Erich Strohmaier, Lawrence Berkeley National Laboratory and TOP500
Wu Feng, Virginia Tech and Green500

with,
Natalie Bates, EE HPC Working Group,
Michael Patterson, Intel and The Green Grid
And others on the Compute System Metrics Team

SETTING TRENDS FOR ENERGY-EFFICIENT SUPERCOMPUTING

SC13 Workshop; November 2013; Denver

METHODOLOGY: Erich Strohmaier

UNIFY AND IMPROVE METHODOLOGY

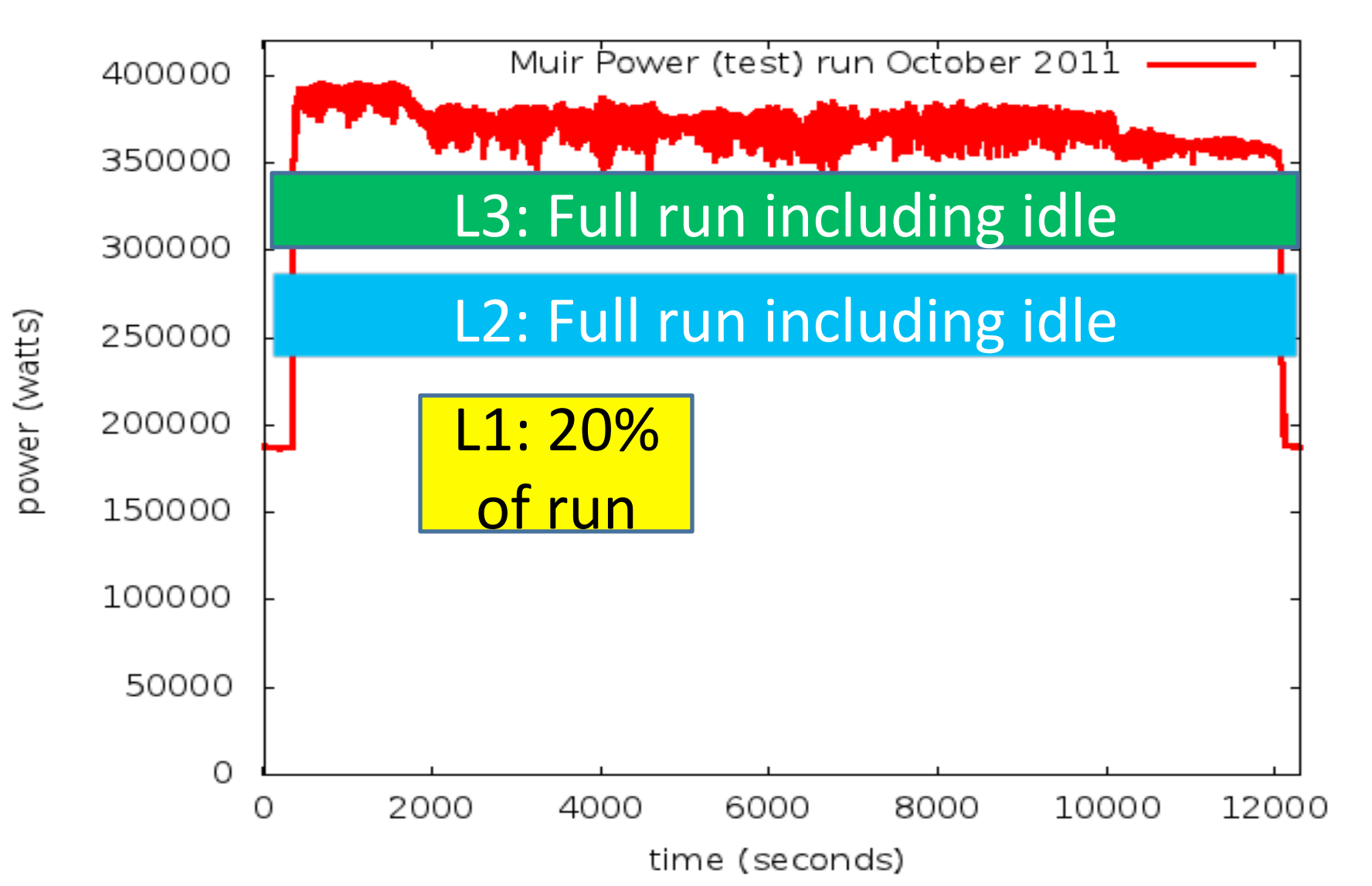
- Green500 & TOP500 power-measurement methodology issues and concerns
  - Variation in start/stop times as well as sampling rates
  - Node, rack or system level measurements
  - What to include in the measurement (e.g., integrated cooling)
- Flexible, but compromises consistency between submissions

