Metrics for HPC Data Center Power Proportionality and Efficiency

Natalie Bates

The First Workshop on HPC Power Management: Measuring Effectiveness

22 September, 2015

Learn from each other

What makes a good metric?

- It matters.
- But, metrics can have limitations.

“I don't see a real value in PUE.”

Gordon Lane, Facilities Coordinator at Petro Canada
LinkedIn Discussion on Power Usage Effectiveness (PUE)

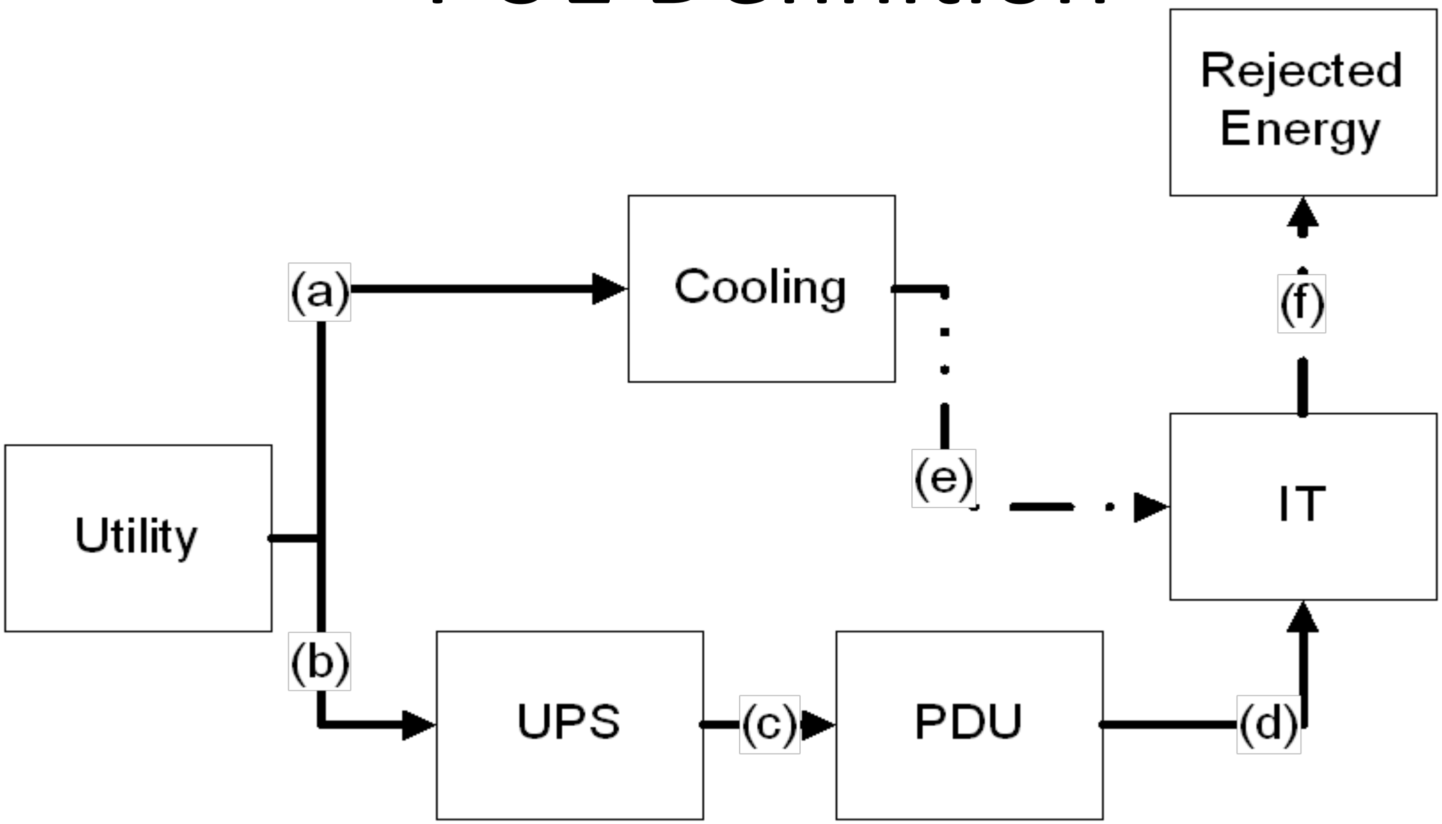
PUE

- Introduced in 2006 by Malone and Belady
- Developed and agreed to by EU Code of Conduct, DOE, EPA, Green Grid, ASHRAE, etc…
- Has led Energy Efficiency drive in Data Centers
  - PUE Average in 2007 ~ 2.5 … and in 2013 ~ 1.8
  - Best in Class 2013:

