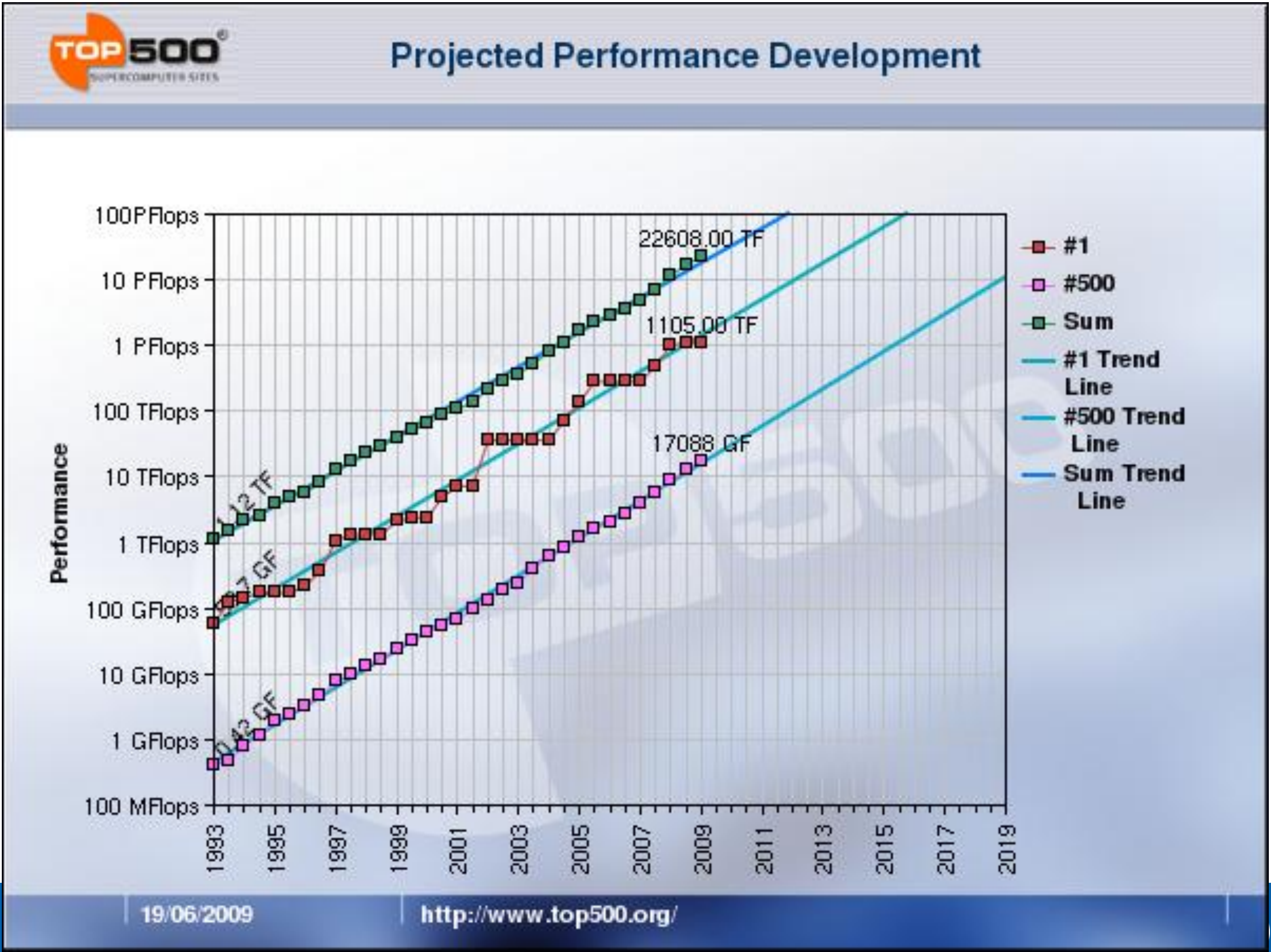
Data Center design and efficiency; challenges today and in the future

Michael K Patterson, PhD, PE
Senior Power & Thermal Architect
Eco-Technology Program Office
Intel Architecture Group

April, 2010

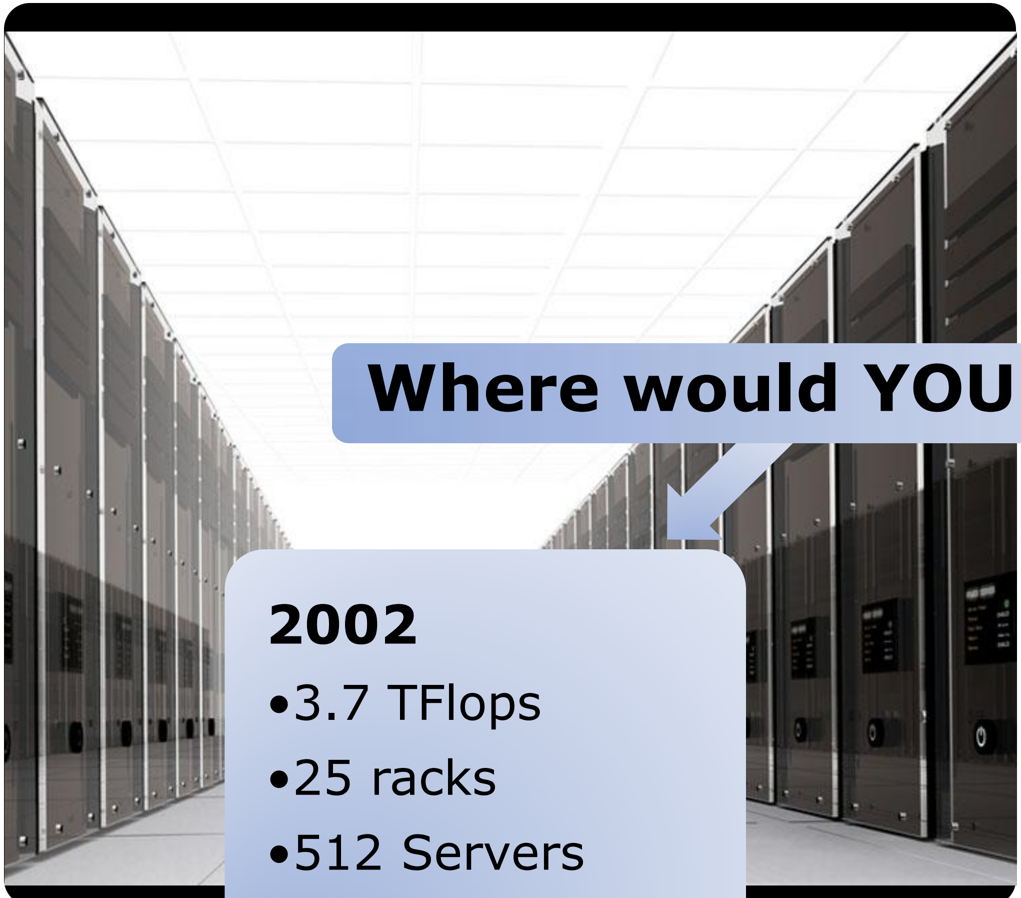
Legal Disclaimer



http://www.top500.org/lists/2009/06/performance_development

Moore's Law …

Alive and Well in the Data Center



Where would YOU rather be?

2002

- 3.7 TFlops
- 25 racks
- 512 Servers
- 90 m2
- 128 kW

2007

- 3.7 TFlops
- 1 rack
- 53 Servers
- 4 m2
- 21 kW

But what does this mean to our HPC challenges?

As we have commoditized the server, we have un-commoditized the data center.

*This is NOT a "crisis", instead it's an **opportunity** to differentiate.*

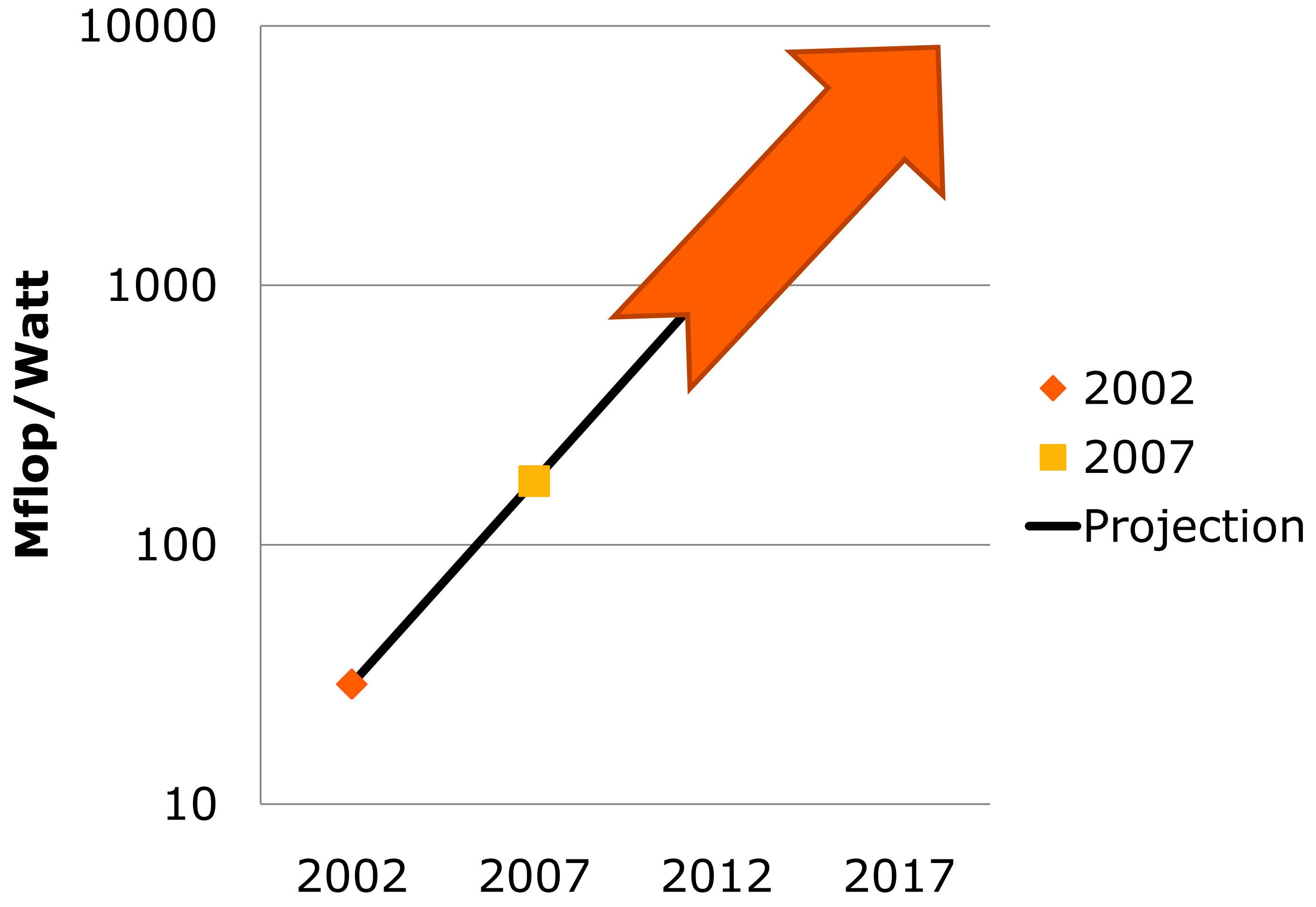
Those who apply the right planning and engineering will prosper, those who do not will suffer.

Good engineering will allow higher density, lower TCO, and more space for more compute

2010

- 6 TFlops
- 1 rack
- 64 Servers
- 4 m^2
- 30 kW

Extrapolations can be fun, *but dangerous!*



An Exascale machine will take 120-150 MW unless we do something different, Intel currently working this challenge

A back of the envelope calculation.. (in other words not an Intel projection…but more a thought experiment)

$$2010 \Rightarrow \frac{1 \text{ petaflop}}{521 \text{ kW}}$$

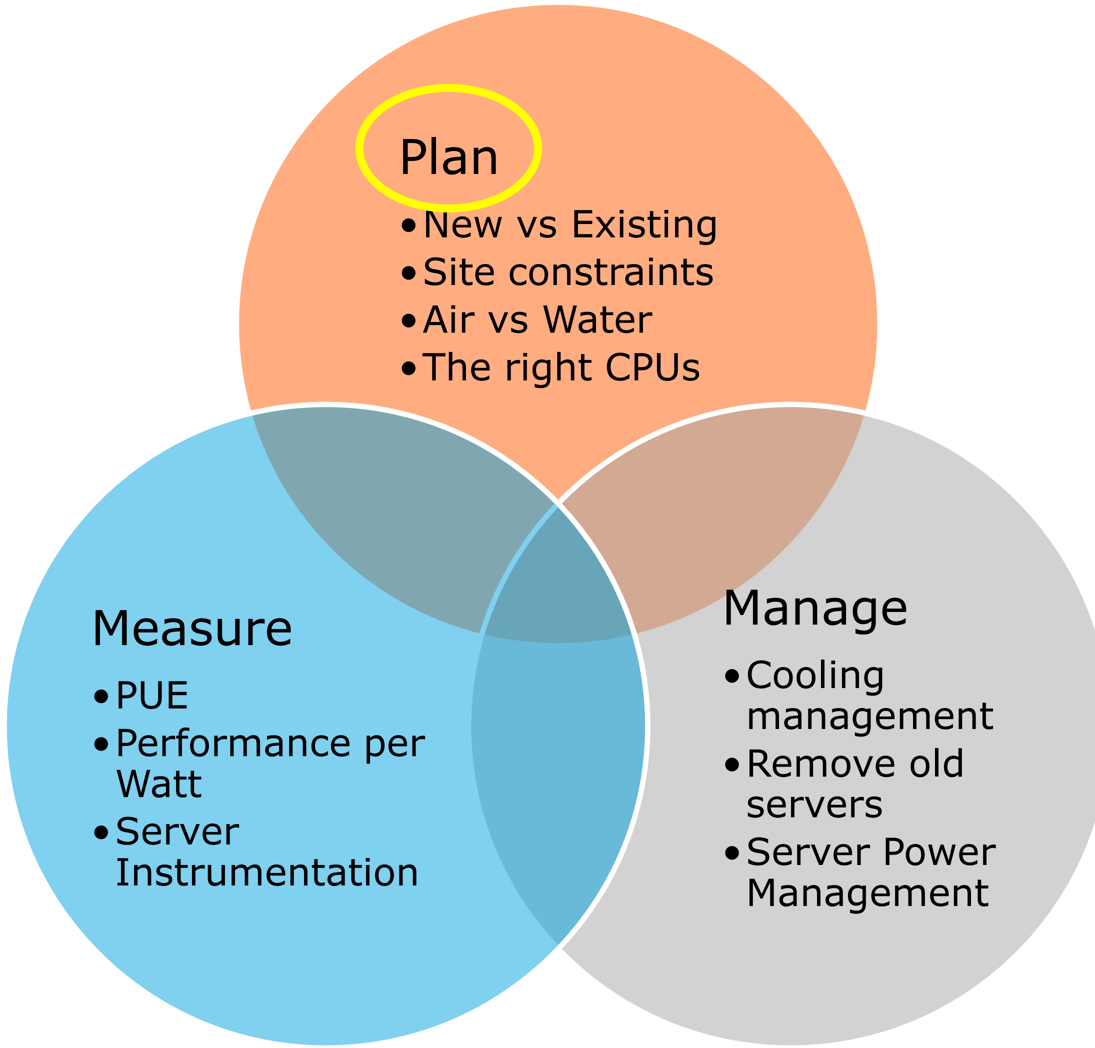
$$2008 \Rightarrow \frac{3.7 \text{ teraflops}}{21 \text{ kW}}$$

$$2002 \Rightarrow \frac{3.7 \text{ teraflops}}{128 \text{ kW}}$$

So at 0.10 CHF per kW-hr and a PUE of 2.0 the system will cost 900,000 CHF per year to run.

Each 0.1 reduction in PUE saves 45,000 CHF every year!

How to succeed going forward?



Technology to meet the HPC Challenges

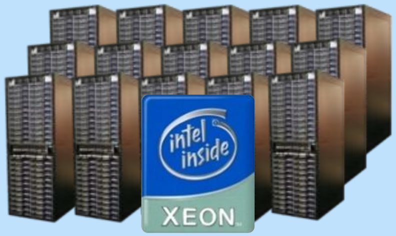
Westmere EP	Nehalem EX

- Cost Sensitive
- CPU Bound
- Memory Bandwidth
- Memory Capacity

EP Server Refresh: 2010

Increasing your performance while decreasing your Energy cost

2005



15 racks of Intel® Xeon® Single Core Servers

2010

Efficiency Refresh 1:1

15 racks of Intel® Xeon® 5600 Based Servers

Up to 15x Performance

8% Annual Energy Costs Estimated Reduction

2006

5 racks of Intel® Xeon® Dual Core Servers

2010

Performance Refresh 1:1

5 racks of Intel® Xeon® 5600 Based Servers

Up to 5x Performance

10% Annual Energy Costs Estimated Reduction

Source: Intel estimates as of Jan 2010. Performance comparison using SPECjbb2005 bops (business operations per second). Results have been estimated based on internal Intel analysis and are provided for informational purposes only. Any difference in system hardware or software design or configuration may affect actual performance. For detailed calculations, configurations and assumptions refer to the legal information slide in backup.

