Metadata for Machine Analytics in High-Performance Computing

Line C. Pouchard, Stephen Poole

Extreme Scale Systems Center, Computer Science and Mathematics

Oak Ridge National Laboratory

Why collect machine health data?

- Monitor power consumption
- Improve resource management and scheduling
- Detect application signatures (power, intrusion detection)

What data to collect?

- Temperatures, voltages, current
- Fans, humidity
- What runs where when and for whom?

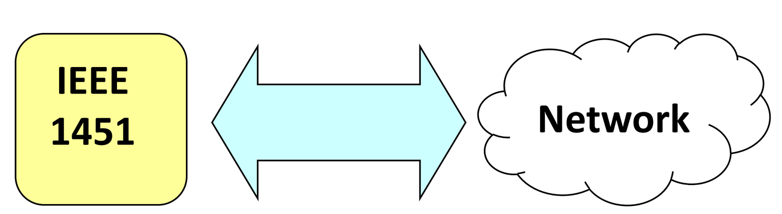
Data Collection Challenges

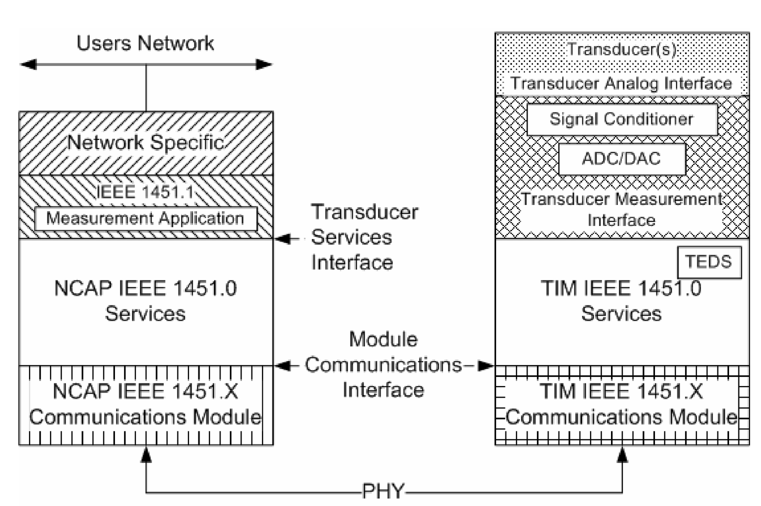
- Massive amount of data streaming from potentially thousands of nodes
- Because of volume, data is analyzed on the fly, only the results are persisted, precluding forensic analysis
- Sensor data variables not always accessible
- Reliance on proprietary instrumentation with no verification and validation.
- Calibration is sometimes ignored

Metadata Challenges

- Heterogeneous sensors and sources
- Many forms of instrumentation with output in multiple formats
- Sensor output embedded in (proprietary) system monitoring tools like CRMS.
- System sensor data is represented with numerous schemas and data models
- No single specification/package satisfies the goal of providing a metadata model suitable to analyze sensor data from all manufacturers.

IEEE 1451 Transducer Electronic Data Sheet





Data structure of a TEDS

- Unsigned integer 32, 4 octets
- MetaTEDS (internal timeout value)
- Transducer Channel (sensor metadata)
- User's Transducer Name
- Frequency response
- Calibration
- Transfer function
- Command (sensor control)
- Geo-location

