

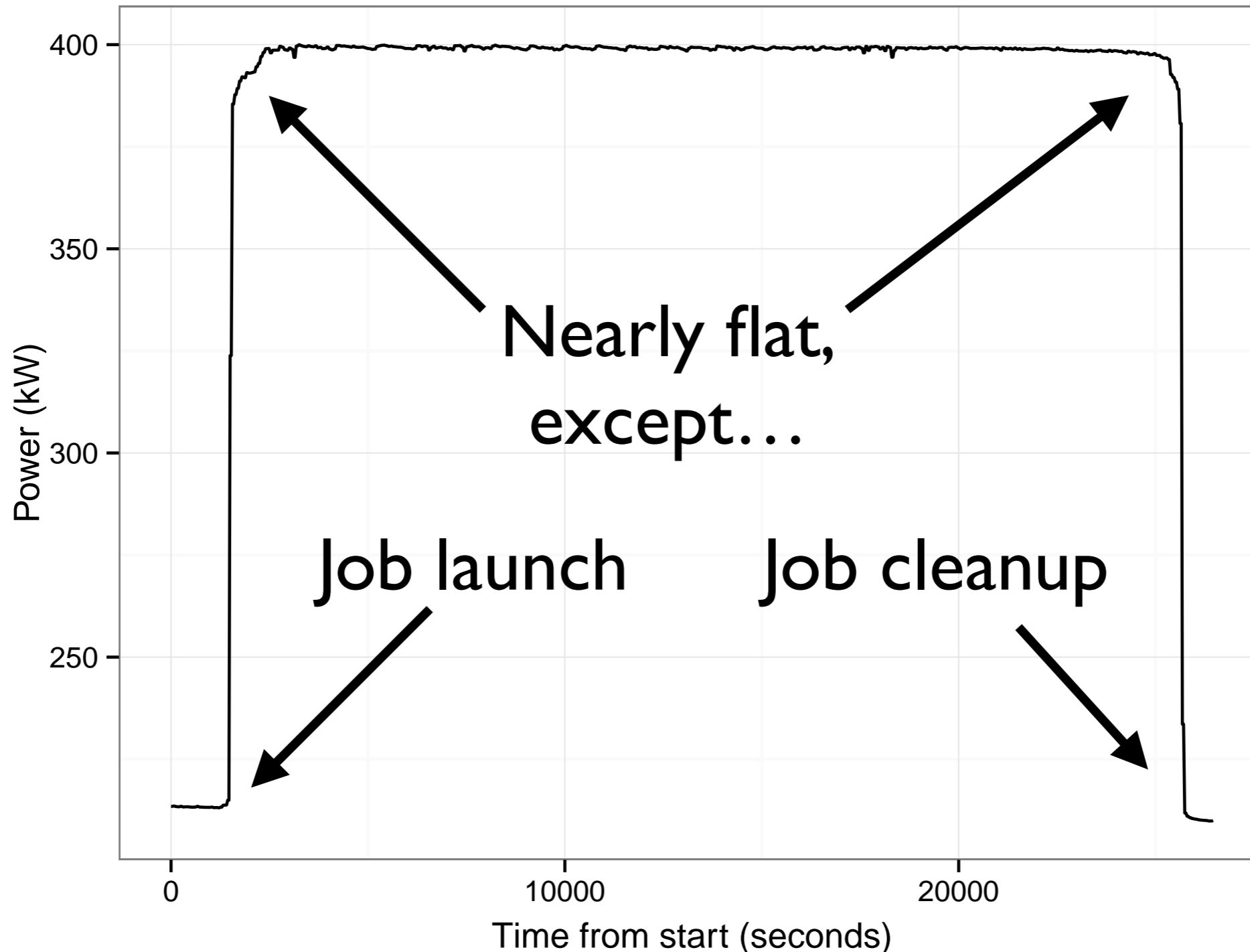
Re-Visiting Power Measurement for the Green500

