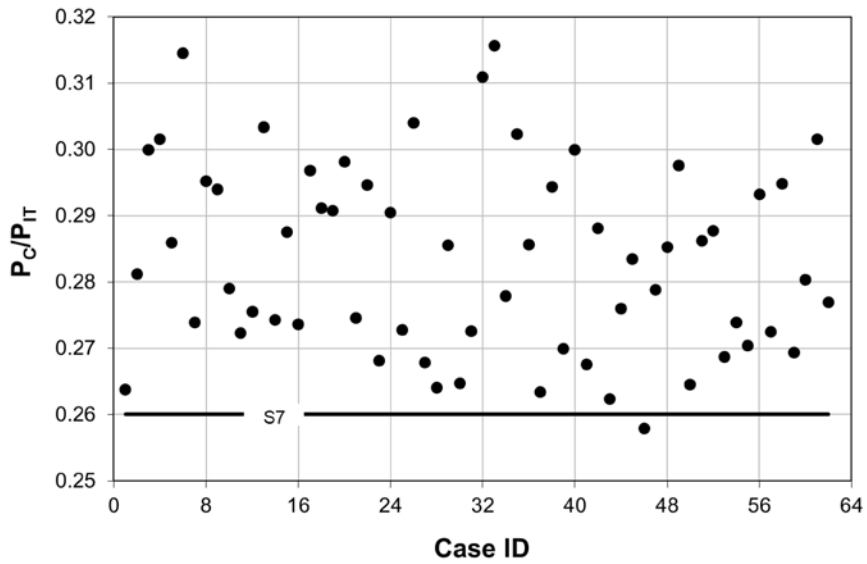
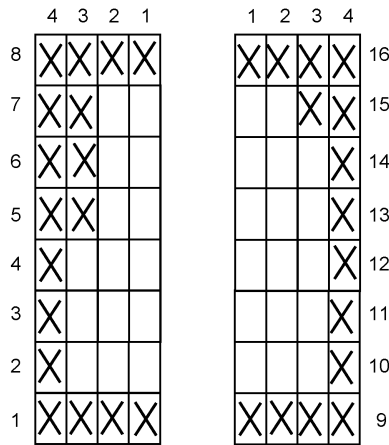


Figure C.2 - Normalized Energy Consumption Results for all DOE Cases

Scenario 7 – Turn off hottest



Best DOE Case - #47

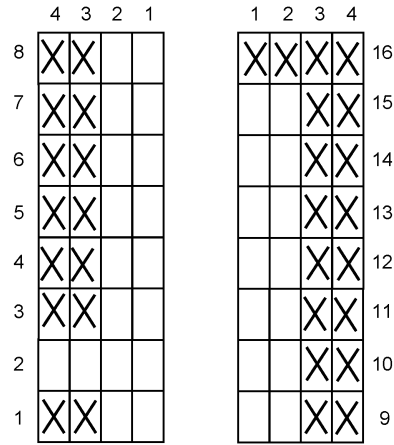


Figure C.3 - Comparison of Load Placement between Scenario 7 and Best DOE Case

9 References

- Abdelmaksoud, W., Dang, T.Q., Khalifa, H.E., Elhadidi, B., Schmidt, R., and Iyengar, M., 2010, “Experimental and Computational Study of Perforated Floor Tile in Data Centers,” *Proceedings of IEEE ITherm*, Las Vegas, NV, June 2 - 5.
- Abdelmaksoud, W., Khalifa, H.E., Dang, T.Q., Schmidt, R., and Iyengar, M., 2010, “Improved CFD Modeling tranof a Small Data Center Test Cell,” *Proceedings of IEEE ITherm*, Las Vegas, NV, June 2 – 5.
- Arndt, R.E.A., Long, D.F., and Glauser, M.N., 1997, ”The Proper Orthogonal Decomposition of Pressure Fluctuations Surrounding a Turbulent Jet,” *Journal of Fluid Mechanics*, Vol. 340, pp. 1 – 33.
- ASHRAE, 2000, *HVAC Systems and Equipment*, Atlanta: American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc.
- ASHRAE, 2004a, “ANSI/ASHRAE Standard 62.1-2004 – Ventilation for Acceptable Indoor Air Quality,” Atlanta: American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc.

ASHRAE, 2004b, “ANSI/ASHRAE Standard 90.1-2004 – Energy Standard for Buildings except Low-Rise Residential Buildings,” Atlanta: American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc.

ASHRAE, 2008a, “ASHRAE Environmental Guidelines for Datacom Equipment: Expanding the Recommended Environmental Envelope,” Atlanta: American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc.

ASHRAE, 2008b, *Best Practices for Datacom Facility Energy Efficiency*, Atlanta: American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc.

ASHRAE, 2008c, *High Density Data Centers: Case Studies and Best Practices*, Atlanta: American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc.

ASHRAE, 2011a, “2011 Thermal Guidelines for Data Processing Environments – Expanded Data Center Classes and Usage Guidance,” Atlanta: American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc.

ASHRAE, 2011b, *ASHRAE Technical Committee 9.9 Mission Critical Facilities, Technology Spaces and Electronic Equipment*, <http://tc99.ashraetcs.org/>.

Applied Math Modeling Inc., 2011, *CoolSim*, <http://www.koolsim.com>.

