Towards a Roadmap for HPC Energy Efficiency

International Conference on Energy-Aware High Performance Computing

September 11, 2012

Natalie Bates

Future Exascale Power Challenge

Where do we get a 1000x improvement in performance with only a 10x increase in power?

?

5

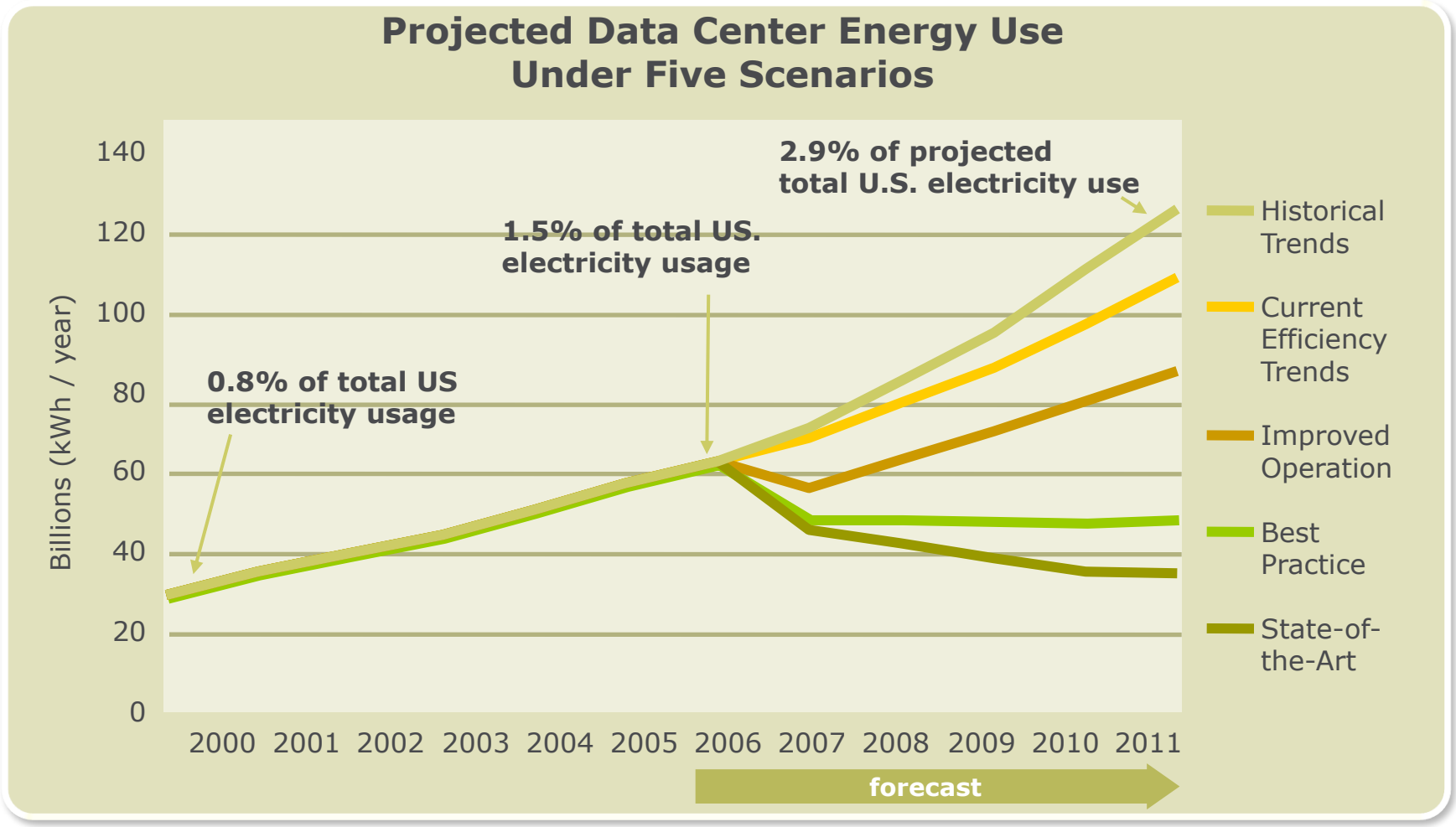
How do you achieve this in 10 years with a finite development budget?