NREL= 1.06, LRZ= 1.15, NCAR~1.2,

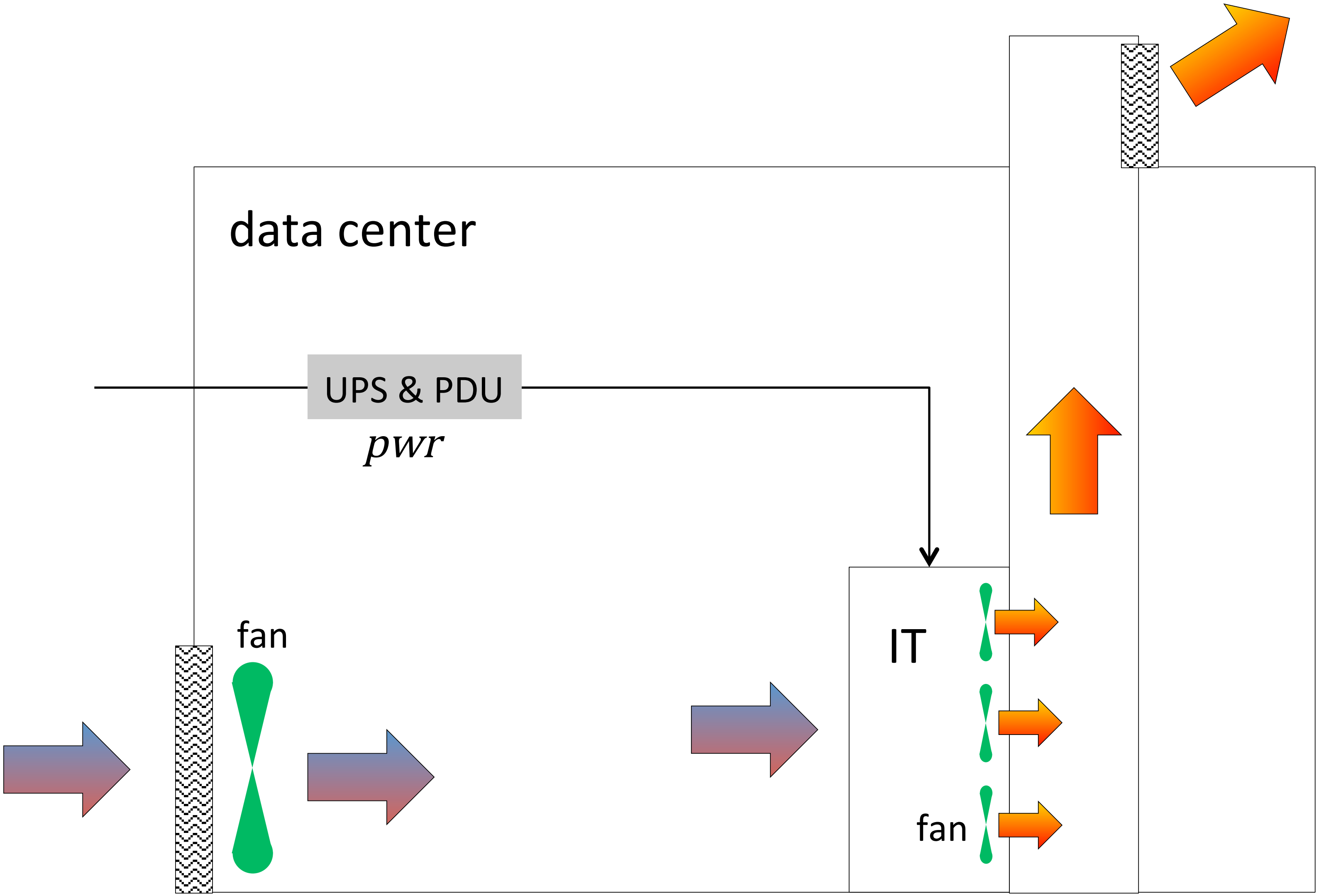
ORNL= 1.25, TU Dresden < 1.3

PUE Definition



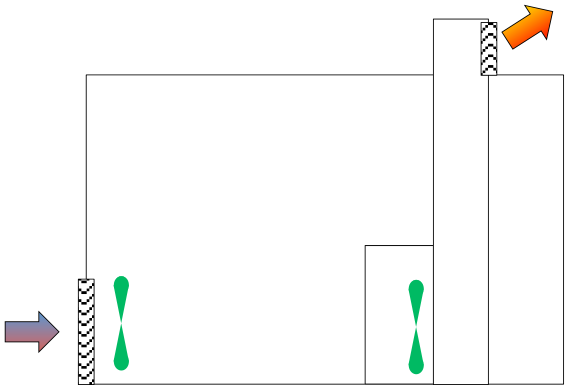
$$PUE = \frac{Total\ Energy}{IT\ Energy} = \frac{Cooling + PowerDistribution + Misc + IT}{IT} = \frac{a+b}{d}$$

but PUE isn't perfect, consider.....



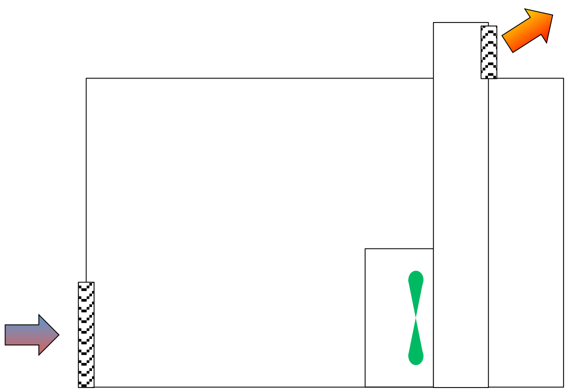
Three variations...

a) both fans



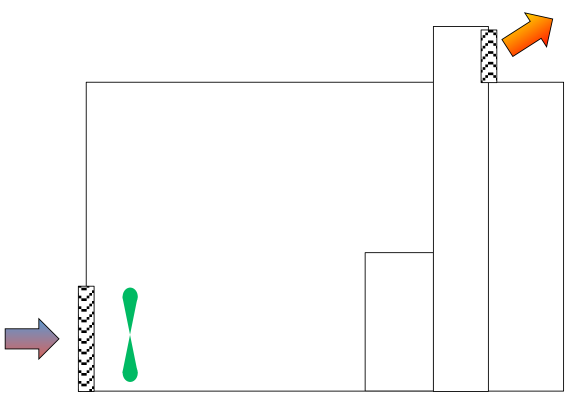
$$PUE_a = pwr + fan_{DC} + (IT + fan_{IT})/(IT + fan_{IT})$$

b) IT fans only



$$PUE_b = pwr + (IT + fan_{IT})/(IT + fan_{IT})$$

c) bldg fan only



$$PUE_c = pwr + fan_{DC} + IT/IT$$

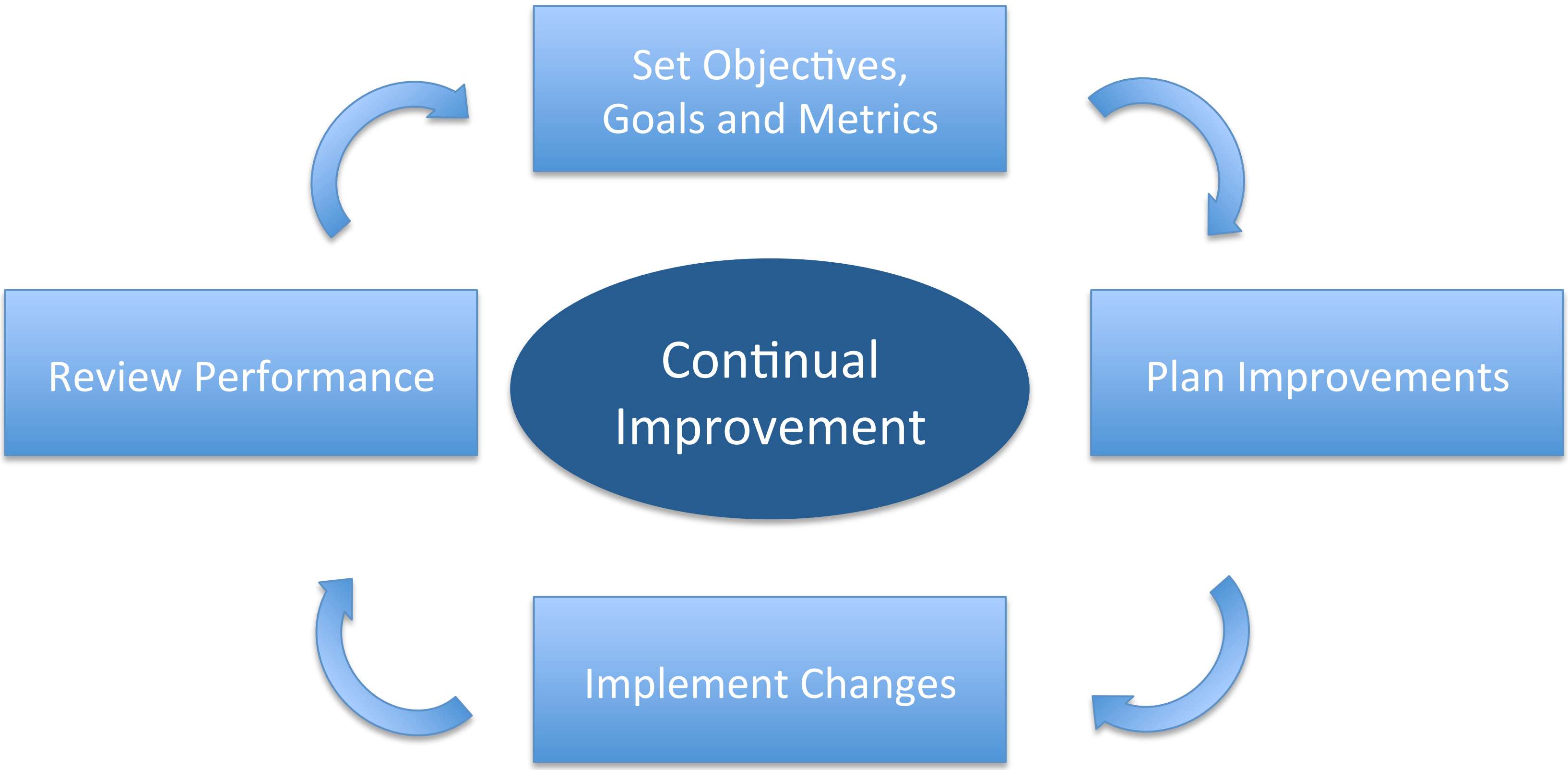
$PUE_b < PUE_a < PUE_c$ but is (b) best?
We don't know....

What makes a good metric?

