Data Center Power and Cooling Issues and Future Designs

Bob Pereira / Ken Baker
Advance Technology Manager
ISS Rack & Power

HP WORLD 2003
Solutions and Technology Conference & Expo

Why are we talking today?

- Due to server and storage densification, yesterday’s data center infrastructure and design practices cannot keep pace with the growing demands for power and cooling
- The high density architectures and products we talked about for the last two years are here today
- Successful installations and implementations require careful, coordinated planning
- We, the new hp, need to be looking beyond our servers and storage boxes to become more knowledgeable about the impact that our high density solutions are causing in order to provide complete solutions for our customers

on the customer front

- What are the right questions to be asking?
  - Have I engaged facility engineering in planning for growth and hardware deployments?
  - Have I planned for growth in power and heat density?
  - Is my data center configured to provide sufficient airflow to the new servers & storage products?
- You own and manage the environment !

successful high density deployments

Customer IT professionals

+ Customer Facilities Engineers

+ *hp*

When you involve the facilities engineers early and often in your data center expansion and build outs, you minimize risk and maximize success.

power trends

historical server comparison

The original SystemPro

- First ProLiant PC Server
- 33MHz 386 processor
- 8 Mb RAM, 210 Mb HDD
- $36,000
- 400W

BL 10e blade

- First ProLiant blade server
- 900 MHz processor
- 512 Mb RAM, 40 Gb HDD
- $1,800
- 20-25W

Power consumption

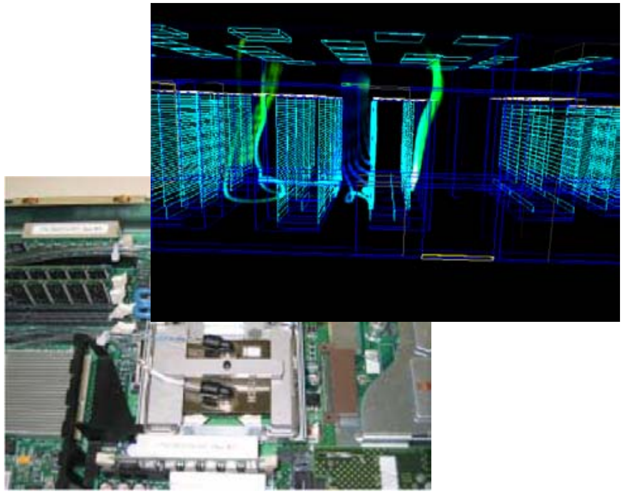
- Power supply *capacities* have not grown substantially
  - 200 -1200 watts
- Power supply *physical* size is decreasing
- Power supply efficiency is improving
  - blade driven

State of the industry

- Two years ago
  - Average servers per rack = 4 - 6
  - Average watts per rack = 1500 - 3000
- Today
  - Average servers per rack = 8 – 12
  - Not likely to change for ala carte servers
  - Average watts per rack = 5000 - 6000
  - Blades can drive 10 - 18KW implementations

Tomorrow’s datacenter

- Intel reports processor density above 140 watts
- Server density will rise to 15-18 per rack to offset new data center construction costs
  - Majority of customers are not constructing new datacenters; want more life out of existing space
- Commercialization of augmented cooling technologies must be considered
  - Water cooling
  - Heat pipes
  - The “Smart” Datacenter



Rack server trends

HP WORLD 2003 Solutions and Technology Conference & Expo

Generation over generation density growth rates

DL360 G2 250 W 1.2A @ 208V 848 BTU/hr	**DL360 G3** 325 W 1.6 A @ 208V 1108 BTU/hr	**30%** 2p, 4gb, 2hdd, 1pci	1U
DL380G2 362 W 1.8 A @ 208V 1233 BTU/hr	**DL380 G3** 425 W 2.1 A @ 208V 1450 BTU/hr	**17%** 2p, 4gb, 6hdd, 2pci	2U
DL580 G1 471 W 2.3 A @ 208V 1608 BTU/hr	**DL580 G2** 812 W 4.0 A @ 208V 2770 BTU/hr	**72%** 4p, 8gb, 4hdd, 3pci	4U

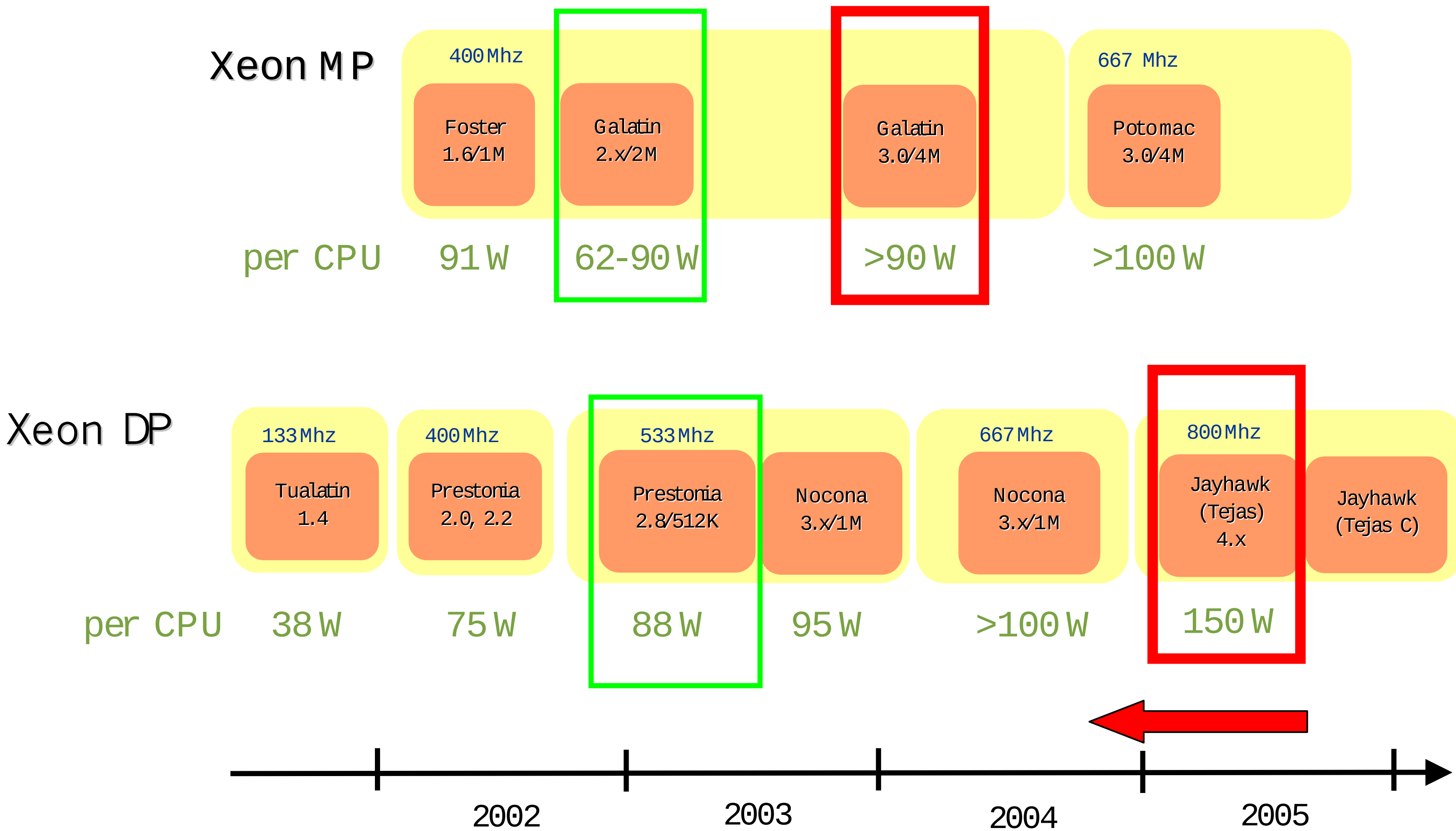
Why is this happening?

Intel Ships 1 Billionth Processor as Company Reaches 25th Anniversary of Intel Architecture; Second Billion Could Come as Soon as 2007

Based on combined desktop, laptop and server shipments, Mercury Research calculates that Intel reached this milestone in April, roughly 25 years after the debut of the first 8086 microprocessor on June 8, 1978

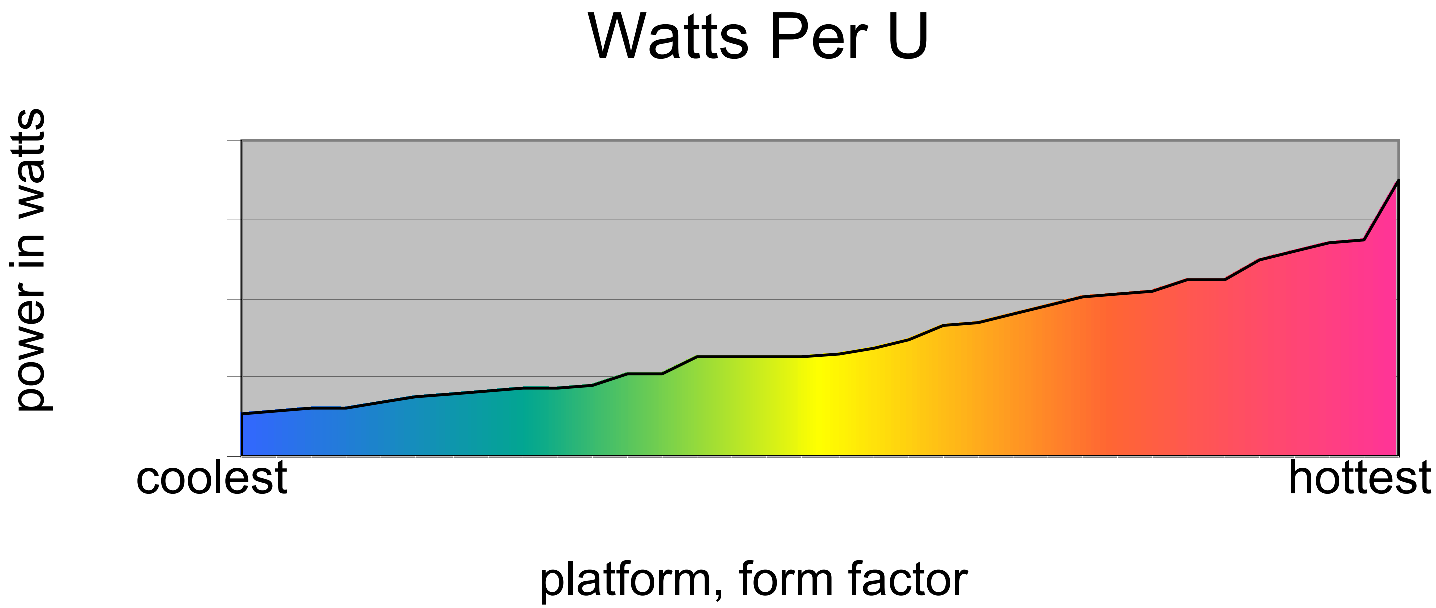
Mercury Research calculates that the next billion X86 CPUs could ship far faster than the first billion processors and could **come as early as 2007**.

Processor trends

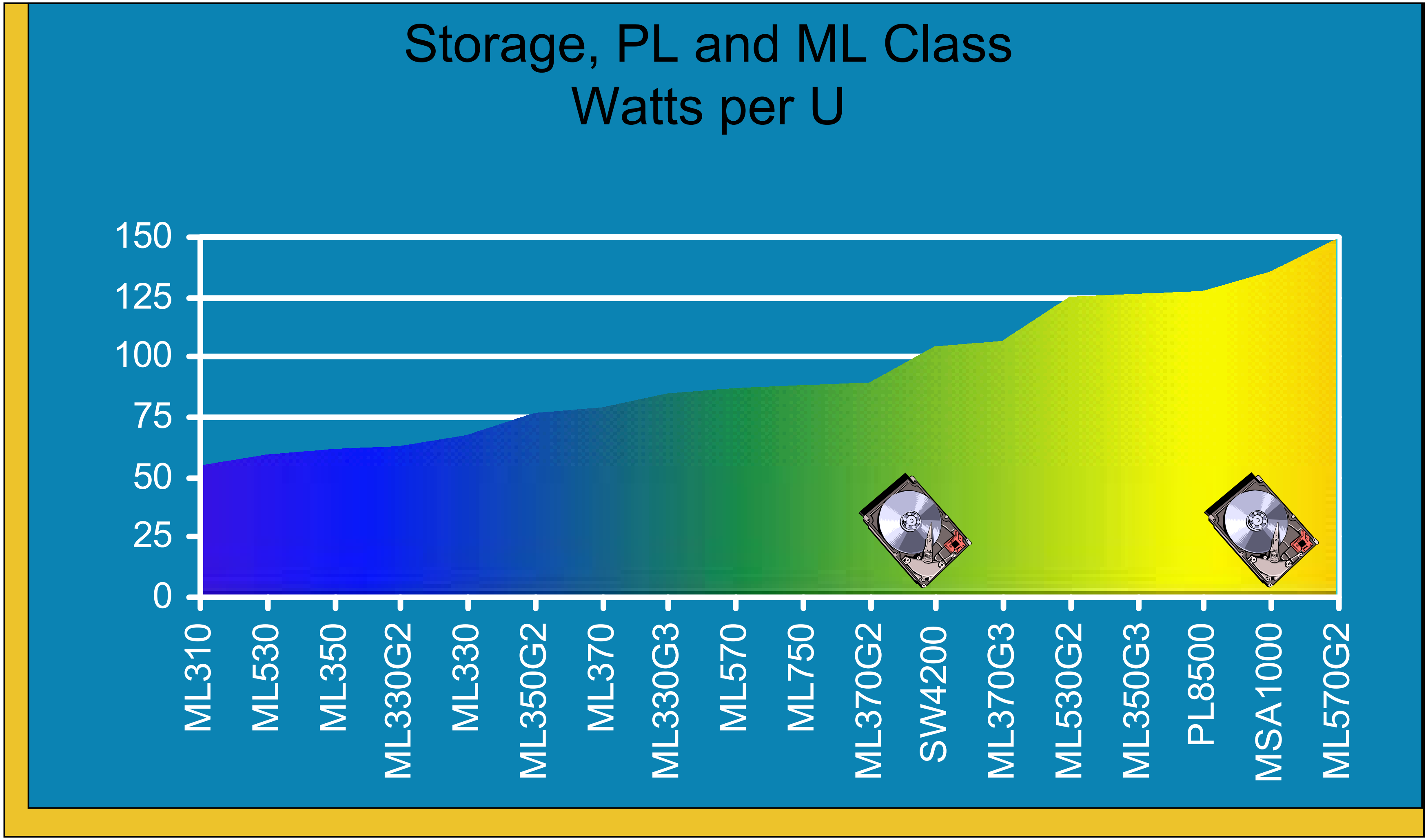


Combining all of the trends

Overall power density per Rack Unit "U" is becoming the key metric, capturing all of the key contributors to the densification.