- Bash, C., Patel, C., and Sharma, R., 2006, "Dynamic Thermal Management of Air Cooled Data Centers," *Proceedings of IEEE ITherm*, San Diego, CA, May 30 – June 2.
- Bejan, A., 1982, *Entropy Generation through Heat and Fluid Flow*. Wiley: New York.
- Bejan, A., 1995, *Entropy Generation Minimization*, CRC Press: Boca Raton, Florida.
- Bejan, A. and Ledezma, G.A., 1996, "Thermodynamic Optimization of Cooling Techniques for Electronic Packages." *Int. J. Heat Mass Transfer*, Vol. 39 (6), pp. 1213 - 1221.
- Boucher, T., Auslander, D., Bash, C., Federspiel, C., and Patel, C., 2004, "Viability of Dynamic Cooling Control in a Data Center Environment," *Proceedings of IEEE ITherm*, Las Vegas, Nevada, June 1 – 4.
- Braun, J.E., Mitchell, J.W., and Klein, S.A., 1987, "Performance and Control Characteristics of a Large Cooling System," *ASHRAE Transactions*, Vol. 93(1), pp. 1830-1852.
- Braun, J.E., Klein, S.A. and Mitchell, J.W., 1989. "Effectiveness Models for Cooling Tower and Cooling Coils," *ASHRAE Transactions*, Vol. 95(2), pp. 164 - 174.

Breen, T.J., Walsh, E., Punch, J., Shah, A., and Bash, C., 2010, "From Chip to Cooling Tower Data Center Modeling: Part I Influence of Server Inlet Temperature and Temperature Rise across Cabinet," *Proceedings of IEEE ITherm*, Las Vegas, NV, June 2 – 5.

Breen, T.J., Walsh, E., Punch, J., Shah, A., Bash, C., Kumari, N., and Cader, T., 2011, "From Chip to Cooling Tower Data Center Modeling: Chip Leakage Power and Its Impact on Cooling Infrastructure Energy Efficiency," *Proceedings of ASME InterPACK*, Portland, OR, July 6 - 8.

Bui-Thanh, T., Damodaran, M., and Willcox, K., 2003, "Proper Orthogonal Decomposition for Parametric Applications in Transonic Aerodynamics," *Proceedings of 15th AIAA Computational Fluid Dynamics Conference*, Orlando, FL.

Cruz, E., Joshi, Y., Iyengar, M., and Schmidt, R., 2009, "Comparison of Numerical Modeling to Experimental Data in a Small Data Center Test Cell," *Proceedings of ASME InterPACK*, San Francisco, CA, July 19 – 23.

DataCenterDynamics, 2011, <http://www.datacenterdynamics.com/>

- Davidson, T., 2009, "Dehumidification of Outdoor Air in Datacom Environments for Air-Side Economizer Operation" *ASHRAE Transactions*, Vol. 115(1), pp. 71 – 82.
- Demetriou, D.W., Khalifa, H.E., Iyengar, M., and Schmidt, R., 2011a, "Development and Experimental Validation of a Thermo-Hydraulic Model for Data Centers," *HVAC&R Research*, Vol. 17 (4), pp. 540 - 555.
- Demetriou, D.W., Khalifa, H.E., Iyengar, M., and Schmidt, R., 2011b, "Development of a Dynamic Energy Use Modeling Tool for Data Center Cooling Infrastructure," *ASHRAE Transactions*, Vol. 117 (2).
- Dygert, R., Russo, J., Dang, T., and Khalifa, H.E., 2009, "Modeling of the Human Body to Study the Personal Micro-Environment," *ASHRAE Transactions*, Vol. 115(2), pp. 407 – 420.
- Efe, M.O. and Ozbay, H., 2003, "Proper Orthogonal Decomposition for Reduced Order Modeling: 2D Heat Flow," *Proceedings of IEEE Conference on Control Applications*, Vol. 2, pp. 1273 – 1277.
- Elhadidi, B. and Khalifa, H.E., 2005, "Application of Proper Orthogonal Decomposition to Indoor Airflows," *ASHRAE Transactions*, Vol. 111(1), pp. 625 – 634.

Emerson, 2011, "Precision Cooling,"

<http://www.emersonnetworkpower.com/enUS/Products/PrecisionCooling>

Everson, R. and Sirovich, L., 1995, "Karhunen-Loeve Procedure for Gappy Data,"

Journal of Optical Society of America, Vol. 12(8), pp. 1657 – 1664.

Fink, J., 2008, "Impact of High Density Hot Aisles on IT Personnel Work Conditions,"

American Power Conversion White Paper #135.

Furihata, Y., Hayama, H., Enai, M., and Mori, T., 2003, "Efficient Cooling System for IT

Equipment in a Data Center," *Proceedings of IEEE INTELEC*, Pacifico

Yokohama, Japan, October 19 - 23.

Gondipalli, S., Sammakia, B., Bhopte, S., Schmidt, R., Iyengar, M., and Murray, B.,

2009, "Optimization of Cold Aisle Isolation Designs for a Data Center with Roofs and Doors Using Slits," *Proceedings of ASME InterPACK*, San Francisco, CA, July 19 – 23.

Gondipalli, S., Ibrahim, M., Bhopte, S., Sammakia, B., Murray, B., Ghose, K., Iyengar,

M., and Schmidt, R., 2010, "Numerical Modeling of Data Center with Transient Boundary Conditions," *Proceedings of IEEE ITherm*, Las Vegas, NV, June 2 – 5.

Google, 2011, "Google Data Center," <http://www.google.com/about/datacenters/>

The Green Grid, 2007, “The Green Grid Data Center Power Efficiency Metrics: PUE and DCiE,” White paper #6.

The Green Grid, 2008a, “The Green Grid Metrics: Data Center Infrastructure Efficiency (DCiE) Detailed Analysis,” White paper #14.

The Green Grid, 2008b, “The Green Grid Productivity Indicator,” White paper #15.

The Green Grid, 2008c, “Seven Strategies to Improve Data Center Cooling Efficiency,” White paper #11.

The Green Grid, 2010, “ERE: A Metric for Measuring the Benefit of Reuse Energy from a Data Center,” White paper #29.

The Green Grid, 2011, *The Green Grid*, <http://www.thegreengrid.org/>.

Hayama, H., Enai, M., Mori, T., and Kishita, M., 2003, “Planning of Air-Conditions Circulation Systems for Data Centers,” *Proceedings of IEEE INTELEC*, Pacifico Yokohama, Japan, October 19 - 23.

