Reducing Costs in the Data Center by Going Green

Share Tech Conference August 2010

Wayne Duquaine

Grandview Systems

Phone: 707-829-9633
E-mail: grandvu@sonic.net

Outline

- Overview of Electrical Energy Usage Stats
- Primary Energy Users in the Data Center
- Key Technical Terms Used for Power and Motor and Monitoring/Control
- Questions You Need To Ask When Evaluating New Equipment

Electrical Energy and Going Green

- Data Centers consumed 61 billion KiloWatt Hours of electricity in 2006, costing $ 4.5 Billion *
  - » 1.5 % of all U.S. Energy
  - » Enough to power 6.1 million homes

- Data Center power usage is doubling every 5 years.
  Data Center power costs more than double every 5 years.

- Using simple practices, typical data center power use can be easily reduced 10 %.
  - » More intense practices can reduce power usage 40-50 %
    i.e. cut your energy bill in half !

* Sources: DOE and EPA

What is a Computer ?

A Calculation Engine

A Heat Generation Engine

- Every kilowatt of power you pump into a CPU ultimately comes out as waste heat.

- Result = you need HVAC (Heating/Ventilation/Air Conditioning) to get rid of that heat

- Example:

	Blade H (14) PC Server	z10 E12 model
KiloWatts in:	4.208	9.52
BTUs out:	14,352/hr	32,500/hr

HVAC in the Data Center

- Every Server cabinet (Mainframe or PC) has cooling Fans
  - » Motor is used to drive the fan
- Chillers/Air Conditioners require a compressor/condenser for cooling.
  - » Compressor requires a fan to cool the refrigerant liquid.
  - » Motor is used to drive the fan
- Pumps must be used to pump chilled water.
  - » Motor is used to drive the pump
- Bottom line: motors are all over the place in a data center.
  - » The motors and your CPU(s) drive up your energy bill.
  - » You are not just running a computer center, you are running a motor control center as well !

HVAC continued

- HVAC is responsible for 1/3 or more of the Data Center energy bill
- In some shops, HVAC energy costs exceed the energy costs of the (PC) Servers
- Use of PC Server Blade technology can exacerbate the HVAC cost problem, since it reduces floor space, but increases heat density, creating a heavier load on the HVAC system.

Data Center: How Power is Used

Data Servers

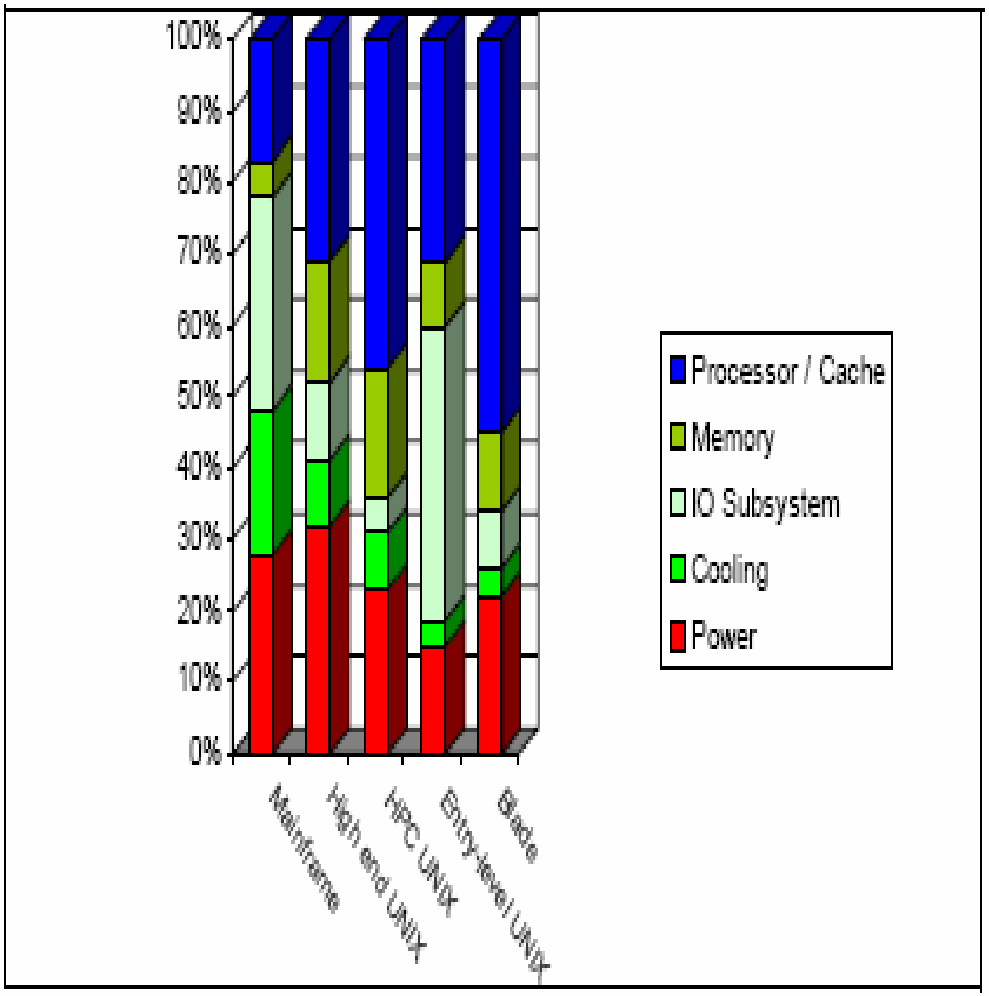
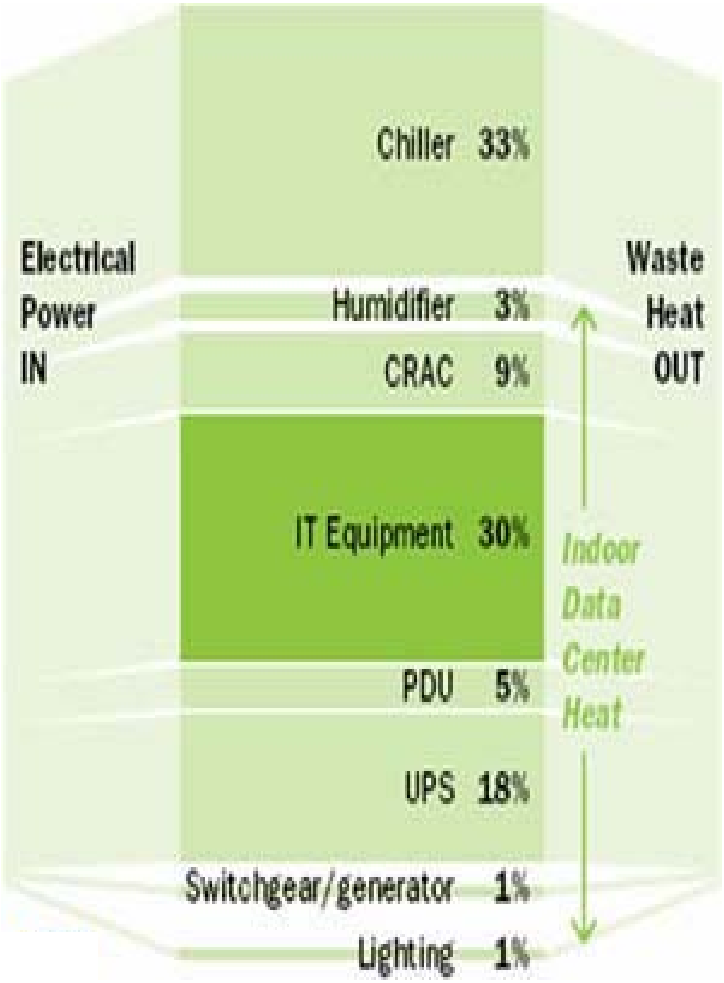