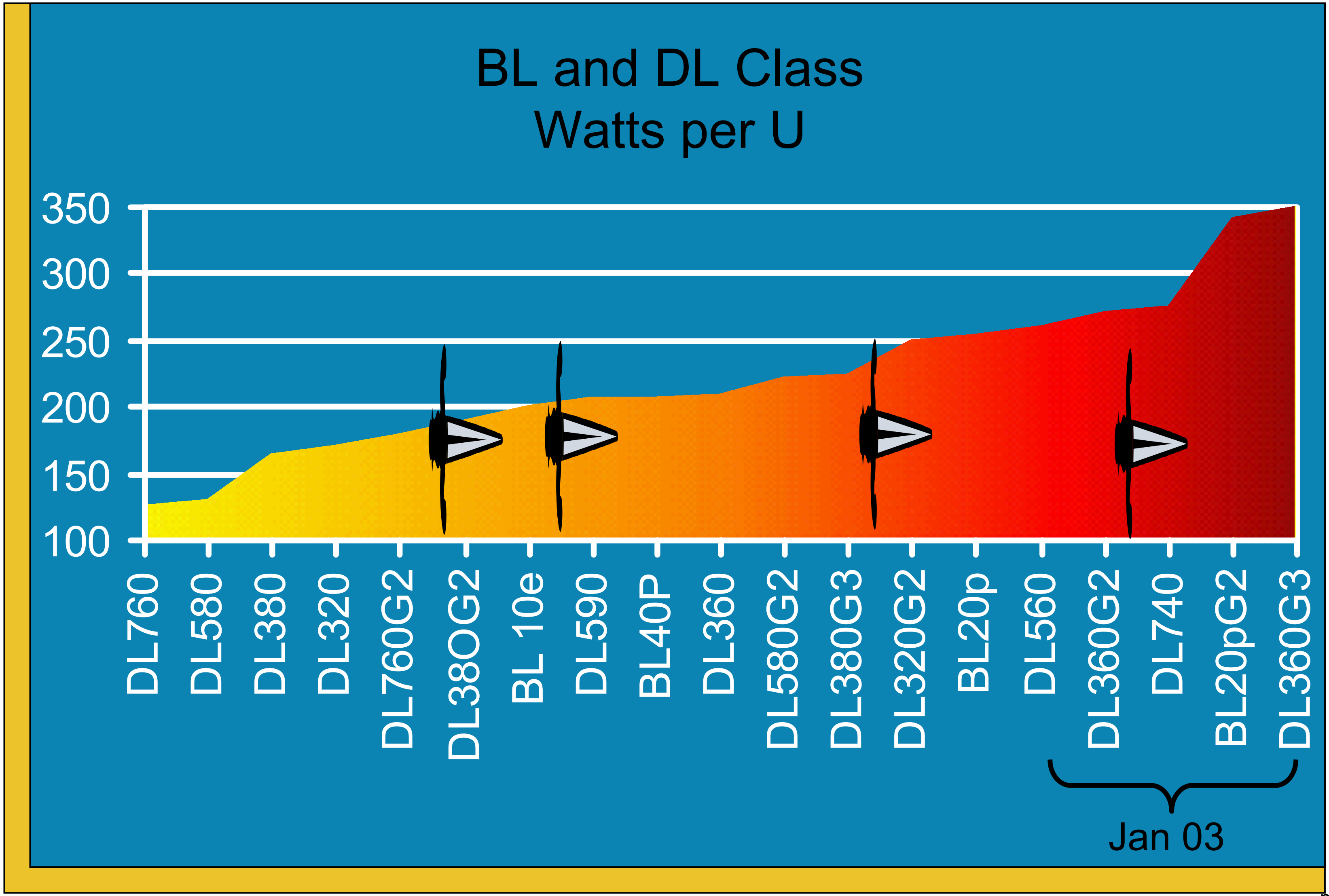


Server Power Density Per Rack Unit



Server power density per rack unit

HP WORLD 2003 Solutions and Technology Conference & Expo



Cooling trends

- 2 years ago
  – Average BTU load was 10-15K BTU’s(1.2 tons)
  – Average U size was 5-7U
- Today
  – Peak BTU load is approaching 28K BTU’s(2.5 tons)
  – Average U size is moving down to 3U
- 2005-2007
  – Heat densities of +50,000 BTU’s
  – Advancements in cooling the room will have to be made

Summary

- Processor
  - x86 architecture does not offer opportunities to reduce power consumption
  - Expect 15%-25% generation over generation power increases
- Servers
  - Physical density in ala carte servers to slow
  - Power consumption to rise with processors
  - Augmented cooling technologies will emerge
- Datacenters
  - Customers are demanding longer life cycles for facilities than they have today
- HP is committed to solving the really hard problems

Delivering Power for Hyper Dense Data Centers

Delivering More Power

- Overcoming the “desktop mentality”
  - Standard 120V, 10A PC power cord is all you need
  - Open power receptacles indicates available unlimited power

- Powering the infrastructure.
  - Must Break the 15A and 20A branch barrier
  - Options include multiple 30A and 3 Phase branches

- More power requires larger hardware

beyond the desktop

- NEMA L6-30P, 208V, 30A, 3w

- NEMA L15-30P 208V, 30A 3ph, 4w

larger copper cabling

- SJ/SO 3 x 10Awg
- SJ/SO 4 x 10Awg

Connector standards, much easier than it looks

- NEMA: North American Electric Manufacturers Assoc.
- L: twist locking connection (absence = standard push in)
- 5 or 6 5 = rated up to 125V
  6 = rated up to 250V
- 15,20,30 rated current in amps
- P/R P = plug (male)
  R = receptacle (female)

- IEC: International Electrotechnical Commission
- 320 General purpose household connectors
- C13/14 style sheets for 10A connectors
  odd = females even = males

  C19/20 style sheets for 16A connectors

- 309 industrial grade connectors pin and sleeve connectors

Beyond the desktop

- IEC 309 Pin/Sleeve Plugs
  - 16A 1Ø
  - 32A 1Ø
  - 32A 3Ø

- larger copper cabling
  - H05V - 3x4.0mm
  - H05V – 5x4.0mm

Power density issues

- On the electrical side, why is deploying full racks of servers a problem?
  - total rack load not the problem
- Issue lies with how power is distributed
  - line cord/distribution outlet relationship and restrictions

Commercial branch circuits in North America

Power limitations associated with each common branch circuit type

Branch voltage	branch circuit size	max load 80% per NEC	maximum available power
120V	20A	16A	1920 VA
120V	30A	24A	2880 VA
208V	20A	16A	3328 VA
208V	30A	24A	4992 VA
208V	20A 3ø	27A	5760 VA
208V	30A 3ø	42A	8650 VA

Today’s power densities dictate leaving the 120V infrastructure, using multiple 208V 30A feeds and looking forward to using 3 phase power to meet the density demands.

What Key Metric is Changing Here?

Distribution

Distribution panel efficiency

- 150kVA, 84 pole panel locations
- 208V distribution requires 2 poles breakers, 42 2-pole
- Appears to be plenty of overall power!

Power requirements




ProLiant Server	**DL 360G2**	**DL 380G2**	**DL 580G2**	**BL 10e**
Enclosures per 42u Rack	42	21	10	14
No. of CPUs	2P × 42 = 84P	2P × 21 = 42P	4P × 10 = 40P	14 × 20p = 280P
Power	42 × 270W = **11.34kW**	21 × 400W = **8.4 kW**	10 × 790W = **7.9** kW	14 × 670W = **9.3** kW
Heat Load**	38,670 BTU/hr	28,644 BTU/hr	26940 BTU/hr	31985 BTU/hr

Commercial branch circuits in North America

HP WORLD 2003
Solutions and Technology Conference & Expo

branch voltage	branch circuit size	max load 80% per NEC	maximum available power
120V	20A	16A	1920 VA
120V	30A	24A	2880 VA
208V	20A	16A	3328 VA
208V	30A	24A	4992 VA
208V	20A 3ø	27A	5760 VA
208V	30A 3ø	42A	8650 VA

Example of distribution limitations

- 24A PDU 208V single/bi-phase PDU limited to 4992 VA
- Subject cabinet of 21 DL380G2 would require 8560 VA total power or two 24A PDUs
- Redundancy is required, that doubles the PDU count to 4
- 4 breakers (8 poles) per cabinet (4 PDUs)
- 10 cabinets per 84 pole panel
- 210 servers provided with power

Increasing power capacity

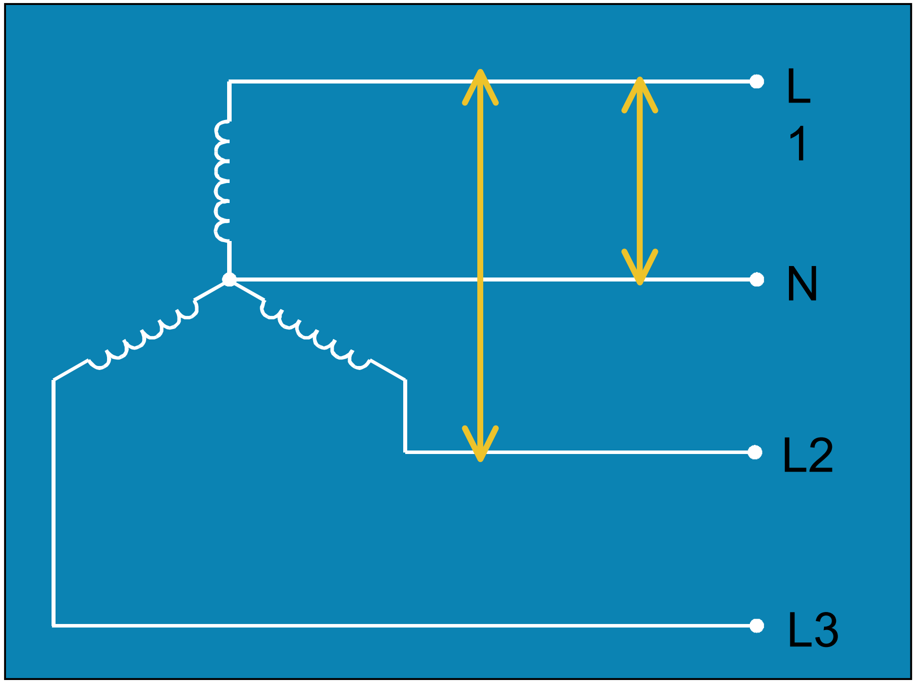
- Why not install larger PDU?
  - any larger 1/2 phase PDU eliminates cost effective pluggable solutions (hardwiring)
  - no cost-effective standard connectors > 30A
- Only potential solution
  - move to 3 phase power distribution
  - cost effective pluggable solution > NEMA L15-30P

Some 3 phase power facts

- Most of the electrical power in the world is 3 phase
- 3 phase power is typically 150% more efficient than single phase
- Power delivered to the load is the same at any instance
- 3 phase conductors can be 75% the size of 1 phase conductors for the same power output

What is 3-phase power?

- All power is generated and distributed using a 3Ø system
- LV loads are connected between 1 phase winding and the Neutral or center tap connection
- HV loads are connected phase to phase
- 3Ø loads are connected to all 3 Phases simultaneously



- Phase – Phase Voltage 208V
- Phase – Neutral Voltage 120 V

30A 3 phase PDU power

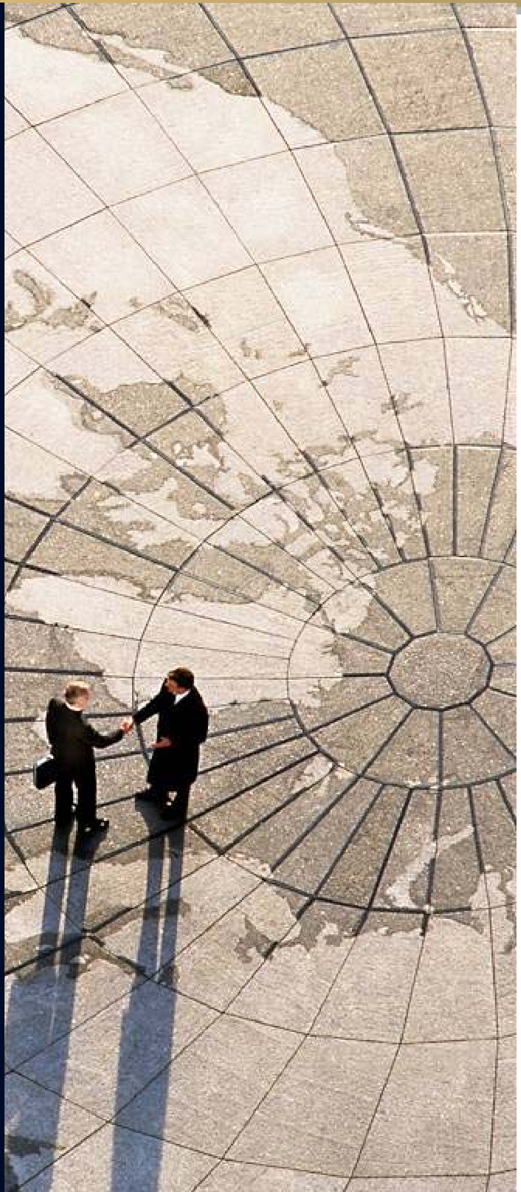
- most efficient way to distribute power
- overall available power in a 30A circuit rises from 4995 VA to 8650VA
- total available current rises from 24 A to 42 A
- fewer overall panel positions used per rack of high density loads

