A Date with Data

How Time Series Information from Sensors & Logs is Revolutionizing HPC Data Center Operations at NERSC

Melissa Romanus[1,2], Elizabeth Bautista [1], Thomas Davis [1], Cary Whitney [1], Norm Bourassa [1]

[1] National Energy Research Scientific Computing Center at Lawrence Berkeley National Laboratory, Berkeley, CA 94720

[2] Rutgers University, Piscataway, NJ 08854

Introduction

- The NERSC Data Center is home to two world-class HPC machines, Edison (#124 Top500 November 2018) and Cori (#12 Top500 November 2018). It will soon house a new supercomputer, Perlmutter, with over 3 times the performance of Cori.
- It is also an extremely energy-efficient data center, requiring no mechanical chillers. Instead, it leverages the cool, dry air of the Bay Area together with evaporative cooling to achieve a Power Usage Efficiency rating of 1.07.
- Operating at this level of efficiency is no small feat. It requires characterizing workloads on HPC resources, optimizing network traffic, controlling air flow through the center, monitoring water cooling systems, mitigating potential bottlenecks, and much more.
- Underlying this extreme efficiency is a sensor and data collection framework – OMNI – that ingests over 25,000 data points per second from sensors and system logs throughout the data center. The result is one of the richest high-performance computing center operational data sets to date.
    - Over 500 billion records
    - 25K new data points per second
    - 125TB of time series data online
    - Query & Analyze 2+ years of data in seconds
- Why keep so much? Ability to analyze information across the entire data center (HPC, network, environmental, etc.), better understand impact of new systems, make optimizations at job-level and facility level, have historical data available when new questions arise
- Backed by open-source technologies specialized for big data ingestion, querying, and indexing (Logstash & Elasticsearch). Exposed via graphical web-based frontend tools Kibana and Grafana for easy data discovery and visualization. Data is exchanged via JSON. Libraries and extensions make it extensible to more full-featured and advanced tools (python, PHP, Perl, Ruby, Java, Javascript).