Figure 3-1 Typical relative power consumption by component for typical systems

Source: IBM Redpiece redp4413

HVAC/Power Supplies/IT



Source: The Green Grid

Rules of Thumb

- Computer Servers: for every $ 1.00 spent on new hardware, an additional $ 0.50 is spent on power and cooling
- HVAC: largest energy user in many Data Centers is the chiller equipment – can consume up to 1/3 of the total energy bill
- Motors: over a motor's lifetime, the cost of energy to run a motor will equal or exceed the original cost of the motor itself
- This stuff is not rocket science: a few pointed questions and an Excel spread sheet can quickly give you ball park numbers.

Summary of the Problem

- We have a self-reinforcing “vicious” cycle:
  - » Computers generate heat.
  - » We add motors to run HVACs to remove that heat.
  - » The motors themselves generate more heat.
  - » We then have to add more motors (or oversize existing motors) to remove the additional heat that the first group of motors created.
  - » All of this is costing us money.

Big Money !

Pieces Needed to Solve the Problem

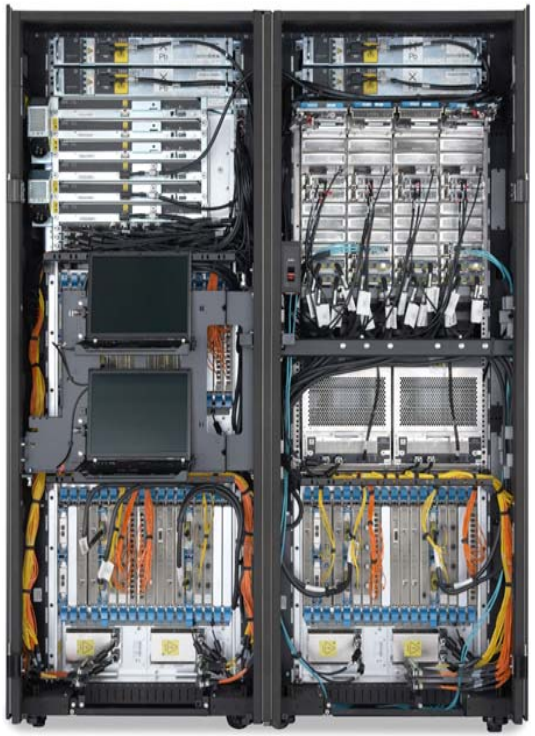
- Computer Server Energy Monitoring
- Motor Monitoring and Control
- Power (supply) Monitoring and Control
- Cooling / Temperature Monitoring and Control
- Virtualization of PC Servers
- Monitoring Tools that Coordinate all of the Above