8

20MW Target - $20M Annual Energy Cost

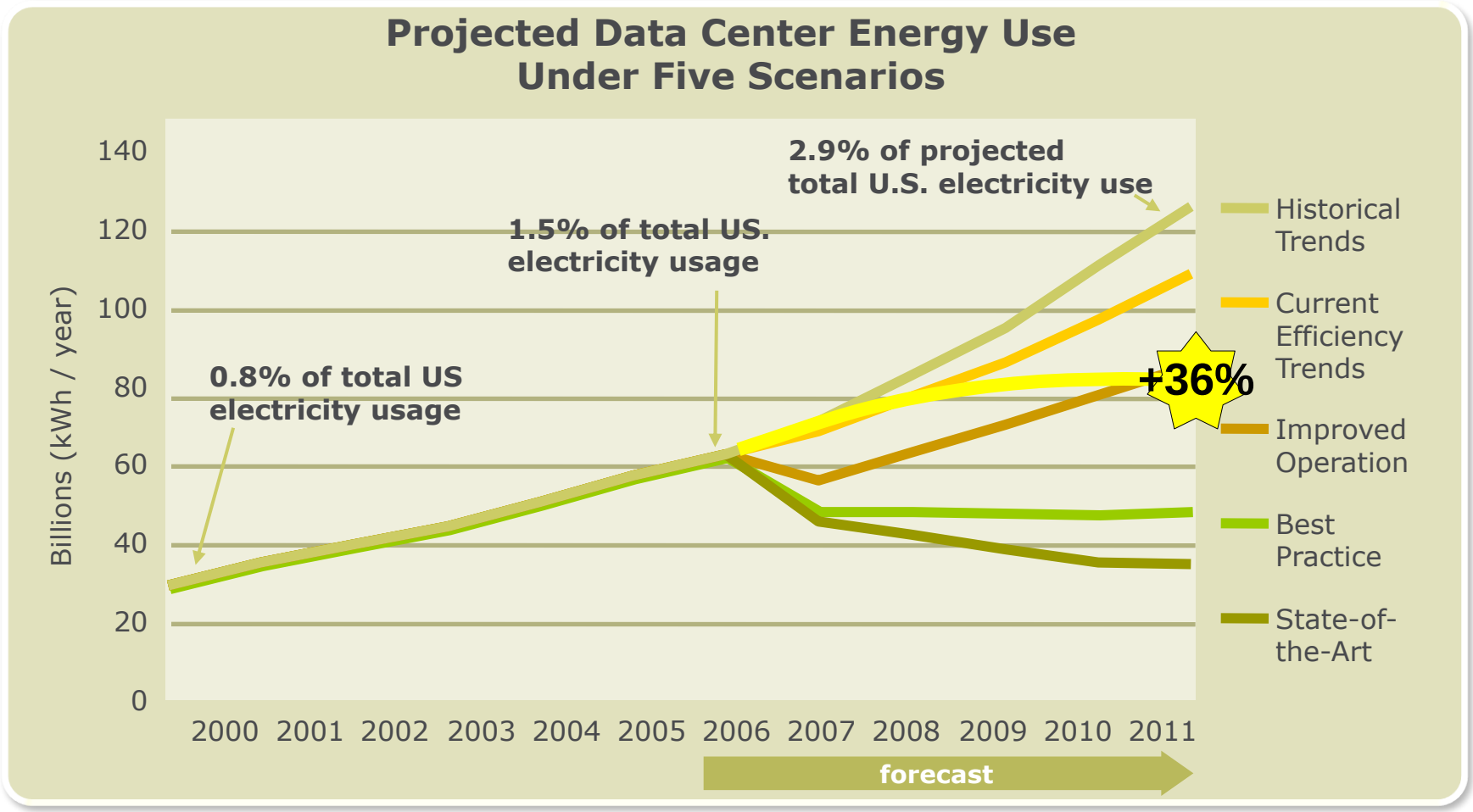
Original material attributable to John Shalf, LBNL

