Platform™

Powering High Performance

Energy optimization of existing datacenters - Save the planet AND you budget!

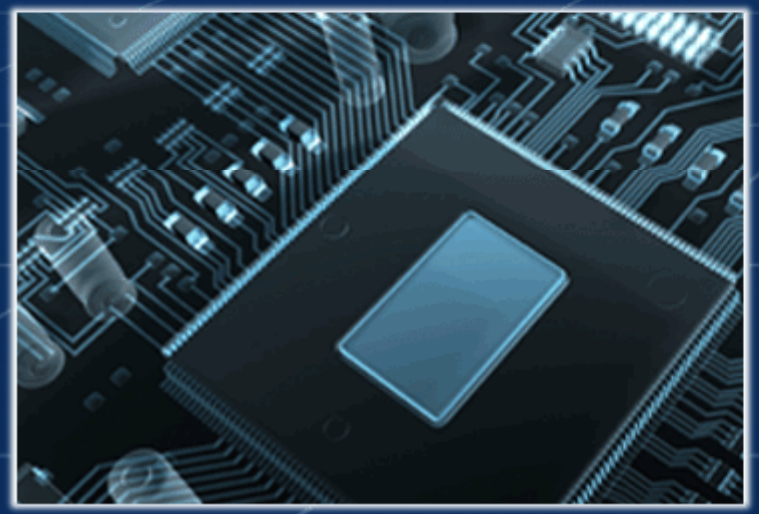
Bernhard Schott bschott@platform.com

OGF25 Catania, 5th March 2009

Agenda

- Energy consumption
  - Why do we care?
  - How much, How fast?
- Platform's path to a GreenHPC Data Center
  - Energy cost optimization methods
- Three easy steps to get on the "path"
  1. Green Workload management by Platform LSF
  2. Green Datacenter Daemon (NEW Product)
  3. Green Monitoring - Visualization of a greener DC
- ROI: Green pays off!
- Summary

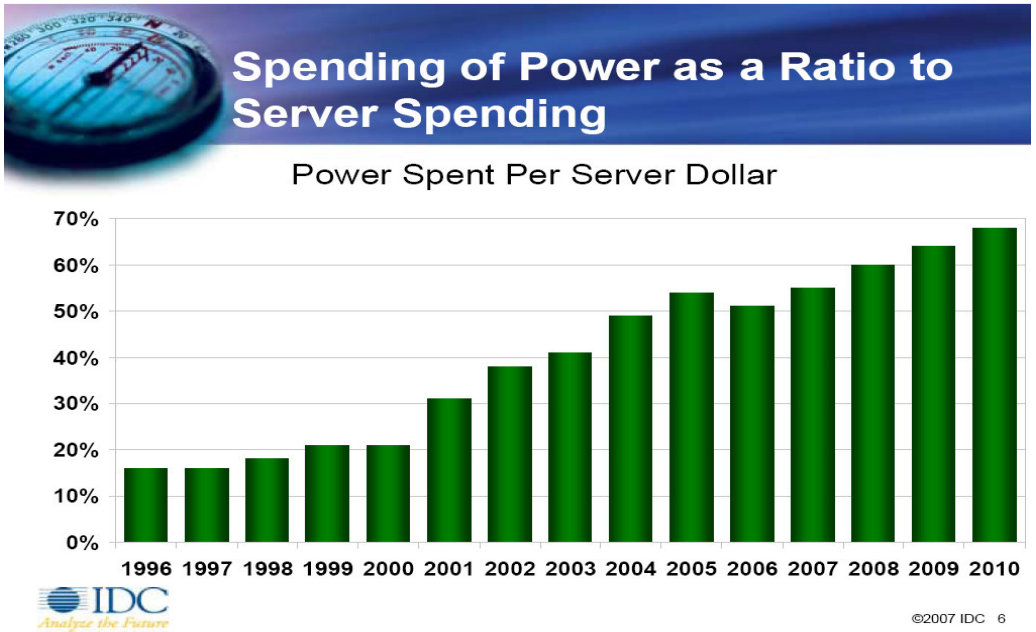
Platform™



Energy consumption – Why do we care?

About Green – What and why

- For data centers, energy cost is a big issue. IDC says
  - "**50¢ is spent to power & cool servers** for every $1 in server spending today; this will increase to 70¢ by 2010"
  - Facility power & cooling one of the top priorities for HPC data centers



- Values
  - Reduced power costs & equivalent carbon emissions
  - "Green" marketing opportunities

Energy Efficiency Importance: EU

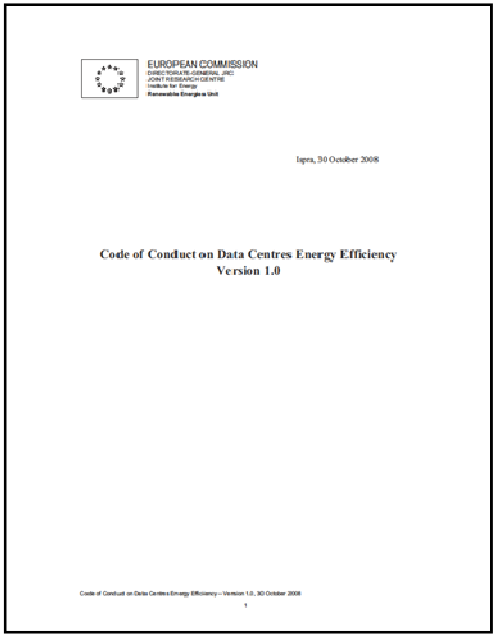
- *"Western European* electricity consumption of 56 TWh per year can be estimated for the year 2007 and is projected to increase to 104 TWh per year by 2020…"*

*commercial sector

- The European Commission started the CoC program in 2007

http://sunbird.jrc.it/energyefficiency/pdf/CoC%20data%20centres%20nov2008/CoC%20DC%20v%201.0%20FINAL.pdf