Computer Power Management Options

- Processor:
  - » Dynamic frequency and voltage scaling (DVS)
    - – Throttling down the CPU frequency can dramatically reduce power usage
    - – Reducing voltage to CPU cores that have gone idle drops power usage
  - » Power capping – Max limit on Watts/BTUs used
    - – Turn off or throttle down one or more CPU cores when power cap limit hit
  - » Increase hardware's MPL – PCs usually only run 30 % utilization
- Memory:
  - » Put idle memory into low-power mode until it is requested
  - » Switching back to active mode typically takes only 1 memory cycle
- DASD Storage
  - » Storage Proliferation is out of control, especially PCs/Web Servers
  - » Better management software and better electronics are coming

Computer Power Monitoring: Mainframe (z10)

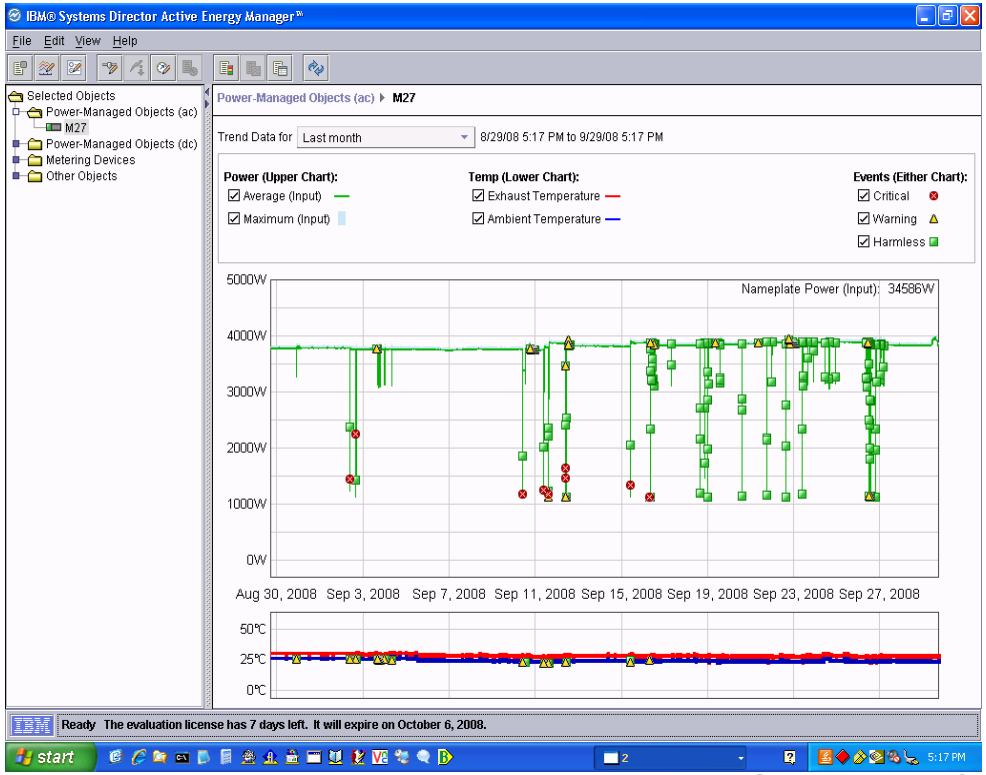
- z10 provides the following monitoring data:
  - » Average Power Use over a 1 minute period
  - » Peak Power Use over a 1 minute period
  - » Exhaust Temperature
  - » Ambient Temperature
  - » "Events" – e.g. changes in fan speeds, …
- No direct control over z10's processor DVS
- Monitoring is done via IBM's System Director Active Energy Manager software



Green Data Ctr

Active Energy Manager (AEM)

- IBM's flagship Computer Power Management Software
  - » Cross platform: mainframes, Linux Server, PC Servers
  - » Power Monitoring for mainframe
  - » Power monitoring and control for PC Servers and Linux boxes
    - – Provides DVS support to tweak CPU power usage
  - » Powerful graphing/trend analysis support (Power, Temperature)
  - » Scripting support
  - » Runs on PCs and Linux
  - » Provides some support for other vendor's equipment – UPS, PDUs, … such as Emerson, Eaton, Liebert …