Panel utilization for racks of DL380G2

PDU Model	24A 1-Phase PDU 120V	24A 2 Phase PDU, 208V	24A 3-Phase, 208V
VA per PDU	2880	4992 VA	8650 VA
PDU per Rack of 21 DL380G2 servers	8 PDU	4 PDU	2 PDU
Breaker Poles required per rack	8 poles (1 poles per PDU)	8 poles (2 poles per PDU)	6 poles (3 poles per PDU)
Racks per 84 pole distribution panel	10 89460VA	10 86730VA	14 121,422VA

Summary/Power issues

- server density is topping out conventional power infrastructure and methods
- we have exceeded the ability of a plug and play solution to be deployed in a high density manner due to power distribution limitations without developing new approaches
- implementing 3 phase power in the rack enclosure is the future

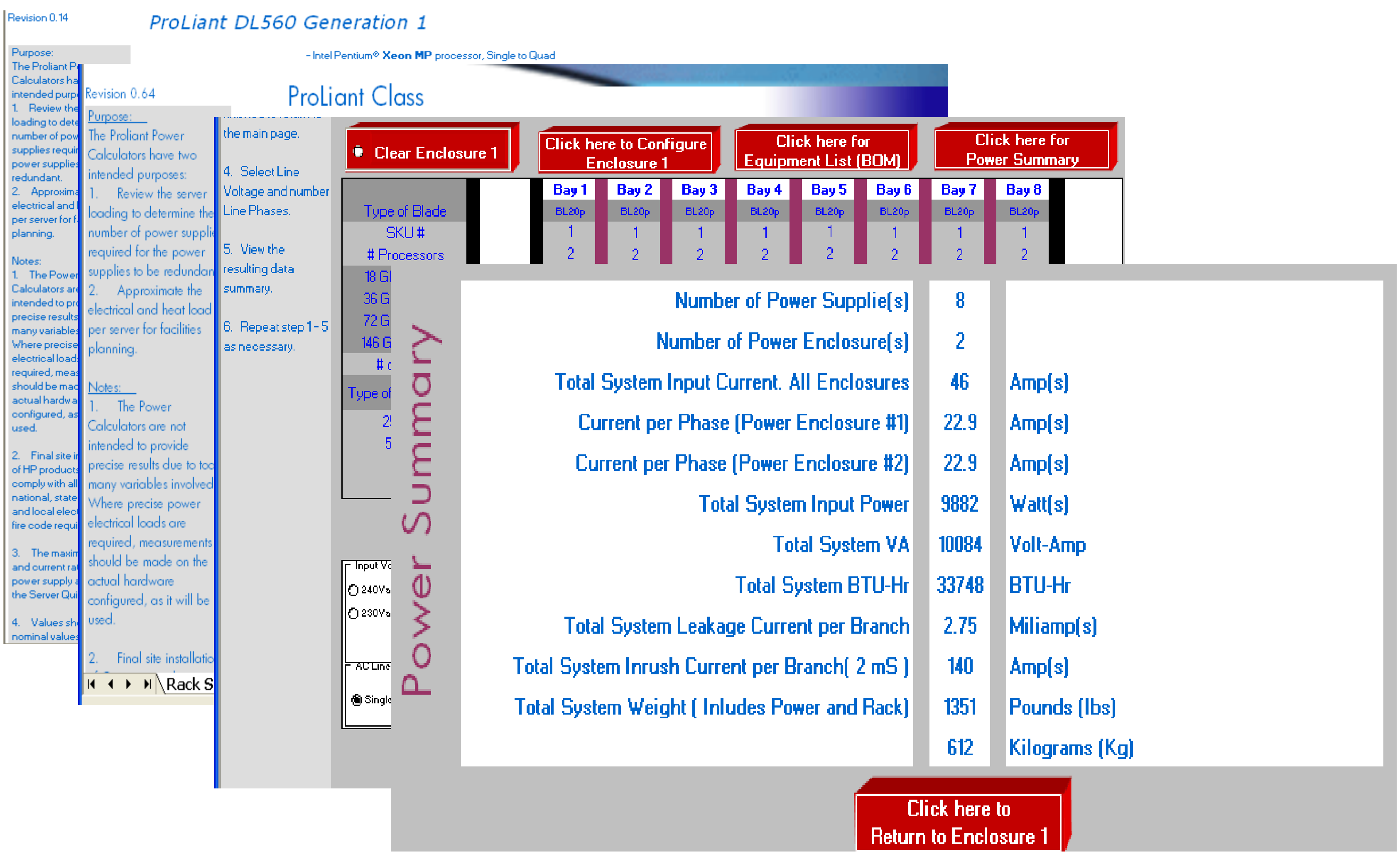


Power Planning

Power planning methods

- Use “name-plate” ratings
  - Worked in yesterdays environment
  - Costly method resulting in wasted infrastructure $
- Use actual power measurements
  - Most accurate approach
  - Numbers are difficult to generate and collect
- Use “ProLiant Power Calculators”
  - Best practice for advanced planning
  - Numbers are more realistic
- Factoring for future growth
  - Rate of change in today’s market continues at 25%-30%

ISS power sizing calculators

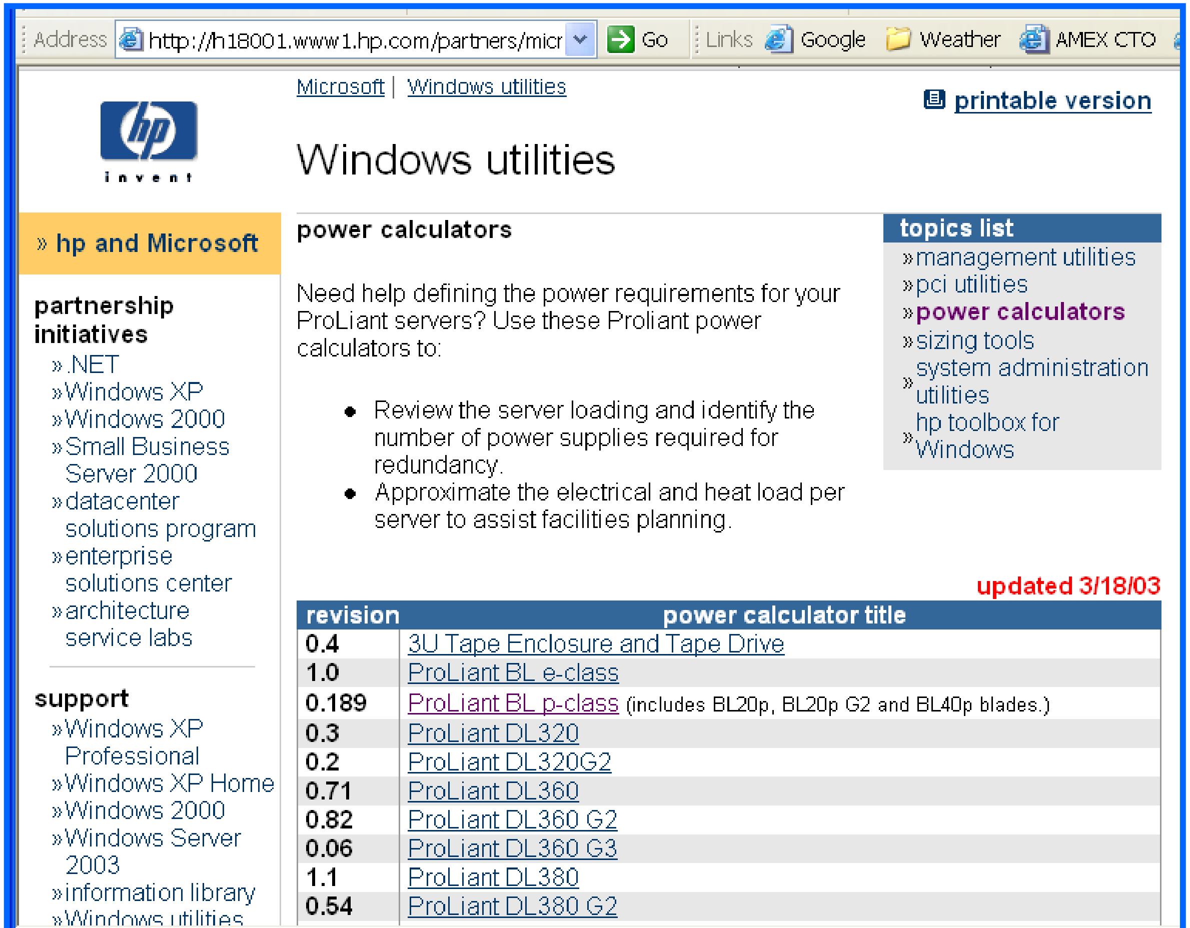


About power calculators

- Based on actual system measurements
- Taken on systems running NT and exercise utilities
- All major system components (CPU, memory and drives) are exercised at 100% duty cycle
- Power results may be higher than your actual configuration leaving you extra headroom
- Calculators can be found on each servers Active Answers page under configuration tools
- New calculator public website

Calculator public website

http://h18001.www1.hp.com/partners/microsoft/utilities/power.html



Address http://h18001.www1.hp.com/partners/micr Go Links Google Weather AMEX CTO

Microsoft | Windows utilities

printable version

Windows utilities

» hp and Microsoft

partnership initiatives
- » .NET
- » Windows XP
- » Windows 2000
- » Small Business Server 2000
- » datacenter solutions program
- » enterprise solutions center
- » architecture service labs

support
- » Windows XP Professional
- » Windows XP Home
- » Windows 2000
- » Windows Server 2003
- » information library
- » Windows utilities

power calculators

Need help defining the power requirements for your ProLiant servers? Use these Proliant power calculators to:

- Review the server loading and identify the number of power supplies required for redundancy.
- Approximate the electrical and heat load per server to assist facilities planning.

topics list
- » management utilities
- » pci utilities
- » power calculators
- » sizing tools
- » system administration utilities
- » hp toolbox for Windows

updated 3/18/03

revision	power calculator title
0.4	3U Tape Enclosure and Tape Drive
1.0	ProLiant BL e-class
0.189	ProLiant BL p-class (includes BL20p, BL20p G2 and BL40p blades.)
0.3	ProLiant DL320
0.2	ProLiant DL320G2
0.71	ProLiant DL360
0.82	ProLiant DL360 G2
0.06	ProLiant DL360 G3
1.1	ProLiant DL380
0.54	ProLiant DL380 G2

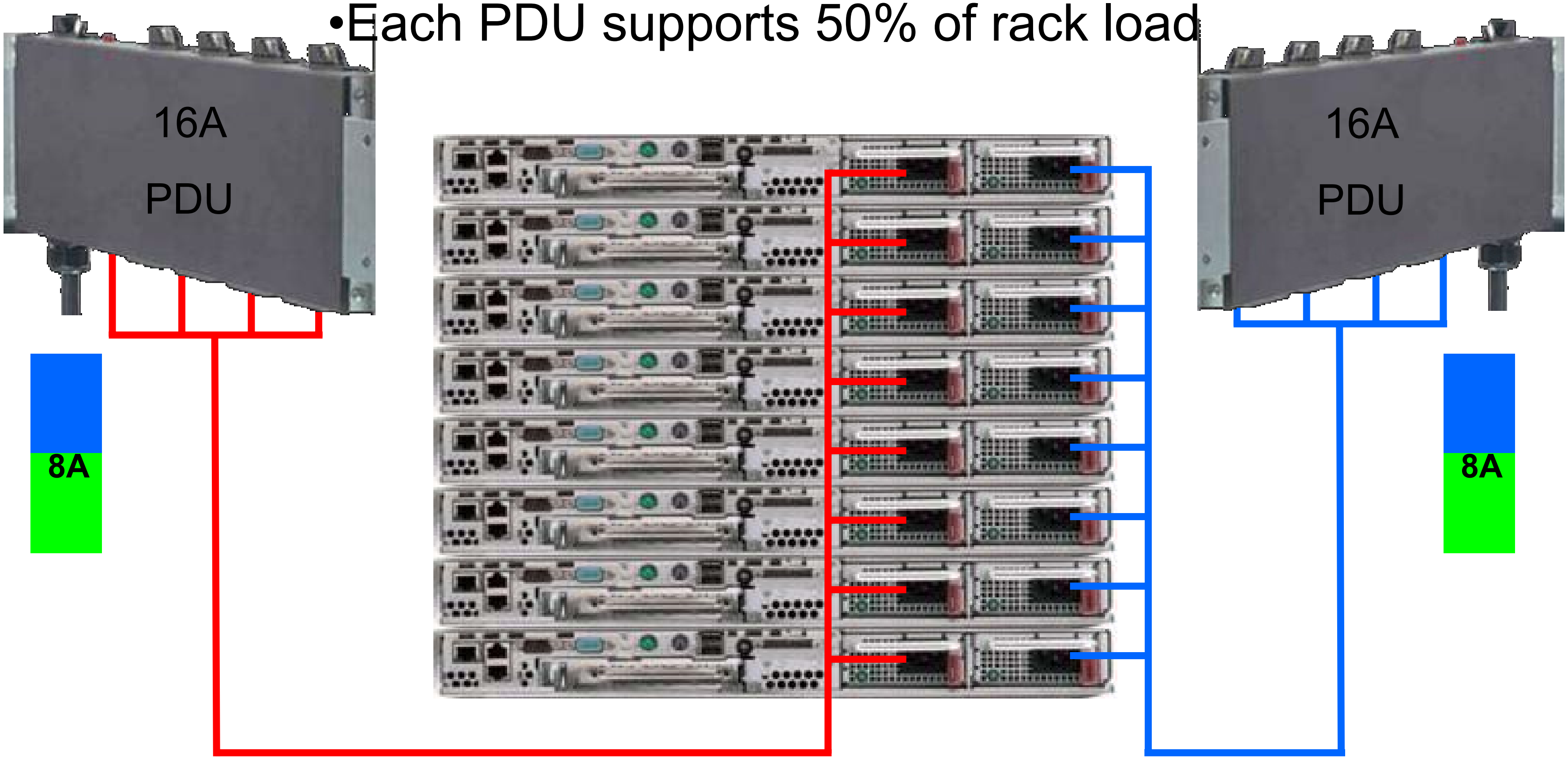
Redundant power planning

- What does 1+1, N+1, N+N mean
- What effect does redundancy power planning
- Do redundant power supplies load share
- How do I size and select PDUs for redundant power schemes
- How much load is safe on a PDU