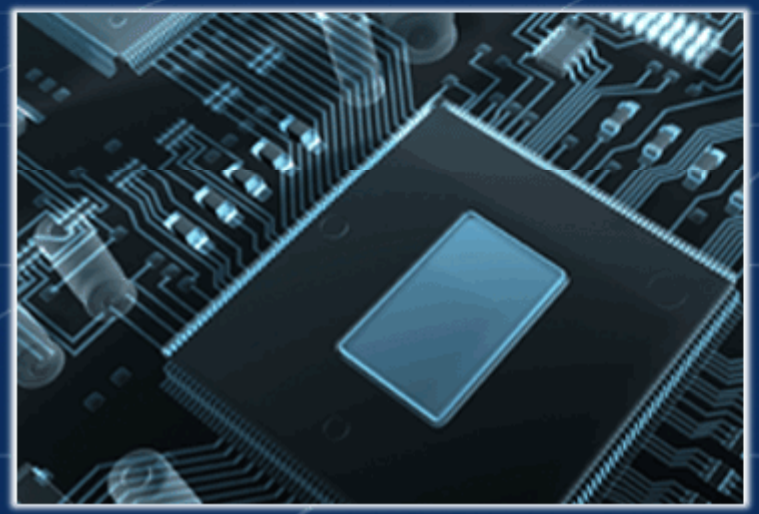
- The Code of Conduct on Data Centres describes steps to improve energy efficiency
- CoC compliance is voluntarily now, eventually becoming compulsory


EUROPEAN COMMISSION

Ispra, 30 October 2008

Code of Conduct on Data Centres Energy Efficiency
Version 1.0

Platform™



Tune your Data Center from Red to Green

Energy Consumption Optimization Path

Smart Consumption
– save energy/CO_2 AND money

- Understand and control thermodynamic properties of workload and datacenter

Intuitively switching off some machines may not be target leading to optimized power costs

- Arbitrarily selected machine may yield insufficient saving
- Increases job latencies by start-up time
- Some sites reported system (20%) and hardware (1%) failures after power cycling, requiring administrator intervention. No hard evidence that this would happen under controlled conditions.

  Please contribute your experience!

Energy Consumption Optimization Path

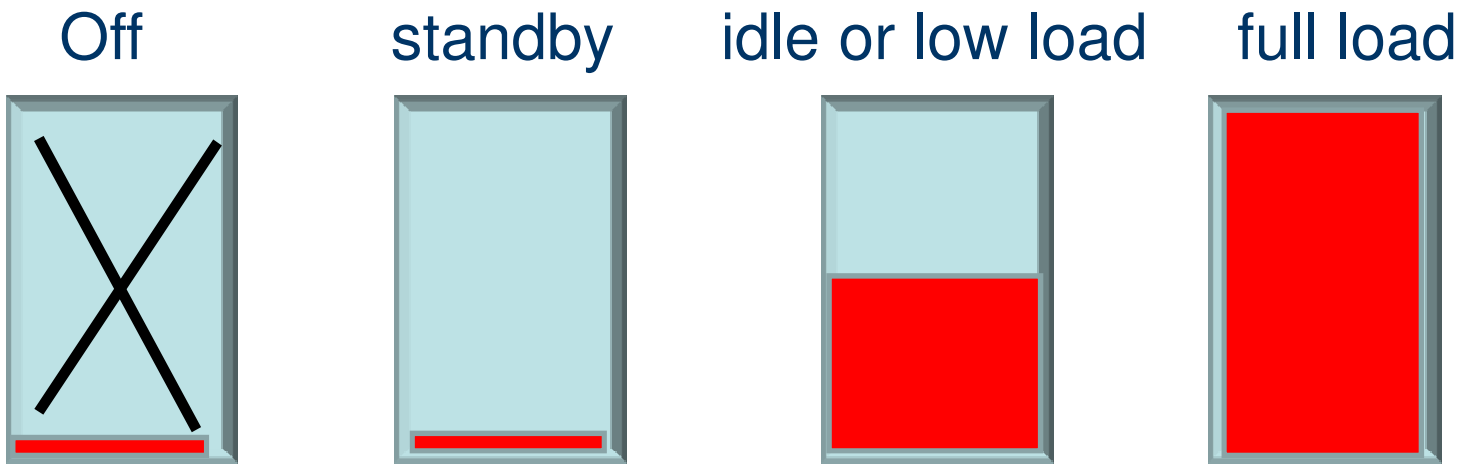
Platform LSF & GDD smart policies provide scheduling and control actions

- Use workload control before employing HW power down
- Profile applications with regards to energy consumption
- Profile HW with regards to energy consumption
- “Switch off” when appropriate to business demand policies