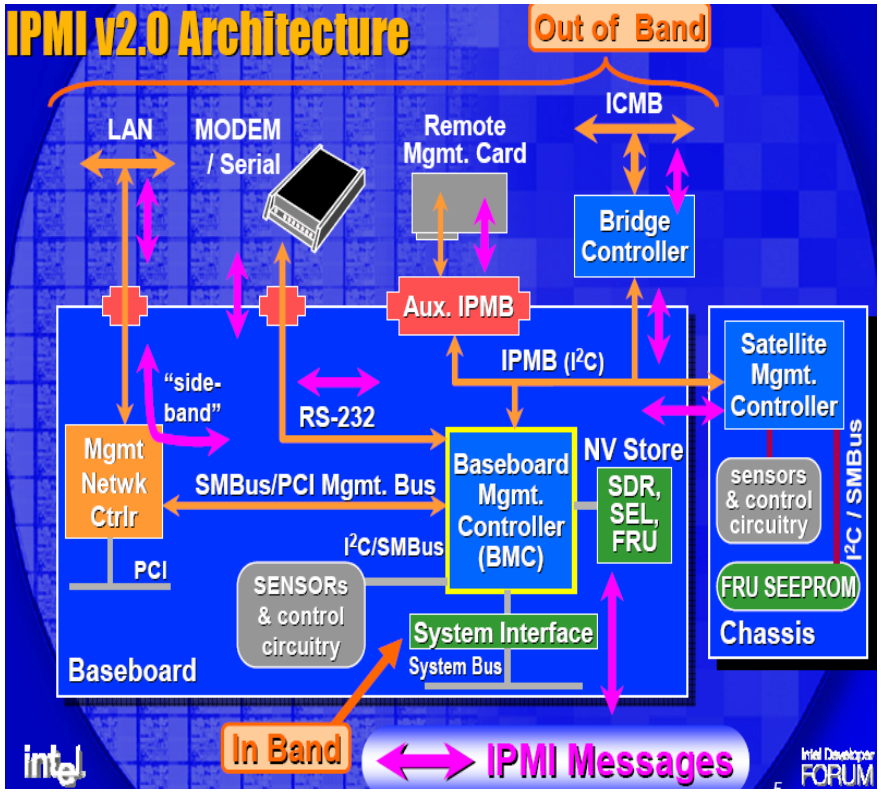


Computer Power Monitoring: PC Servers, Linux

- PC Based Power Monitoring/Control
  - » PCs provide both monitoring and control of power usage
  - » ACPI (Advanced Configuration and Power Interface) – Base OS software provided by Microsoft that provides access to PMBus and CPU monitoring hardware
  - » IBM AEM – DVS for their PC Servers
  - » HP Insight Power Manager – DVS and CPU Capping
  - » Verdiem – Power Mgmt software across multiple PC Vendors
  - » Several software vendors now providing simple scripting to ACPI
- Linux Based Power Monitoring/Control
  - » OSPM (Operating System directed Power Management) – Linux provided support for ACPI
  - » IBM AEM – DVS and CPU Capping for their pSeries (AIX/Linux)
- IPMI (Intelligent Platform Management Interface) - support for cross platform

Computer Power Monitoring: IPMI

- Message based interface to access/control a hardware platform's monitoring/control functions
- Endorsed by over 150 vendors, including all PC and most Linux vendors. Coordinated by Intel
- Supported on Windows and Linux. Open Source is available (OpenIPMI ships with Linux)
- Adding XML/SOAP support to enable access via Web Services



Computer Monitoring: PMBus

- Every PC (Server, Personal, Laptop) today has PMBus built in.

- Whenever you hear the fans in your laptop spinning up or down, it is because the internal PMBus is kicking in
  - » The cheap versions have fixed fan speeds. (poor energy efficiency)
  - » The better versions have variable speed fans that adjust the fan speed up and down for optimum power use.

- What is PMBus (Power Management Bus)
  - » Two-wire (twisted pair) digital bus, based on I2C sensor standard
  - » Has commands to: request fan speed and component's temperature, vary fan speed up or down
  - » Typically implemented with a single, small chip on PC Motherboard

- New Software is now allowing customer access to these controls (ACPI, IPMI software)

Motor Monitoring + Control - Overview

- Motors make up ¼ to 1/3 your Data Center energy bill. Most of the HVAC bill is due to motors.

- Motor Control is all about controlling current (going into motor). Motor windings are inductive, so it is current (amps) that affects them most.
  - » Electrical Power = Watts = Volts * Amps
  - » Amps = Current

- The type of motors your equipment uses can dramatically affect your power bill.

- Many new motors, for the last 8+ years are “intelligent”, and can be both electronically monitored and controlled
  - » Small Micro-controllers (MCUs) and Sensors are attached to the motor