Healey, C., VanGilder, J., Sheffer, Z., and Zhang, X., 2011, “Potential-Flow Modeling for Data Center Applications,” *Proceedings of ASME InterPACK*, Portland, OR, July 6 – 8.

- Hellmer, B., 2010, "Consumption Analysis of Teleco and Data Center Cooling and Humidification Options," *ASHRAE Transactions*, Vol. 116(1), pp. 118 – 133.
- Herrlin, M., 2005, "Rack Cooling Effectiveness in Data Centers and Telecom Central Offices: The Rack Cooling Index (RCI)," *ASHRAE Transactions*, Vol. 111(2), pp. 725 – 731.
- Herrlin, M., 2008, "Airflow and Cooling Performance of Data Centers: Two Performance Metrics," *ASHRAE Transactions*, Vol. 114(2), pp. 182 – 187.
- Herrlin, M. and Compiano, C., 2010, "Top-Level Energy and Environmental Dashboard for Data Center Modeling," *ASHRAE Transactions*, Vol. 116(1), pp. 19 - 23.
- Holmes, P., Lumley, J., and Berkooz, G., 1996, "Turbulence, Coherent Structures, Dynamical Systems and Symmetry", *Cambridge Monographs on Mechanics*, Cambridge University Press.
- IBM, 2011, "IBM Tivoli Software," <http://www-01.ibm.com/software/tivoli/>
- Ibrahim, M., Gondipalli, S., Bhopte, S., Sammakia, B., Murray, B., Ghose, K., Iyengar, M., and Schmidt, R., 2010, "Numerical Modeling Approach to Dynamic Data Center Cooling," *Proceedings of IEEE ITherm*, Las Vegas, NV, June 2 - 5.

Incropera, F. and DeWitt, D., 2002, *Fundamentals of Heat and Mass Transfer*, 5th ed.,
New York: Wiley.

Innovative Research Inc., 2011, *TileFlow*, <http://www.inres.com/>.

Intel, 2007, "Data Center Energy Efficiency Research," Research at Intel Day, Mountain
View, CA, June 20.

Iyengar, M. and Schmidt, R., 2007, "Analytical Modeling of Energy Consumption and
Thermal Performance of Data Center Cooling Systems – From the Chip to the
Environment," *Proceedings of ASME InterPACK*, Vancouver, British Columbia,
Canada, July 8 - 12.

Iyengar, M. and Schmidt, R., 2007, "Analytical Modeling for Thermodynamic
Characterization of Data Center Systems," *Journal of Electronic Packaging*, Vol.
131, pp. 021009-1 – 021009-9.

Iyengar, M., Schmidt, R., and Caricari, J., 2010, "Reducing Energy Usage in Data
Centers through Control of Room Air Conditioning Units," *Proceedings of IEEE
ITherm*, Las Vegas, NV, June 2 – 5.

Iyengar, M., Schmidt, R.R., Hamann, H., and VanGilder, J., 2007, “Comparison Between Numerical and Experimental Temperature Distributions in a Small Data Center Test Cell,” *Proceedings of ASME InterPACK*, Vancouver, British Columbia, Canada, July 8 - 12.

Janna, W., 1998, Design of Fluid Thermal Systems, 2nd Edition, New York: CL-Engineering.

Karlsson, J. and Moshfegh, B., 2005, “Investigation of Indoor Climate and Power Usage in a Data Center,” *Energy and Buildings*, Vol. 37, pp. 1075 – 1083.

Khalifa, H.E. and Demetriou, D.W., 2011, “Enclosed Aisle Data Center Cooling System,” U.S. Patent 13/191,509.

Khalifa, H.E., Demetriou, D.W., 2010, “Energy Optimization of Air-Cooled Data Centers,” *Journal of Thermal Sciences and Engineering Applications*, Vol. 2, pp. 041005-1 - 041005-13.

Khalifa, H.E., Elhadidi, B., and Dannenhoffer, J., 2007, “Efficient Coupling of Multizone and CFD Indoor Flow Models through Proper Orthogonal Decomposition,” *ASHRAE Transactions*, Vol. 113(1), pp. 282 – 289.

Khankari, K., 2009, "Rack Enclosures: A Crucial Link in Airflow Management in Data Centers," *ASHRAE Journal*, August, pp. 48 - 51.

Klein, S., 2002, *TRNSYS – A Transient System Simulation Program*, Engineering Experiment Station Report 38-13, Solar Energy Laboratory, University of Wisconsin-Madison, Madison, WI.

Koomey, J., 2011, "Growth in Data Center Electricity Use 2005 to 2010," A Report by Analytical Press, completed at the request of *The New York Times*.

Lawrence Berkeley National Labs, 2007, "Benchmarking: data centers-Charts," <http://hightech.lbl.gov/benchmarking-dc-charts.html>.

Lawrence Berkeley National Labs, 2005, "High Performance Buildings: Data Center's Uninterruptible Power Supplies (UPS)," Berkeley, CA.

Lopez, V. and Hamann, H., 2011, "A Numerical Technique for the Approximation of Thermal Zones," *Proceedings of ASME InterPACK*, Portland, OR, July 6 – 8.

Li, H. and Hamann, H., 2011, "A Statistical Approach to Thermal Zone Mapping," *Proceedings of ASME InterPACK*, Portland, OR, July 6 – 8.