Energy consumption versus load

Energy consumption of hosts versus load

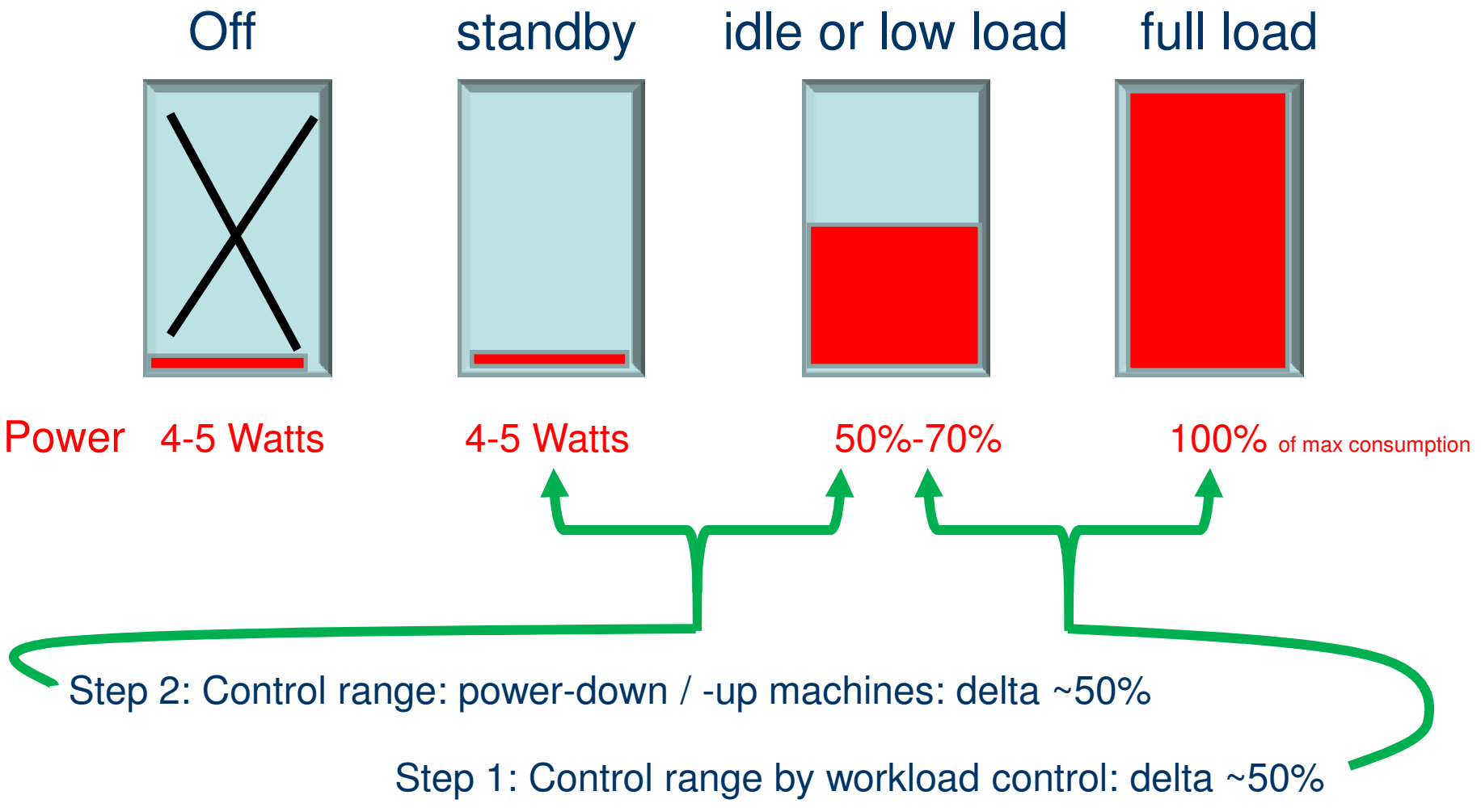
- Energy consumption of hosts vary according to their operational modes
  - Service type applications load hosts lowly
  - Computational type applications load hosts fully



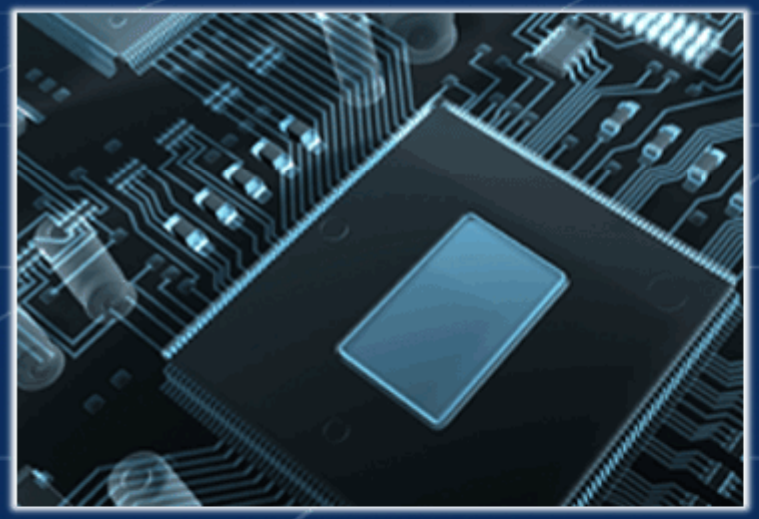
	Off	standby	idle or low load	full load
Power	4-5 Watts	4-5 Watts	50%-70%	100% of max consumption