ADD QUALITY LEVELS AND REFINE ASPECTS

- Three quality levels
  - Level 1 (L1): basic measurement
  - Level 2 (L2): reasonable effort
  - Level 3 (L3): current best

- Four aspects for each level
  - Aspect 1: frequency and time extent of measurement
  - Aspect 2: system fraction actually measured
  - Aspect 3: subsystems included
  - Aspect 4: power measurement location

Aspect 1: Time Extent



Aspect 1: Sampled Data Frequency

Level 3: (L3)

- "Continuously integrated" energy (≥ 120 samples per second)

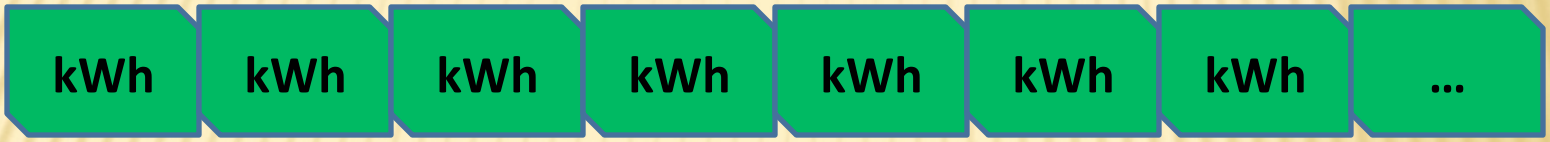
Level 1 and Level 2 (L1 and L2)

- Average power at least once per second

These are *sampling* rates. Data at this rate is typically not seen directly, it's internal to the device.

Aspect 1: Reported Data Requirements

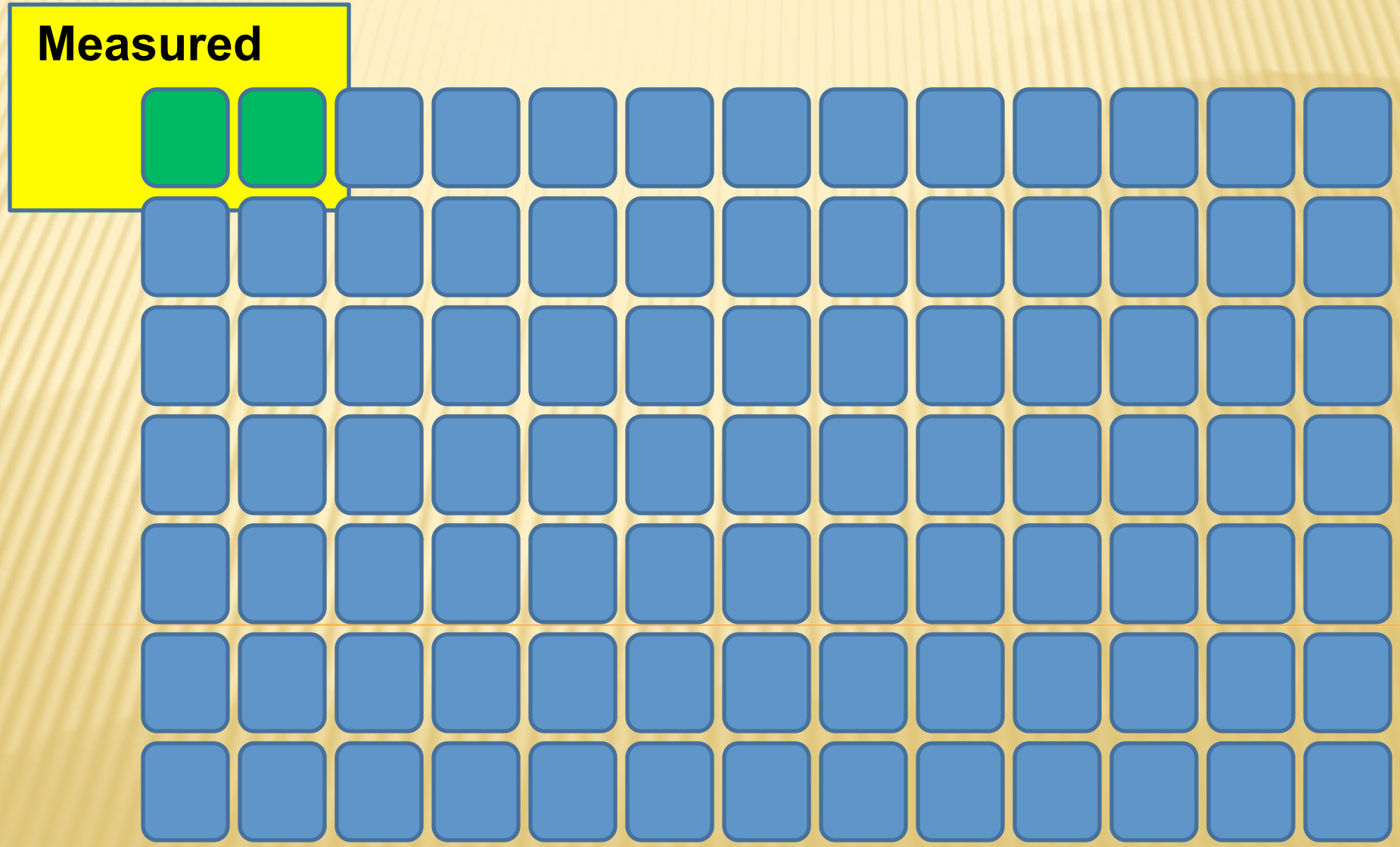
L3: at least 10 reported integrated energy values



