**Andreas Bieswanger**

**STSM, System z Firmware Development & GreenIT**

# Energy Management for IBM zEnterprise™ 196

Session 8669


February 28th, 2011

# agenda

| | |
|---|---|
| *1* | **zEnterprise Energy Efficiency Improvements** |
| *2* | **zEnterprise Energy Management Controls** |
| *3* | **Unified Resource Manager - Energy Monitoring and Management** |
| *4* | **IBM Energy Management Stack Integration** |

# IBM zEnterprise System – Best in Class Systems and Software Technologies

*A system of systems that unifies IT for predictable service delivery*

## Unified management for a smarter system: **zEnterprise Unified Resource Manager**

- Unifies management of resources, extending IBM System z® qualities of service end-to-end across workloads
- Provides platform, hardware and workload management

## The world's fastest and most scalable system: **IBM zEnterprise™ 196 (z196)**

- Ideal for large scale data and transaction serving and mission critical applications
- Most efficient platform for Large-scale Linux® consolidation
- Leveraging a large portfolio of z/OS® and Linux on System z applications
- Capable of massive scale up, over 50 Billion Instructions per Second (BIPS)


## Scale out to a trillion instructions per second: **IBM zEnterprise BladeCenter® Extension (zBX)**

- Selected IBM POWER7® blades and IBM System x® Blades[1] for tens of thousands of AIX® and Linux applications
- High performance optimizers and appliances to accelerate time to insight and reduce cost
- Dedicated high performance private network

1 All statements regarding IBM future direction and intent are subject to change or withdrawal without notice, and represents goals and objectives only.

# Goals for Energy Management

**Cost Reduction and Avoidance**

- Identify opportunities for energy cost reduction *(Operating Expenses)*
- Delay facility expansion due to energy or cooling constraints *(Capital Expenses)*

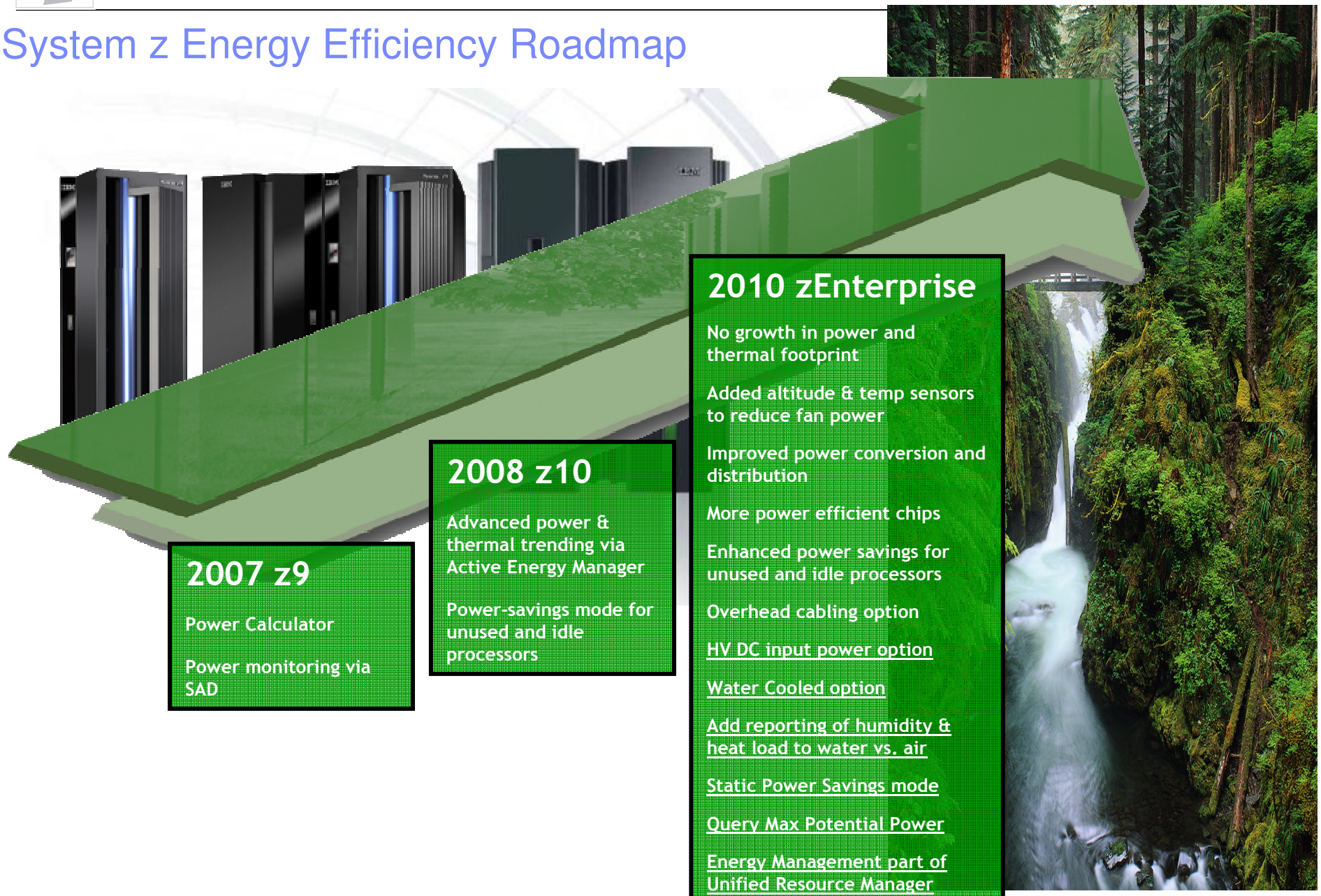
**Remove Operational Barriers**

- Manage power and cooling capacity to enable growth and flexibility
- Power Control (Capping, Power Saving)

**Manage Risk and Streamline Compliance**

- Document and validate energy efficiency gains to stakeholders

# System z Energy Efficiency Roadmap



## 2007 z9

Power Calculator

Power monitoring via SAD

## 2008 z10

Advanced power & thermal trending via Active Energy Manager

Power-savings mode for unused and idle processors

## 2010 zEnterprise

No growth in power and thermal footprint

Added altitude & temp sensors to reduce fan power

Improved power conversion and distribution

More power efficient chips

Enhanced power savings for unused and idle processors

Overhead cabling option

HV DC input power option

Water Cooled option

Add reporting of humidity & heat load to water vs. air

Static Power Savings mode

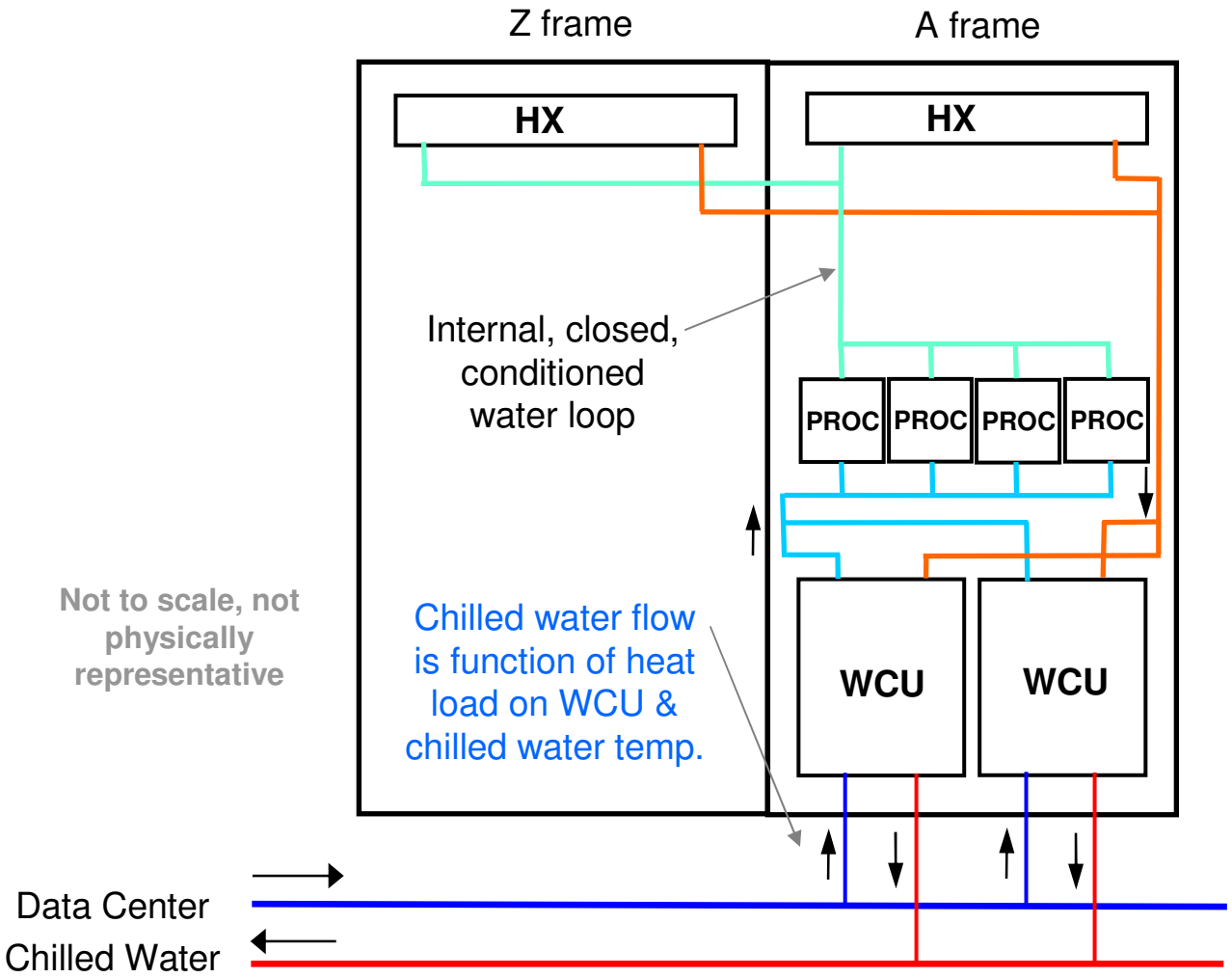
Query Max Potential Power

Energy Management part of Unified Resource Manager

# z196 High Voltage DC

- A direct HV DC datacenter power design can improve data center energy efficiency by removing the need for an additional DC to AC inversion step.
- System bulk power for all z196 systems is updated to support HV DC so the only difference in shipped HW to implement the option is the DC line cord:
  - This adds DC line cord feature codes
  - Nominal DC supply voltage supported will be:
    - 380V – 520V (absolute min 330V, absolute max 550V)
    - New technology, multiple proposed “standards”
    - Support both ground referenced and dual polarity HV DC supply
    - As defined will support -380V to -520V, +/-190 to +/-260V, +380V, etc.
- System saves approximately 3% input power when run on HV DC