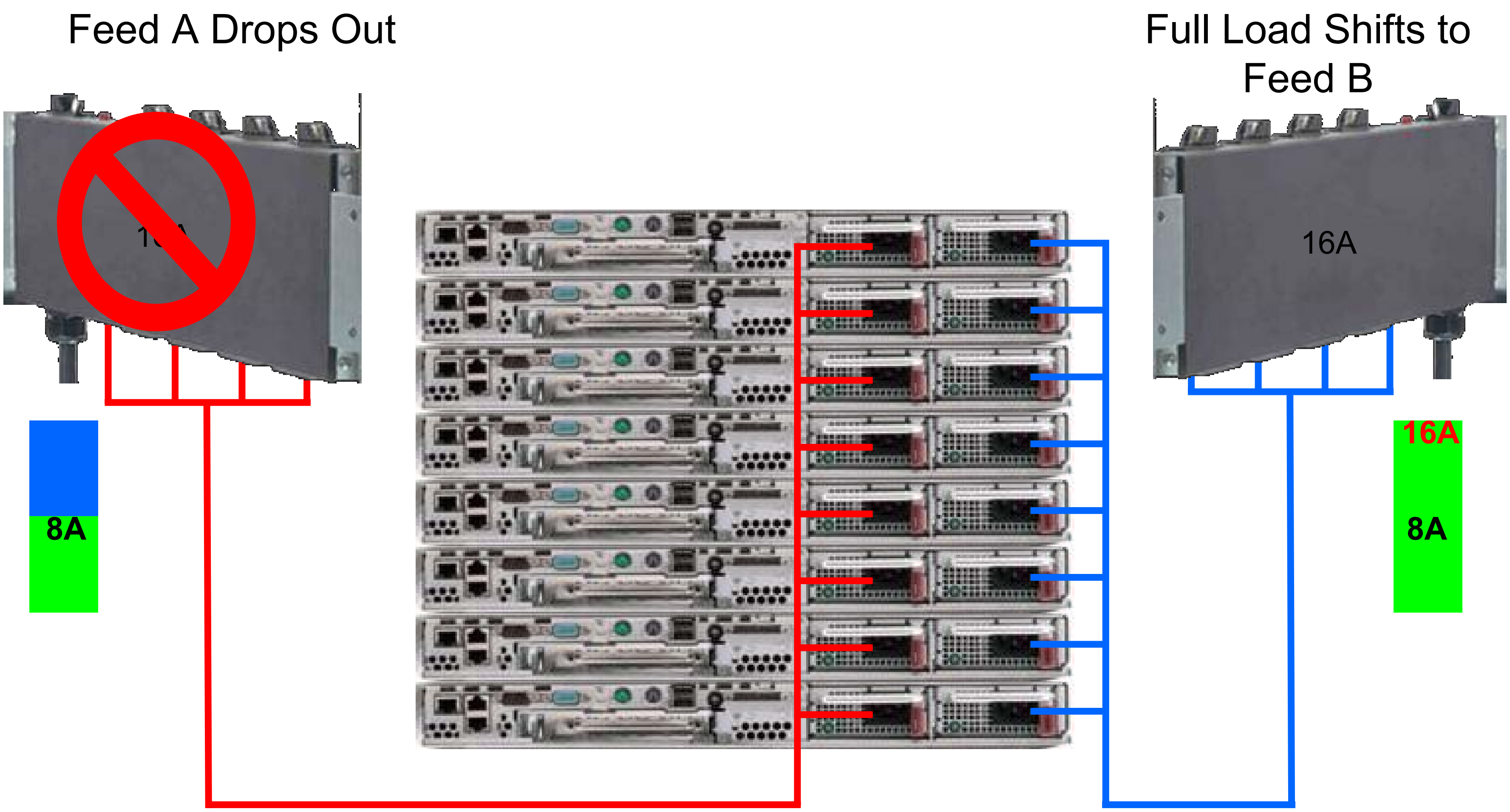
Normal 1+1 Redundant Power Operation

•Each P/S supports 50% of server load

•Each PDU supports 50% of rack load

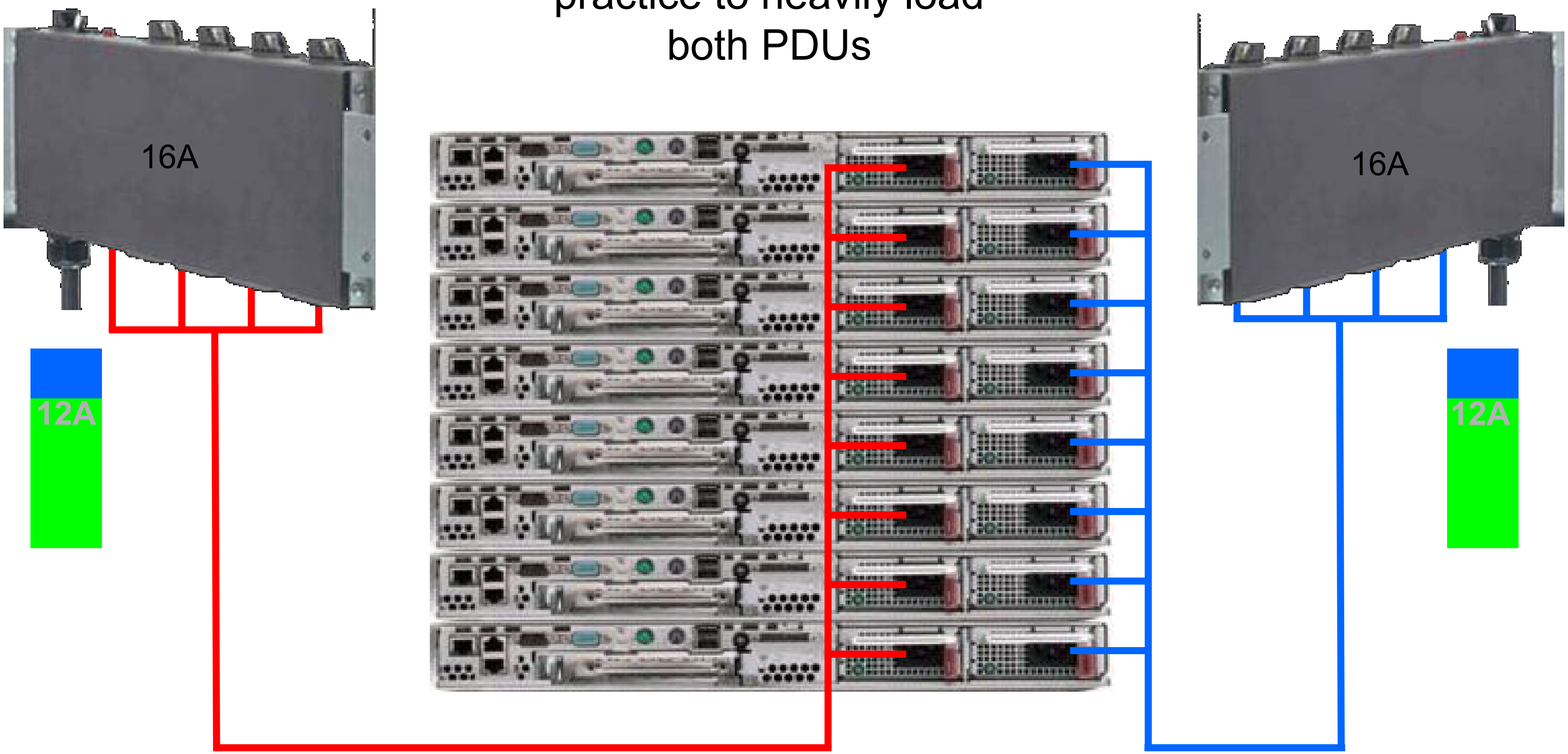


Redundant Feed Failure

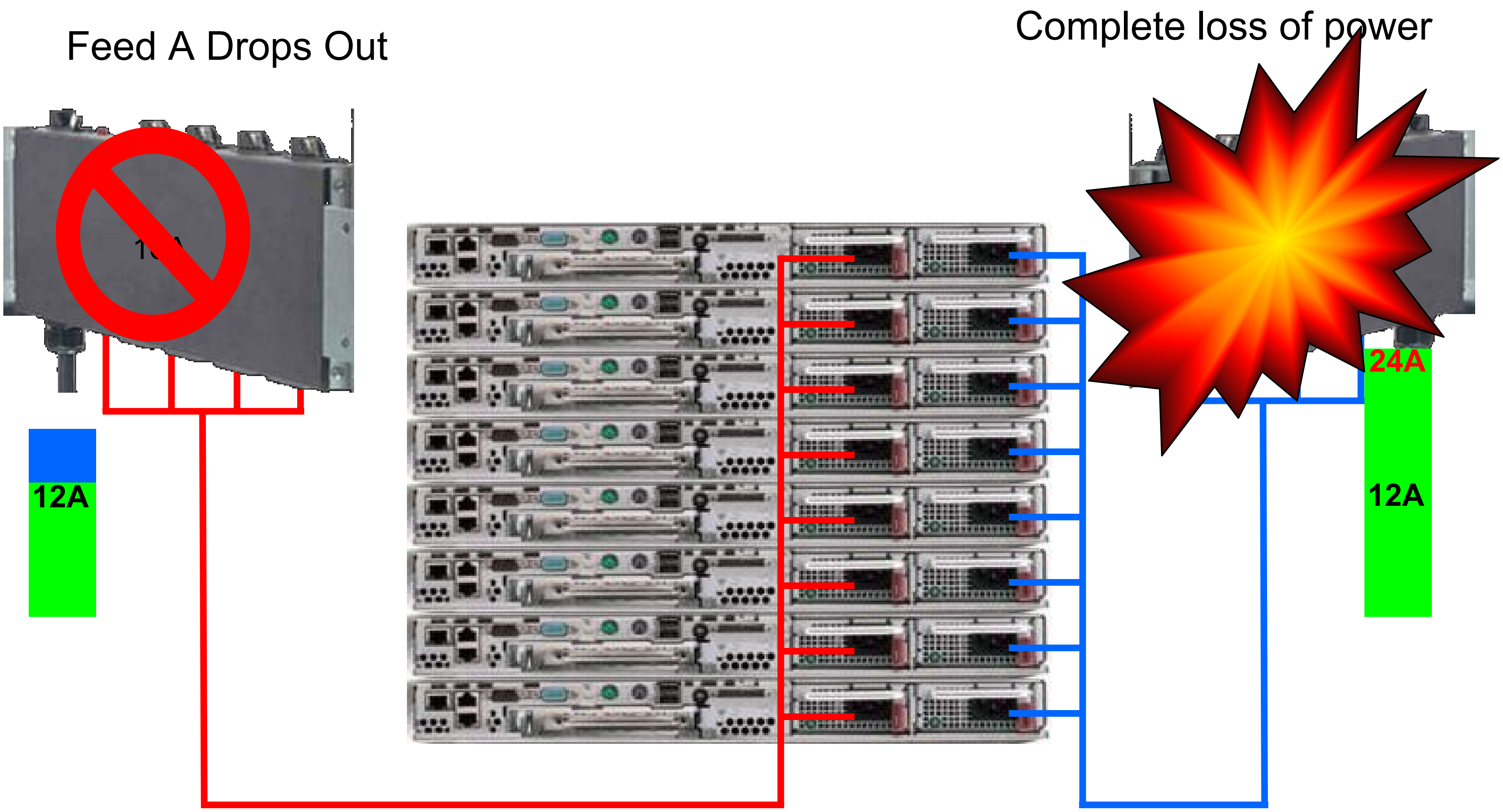


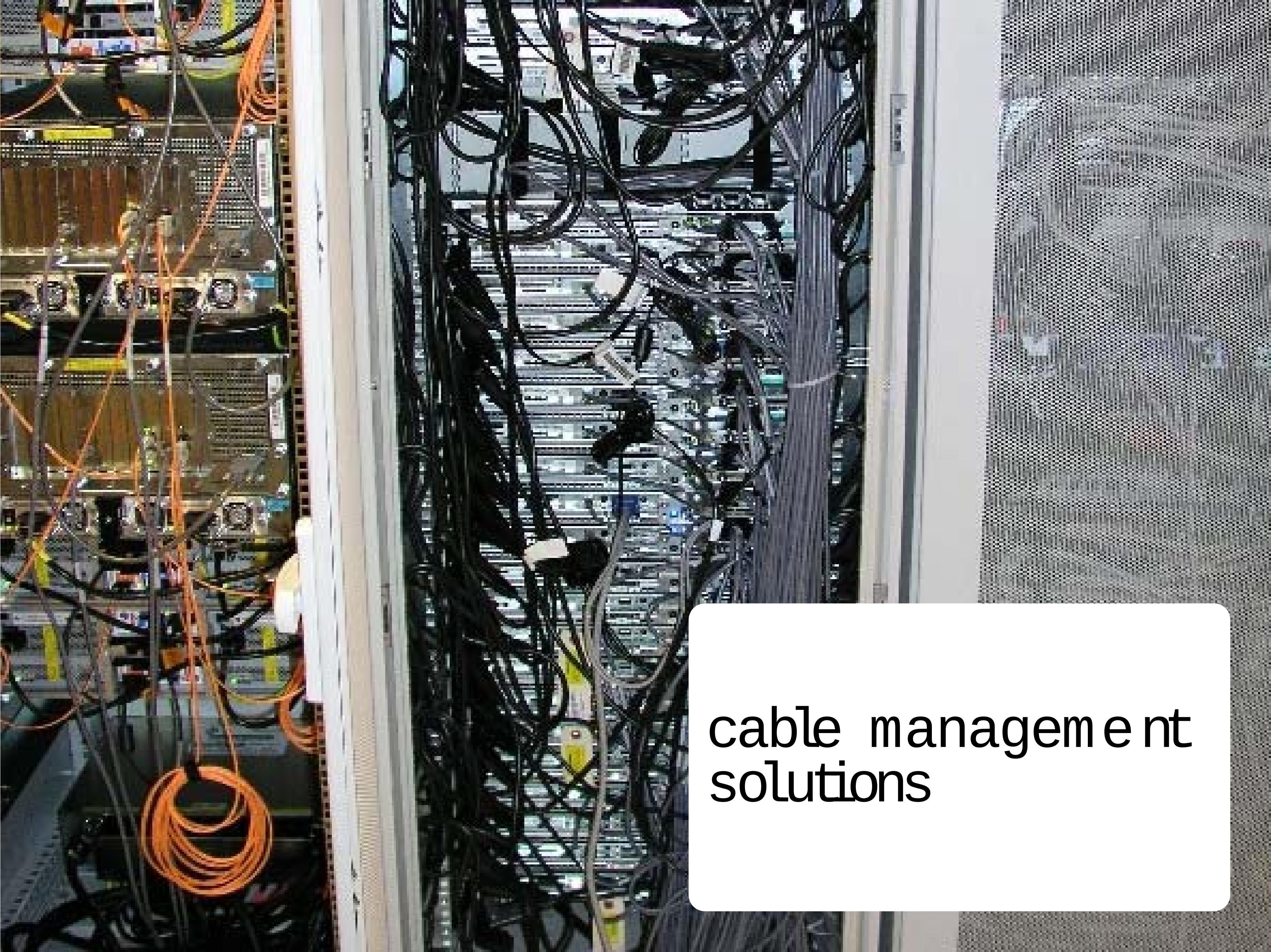
Disaster waiting to happen

Not an uncommon practice to heavily load both PDUs



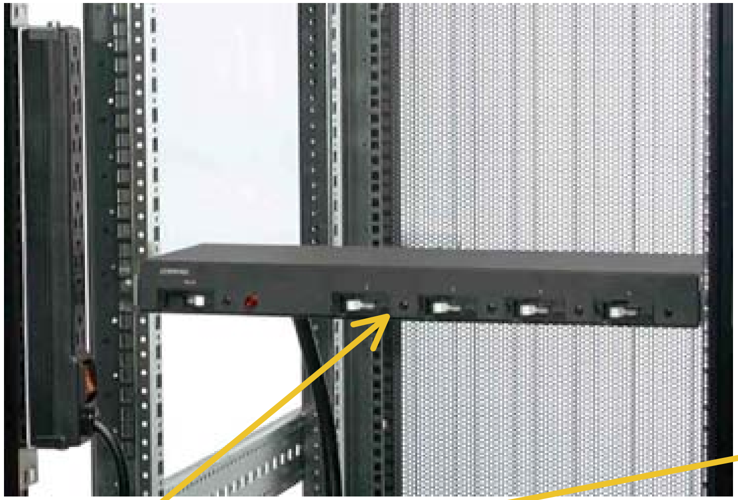
Disaster strikes

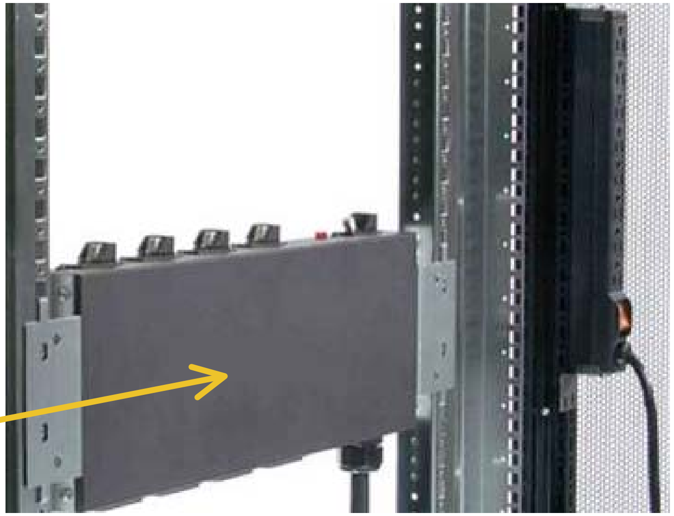




cable management solutions

Power cable reduction - Modular PDU



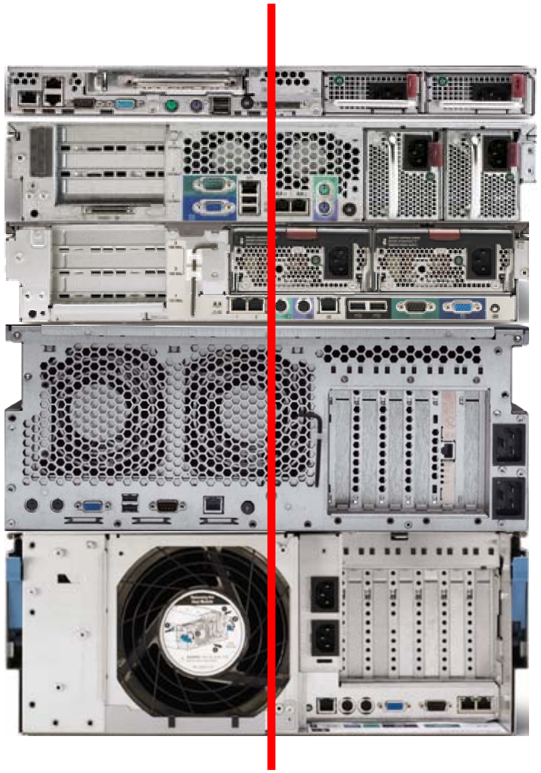


- 1U / 0U mounting options for flexibility
- extension bars provide **better outlet accessibility**
- extension bars allow the use of **shorter power cables**
- 142257-007 option is a must, providing fifteen 1.3m IEC-IEC power cables

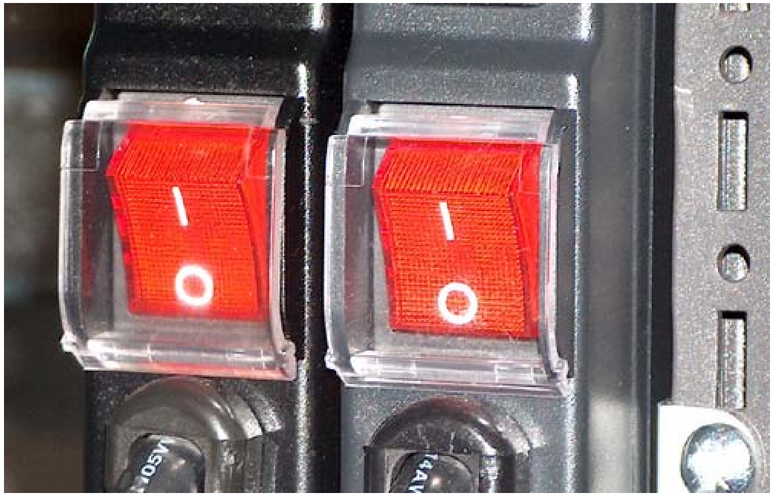
Platform Developments

- New platform power supply locations are all right justified
- This drives power infrastructure to one side of rack
- This cable management development drove the demand for new Extension Bar mounting brackets

Current DL 2P, 4P, 8P Platforms



Modular PDU enhancements



New mounting brackets

- mounts 2 bars side by side
- same side rack mounting
- symmetrical to allow left or right side mounting
- up to 10 mounted on one side

New switch covers

- **No Cost Spares Number 337116-001**

HP Factory Integration Services

Customer rack without HP Staging & Integration