Server Refresh Benefits

(Single Core)

2005

100
Intel® Xeon®
Single Core
4-socket
Servers

2010

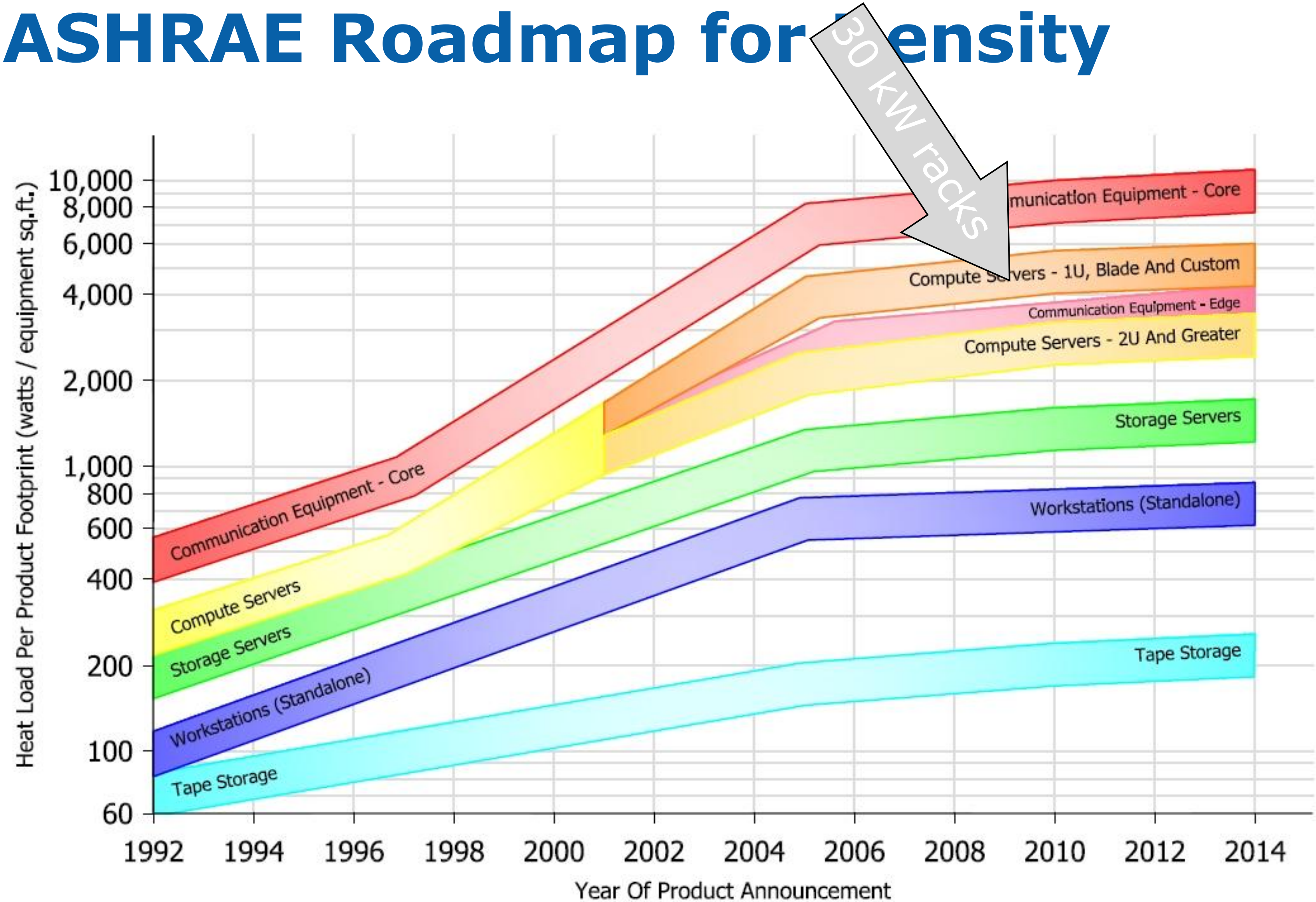
Performance
Refresh
1:1

100
Intel® Xeon®
7500 Based
Servers

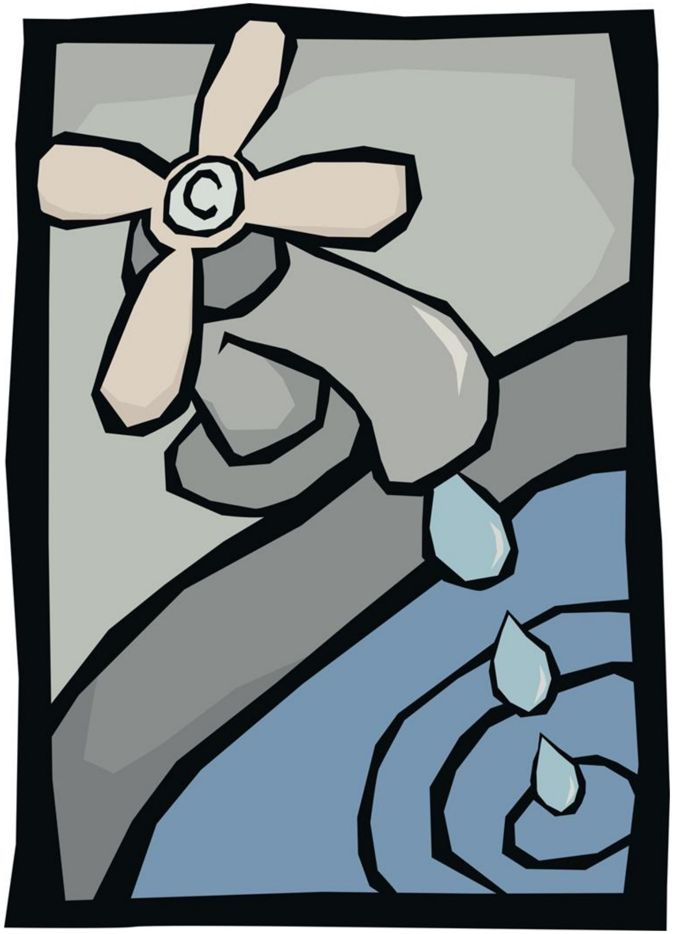
Estimated
Up to
20x
Performance

Source: Intel estimates as of February 2010. Performance comparison using internal workload. Results have been estimated based on internal Intel analysis and are provided for informational purposes only. Any difference in system hardware or software design or configuration may affect actual performance. For detailed calculations, configurations and assumptions refer to the legal information slide in backup.

ASHRAE Roadmap for Density

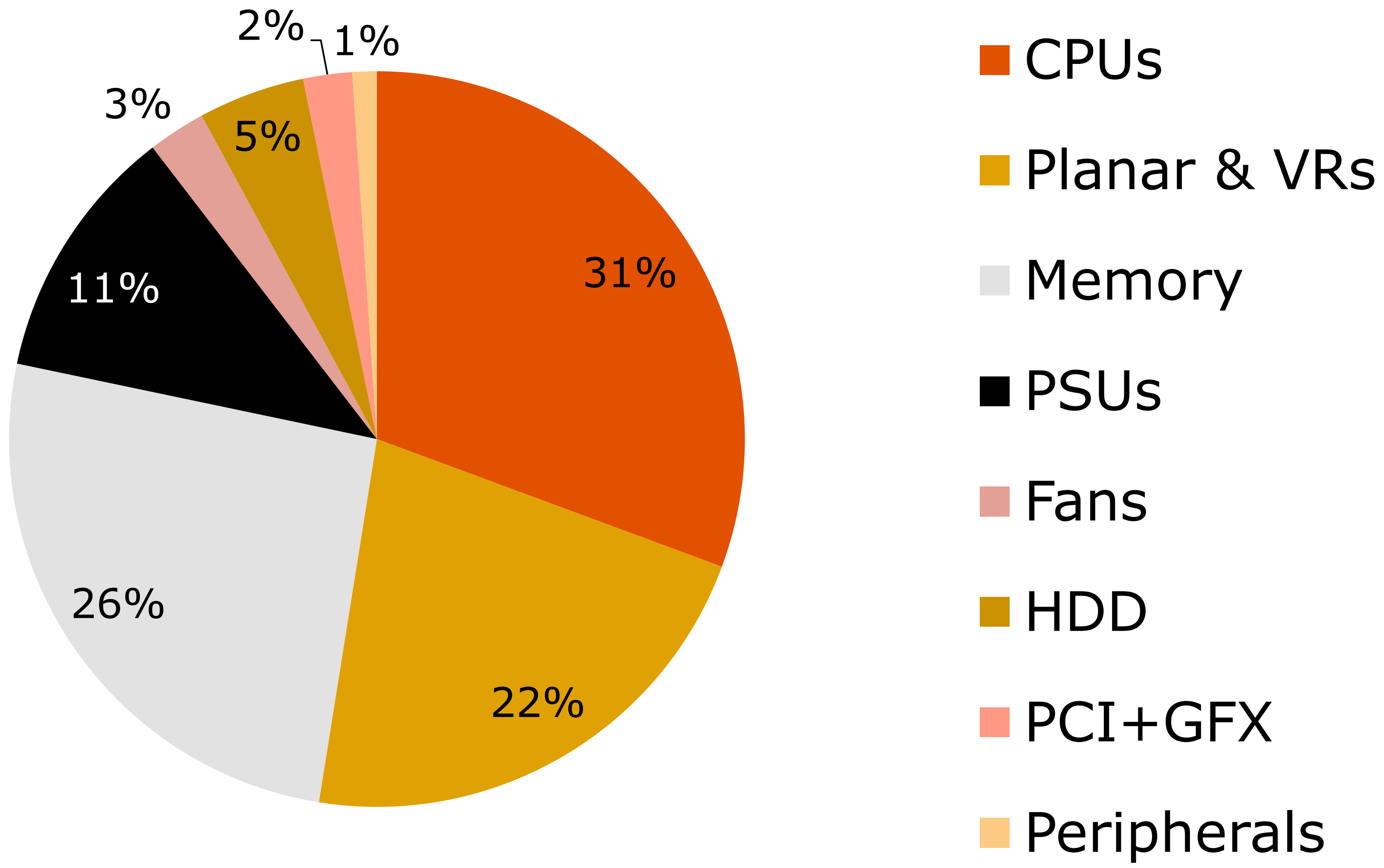


But we are heading for liquid cooling, aren't we?

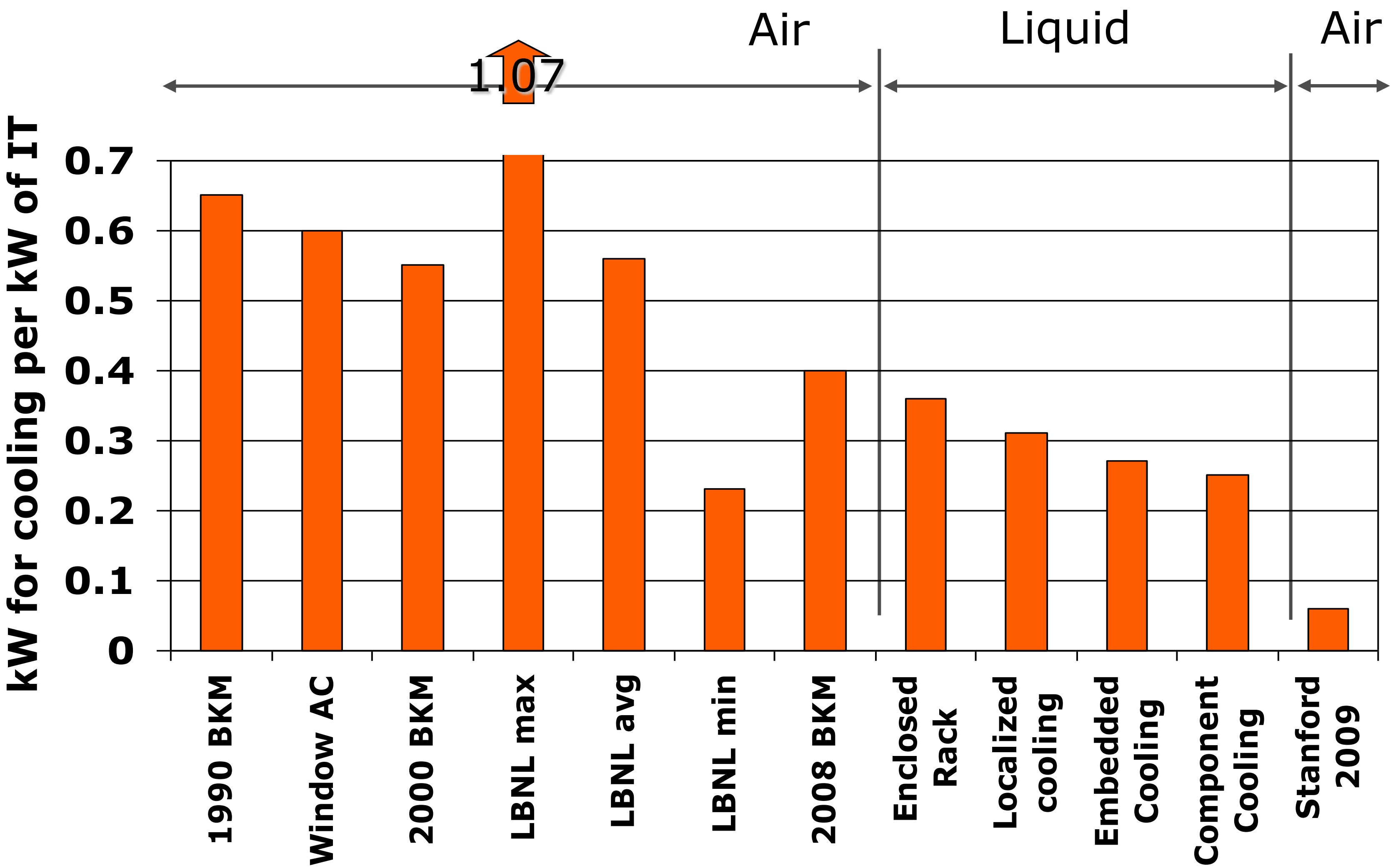


- Only maybe…
- Consider what we are liquid-cooling? Rack, server, component?
- Consider TCO
  - Different answer for different sites