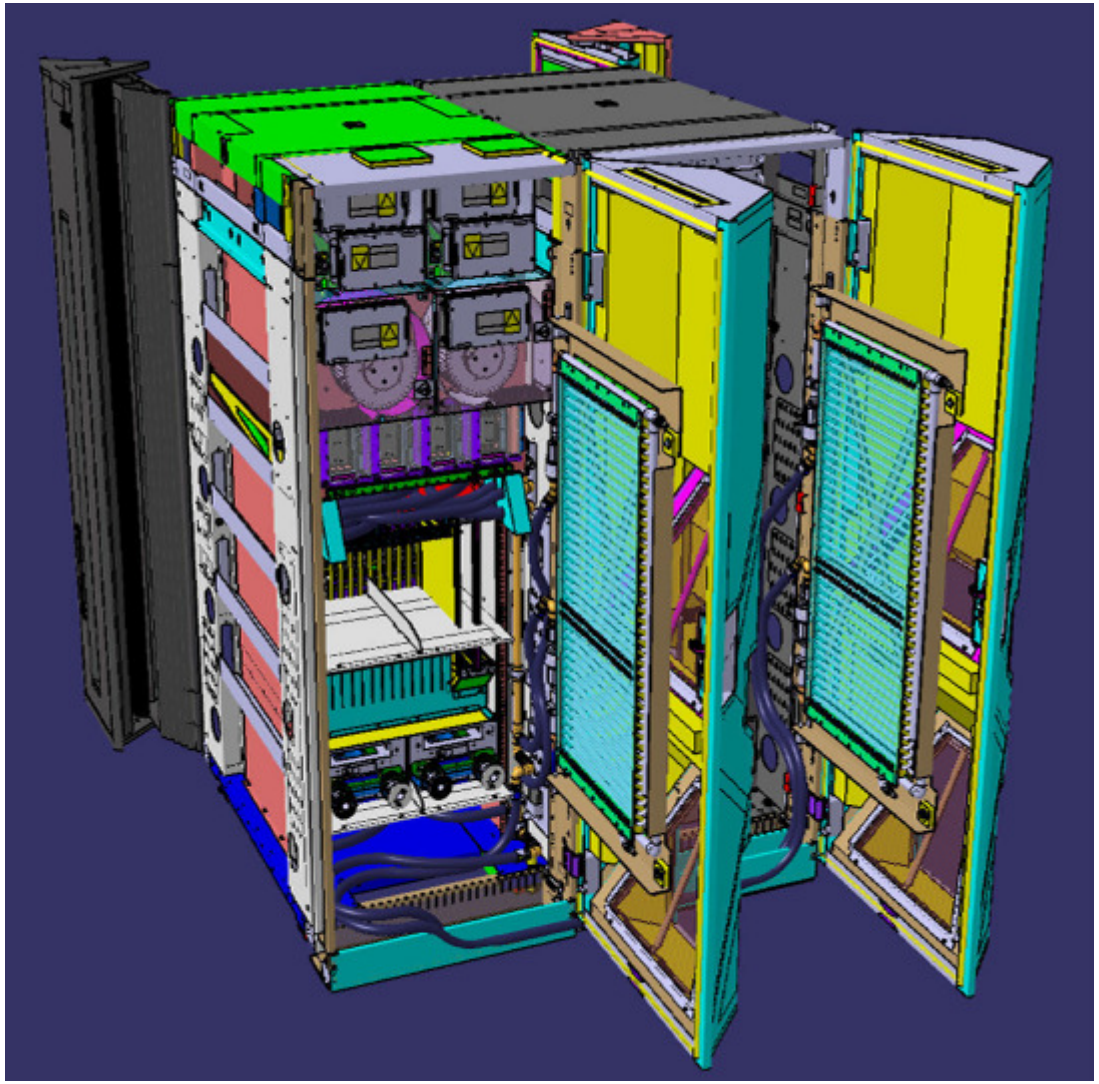
# z196 Water Cooling Option




- **Water cooled cold plate on processor MCM in each processor book**
- **2N Water Conditioning Unit (WCU) with independent chilled water connections**
- **One WCU can support system**
- **Heat Exchanger (HX) removes heat from exhaust air at back of both frames**
- **Typically ~70% of system heat load is removed to water.**
- **Air cooling back-up mode for maximum robustness (all heat load to air if lose chilled water in to WCU's)**

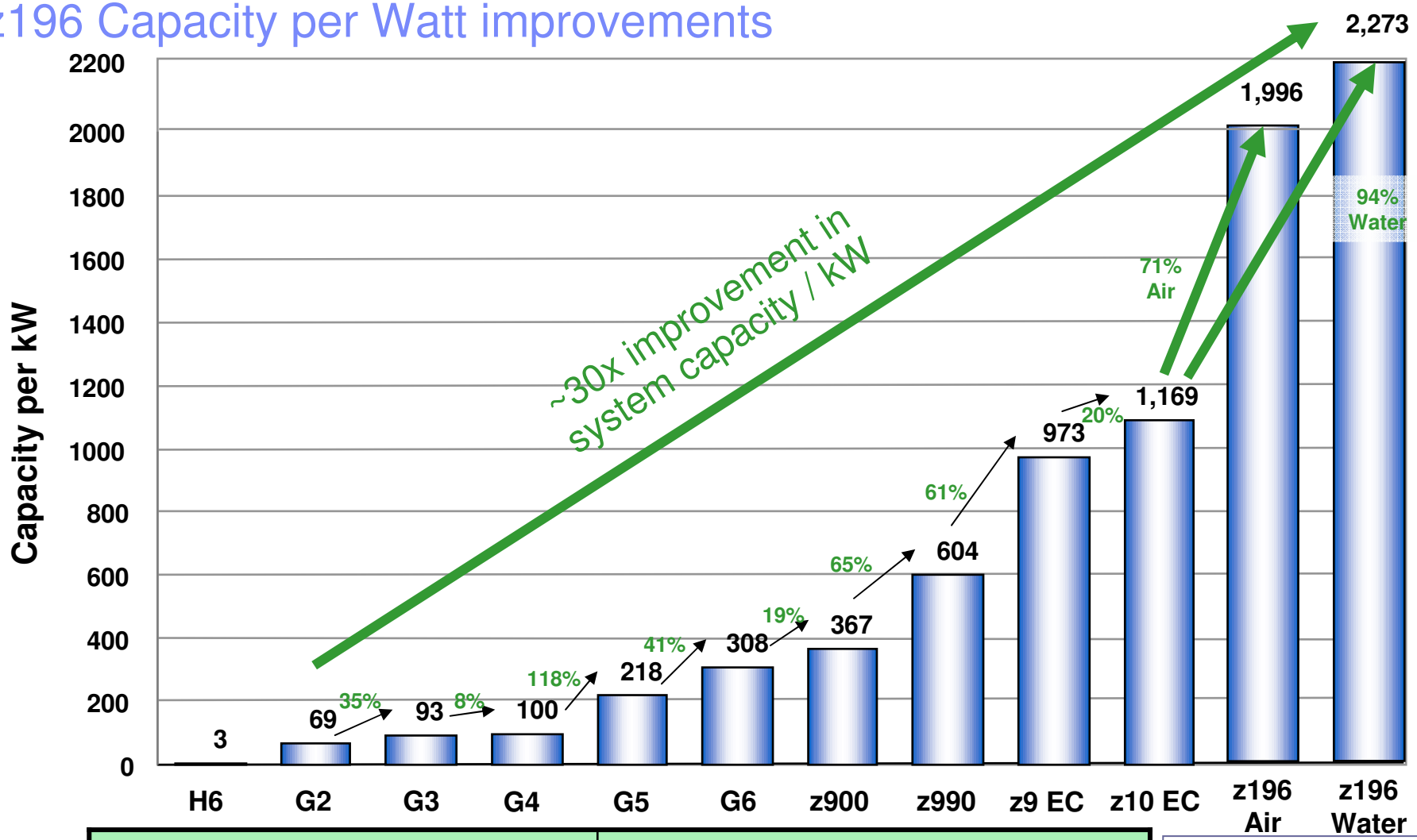
2N building chilled water lines will have better RAS – single facility supply/return shown here.

# Z196 Water Cooling Option



- Reduce max air heat load to less than 10 kW (about 5 kW typical)
- Input energy saving 2 kW
- Additional power saving in data center typically about 3 kW
  (water cooling efficiency higher than air cooling efficiency)

# z196 Capacity per Watt improvements




| 15 years of CMOS: G2 to z196 * | | Net Effect: G2 to z196 * | |
|---|---|---|---|
| Power Increase: | **17%** per year | Performance increased by: | **~300x** |
| Performance increase: | **46%** per year | **Performance / kWatt increased by:** | **~30x** |
| Power density increase: | **13%** per year | Performance / sq ft increased by: | **~190x** |

Note: Capacity/kWatt assumes hot room, max plugged I/O power, max memory power and all engines turned on. Real world max capacity system is about 3/4 of this.