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

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

Workload Phase: A classic HPL Profile

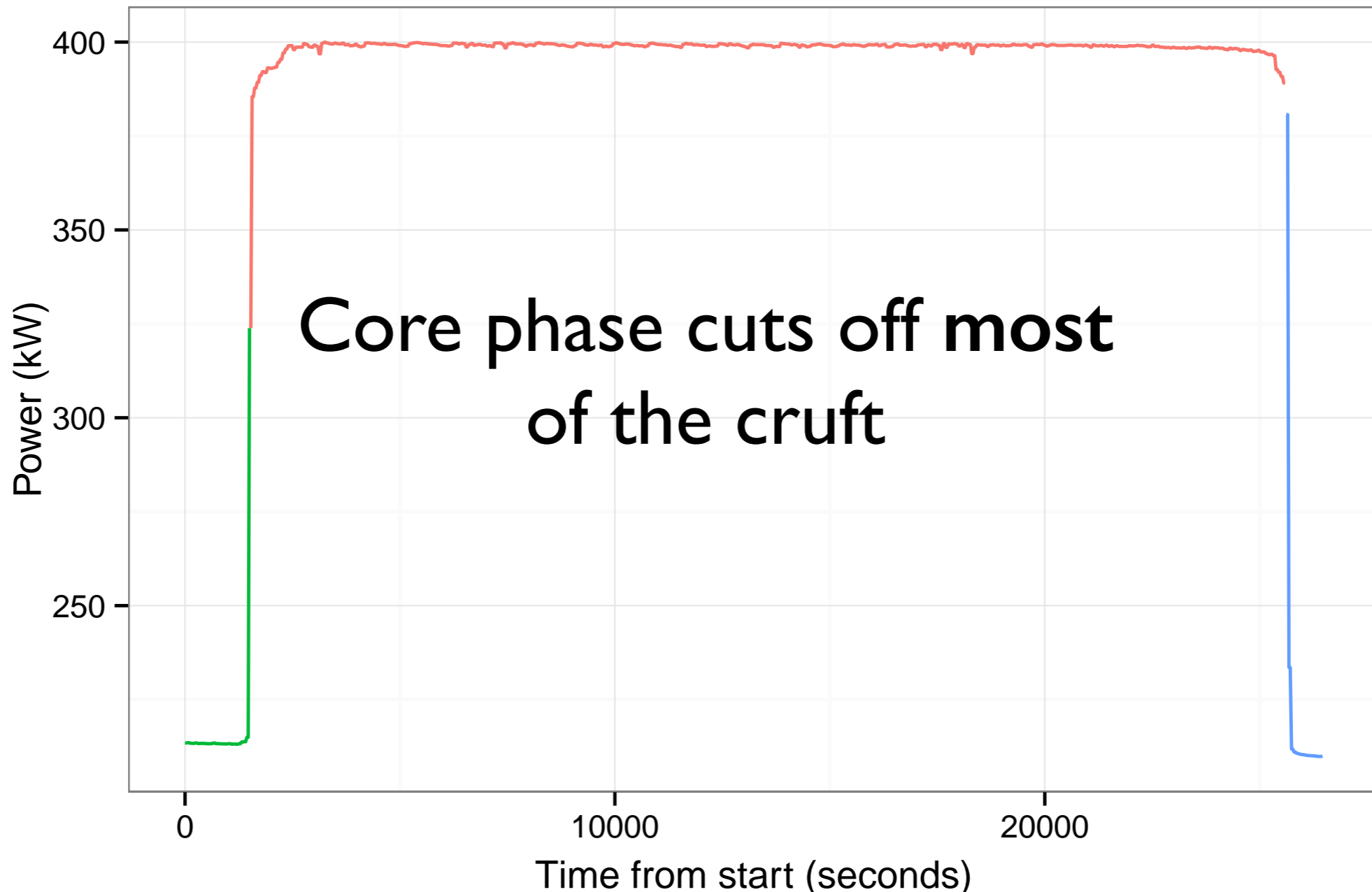


The Core Phase

- 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

The Core Phase: Linpack Example

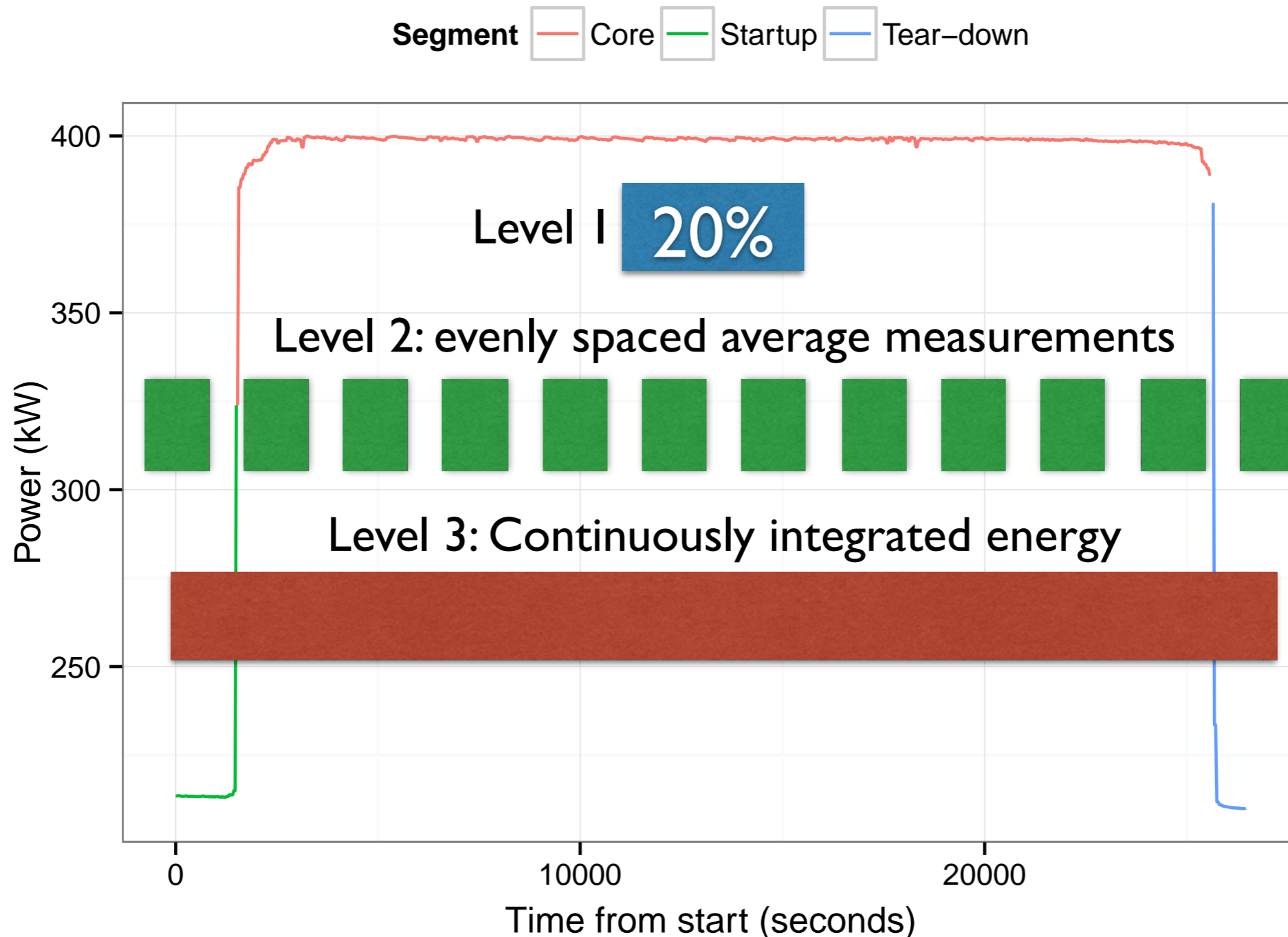
