

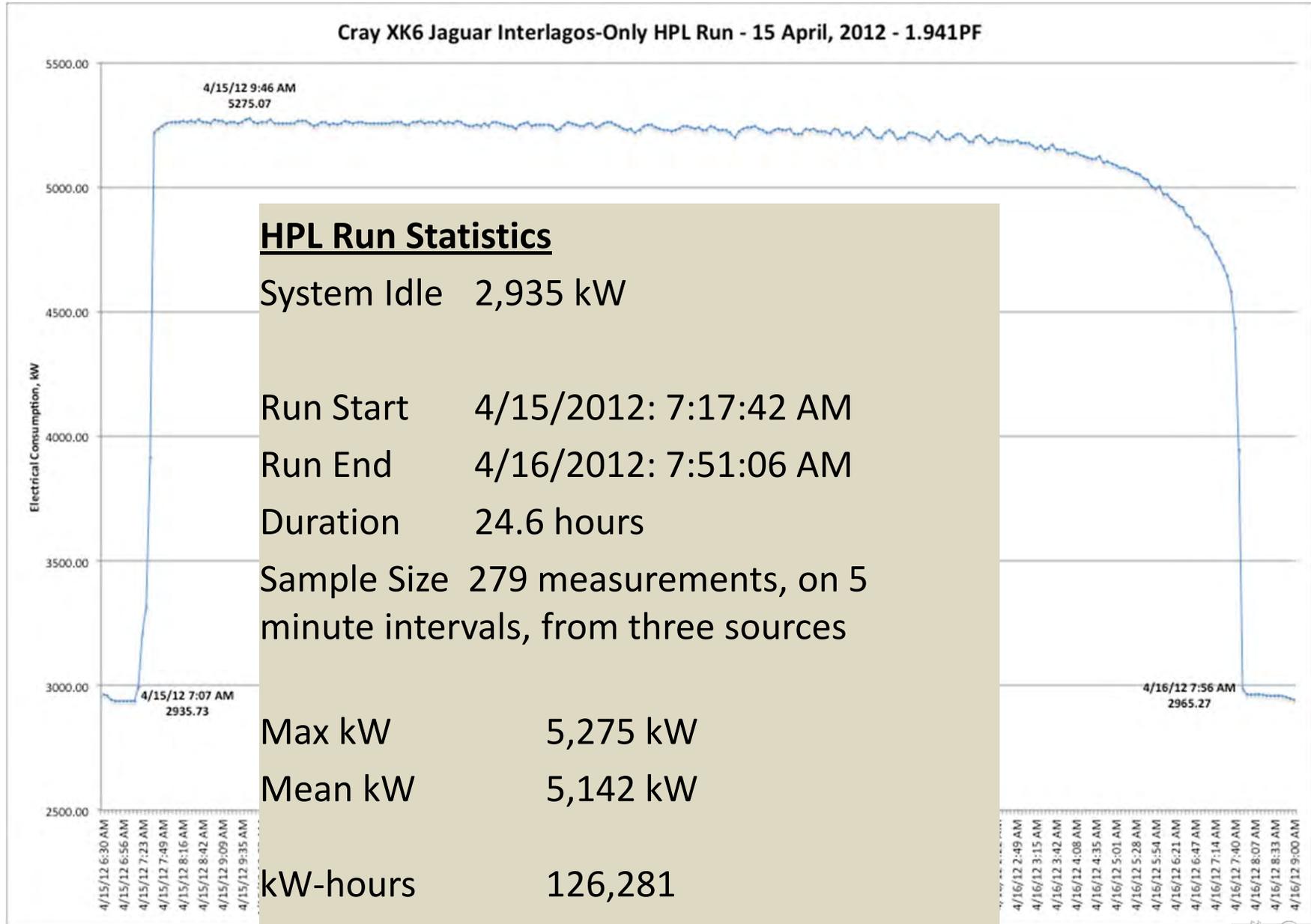
# ORNL's "Titan" System

- Upgrade of Jaguar from Cray XT5 to XK6
- Cray Linux Environment operating system
- Gemini interconnect
  - 3-D Torus
  - Globally addressable memory
  - Advanced synchronization features
- AMD Opteron 6274 processors (Interlagos)
- New accelerated node design using NVIDIA multi-core accelerators
  - 2011: 960 NVIDIA x2090 "Fermi" GPUs
  - 2012: 14,592 NVIDIA "Kepler" GPUs
- 20+ PFlops peak system performance
- 600 TB DDR3 mem. + 88 TB GDDR5 mem