Note: z196 data is best available 8/27/2010

# Three fundamentals of energy management


## Measure/Trend Power Consumption

- Determine the power being consumed now
- Trending energy and thermals over extended periods of time


## Allocate Power Correctly

- Rightsizing of power and cooling allocations
- Enables deployment of more servers within the physical limits of a data center


## Reduce power consumed

- Reduce power in periods of low utilization to reduce energy cost

# Power Estimation Tool for z196

**Configuration**

| | |
|---|---|
| Model | M32 |
| Voltage group | 208 to 240V group |
| Line voltage | 208 |
| Water cooled (FC 0159) | Yes |
| Room temp | <28C (82F) |
| Installed altitude (in feet) | 0 |
| Workload | Normal power save |
| Flexible memory | No |
| Balanced power (FC 3003) | No |
| Cargo cages (FC see help) | 1 |
| Cayuga drawers (FC 4000 or 4004) | 1 |

**CEC data**

| Card Name | FC | Quantity | Max |
|---|---|---|---|
| Customer configured processors | | 9 | 32 |
| Base SAPs | | 6 | |
| Ordered memory | | 32 | 1520 |
| IBT-2 Fan-out Card for Copper | 0162 | 0 | 16 |
| IBT+K42-2 Fan-out Card for Optics | 0163 | 0 | 16 |

**Cargo Cages**

| Card Name | FC | Quantity | Max |
|---|---|---|---|
| ISC-daughter card | 0218 | 0 | 24 |
| IB-MP Daughter Card | 0326 | 0 | 24 |
| Crypto Express3 | 0864 | 0 | 8 |
| 16 port ESCON (old QZ26) | 2323 | 0 | 36 |
| 16 port ESCON (new QZ2S) | 2323 | 10 | 36 |
| FICON Express4 10KM LX | 3321 | 4 | 36 |
| FICON Express4 SX | 3322 | 0 | 36 |
| FICON Express4 4KM LX | 3324 | 0 | 36 |
| FICON Express8 10KM LX | 3325 | 0 | 36 |
| FICON Express8 SX | 3326 | 0 | 36 |
| OSA-Express3 GbE LX | 3362 | 0 | 24 |
| OSA-Express3 GbE SX | 3363 | 0 | 24 |
| OSA-Express2 GbE LX | 3364 | 0 | 24 |
| OSA-Express2 GbE SX | 3365 | 0 | 24 |
| OSA Express2 1000base-T | 3366 | 0 | 24 |
| OSA-Express3 1000base-T | 3367 | 10 | 24 |
| OSA-Express3 10 GbE LR | 3370 | 0 | 24 |
| OSA-Express3 10 GbE SR | 3371 | 0 | 24 |
| Power Sequence Controller | 6501 | 0 | 2 |

**Results**

| | |
|---|---|
| System total heat load: | 20677 BTU/hr |
| Utility input power: | 6064 W |

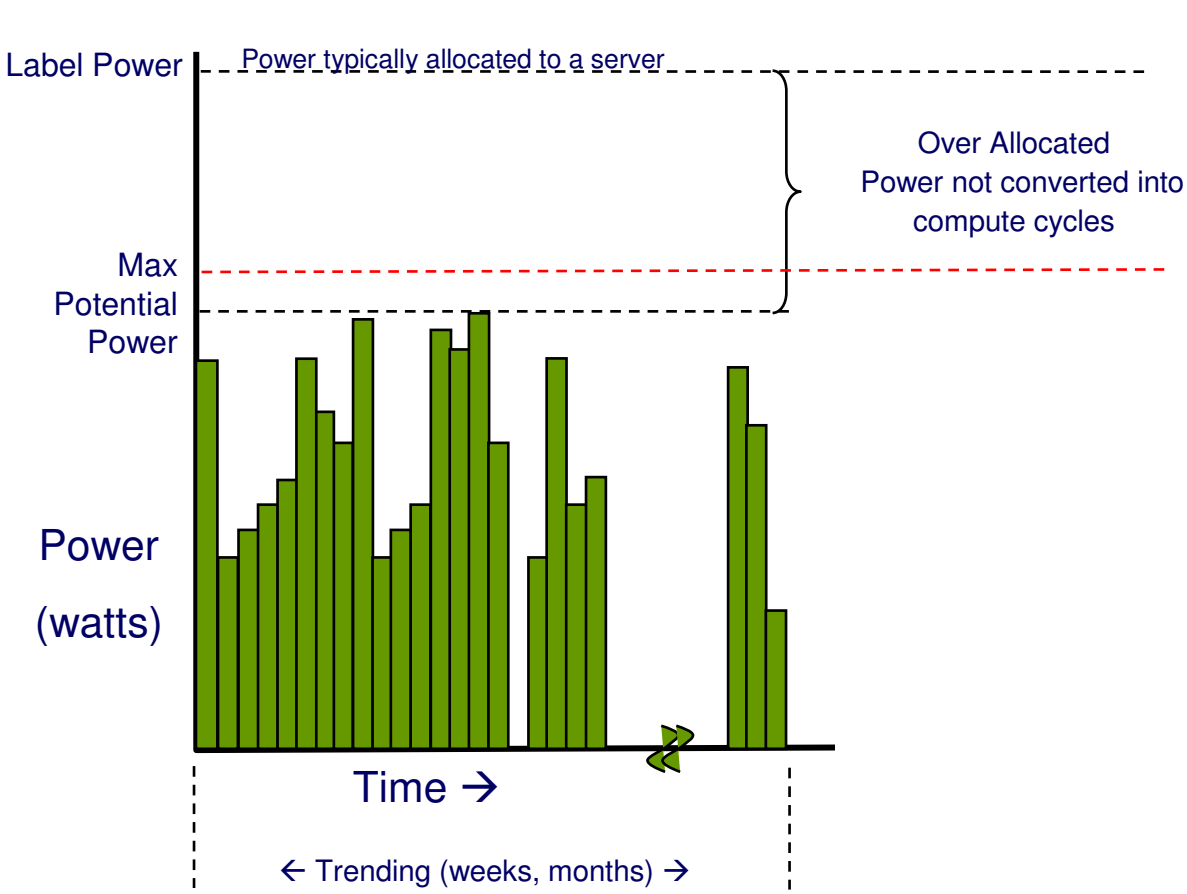
Calculate Clear

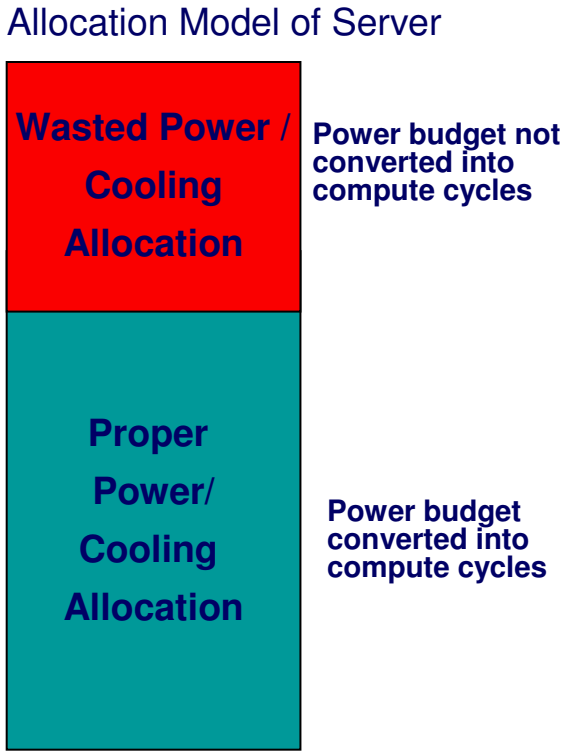
https://www-304.ibm.com/servers/resourcelink/hom03010.nsf/pages/pet2817v2110?opendocument

# Max Potential Power

- **Main use cases**
  - Allows reducing power allocation for system since you know the maximum power system can draw even with faults and hot room
  - Allows facility and system people without knowledge of z system configuration and use details to query max possible power of system
  - Looks like power capping to higher level management tools
- **Base mechanism: Calculation of max potential power based on**
  - System configuration
  - Altitude (absolute pressure sensors in bulk power subsystem)
  - Hot room environment
  - Highest single fault service scenario power condition for this configuration
  - Reasonable tolerances
- **Max Potential Power should be used in conjunction with the System z Power Estimation Tool which allows pre-planning for power and cooling needs**

# Optimize Power/Cooling Allocation with Max Potential Power



# Static Power Saving Mode

- **Main use cases**
  - Periods of low utilization
  - CBU Systems: Systems used for disaster recovery
- **Base mechanism**
  - Build upon existing RAS functions (frequency/voltage variation) implemented originally for MRU failures (since z900)
  - Use frequency and voltage reduction to reduce energy consumption of CEC
  - Only explicitly triggered by customer. No autonomic changes done “under the cover”
- **Power Savings Mode expectations**
  - Frequency reduction: 17%
  - Processor voltage reduction: 9%
  - Expected system power savings: 15%-20% (configuration dependent)
- For air-cooled systems entering power save is limited to once a day.
- Update to “**STSI: SYSIB 1.2.1 (Basic-Machine CPU) Performance-Reduction Indicator**” to reflect entering and leaving power save mode

# … value made possible by the Unified Resource Manager

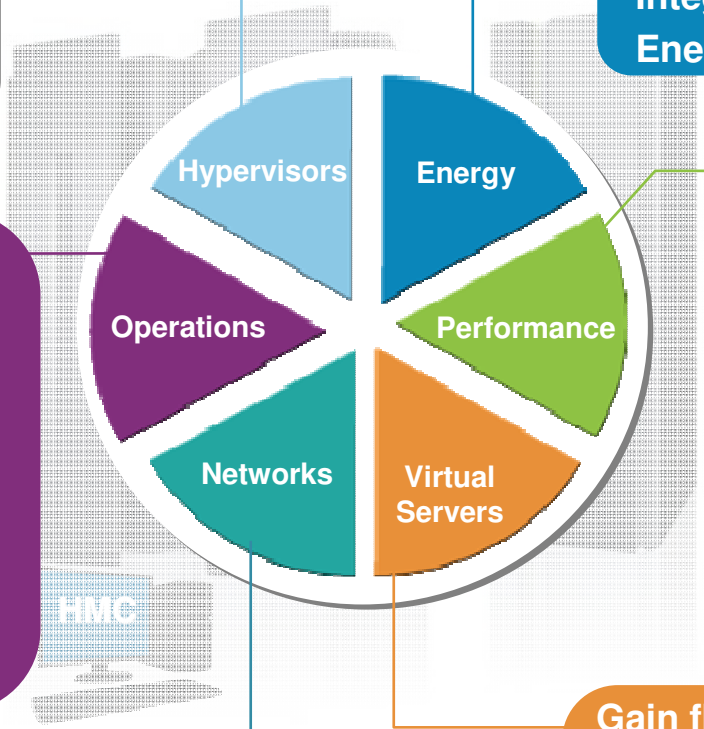
**Simplified installation of hypervisors**