Segment — Core — Startup — Tear-down



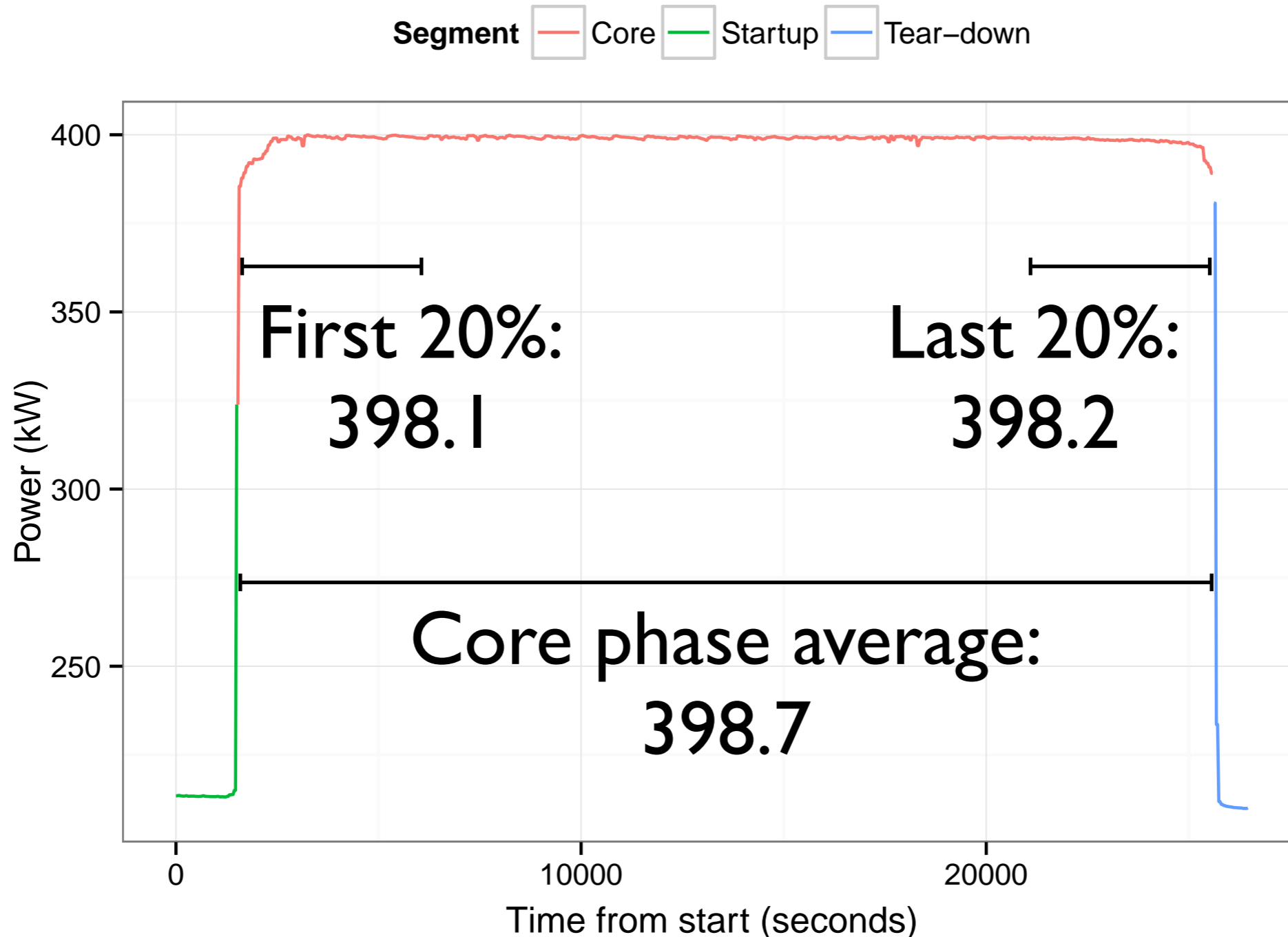
**Core phase cuts off most
of the cruft**

What do we require now?