Titan Specs	
Compute Nodes	18,688
Login & I/O Nodes	512
Memory per node	32 GB + 6 GB
# of Fermi chips (2012)	960
# of NVIDIA "Kepler" (2013)	14,592
Total System Memory	688 TB
Total System Peak Performance	20+ Petaflops
Liquid cooling at the cabinet level	Cray EcoPHLex

# April 15, 2012 Top 500 Submission: Jaguar



# Assessment of April 15, 2012 HPL Submission

## Energy Efficient HPC System Workload Power Measurement Methodology

### – Aspect 1: **Level 2, Level 3 is available**

- Eaton IQ Analyzer sampling at up to 8 times per second
- Total energy is available directly from the unit
- Sample is the instantaneous measurement at that time
- The typical measurement interval for historical purposes is a 5-minute sample. Shorter measurement period of 1-minute samples are frequently used for analysis of consumption during full machine runs (HPL and others).

### – Aspect 2: **Level 3**

- All 200 cabinets were measured from three main switchboards.

### – Aspect 3: **Level 3**

- Power metering is at the three switchboards, not at the individual devices.
- More accurate assessment of total consumption, including line losses in the 200+48 480V branch circuits.
- All 48 Liebert XDPs included in the measurement (Cray EcoPhlex closed loop cooling system)
- Not included in the measurement:
  - Chilled water cost
  - External parallel file system
  - External login nodes

Report and Analyze	Level 1	Level 2	Level 3
Aspect 1: requirements of measured values for ac measurement	1 instantaneous power measurement per second	1 instantaneous power measurement per second	continuously integrated total energy
Aspect 1: requirements of reported values for submission	one average power covering at least 20% of the run	time series of equally-spaced averaged power values	time series of equally spaced total energy values
Aspect 2: machine fraction	at least 1/64 of the machine or 1 kW	at least ? of the machine or 10 kW	whole machine
Aspect 3:	subsystems included	subsystems included	subsystems included
Aspect 3:	Point in power distribution where measurement is taken	Point in power distribution where measurement is taken	Point in power distribution where measurement is taken
required analyzed values for submission	core phase average power	core phase average power and whole application average power	core phase average power and whole application average power



# Metering Capabilities for HPC Systems at ORNL



- Every electrical service delivery system (switchboard, panel, PDU, RDU) is metered throughout the computer facility as part of the cost recovery mechanism for the facility.
- Metering at existing switchboards using Eaton IQ Analyzer (installed on main switchboards in 2009)
- IQ A Metered/Monitored Parameters
  - rms sensing.
  - Phase neutral, and ground currents.
  - Volts: L-L, L-N, Avg. L-L, Avg. L-N, N-G.
  - Power: real, reactive, apparent (system and per phase).
  - Frequency.
  - Power factor: apparent and displacement (system and per phase).
  - Energy and demand (forward, reverse, net) real, reactive apparent at four different utility rates.
  - Individual current and voltage harmonics: magnitude, phase angle.
  - % THD: current and voltage.
  - Waveform capture.
  - ANSI C12.20 Class 0.5% revenue metering accuracy, ANSI C12.16, IEC687 Class 0.5%.
- New Capabilities (2012)
  - XFMR\_S36/MSB14 (3.0MVA transformer/switchboard pair) are metered by Schneider Electric CM4000 PowerLogic Circuit Monitor
    - Highly accurate power quality monitor for critical energy systems. Substantially higher performance/capability than original equipment.
    - Provides mechanism for measuring and comparing features against original Eaton baseline, especially potentially troubling aspects including harmonics. Adds very accurate voltage transient and flicker analysis features.
  - Individual cabinet meters on two of the Cray XK6 cabinets. One meter on a non-accelerated XK6 cabinet, and a second meter on a NVIDIA Kepler-accelerated cabinet.

# Recommendations

- Define the boundary of the system for measurement:
  - Disks?
  - Storage Area Network?
  - Cooling, Pumps, Chillers?
  - Transformers, UPS, AC-DC conversion?
- Remember that power measurement is a tool, not an end in itself. We use this to help inform choices, not dictate decisions.