Past Pending Crisis



EPA Report to Congress of Server and Data Center Energy Efficiency, 2007

And Opportunity for Improvement

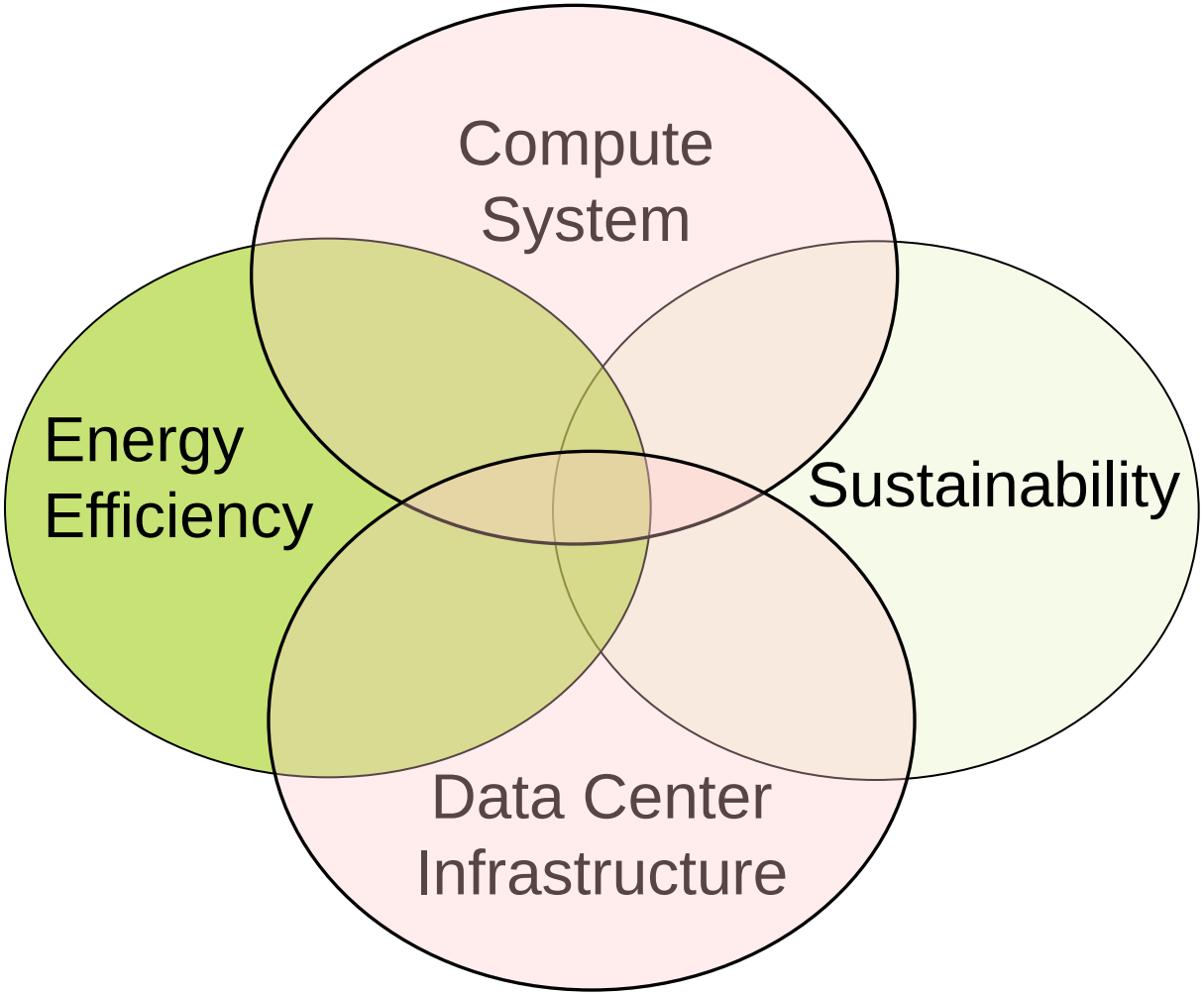


Koomey, 2011, 36% growth

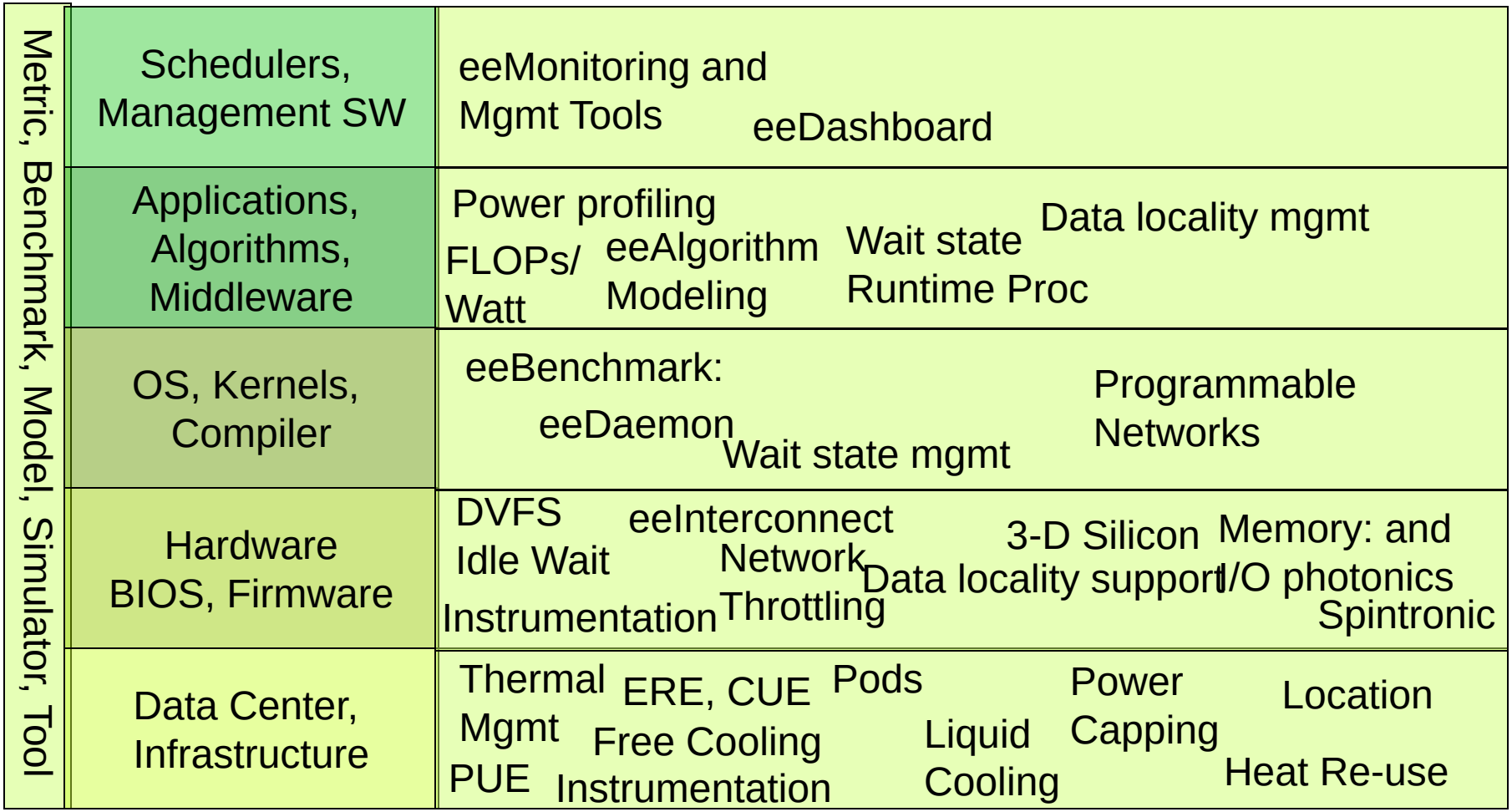
Grace Hopper Inspiration

nersc.gov

High Performance Computing, Energy Efficiency and Sustainability



Energy-efficiency Roadmap



Energy Efficient HPC Working Group

- Driving energy conservation measures and energy efficient design in HPC
- Forum for sharing of information (peer-to-peer exchange) and collective action
- Open to all interested parties

EE HPC WG Website
http://eehpcwg.lbl.gov
Email
energyefficientHPCWG@gmail.com
Energy Efficient HPC Linked-in Group
http://www.linkedin.com/groups?gid=2494186&trk=myg_ugrp_ovr

With a lot of support from Lawrence Berkeley National Laboratory

Membership

- Science, research and engineering focus
- 260 members and growing
- International- members from ~20 countries
- Approximately 50% government labs, 30% vendors and 20% academe
  - United States Department of Energy Laboratories
- Only membership criteria is ‘interest’ and willingness to receive a few emails/month
- Bi-monthly general membership meeting and monthly informational webinars

Teams and Leaders

- EE HPC WG
  - Natalie Bates *(LBNL)*
  - Dale Sartor *(LBNL)*
- System Team
  - Erich Strohmaier *(LBNL)*
  - John Shalf *(LBNL)*
- Infrastructure Team
  - Bill Tschudi *(LBNL)*
  - Dave Martinez *(SNL)*
- Conferences (and Outreach) Team
  - Anna Maria Bailey *(LLNL)*
  - Marriann Silviera *(LLNL)*

Technical Initiatives and Outreach

- Infrastructure Team
  - Liquid Cooling Guidelines
  - Metrics: ERE, Total PUE and CUE
  - *Energy Efficiency Dashboards**
- System Team
  - Workload-based Energy Efficiency Metrics
  - *Measurement, Monitoring and Management**
- Conferences (and Outreach) Team
  - Membership
  - Monthly webinar
  - Workshops, Birds of Feather, Papers, Talks