- Lui, Y., 2010, "Waterside and Airside Economizers Design Considerations for Data Center Facilities," *ASHRAE Transactions*, Vol. 116(1), pp. 98 – 108.
- Lumley, J., 1981, *Coherent Structures on Turbulence, in Transition and Turbulence*, R.E. Meyer, ed., New York: Academic Press.
- Ly, H.V. and Tran, H.T., 2001, "Modeling and Control of Physical Processes using Proper Orthogonal Decomposition," *Mathematical and Computer Modeling*, Vol. 33, pp. 223 – 236.
- Malone, C. and Belady, C., 2008, "Optimizing Data Center TCO: Efficiency Metrics and an Infrastructure Cost Model," *ASHRAE Transactions*, Vol. 114(1), pp. 44 – 50.
- Mathworks Inc, 2008, *Simulink*, Version 7.1 (R2008a), Natick, MA.
- Merkel, F., 1925. "Verdunstungskühlung (Evaporative cooling)," *VDI Zeitschrift*, Vol. 70, pp. 123–128.
- Montgomery, D, 2001, Design and Analysis of Experiments, 4th Edition, Wiley: New York.
- Moore, J., 2005, Automated Cost-Aware Data Center Management, PhD Dissertation, Duke University, Durham, NC.

- Moore, J., Chase, J., Ranganathan, P., and Sharma, R., 2005, “Making Scheduling “Cool”: Temperature-Aware Workload Placement in Data Centers,” *Proceedings of USENIX Annual Technical Conference*, Anaheim, CA, April 10 – 15, pp. 61 – 74.
- Moore, J., Chase, J., and Ranganathan, P., 2006a, “ConSil: Low-cost Thermal Mapping of Data Centers,” *Proceedings of SysML*, Saint Malo, France, June 27.
- Moore, J., Chase, J.S., and Ranganathan, P., 2006b, “Weatherman: Automated, Online and Predictive Thermal Mapping and Management for Data Centers,” *Proceedings of IEEE ICAC*, Dublin, Ireland, June 12 – 16, pp. 155 -164.
- Moss, D., 2009a, “Universal Application and Benefits of Air Containment: A Practical Guide,” Dell Technical White Paper.
- Moss, D., 2009b, “The Energy Advantages of Containment Systems,” Dell Technical White Paper.
- Mukherjee, T., Banerjee, A., Varsamopoulos, G., and Gupta, S., 2009, “Spatio-Temporal Thermal-Aware Job Scheduling to Minimize Energy Consumption in Virtualized Heterogeneous Data Centers,” *Computer Networks*, Vol. 53(17), pp. 2888 – 2904.

- Muller, C., 2010, "What's Creeping Around in Your Data Center," *ASHRAE Transactions*, Vol. 116(1), pp. 207 – 222.
- Neimann, J., 2008, "Hot Aisle v Cold Aisle Containment," American Power Conversion White Paper #135.
- Noh, M., Song, K., and Chun, S.K., 1998, "The Cooling Characteristic on the Air Supply and Return Flow System in the Telecommunication Cabinet Room," *Proceedings of IEEE INTELEC*, San Francisco, CA, October 4 – 8, pp. 777 – 784.
- Patel, C., Bash, C., and Belady, C., 2001, "Computational Fluid Dynamics Modeling of High Compute Density Data Centers to Assure System Inlet Air Specifications," *Proceedings of ASME InterPACK*, Kauai, HI, July 8 - 13.
- Patel, C., Sharma, R., Bash, C., and Graupner, S., 2002, "Energy Aware Grid: Global Workload Placement based on Energy Efficiency," *Hewlett-Packard Company Report HPL-2002-329*.
- Patterson, M., Atwood, D., and Miner, J., 2009, "Evaluation of Air-Side Economizer Use in a Compute-Intensive Data Center," *Proceedings of ASME InterPACK*, San Francisco, CA, July 19 – 23.

Pelley, S., Meisner, D., Wenisch, T., and VanGilder, J., 2009, "Understanding and Abstracting Total Data Center Power," *Proceedings of the 2009 Workshop on Energy Efficient Design (WEED)*, University of Michigan, Ann Arbor, MI, June.

Podvin, B. and Lumley, J., 1998, "reconstructing the Flow in the Wall Region from Wall Sensors," *Physics of Fluids*, Vol. 10(5), pp. 1182 – 1190.

Rambo, J. and Joshi, Y., 2007, "Modeling of Data Center Airflow and Heat Transfer: State of the Art and Future Trends," *Distributed Parallel Databases*, Vol. 21(3), pp. 193 – 225.

Rambo, J. and Joshi, Y., 2005, "Reduced Order Modeling of Steady Turbulent Flows using the POD," *Proceedings of ASME Summer Heat Transfer Conference*, San Francisco, CA, July 17 – 22.

Rasmussen, N. and Torell, W., 2007, "Data Center Projects: Establishing a Floor Plan," American Power Conversion White Paper #144.

Rolia, J., Singhal, S., and Friedrich, R., 2000, "Adaptive Internet Data Centers," *Proceedings of International Conference on Advances in Infrastructure for Electronic Business, Science and Education on the Internet*.

- Russo, J., 2010, A Detailed and Systematic Investigation of Personal Ventilation Systems, Ph.D. Dissertation, Syracuse University, Syracuse, NY.
- Salim, M. and Tozer, R., 2010, “Data Centers’ Energy Auditing and Benchmarking-Progress Update,” *ASHRAE Transactions*, Vol. 116(1), pp. 109 - 117.
- Samadiani, E., Joshi, Y., Hamann, H., Iyengar, M., Kamalsy, S., and Lacey, J., 2009, “Reduced Order Thermal Modeling of Data Centers Via Distributed Sensor Data,” *Proceedings of ASME InterPACK*, San Francisco, CA, July 19 - 23.
- Schmidt, R., Karki, K., Kelkar, K., Radmehr, A., and Patankar, S., 2001, “Measurements and Predictions of the Flow Distribution through Perforated Tiles in raised Floor Data Centers,” *Proceedings of ASME InterPACK*, Kauai, HI, July 8 - 13.
- Schmidt, R., 2004, “Thermal Profile of a High-Density Data Center—Methodology to Thermally Characterize a Data Center,” *ASHRAE Transactions*, Vol. 110(2), pp. 635 – 642.
- Schmidt, R., Karki, K., and Pantankar, S., 2004, “Raised-floor Data Center: Perforated Tile Flow Rates for Various Tile Layouts,” *Proceedings of IEEE ITherm*, Las Vegas, NV, June 1 – 4.

