Re-Visiting Power Measurement for the Green500

Thomas R. W. Scogland (LLNL/CASC, Green500)

Http://www.green500.org

New Considerations for the Level 1 Measurement Methodology The Green500 List and its Continuing Evolution BoF, November 2014

EE HPC Working Group http://eehpcwg.lbl.gov/

Level I Requirements

- Workload phase: Measure at least 20% of the middle 80% of the core phase
- Machine fraction: Measure at least 1/64th of the system or 1kW, whichever is greater
- Subsystems measured: Measure the compute components, network, storage and other subsystems are not required



EE HPC Working Group http://eehpcwg.lbl.gov/





Working Group http://eehpcwg.lbl.gov/

EE HPC

The Green500 List and its Continuing Evolution BoF, November 2014

The Core Phase

- The time period under test
- Possible core phases:
 - Job scheduling -> Job completion
 - Application start -> application end
 - Benchmark start -> benchmark end
- Any is valid, so long as it matches your other metrics



EE HPC Working Group



EE HPC

http://www.green500.org

What do we require now?

EE HPC Working Group



Workload Timing by Measurement Level



Working Group

EE HPC



Power Variability



EE HPC



Why Change the Requirement?

EE HPC Working Group http://eehpcwg.lbl.gov/



Newer system designs have a different pattern.



EE HPC Working Group

Piz Daint (GPU accelerated) Linpack Profile



Working Group http://eehpcwg.lbl.gov/

EE HPC

The Green500 List and its Continuing Evolution BoF, November 2014

11 **EGR** http://www.green500.org

Core Phase Averaged for Piz Daint



EE HPC

12 http://www.green500.org

Core Phase Averaged for Piz Daint



EE HPC

http://www.green500.org

EGR

13

What do we propose?

I4 **FGREEN** 500€ http://www.green500.org

EE HPC Working Group

Workload Timing by Measurement Level



EE HPC

EGR http://www.green500.org

15

Measurement Fraction

- Level I requires 1/64th of the machine
- Which 64th of the machine?



EE HPC Working Group

Variability Across Levels: SuperMUC

Quality Level	Mflops/Watt full run	Efficiency Drop From Level 1
L1 (compute only)	1055	0
L2 (>10kW) (compute and interconnect)	1011	44 (~4%)
L2 (>1/8) (compute and interconnect)	994	61 (~6%)
L3 (compute, interconnect, storage, cooling, power distribution)	887	168 (~16%)



EE HPC Working Group http://eehpcwg.lbl.gov/

A Power-Measurement Methodology for Large-Scale, High-Performance Computing, International Conference on Performance Engineering, March 2014

Subsystem Contribution

- Networks have been considered "in the noise" by Level 1 to this point
- We have increasing reports of the network contributing 10-20% of overall power use





A Power-Measurement Methodology for Large-Scale, High-Performance Computing, International Conference on Performance Engineering, March 2014

Conclusions

- Our current requirements for level I are no longer sufficient
- We propose raising the requirements of Level I:
 - Measurement phase: 100% of the core phase
 - System fraction: I/I6th or more

EE HPC

Working Group

http://eehpcwg.lbl.gov/

• Subsystems included: Compute and networking