Customer rack with HP Staging & Integration

Best cable management service in the industry!
We do it right!

Cooling Trends

- 2 years ago
  - average BTU load was 10-15K BTU's
  - average U size was 5-7U
- Today
  - average BTU load is approaching 28K BTU's
  - average U size is moving down to 3U

cooling trends

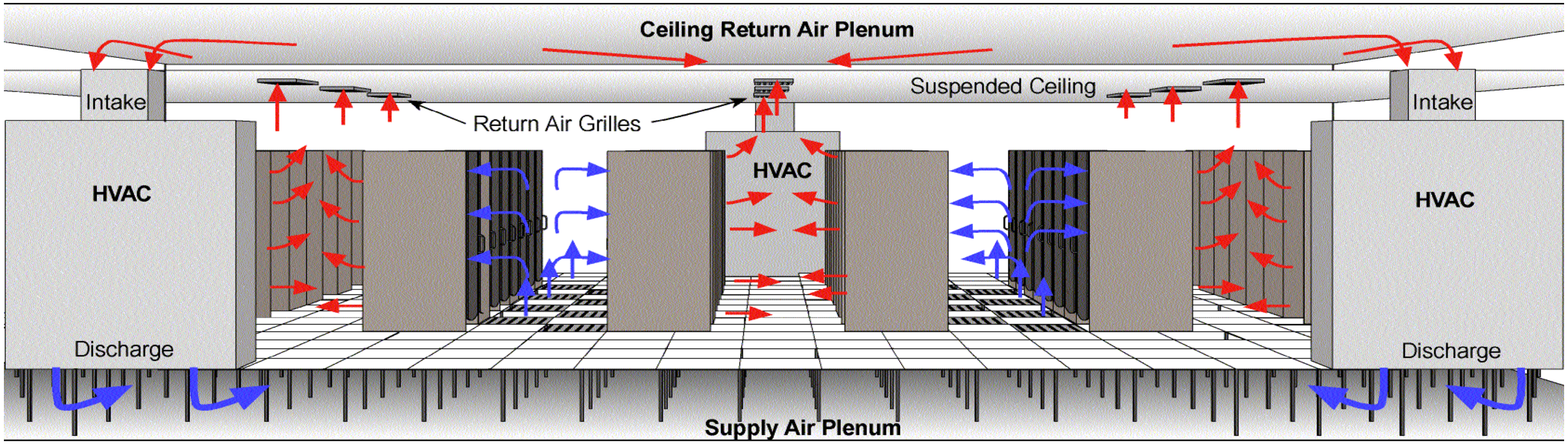
- as power consumption grows, so does the thermal demands
- cooling needs must be expressed in tons of AC today
- largest issue is not finding the cooling media, but finding a way to get it where it belongs
- driving new technologies to deliver cooled media

Cooling Trends (cont.)

- as power consumption grows, so does the thermal demands
- cooling needs must be expressed in tons of AC today
- largest issue is not finding the cooling media, but finding a way to get it where it belongs
- driving new technologies to deliver cooled media

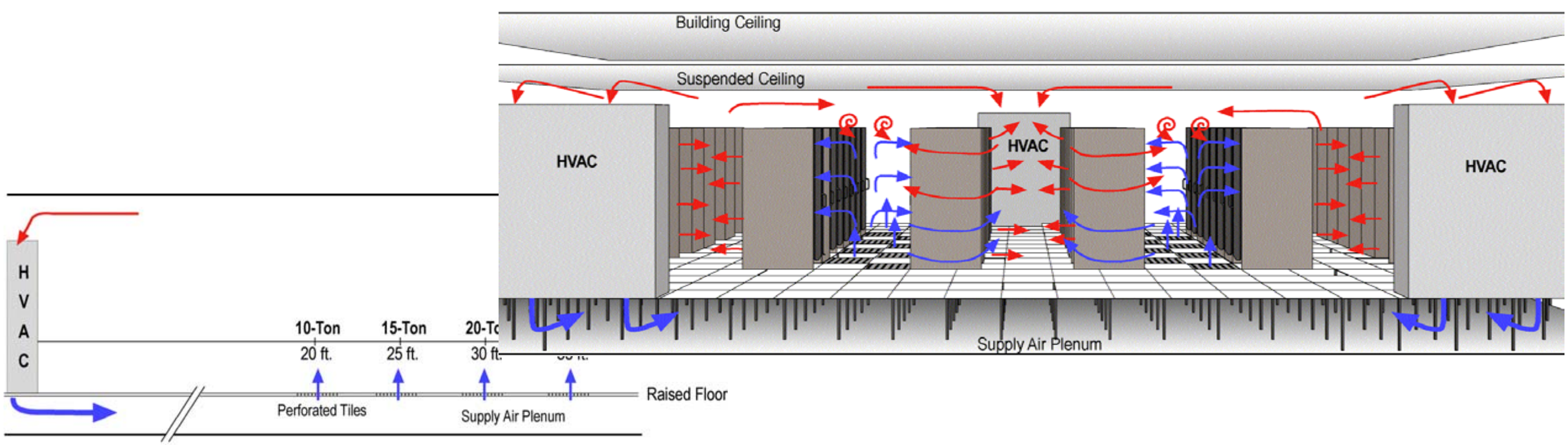
Data Center Air-Conditioning

- One ton of air-conditioning is 12,000BTU's
- How big is a ton of air-conditioning?
  - 30 ton Air Conditioner takes roughly the same floor space as five 42U cabinets
- Problem is not production, it is delivery



Air-conditioning

- Issues we never thought we would have to tackle
  - How many servers can a ton of A/C cool? 3500W worth
  - How much airflow does a densely populated cabinet require? 1200-1600 CFM
  - How does the warmed air get back to the CRAC?
  - How do we maximize A/C efficiency?