- Schmidt, R., Cruz, E., and Iyengar, M., 2005, "Challenges of Data Center Thermal Management," *IBM Journal of Research and Development*, Vol. 49, pp. 709 – 723.
- Schmidt, R., Iyengar, M., Beaty, D., and Shrivastava, S., 2005, "Thermal Profile of a High Density Data Center – Hot Spot Heat Fluxes of 512 Watts/ft²," *ASHRAE Transactions*, Vol. 111(2), pp. 765 – 777.
- Schmidt, R. and Iyengar, M., 2006, "Best Practices for Data Center Thermal and Energy Management – Review of Literature," *ASHRAE Transactions*, Vol. 113(1), pp. 206 – 218.
- Schmidt, R., Vallury, A., and Iyengar, M., 2011, "Energy Savings through Hot and Cold Aisle Containment Configurations for Air Cooled Server in Data Centers," *Proceedings of InterPACK*, Portland, OR, July 6 – 8.
- Shah, A., Bash, C., Kumari, N., Cader, T., Breen, T.J., Walsh, E., and Punch, J., 2011, "One the Need for Energy Efficient Metrics that Span Integrated IT-Facilities Infrastructures," *Proceedings of ASME InterPACK*, Portland, OR, July 6 – 8.
- Sharma, R., Bash, C., and Patel, C., 2002, "Dimensionless Parameters for Evaluation of Thermal Design and Performance of Large-Scale Data Centers," *Proceedings IEEE ITherm*, San Diego, CA, May 29 – June 1.

- Sharma, R., Bash, C., Patel, C., Friedrich, R., and Chase, J., 2005, "Balance of Power: Dynamic Thermal Management for Internet Data Centers," *Internet Computing*, Vol. 9(1), pp. 42 – 49.
- Shih, T.H., Liou, W.W., Shabbir, A., Yang, Z., and Zhu, J., 1995. "A New k- ϵ Eddy Viscosity Model for High Reynolds Number Turbulent Flows," *Computer Fluids*, Vol. 24, pp. 227-238.
- Shrivastava, S., Iyengar, M., Sammakia, B., Schmidt, R., and VanGilder, J., 2006, "Experimental Numerical Comparison for a High Density Data Center: Hot Spot Heat Fluxes in Excess of 500 W/ft²," *Proceedings of IEEE ITherm*, San Diego, CA, May 30 – June 2.
- Shrivastava, S., VanGilder, J., and Sammakia, B., 2006, "A Statistical Prediction of Cold Aisle End Airflow Boundary Conditions," *Proceedings of IEEE ITherm*, San Diego, CA, May 30 – June 2.
- Shrivastava, S., VanGilder, J., and Sammakia, B., 2007, "Prediction of Cold Aisle End Airflow Boundary Conditions Using Regression Modeling," *IEEE Transactions on Components and Packaging Technologies*, Vol. 30(4), pp. 866 – 874.
- Shrivastava, S., 2008, Cooling Analysis of Data Centers: CFD Modeling and Real-Time Calculators, PhD Dissertation, Binghamton University, Binghamton, NY.

Simpson, W.M. and Sherwood, T.K., 1946, "Performance of Small Mechanical Draft Cooling Towers," *Refrigerating Engineering*, Vol. 52, pp. 525–543.

Sirovich, L., 1987, "Turbulence and the Dynamics of Coherent Structures, Part 1: Coherent Structures," *Quarterly of Applied Mathematics*, Vol. XLV(3), pp. 561 – 571.

Sirovich, L. and Everson, R., 1992, "Management and Analysis of Large Scientific Datasets," *Journal of Supercomputer Applications*," Vol. 6(1), pp. 50 -68.

Sisk, D., Khalell, M., Marquez, A., Hatley, D., Cader, T., and Schmidt, R., 2009, "Real-Time Data Center Energy Efficiency at Pacific Northwest National Laboratory," *ASHRAE Transactions*, Vol. 115(1), pp. 242 – 253.

Song, Z., Murray, B., and Sammakia, B., 2011, "Multivariate Prediction of Airflow and Temperature Distributions using Artificial Neural Networks," *Proceedings of ASME InterPACK*, Portland, OR, July 6 - 8.

Sorell, V., 2007, "OA Economizers for Data Centers," *ASHRAE Journal*, December, pp. 32 - 37.

Stahl, L. and Belady, C., 2001, "Designing and Alternative to Conventional Room Cooling," *Proceeding of INTELEC*, Edinburgh, Scotland, October 14 – 18.

- Stein, J., 2009, "Waterside Economizing in Data Centers: Design and Control Considerations," *ASHRAE Transactions*, Vol. 115(2), pp. 192 – 200.
- Tang, Q., Mukherjee, T., Gupta, S.K.S., and Crayton, P., 2006, "Sensor-Based Fast Thermal Evaluation Model for Energy Efficient High-Performance Datacenters," *Proceedings of IEEE Intelligent Sensing & Information*, Bangalore, India, December 15 – 18, pp. 203-208.
- Tang, Q., Gupta, S.K.S., Stanzione, D., and Crayton, P., 2006, "Thermal-Aware Task Scheduling to Minimize Energy Usage of Blade Server Based Data Centers," *Proceedings of IEEE DASC*, Indianapolis, IN, September 29 – Oct 1.
- Tang, Q., Gupta, S.K.S., and Varsamopoulou, G., 2007, "Thermal-Aware Task Scheduling for Data Centers through Minimizing Heat Recirculation," *Proceedings of IEEE Cluster*, Austin, TX, September 17 - 21.
- Tang, Q., Gupta, S., and Varsamopoulos, G., 2008, "Energy-Efficient Thermal Aware Task Scheduling for Homogenous High-Performance Computing Data Centers: A Cyber-Physical Approach," *IEEE Transactions on Parallel and Distributed Systems*, Vol. 19(11), pp. 1458 – 1472.