Typical Platform Power Distribution-HPC

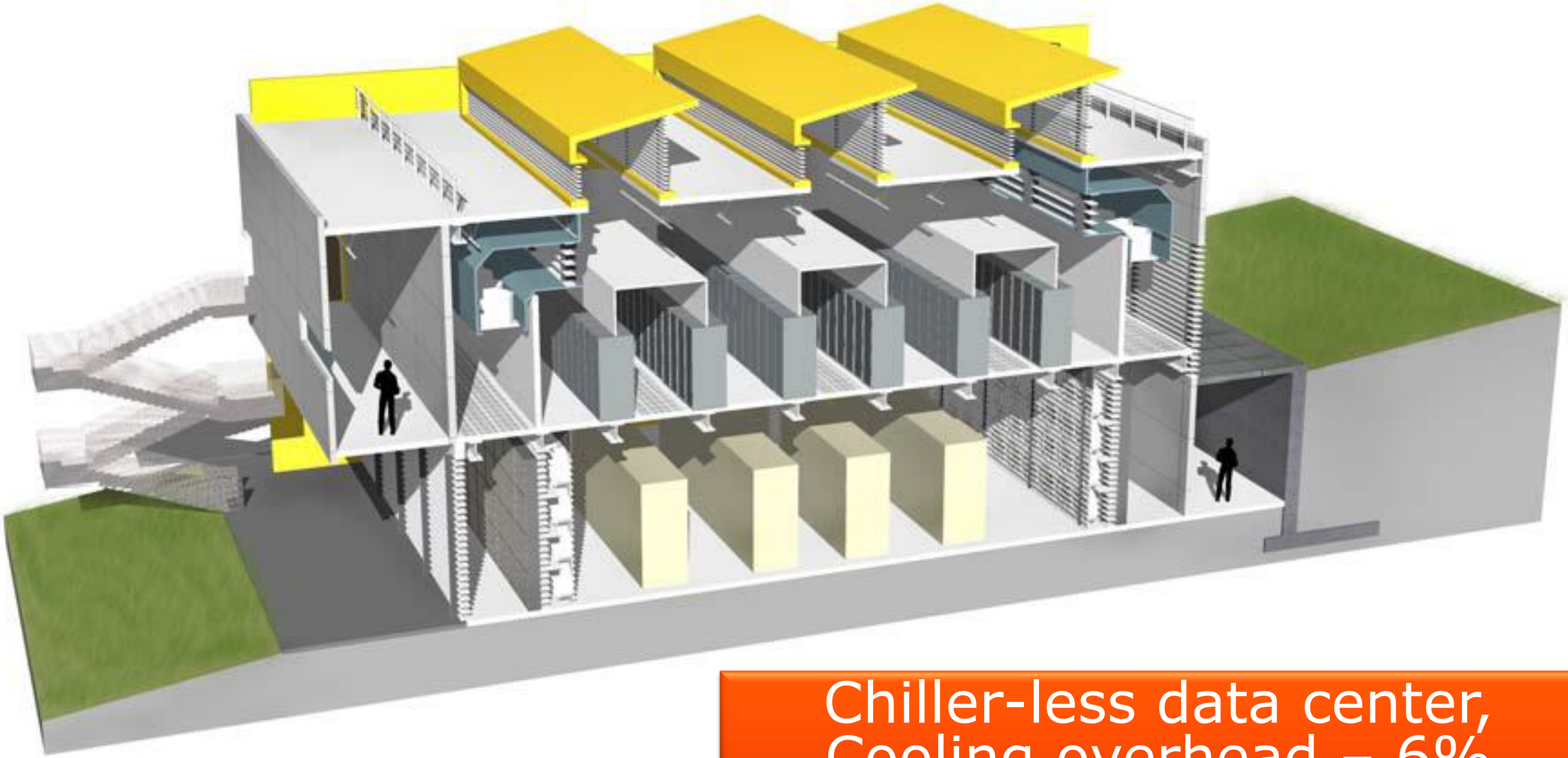


Cooling overhead by technology



Source: compilation of data from LBNL, Liebert, IDC, Sears, Intel

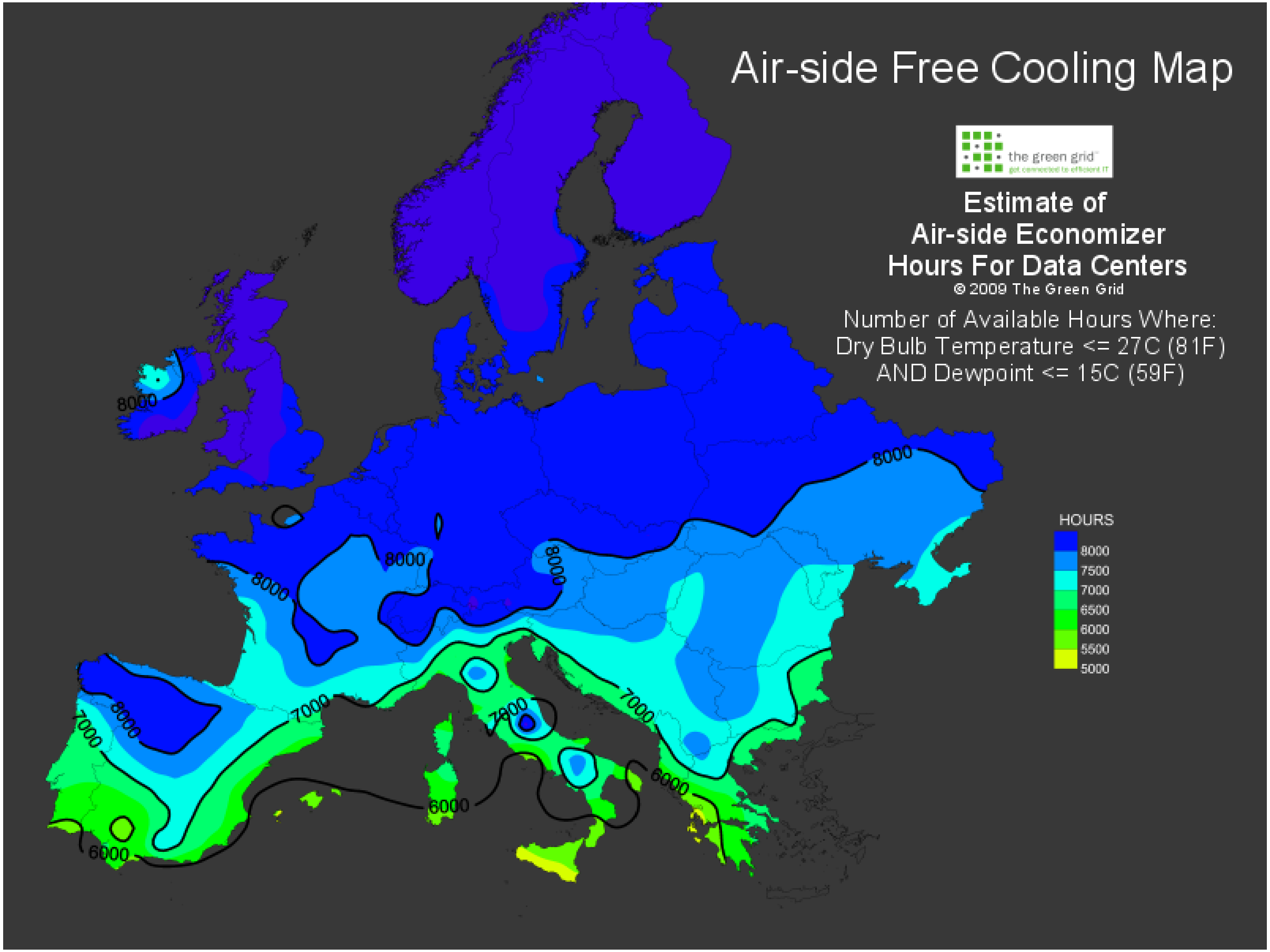
Scientific Research Computing Facility
Stanford University

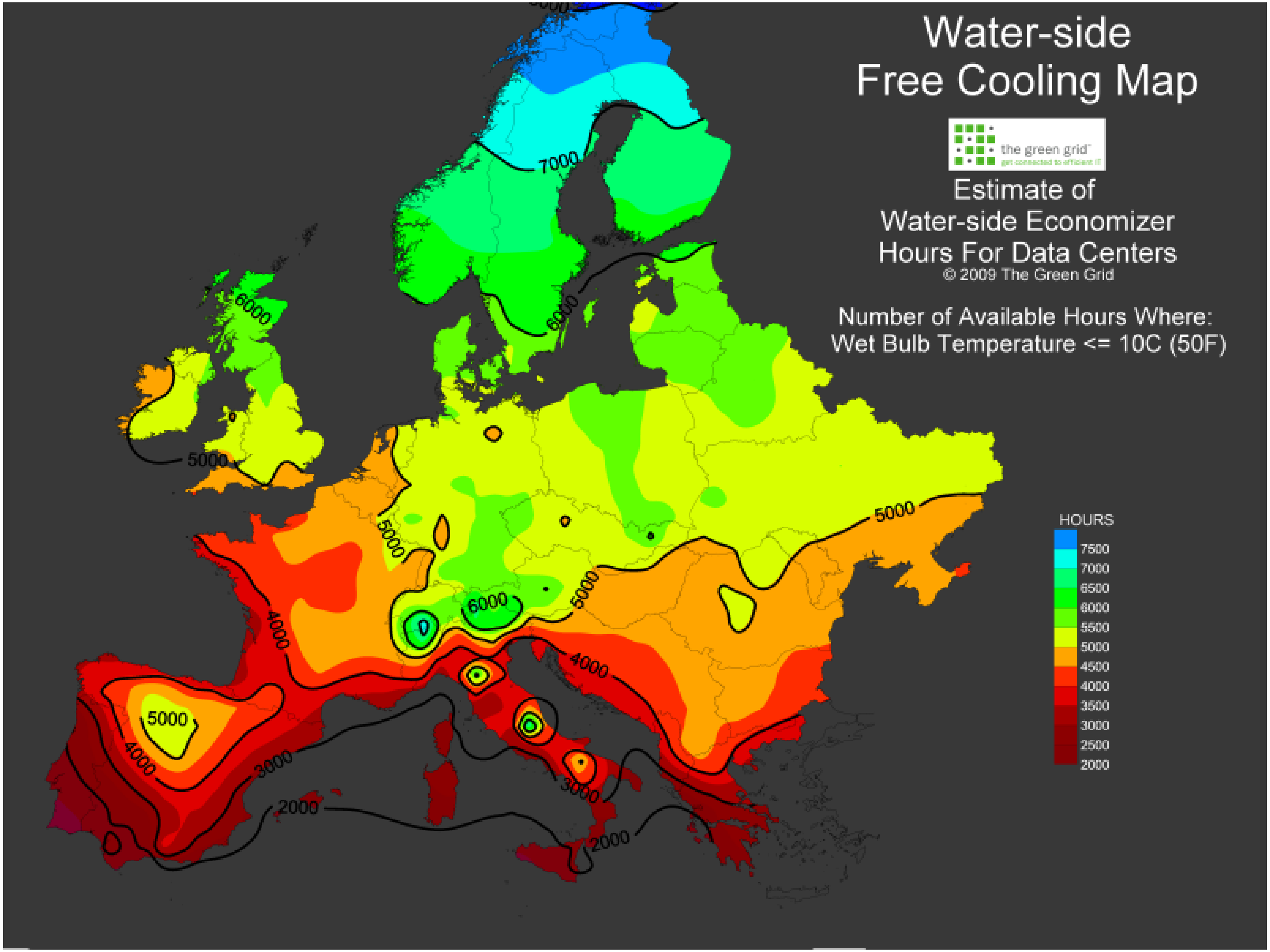


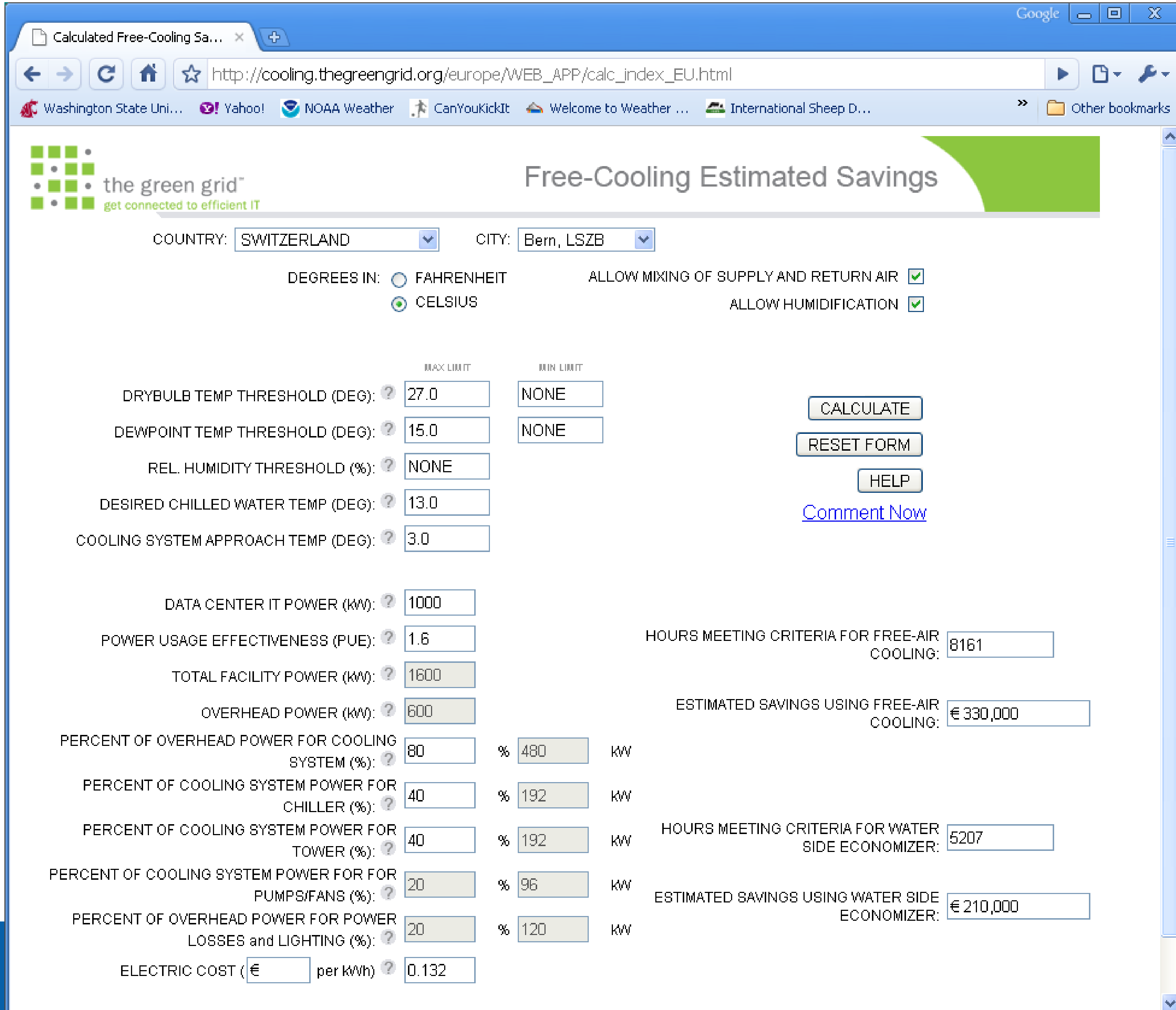
Chiller-less data center,
Cooling overhead – 6%

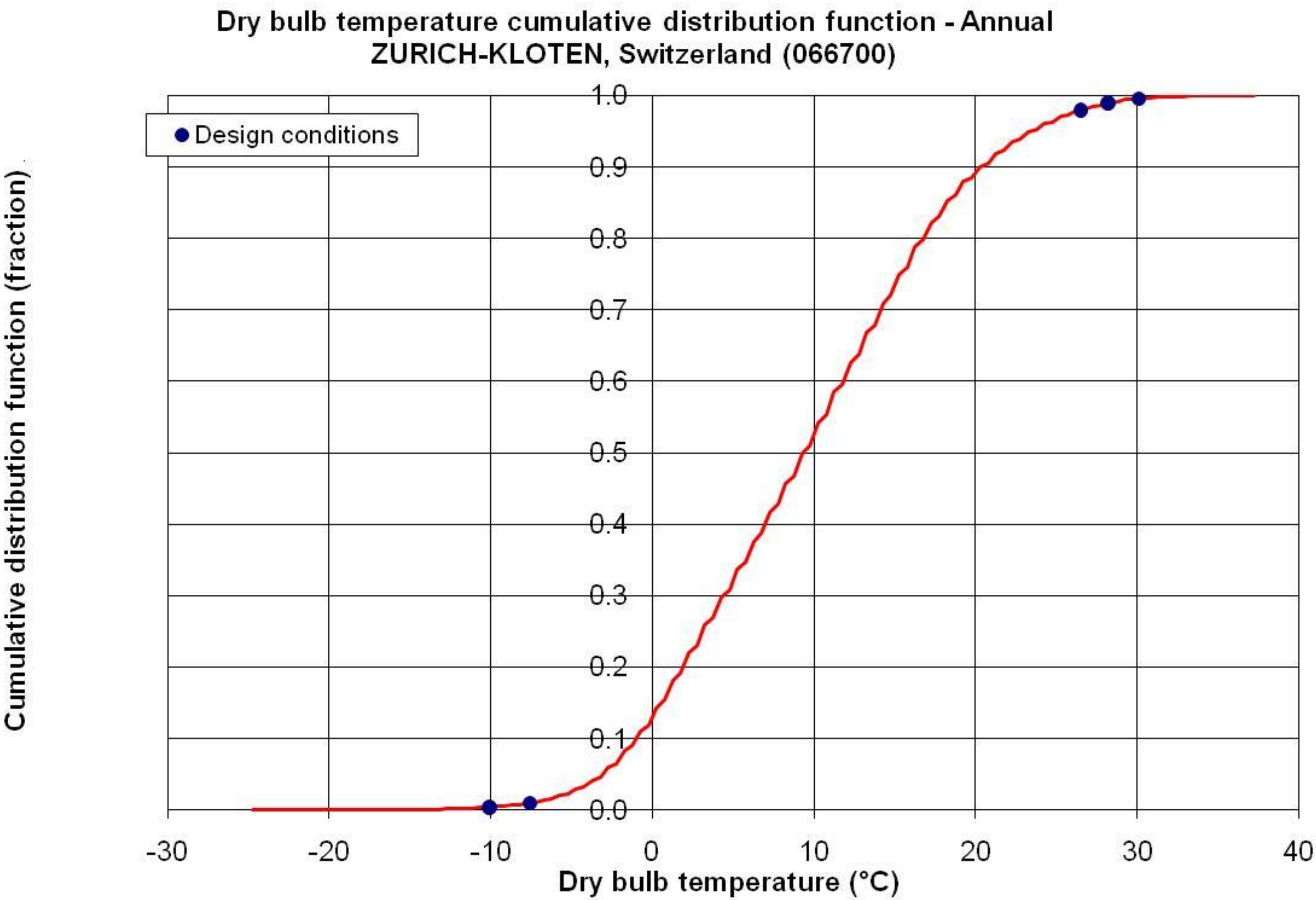
Cooling Limit *very rough* estimates +/-