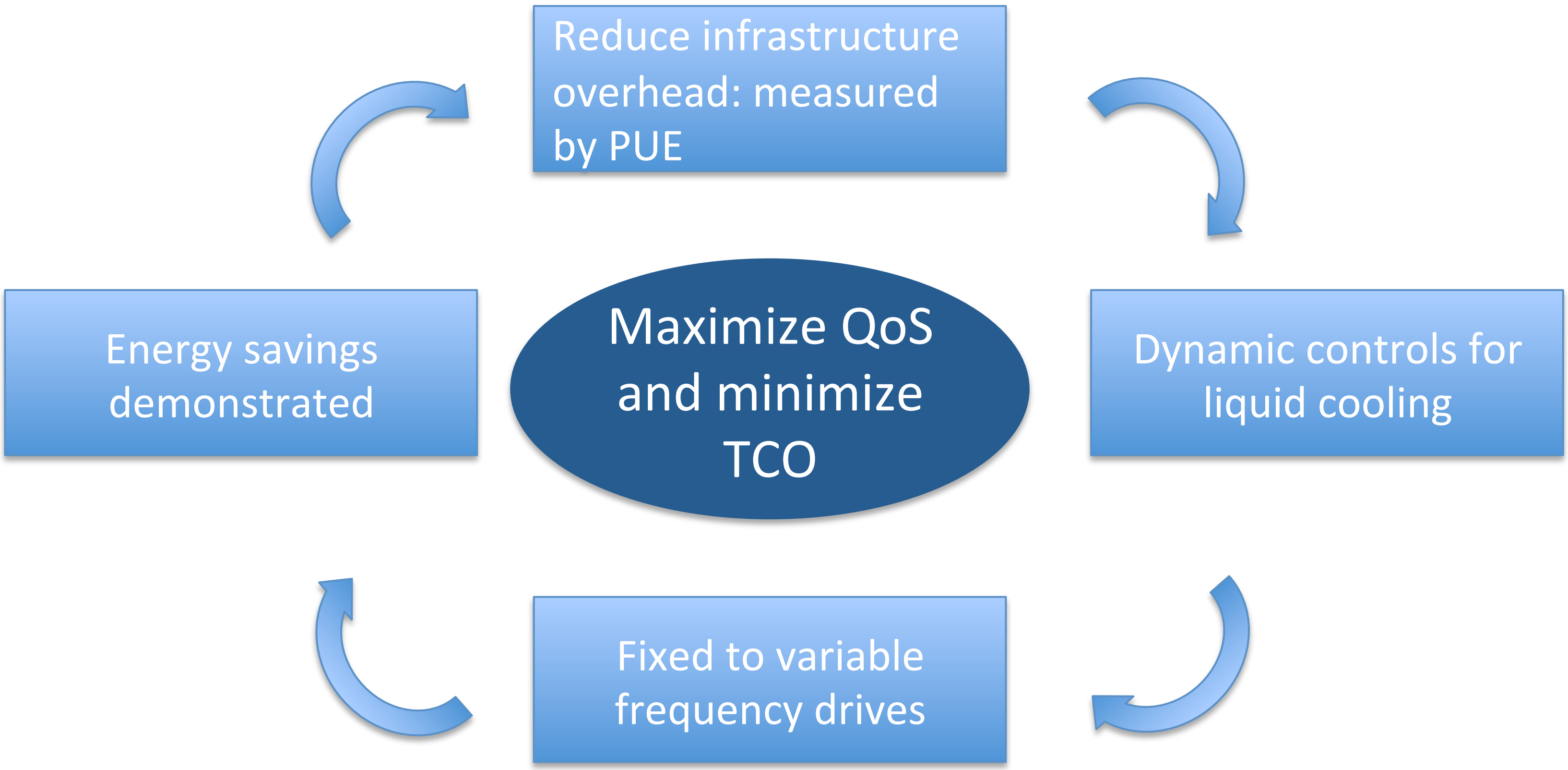
- Simple, measurable, actionable.
- Tied to an objective and links with a clear goal.
- It matters.

- Metrics are indicators to be used as part of a continual improvement process.

Continual improvement process



Example “based on real story”



Energy efficiency improvements.
No silver bullet…

… lots of silver BBs.

Facebook BBs include data centers

Facebook Greenfield Datacenter

Goal

- Design and build the most efficient datacenter eco-system possible

Control

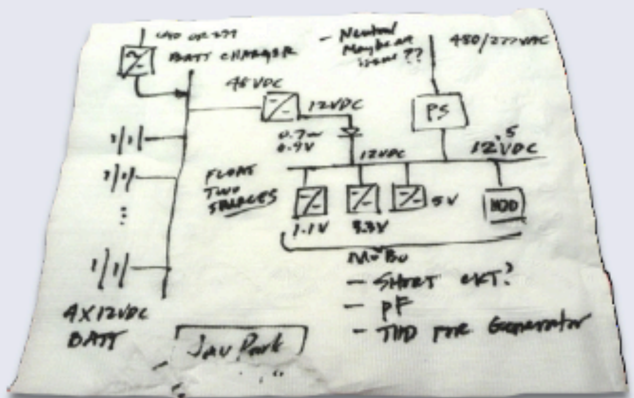
- Application
- Server configuration
- Datacenter design

Sites

- Prineville, OR
- Forest City, NC
- Luleå, Sweden
- Altoona, IA

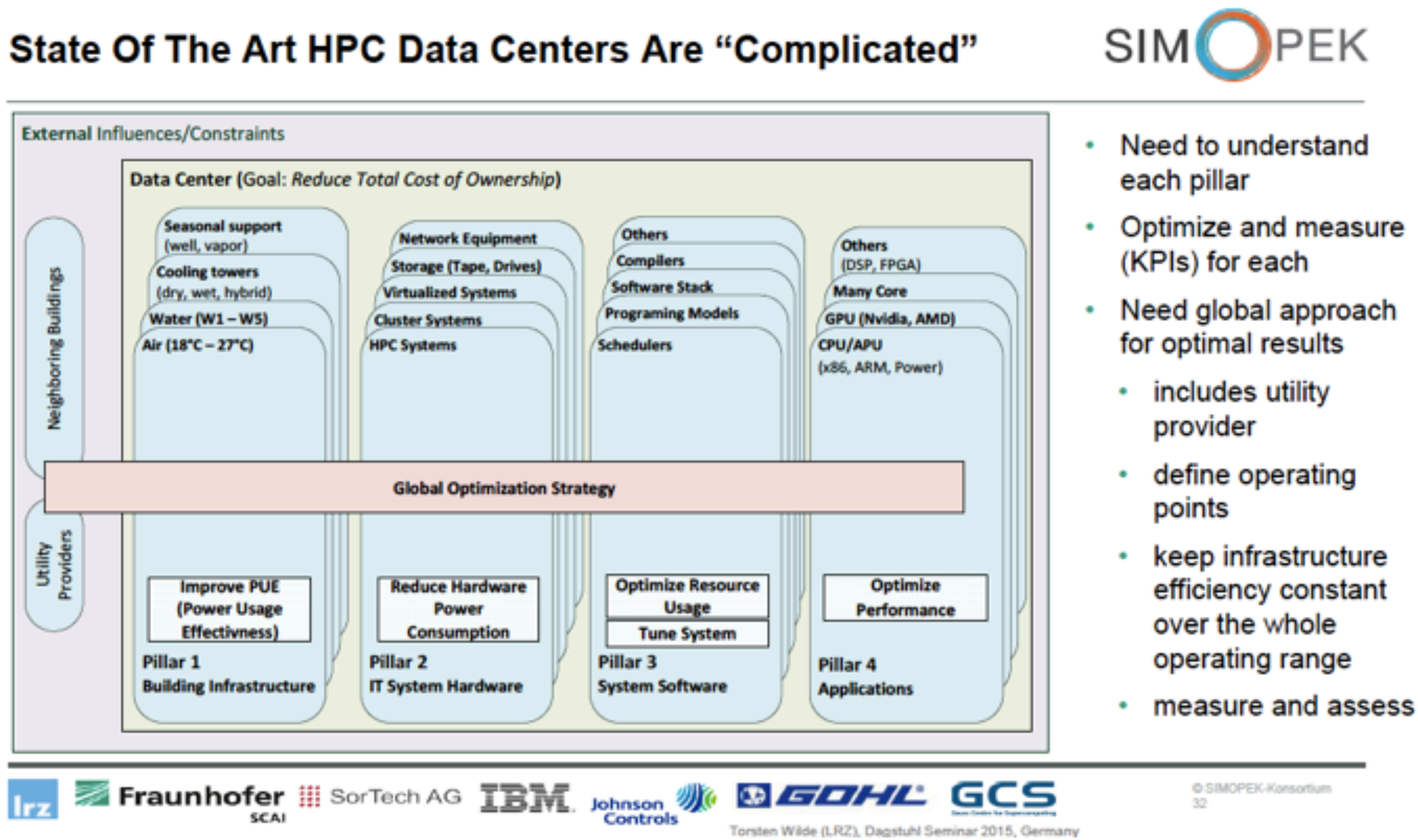
Look for synergies

Takeaways



- Take a clean sheet approach
- Efficiency as a design goal
- Look for synergies across the whole cross functional stack
  - data center
  - rack
  - server
  - network
  - application
- Open source