- Tang, Q., 2009, Thermally-Aware Scheduling in Environmentally Coupled Cyber-Physical Distributed Systems, PhD Dissertation, Arizona State University, Tempe, AZ.
- Tozer, R., Kurkjian, C., and Salim, M., 2009, “Air Management Metrics in Data Centers,” *ASHRAE Transactions*, Vol. 115(1), pp. 63 - 70.
- Tschudi, W., 2007, “Data Center Economizer Contamination and Humidity Study,” LBNL/PG&E Report,
http://hightech.lbl.gov/documents/DATA_CENTER/EconomizerDemoReportMarch13.pdf
- United States Department of Energy, 2009, 2008 Buildings Energy Data Book, Prepared for the Buildings Technologies Program, Energy Efficiency and Renewable Energy by D&R International Ltd.
- United States Energy Information Administration, 2010, *Electricity Statistics and Analysis*, <http://www.eia.doe.gov/fuelelectric.html>
- United States Environmental Protection Agency, 2007, “Report to Congress on Data Center Energy Efficiency,” Public Law 109-431.
- The Uptime Institute, 2011, *The Uptime Institute*, <http://www.uptimeinstitute.org>.

The Uptime Institute, 2009, "Data Center Site Infrastructure Tier Standard: Topology,"

Prepared by the Uptime Institute.

VanGilder, J. and Schmidt, R., 2005, "Airflow Uniformity through Perforated Tiles in a Raised-Floor Data Center," *Proceedings of ASME InterPACK*, San Francisco, CA, July 17 - 22.

VanGilder, J. and Shrivastava, S., 2006, "Real-Time Prediction of Rack-Cooling Performance," *ASHRAE Transactions*, Vol. 112(2), pp. 151 - 162.

VanGilder, J. and Shrivastava, S., 2007, "Capture Index: An Airflow-Based Rack Cooling Performance Index," *ASHRAE Transactions*, Vol. 113(1), pp. 126 - 136.

VanGilder, J., Zhang, X., and Shrivastava, S., 2007, "Partially Decoupled Aisle Method for Estimating Rack-Cooling Performance in Near-Real Time," *Proceedings of ASME InterPACK*, Vancouver, British Columbia, Canada, July 8 – 12.

VanGilder, J. and Zhang, X., 2008, "Coarse-Grid CFD: The Effect of Grid Size on Data Center Modeling," *ASHRAE Transactions*, Vol. 114 (2), pp. 166 – 181.

Villa, H., 2010, "High Density Cooling Solutions – Taking IT to the Next Level: The Cold Aisle Containment Alternative," *ASHRAE Transactions*, Vol. 116(1), pp. 32 – 36.

Walsh, E., Breen, T.J., Punch, J., Shah, A., and Bash, C., 2010, “From Chip to Cooling Tower Data Center Modeling: Part II Influence of Chip Temperature Control Philosophy,” *Proceedings of IEEE ITherm*, Las Vegas, NV, June 2 – 5.

Wilman, F., 2007, “Fresh Air Cooling in Data Centers – Overview of approach showing potential savings in operational costs,” Future-Tech White Paper.

Yarlanki, S., Das, R., Hamann, H., Lopez, V., and Stepanchuk, A., 2011, “Employing Thermal Zones for Energy Optimization in Data Centers,” *Proceedings of ASME InterPACK*, Portland, OR, July 6 – 8.

Zhang, X., VanGilder, J., Iyengar, M., and Schmidt, R.R., 2008, “Effect of Rack Modeling on the Numerical Results of a Data center Test Cell,” *Proceedings of IEEE ITherm*, Orlando, FL, May 28 – 31.

Dustin W. Demetriou

Curriculum Vitae

Mechanical and Aerospace Engineering
Syracuse University
263 Link Hall
Syracuse, NY 13244

203 S. Beech St.
Syracuse, NY 13210
845 -313-5626 (cell)
dustin.demetriou@gmail.com
DOB: October 28, 1984

Research Interests

Energy conversion systems, HVAC&R, energy efficient environmental control systems for buildings and high-density data centers, computational fluid dynamics

Education

Doctor of Philosophy, Mechanical and Aerospace Engineering
Syracuse University, Syracuse, NY
Thesis: Thermally aware, energy-based techniques for improving data center energy efficiency *Feb. 2012*
Advisor: H. Ezzat Khalifa

Masters of Science, Mechanical and Aerospace Engineering
Syracuse University, Syracuse, NY
Capstone: Distributed demand controlled ventilation for improving IAQ and reducing energy consumption in office buildings *May 2008*
Advisor: H. Ezzat Khalifa

Bachelor of Science, Mechanical Engineering, (Magna cum laude) *May 2006*
Manhattan College, Riverdale, NY

Licenses and Certifications

- Completed all requirements and passed the Fundamentals of Engineering (FE) exam in New York State

Professional Experience

Syracuse University

Ph.D. Candidate, Mechanical and Aerospace Engineering *2008 - 2012*
Teaching Assistant, MAE 251: Thermodynamics *Spring 2012*
Teaching Assistant, MAE 551: Energy Conversion *Spring 2011*
Teaching Assistant, MAE 251: Thermodynamics *Spring 2011*
Research Assistant, Mechanical and Aerospace Engineering *2008 -*
Research Assistant, Mechanical and Aerospace Engineering *2006 -2008*

New York Power Authority

Developmental Intern, Energy Services and Technologies Group *Summer 2006*

Research Experience

Doctoral Research (2008 – Present)

Mechanical and Aerospace Engineering, Syracuse University,

Advisor: H. Ezzat Khalifa

- Developing thermally aware, energy-based virtualization strategies for IT data centers, which can improve energy efficiency, while maintaining IT equipment reliability.
- Experimentally investigating transient thermal behavior of data center equipment.