**Under Construction*

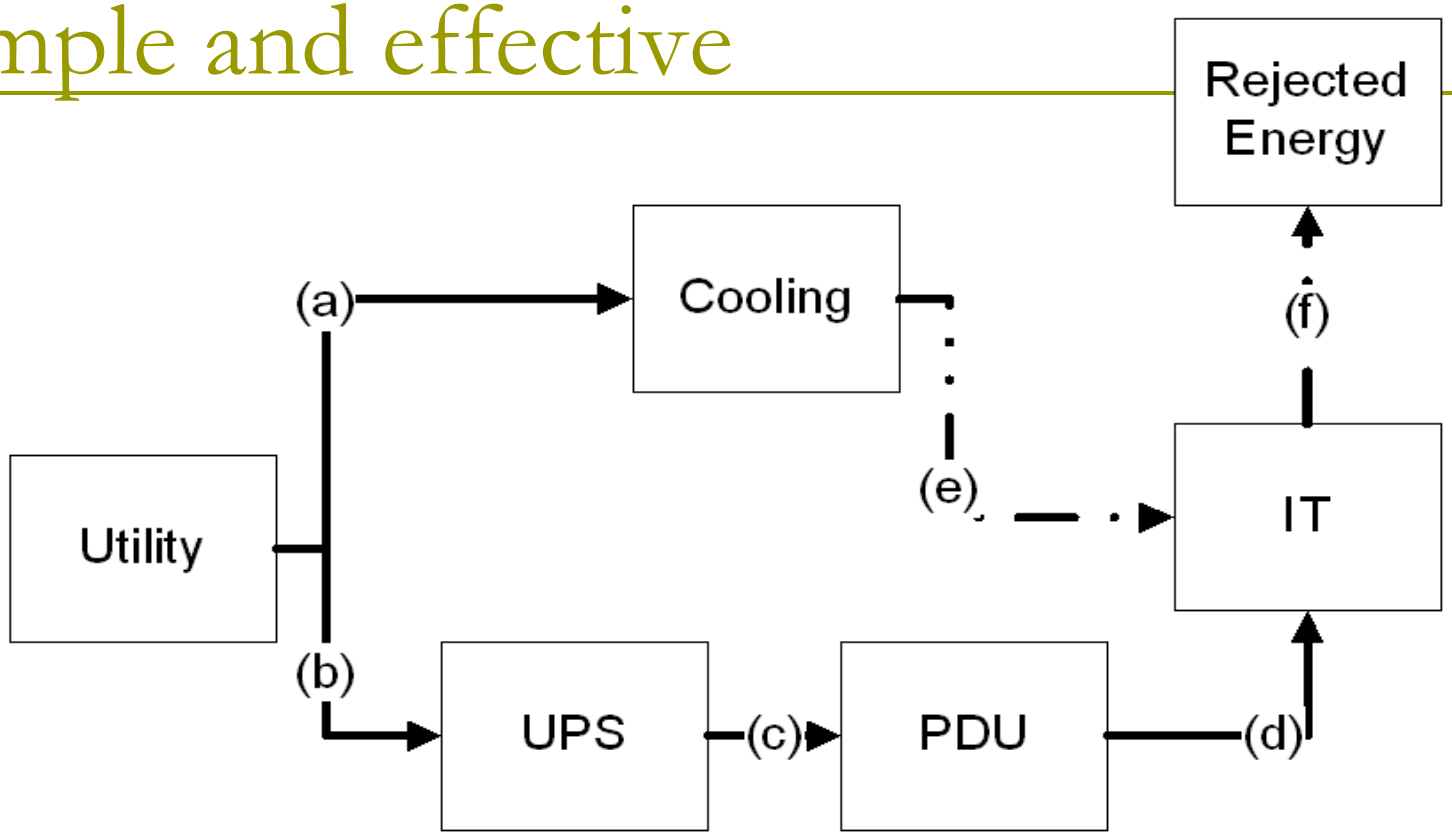
Energy Efficient Liquid Cooling

- Eliminate or dramatically reduce use of compressor cooling (chillers)
- Standardize temperature requirements
  - common design point: system and datacenter
- Ensure practicality
  - Collaboration with HPC vendor community to develop attainable recommended limits
- Industry endorsement
  - Collaboration with ASHRAE to adopt recommendations in new thermal guidelines

Analysis and Results

- Analysis
  - US DOE National Lab climate conditions for cooling tower and evaporative cooling
  - Model heat transfer from processor to atmosphere and determine thermal margins
- Technical Result
  - Direct liquid cooling using cooling towers producing water supplied at 32°C
  - Direct liquid cooling using only dry coolers producing water supplied at 43°C
- Initiative Result
  - ASHRAE TC9.9 Liquid Cooling Thermal Guideline

Power Usage Effectiveness (PUE) – simple and effective



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

The Green Grid, www.thegreengrid.org

PUE: All about the "1"

	PUE
EPA Energy Star Average – reported in 2009	1.91
Intel Jones Farm, Hillsboro	1.41
ORNL CSB	1.25
T-Systems & Intel DC2020 Test Lab, Munich	1.24
Google	1.16
Leibniz Supercomputing Centre (LRZ)	1.15
National Center for Atmospheric Research (NCAR)	1.10
Yahoo, Lockport	1.08
Facebook, Prineville	1.07
National Renewable Energy Laboratory (NREL)	1.06

PUE reflect reported as well as calculated numbers

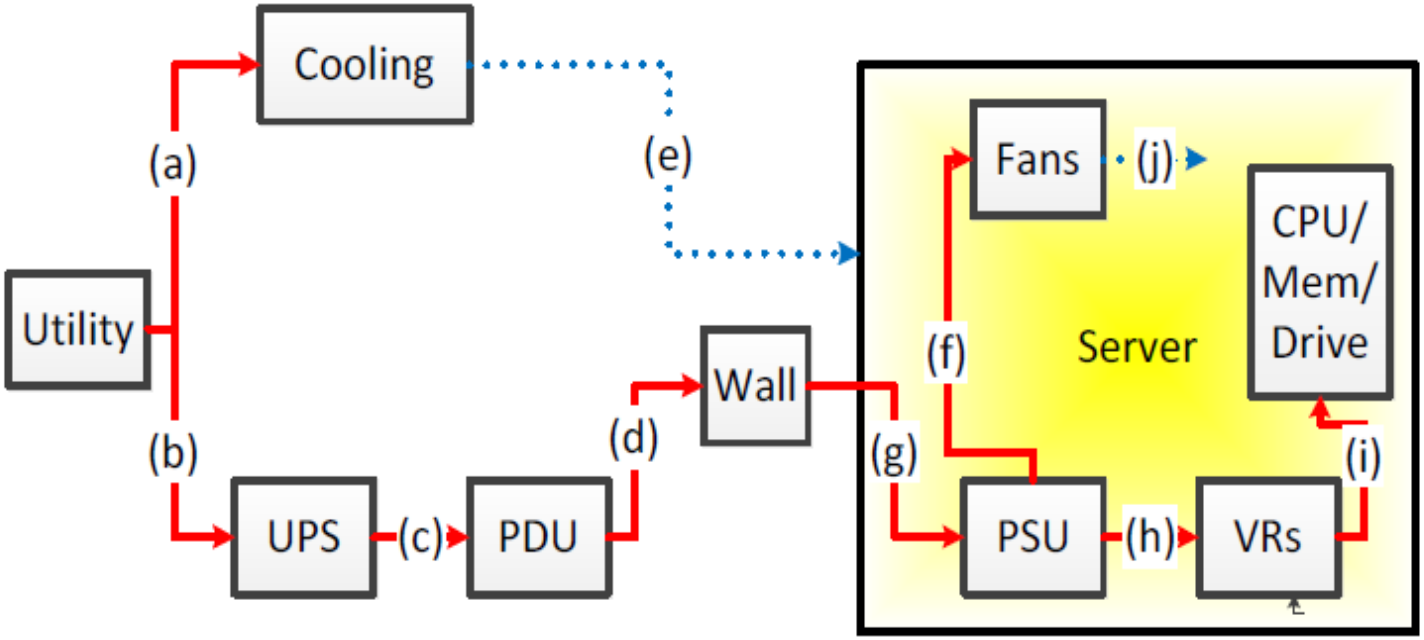
Refining PUE for better comparison - TotalPUE