Energy management control range

- Energy consumption of hosts versus load



Platform™



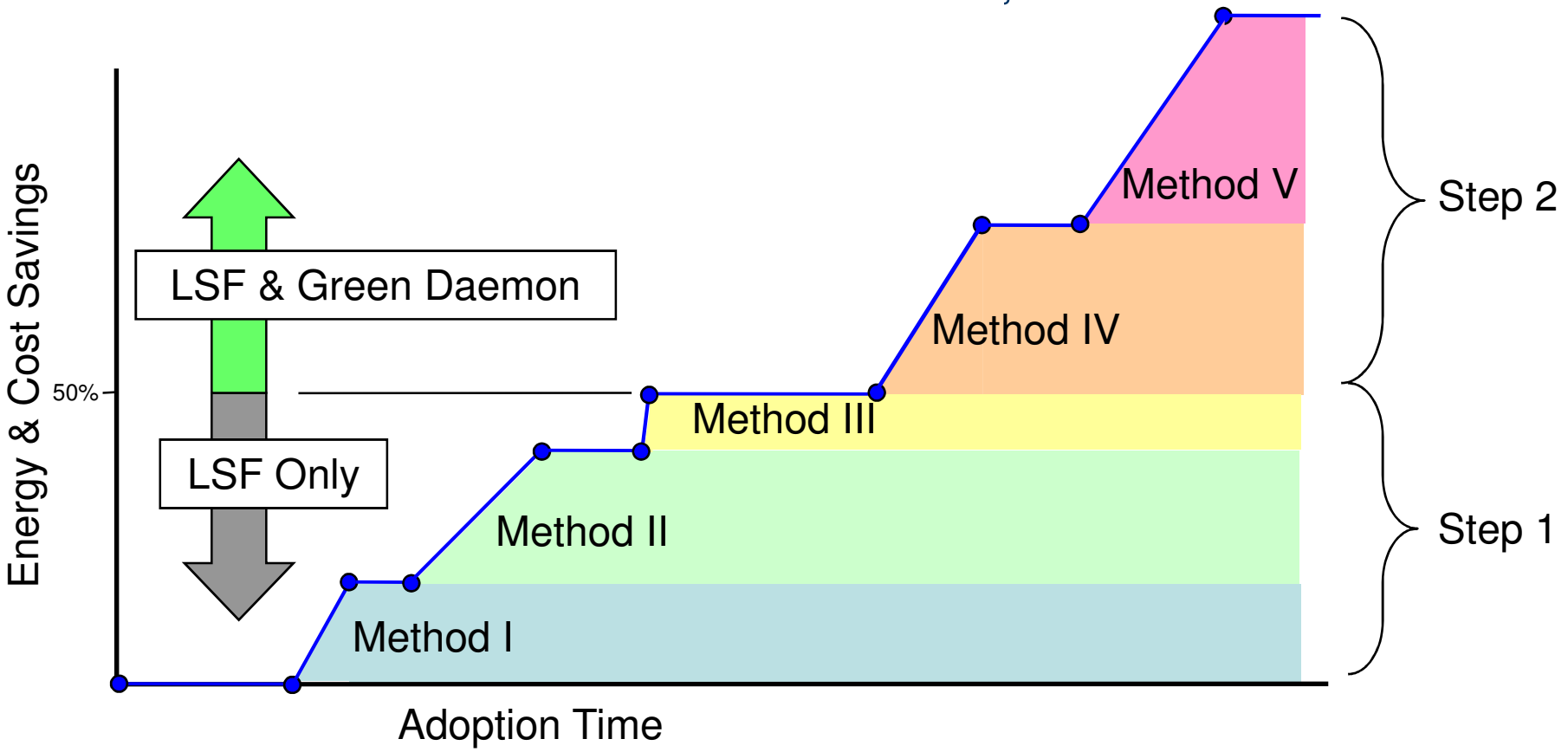
Three Steps to Green-HPC

3 Steps to Green-HPC

- Three easy steps to get on the “path”
  1. Green Workload management by Platform LSF
     - Applying Green workload management methods
  2. Green Datacenter Daemon (NEW Product)
     - Doubling the “Green” energy control range
  3. Green Monitoring - Visualization of a greener DC
     - Keeping in control – proofing the benefit
     - Reporting energy savings according to regulations

Energy Optimization Pathway

- Multiple independent methods
- No requirement for methods1-3 or 4-5 to be in order
- Recommended methods1-3 before 4,5



Platform Implementing Energy optimization

5 Independent Methods for Energy Cost Optimization

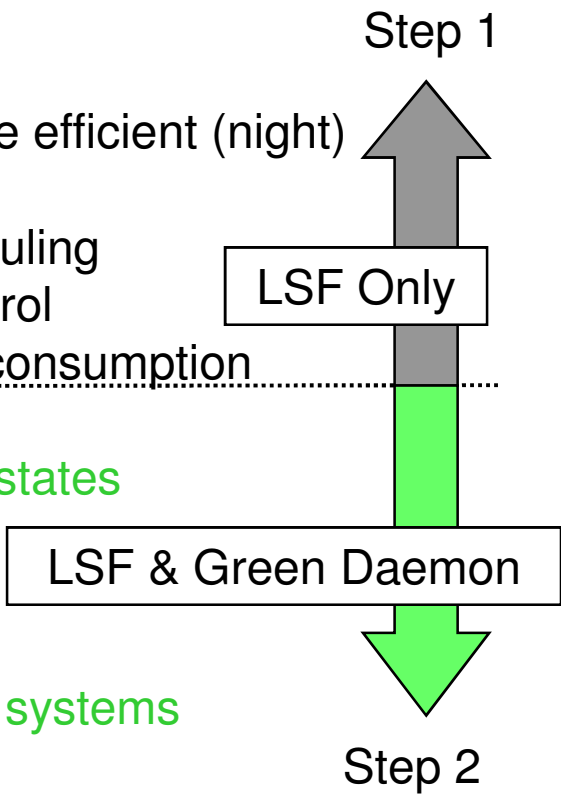
I. Energy Cost Optimization
- Use of energy according to supplier tariffs
- Shift of workload to low cost tariffs

II. Energy Efficiency Optimization
- Use most efficient host first
- Place workload (= generate heat) when cooling more efficient (night)

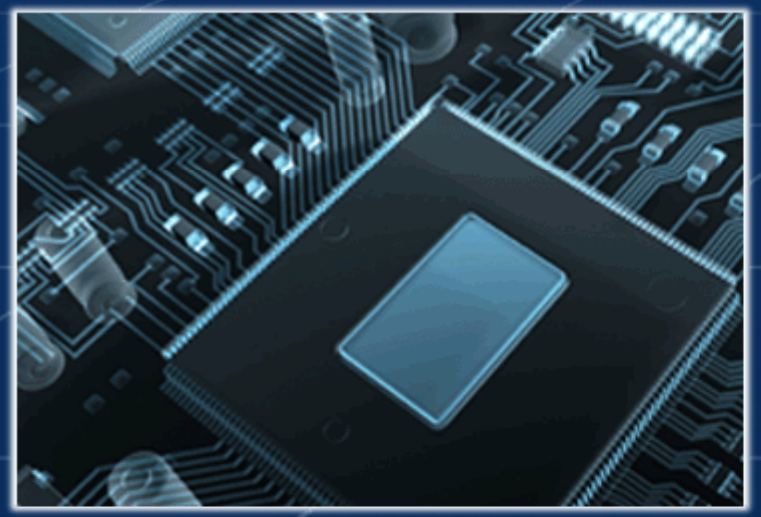
III. Hot Spot Control
- Use coldest host first / app. profile dependent scheduling
- Less cooling headroom needed due to hot spot control
- Reducing operational risk && lower cooling energy consumption

IV. Transient Energy & Performance Optimization
- Extends energy control range to hibernate and “off” states
- Full cost optimization of power for computing
- Balance performance versus energy savings

V. Full Thermodynamic Optimization
- Linking workload-, server power- and CRAC-control systems
- Dynamic CRAC parameterization
- Reducing cooling loads