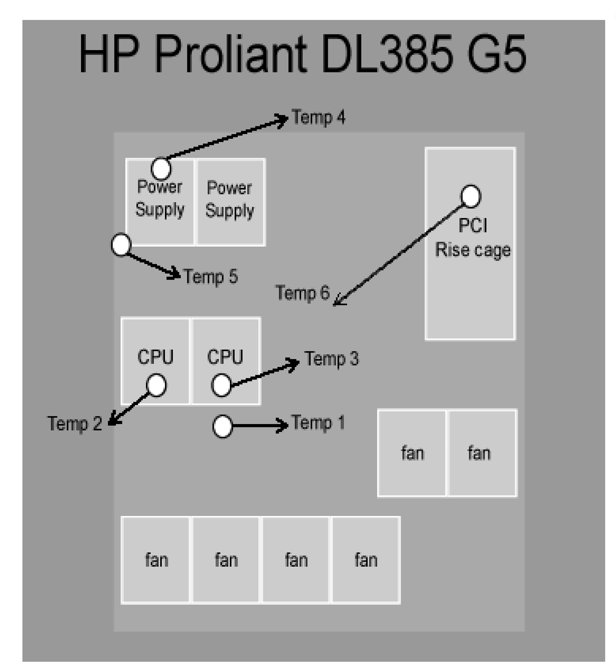
- TIM: Transducer Interface Model
- NCAP: Network Capable Application Processor
- PHY: Physical Connections
- ADC: Analog-to-Digital Conversion
- DAC: Digital-to-Analog Conversion

Intelligent Platform Management Interface (IPMI)

- Adopted by over 200 equipment manufacturers, but nor Cray
- Open source specification, requires drivers and proprietary implementations
- Sensor data is queried via ipmitool –job outputs data once per minute (not customizable, calibration is tricky)

IPMI Output

VRM 1	0 unspecified	cr\par
VRM 2	0 unspecified	cr\par
Fan 1	50.18 unspecifi	nc\par
Fan 2	50.18 unspecifi	nc\par
Fan 3	39.98 unspecifi	nc\par
Fan 4	39.98 unspecifi	nc\par
Fan 5	39.98 unspecifi	nc\par
Fan 6	39.98 unspecifi	nc\par
Temp 1	18 degrees C	ok\par
Temp 2	34 degrees C	ok\par
Temp 3	35 degrees C	ok\par
Temp 4	34 degrees C	ok\par
Temp 5	40 degrees C	ok\par
Temp 6	40 degrees C	ok\par