Cooling Architecture	Limits	Comments
Open data center	4 kW	Least efficienct
Hot aisle/cold aisle	8-12 kW	Reguires specific layouts for higher end
Cold aisle containment	30 kW	Best for legacy DCs
Hot aisle containment or chimney cabinets	30 kW	Best for new construction
Liquid cooled rack enclosure	35 kW	Ideal for legacy DCs with limited airflow
Liquid rear doors	15-30 kW	Can make rack room-nuetral
In row coolers	20-35 kW	Space implications but rack independent
Close coupled coolers	10-25 kW	Overhead has space advantages
CPU liquid cooling	80 kW	Caution! must we cool the rest with air?
Integrated liquid cooling	100 kW	Benefits in Interconnect

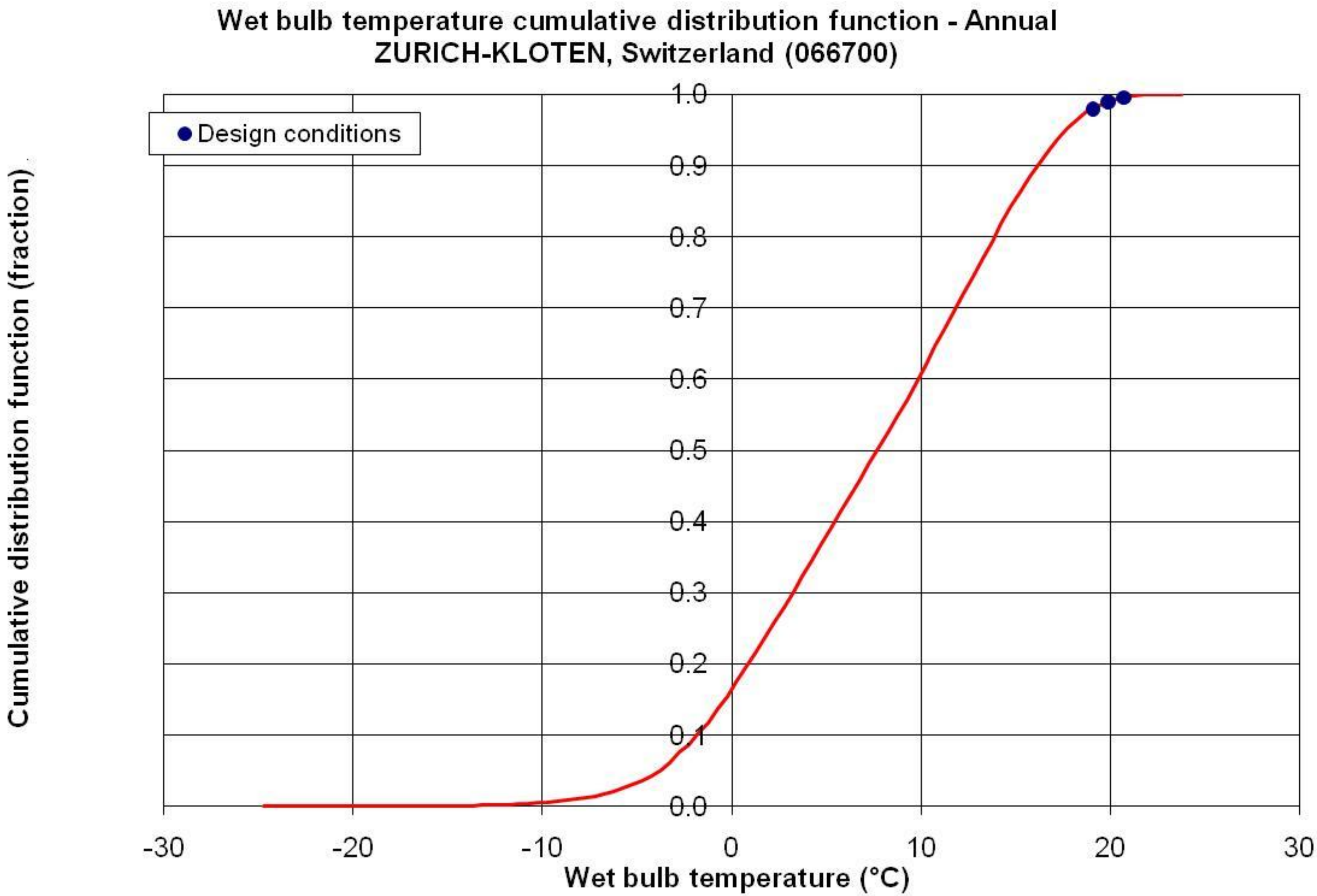

Air-side Free Cooling Map
the green grid
Estimate of
Air-side Economizer
Hours For Data Centers
© 2009 The Green Grid
Number of Available Hours Where:
Dry Bulb Temperature <= 27C (81F)
AND Dewpoint <= 15C (59F)
HOURS
8000
7500
7000
6500
6000
5500
5000
8000
7000
6000


Water-side
Free Cooling Map
the green grid
Estimate of
Water-side Economizer
Hours For Data Centers
© 2009 The Green Grid
Number of Available Hours Where:
Wet Bulb Temperature <= 10C (50F)
HOURS
7500
7000
6500
6000
5500
5000
4500
4000
3500
3000
2500
2000
7000
6000
6000
5000
5000
5000
6000
5000
5000
4000
4000
4000
5000
4000
3000
3000
2000
2000

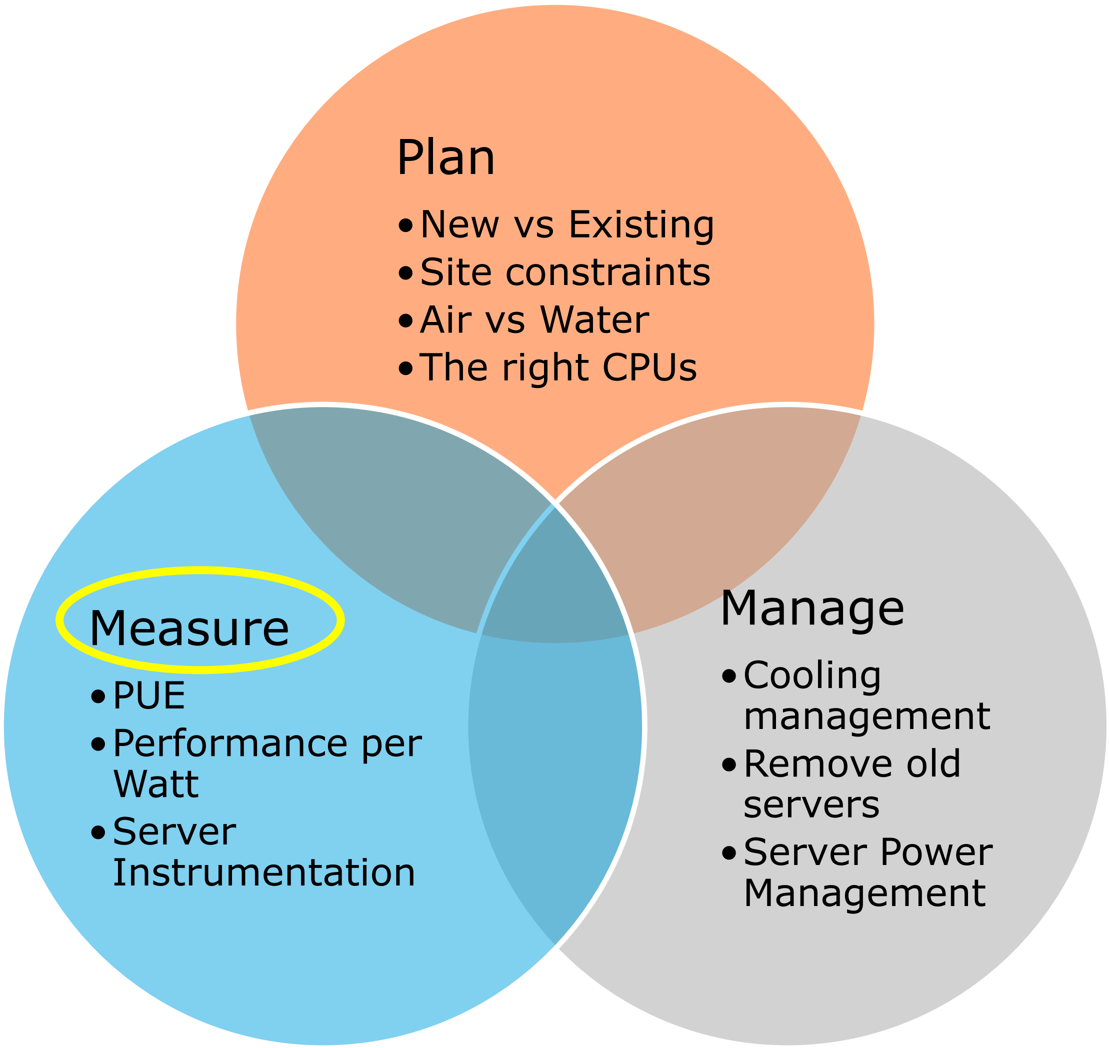

Calculated Free-Cooling Sa...
http://cooling.thegreengrid.org/europe/WEB_APP/calc_index_EU.html
Washington State Uni...
Yahoo!
NOAA Weather
CanYouKickIt
Welcome to Weather ...
International Sheep D...
Other bookmarks
the green grid
get connected to efficient IT
Free-Cooling Estimated Savings
COUNTRY: SWITZERLAND
CITY: Bern, LSZB
DEGREES IN: FAHRENHEIT
CELSIUS
ALLOW MIXING OF SUPPLY AND RETURN AIR
ALLOW HUMIDIFICATION
MAX LIMIT
MIN LIMIT
DRYBULB TEMP THRESHOLD (DEG): 27.0 NONE
DEWPOINT TEMP THRESHOLD (DEG): 15.0 NONE
REL. HUMIDITY THRESHOLD (%): NONE
DESIRED CHILLED WATER TEMP (DEG): 13.0
COOLING SYSTEM APPROACH TEMP (DEG): 3.0
CALCULATE
RESET FORM
HELP
Comment Now
DATA CENTER IT POWER (kW): 1000
POWER USAGE EFFECTIVENESS (PUE): 1.6
TOTAL FACILITY POWER (kW): 1600
OVERHEAD POWER (kW): 600
PERCENT OF OVERHEAD POWER FOR COOLING SYSTEM (%): 80 % 480 kW
PERCENT OF COOLING SYSTEM POWER FOR CHILLER (%): 40 % 192 kW
PERCENT OF COOLING SYSTEM POWER FOR TOWER (%): 40 % 192 kW
PERCENT OF COOLING SYSTEM POWER FOR FOR PUMPS/FANS (%): 20 % 96 kW
PERCENT OF OVERHEAD POWER FOR POWER LOSSES and LIGHTING (%): 20 % 120 kW
ELECTRIC COST (€ per kWh) 0.132
HOURS MEETING CRITERIA FOR FREE-AIR COOLING: 8161
ESTIMATED SAVINGS USING FREE-AIR COOLING: € 330,000
HOURS MEETING CRITERIA FOR WATER SIDE ECONOMIZER: 5207
ESTIMATED SAVINGS USING WATER SIDE ECONOMIZER: € 210,000



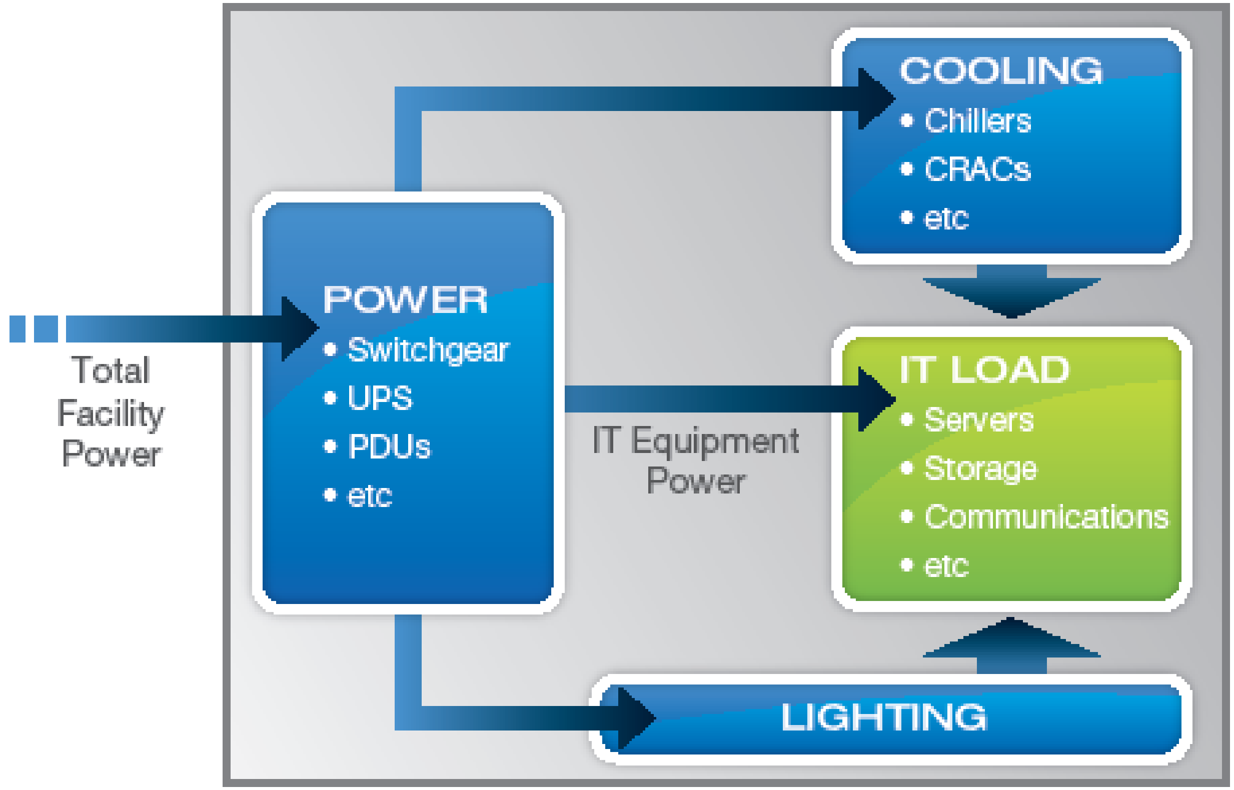
Dry bulb temperature cumulative distribution function - Annual
ZURICH-KLOTEN, Switzerland (066700)
Design conditions
Cumulative distribution function (fraction)
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-30
-20
-10
0
10
20
30
40
Dry bulb temperature (°C)


Wet bulb temperature cumulative distribution function - Annual
ZURICH-KLOTEN, Switzerland (066700)
Design conditions
Cumulative distribution function (fraction)
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-30
-20
-10
0
10
20
30
Wet bulb temperature (°C)

How to succeed going forward?