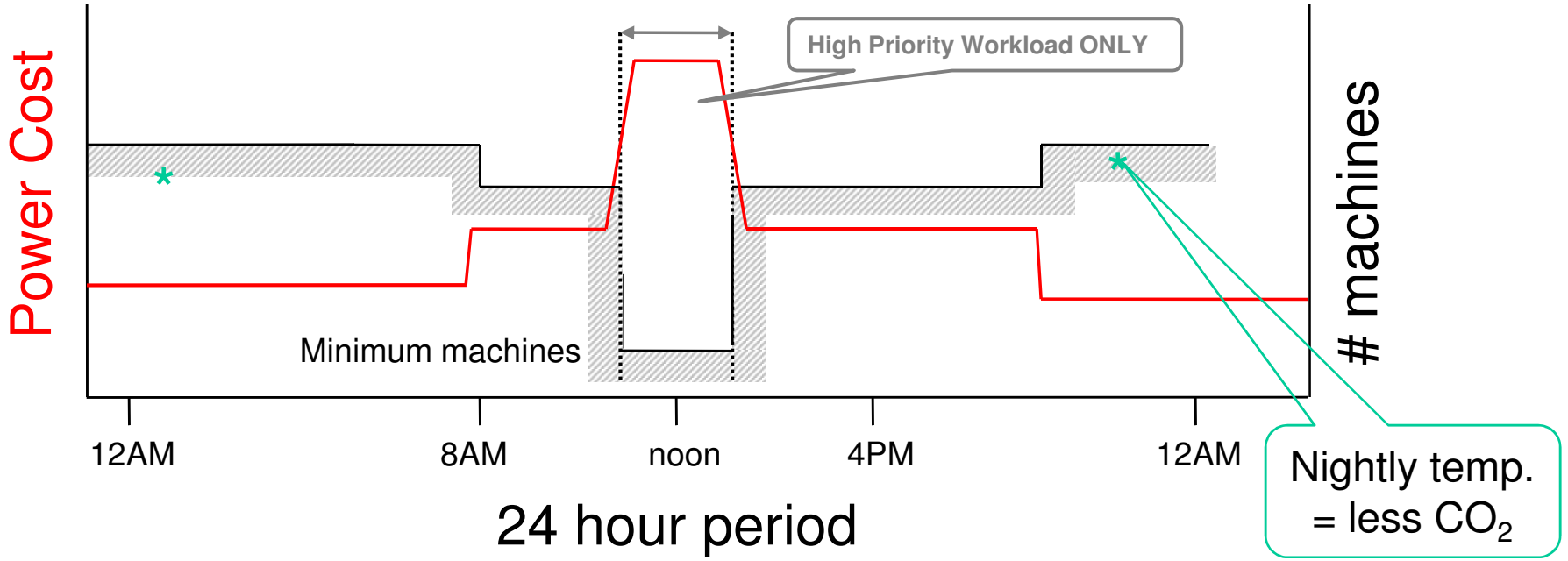
Platform™



Step 1:
Green Workload management by Platform LSF
-
Applying Green workload management methods

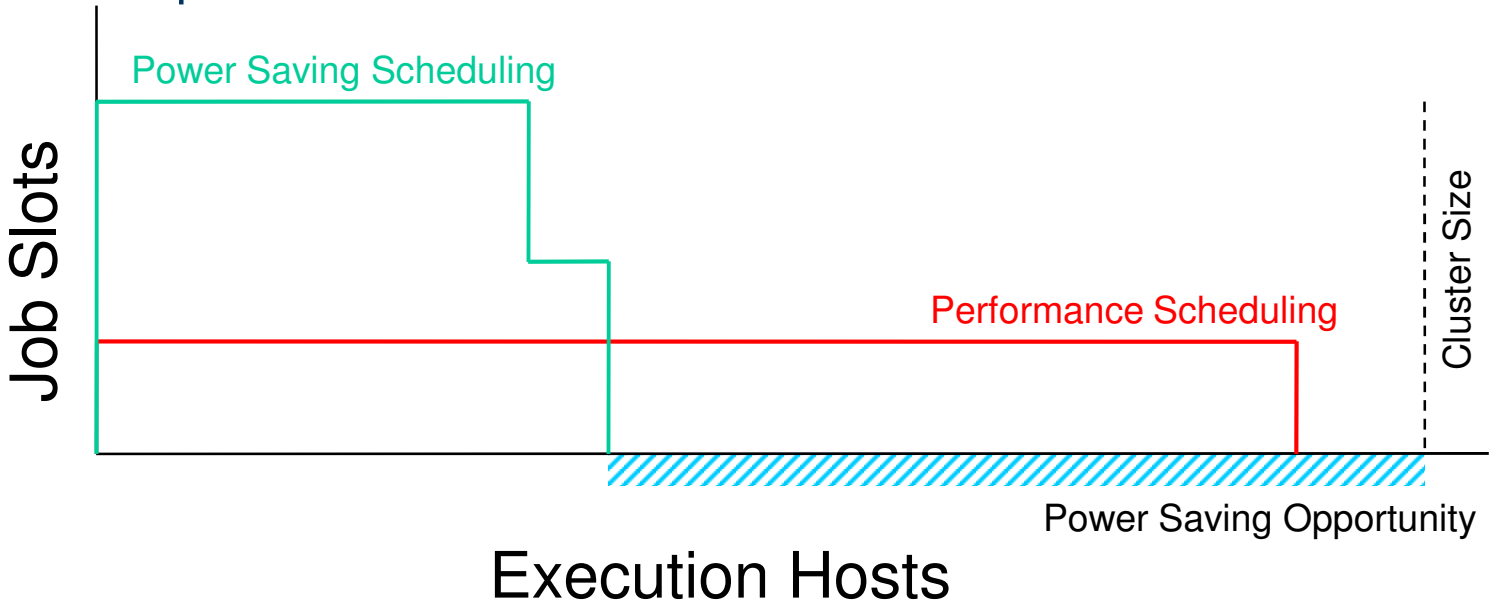
Schedule Power Requirements

- Time based configuration awareness
  - Low priority work to be scheduled when power is least expensive. Throughput can be higher during these times for no increase in power bill.
  - High priority work to run regardless of time of day
  - If workload is understood / predictable then provide an anticipatory level of min/max machines available