**Gain significant time to market with improved speed of deployment**

**Insight into energy consumption and environmental data**

**Integrated energy management**

**Energy cost savings**




**Save time, cost and simplify asset management**

**Decrease problem determination and resolution time for cross-platform resources**

**Improve and simplify cross-platform availability procedures**

**Enable broader and more granular view of resource consumption**

**Allow critical workloads to receive resources and priority based on goal-oriented policies established by business requirements**

**Smart business adjustments based on workload insight**

**Provide deep insight into how IT resources are being used**

**Factory installed and configured network**

**Improved network security with lower latency, less complexity, no encryption/decryption**

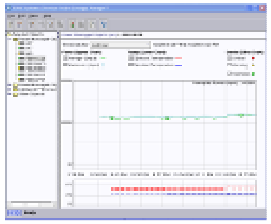
**Gain flexibility, consistency and uniformity of virtualization**

**Provide the business with faster time to market**

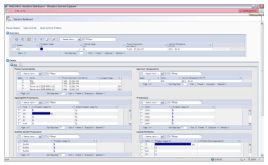
**Simplified network management for applications**

# zEnterprise Energy Management Structure

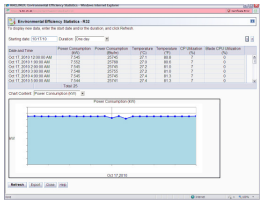
Tivoli Monitoring
for Energy Management



Active Energy Manager

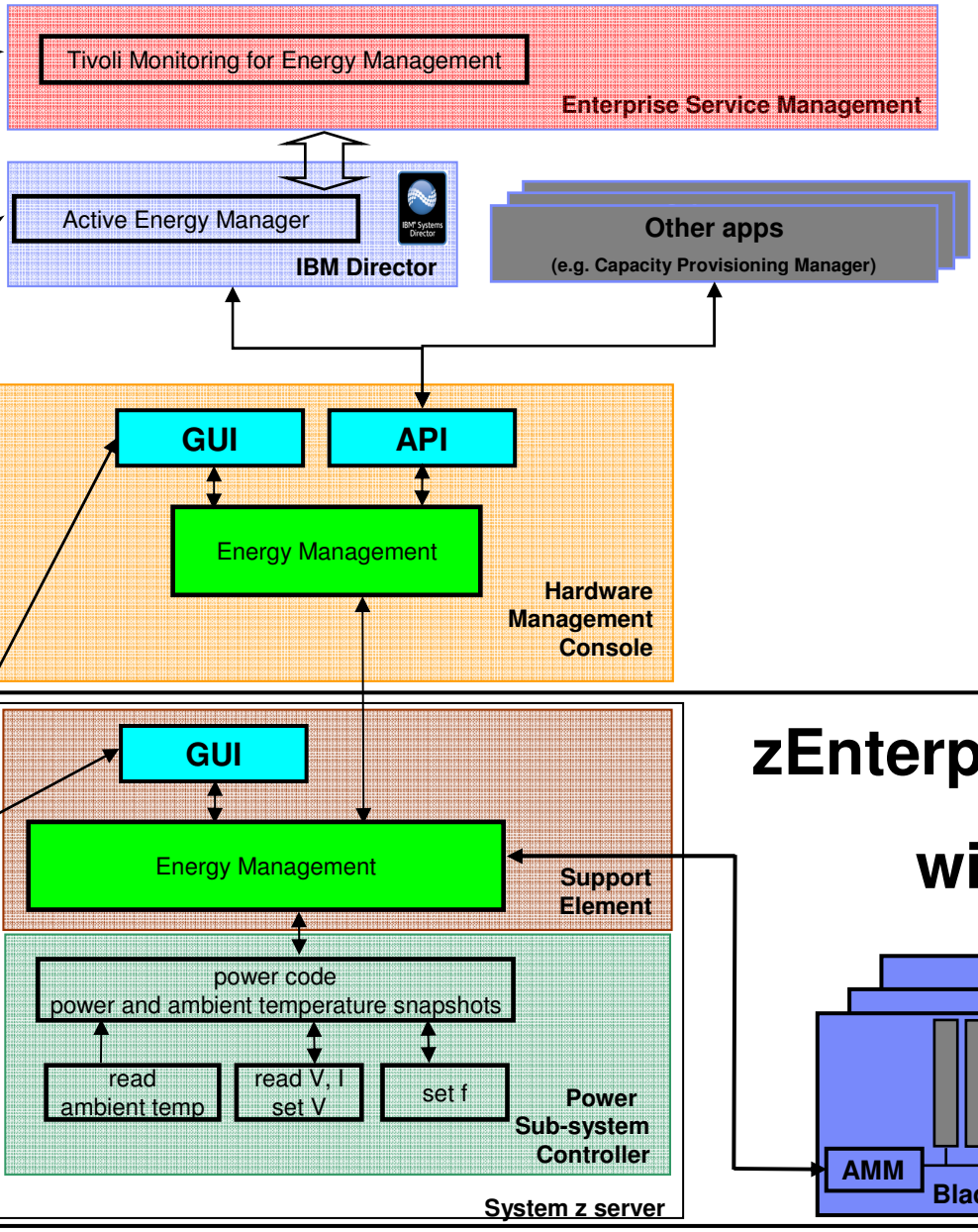


HMC/SE
Monitoring Dashboard



Environmental Efficiency
Statistics Display

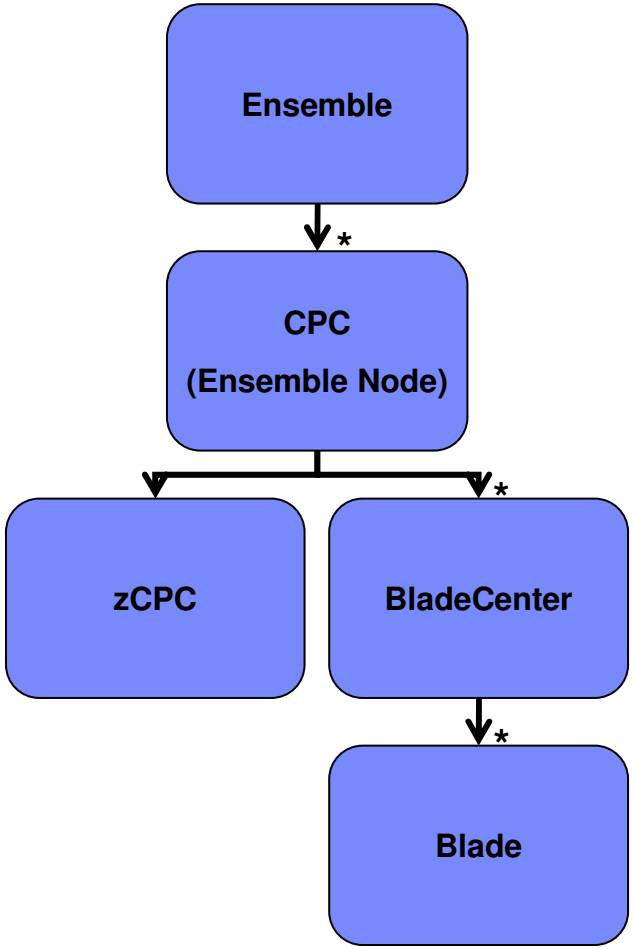
Energy Management
Details Panels



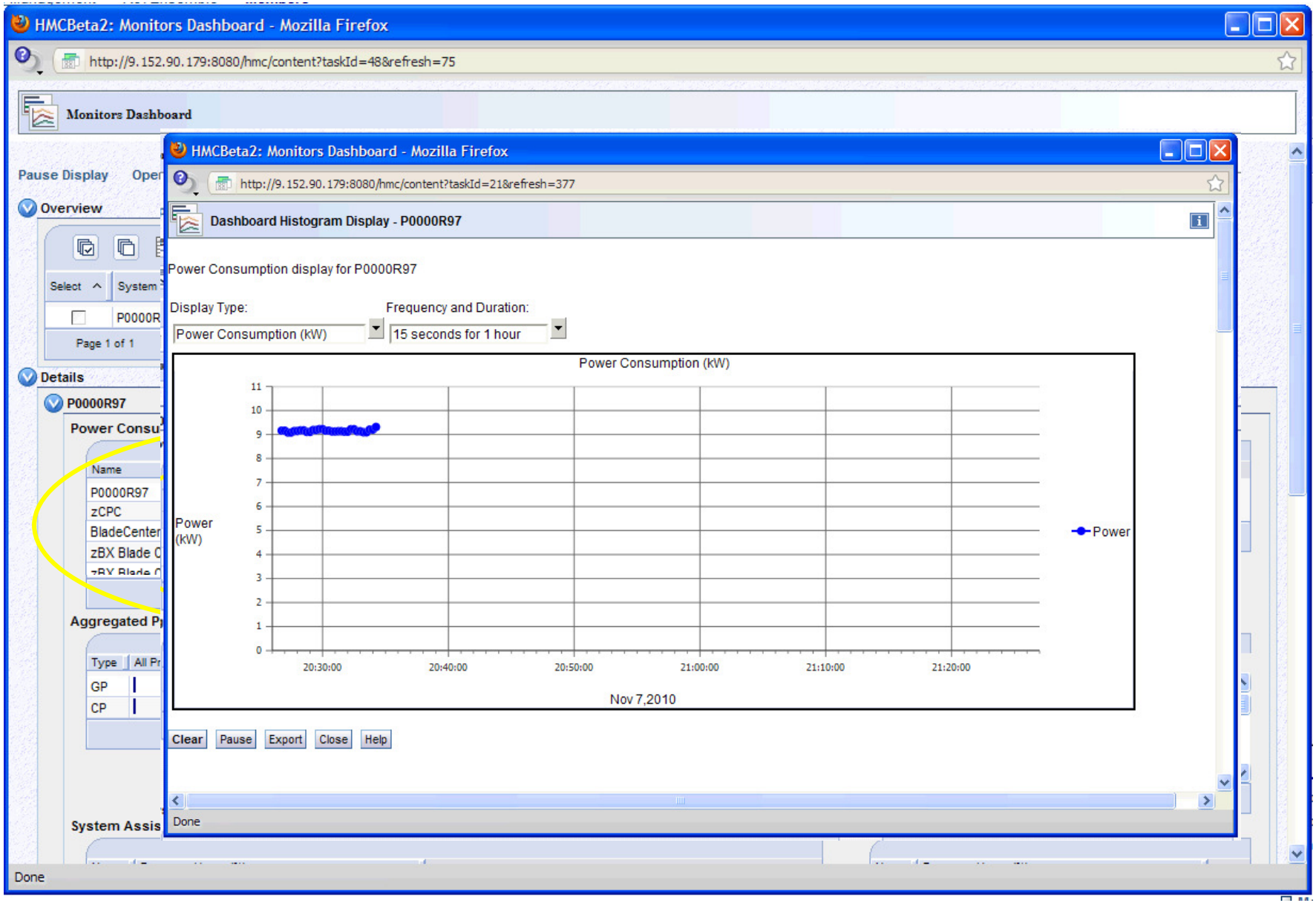
# Unified Resource Manager - Energy Monitoring Overview