cooling requirements

ProLiant Server	DL 360G3	DL 380G3	DL 580G2	BL 10e
Enclosures per 42U Rack	42	21	10	14
Air Flow per 42U Rack	1680 CFM	840 CFM	TBD	TBD
Power	42 × 325W = **13.65kW**	21 × 500W = **10.5** kW	10 × 790W = **7.9** kW	14 × 670W = **9.3** kW
Heat Load**	46,590 BTU/hr	35,840 BTU/hr	26940 BTU/hr	31985 BTU/hr
Cooling Requirement***	3.9 tons	3.0 tons	2.2 tons	2.6tons

airflow data for servers

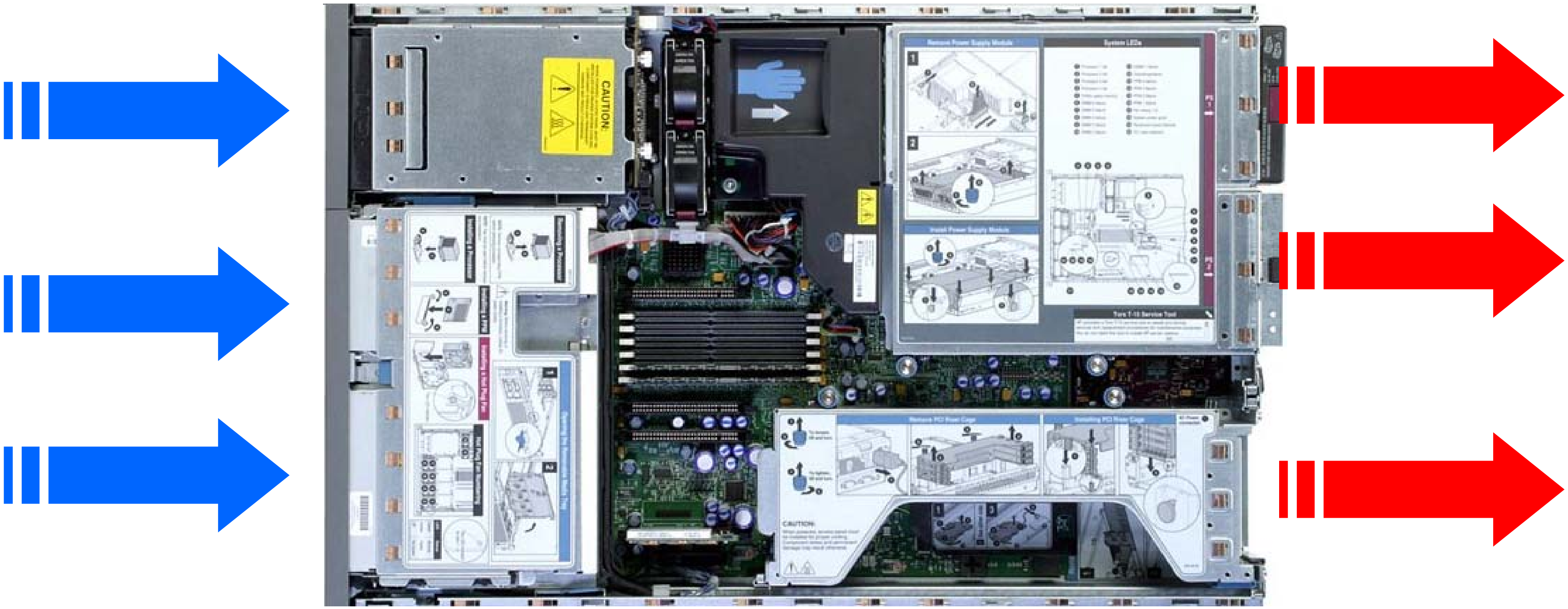
Model	U Size	Low Speed	High Speed
DL320	**1U**	**Pending**	**Pending**
DL360G2	**1U**	**Pending**	**Pending**
DL360G3	**1U**	**40CFM**	**64 CFM**
DL380G2	**2U**	**30 CFM**	**43 CFM**
DL380G3	**2U**	**40 CFM**	**61 CFM**
DL560 G1	**2U**	**32 CFM**	**52 CFM**
DL580 G1	**4U**	**58 CFM**	**58 CFM**
DL580 G2	**4U**	**Pending**	**Pending**
DL740 G1	**4U**	**114 CFM**	**192 CFM**
DL760G1	**7U**	**Pending**	**Pending**
DL760 G2	**7U**	**120 CFM**	**208 CFM**

Model	U Size	Low Speed	High Speed
BL10e	**3U**	**Pending**	**Pending**
BL20p G1	**6U x 1**	**22.7 CFM**	**22.7 CFM**
BL20p G2	**6U x 1**	**22.4 CFM**	**36.7 CFM**
BL40pG1	**6U x 4**	**72.6 CFM**	**99.5 CFM**
Patch Panel	**6U x 1**	**21.8 CFM**	**21.8 CFM**
GbE Sw	**6U x 1**	**21.1 CFM**	**21.1 CFM**
Pwr Encl	**3U**	**Pending**	**Pending**

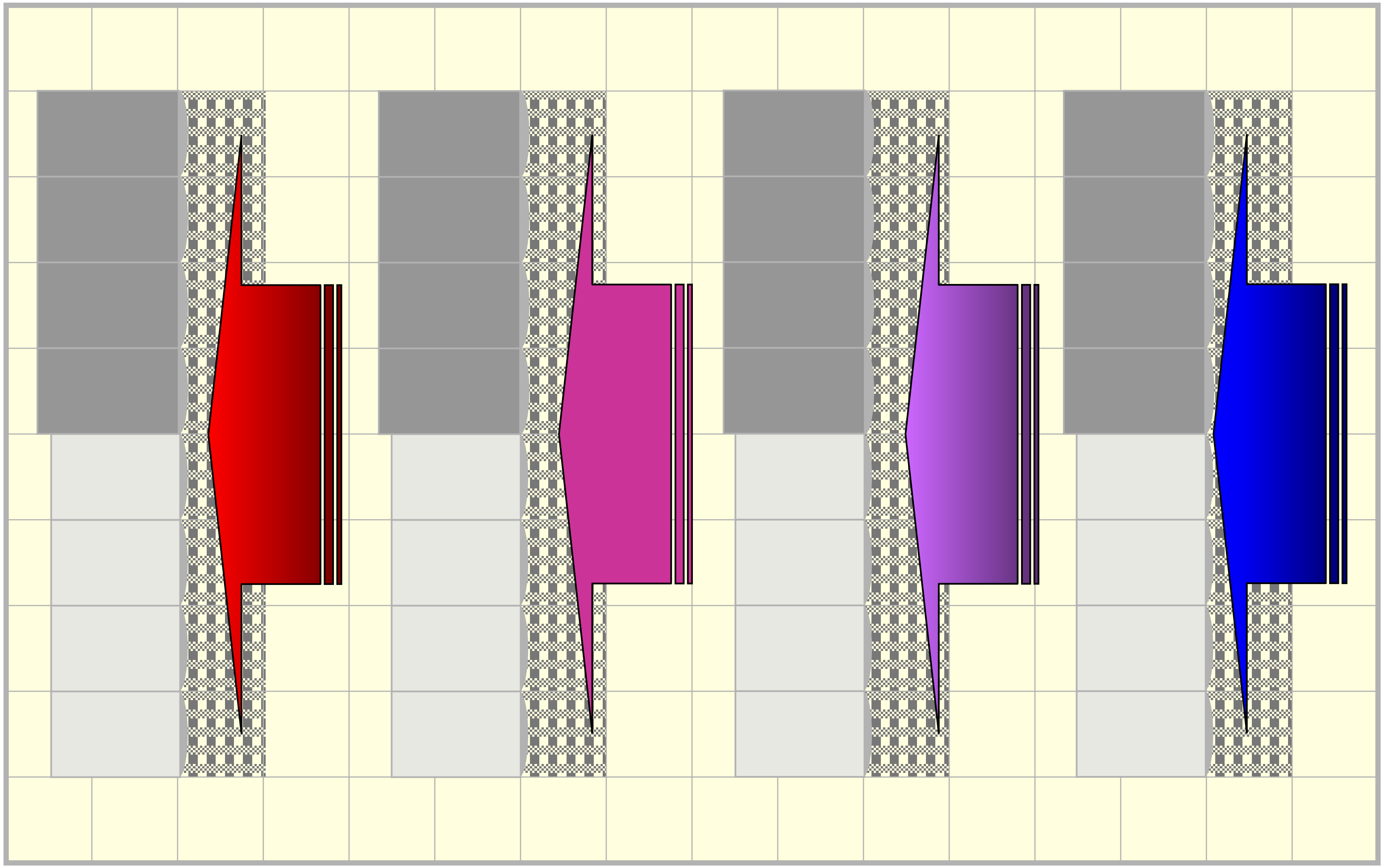
Critical Nature of Airflow

If tremendous design effort goes into designing each server for optimized airflow,

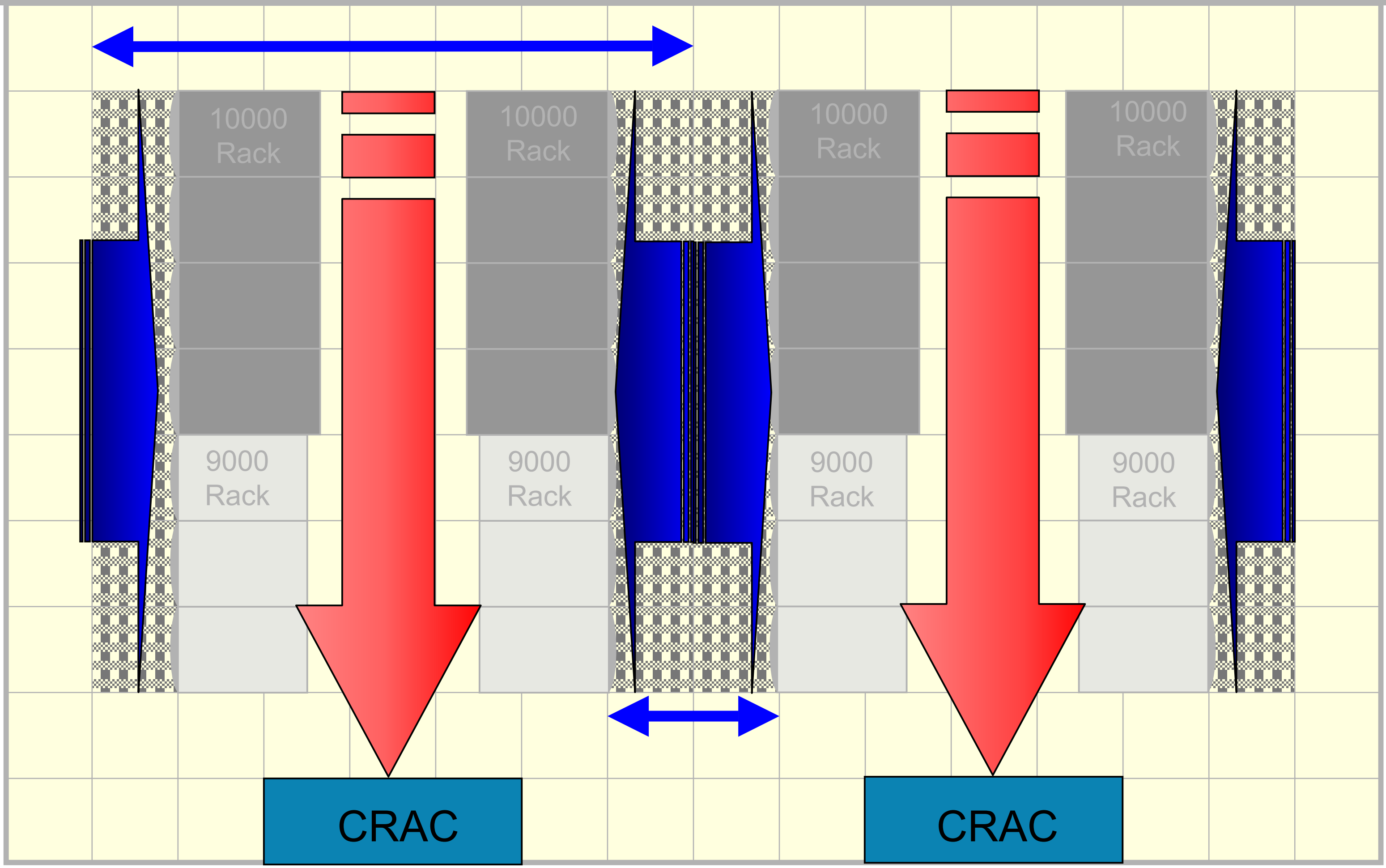
How much effort should be spent on airflow at the rack and data center level?



Yesterday’s raised floor layout method



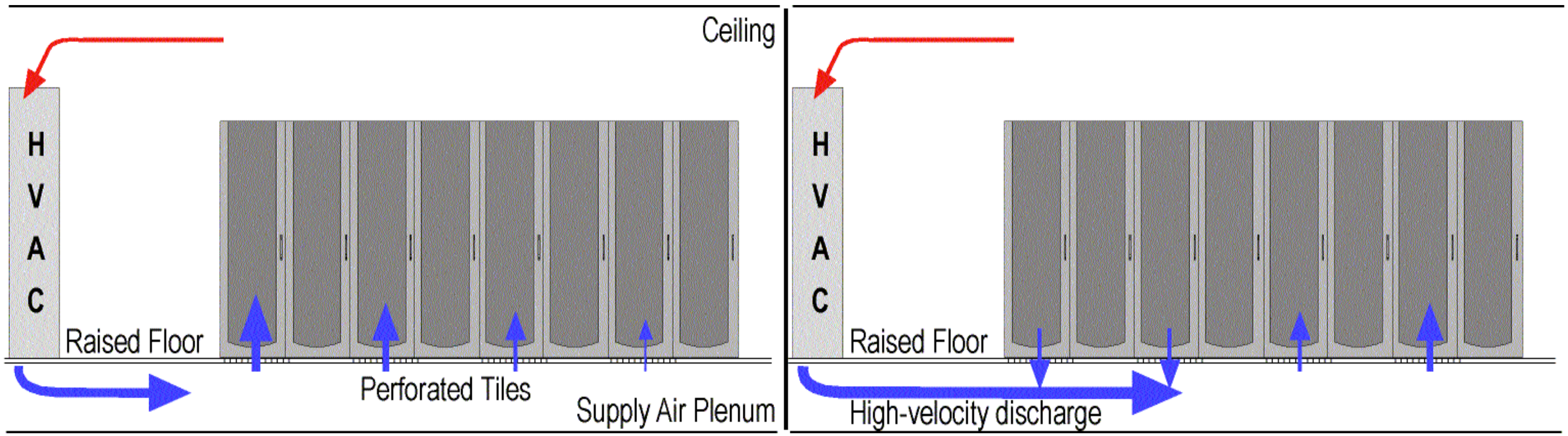
Today's Hot Isle / Cold Isle Layout



- Cold Isle spacing should be 48" minimum, two full tiles.
- Center Cold Isle to Center Cold Isle should be a minimum of 14' or 7 full tiles.
- CRAC units should be perpendicular to rows and centered on hot isle, blowing in same direction.