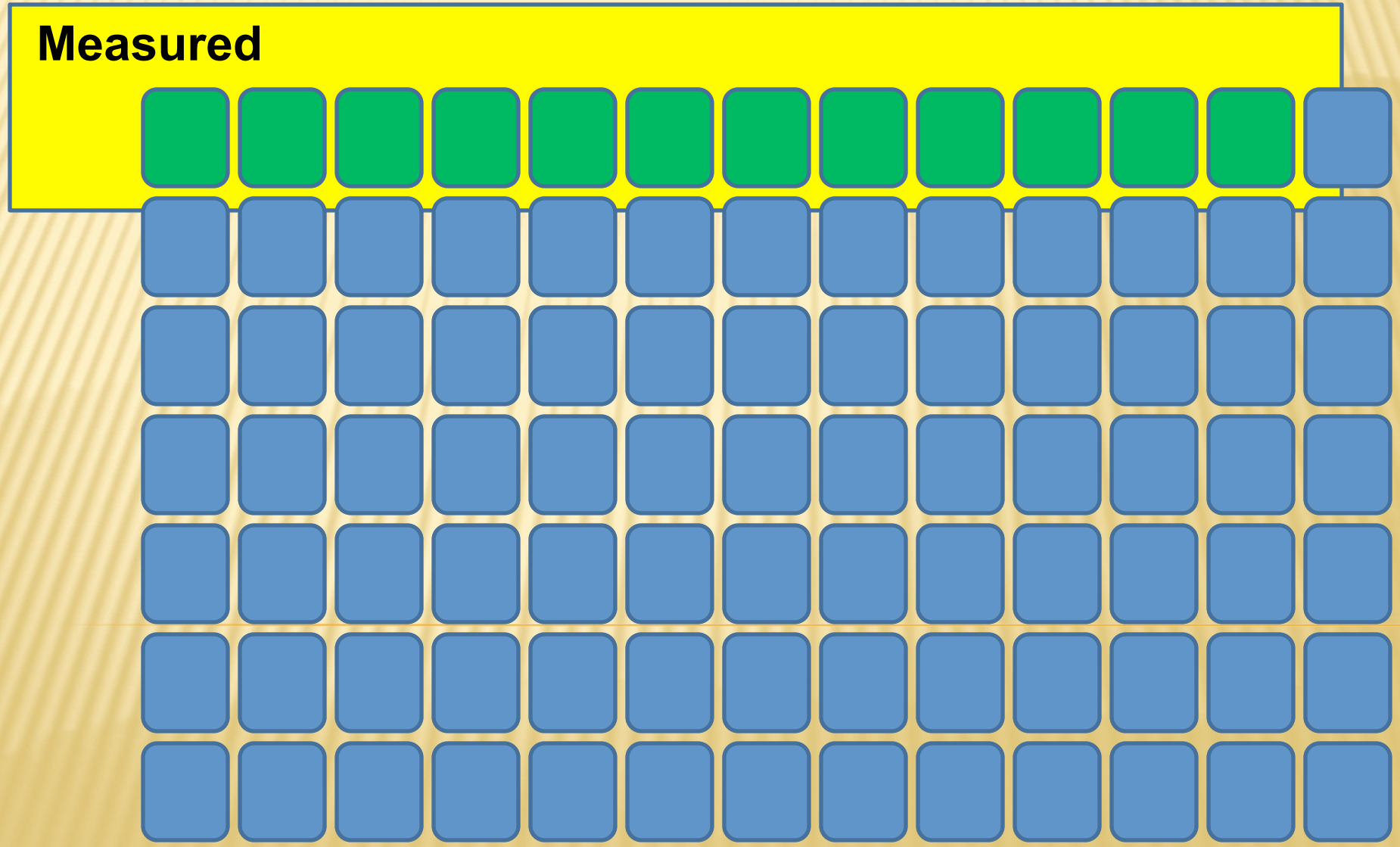
L2: at least 10 power averaged values

L1: at least one power averaged value

Aspect 2: Machine Fraction
L1: at least 1/64 or 1 kW