**Monitoring data available at**

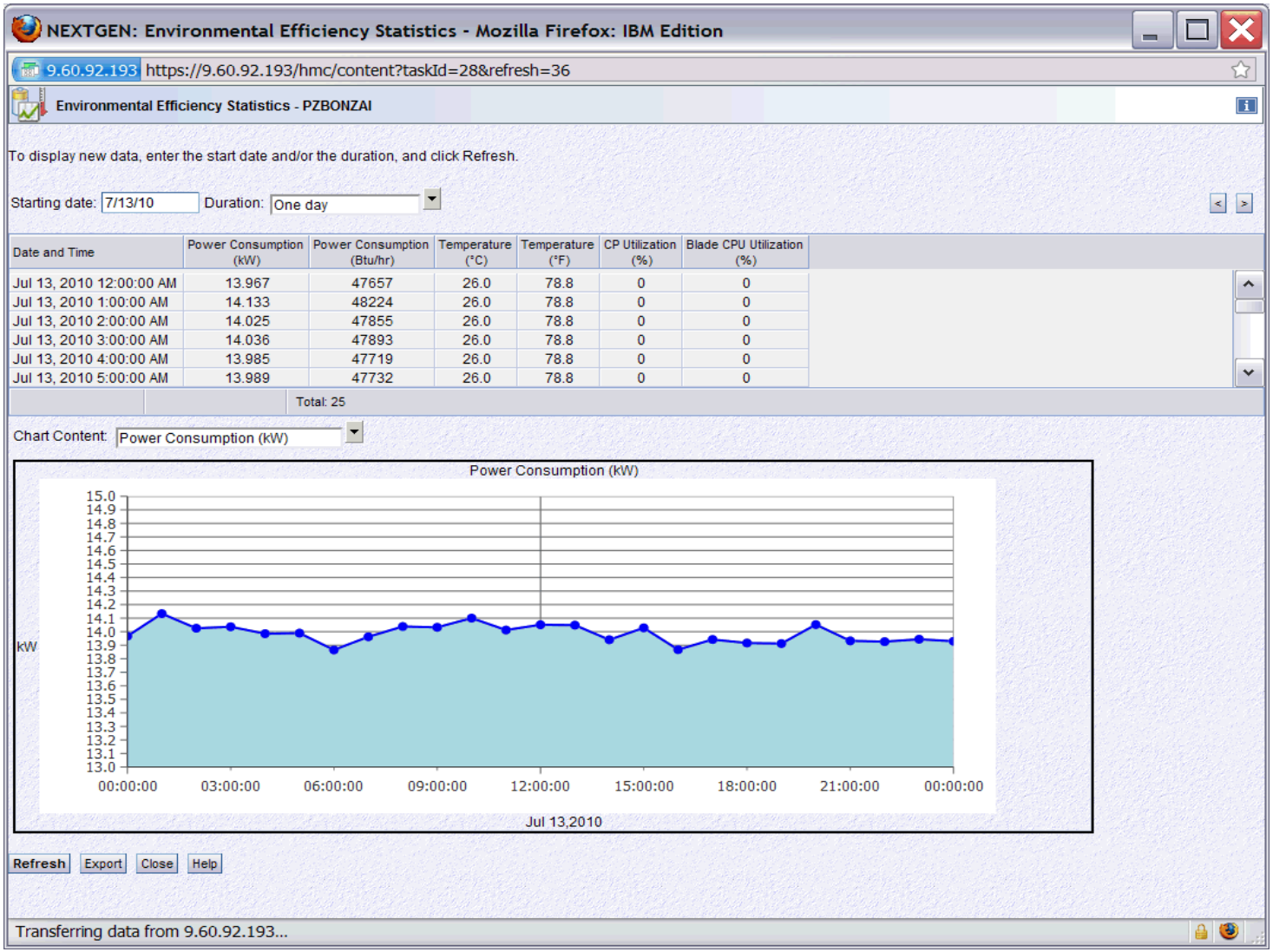
- Monitors Dashboard
- Environmental Efficiency Statistics
- Additional detailed data provided for
  - **Blade,**
    - Energy and environmental data
  - **BladeCenter**
    - Aggregated energy and environmental data
  - **zCPC**
    - Energy and environmental data
    - Max potential power
  - **CPC**
    - Aggregated energy and environmental data
  - **Ensemble**
    - Aggregated energy data