Workload Timing by Measurement Level



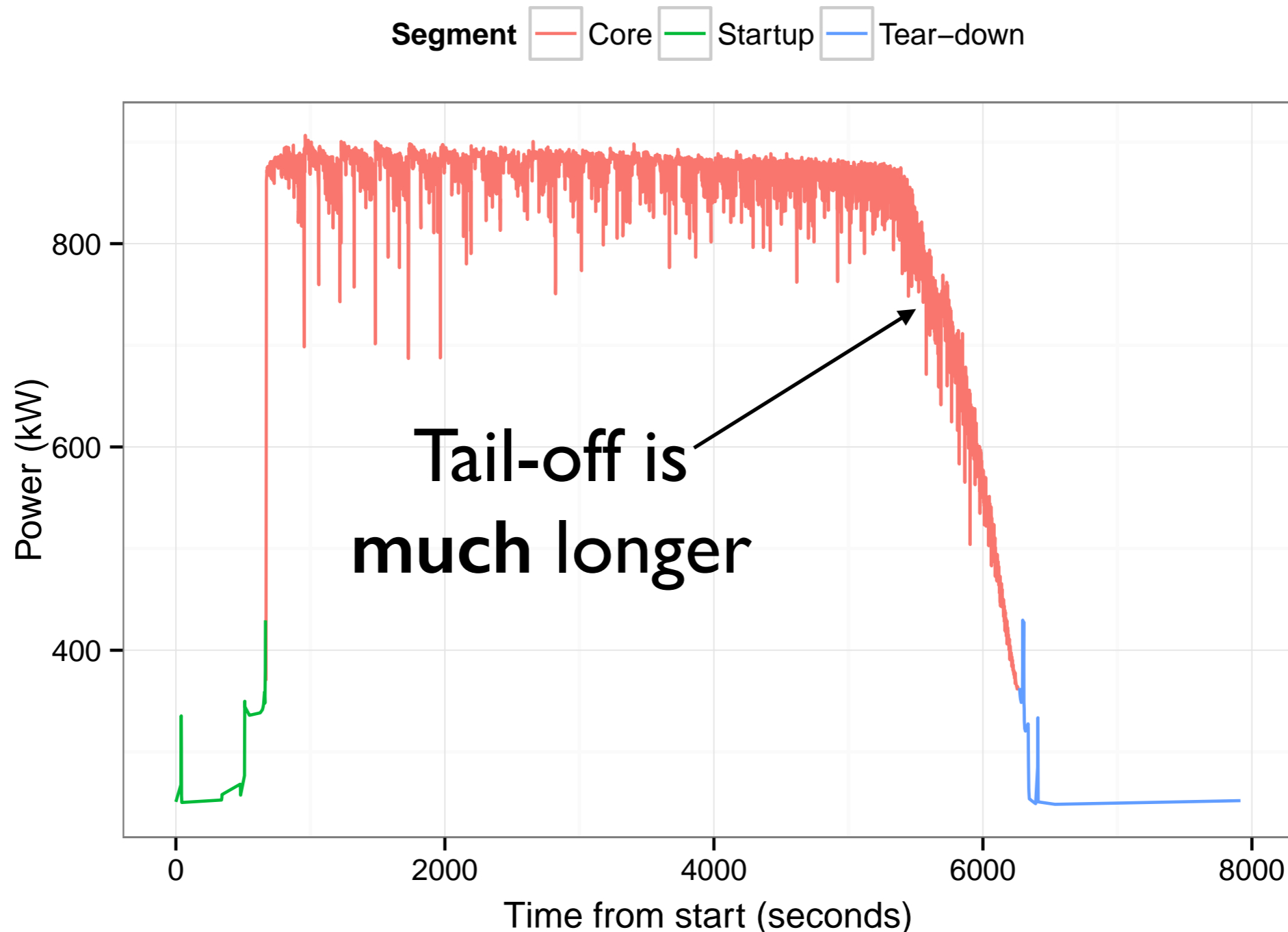
Power Variability



Why Change the Requirement?