Motors – Basics - 1

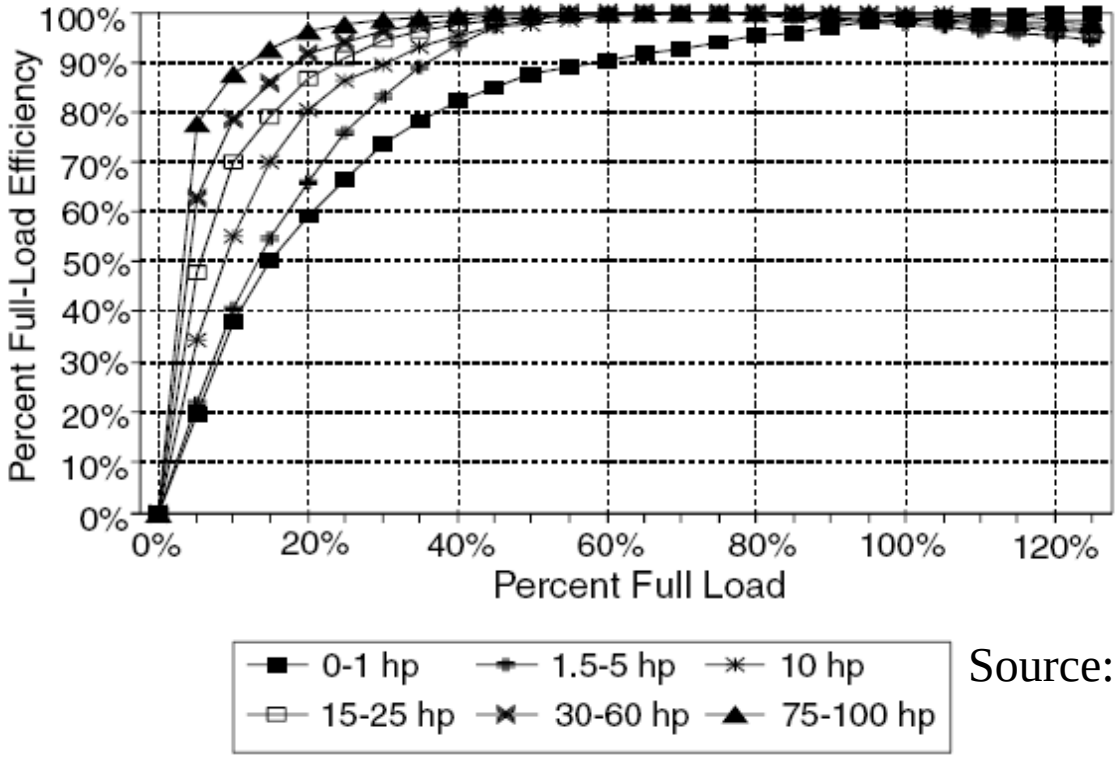
- Dumb Motors
  - » Brushed DC (BDC) Motors and simple AC motors.
  - » 90 % of current motors in use today are “dumb” motors.
  - » Are very cheap to build and buy.
  - » Dumb motors in operation typically only 48-49 % efficient – I.e. over half the power used to run the motor gets burned up as waste heat.
  - » 50 % of all U.S. electrical energy is spent driving some kind of motor.

Motors – Basics - 2

- Intelligent Motors aka "Variable Frequency Drives" (VFD)
  - » Brushless DC (BLDC) or ECM AC motors most common.
  - » Efficiency ranges from 65 – 94 %
  - » Micro-controller (MCU) based. The MCU electronically controls all the major operations (commutation) of the motor.
  - » Are about 1.25 to 1.4 times more expensive than a "dumb" motor.
  - » Many intelligent motors provide cabling to monitor and control the motor
- Remember – over the motor's lifetime you will spend as much or more on energy to run the motor, than the motor's original cost.

Motors – Right-Sizing

- Over-sizing Motors can work against you.
- For most motors, running them with less than 40 % load, causes efficiency to drop off quickly, wasting energy.
- Objective is to keep it running between 75-100% Load.
  - » For many motors (Fans, …), reducing speed by 20 % (during lower demand) will reduce motor's power needs by 50% (VFD vs fixed speed)



Source: US EIA

Power Supplies

- The humble power supply is quite ubiquitous (and can also be a major power waster)
  - » 3.1 Billion Power Supplies in the U.S. *
  - » Typical efficiency is 30-60 %
  - » Energy lost each year from power supply losses (inefficiency) is 110-150 billion KW hours/year = 3 - 4 % of total electrical usage in U.S.
- Typical Computer Room Power Supply is 70-95 % efficient
  - » Switch Mode Power Supply (SMPS) is most common
- In large PC Server clusters, the combination of HVAC (fan motors, etc) and Power Supplies can be 50 % of entire energy bill (e.g. Google)
  - » Google had to go out and start building their own custom power supplies

* Numbers courtesy of Texas Instruments

Power Supplies - contd

- Almost none of your AC power goes directly into your computers nor HVAC motors
  - » It gets sliced, and diced, and inverted, and twisted, bent, and spindled.
  - » The inefficiency of your power supplies (inverters, et al) can cost you 10-25% or more in energy losses

- Key Terms
  - » Inverter - converts AC to DC, or DC to AC, or DC to DC
  - » Boost Convertor - increases voltage up
  - » Buck Convertor - decreases voltage down (e.g 5.0 V to 3.3 V)
  - » PFC – Power Factor Correction – cleans up your power, reduces losses

- Inverters, Buck, and Boost - heavily used in PC Servers. Inverters - heavily used in HVAC motor drivers.

Power Converter Basics

- Conventional (“dumb”) power supplies typically 30-60% efficient
- Intelligent power supplies (SMPS) are typically 70-95% efficient
- Intelligent power supplies have on-board micro-controllers (MCUs) that allow the SMPS to be electronically monitored and controlled.

Cooling in the Data Center - Overview

- Typical Rack PC Server consumes 10-20 kiloWatts and requires an additional 10-20 kiloWatts for Power and Cooling

- Key Objective when Cooling in a Data Center - need to focus on "effective air handling"
  - » Optimize the delivery of cool air
  - » Optimize the collection of waste heat
  - » Minimize the interaction (mixing) between cool air and warm air
    - – Mixing warm air and cold air kills your cooling efficiency

Cooling in the Data Center - 2

- Effective air handling has a major impact on energy efficiency
  - » good versus bad practices yield 5:1 difference in effective air distribution

- Objective should be monitoring hot spots and directing cold air to them, rather than shotgun “blowing cold air everywhere”