PUE: Power Usage Effectiveness



$$PUE = \frac{\text{Total Facility Power}}{\text{IT Equipment Power}}$$

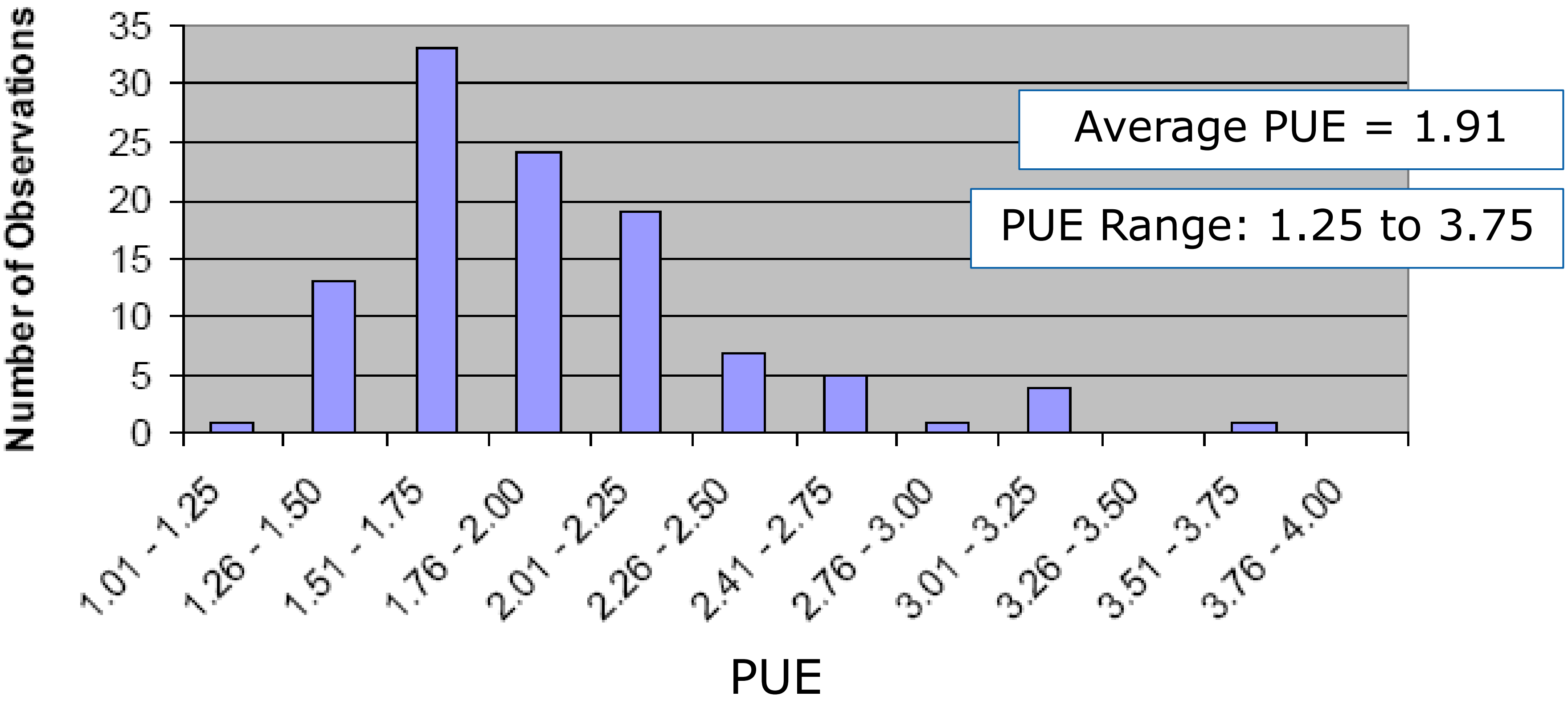
$$DCIE = \frac{1}{PUE} \times 100\%$$

Data Center Efficiency

$$\text{PUE} = \frac{\text{Total Facility Power}}{\text{IT Equipment Power}}$$

- Expressed as Power Usage Effectiveness (PUE)
  - Ideal PUE = 1.0
- Typical data center today = 2.0 →3.0+
- Best Practices ~1.5
- World Class ~1.2-1.3

EPA Energy Star for Data Centers first years results….



What to do? State of the art needed today for better PUEs

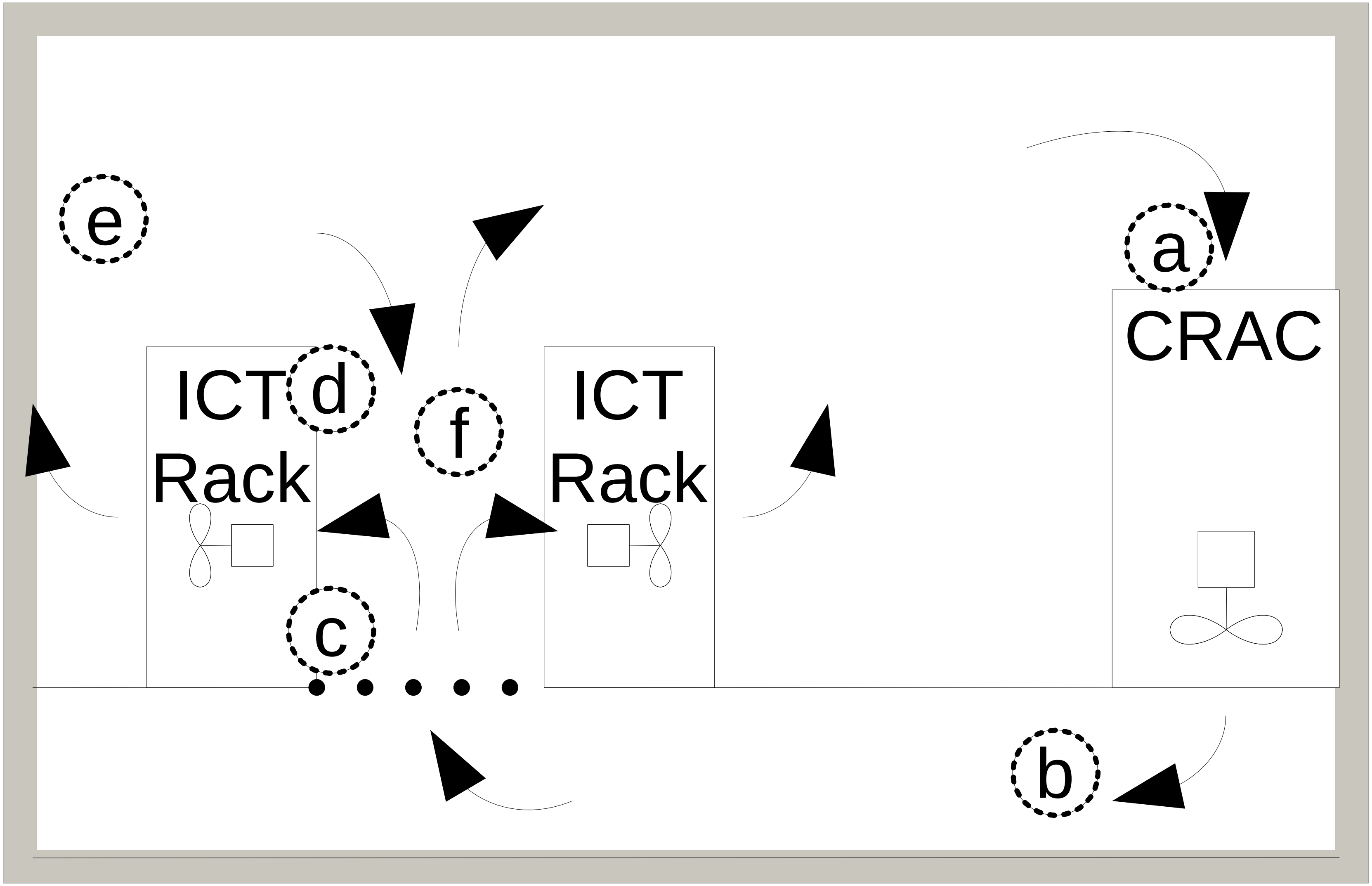
- ***Cooling***
  - Airflow management
  - Liquid cooled racks or rear doors
  - Close-coupled cooling
  - Economizers
    - Air, water, evaporative
  - ASHRAE Extended Envelope*
    - 18-27C
    - 5.5-15C dewpoint
  - Airflow management

- ***Power***
  - High efficiency power architecture, primarily spend the money for better components
    - Climate Savers Computing Initiative
  - 208Vac or 230Vac
  - 480/277 only in the very high end
  - Smart UPS (only on what you need)
  - DC power and HPC don't align due to limited UPS use

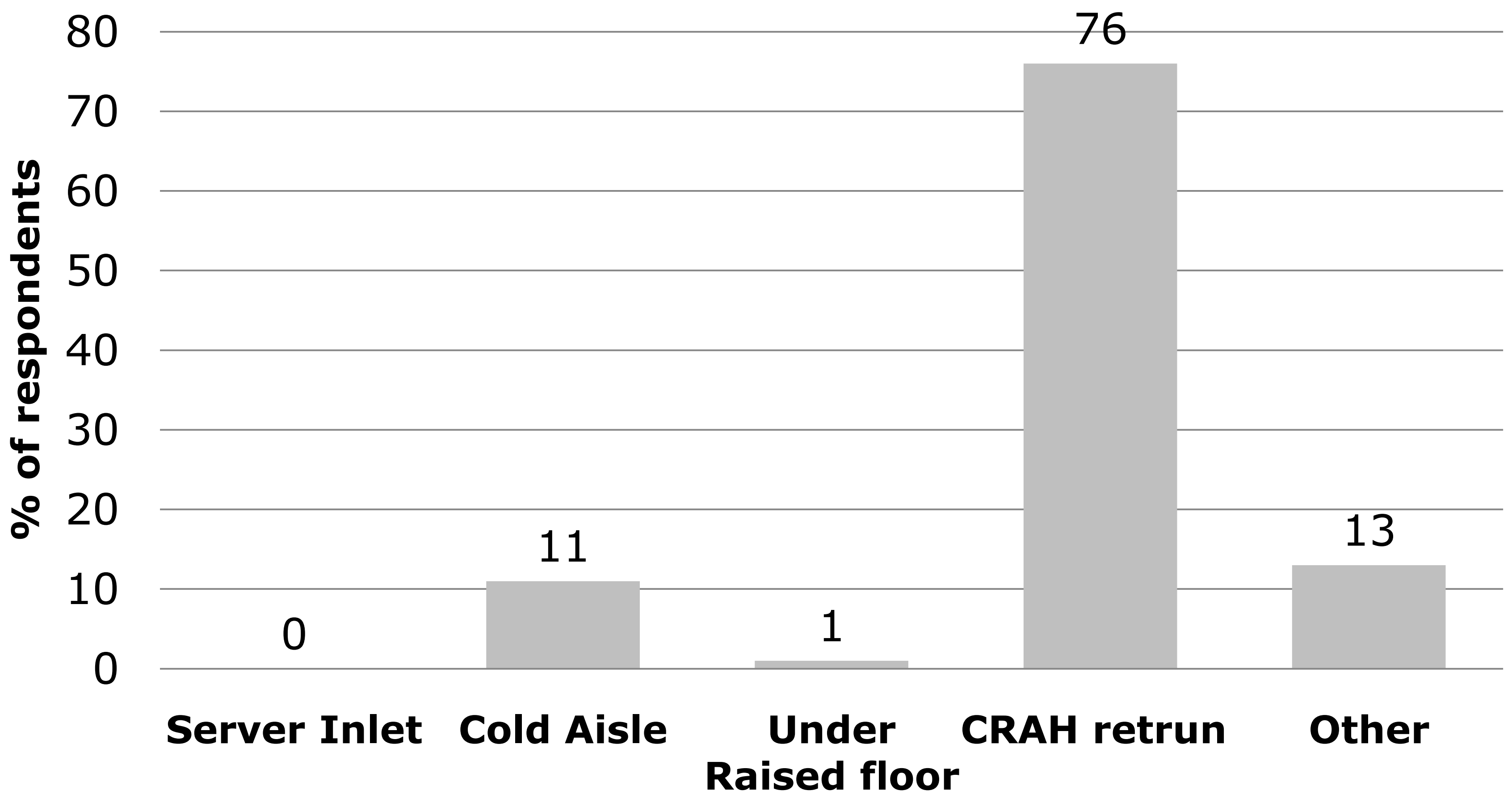
It's really all about good engineering!

*http://tc99.ashraetcs.org/documents/ASHRAE_Extended_Environmental_Envelope_Final_Aug_1_2008.pdf

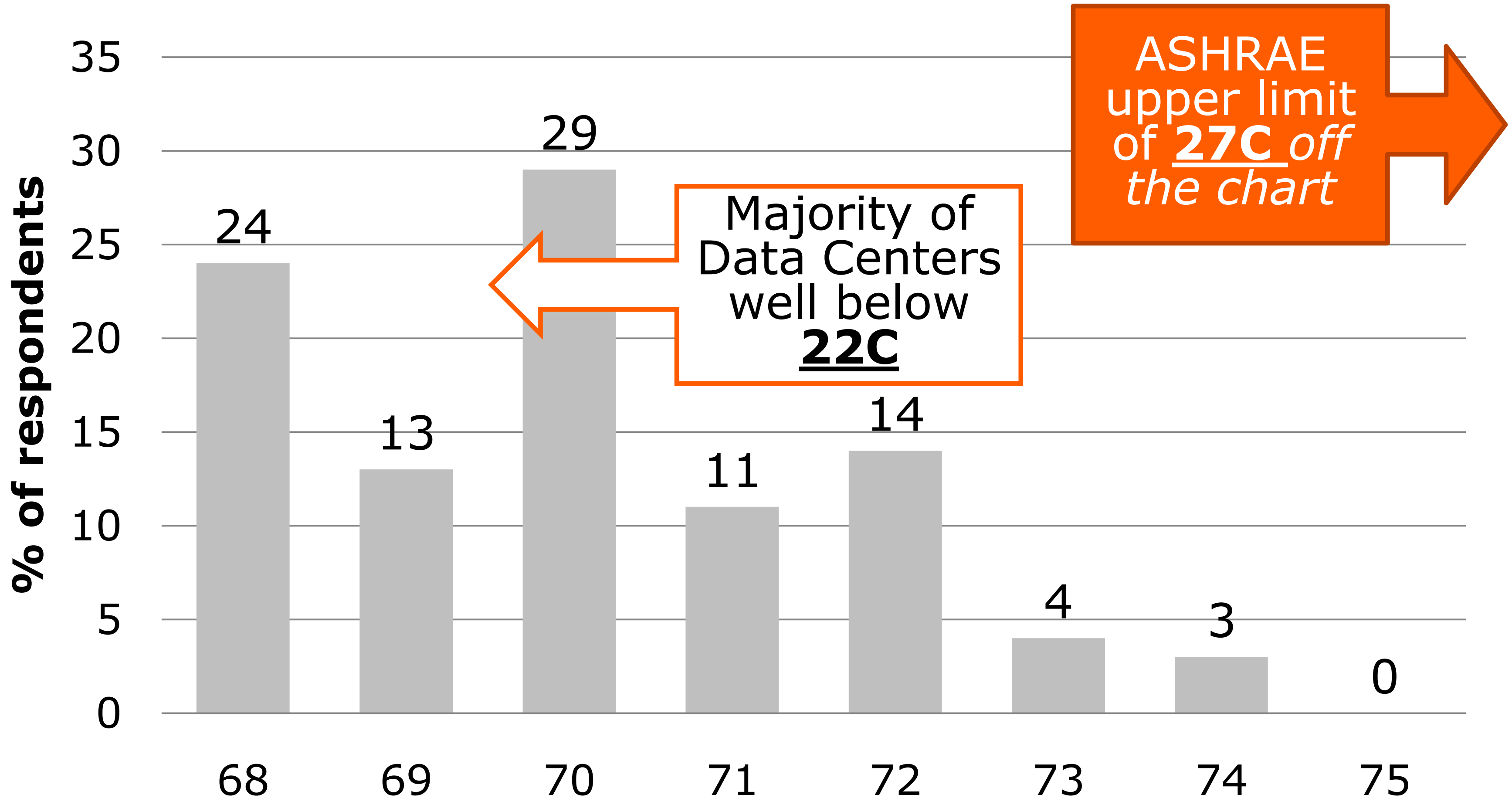
Typical DC & control strategies



CRAH/CRAC Set-Point Location
(from a recent Liebert DCUG survey)

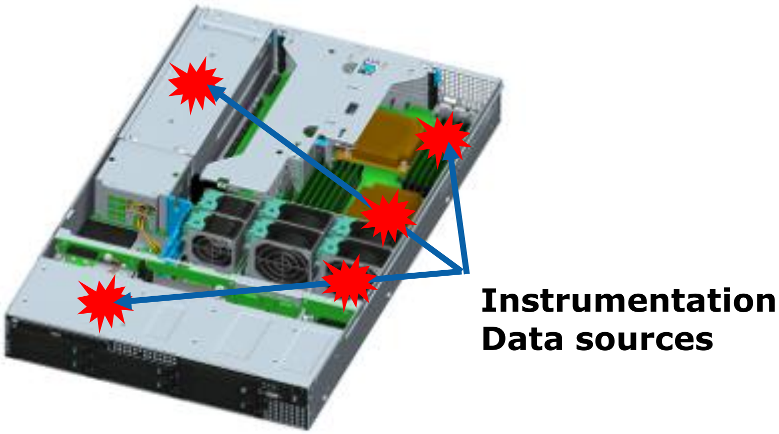


CRAH/CRAC Set-Point Values ^{o}F
(from a recent Liebert DCUG survey)