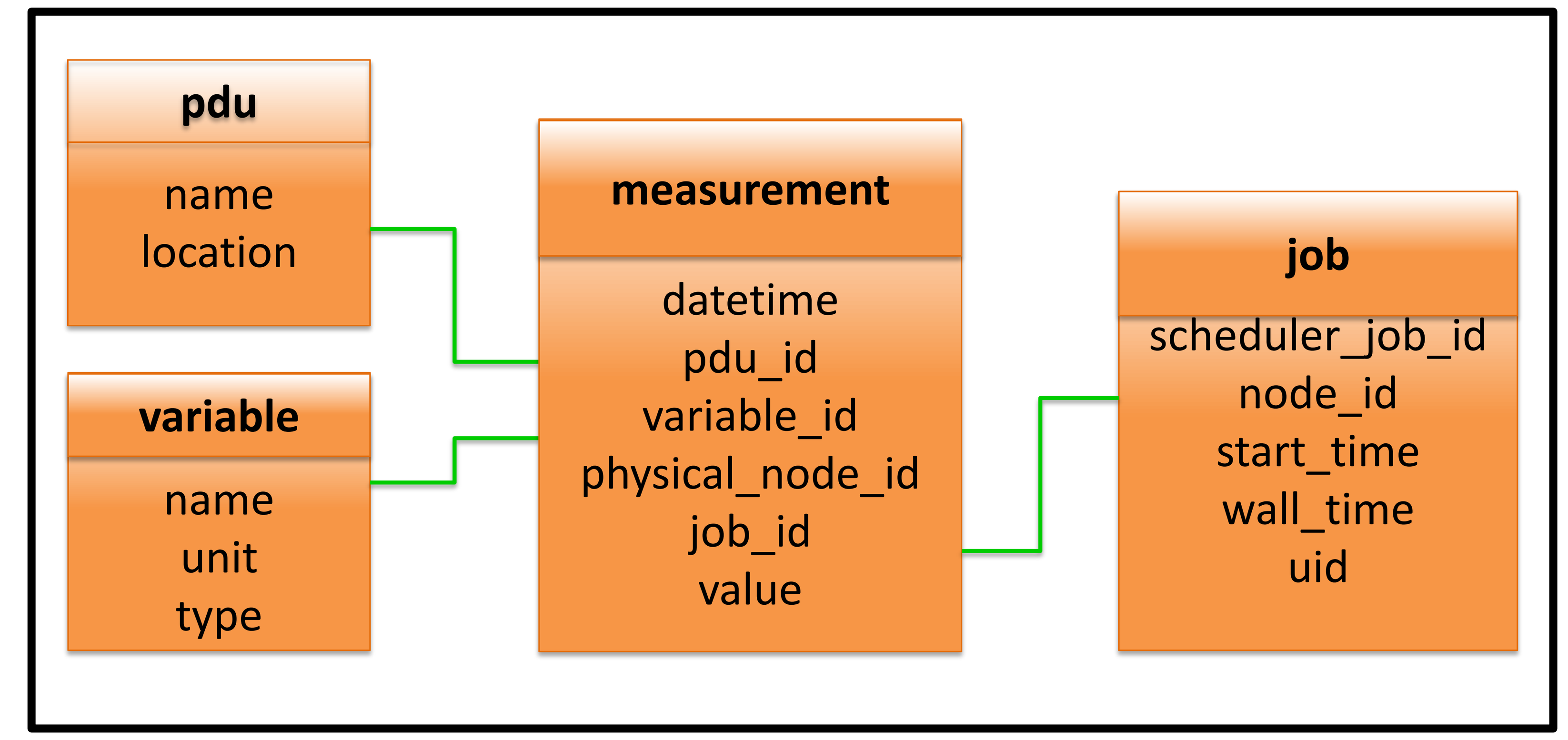


System Environment Data Collection (SEDC)