Platform Scheduling for Green-HPC

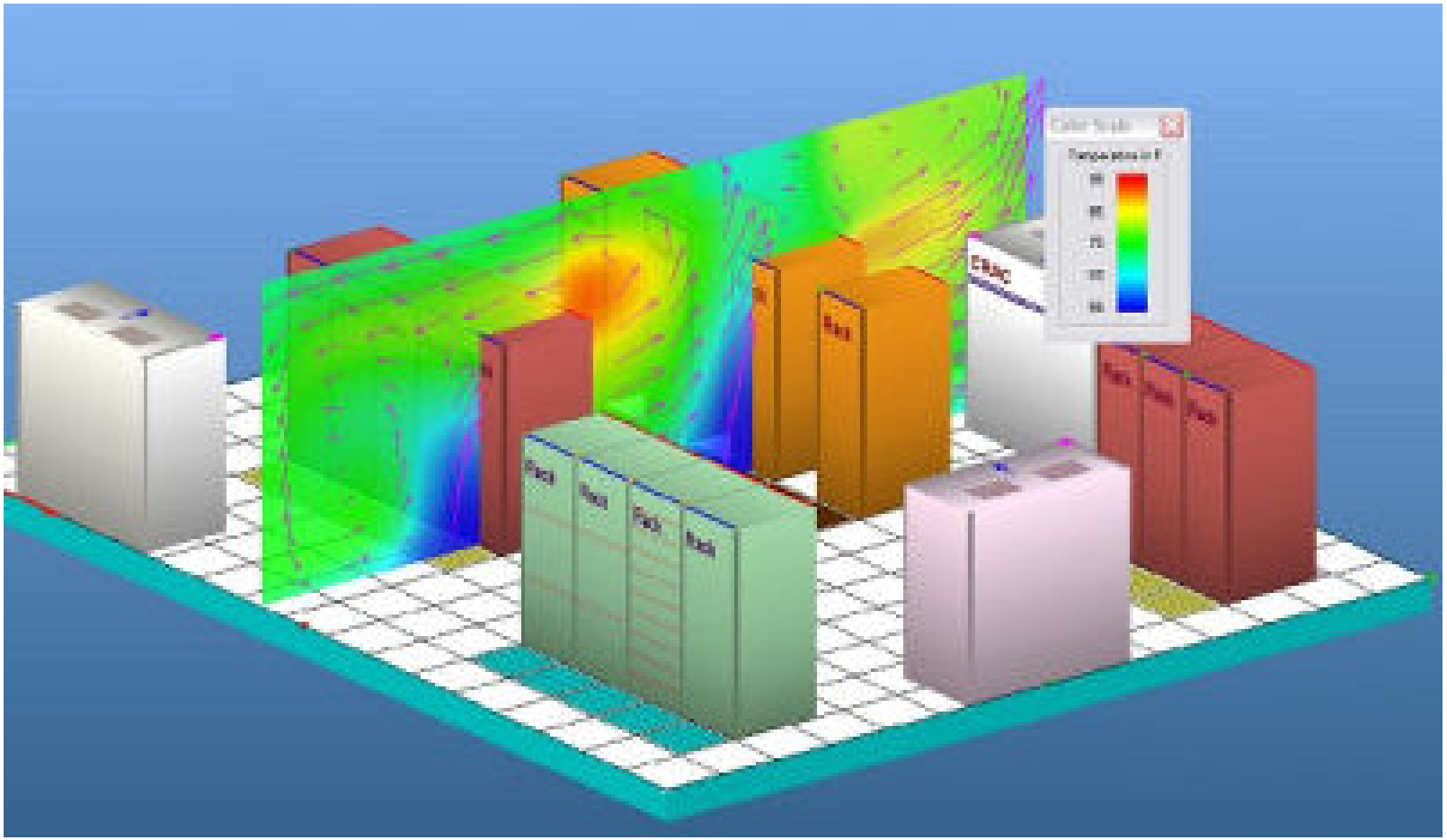
- Maximize opportunity to shutdown unused servers
  - Pack jobs to fill running servers
  - Select servers with higher job slot counts and coolest temperature



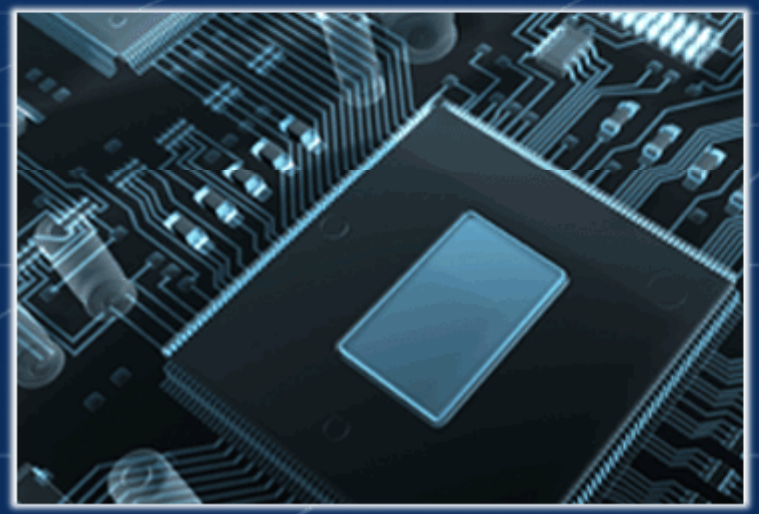
- But: I have zero free hosts in my cluster!
- So: it might save money to have more machines in order to switch some off. Compare to ROI spreadsheet!

Platform Spatial Scheduling

- Minimization of “hot spots” in a datacenter allows all CRAC systems to run at lower capacity



Platform™



Step 2:
Green Datacenter Daemon

Doubling the “Green” energy control range

Platform Green Datacenter Daemon

- Architecture
  - Daemon based solution
  - One Green Datacenter Daemon per LSF cluster
  - Standard LSF + configuration to support the goal
- Supported Hardware / OS
  - Intel and AMD hardware
    - Practically all HW vendors
  - Linux (kernel 2.6, glibc> 2.3), Windows 200x Server
- Features
  - iLO, IPMI, scriptable & reliable power ON/OFF method
  - Green Datacenter Daemon (GDD) runs on an LSF host and performs system queries, feeds metrics into LSF
- Optional
  - PMC (part of LSF) Installation for visualization