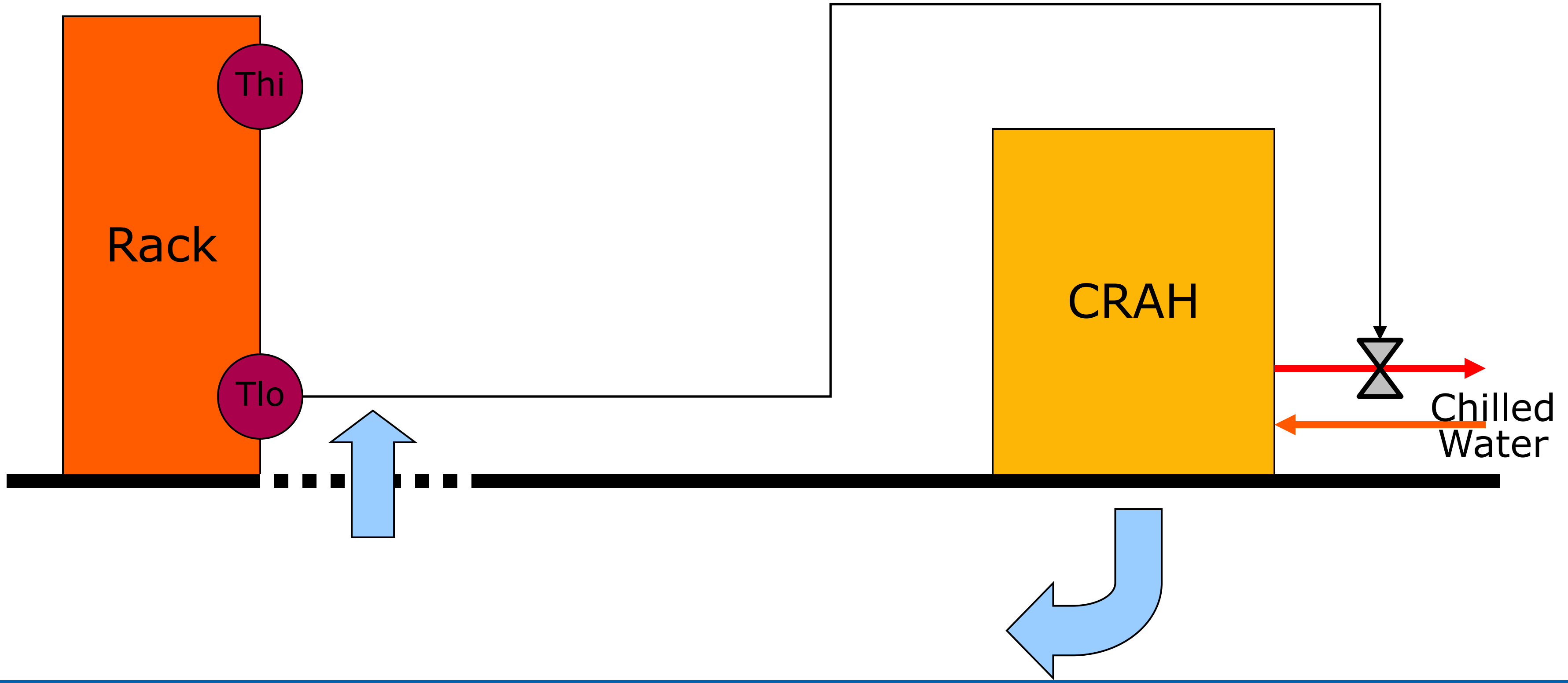
Server Instrumentation: Your System Efficiency Control Panel

- Instrumentation allows a device to report data about itself and provides ability to toggle control
- Server instrumentation today controls:
  - System power
  - Temperature
  - P & T states
  - Variable speed tachometer
  - Perf counters, ACPI, OSPM, and more

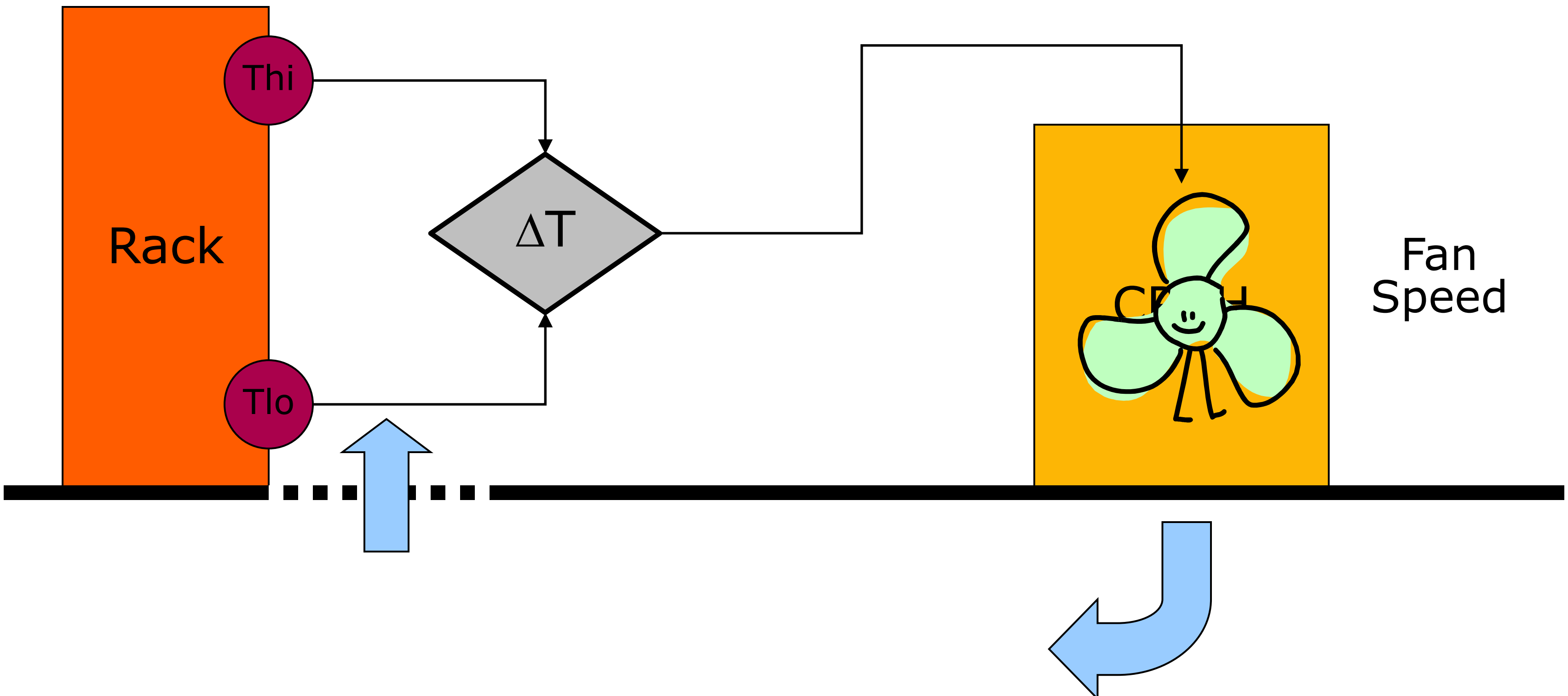


Instrumentation Provides Data Foundation For Server Management

Logical Controls Architecture
Chilled Water Valve Position = f(T_{lo})

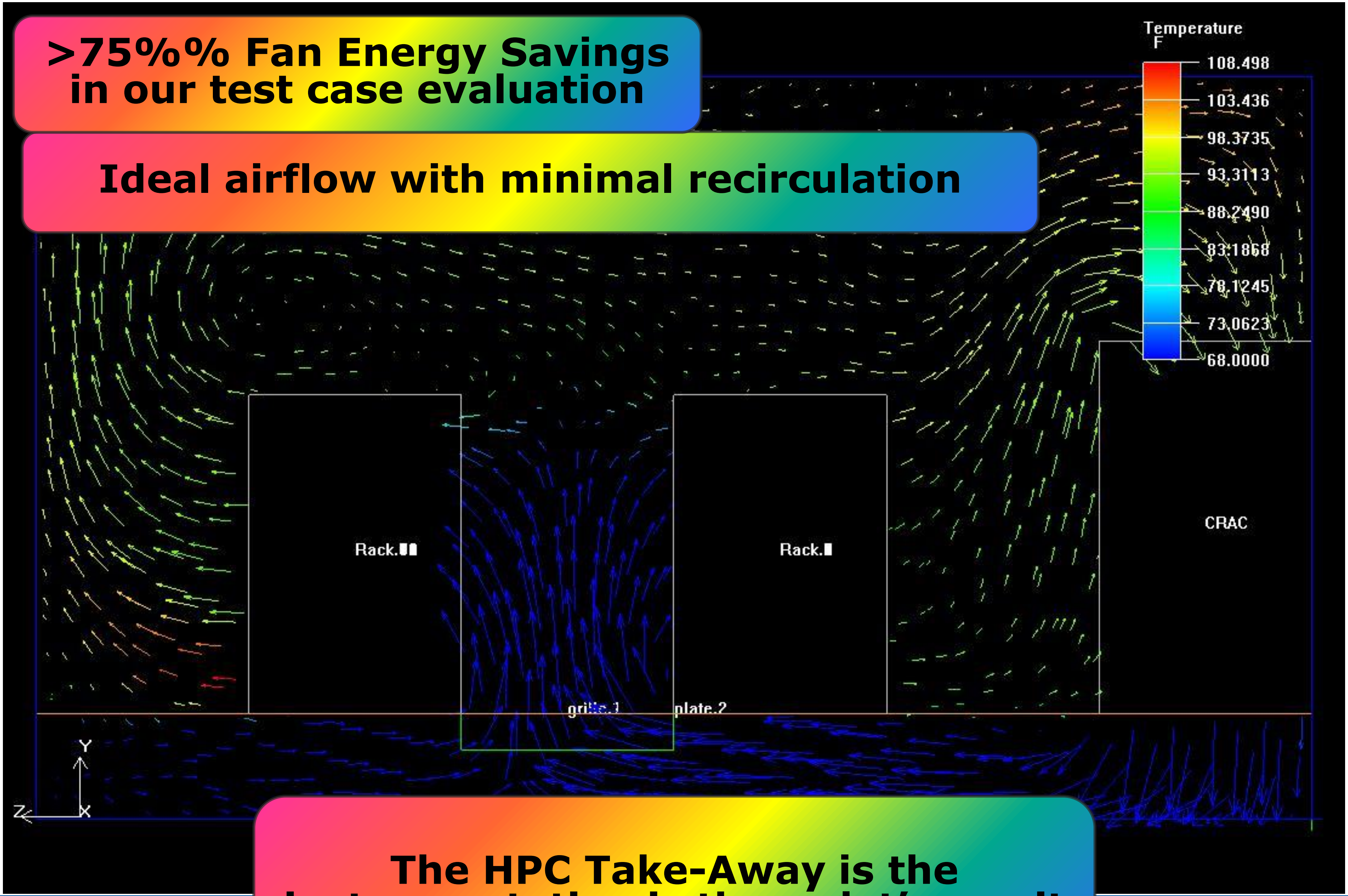


Logical Controls Architecture
CRAH Fan Speed = $f(T_{hi} - T_{lo})$

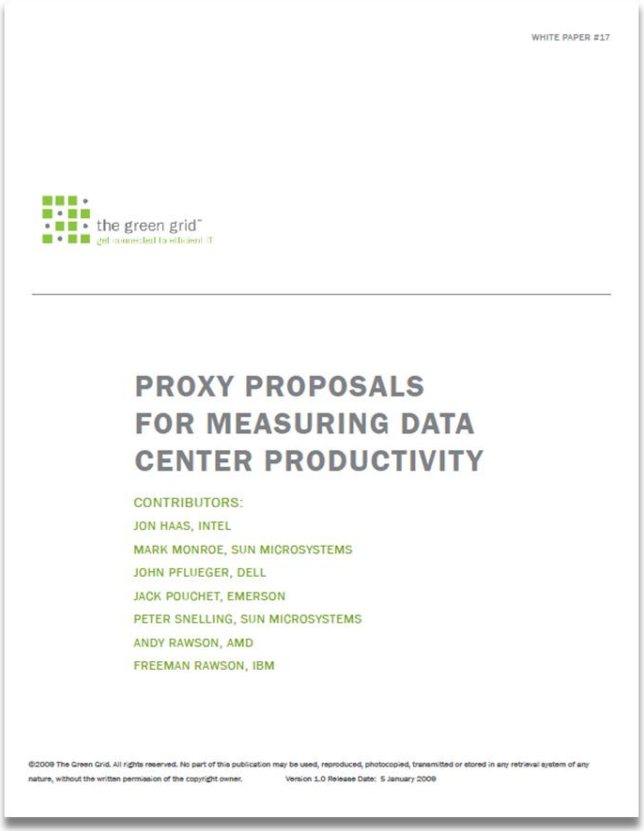


>75%% Fan Energy Savings in our test case evaluation

Ideal airflow with minimal recirculation



The HPC Take-Away is the instrumentation is there, let’s use it

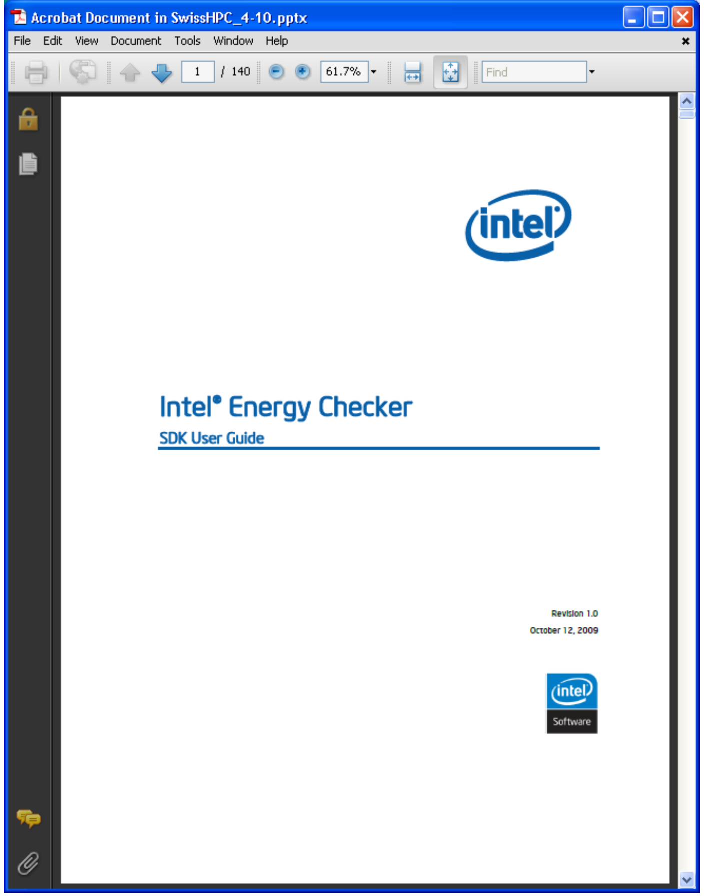

WHITE PAPER #17

the green grid

PROXY PROPOSALS FOR MEASURING DATA CENTER PRODUCTIVITY

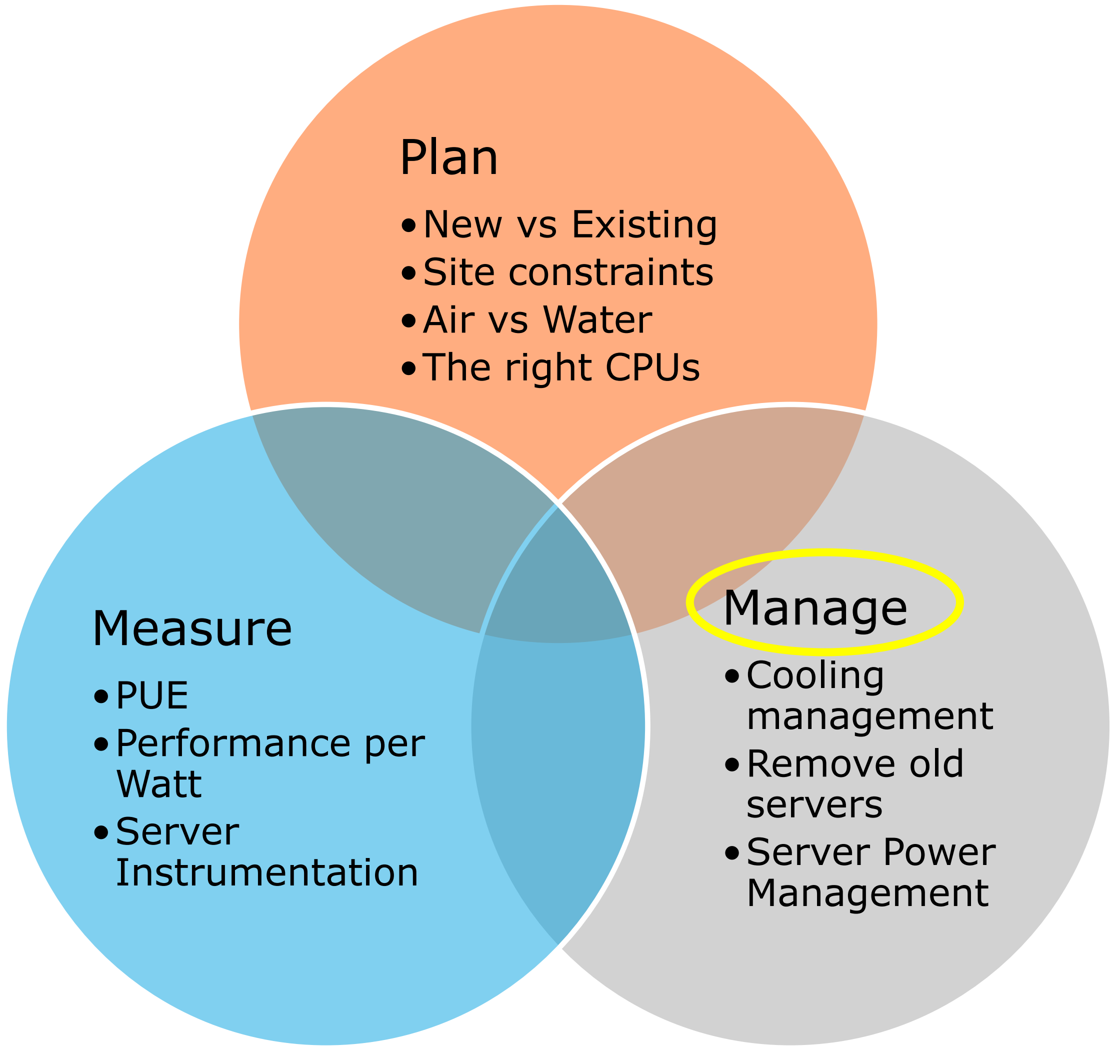
CONTRIBUTORS:
JON HAAS, INTEL
MARK MONROE, SUN MICROSYSTEMS
JOHN PFLUEGER, DELL
JACK POUCHET, EMERSON
PETER SNELLING, SUN MICROSYSTEMS
ANDY RAWSON, AMD
FREEMAN RAWSON, IBM