- PUE does not account for cooling and power distribution losses inside the compute system
- ITPUE captures support inefficiencies in fans, liquid cooling, power supplies, etc.
- TUE provides true ratio of total energy, (including internal and external support energy uses)
- TUE preferred metric for inter-site comparison

$$TUE = \quad PUE \times ITPUE = \frac{a+b}{d} \times \frac{f+g}{i} = \frac{a+b}{i}$$

EE HPC WG Sub-team proposal

Combine PUE and ITUE for TUE

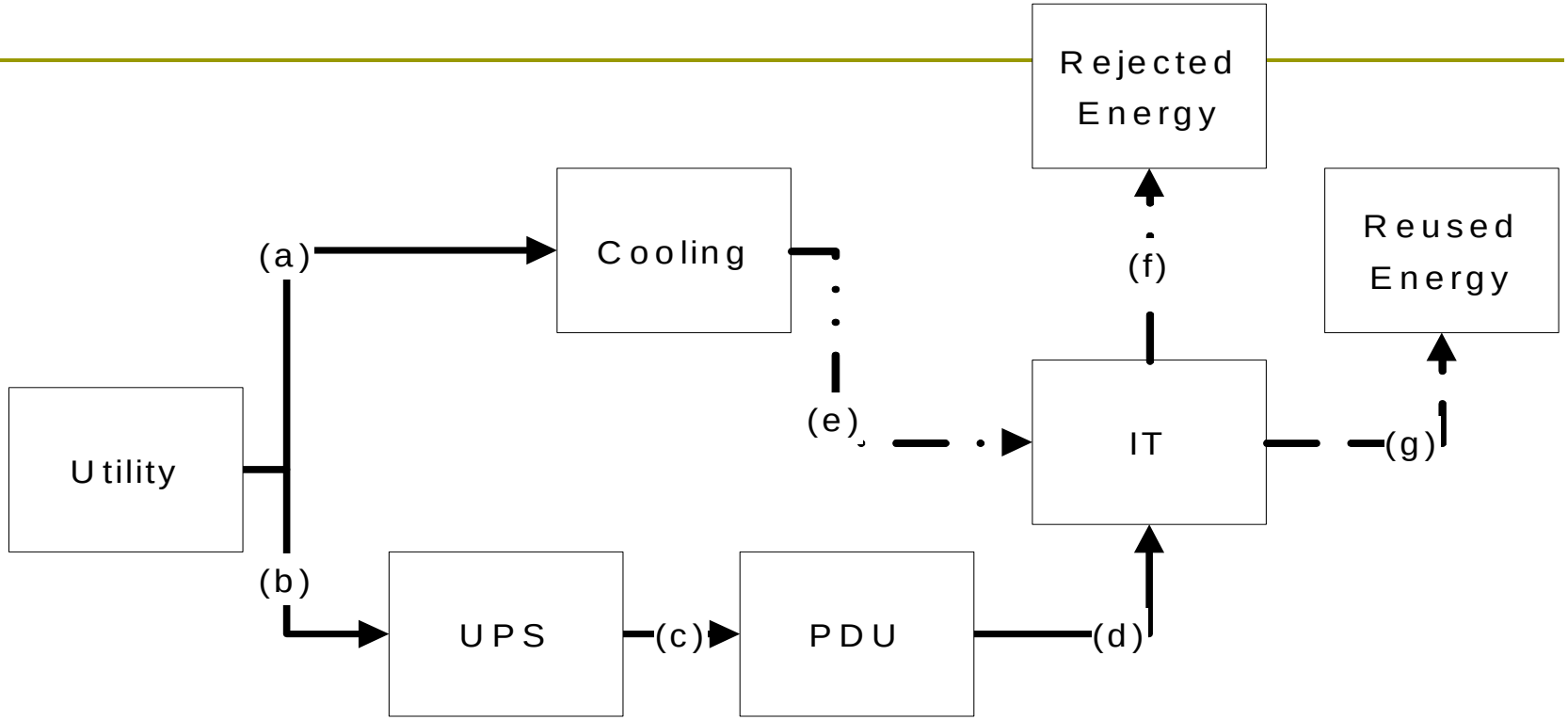


$$TUE = DCPUE \times ITPUE = \frac{a+b}{d} \times \frac{f+g}{i} = \frac{a+b}{i}$$

"I am re-using waste heat from my data center on another part of my site and my PUE is 0.8!"

"I am re-using waste heat from my data center on another part of my site and my PUE is 0.8!"