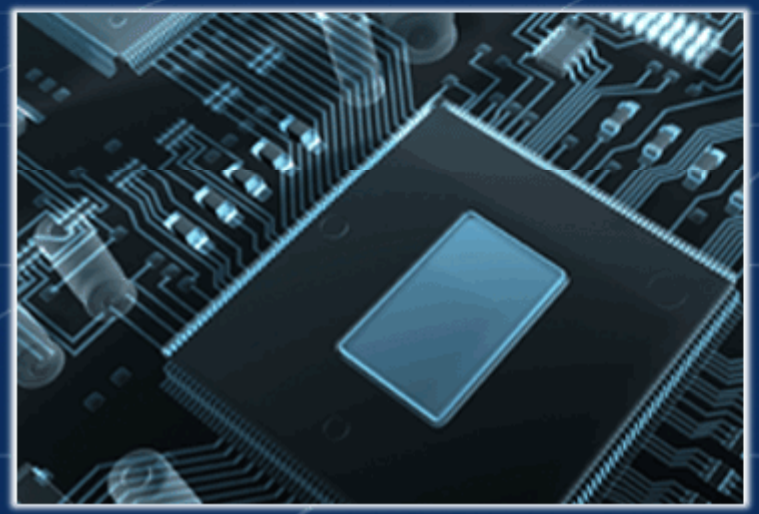
Green Datacenter Daemon

- Monitors LSF for job & host load
- Triggers power on/off actions based on runnable workload
- Estimates power consumption and savings over time
- Provides a CLI for status and manual control of power
  - Status: # of machines up / down, Current estimated power consumption, current savings per time
  - Control: power on/off specific hosts
  - Control: remove specific hosts from power control

What about MTBF decrease?

- Power cycling compute hardware might cause increased failures or require manual intervention
- What does Platform's GreenIT technology do to minimize the impact of this effect?
  - Power thrash prevention
    - Minimum duration of downtime and uptime per host
    - Minimum number of runable pending before power action
    - Maximum number of power cycles per server
    - Power action in groups of servers
  - Logging
    - If a server is sent a power ON signal but does not join the LSF cluster within a certain time, that server is labelled as requiring manual intervention and removed from power control

Platform™

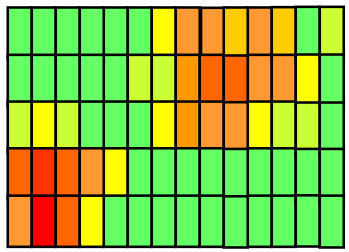


Step 3:
Green Monitoring - Visualization

Report energy savings according to Energy Star or CoC

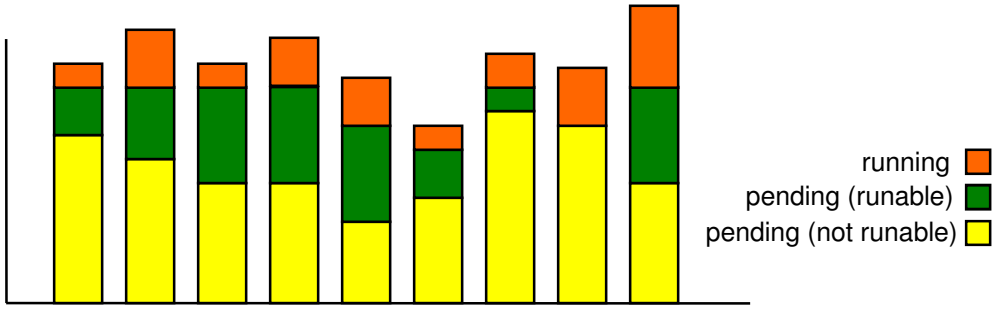
Platform Visualization

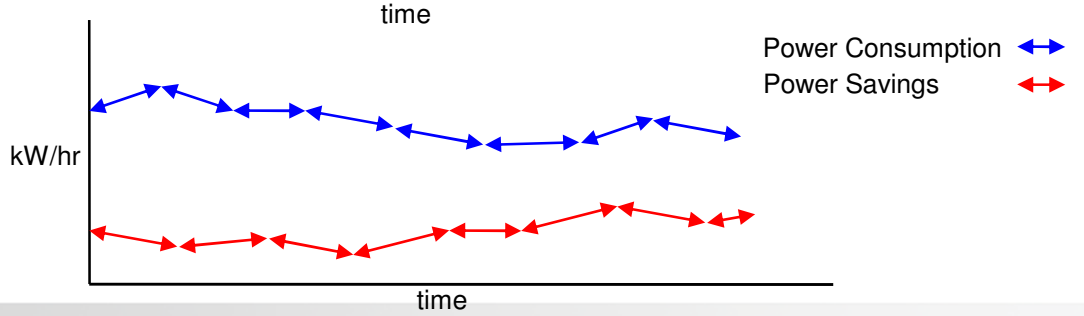
- How does an administrator know what is happening in relation to power on the cluster?
  - Platform GreenIT provides an interface for visualization
    - Hosts powered up/down
    - Pending jobs
    - Host temperature (datacenter wide, per rack)
    - Fan Speeds
    - Power consumption (kW/h)



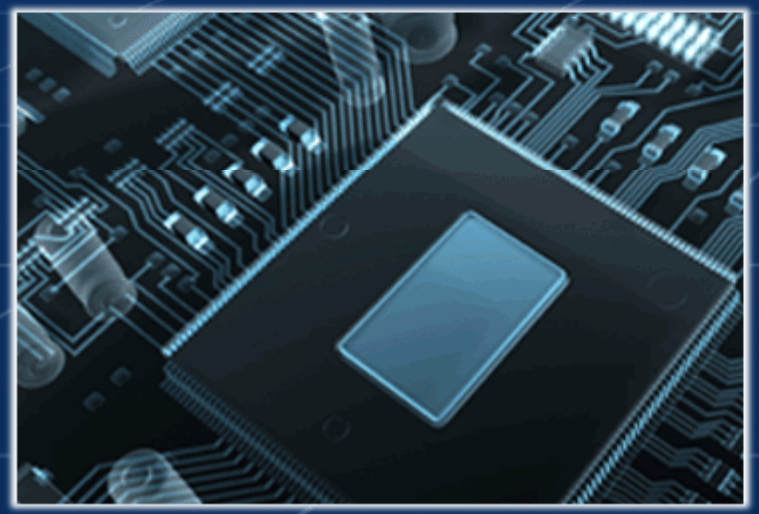
hosts

2D Datacenter temperature map





Platform™



Green pays off – do the math!