Four pillar framework



Slide curtesy of Torsten Wilde (LRZ) [www.simopek.de]; HPC Data Center Infrastructure Challenges Under A Power Bound, Dagstuhl Seminar 2015, Germany
Open Access 4 Pillar Framework Paper: http://www.springerlink.com/openurl.asp?genre=article&id=doi:10.1007/s00450-013-0244-6

Green IT at Leibniz Supercomputing Centre

The Four Pillar Model for Energy Efficient Data Center Operations

↳ Building Infrastructure	↳ Hardware	↳ Management Software	↳ Applications
▪ Reduction of losses in the power supply chain ▪ Improved technologies for cooling ▪ Reuse of waste heat	▪ Latest semiconductor technology ▪ Power saving mechanisms of today's CPUs ▪ Compute accelerators	▪ Monitoring of power consumption in IT and building infrastructures ▪ Virtualisation ▪ Smart control of hardware power saving modes	▪ Efficient algorithms ▪ Optimized libraries ▪ Efficient programming paradigms

https://www.lrz.de/wir/green-it_en/

Useful constructs for energy efficiency improvements

- Four pillar framework
- Power, space and cooling (PSC) capability
- Stranded and trapped capacity

PSC Capability: the availability of power, space, and cooling (PSC) to support HPC systems. Generally an HPC data center is built with the ability to expand, i.e., more PSC capability than initially required.

Stranded and trapped capacity

- Stranded Capacity: That which results from the HPC data center having uneven capability of power, space or cooling (PSC) resources to support additional HPC systems.

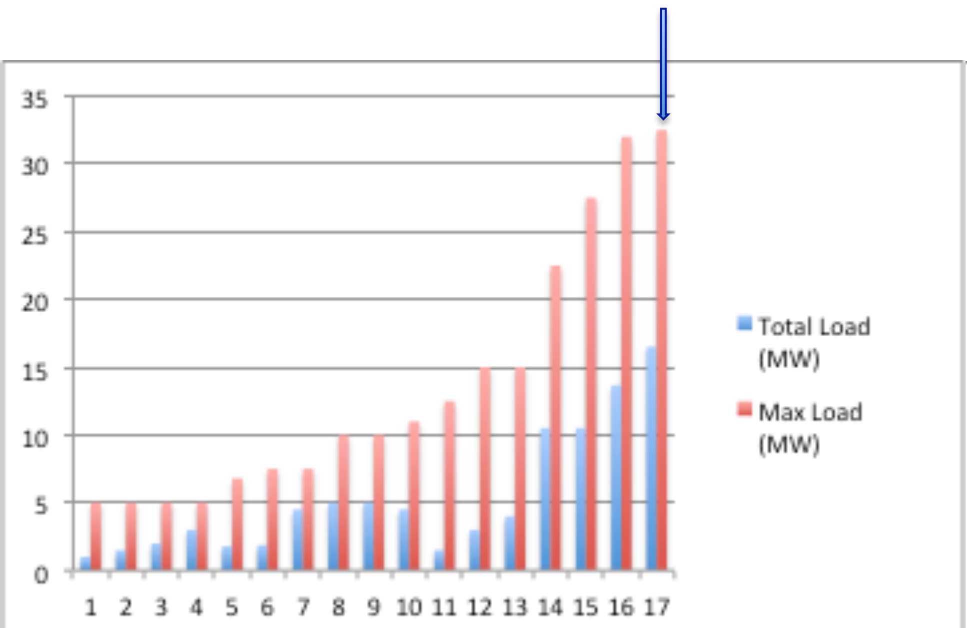
- Trapped Capacity: occurs when the workload does not use the intended PSC that has been allocated to it.

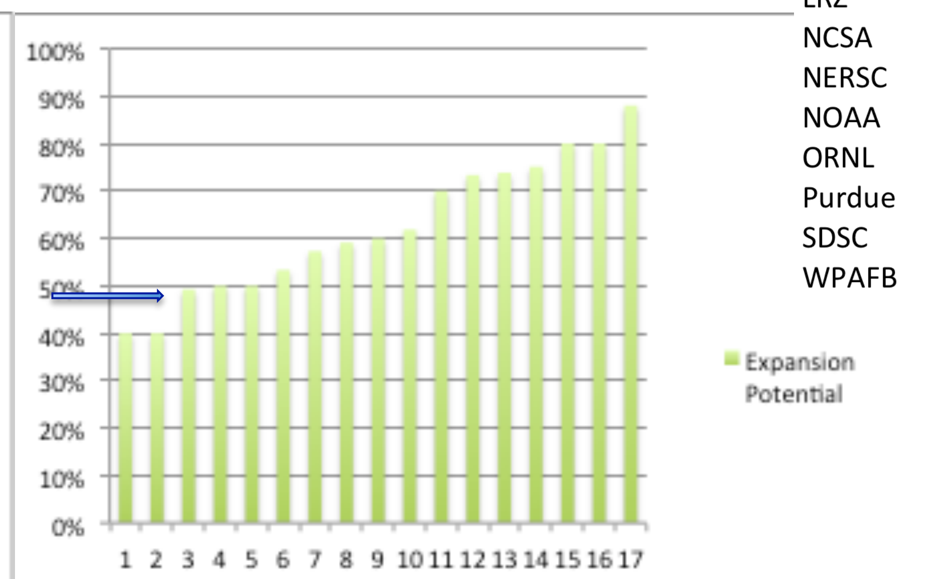
Lots of PSC capability…

Capability to more than double load

- Total load = as defined by PUE
- Max load = theoretical peak energy
- Based on EE HPC WG Survey, 2014

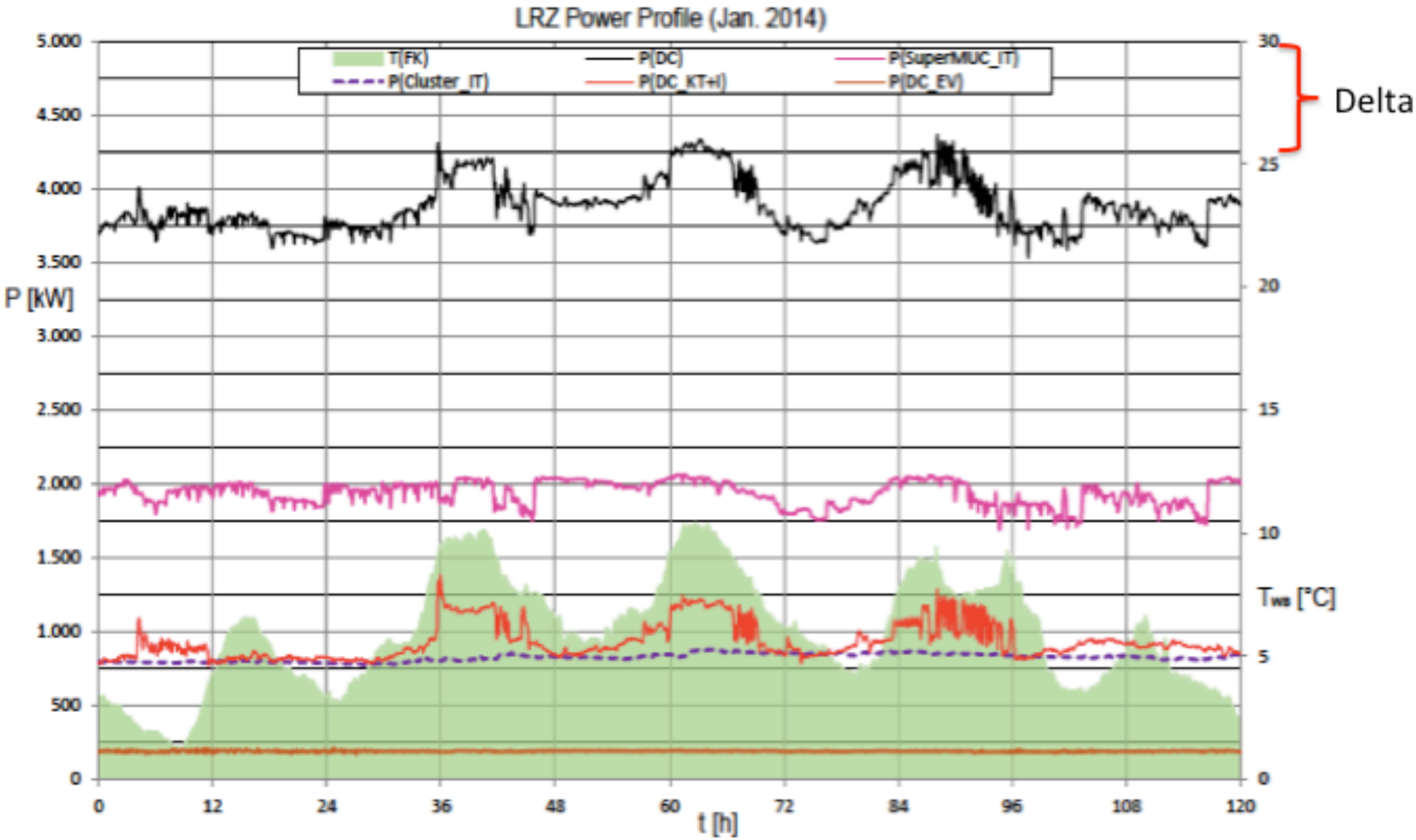
Sites
ANL
CEA
CINECA
CSCS
ECMWF
HLRS
KTH
LANL
LLNL
LRZ
NCSA
NERSC
NOAA
ORNL
Purdue
SDSC
WPAFB