©2009 The Green Grid. All rights reserved. No part of this publication may be used, reproduced, photocopied, transmitted or stored in any retrieval system of any nature, without the written permission of the copyright owner. Version 1.0 Release Date: 5 January 2009

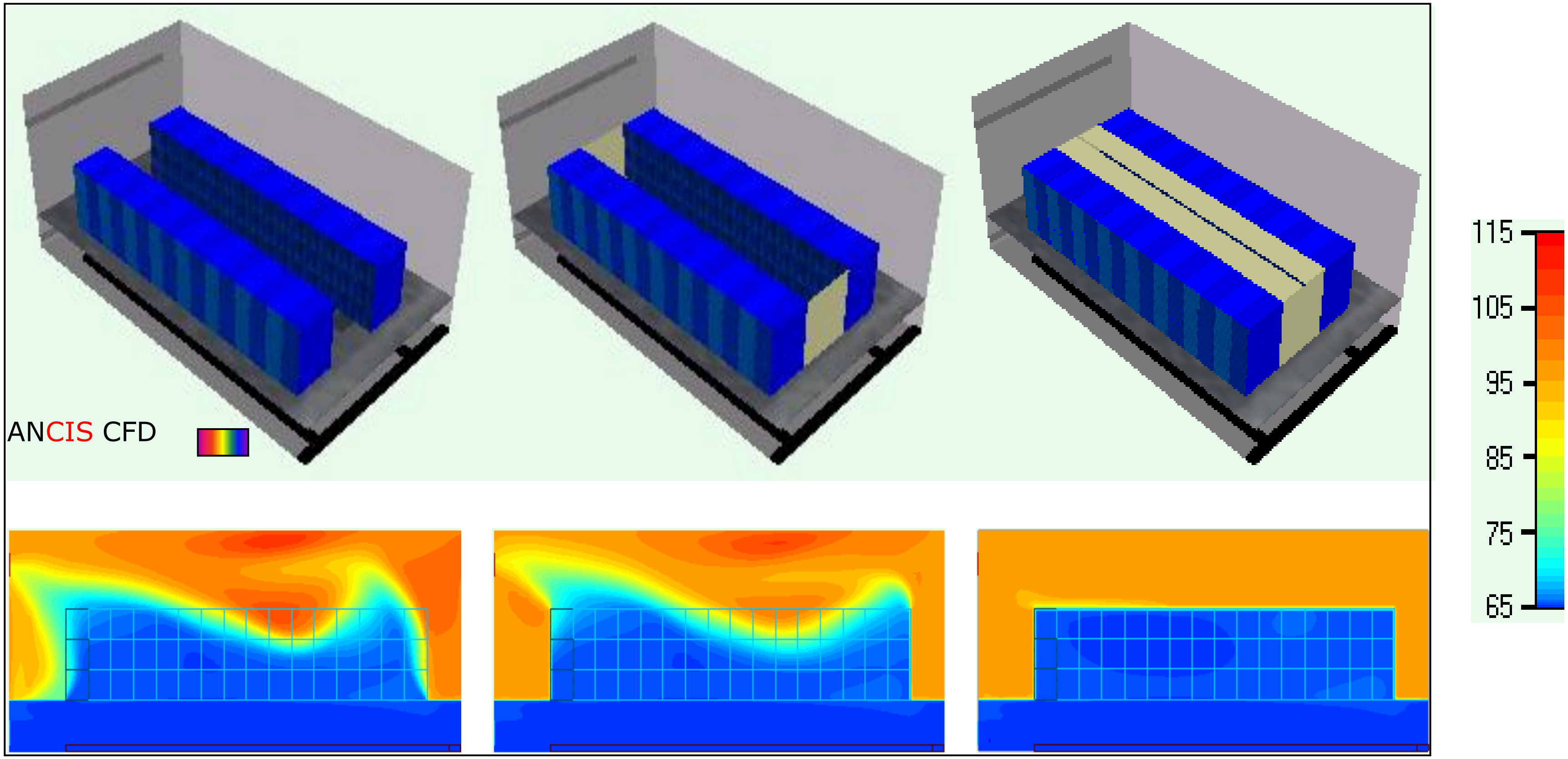


- The Green Grid is working on ***Performance per Watt*** metrics
- Current Activity with Pacific Northwest National Labs – HPC focus
- Explore Intel’s Energy Checker

How to succeed going forward?

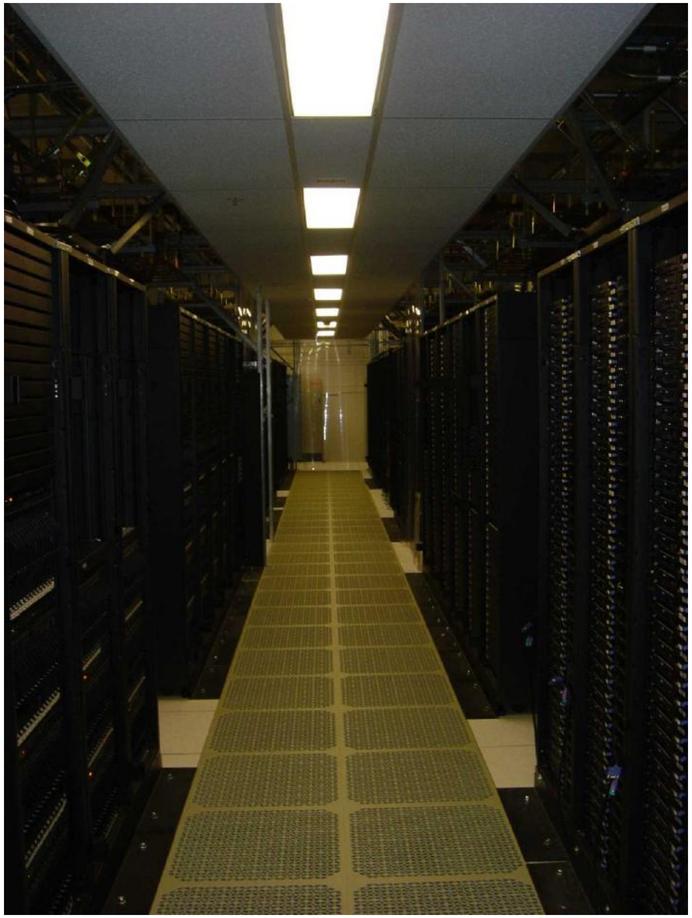


Why is Hot/Cold Air Segregation Important?



Open Semi-Enclosed Enclosed

Intel's SC4 retrofit; cold aisle containment

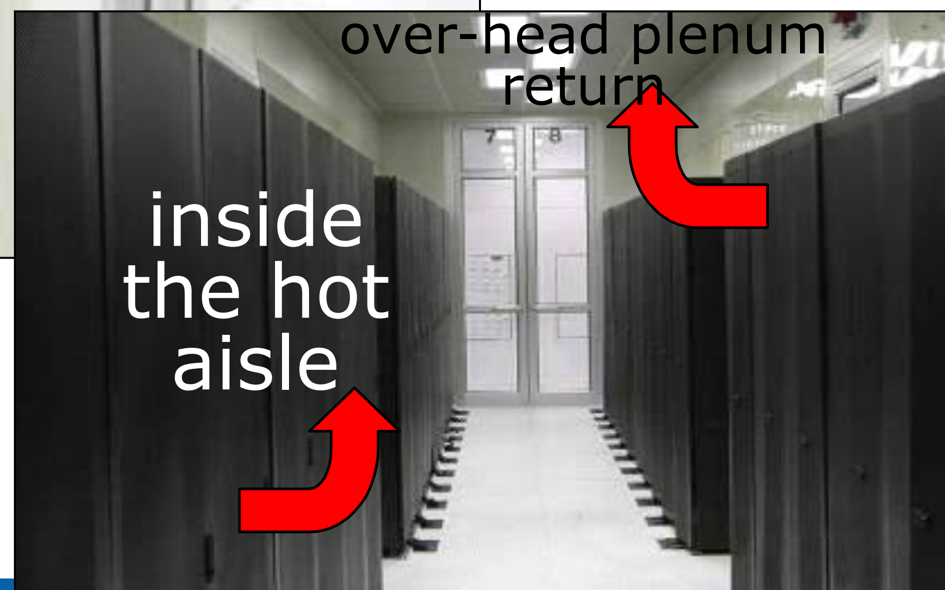


14 - 22 kW Racks

6000 w/sq m of raised floor

State-of-the-art air-flow management

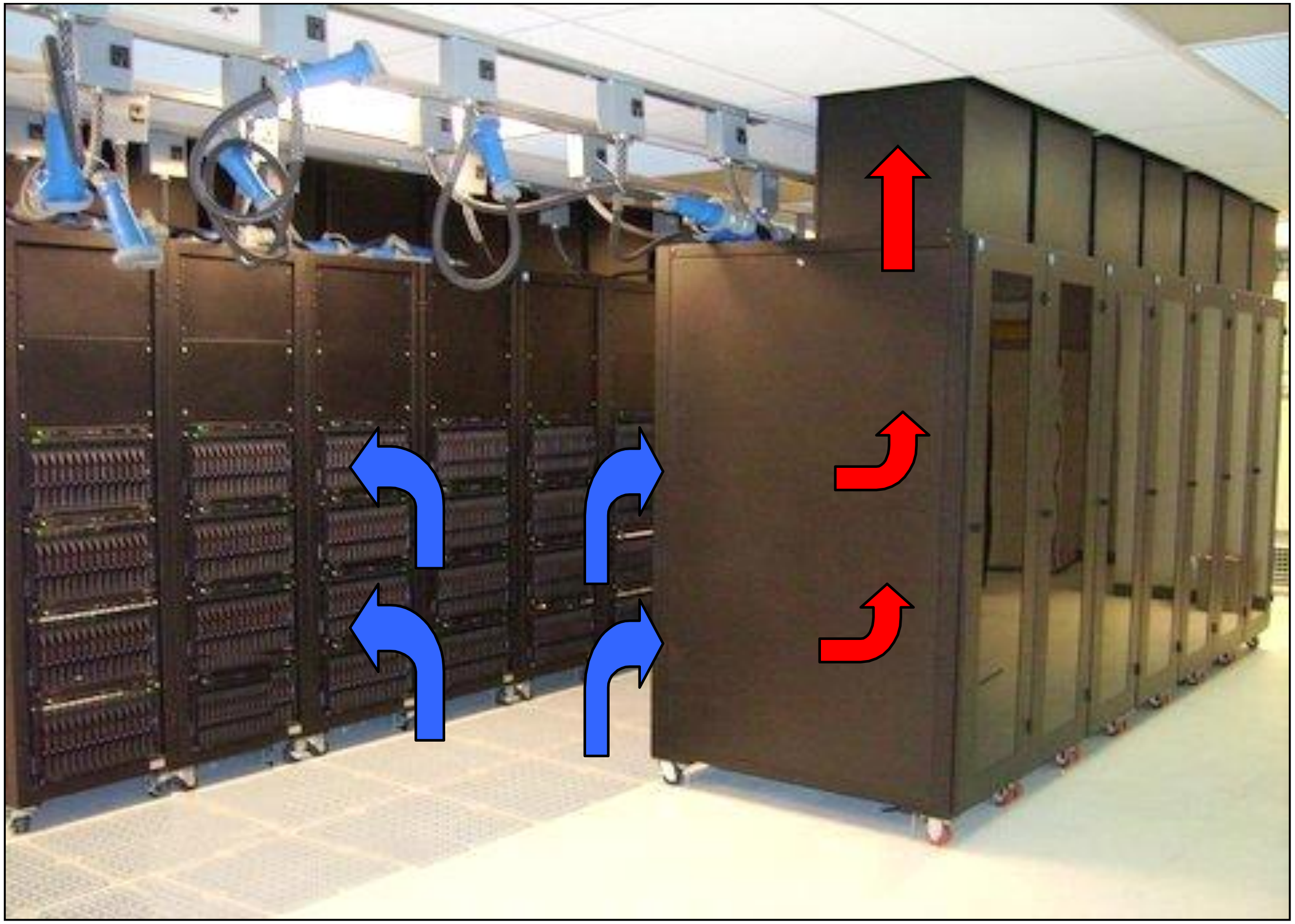
- Entire space becomes the cold aisle
- Raised floor optional



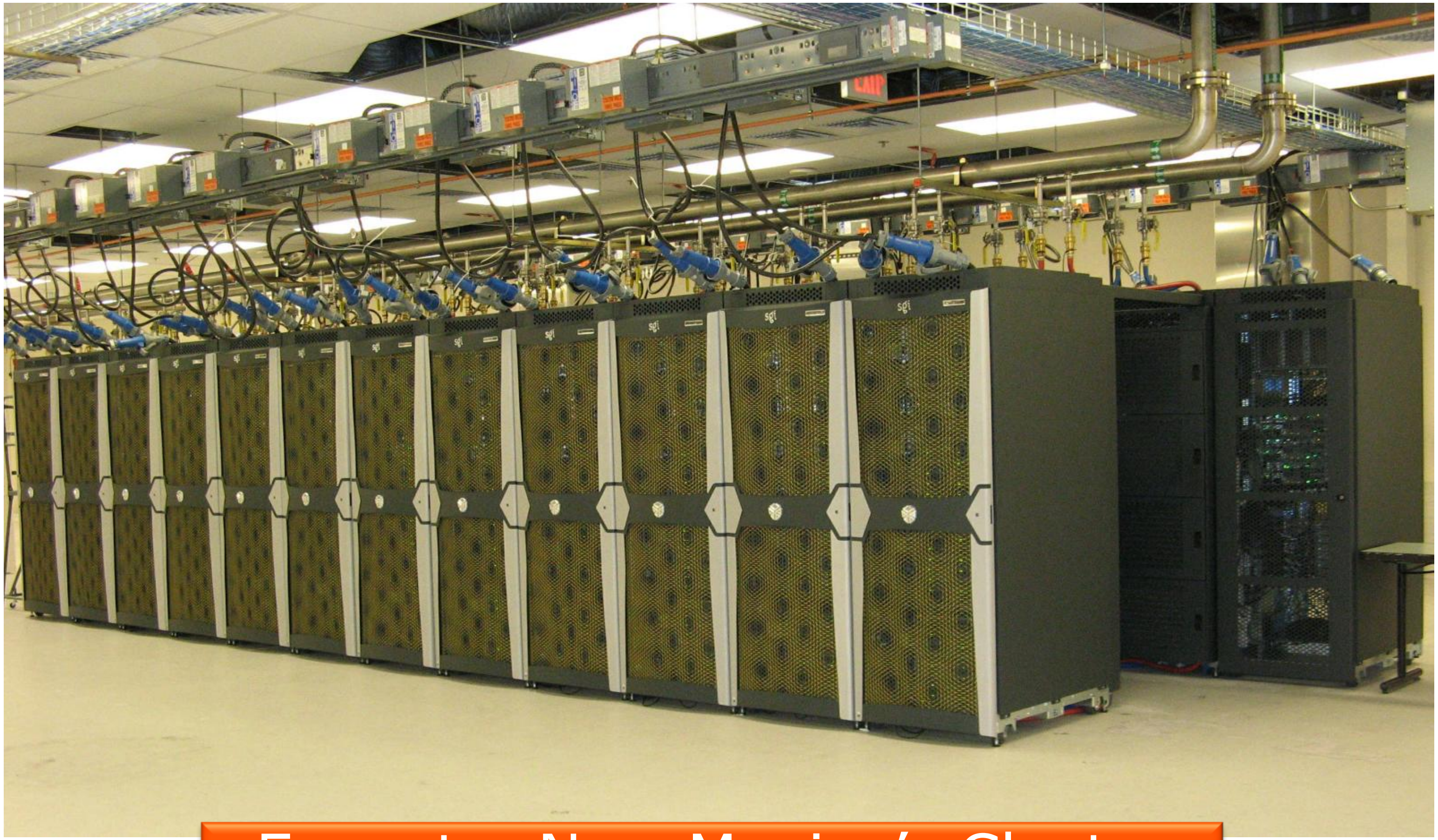
- Best for new construction
- Very efficient when combined with economizer
- Self balancing, low fan energy