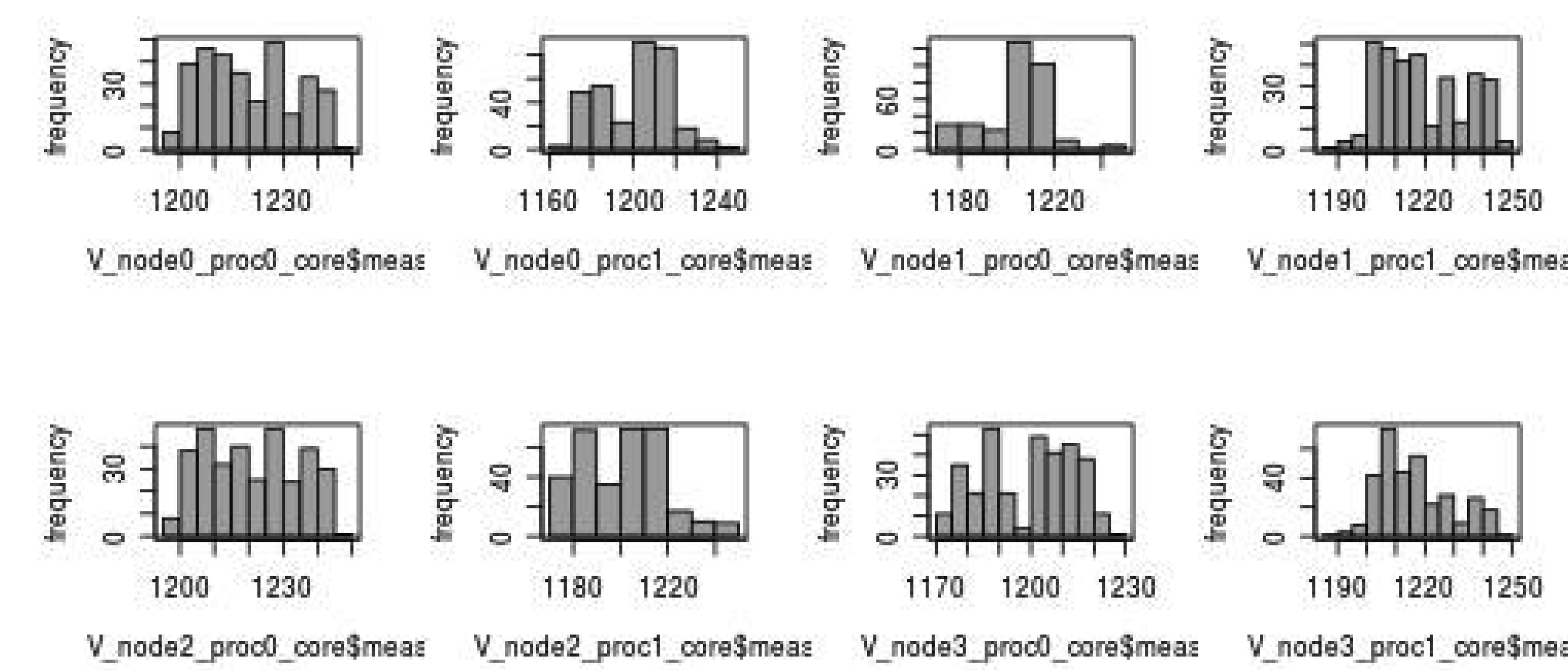
```
c0-5c0s0   L0_V_XT5_TYCO_BUS           | 12763
c0-5c0s0   L0_V_XT5_12V_BUS            | 12689
c0-5c0s0   L0_V_XT5_DISC_SW_IN         | 12763
c0-5c0s0   L0_V_XT5_DISC_SW_MID        | 12763
c0-5c0s0   L0_V_XT5_MEZZ_LDT           |  1228
c0-5c0s0   L0_V_XT5_MEZZ_AVDD          |  2468
c0-5c0s0   L0_V_XT5_MEZZ_CORE          |  1602
c0-5c0s0   L0_V_XT5_MEZZ_5V_BIAS       |  4984
c0-5c0s0   L0_V_XT5_NODE0_VDDA         |  2435
c0-5c0s0   L0_V_XT5_NODE1_VDDA         |  2473
c0-5c0s0   L0_V_XT5_NODE2_VDDA         |  2444
c0-5c0s0   L0_V_XT5_NODE3_VDDA         |  2426
c0-5c0s0   L0_V_XT5_NODE0_12V          | 12744
c0-5c0s0   L0_V_XT5_NODE1_12V          | 12672
c0-5c0s0   L0_V_XT5_NODE2_12V          | 12690
c0-5c0s0   L0_V_XT5_NODE3_12V          | 12690
c0-5c0s0   L0_V_XT5_NODE0_PROC0_VDDIO  |  1792
c0-5c0s0   L0_V_XT5_NODE0_PROC1_VDDIO  |  1790
c0-5c0s0   L0_V_XT5_NODE1_PROC0_VDDIO  |  1795
c0-5c0s0   L0_V_XT5_NODE1_PROC1_VDDIO  |  1798
c0-5c0s0   L0_V_XT5_NODE2_PROC0_VDDIO  |  1792
c0-5c0s0   L0_V_XT5_NODE2_PROC1_VDDIO  |  1798
c0-5c0s0   L0_V_XT5_NODE3_PROC0_VDDIO  |  1798
c0-5c0s0   L0_V_XT5_NODE3_PROC1_VDDIO  |  1795
c0-5c0s0   L0_V_XT5_NODE0_PROC0_VTT    |   890
c0-5c0s0   L0_V_XT5_NODE0_PROC1_VTT    |   899
c0-5c0s0   L0_V_XT5_NODE1_PROC0_VTT    |   896
c0-5c0s0   L0_V_XT5_NODE1_PROC1_VTT    |   908
c0-5c0s0   L0_V_XT5_NODE2_PROC0_VTT    |   902
c0-5c0s0   L0_V_XT5_NODE2_PROC1_VTT    |   902
c0-5c0s0   L0_V_XT5_NODE3_PROC0_VTT    |   905
c0-5c0s0   L0_V_XT5_NODE3_PROC1_VTT    |   899
c0-5c0s0   L0_V_XT5_NODE0_PROC0_VDDNB  |  1207
```

Data Volumes per day XT5	about 10 MB per node	4800 physical nodes
controller node	13 Voltages	13 Temperatures
worker node	56 voltages	55 Temperatures