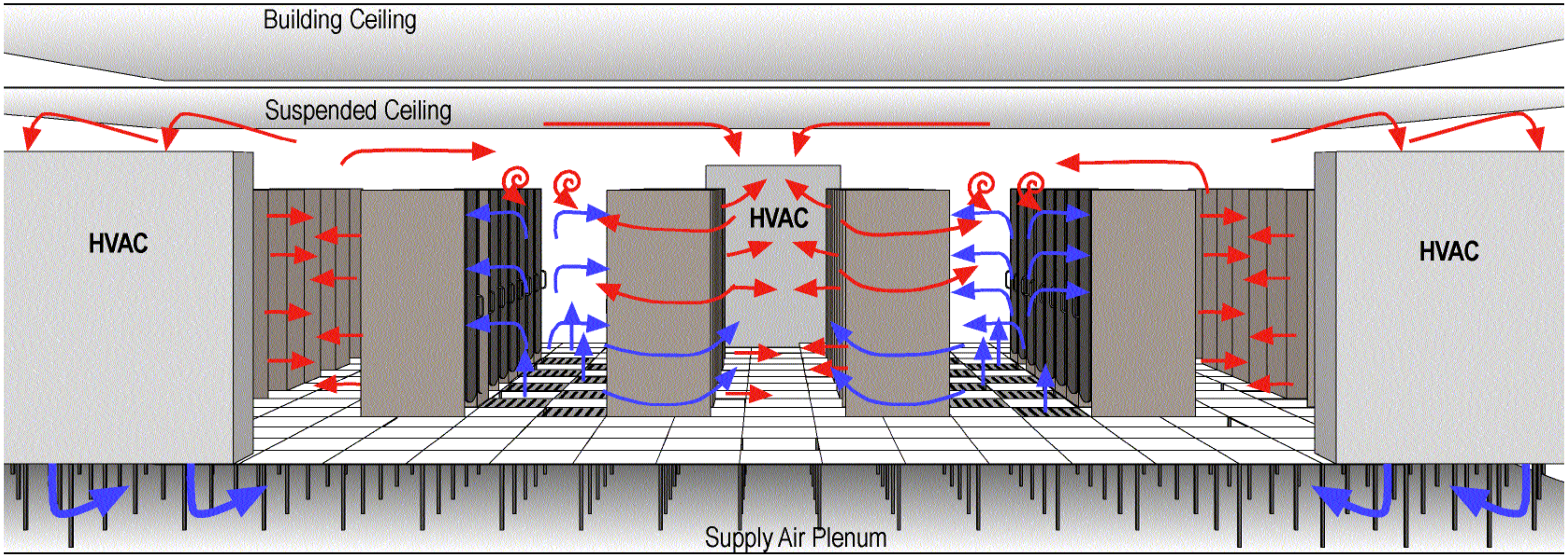
Issues of poor CRAC placement

If CRAC is placed to close to cabinets very low flow or even negative air pressure (vacuum) can be created, robbing system of cooling



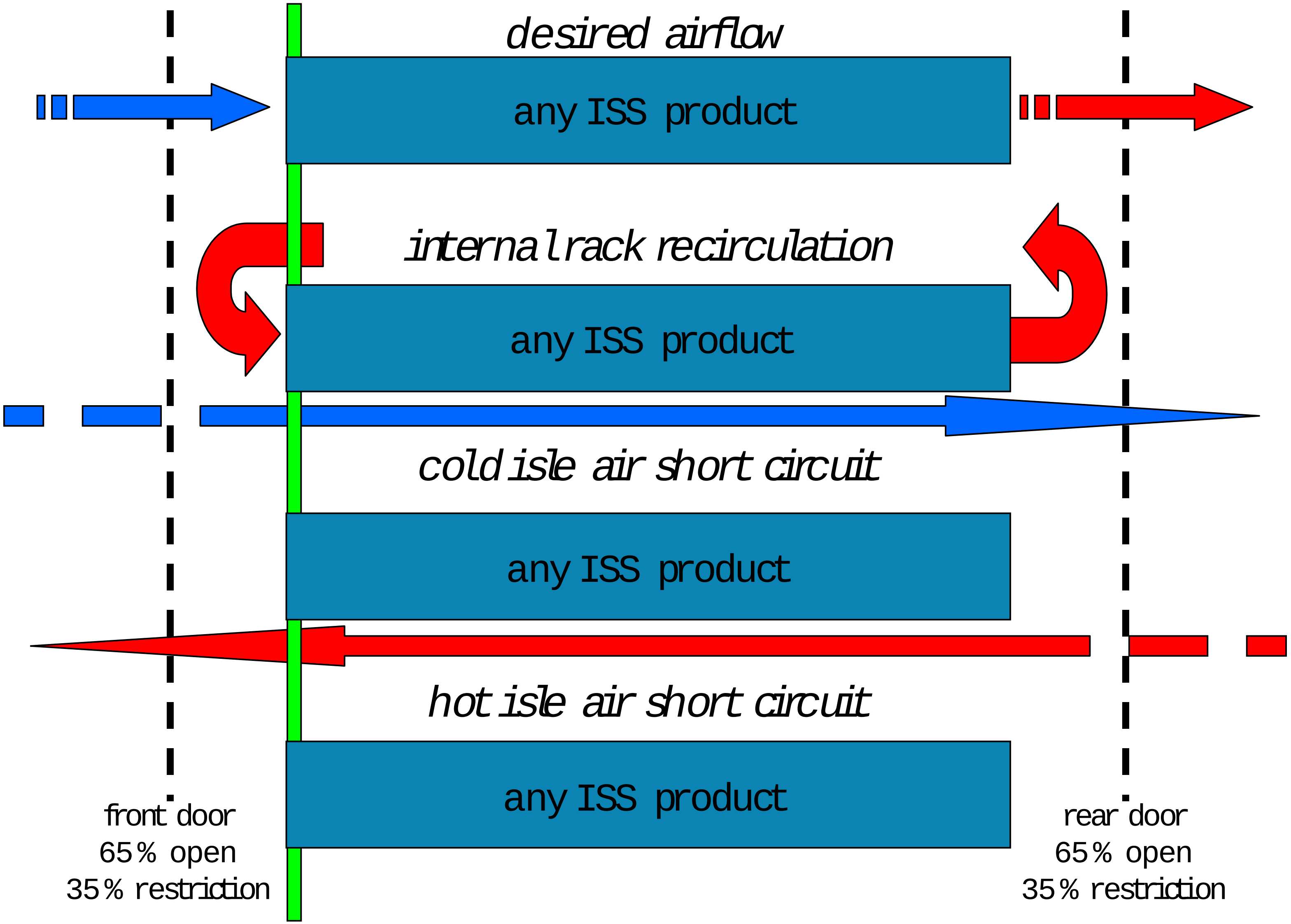
combined with end of row recirculation, this can create some big hot spots

The negative affects of mixed air



Contributors: low ceiling - low return air volume
blanking panels not used
internal rack recirculation
end of row recirculation
low volume perf tiles
low supply plenum pressure

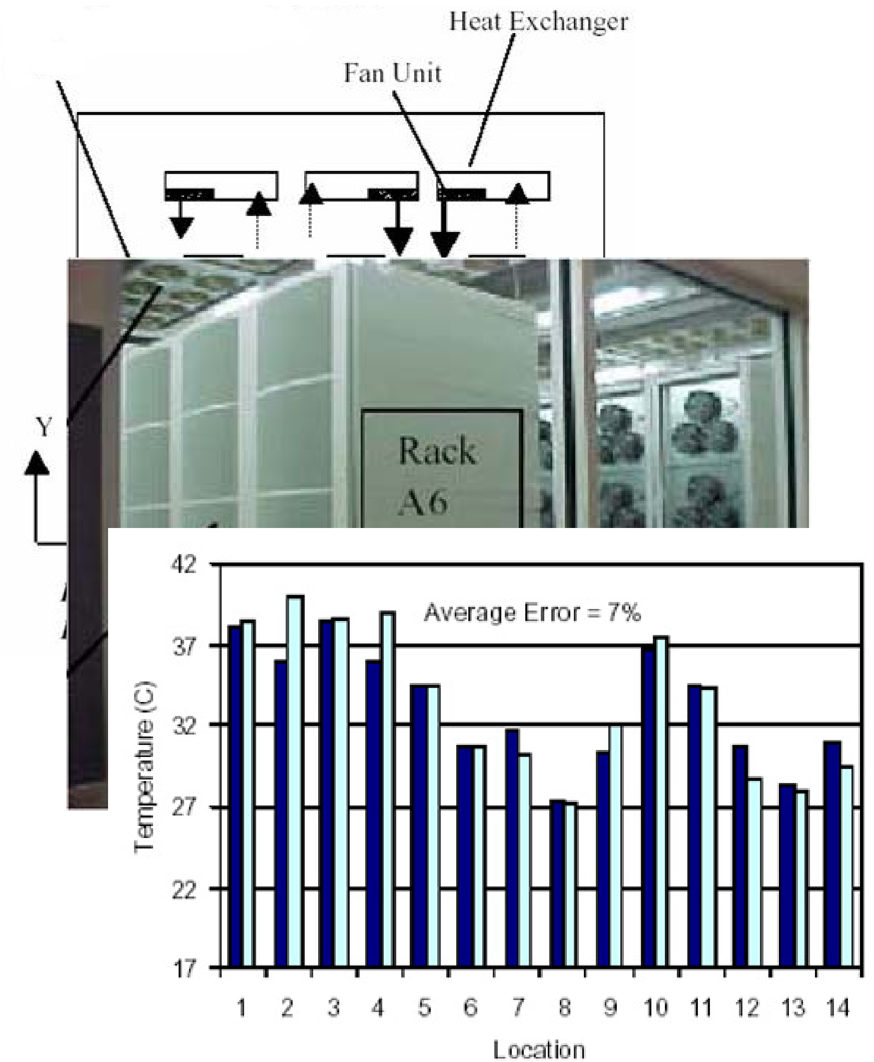
Blanking panels, not for looks any more



Data Center Research

hp "cool labs"

- Research on analysis of data center with CFD theoretical model
- Constructed physical model of to prove theory
- Physical model analysis proved theoretical model assumptions analyzed
- Accuracy proven to be within 7%



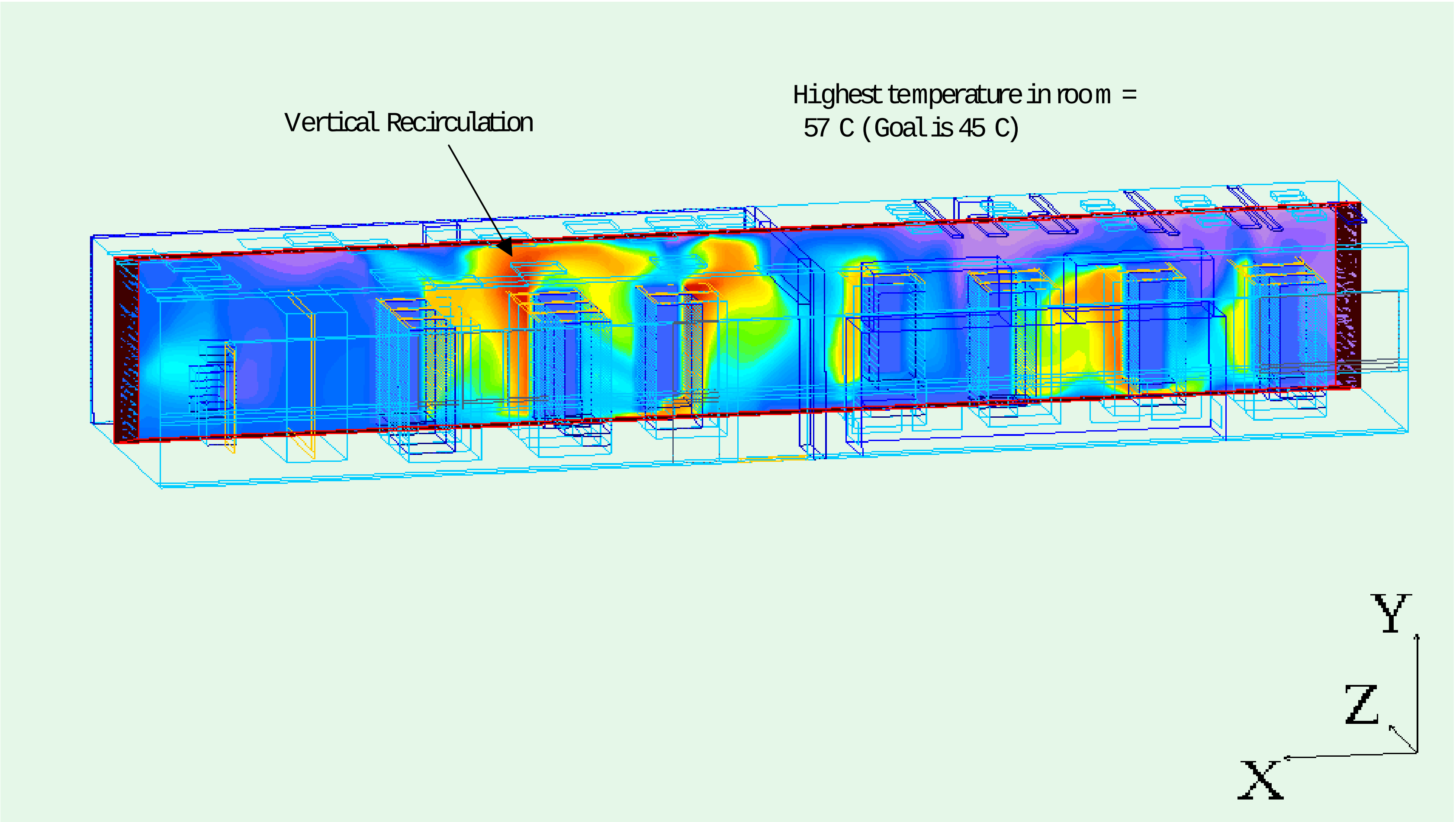
hp data center validation services

Current setup,

- Determine need for detailed validation services
- Engage the data center services team
- Front end information from the customer is compared against criteria document
- If the data center exceeds the key criteria the services team suggests a detailed 3D CFD analysis.
- Rule of thumb, data centers exceeding 100 W per sq ft are primary candidates for detailed analysis.