… possibly stranded capacity.

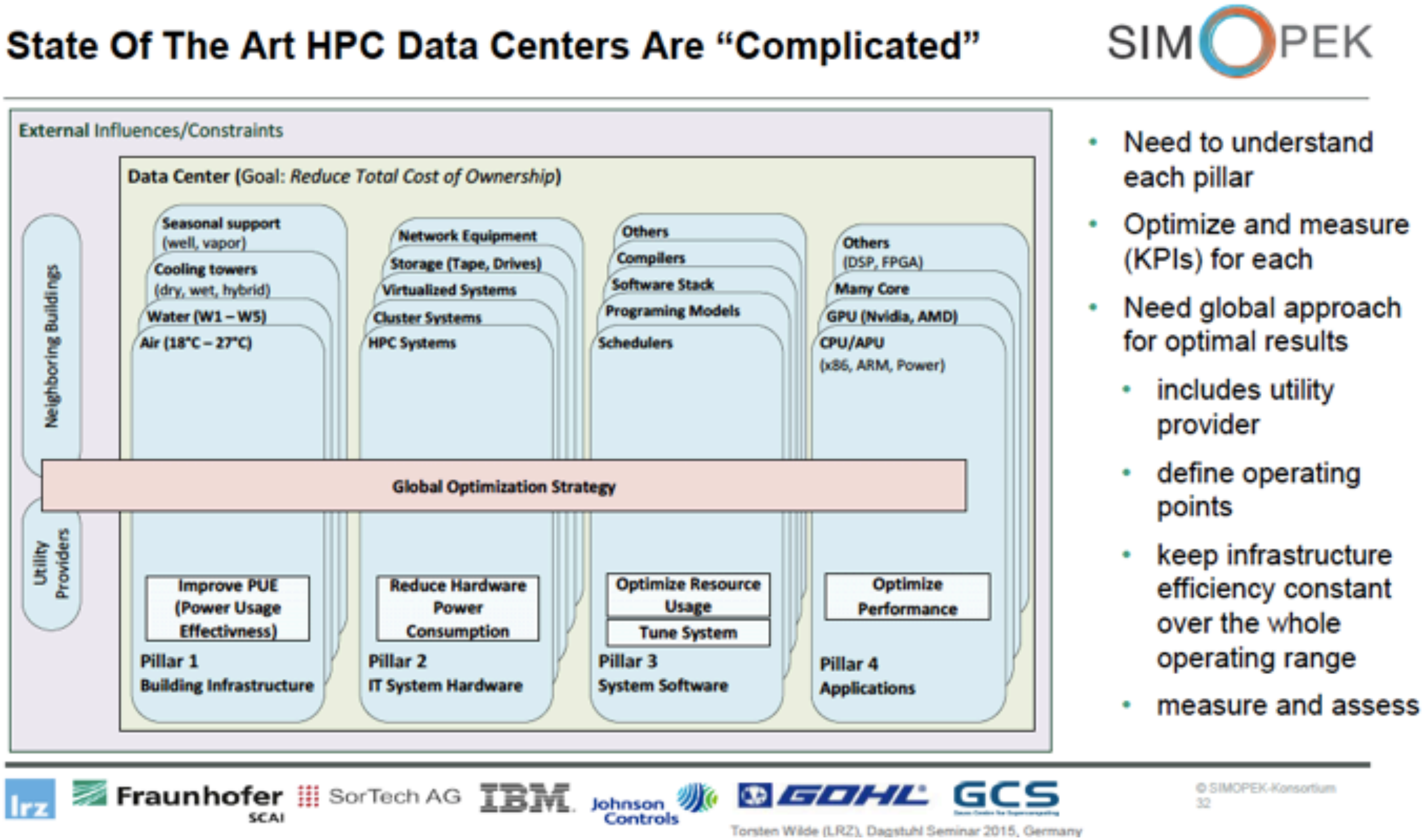
Typical load less than fit-up



Data provided by Detlef Labrenz and Torsten Wilde from Leibniz Supercomputing Centre (LRZ)