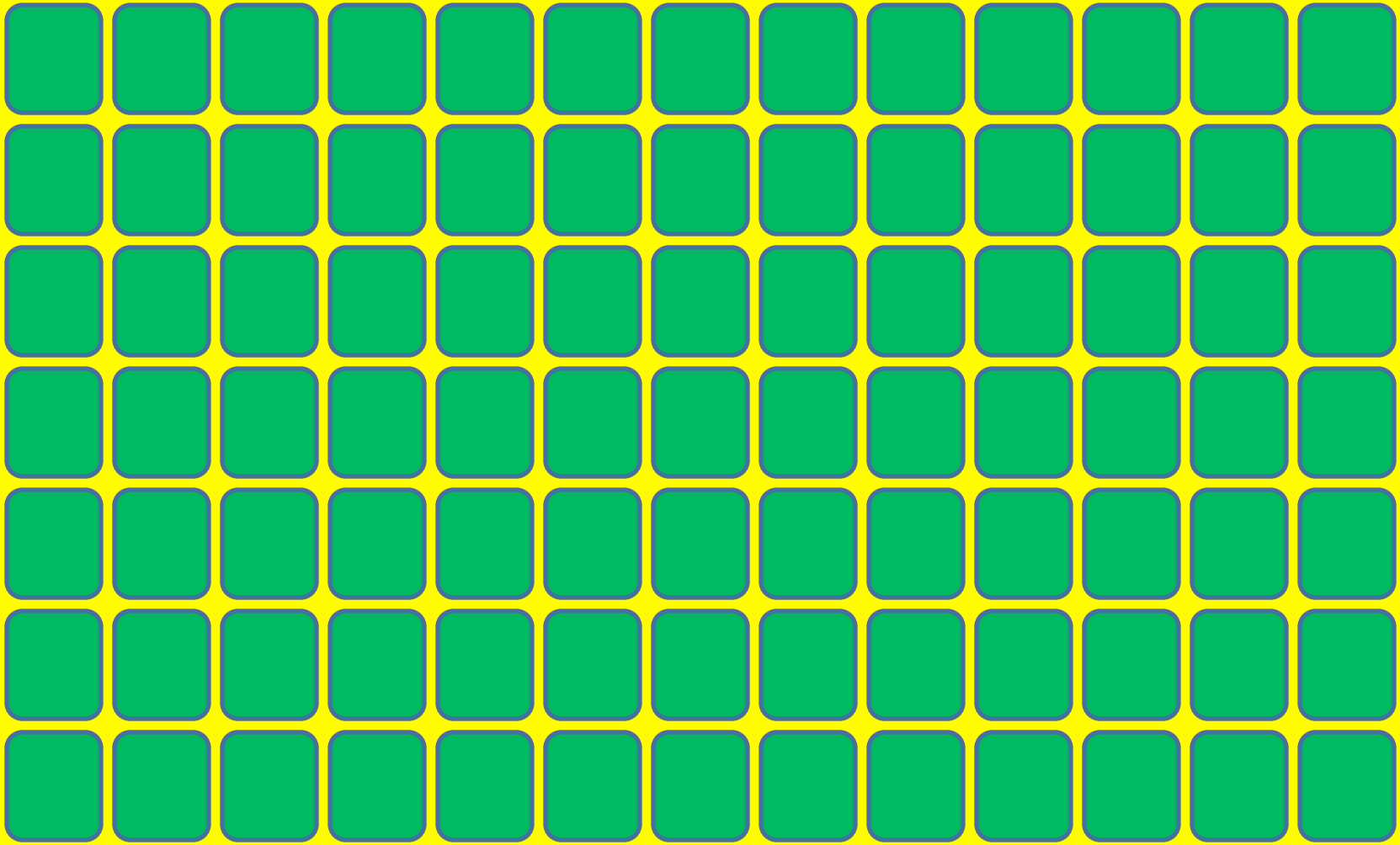


Aspect 2: Machine Fraction
L2: at least 1/8 or 10 kW



Aspect 2: Machine Fraction
L3: whole machine

Measured