Energy ROI on GreenHPC

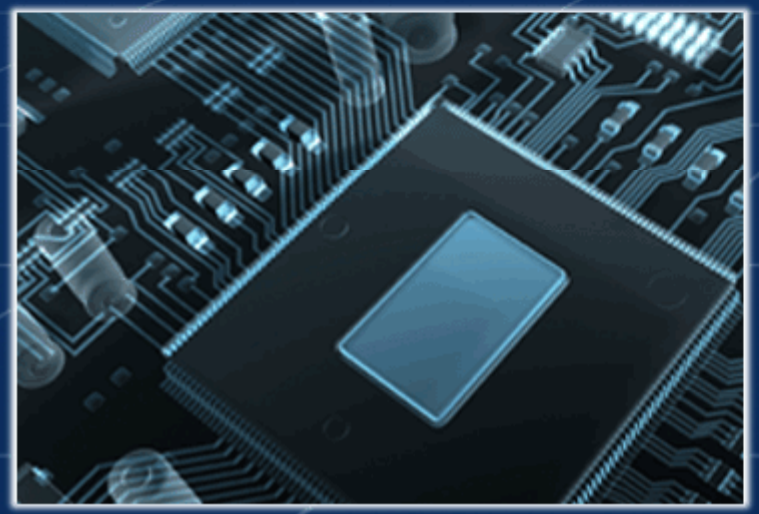
Do the Math: input data on servers, energy consumption and energy tariffs, adapted to your specific use case.

- Power consumption of (compute cluster) servers x runtime
- Add overhead by cooling, UPS, PDU, etc.
  - "cooling & losses" 200% of "direct power". Your data center may be better or worse – so adjust this value in the spreadsheet accordingly!

	Power Consumption				Energy costs time dependent tariff					Save energy costs by tariffs			Reduced energy consumption			
					h / day	night tariff h	day tariff h	peak tariff h			workload shift from		savings on cooling		total savings	
	cooling & losses* ratio		200%		24	8	14	2	total	save on peak	peak to day	day to night	raise intake temp*	shift to night*	EUR	%
						night tariff €	day tariff €	peak tariff €	average	3%	6%	6%	6%	6%		
	server	cooling & losses*	total	h	kWh	0,08 €	0,14 €	0,32 €	0,14 €							
per day	333W	667W	1,0kW	24	24	0,64 €	1,96 €	0,64 €	3,24 €	0,02 €	0,02 €	0,05 €	0,19 €	0,18 €		
			week													
per week	333W	667W	1,0kW	168	168	4,48 €	13,72 €	4,48 €	22,68 €	0,13 €	0,15 €	0,35 €	1,32 €	1,24 €		
			month													
per month	333W	667W	1,0kW	732	732	19,52 €	59,78 €	19,52 €	98,82 €	0,59 €	0,66 €	1,54 €	5,76 €	5,42 €		
# servers /	per year		year													
1	333W	667W	1,0kW	8760	8.760	233,60 €	715,40 €	233,60 €	1.183 €	7,01 €	7,88 €	18,40 €	68,96 €	64,82 €	167,07 €	14,13%
10	3.333W	6.667W	10kW	8760	87.600	2.336 €	7.154 €	2.336 €	11.826 €	70 €	79 €	184 €	689,59 €	648,21 €	1.670,68 €	14,13%
100	33.333W	66.667W	100kW	8760	876.000	23.360 €	71.540 €	23.360 €	118.260 €	701 €	788 €	1.840 €	6.895,87 €	6.482,12 €	16.706,79 €	14,13%
1000	333.333W	666.667W	1.000kW	8760	8.760.000	233.600 €	715.400 €	233.600 €	1.182.600 €	7.008 €	7.884 €	18.396 €	68.958,72 €	64.821,20 €	167.067,92 €	14,13%
6000	2.000.000W	4.000.000W	6.000kW	8760	52.560.000	1.401.600 €	4.292.400 €	1.401.600 €	7.095.600 €	42.048 €	47.304 €	110.376 €	413.752,32 €	388.927,18 €	1.002.407,50 €	14,13%

- Enter your energy tariffs: night, day, peak, per day, special rates/period, flat,..
- Current total energy price calculated should (about) match energy bill

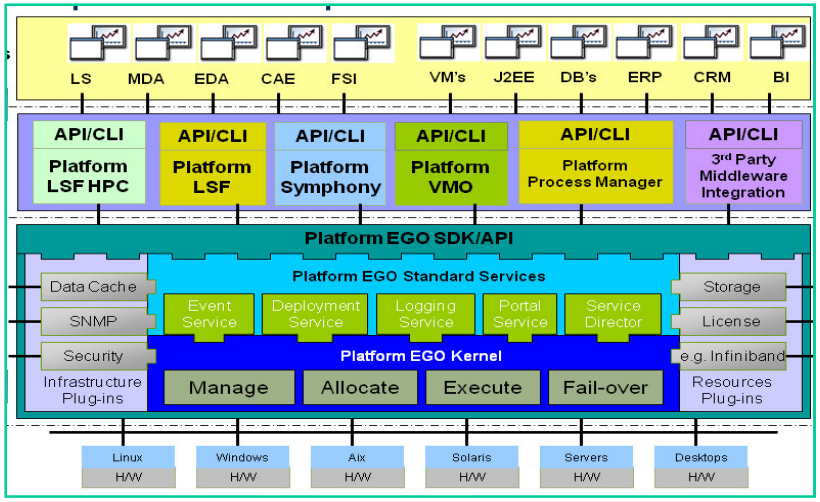
Platform™



Summary

Energy optimization of existing datacenters

- Smart workload management AND HW power management yield maximum consumption control range
- Platform Computing Grid Technology supports Energy and CO_2 reduction without major investments.
- Planned research: agile thermodynamic handling of datacenters (collaboration with FZK, R.Berlich, M.Kunze)
- Ongoing: Storage power management integration.
- Need: Improve IPMI and related standards *(implementation!)*
- Looking forward to discuss and share with you!



Bernhard Schott
Dipl. Phys.
EU-Research Program Manager
Platform Computing GmbH

Mobile +49 (0) 171 6915 405
Email: bschott@platform.com
Skype: bernhard_schott
Web: http://www.platform.com/