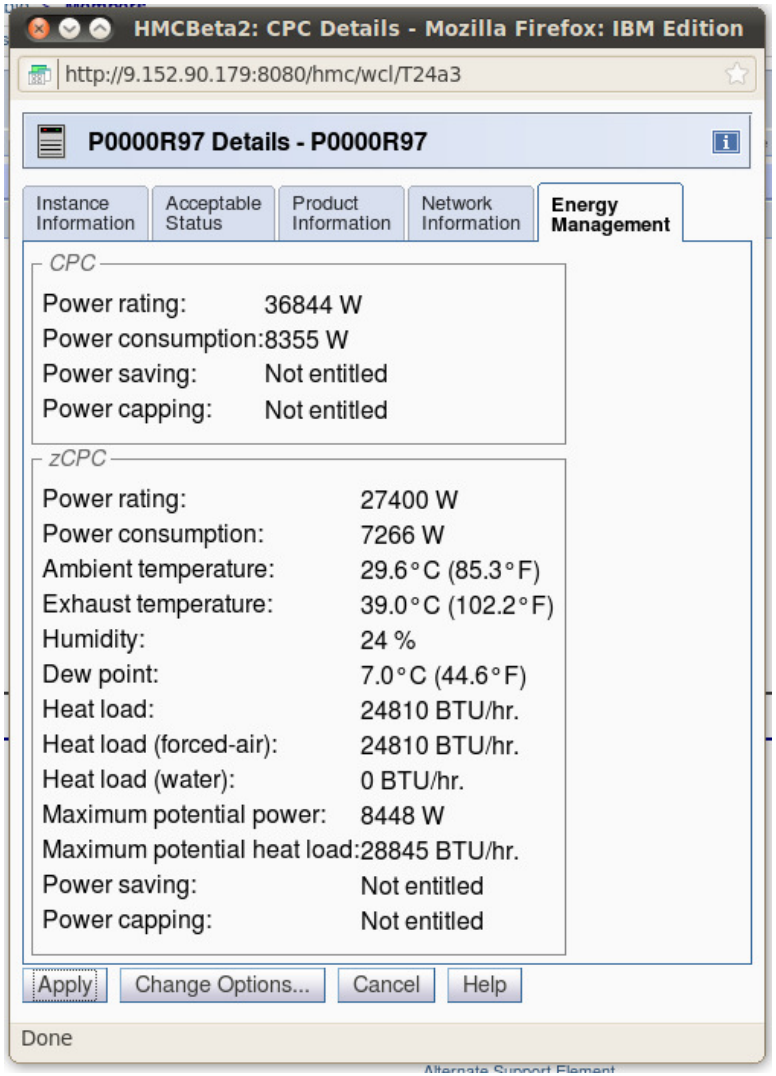
# zEnterprise Monitors Dashboard



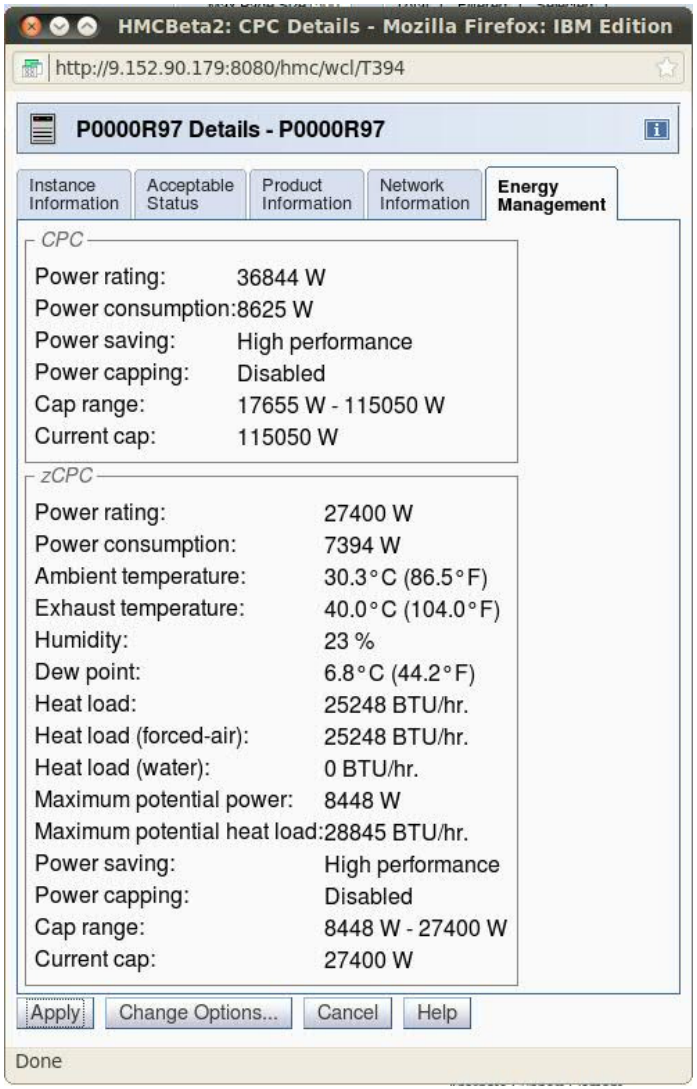
# Environmental Efficiency Statistics



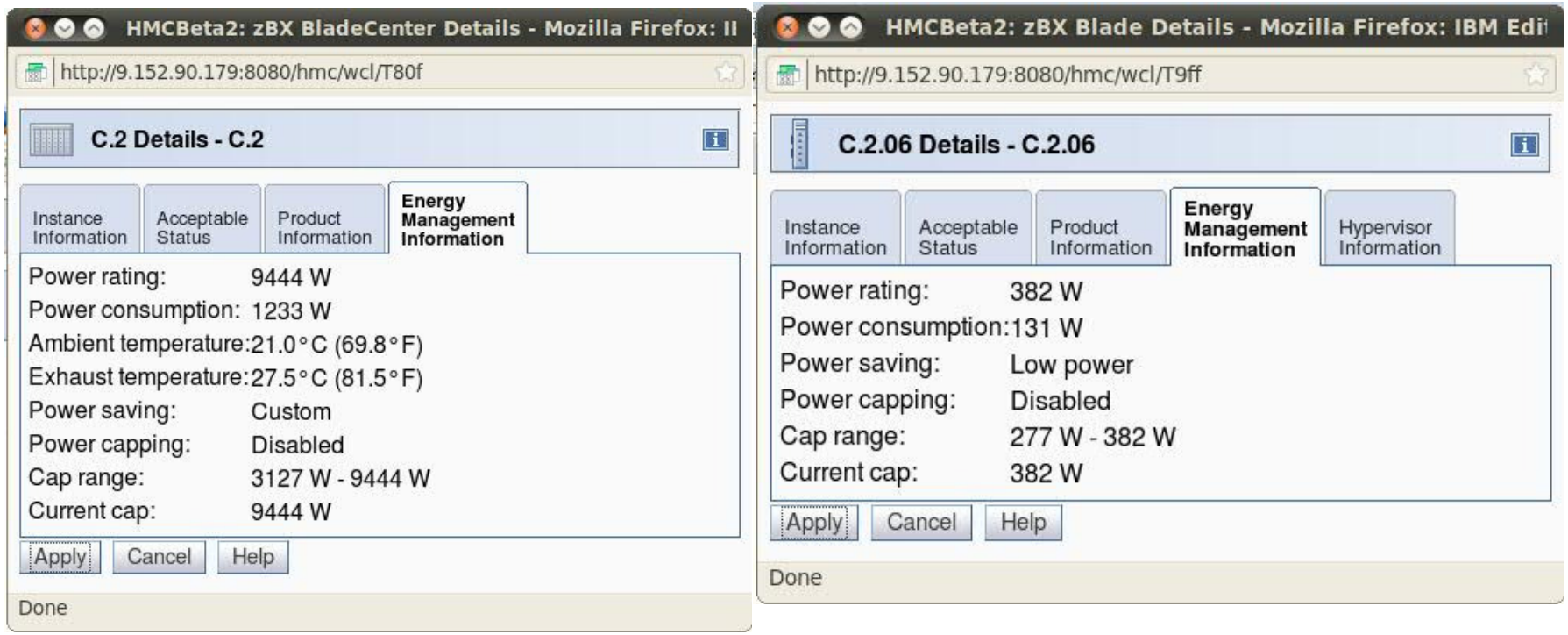
# Energy Management Information - CPC and zCPC



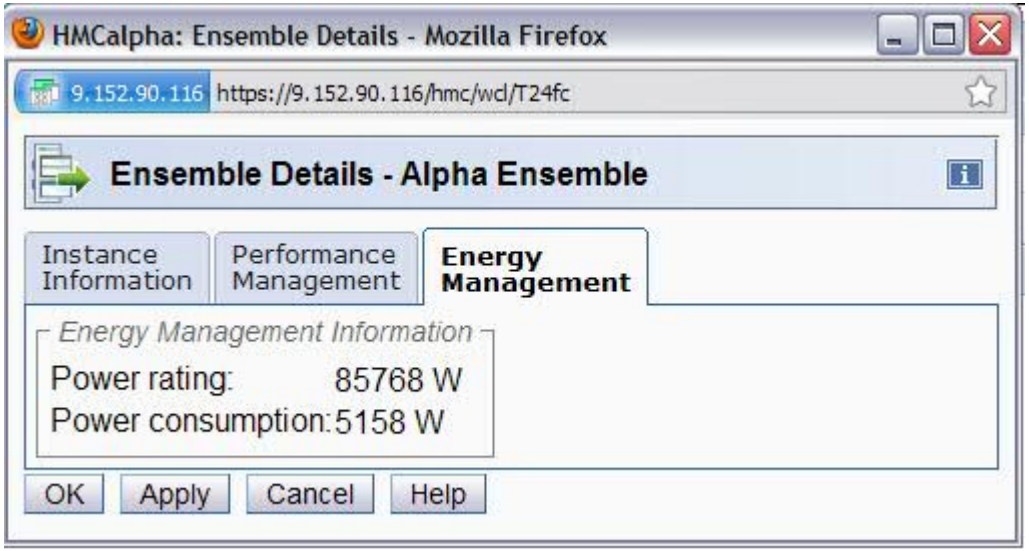
# Energy Management Information - CPC and zCPC