Aspect 3: Subsystem Inclusion

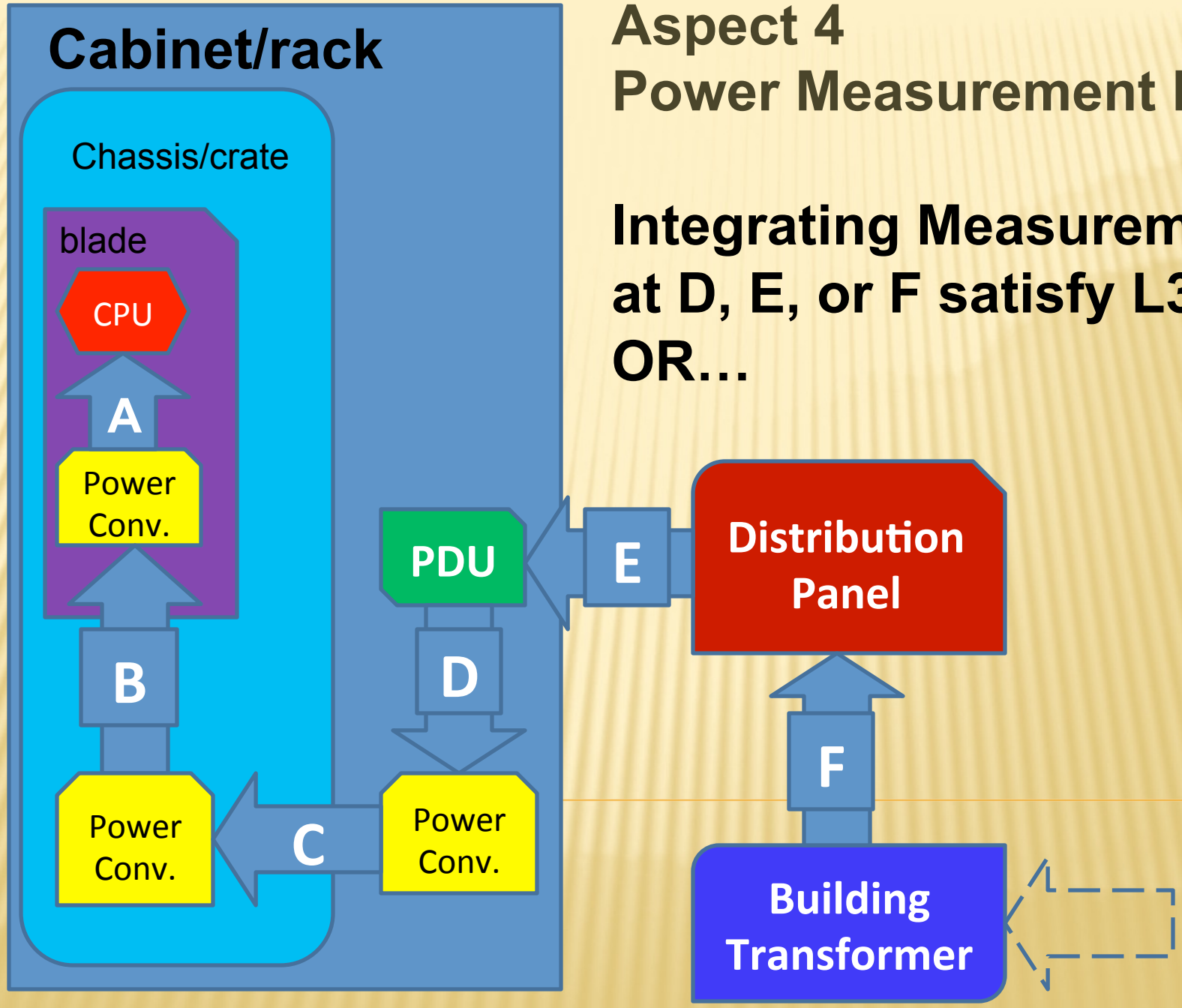
General philosophy: include all parts of computational system that participate in the workload