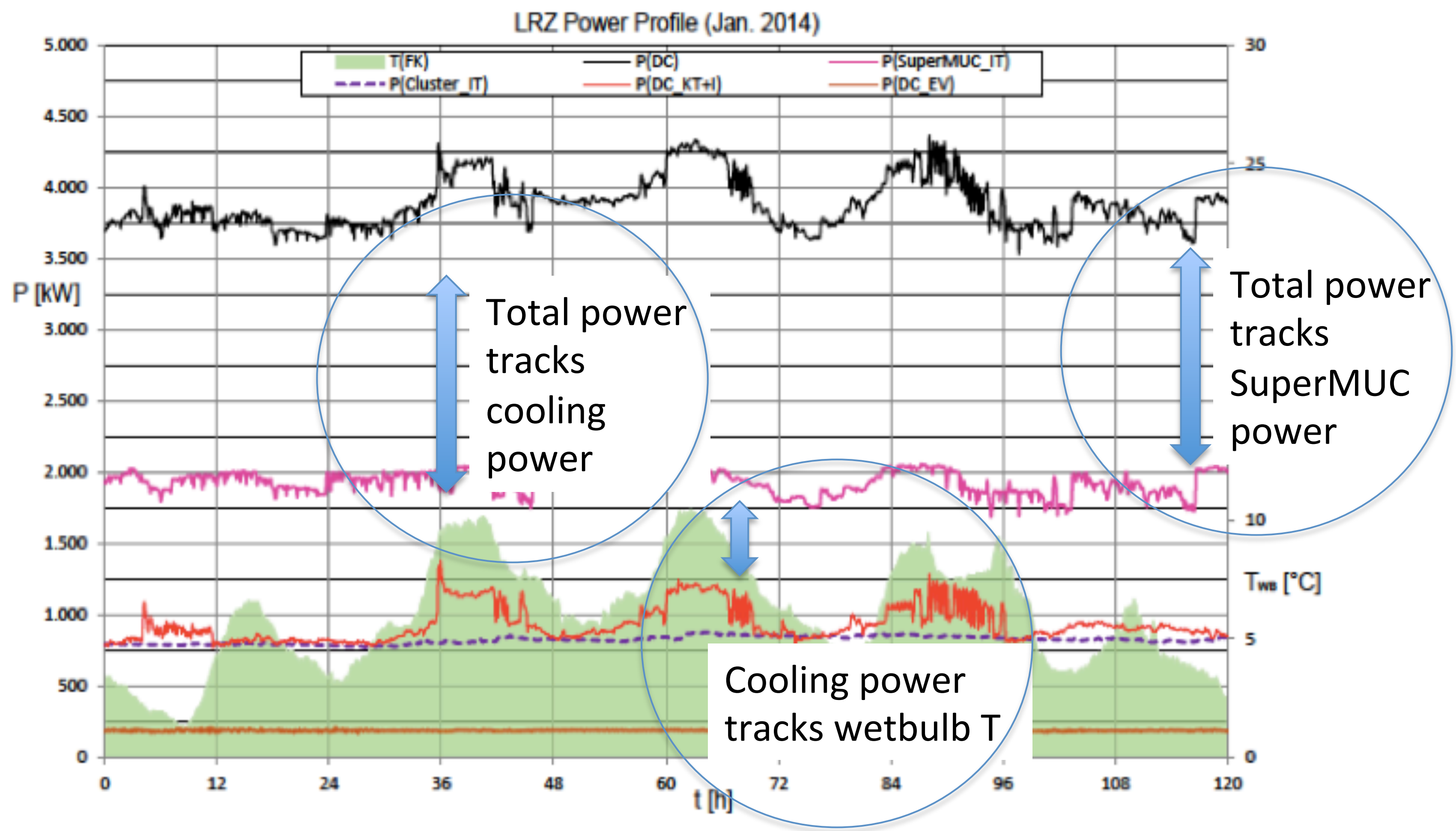

OH, THAT'S JUST GREAT! IT'S NOT BAD ENOUGH I'M STRANDED HERE! NOW I'M TRAPPED IN AN INVISIBLE BOX!
Bill Whitehead

Four pillar framework



Slide curtesy of Torsten Wilde (LRZ) [www.simopek.de]; HPC Data Center Infrastructure Challenges Under A Power Bound, Dagstuhl Seminar 2015, Germany
Open Access 4 Pillar Framework Paper: http://www.springerlink.com/openurl.asp?genre=article&id=doi:10.1007/s00450-013-0244-6

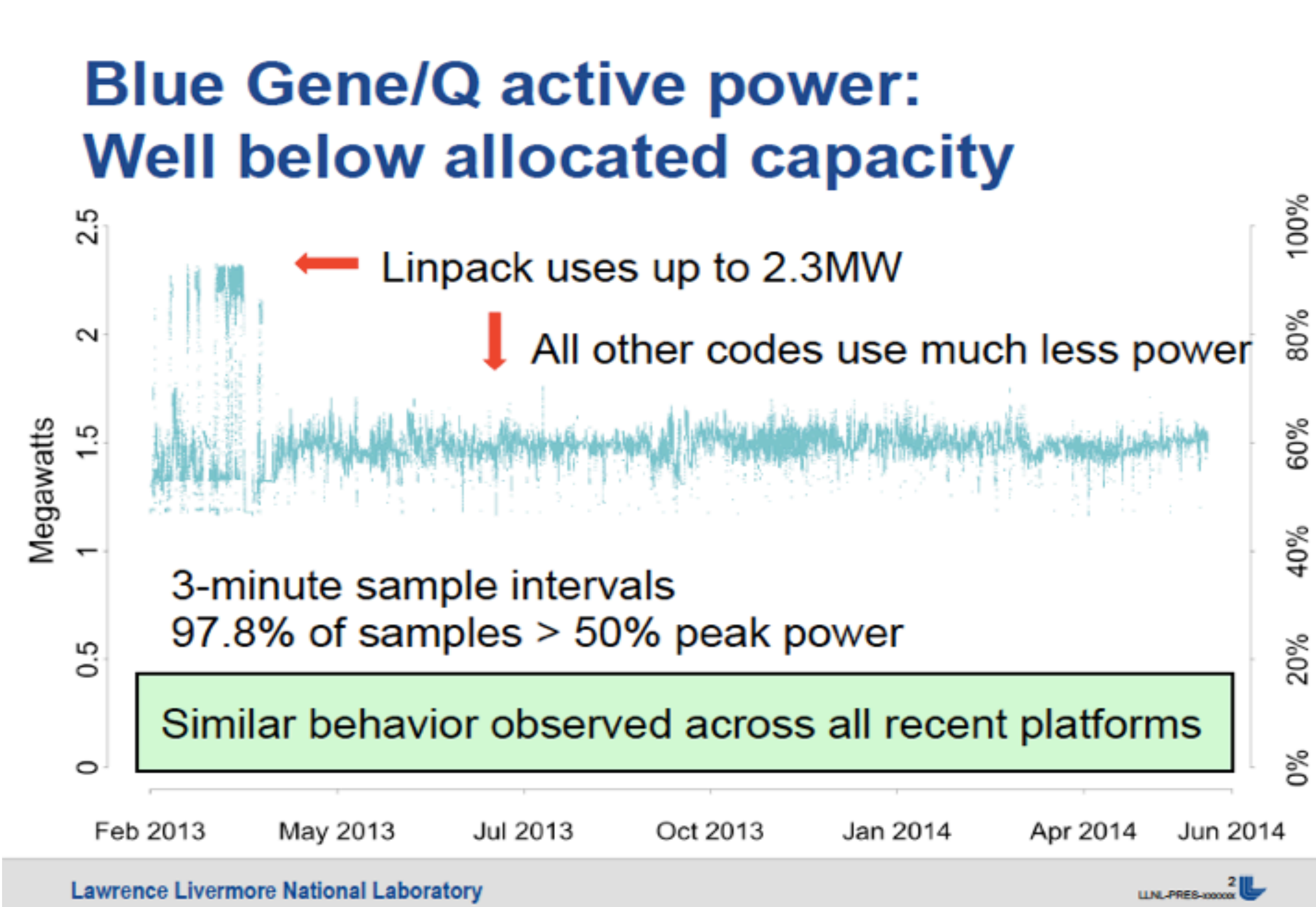
Investigating improvement opportunities



Challenges for energy efficiency when operating complex cooling infrastructures

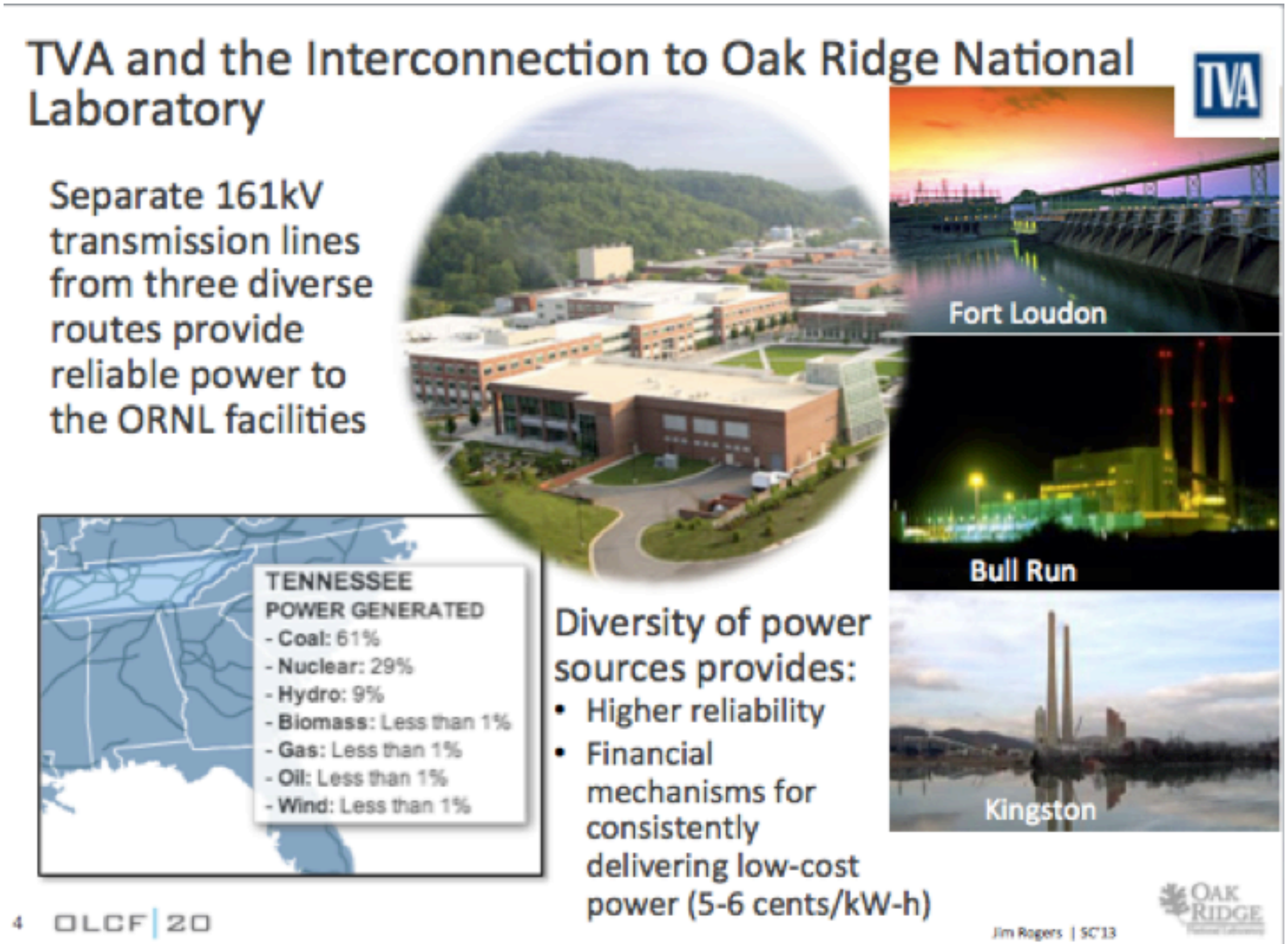
Data provided by Detlef Labrenz and Torsten Wilde from Leibniz Supercomputing Centre (LRZ)