Masters Research (2006 – 2008)

Mechanical and Aerospace Engineering, Syracuse University

Advisor: H. Ezzat Khalifa

- Developed and experimentally validated a software tool that gives operators, designers and researchers the ability to evaluate the dynamic, off-design energy use of a data center's cooling infrastructure.
- Evaluated energy savings technologies in data centers, including, water economizers, air economizers, evaporative cooling and tri-generation systems.

Mechanical and Aerospace Engineering, Syracuse University,

Advisor: H. Ezzat Khalifa

- Developed a coupled airflow, contaminant transport and building energy consumption simulation tool.
- Studied distributed demand control ventilation systems, which can reduce energy consumption and improve the indoor air quality of all individuals in office buildings.

Industrial Research (2009-Present)

Analysis and Design Center, Syracuse University/IBM

Supervisors: H. Ezzat Khalifa (SU) and Roger Schmidt (IBM)

- Developed models and performed energy assessments of new energy efficient data center designs and configurations.

Industrial Research (2006)

Energy Services and Technologies Group, New York Power Authority

Supervisor: Rino Trovato

- Designed a solar thermal, process hot water system, including the selection of solar collectors piping and control systems. Performed site visits, site audits, equipment verification, monitoring and data collection.
- Prepared drawings, design documents, specifications and grant applications.

Leadership Experience

Research Mentor (Summer 2009, Summer 2010)

Mechanical and Aerospace Engineering, Syracuse University

- Mentored visiting undergraduate research students on various projects related to data centers

Vice President (2005 – 2006)

Pi Tau Sigma, Manhattan College

- Developed and organized an undergraduate tutoring program for Mechanical Engineering students.

Undergraduate Tutor (2004 – 2006)

Mechanical Engineering, Manhattan College

- Provided tutoring to undergraduate students in subjects related to Mechanical Engineering, including Heat Transfer, Thermodynamics, Calculus and Physics.

Honors and Awards

- Syracuse Center of Excellence Annual Symposium Poster Competition Winner *2011*
- ASME - Best Paper in Data Centers and Energy Efficient Electronic Systems, InterPACK *2011*
- Nunan Lecture and Research Day Poster Winner – Applicability of Research to Business and Industry *2010*
- Nunan Lecture and Research Day Poster Competition Winner – Mechanical and Aerospace Engineering *2010*
- Building Simulation: An International Journal – Best Paper Award *2009*
- Syracuse University Graduate Student Commencement Speaker *2008*
- Nunan Lecture and Research Day Poster Competition Winner – Mechanical and Aerospace Engineering *2008*
- The Brother Aubert Medal for Mechanical Engineering, Manhattan College *2006*
- Tau Beta Pi, Manhattan College *2004*
- Pi Tau Sigma, Manhattan College *2004*
- Epsilon Sigma Pi, Manhattan College *2006*

Technical Skills

- Computer programming: Matlab, Fortran, C++, EES
- Thermodynamic modeling: TRNSYS, Simhydraulic, Fathom
- Computational fluid dynamics: Fluent, Gambit, Pointwise
- Experimental fluid dynamics: PIV, Hot wire anemometry, Tracer gas

Professional Affiliations

- American Society of Heating, Refrigeration and Air Conditioning Engineers *2007 -*
- American Society of Mechanical Engineers *2006 -*
- International Society of Indoor Air Quality and Climate *2008 -*

Professional Activities

- Paper Peer-Reviewer – ASME InterPACK *2011*
- Paper Peer-Reviewer – IAQVEC Conference *2010*

Journal Publications

1. **Demetriou, D.W.**, Khalifa, H.E. (2012). "Thermally Aware, Energy-based Load Placement: Part II – Systematic Investigation of Practical Load Placement in Data Centers." *To be submitted to the Journal of Electronic Packaging*.
2. **Demetriou, D.W.**, Khalifa, H.E. (2012) "Thermally Aware, Energy-based Load Placement: Part I – Development of a Reduced Order Model using POD." *To be submitted to the Journal of Electronic Packaging*.
3. **Demetriou, D.W.**, Khalifa, H.E. (2012). "Optimization of Enclosed Aisle Data Centers Using Bypass Recirculation," *Accepted for publication in the Journal of Electronic Packaging*.
4. **Demetriou, D.W.**, Khalifa, H.E., Iyengar, M., Schmidt, R. (2011). "Development and Experimental Validation of a Thermo-Hydraulic Model for Data Centers," *HVAC&R Research*, Vol. 17 (4), pp. 540 - 555.
5. **Demetriou, D.W.**, Khalifa, H.E., Iyengar, M., Schmidt, R. (2011). "Development of a Dynamic Energy Use Modeling Tool for Data Center Cooling Infrastructure," *ASHRAE Transactions*, Vol. 117 (2), pp. 3 - 9.
6. Dhummi, V., **Demetriou, D.W.**, Palanhandalam-Madapusi, H.J., Khalifa, H.E., Isik, C. (2011). "Robust Occupancy-Based Distributed Demand Control Ventilation," *International Journal of Ventilation*, Vol. 9 (4), pp. 359 - 369.
7. Khalifa, H.E., **Demetriou, D.W.** (2010). "Energy Optimization of Air-Cooled Data Centers," *Journal of Thermal Sciences and Engineering Applications*, Vol. 2, pp. 041005-1 - 041005-13.
8. **Demetriou, D.W.**, Khalifa, H.E. (2009). "Evaluation of Distributed Environmental Control Systems for Improving IAQ and Reducing Energy Consumption in Office Buildings," *Building Simulation: An International Journal*, Vol. 2 (3), pp. 197 - 214.