(gross load over gross area)

Dreamworks data center before "Static Smart Cooling"



Computational Fluid Dynamics (CFD)

Detailed CFD analysis first models the customer data center’s

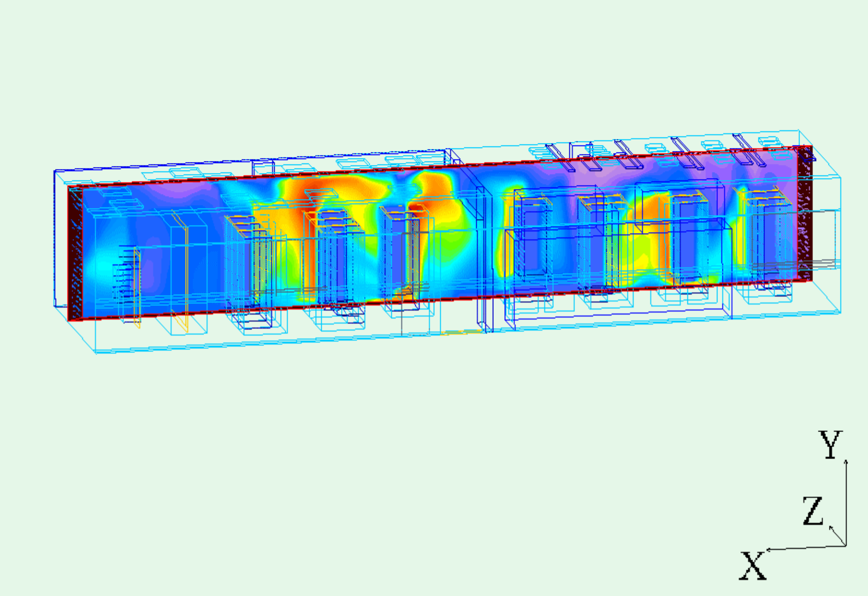
- physical layout
- air conditioning resources
- enterprise infrastructure equipment

Then provisions

- air conditioning resources
- CRAC unit settings,
- perforated floor tile layout
- return air vents / supply vents (if applicable)
- heat load distribution

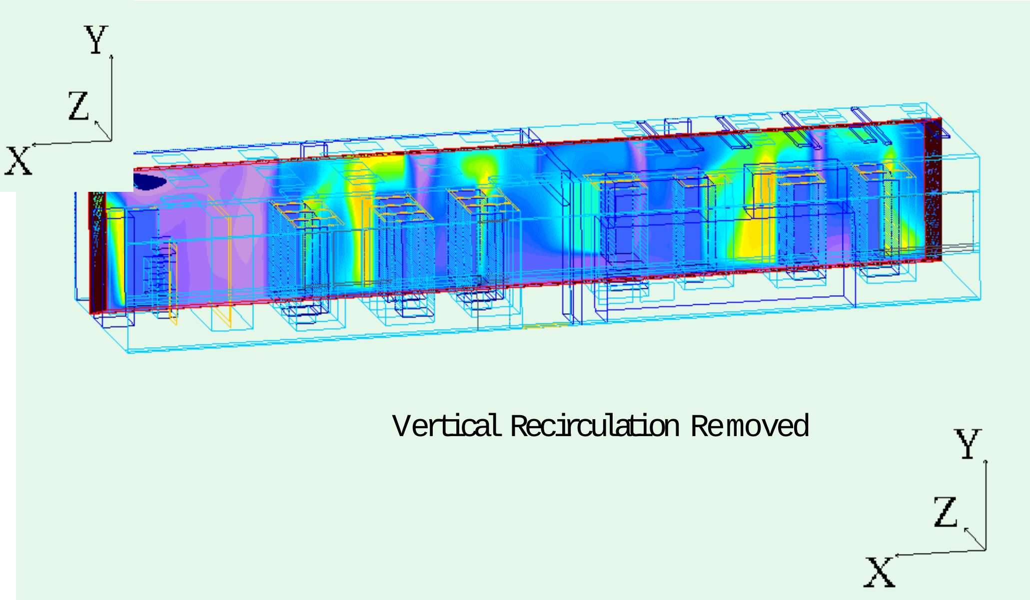
To provide the best possible support for the customer’s enterprise

Data Center after "Static Smart Cooling"



Before

Highest temperature in room = 47 C (HPL goal was 45 C)



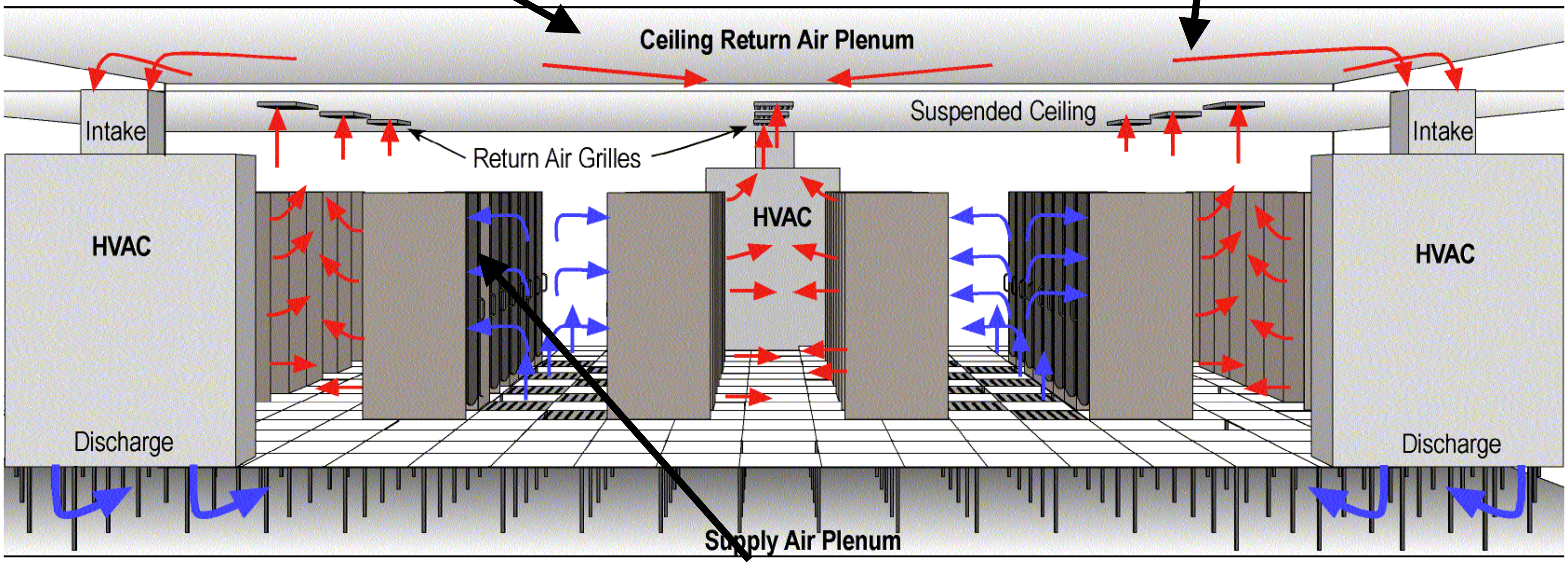
After

Vertical Recirculation Removed

Optimized cooling configurations

ceiling return air plenum

critically placed return grilles



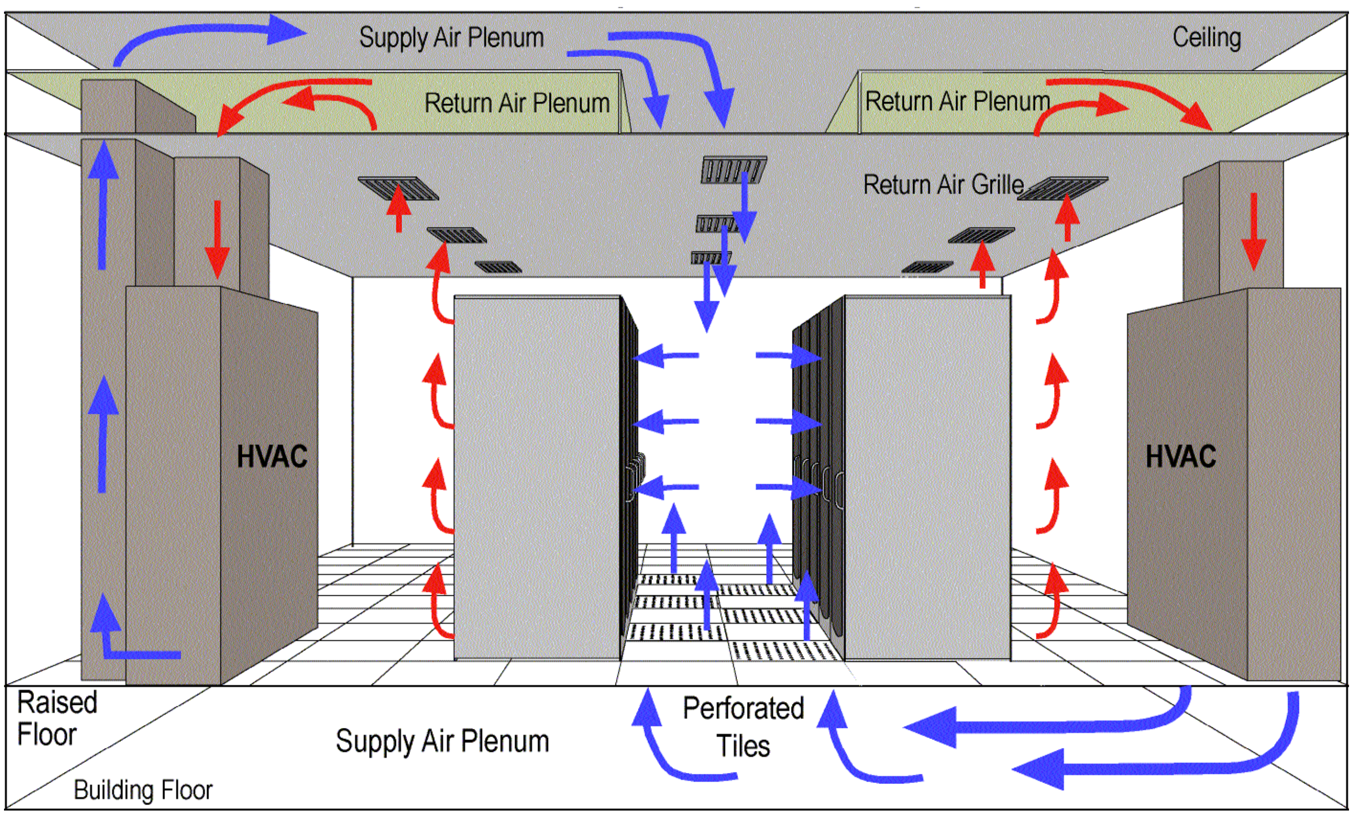
rack blanking panels

Just 3 of many areas that can be optimized!

What Key Metric is Changing Here?

Distribution

dual supply plenum configuration



Recommendations

- When arranging cabinets in a data center, arrange front to front, back to back
- Use “Blanking Panels” to fill all empty space in racks. This prevents the short circuiting of cold air to the hot aisle.
- Calculate air conditioning based on “Sensible Capacity” not rated tonnage.
- **<u>Map out a maximum load on the facility and keep to it</u>**
  - May involve empty slots in racks

Recommendations (cont)

- Avoid creating hot spots
  - Work to balance load in your data center to maximize HVAC capabilities
- Biggest mistake in laying out equipment in data centers is arranging by component type
  - Racks and rows full of servers (high heat loads)
  - Racks and rows full of Storage

questions

References

1. Patel, C.D., Bash, C.E., Belady, C., Stahl, L., Sullivan, D., July 2001, "Computational fluid dynamics modeling of high compute density data centers to assure system inlet air specifications", *Proceedings of IPACK'01 – The PacificRim/ASME International Electronics Packaging Technical Conference and Exhibition*, Kauai, Hawaii.

[Modeling Validation]

2. Stahl, L., Belady, C. L., Oct 2001, "Designing an Alternative to Conventional Room Cooling", *Proceedings of the INTELEC'01 International Telecommunications Energy Conference*, Edinburgh, Scotland ……………………………… [Infrastructure]

3. Friedrich, R., Patel, C.D., Jan 2002, "Towards planetary scale computing - technical challenges for next generation Internet computing", *THERMES 2002*, Santa Fe, New Mexico……………………………………………..[Pervasive Computing and Data Centers]

4. Patel, C.D., Sharma, R.K, Bash, C.E., Beitelmal, A, May 2002, "Thermal Considerations in Cooling Large Scale High Compute Density Data Centers", *ITherm 2002 - Eighth Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems*, San Diego, California…………………….
[Static Provisioning]

5. Sharma, R.K, Bash, C.E., Patel, C.D., *Dimensionless Parameters for Evaluation of Thermal Design and Performance in Large Scale Data Centers*, 8th AIAA/ASME Joint Thermophysics and Heat Transfer Conference, St. Louis, June 2002.

[Data Center Figure of Merit, non-dimensional numbers for data centers]

6. DeLorenzo, D., Thermal Trends and Inflection, 7X24 Exchange 2002, Orlando, Florida [Trends]