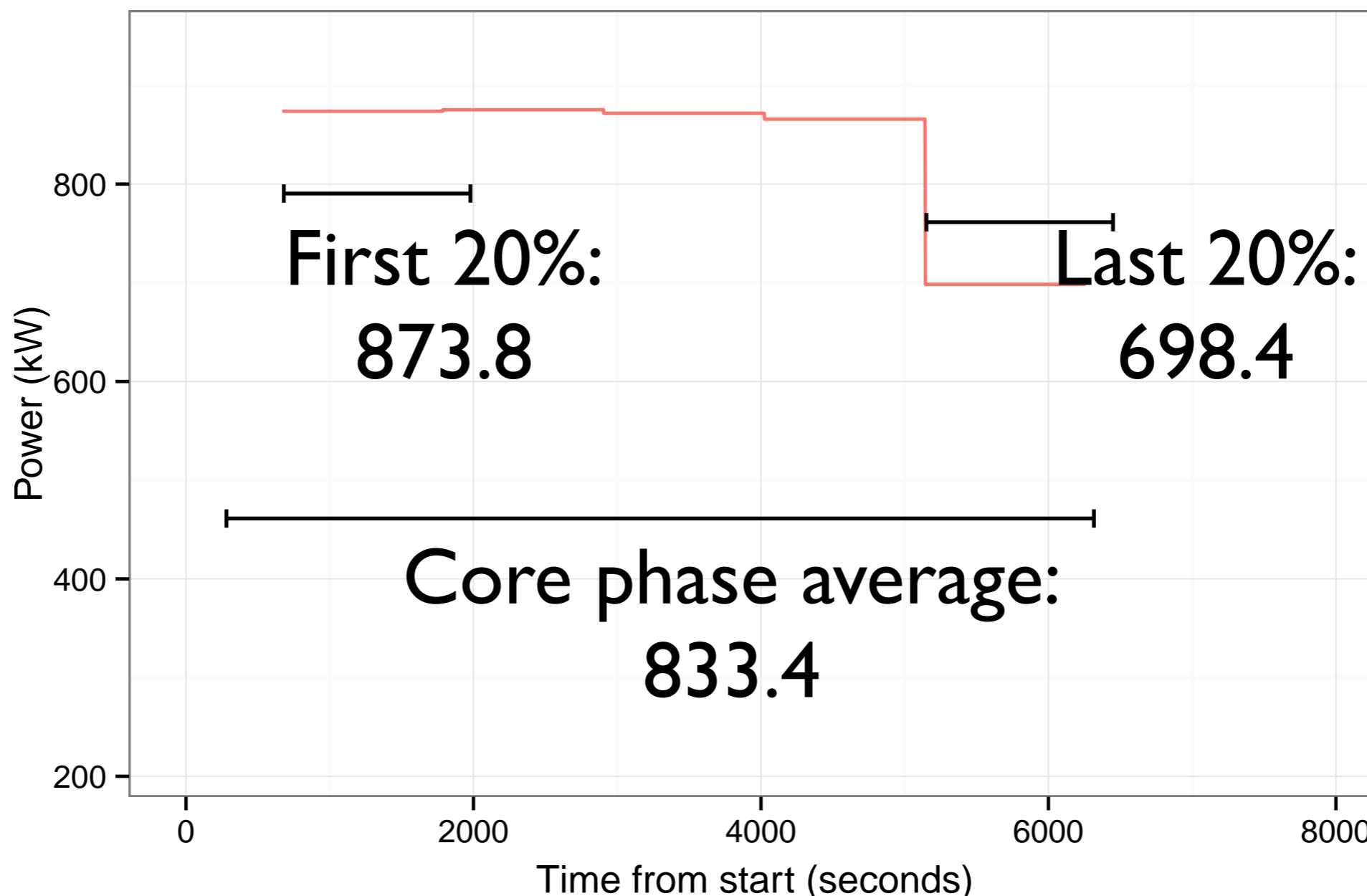
Newer system designs
have a different pattern.