Cooling - Terms

- Central Chiller – Centralized air-conditioning/chilling that is then fed into the computer room
- CRAC – Computer Room Air Conditioner
  Both cools and pumps the air
- CRAH – Computer Room Air Handler
  Just boosts up the air flow
- RDHX – Read Door Heat Exchanger
  Water cooling integrated into rear door of server rack
- Major Cooling Techniques
  - » In Row
  - » In Rack
  - » In Server

Data Center: In-Row Cooling

- Based upon concept of alternating hot aisles and cold aisles
  - » Objective is to avoid mixing the rack's hot air exhaust with the cool air coming into the racks
  - » Server racks are laid out with alternating rows:
    - – One row has the all the racks air intakes pulling in cold air
    - – The other row has all the racks exhaust air collected

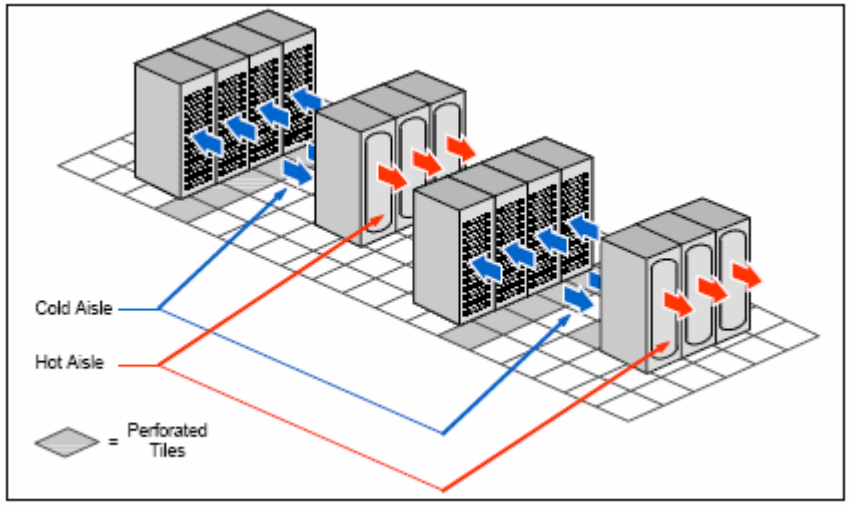


Figure 4-6 Hot and cold aisle configuration

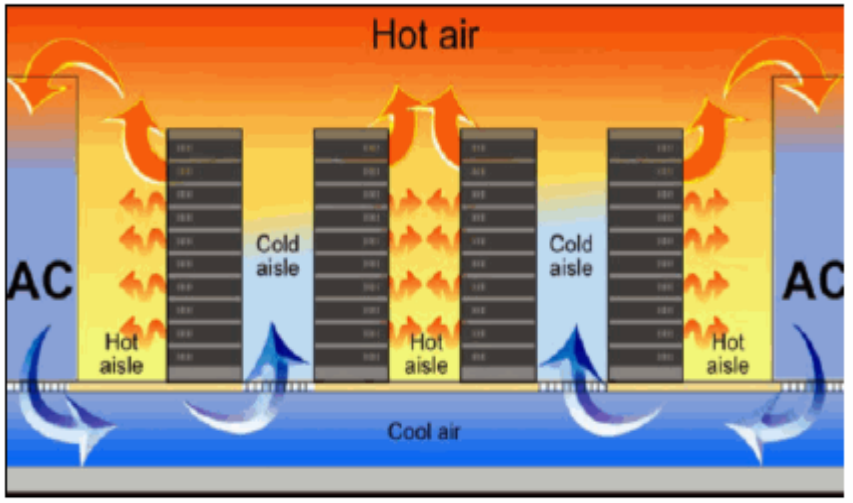


Figure 4-7 Thermal flow of hot and cold air

Source: IBM

- Maximize air management, separate hot versus cold flows
  - » Best cooling requires a high delta-T (difference between hot return air and cold supply air)

Data Center: In-Rack Cooling

- Standard In-Row cooling can handle cooling loads of 5k BTU to 10K BTU for server racks, but things get to be a stretch when cooling 20k BTU server rack loads
  - » Many high end PC Server Blade configurations reach 20-25 K BTUs
- For high end servers, “In-Rack” cooling, to augment the In-Row cooling is required.

- In-Rack cooling can be:
  - » Special ducting and fans within the rack cabinet to boost air flow, OR
  - » External cold water feeding cooling coils inside the rack e.g. Rear Door Heat Exchangers

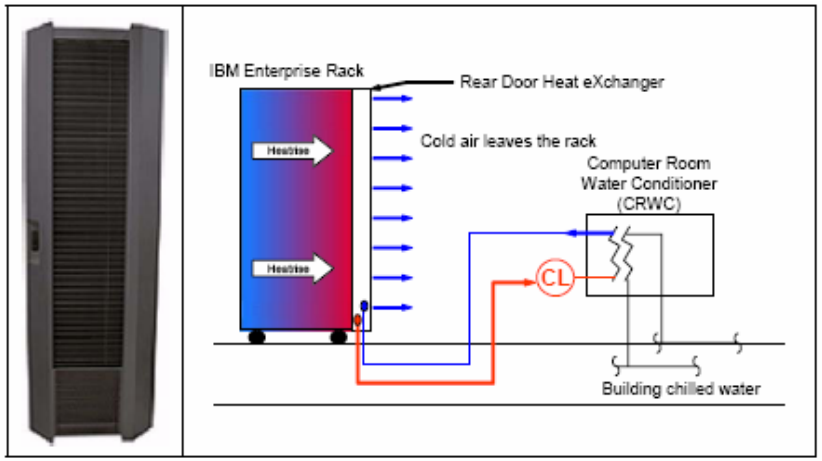


Figure 2-62 Rear Door Heat eXchanger (left) and functional diagram

Source: IBM

Data Center: In-Server Cooling

- Put special cooling inside the server “blade” itself
  - » Water cooling the CPU (water cooled cooler on top of CPU instead of fan)
  - » Spraying the Server CPU components with an inert refrigerant spray that is then collected, cooled, and re-circulated.