Energy Re-use Effectiveness



$$ERE = \frac{Total\ Energy - Reuse\ Energy}{IT\ Energy}$$

$$= \frac{Cooling + PowerDistribution + Misc + IT - Reuse}{IT} = \frac{a + b - g}{d}$$

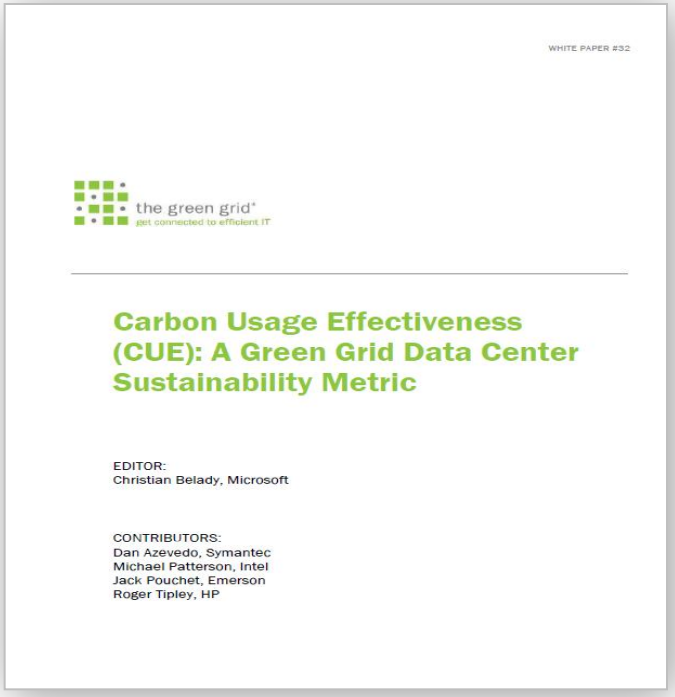
PUE & ERE resorted….

	PUE	Energy Reuse
EPA Energy Star Average	1.91	
Intel Jones Farm, Hillsboro	1.41	
T-Systems & Intel DC2020 Test Lab, Munich	1.24	
Google	1.16	
NCAR	1.10	
Yahoo, Lockport	1.08	
Facebook, Prineville	1.07	
Leibniz Supercomputing Centre (LRZ)	1.15	☑ ERE <1.0
National Renewable Energy Laboratory (NREL)	1.06	☑ ERE <1.0

Carbon Usage Effectiveness (CUE)

$$CUE = \frac{Total\ CO\ emissions\ caused\ by\ the\ Total\ Data\ Center\ Energy}{IT\ Energy}$$

- Ideal value is 0.0
- Example, the Nordic HPC Data Center in Iceland is powered by renewable energy – CUE ~ 0.0



What is Needed

- Form a basis for evaluating energy efficiency of individual systems, product lines, architectures and vendors
- Target architecture design and procurement decision making process

Agreement in Principal

- Collaboration between Top500, Green500, Green Grid and EE HPC WG
- Evaluate and improve methodology, metrics, and drive towards convergence on workloads
- Report progress at ISC and SC

Workloads

- Leverage well-established benchmarks
- Must exercise the HPC system to the fullest capability possible
- Measure behavior of key system components including compute, memory, interconnect fabric, storage and external I/O
- Use High Performance LINPACK (HPL) for exercising (mostly) compute sub-system

Methodology

I get the Flops…

but, per Whatt?

Complexities and Issues

- Fuzzy lines between the computer system and the data center, e.g., fans, cooling systems
- Shared resources, e.g., storage and networking
- Data center not instrumented for computer system level measurement
- Measurement tool limitations, e.g., frequency, power verses energy
- dc system level measurements don’t include power supply losses

Proposed Improvements

- Current power measurement methodology is very flexible, but compromises consistency
- Proposal is to keep flexibility, but keep track of rules used and quality of power measurement
- Levels of power measurement quality
  - L3 = current best capability (LLNL and LRZ)
  - L1 = Green500 methodology
  - ↑ quality: more of the system, higher sampling rate, more of the HPL run
  - Common rules for system boundary, power measurement point and start/stop times
  - Vision is to continuously 'raise the bar'