Piz Daint (GPU accelerated) Linpack Profile



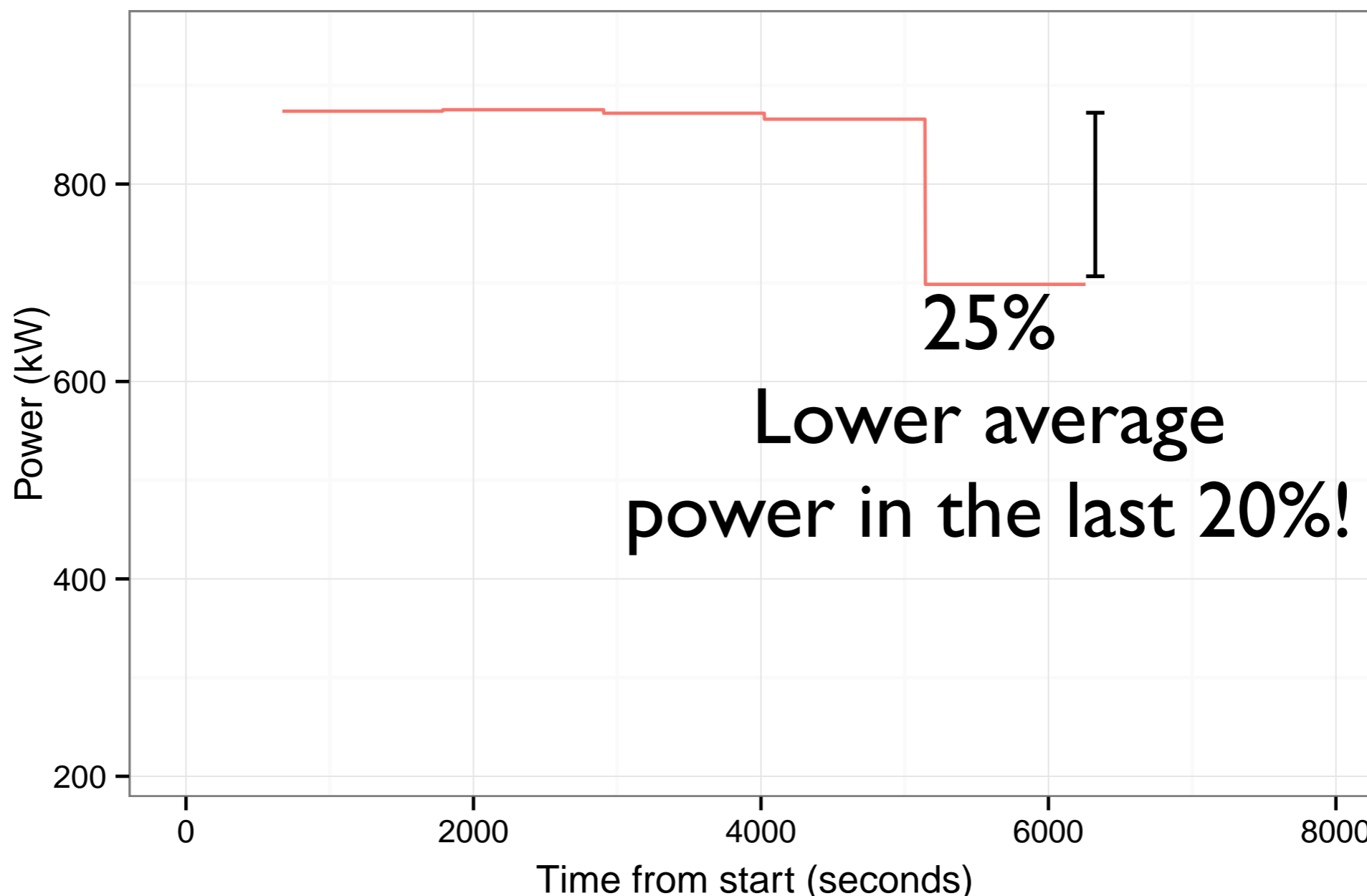
Core Phase Averaged for Piz Daint

Segment — Core — Startup — Tear-down



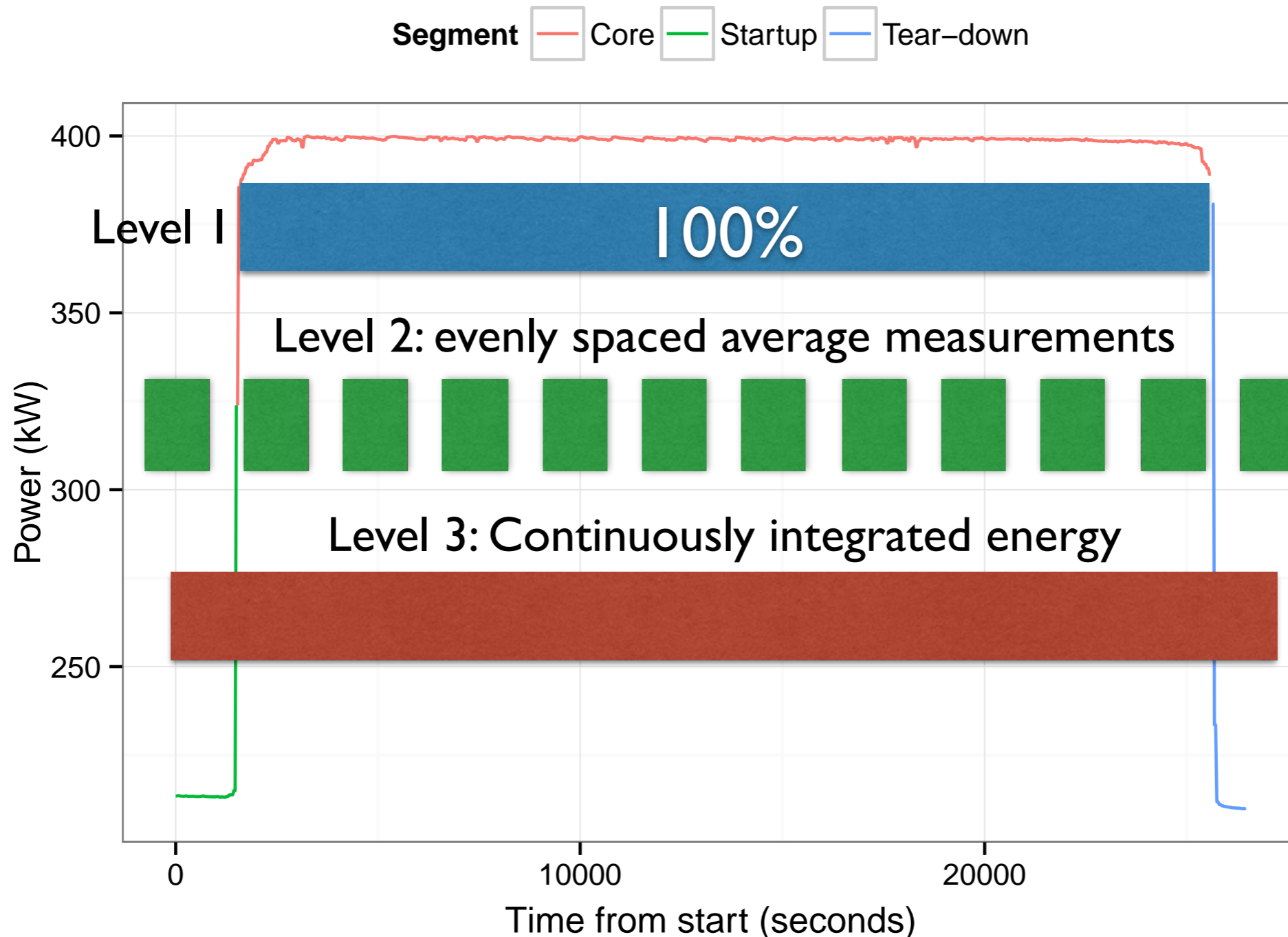
Core Phase Averaged for Piz Daint

Segment — Core — Startup — Tear-down



What do we propose?

Workload Timing by Measurement Level



Measurement Fraction

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

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%)

Subsystem Contribution

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

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: 1/16th or more
 - Subsystems included: Compute and networking