Temperature Monitoring

- “You can’t manage what you don’t measure”
- Need to have sensor monitoring of:
  - » Temperature of each aisle, at various point along the aisle
  - » Temperature of each server rack – both at the front/bottom (cold air intake) and at the top/rear (hot air exhaust)
- This data then needs to be fed into your Data Center monitoring applications
- Temperature sensors are becoming dirt cheap, and many have either wired or wireless transmission capability

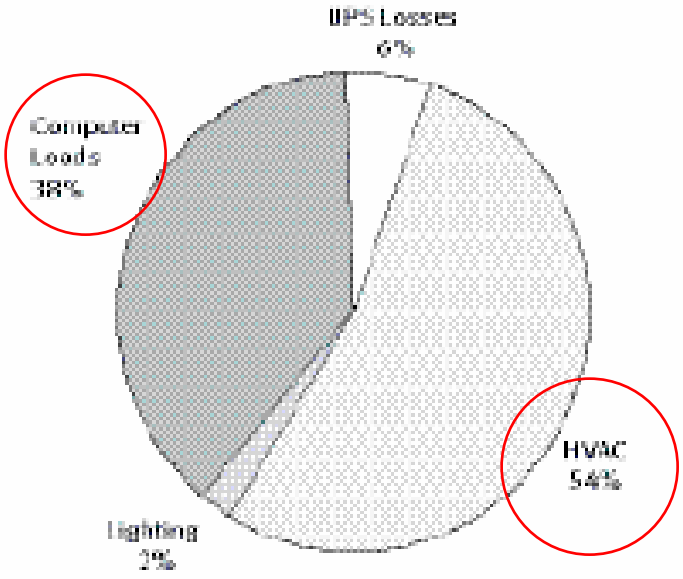
Measuring Overall Efficiency

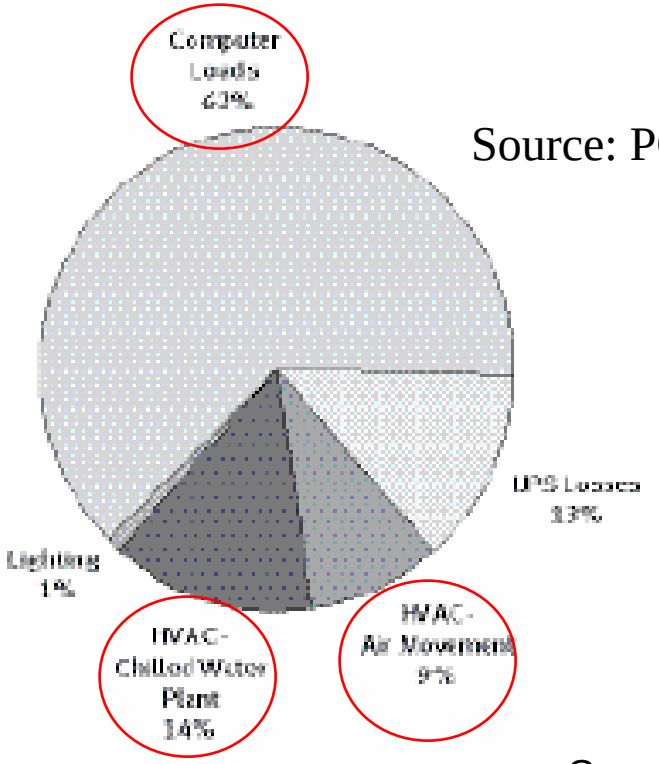
- DCiE – Data Center Infrastructure Efficiency
  = Energy for IT Equipment / Total Energy for Data Center
  Should be > 0.5 0.7-0.85 is Very Good
- PUE – Power Utilization Effectiveness = 1 / DCiE Should be < 2.0

- Example:

Data Center	Computer Load %	HVAC %	Rating
# 1	38 %	54 %	Poor
# 2	63 %	23 %	Excellent

FIGURE 4
ELECTRICITY CONSUMPTION DISTRIBUTION FOR TWO DATA CENTERS





Source: PGE

What can we infer about this Data Center's HVAC setup ?

Data Center Cooling: "Aggressive Techniques"

- The previous techniques focused on the "low hanging fruit" and can easily yield 25 % or better savings.