Contributors to trapped capacity



Content provided courtesy of Barry Rountree, LLNL

Electricity provider impacts



Content provided courtesy of Jim Rogers, ORNL at SC13 EE HPC WG Workshop

Stiff and accommodating

Preliminary Model Results

- Current ORNL MV system model used; TVA HV system model contained basic elements needed for preliminary check.
- Worst case: 30 MW Load at ORNL cycled every 3/4 of cycle for 5 seconds
  - Out of step relay operation within a 30 cycle window:
    - If any out-of-step event is seen by the relay it will start a timer. If the event continues and is seen again within the 30 cycle timer window the line will trip due to load swings and/or other additional contingencies which created the instability .
  - Model was run with .99 pf and approximately .71 pf
- Results indicated that these load and pf swings *would not* impact out of step relay
  - Voltage drop at ORNL HPC system was high for .71 pf as would be expected
- Future collaboration with TVA to update their model
- Based on this preliminary analysis, the HV transmission system in the ORNL area is very stiff and can accommodate fast load switches as might be produced by HPC
- Other areas of the country where HV generation and transmission systems aren't as stiff should perform a similar analysis to verify that HPC loads will not have detrimental effect on the utilities' system.

Content provided courtesy of Jim Rogers, ORNL at SC13 EE HPC WG Workshop

Sensitivities and flexibility

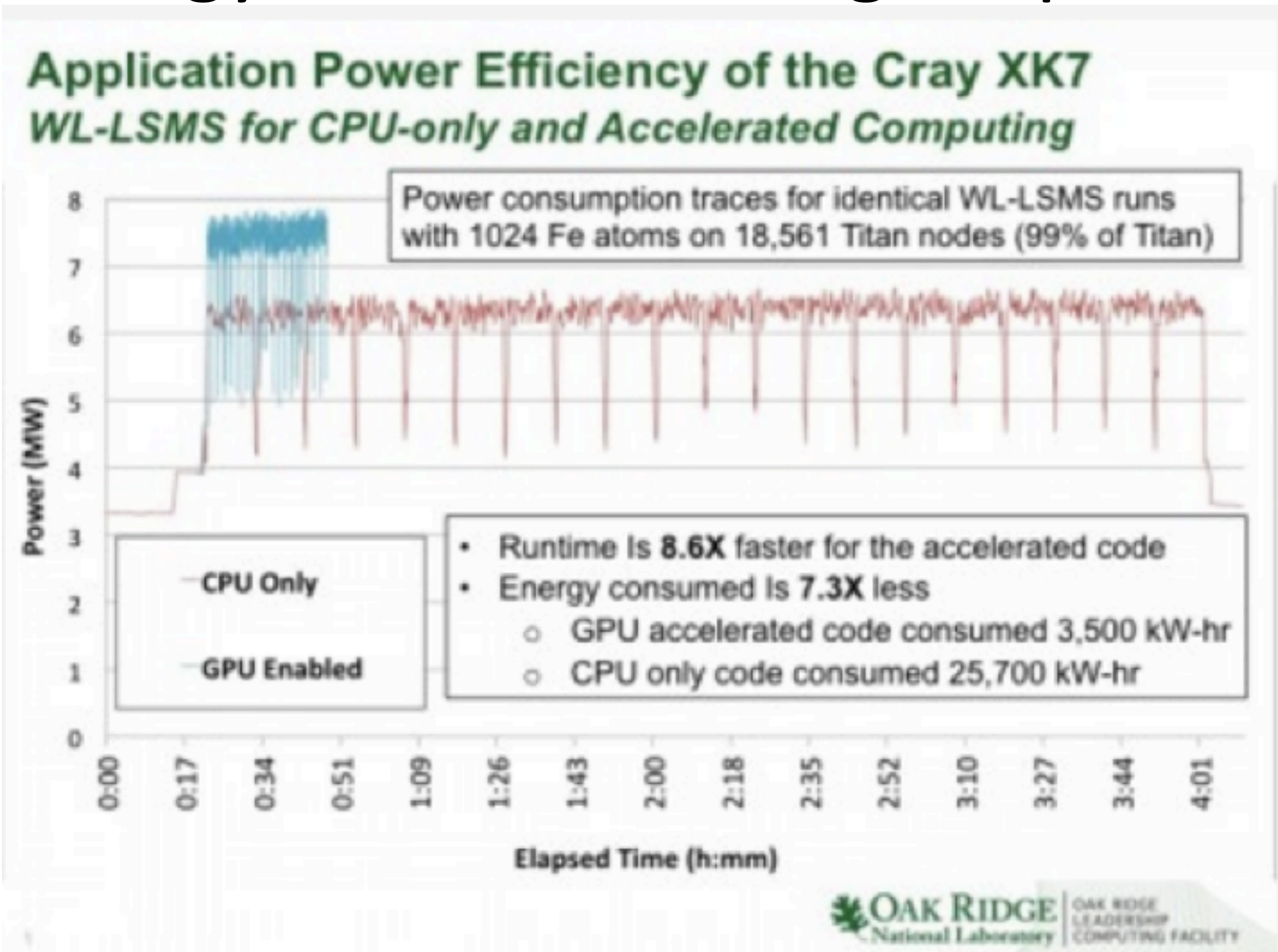
lrz LRZ Power Contract

- Power Provider: „Stadtwerke München“
- 100% Power from renewable energies
  - End of each year we get a document listing the renewable energy mix used → mostly water power stations in 2013

- Just signed new 3 years power contract with Stadtwerke München for the timeframe 2015 – 2017 with the following features
  - 100% Power from renewable energies
  - 10% variation in annual energy consumption
  - 4 power price tranches per year → watch European power stock exchange
  - Power provider should be at least three days ahead informed when large power consumption fluctuations are foreseeable (e.g., SuperMUC maintenance windows)

Content provided courtesy of Herbert Huber, LRZ at SC13 EE HPC WG Workshop

Energy efficient and higher power



Source: https://www.olcf.ornl.gov/2014/08/20/titan-proves-to-be-more-energy-efficient-than-its-predecessor/

Key take-aways

- Continually improve
- Use metrics that ‘matter’
- Understand four pillar synergies
- Find stranded and trapped capacity

No single metric,

- Identify and prioritize HPC center energy parameters for dashboards
- Identify potential stakeholder(s) for each of the energy parameters
- Document recommendations to assist the HPC community to choose the parameters they want to monitor and manage

… but a list to choose from.