Chimney Cabinets

- Variant of hot aisle enclosures
- Works w/ or w/o raised floor



Liquid Cooled 30kW racks



Encanto, New Mexico's Cluster

Air-cooled 30kW cabinets

w/ chimney cabinets

Air or Water? Doesn't really matter

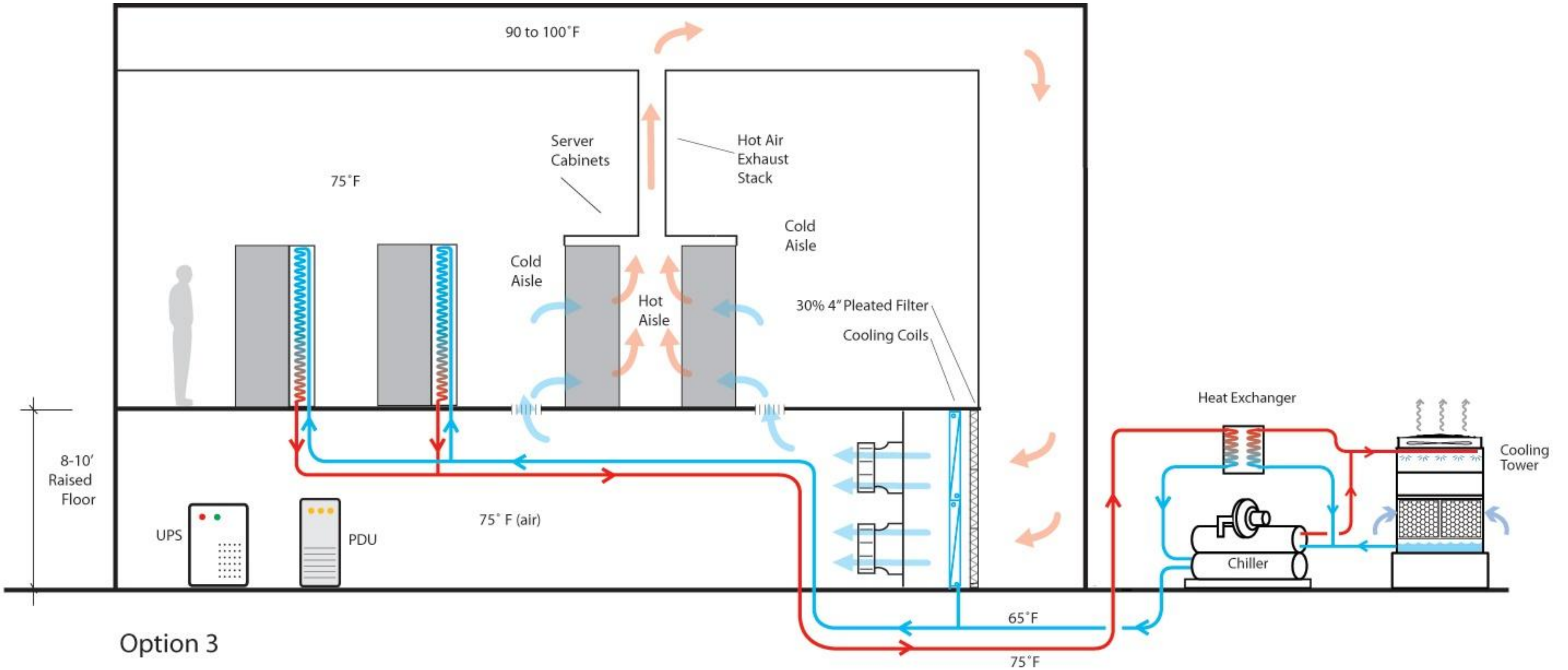
Liquid Cooled Racks

- Many shapes and sizes
- Applicable in retrofit data centers as well as new
- Ideal for extending legacy DCs where there is cooling, but limited airflow

NCAR-Wyoming Supercomputing Center

Graphic courtesy of H+L Architects & NCAR

NWSC Cooling Plan

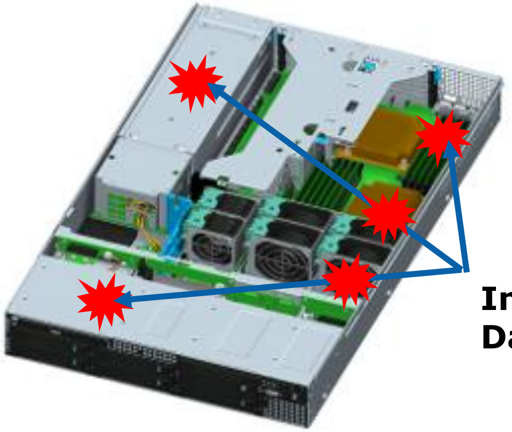


Design by Rumsey Engineers, Inc.

Intel's Commitment to Instrumentation

- Intel has been delivering and extending instrumentation in its processors and chipsets for years
- With Nehalem-EP, Intel is delivering capabilities that work with other component level instrumentation to extend instrumentation to the platform and data center

Capability	Intel Delivers
Power Reporting	Node Manager
Power Capping	Node Manager
Console Management	Data Center Manager
Simplified Management	Data Center Management Interface



Instrumentation Data sources

Delivering Instrumentation Capabilities That Address Data Center Issues

Introducing Intel Dynamic Node Manager: Streamlined Server Instrumentation

Hardware and Software Solution Delivered on System Processor

- Streamlines instrumentation design delivery
- Works with server hardware, BIOS, and OS to report and cap system power
- System budgets, specified by external management agent, are communicated over standard interface and can be updated real time
- Available on Intel next generation server platforms Nehalem-EP

- Improves rack density
- Increases performance density
- Delivers load balancing improvements
- Leads to higher availability and TCO

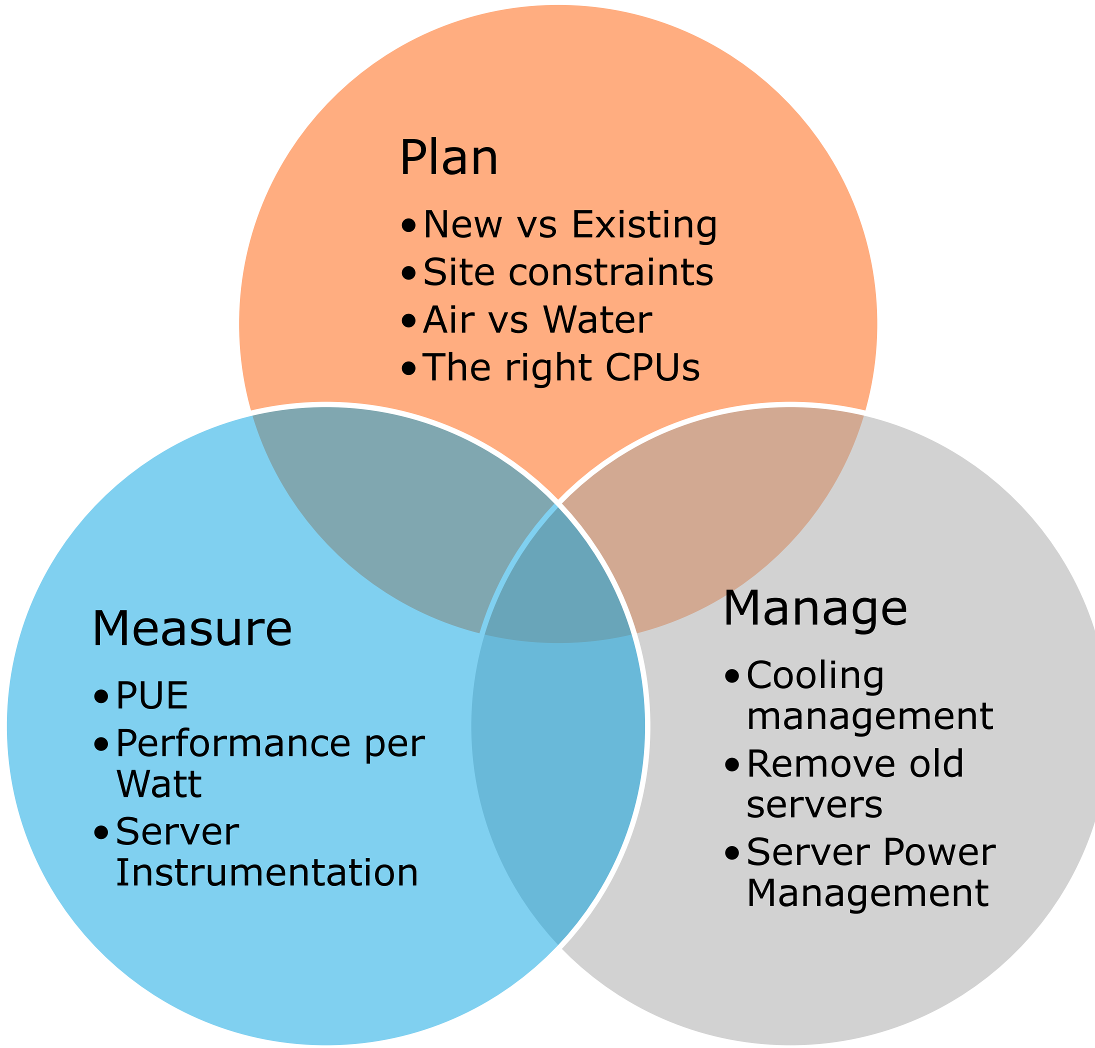
Planning for the future. State of the art needed tomorrow to maintain affordability

- *Cooling*
  - Depends on the density split (<40kW, >60kW)
  - Airflow management
  - Liquid cooled racks
  - Economizers
    - Air, water, evaporative
  - ASHRAE Extended Envelope
    - 18-27C
    - 5.5-15C dewpoint
  - Smarter controls, adaptable cooling

- *Power*
  - Purpose built power architecture
  - More 480/277Vac
  - 208Vac or 230Vac still an option
  - Smart UPS (only on what you need)
  - DC power and HPC don't align due to limited UPS use

It's ***still*** really all about good engineering!

How to succeed going forward?



Thank You! Questions?

back-up

Economizer resources

Proceedings of IPACK2009
InterPACK'09
July 19-23, 2009, San Francisco, California, USA

IPACK2009-89358

EVALUATION OF AIR-SIDE ECONOMIZER USE IN A COMPUTE-INTENSIVE DATA CENTER

Michael K Patterson
Eco-Technology Program Office
Digital Enterprise Group
Intel
Hillsboro, Oregon, USA

Don Atwood
Regional Data Center Manager
Information Technology
Intel
Rio Rancho, New Mexico, USA

John G Miner
IT Innovation Center
Intel
Rio Rancho, New Mexico, USA

ABSTRACT

Moore's Law continues to drive increased compute capability and greater performance per watt in today's and future server platforms. However the increased demand for compute services has outstripped these gains and the energy consumption in the data center continues to rise. The challenge for the data center operator is to limit the operational costs and reduce the energy required to run the Information and Communications Technology (ICT) equipment and the supporting infrastructure.

The cooling systems can represent a large portion of the energy use in the support infrastructure. There is significant focus in industry today on applying advanced cooling technologies to reduce this energy. One potential solution is the use of air-side economizers in the cooling system. This technology can provide a reduction in cooling energy by being able to maintain the required temperatures in the data center with the mechanical refrigeration turned off, significantly reducing the PUE for the data center.

This paper reviews recent industry activities around the recommended environmental conditions in the data center, the impact to the ICT equipment of air-side economizers, where they can best be applied, and provides data from a case study recently concluded at Intel's site in New Mexico. In that case study servers from an engineering compute data center were split into a standard configuration (closed system, tight temperature control) and a very aggressive air-side economization section (open system, significant out-door air quantities, moderate temperature control). Both sections performed equally well over a year long on-line test, with significant energy savings potential demonstrated by economizer side.

The American Society of Air-conditioning Heating and Refrigerating Engineers (ASHRAE) has recently published new ICT-vendor consensus-based recommendations for the environmental conditions in data centers. These new limits are discussed in light of the successful experiment run in New Mexico as the revised operational envelop allows a far greater number of hours per year when a data center can be run in "free-cooling" mode to obtain the energy savings.

Server design features as well as lessons learned from the experiment and their applicability to the potential use of air-side economizers is also discussed.

INTRODUCTION AND ANALYSIS

Server design continues to evolve and improve, reducing the energy per unit of computation. Moore's Law is continuing to provide society with the best energy efficiency gains of any industry or technology sector in history. Yet these gains come with challenges. Namely the power density available in servers

Gaseous and Particulate Contamination Guidelines for Data Centers

Whitepaper prepared by ASHRAE Technical Committee (TC) 9.9
Mission Critical Facilities, Technology Spaces, and Electronic Equipment

This ASHRAE white paper on data center airborne contamination was developed by members of the TC 9.9 committee representing the following IT equipment manufacturers: AMD, Cisco, Cray, Dell, EMC, Hitachi, HP, IBM, Intel, Seagate, SGI, and Sun.

Executive Summary

ASHRAE TC 9.9 committee recently published the 2008 ASHRAE Environmental Guidelines for Datacom Equipment which extended the temperature-humidity envelope to provide greater flexibility in data center facility operations, particularly with the goal of reducing energy consumption. The recommended temperature limits are from 18°C (64.4°F) to 27°C (80.6°F). The humidity is limited to less than 60% with the lower and upper dew point temperatures of 5.5°C (41.9°F) and 15°C (59°F).

The recent increase in the rate of hardware failures in data centers high in sulfur-bearing gases, highlighted by the number of recent publications on the subject, led to the need for this white paper that recommends that in addition to temperature-humidity control, dust and gaseous contamination should also be monitored and controlled. These additional environmental measures are especially important for data centers located near industries and/or other sources that pollute the environment.

Effects of airborne contaminations on datacenter equipment can be broken into three main categories: Chemical effects, mechanical effects and electrical effects. Two common chemical failure modes are copper creep corrosion on circuit boards and the corrosion of silver metallization in miniature surface mounted components. Mechanical effects include heat sink fouling, optical signal interference, increased friction, etc. Electrical effects include changes in circuit impedance, arcing, etc. It should be noted that the reduction of circuit board feature sizes and the miniaturization of components, necessary to improve hardware performance, also makes the hardware more prone to attack by contamination in the data center environment. Manufacturers are in a constant struggle to maintain the reliability of their hardware with ever shrinking feature sizes, without taking the added costly measure of hardening all their IT equipment, most of which is not installed in corrosive environments where it can be exposed to higher risk of failure.

Most data centers are well designed and are in areas with relatively clean environments and most contamination is benign. Most data centers should not experience particulate or gaseous contamination related hardware failures. This paper is primarily targeted at a minority of data centers which may have harmful environments arising from the ingress of outdoor particulate and/or gaseous contamination. In some rare instances, contamination has been known to be generated within the data center.

It is incumbent on the data center managers to do their part in maintaining hardware reliability by monitoring and controlling the dust and gaseous contamination in their data centers. Data centers must be kept clean to ISO 14644-1 Class 8. This level of cleanliness can generally be achieved by an appropriate filtration scheme as outlined here:

1. The room air may be continuously filtered with MERV 8 filters as recommended by ANSI/ASHRAE Standard 127-2007, Method of Testing for Rating Computer and Data Processing Room Unitary Air Conditioners.
2. Air entering a data center may be filtered with MERV 11 or MERV 13 filters as recommended by ASHRAE book titled, "Particulate and Gaseous Contamination in Datacom Environments".

Sources of dust inside data centers should be reduced. The gaseous contamination should be within the modified ANSI/ISA-71.04-1985 severity level G1 that meets:

1. A copper reactivity rate of less than 300Å/month, and
2. A silver reactivity rate of less than 300Å/month.

For data centers with higher gaseous contamination levels, gas-phase filtration of the inlet air and the air in the data center is highly recommended.

2008 ASHRAE Environmental Guidelines

	High Limit	Low Limit
Temperature	27C (80.6F)	18C (64.4F)
Humidity	Lower of 60% RH or 15C dewpoint	5.5C dewpoint

Opportunities for expanded temperature and humidity operation of data centers to allow less energy use in the cooling system; lowering PUE