Must include:

- Processors, memory, cooling power internal to the machine (fans, etc.)
- Internal Interconnect network
- Login/compile nodes

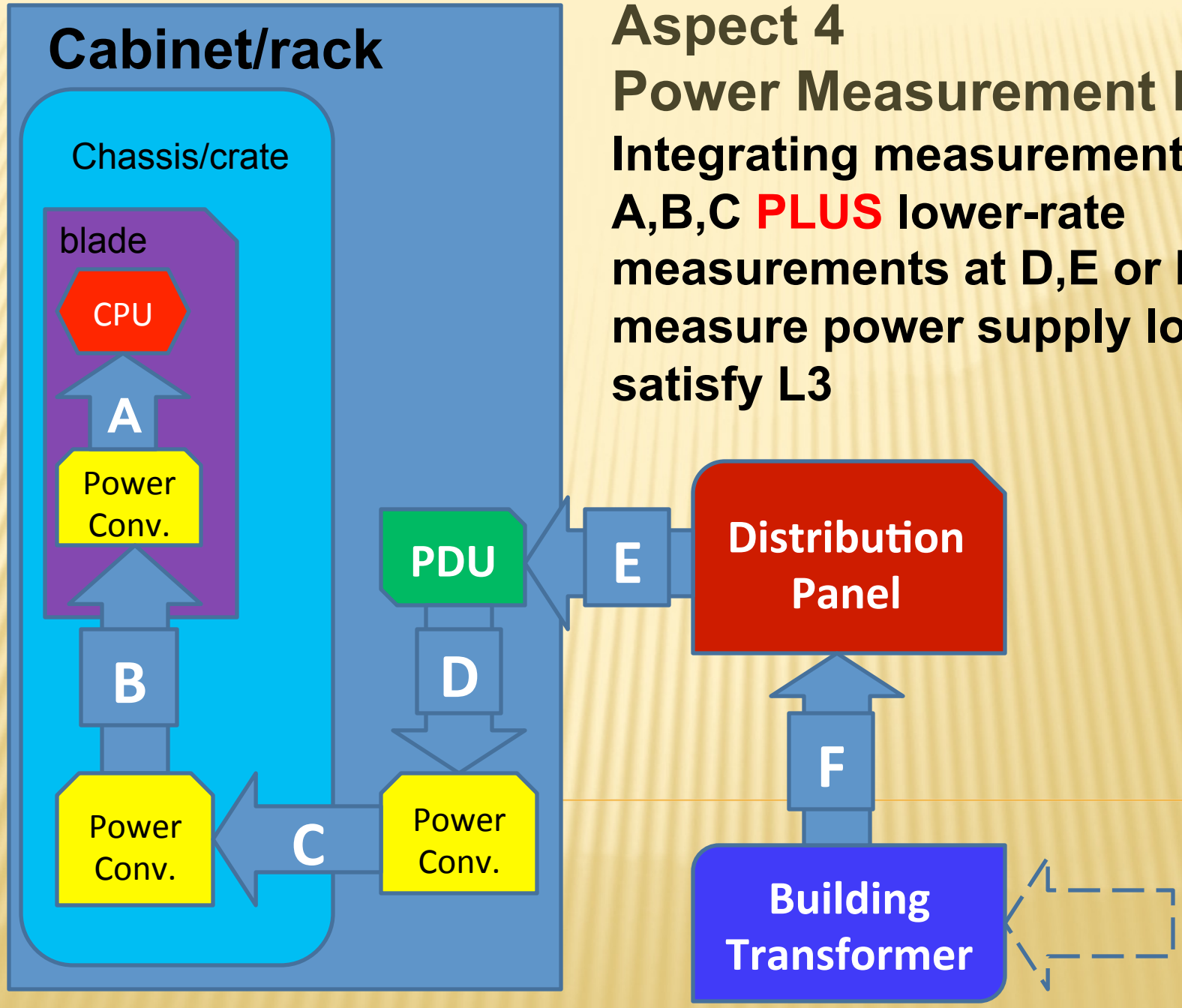
Aspect 4
Power Measurement Point

Integrating Measurements at D, E, or F satisfy L3 OR…



Aspect 4
Power Measurement Point:

Integrating measurements at A,B,C PLUS lower-rate measurements at D,E or F (to measure power supply losses) satisfy L3



ADOPTION: Wu Feng

WHERE TO FIND THE METHODOLOGY

THE GREEN 500™

EEHPC WG: Power Measurement Methodology

Click the link below to download the EEHPC WG Power Measurement Methodology to find out more about Level 2 and Level 3 measurements.

Download the EEHPC WG: Power Measurement Methodology Document (PDF)

JUNE 2013: GREEN500 RELEASES NEW METHODOLOGY

- Green500 accepts higher-precision measurements, denoted as Level 2 and 3
- *"Higher quality measurements… provide much better picture of the real-world costs… as well as a more in-depth picture of how the system handles a Linpack run."* Green500 press release

JUNE 2013 GREEN500 LISTS

Site*	Computer*
DOE/NNSA/LLNL	Sequoia-25 - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom Level 2 measurement data available
Leibniz Rechenzentrum	SuperMUC - iDataPlex DX360M4, Xeon E5-2680 8C 2.70GHz, Infiniband FDR Level 3 measurement data available