The list is stakeholder dependent

- Director – Responsible for the overall center's activity
- Facility Manager – Primarily responsible for the physical infrastructure
- Information Technology Manager – Primarily responsible for the information technologies (hardware & software) in the data center

Facility manager's items

Item	Primary Information	Unit
1	Total power/energy	kW & kWh
2	IT Power /energy	kW & kWh
3	Power Usage Effectiveness -Power	Index
4	Power Usage Effectiveness- Energy	Index
5	Cooling Efficiency	kW/ton
6	Cooling Energy Use	kWh
7	Data center IT equipment cooling diagram	degF/C
8	Temperature (map)	degF/C
9	UPS input / output power /Energy	kW & kWh
10	Data center electrical distribution diagram	
11	CRAC/CRAH/AHU RAT (avg, min, max)	degF/C
12	CRAC/CRAH/AHU SAT (avg, min, max)	degF/C

Systems manager's items

Item	Primary Information	Unit
1	Energy Cost per data processing unit	$/unit
2	Total power/energy	kW & kWh
3	IT Power /energy	kW & kWh
4	Average IT utilization-Compute System	Percent
5	Power Usage Effectiveness – Power	Index
6	Power Usage Effectiveness-Energy	Index
7	IT efficiency[a]	Work output/ W*
8	Data center IT equipment cooling diagram	degF/C

[a] Depends on how each HPC center defines its work output

Director’s items

Item	Primary Information	Unit
1	Total power & energy	kW & kWh
2	Energy cost	$
3	Average IT utilization- Compute System	Percent
4	Power Usage Effectiveness –Power	Index
5	Power Usage Effectiveness- Energy	Index
6	IT efficiency [a]	Work output/Watt

[a] Depends on how each HPC center defines its work output

It is time to update these lists

- Today's workshop and ensuing report
- SC15 BoF "Identifying a Few, High-Leverage Energy Efficiency Metrics"

General Recommendations for High Performance Computing Data Center Energy Management Dashboard Display, Sartor, D. et al. http://doi.ieeecomputersociety.org/10.1109/IPDPSW.2013.272

Mobilize the community to accelerate energy efficient HPC

- Explore innovative approaches
- Peer to peer exchange
- Share best practices
- Take collective action

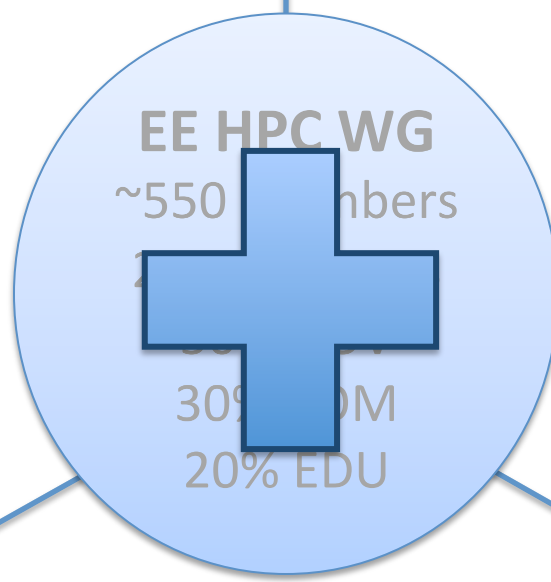
EE HPC WG
~600 Members
20+ countries
50% GOV
30% COM
20% EDU

http://eehpcwg.llnl.gov

natalie.jean.bates@gmail.com

Unique inter-disciplinary And organizational blend

HPC Systems

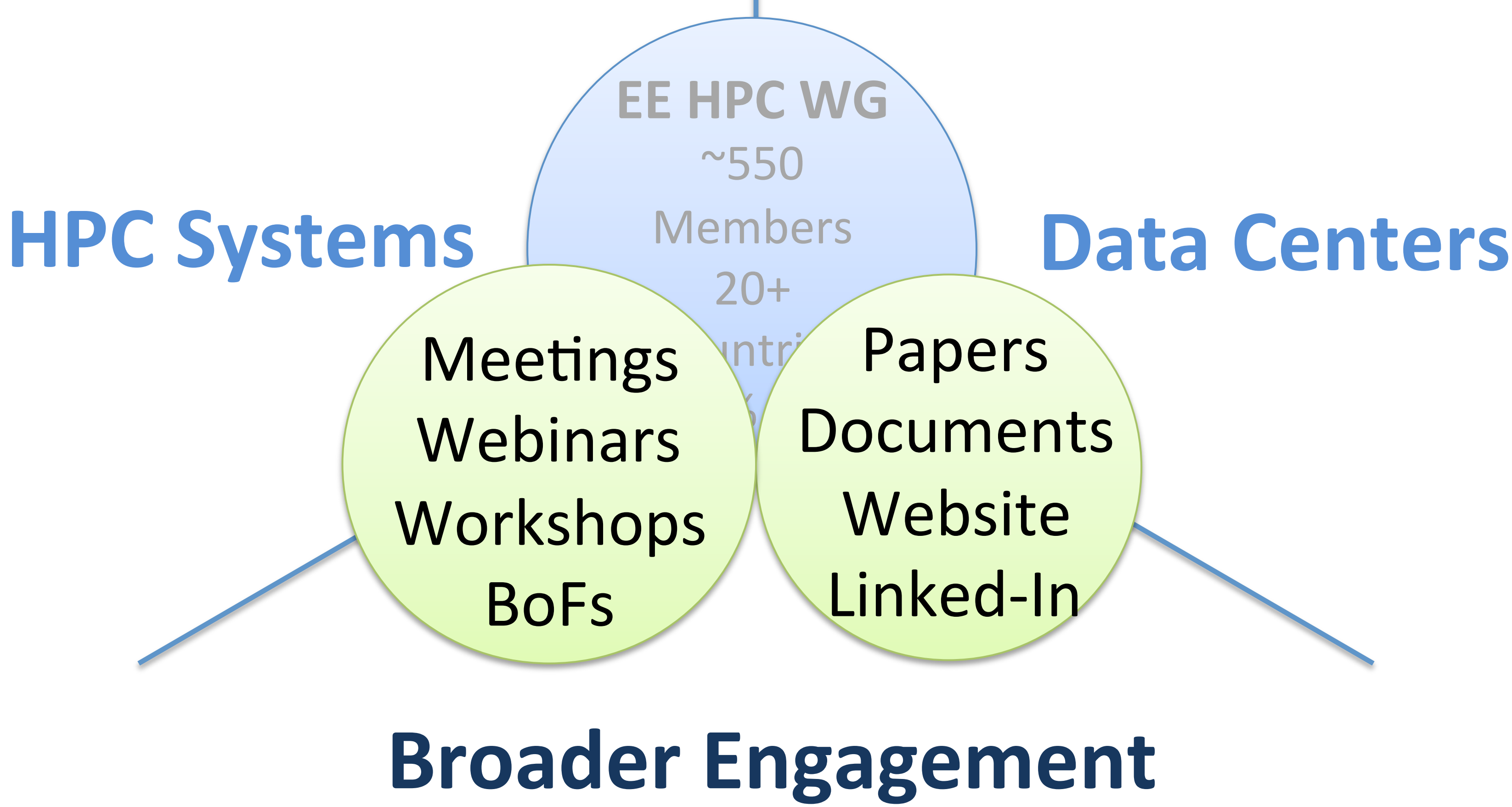


Data Centers

Broader Engagement

Open, virtual and accessible

- Membership criteria: interest in energy efficient HPC



SC15 events

- EE HPC WG Booth
- 6th Annual Workshop for the Energy Efficient HPC Working Group (EE HPC WG) 9-5:30 Monday, November 16
  - Keynote, Justin Rattner, CTO Intel (retired)
  - Opening Remarks, Satoshi Matsuoka, Tokyo Institute of Technology
  - Japanese Supercomputing Center Perspectives
  - Oil & Gas Industry Panel
  - Lessons learned from NREL, Dresden and LANL
- BoF “Identifying a Few, High-Leverage Energy Efficiency Metrics”
- BoF “Dynamic Liquid Cooling, Telemetry and Controls; Opportunity for Improved TCO?”
- BoF “The Green500 List and its Continuing Evolution”
- Paper “Node Variability in Large-Scale Power Measurements: Perspectives from the Green500, Top500 and EEHPCWG”

- Thank you!
- Questions?

http://eehpcwg.llnl.gov

natalie.jean.bates@gmail.com