# Energy Management Information - BladeCenter and Blade

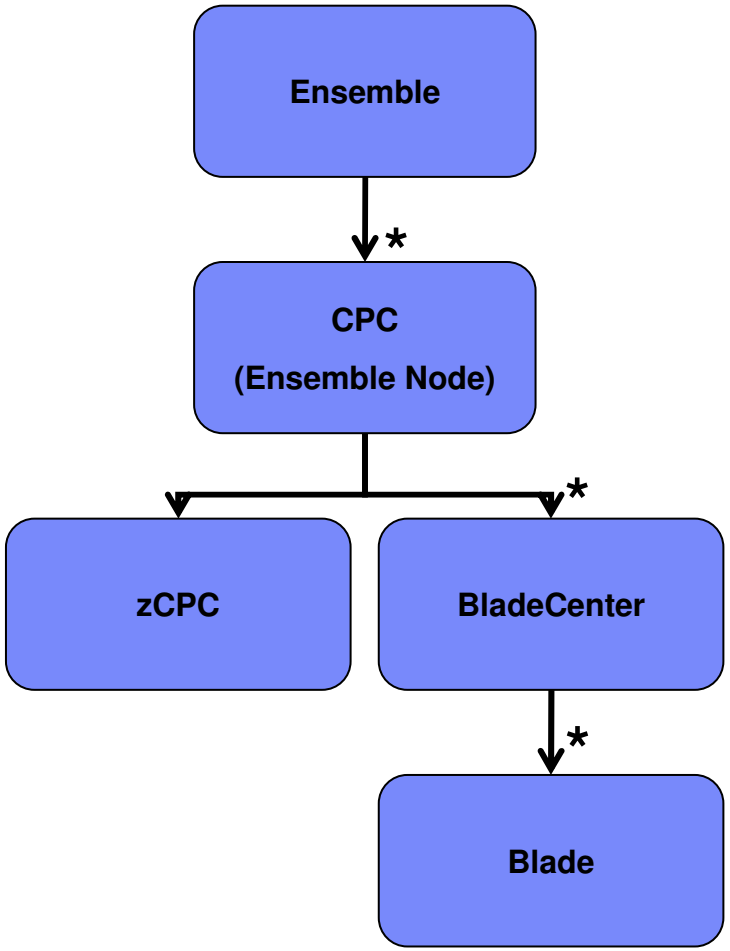


# Energy Management Information - Ensemble

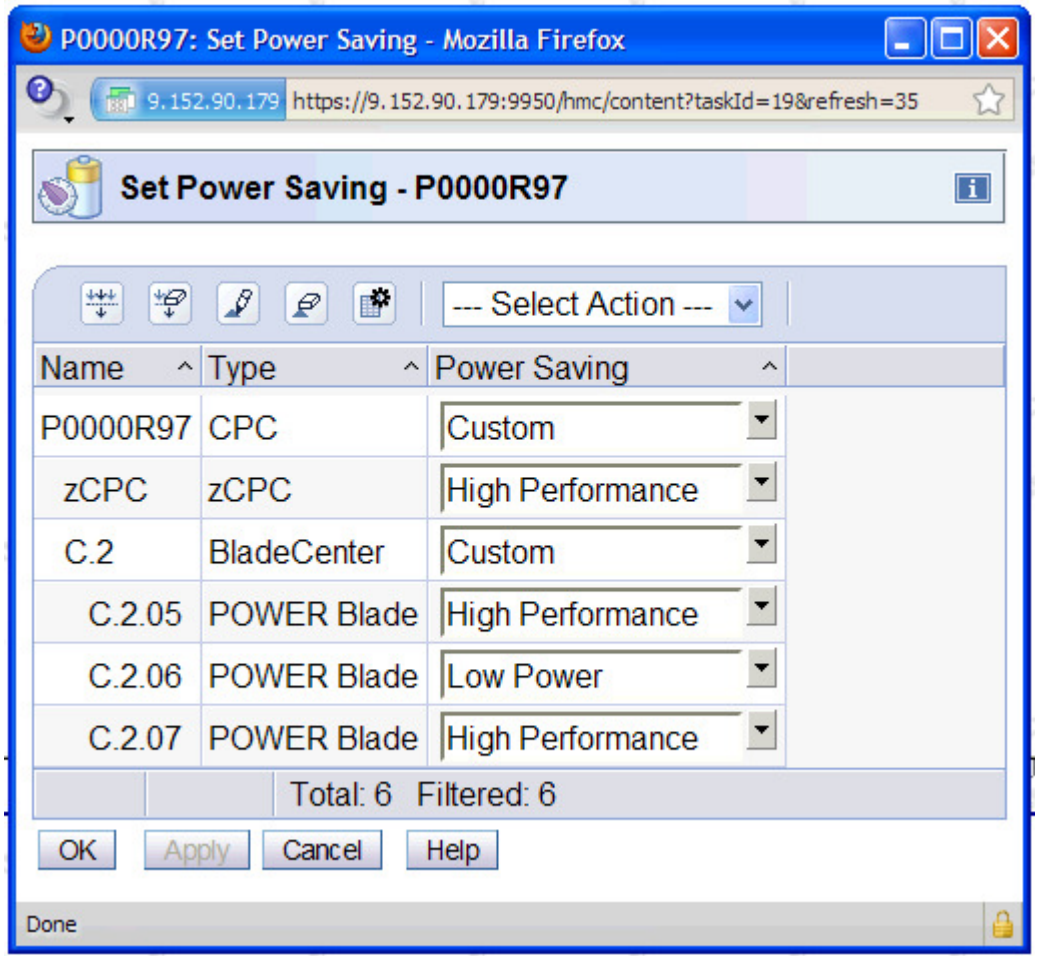


# Unified Resource Manager - Energy Controls Overview

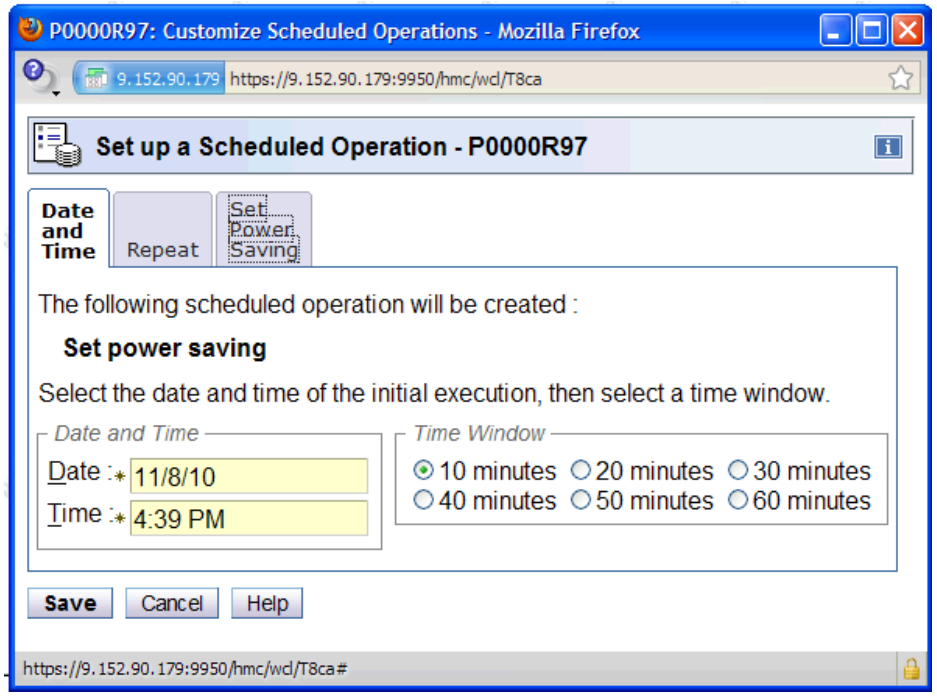
- **zCPC**
  - **Power Save**
- **Blade**
  - **Blade power save**
    for all blades supporting power savings mode
  - **Blade power cap**
- **BladeCenter**
  - **BladeCenter group power save**
    Ensure that all elements in a group (that support it) are in power save or high performance state.
  - **BladeCenter group power cap**
    Ensures that the group power consumption stays at or below the maximum value specified in the group cap using automatic power budget distribution.
- **CPC**
  - **CPC group power save**
  - **CPC group power cap**
    - Uses max potential power as Pcap-min for zCPC