Maui High-Performance Computing Center (MHPCC)	Riptide - iDataPlex DX360M4, Xeon E5-2670 8C 2.600GHz, Infiniband FDR Level 3 measurement data available
Universit Laval	Colosse - Sun Blade x6048, Xeon X5560 2.8 Ghz, Infiniband QDR Level 3 measurement data available

Level 2/3 measurement data available

PROGRESS SINCE ISC BOF

- June'13 Green500 List has four L2/L3 Submissions
- Further testing and review of methodology
  - Swiss Supercomputing Center (CSCS) L3 measurement and attempted submission
  - Peer review of paper

ISSUES TO RESOLVE

- *Refine methodology*
  - *System boundary, e.g., file system*
  - *Environmentals, e.g., air+liquid cooling*
  - *Measurement instrument specification, accuracy and precision*
- *Identify workloads for exercising other sub-systems; e.g., memory, storage, I/O*
- *Still need to decide upon metrics*
  - *Classes of systems (e.g., Top50, Little500, technologies)*
  - *Multiple metrics or a single index*

EXTRA SLIDES

Early Adopters and Testers

- Lawrence Livermore National Laboratory
- Leibniz Supercomputing Center
- Oak Ridge National Laboratory
- Argonne National Laboratory
- Universite Laval, Calcul Quebec, Compute Canada
- University of Jaume
- University of Tennessee
- CEA
- National Center for Atmospheric Research
- Maui High Performance Computing Center
- Swiss National Supercomputing Center