Patents

1. **Demetriou, D.W.**, Khalifa, H.E. (2010). "Enclosed Aisle Data Center Cooling System," U.S. Patent # 13/191509

Conference Publications (Peer-Reviewed)

1. **Demetriou, D.W.**, Khalifa, H.E. (2012). "Effect of Virtualization on the Cooling Infrastructure's Power Consumption in Enclosed Aisle Data Centers," *Abstract accepted for ITherm 2012*, San Diego, CA.
2. **Demetriou, D.W.**, Khalifa, H.E. (2011). "Energy Modeling of Air-Cooled Data Centers: Part I The Optimization of Enclosed Aisle Configurations," In *Proceedings of ASME InterPACK*, Portland, Oregon, July 6 - 8.
3. **Demetriou, D.W.**, Khalifa, H.E. (2011). "Energy Modeling of Air-Cooled Data Centers: Part II The Effect of Recirculation on the Energy Optimization of Open-Aisle, Air-Cooled Data Centers," In *Proceedings of ASME InterPACK*, Portland, Oregon, July 6 - 8.
4. **Demetriou, D.W.**, Khalifa, H.E. (2011). "Evaluation of a Data Center Recirculation Non-Uniformity Metric Using Computational Fluid Dynamics," In *Proceedings of ASME InterPACK*, Portland, Oregon, July 6 - 8.
5. Palanhandalam-Madapusi, H.J., Dhummi, V., **Demetriou, D.W.**, Khalifa, H.E. (2010). "Occupancy-Based Distributed Demand Control Ventilation Strategies," In *Proceedings of 7th International Conference on Indoor Air Quality, Ventilation and Energy Conservation in Buildings*, IAQVEC, Syracuse, NY, August 15 - 18.

6. Dhummi, V., **Demetriou, D.W.**, Palanhandalam-Madapusi, H.J., Khalifa, H.E. (2010). "A Camera-Based Occupancy Detection System for Distributed Demand-Controlled Ventilation," In Proceedings of 10th International World Congress, CLIMA, Antalya, Turkey, May 9 - 12.
7. **Demetriou, D.W.**, Khalifa, H.E. (2009). "Comparison of Distributed Demand Controlled Ventilation in Various United States Climates," In Proceedings of the 9th International Conference and Exhibition: Healthy Buildings, Syracuse, NY, September 13 - 19.
8. **Demetriou, D.W.**, Khalifa, H.E. (2009). "A Fast Executing Model for Whole Building Simulations of Distributed Ventilation Systems," In Proceedings of the 11th International ROOVENT, Busan, South Korea, May 24 - 27.
9. Sivakumar, S., **Demetriou, D.W.**, Palanhandalam-Madapusi, H.J., Dang, T.Q., Khalifa, H.E. (2009). "Improved Indoor Air Quality Through Intelligent Assisted Natural Ventilation," In Proceedings of the 11th International ROOVENT, Busan, South Korea, May 24 - 27.
10. **Demetriou, D.W.**, Ozdemir, O., Khalifa, H.E., Isik, C. (2008). "Distributed Demand Controlled Ventilation for Improving Indoor Air Quality," In Proceedings of the 11th International Conference of Indoor Air Quality and Climate, Copenhagen, Denmark, August 17 -22.

Presentations, Abstracts and Posters

1. **Demetriou, D.W.**, Khalifa, H.E. (2011). "Energy Optimization of Air-Cooled Data Centers," 11th Annual Symposium on Environmental and Energy Systems, Syracuse Center of Excellence, Syracuse, NY.
2. **Demetriou, D.W.**, Khalifa, H.E. (2010). "Optimized Air Cooled Data Centers Using CRAC Bypass," The 1000 Islands Energy Research Forum (TIERF), Alexandria Bay, NY.
3. **Demetriou, D.W.**, Khalifa, H.E. (2010). "Dynamic Energy Simulations of Data Center Cooling Infrastructure," Nunan Lecture and Research Day, Syracuse University, Syracuse, NY.
4. **Demetriou, D.W.**, Gandel, P., Khalifa, H.E. (2010). "Syracuse University's Green Data Center," iDa Singapore Green Data Center Conference. (**Invited**)
5. **Demetriou, D.W.**, Khalifa, H.E. (2009). "Green Data Centers," The 1000 Islands Energy Research Forum (TIERF), Alexandria Bay, NY. (**Invited**)
6. **Demetriou, D.W.** (2009). "Thermo-Hydraulic Modeling of Data Center Cooling Infrastructure," 3rd IBM Conference on Energy Management in Buildings and Data Centers, Yorktown Heights, NY. (**Invited**)
7. **Demetriou, D.W.** (2009). "Distributed Demand Controlled Ventilation," Mechanical and Aerospace Engineering and EQSTAR Seminar Series, Syracuse University, Syracuse, NY. (**Invited**)
8. **Demetriou, D.W.**, Khalifa, H.E. (2009). "Thermodynamics Modeling of Energy Efficiency in High Density Data Centers," Nunan Lecture and Research Day, Syracuse University, Syracuse, NY.
9. **Demetriou, D.W.**, Khalifa, H.E., Isik, C. (2008). "Distributed Demand Controlled Ventilation for Improving Indoor Air Quality," Nunan Lecture and Research Day, Syracuse University, Syracuse, NY.
10. **Demetriou, D.W.**, Ozdemir, O., Khalifa, H.E., Isik, C. (2007). "Distributed Demand Controlled Ventilation for Improving IAQ," Syracuse COE Symposium, Syracuse, NY.
11. Ozdemir, O., **Demetriou, D.W.**, Khalifa, H.E., Isik, C. (2006). "Demand Controlled Ventilation at the Micro-level. Syracuse COE Symposium, Syracuse, NY.