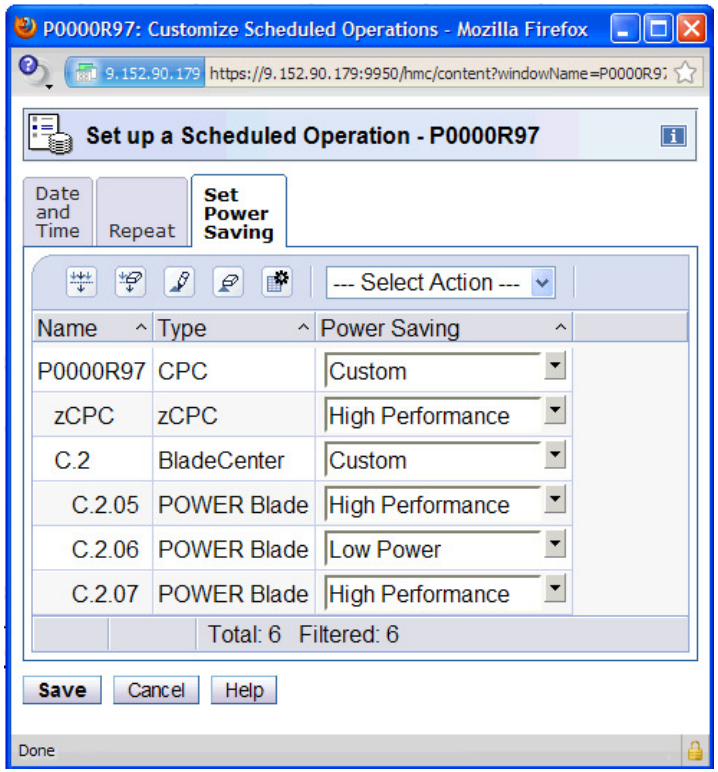


# Set Power Saving Task

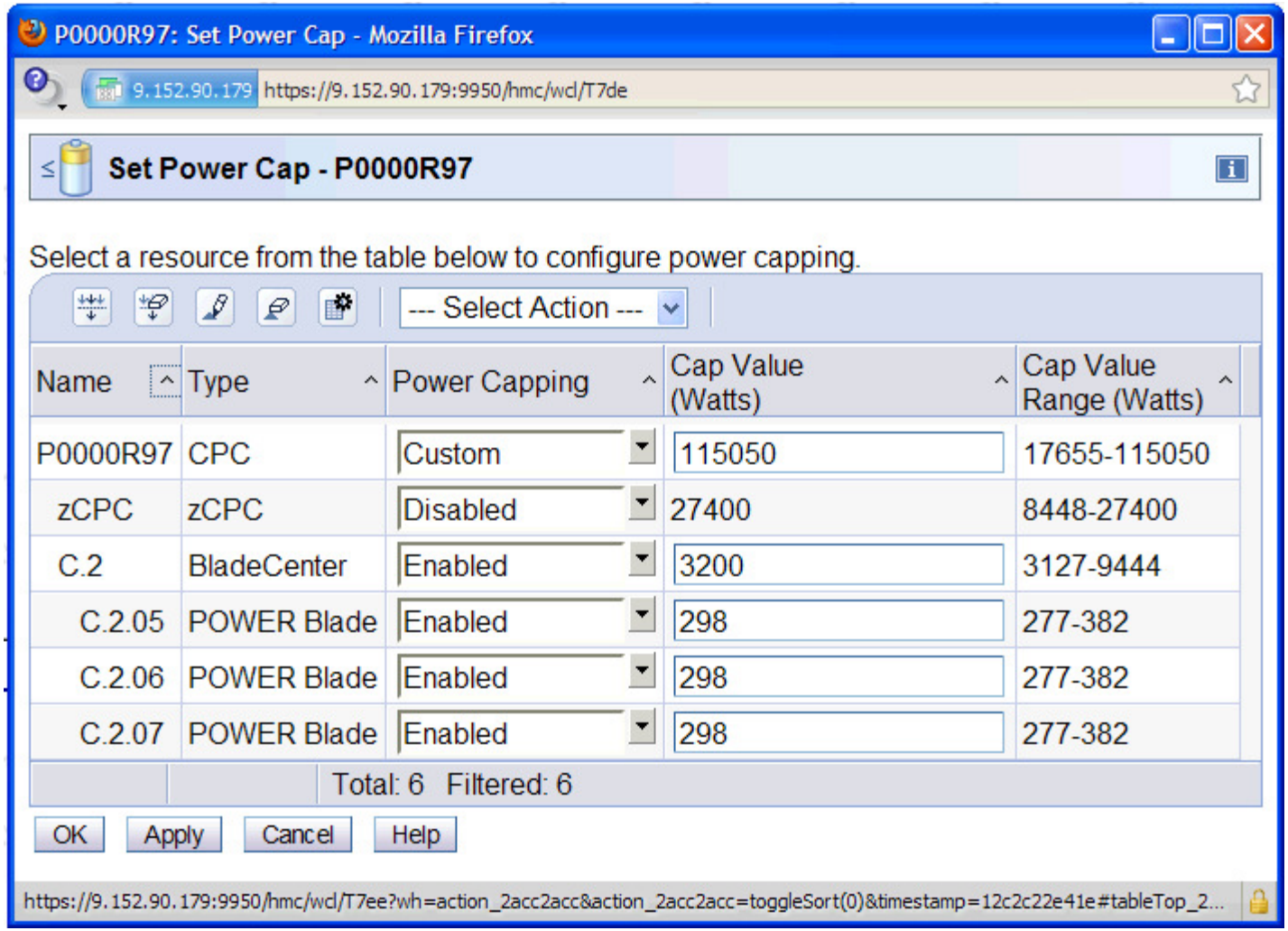


# Energy Management Automation

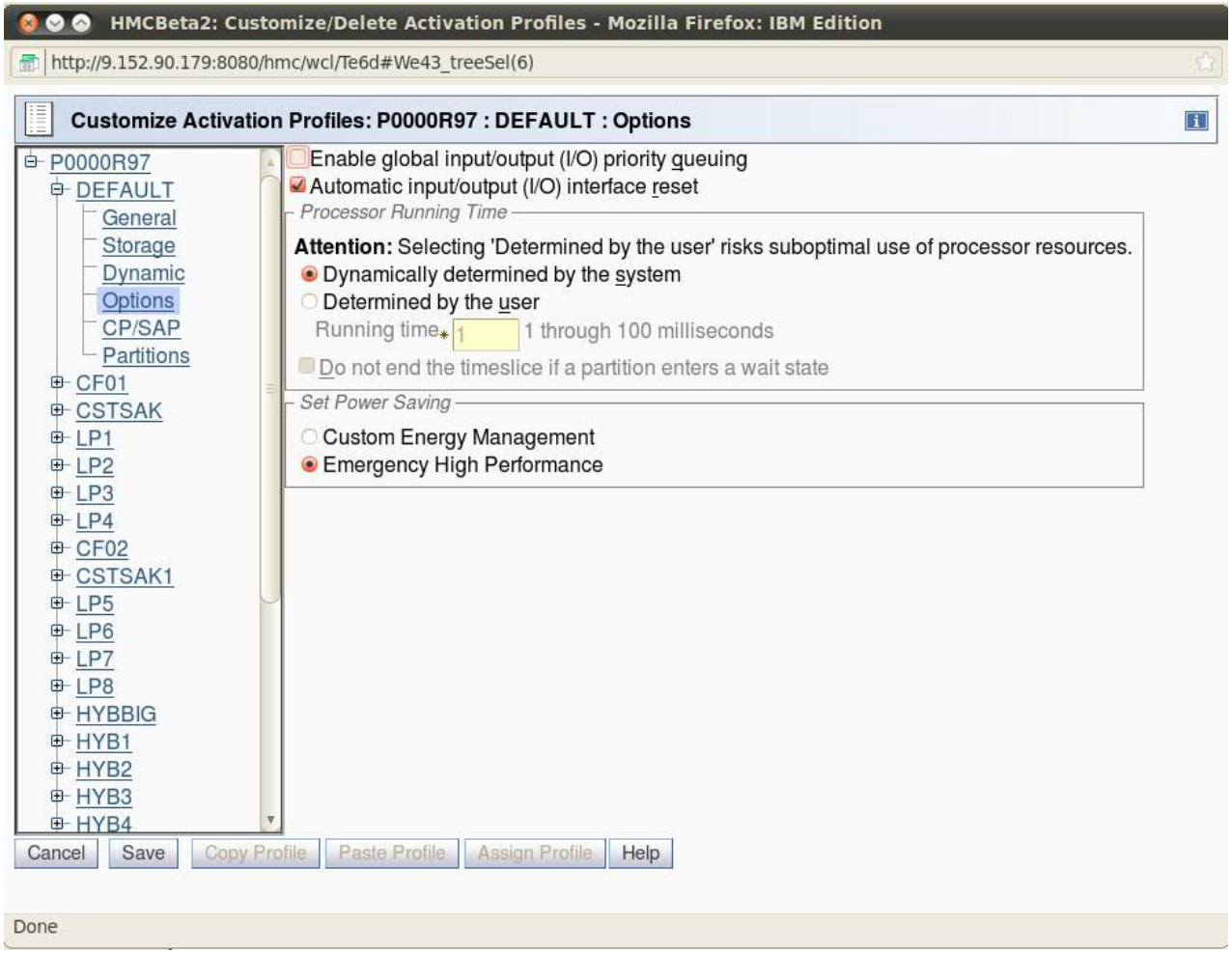


# Set Power Cap



# Set zCPC Power Saving Policy in the Activation Profile

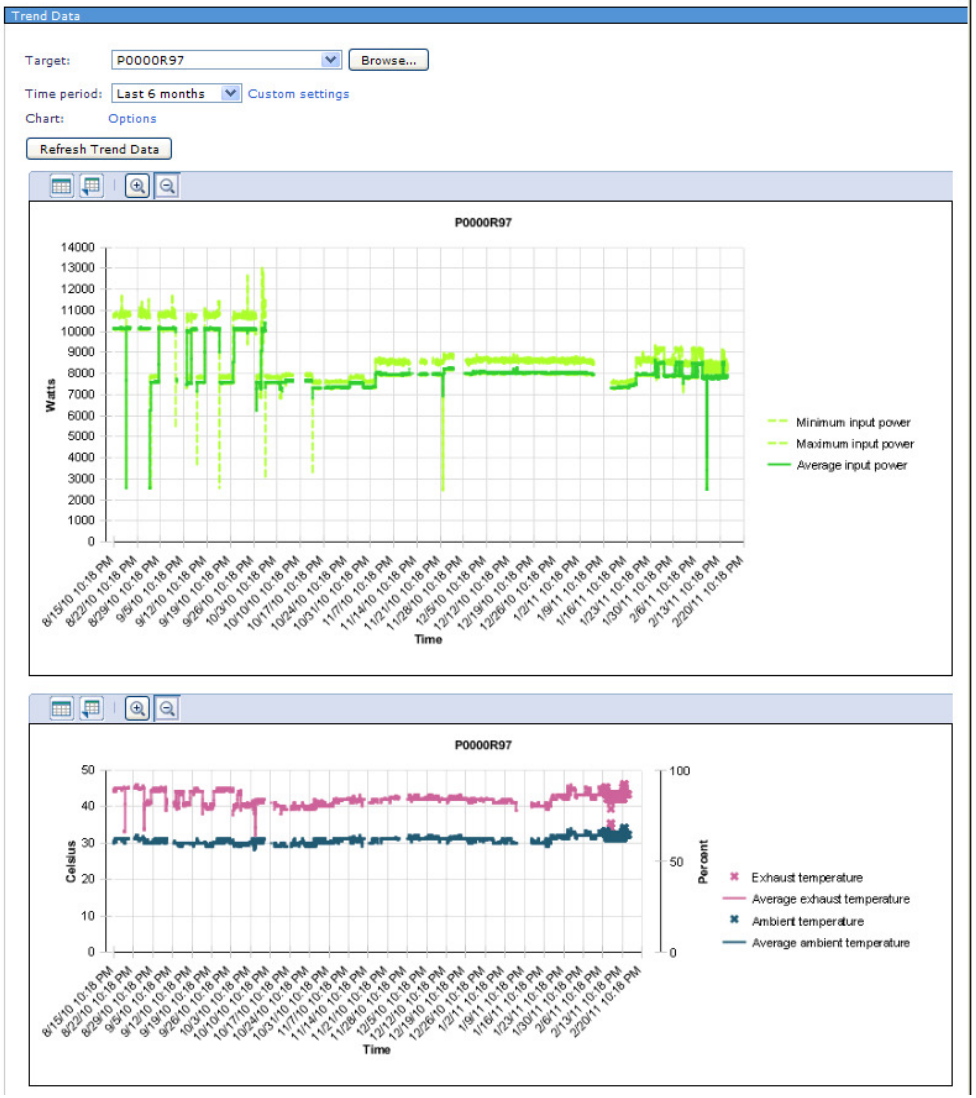


# Active Energy Manager Integration

- IBM System Director Active Energy Manager is an advanced energy manager provided through IBM Systems Director
- AEM monitors, measures and controls energy usage at the data center level
- Support across a large spectrum of IBM and non-IBM systems. System z support available since z10 GA1.
- AEM monitoring functions can be used free of charge.
- Enables to monitor System z in context of a heterogeneous data center.
- AEM 4.3.1 added Power Savings support for both zCEC and zBX

IBM® Systems Director

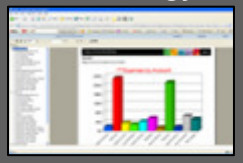
zEnterprise



# IBM Integrated Energy Management

## Tivoli energy management solution

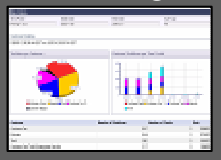
**Financial Accounting for Energy**



**Storage & Data Optimization**

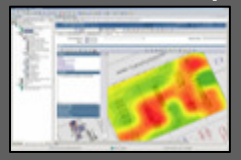
**Energy-Aware Provisioning and Scheduling**

**Energy Dashboard for Business Service Management**



**Optimize Energy Efficiency of Assets**

**Data Center Mapping and Thermal Maps**



## Tivoli Monitoring for Energy Management

Tivoli

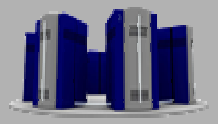
**Enterprise Data Repository**

**Enterprise Energy Optimization & Reporting**

**Enterprise Alerting for IT and Facilities**

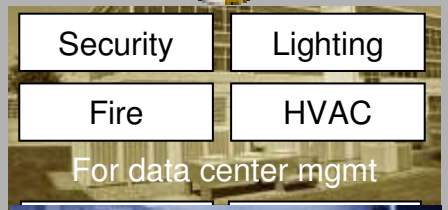
***IT Assets***

**Discover and Manage Non-IBM Systems**



***Enterprise Assets***

***Facility Infrastructure Assets***




## IBM Systems Director Active Energy Manager

IBM® Systems Director

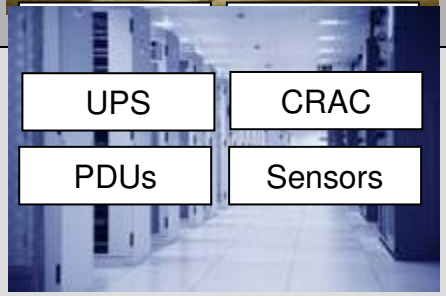
**Active Energy Management**

**Views, Alerting, & Reporting for IBM Systems**

***IT Assets***

**Discover and Manage IBM Systems**

***Data Center Infrastructure Assets***



# zEnterprise 196 – Energy Efficiency and Management Summary

- Significant improvements in energy efficiency
  - Tremendous performance improvement with same energy footprint
- Enables additional efficiency gains
  - Water cooling option
  - Overhead cabling option
  - HV DC power input option
- Energy Monitoring and Management delivered as part of Unified Resource Manager
  - Extensive monitoring of energy consumption and key environmental parameters
  - Integrated Energy Management Controls
- Integration into IBM Energy Management stack through Active Energy Manager


Session 8669