Data Collection Architecture

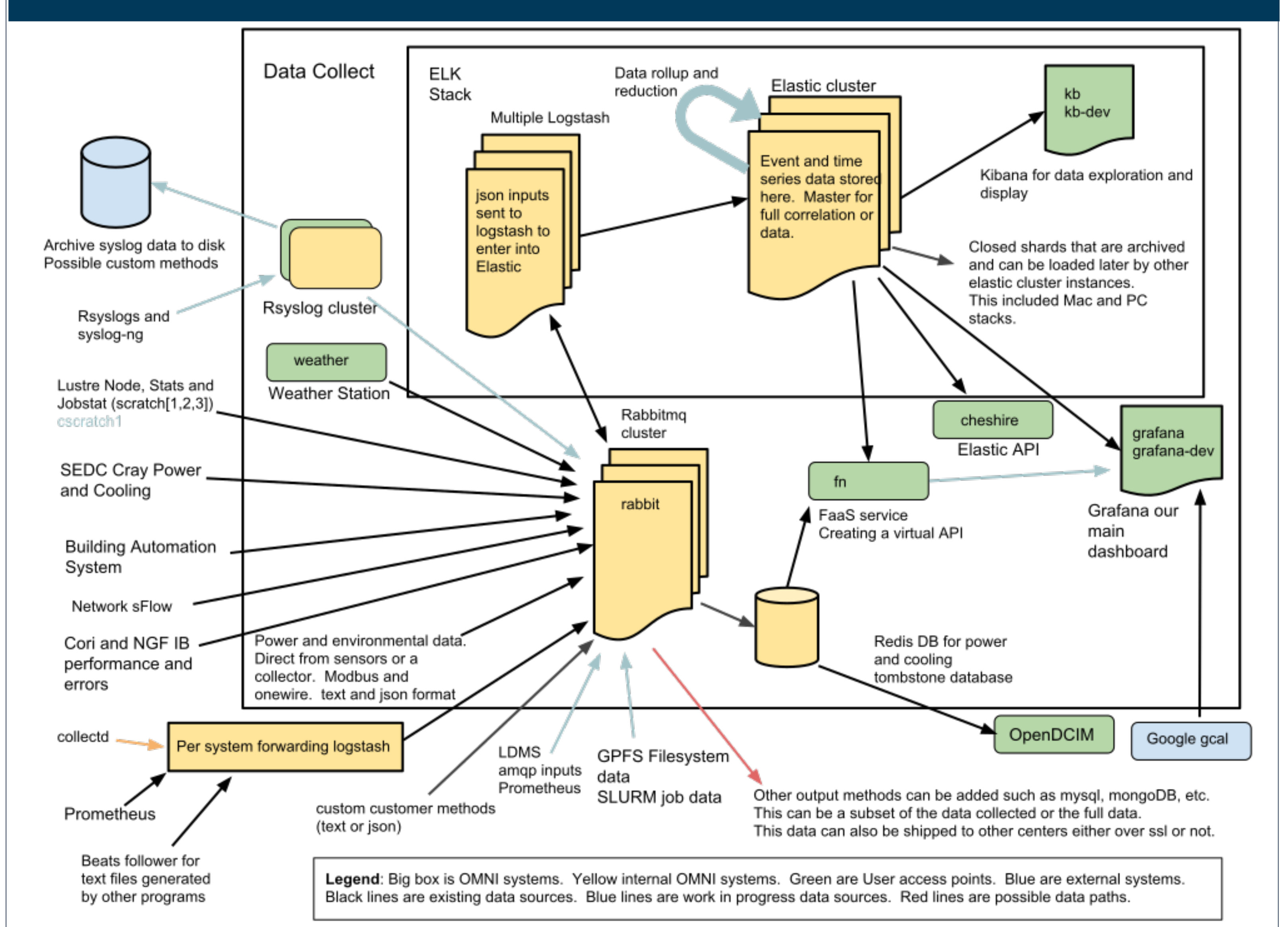


Figure 1, Data Collection Architecture. Sources include weather station, water towers, substation power, ESnet, building management systems, internal power at breakers/PDUs/UPS, Cray system environmental & power data, job characteristics, Lustre filesystem, burst buffer, InfiniBand high-speed network, particle counters, earthquake sensors, and more.