Methodology Testing

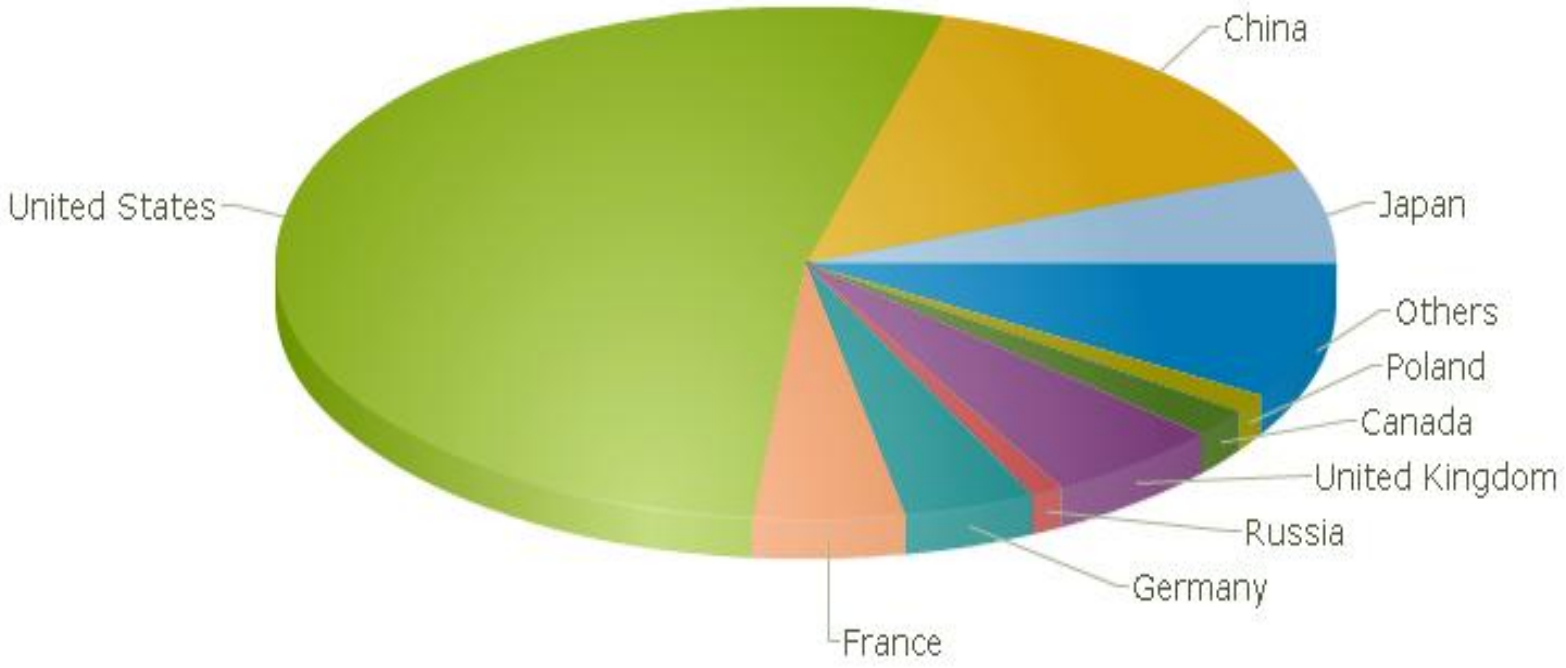
- Alpha Test- ISC'12
  - 5 early adopters
    - Lawrence Livermore National Laboratory, Sequoia
    - Leibniz Supercomputing Center, SuperMUC
    - Oak Ridge National Laboratory, Jaquar
    - Argonne National Laboratory, Mira
    - Université Laval, Colosse
  - Recommendations
    - Define system boundaries
    - ↑ quality = measurements for power distribution unit
    - Define measurement instrument accuracy
    - Capture environmental parameters, e.g., Temp
    - Use a benchmark that runs in an hour or two
- Beta Test- SC'12 Report

SC12 Workshop

- Third Annual Workshop on Energy Efficient High Performance Computing - Redefining System Architecture and Data Centers
- Workshop Speakers:
  - Peter Kogge, University of Notre Dame
  - John Shalf, Lawrence Berkeley National Laboratory
  - Satoshi Matsuoka, Tokyo Institute of Technology
  - Herbert Huber, Leibniz Supercomputing Centre
  - Steve Hammond, National Renewable Energy Laboratory
  - Nicolas Dube, Hewlett Packard
  - Michael Patterson, Intel
  - Bill Tschudi, Lawrence Berkeley National Laboratory
- Sunday, November 11th

My two cents-
Location, location, location…

75% of Top500 Supercomputers are located in three countries



Shift from ‘energy’ as an upper bound

- First and foremost, drive system design for energy efficiency
  - low power processor designs, on and off chip interconnect awareness, data locality management, memory photonics and 3D stacking, optical silicon circuit photonic interconnects
- Build a data center with PUE ~ 1 and ERE << 1
- Use renewable energy sources only
- Site where electricity cost is ~$0.03 kWatt/hour
- Which is better?
  - A 20MW, $20
  - B 60MW, $20 with renewable energy and energy re-use

Grace Hopper Inspiration

nersc.gov

Thank you!

- Questions, comments, critique?

Back-up

The Green Grid: Enterprise and Business Focus

- Mission: To become the global authority on resource efficient data centers and business computing ecosystems
- Membership: Spans from Contributor to Individual Levels
- Benefits: Access to relevant content, tools and resources, consultants, networking

Example: PUE and Cooling

	PUE
Case A: Both building and IT fans.	Medium
Case B: Only IT fans.	Lowest
Case C: Only building fans.	Highest

PUE definition includes IT in numerator and denominator
=> Lower PUE if IT cooling fans

HPC Energy Re-use List

- **9. Do you use any heat re-use techniques: e.g. tri-generation, building heating, commercial heat consumers, other.**
- Less than 30% of answers indicated any use of the waste heat generated in the Data Centre.
- In all cases the generated heat was used for heating buildings or providing the heat for urban heating networks.
- There are however more than 60% of answers indicating that there are plans for the future resulting in either expanding existing systems or implementing new ways of reusing the heat.
- Apart from using the heat directly in buildings there are also plans to use trigeneration and adsorption chillers.

http://www.prace-project.eu/IMG/pdf/hpc-centre-cooling-whitepaper.pdf