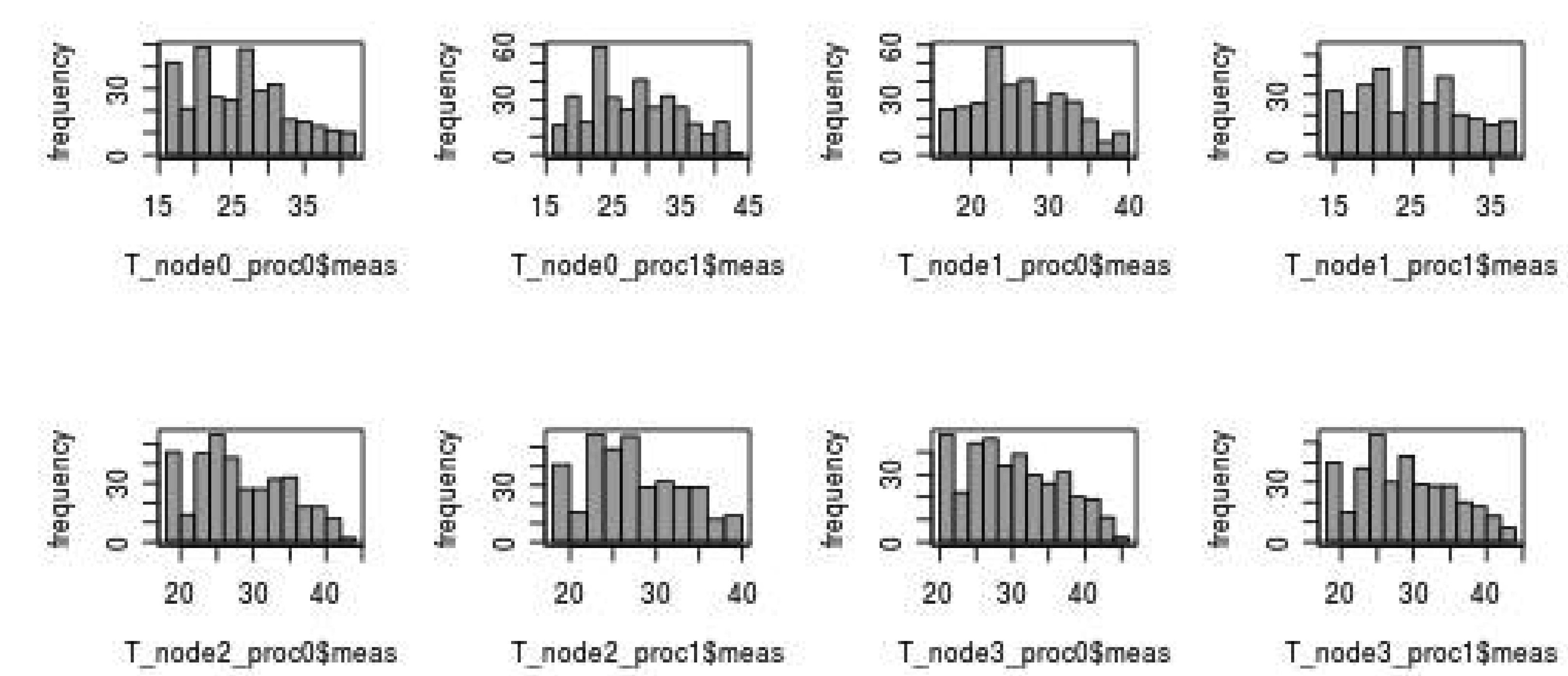
Data Schema



Distribution of voltages at core for one cabinet



Distribution of temperatures at core for one cabinet



This work has been authored in part by the US Department of Energy, Office of Science of the Oak Ridge National Laboratory, managed for the U. S. DOE by UT-Battelle, LLC, under contract No. DE-AC05-00OR22725. It was supported by the United States Department of Defense and used resources of the Extreme Scale Systems Center at Oak Ridge National Laboratory.