- The following techniques can wring out another 20-25 % savings, but are more more expensive and often can require major tear-up of any existing Data Center. They should however, be evaluated when planning a new data center.
  - » Replace CRACs with centralized Chiller plant
  - » Free Cooling via utilizing Evaporative Cooling towers at night
  - » Utilize outside air during Fall/Winter/Spring seasons
  - » Self-generation or Co-generation of electricity on-site

PC Server Virtualization

- PC Servers are typically only 10-30 % utilized.
  - » Three PC Server Blade racks, each 30 % utilized, will burn up more energy and HVAC than your mainframe !

- Adding PC “virtualization” software, e.g. VMware, etc can boost utilization of a PC Server to 90 %

- Adding virtualization also dramatically reduces the number of PC Server racks that you have to buy and deploy

Monitoring Standards

"The nice thing about standards, is that there are so many to choose from"

- AEM – IBM defined standard for CPU Server monitoring.
- IPMI – Industry standard for cross system power/temperature/fan control and monitoring. Based on PMbus.
- ACPI – PC Industry standard for software interface to local computers power/temperature/fan controls (monitor and control). Based on PMbus.
- SNMP (Johnson Controls, other HVAC vendors, …)
- Zigbee Energy Profiles (maybe). Something using 802.15.4 will happen
- . . .

Monitoring - Long Term Trends

- Multiple monitoring standards will continue, and probably proliferate
- Virtually all of the products that implement the key monitoring standards have some form of TCP or HTTP hooks in them
- All of them will ultimately converge to send out their metrics as HTTP data, and allow HTTP based commands for control
- Long-term, your SOA-based App Server will become the point of integration:
  - » Data feeds from each of the different vendor's products will be routed to Portals that support the HTTP datastream for that product
  - » Portals will become the main control interface
- As a customer, beat on your vendors to provide HTTP based support. This will at least start forcing them on the path.

Long Term Technology Trends

- Costs for Sensors to monitor and control this stuff has dropped dramatically in the last 5 years
  - » 1 GHz 32-bit MCU for $ 12.00 (and more horsepower than a 308x mainframe)
  - » Wireless chips for $ 2.00 - 4.00 or less
  - » Silicon Temperature Sensors for $ 1.00 - 3.00 or less
  - » Dramatically driving down cost of making a motor/fan/power supply/… intelligent

- Intelligent Monitoring (MCU monitored)
  - » Explosion of Wireless Sensing is Coming
  - » Sensors everywhere you need them

- Intelligent Control (MCU driven)
  - » Intelligent Motors (e.g. VFD motors)
  - » Intelligent Power Supplies (e.g. SMPS)

Questions you need to ask your PC Vendors

- During the PC Server selection process, you want to get your best bang for the energy buck. The following questions will help.

- What kind of CPU power monitoring options are available?
  - » DVS, Power Capping, …
  - » What protocol/monitor tools for monitoring and/or control are used/available ?
- What kind of cooling options are available ?
  - » In-rack cooling, RDHX, …
- What kind of fan/blower motors are used in the rack for cooling ?
  - » Fixed speed (bad) or Variable speed (good)
  - » Is the fan speed internally controlled (thermistor automatically turns it on) or can it be externally controlled (PMbus, …)
  - » Can the fans be externally monitored ?
- What kind of Power Supplies are used in the Rack
  - » What kind of Power Supply fans are used (Fixed vs Variable speed)
  - » Can they be monitored (see above)
  - » How good is the PFC ?
- What kind of temperature monitoring is available in the rack

Questions you need to ask your HVAC vendors

- Does their equipment use “Variable Speed Drives/Motors” ?
  - » What is the efficiency ratings of their motors ?
- What level of control over the HVAC equipment is available
  - » Monitoring Fans, Pumps, Cold Temperatures, Hot Temperatures, Intake air flow speeds, Exhaust air flow speeds, …
  - » Controlling Fans and Pumps (speed up or down)
- What monitoring software or standards do they support ?

Summary

- There is no single magic bullet
- It is a series of decisions, each affecting different areas
  - » Mainframe Server Power/Cooling
  - » PC Server type (Blade, …) and associated Power/Cooling
  - » HVAC motors and efficiency (VFDs)
  - » Power Supply types and efficiency
  - » Equipment placement (Hot/cold aisles)
  - » HVAC cooling air and heat venting layout
  - » Sensor Placement (Temperature, …)
  - » Monitoring Software
- Is an iterative process

References

- **IBM Active Energy Manager Web site**
  www.ibm.com/systems/management/director/plugins/actengmgr/
- "Technical Overview of Power Management",
  http://www.energystar.gov/index.cfm?c=power_mgt.pr_power_management
- "Guidelines for Energy Efficient Data Centers." The Green Grid. 2007
  http://www.thegreengrid.org/
- DOE – DC Pro Tool Software – Free Software to Assist Data Ctr Planning
  http://www1.eere.energy.gov/industry/datacenters/software.html
- DOE – Creating Energy Efficient Data Centers
  http://www1.eere.energy.gov/industry/saveenergynow/pdfs/doe_data_centers_presentation.pdf
- High Performance Data Centers: A Design Guidelines Source Book
  http://hightech.lbl.gov/documents/DATA_CENTERS/06_DataCenters-PGE.pdf
- ASHRAE Data Center technical guidebooks http://tc99.ashraetcs.org/
- IBM Cool Blue Tutorials, http://www-03.ibm.com/systems/x/advantages/energy/resources.html

EOJ