Case 1: Power Analysis of Cori

395 days
30 days

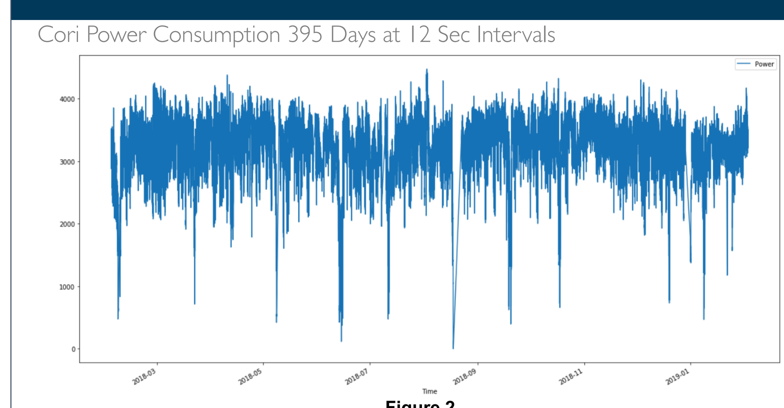


Figure 2

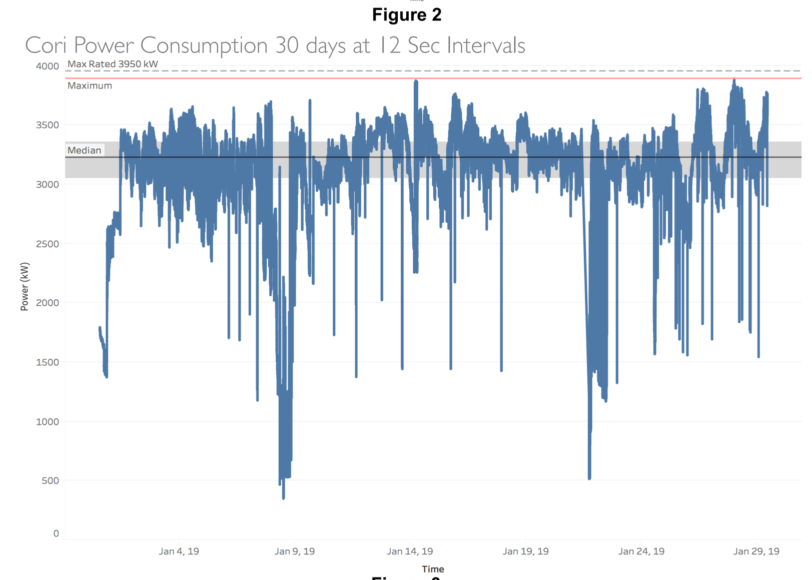


Figure 3

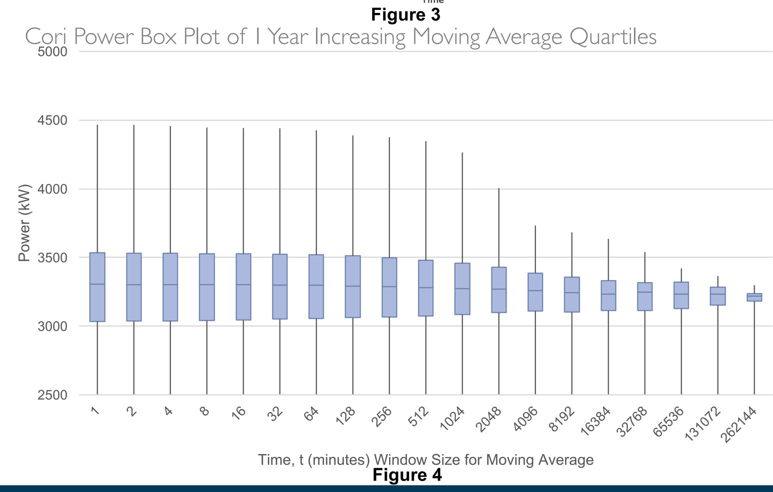


Figure 4

- Better characterization of the power on HPC machines will be the foundation for future machine procurements and increased efficiency via power-aware scheduling
- Power readings from substations are collected from Cori at 1-3 second intervals. At 1 second granularity, over a total duration of 1 year, that's over 31.5 million records! Prior to OMNI, data was typically downsampled to 15m, 30m, or 1hr segments for analysis. With OMNI, we are able to perform higher-granularity analysis of job-level fluctuations in power, such as when a node is draining or a large reservation is starting on the system.
- The data shows that Slurm's scheduling efficiency when allocating system resources to jobs allows Cori to maintain a consistent power load (high-utilization) at all times
- We hope to identify new opportunities toward Machine Learning and Power-Aware Scheduling.

Figures (top to bottom)

Figure 2, Cori Power Consumption 395 Days at 12 sec Intervals. To preserve the 12 sec granularity between measurements, we utilized python with the Elasticsearch and pandas library for ease of access. **Figure 3, Cori Power, 30 days at 12 sec Intervals.** The interquartile range band illustrates Cori's consistent load. **Figure 4, Cori Power Moving Average Quartiles.** At ~1 day (1024 minutes), the data approaches a Gaussian distribution.

Case 2: Perlmutter Mechanical Substation

442 days

- The new Perlmutter system requires a number of facility upgrades to Shyh Wang Hall (home of NERSC), including power, water, and evaporative cooling units.
- When planning for this system, the LBNL Project Planning code dictates that the need for a new substation to power the mechanical upgrades (evaporative cooling air handling units, chilled water, etc.) must be computed by either (1) summing the peak power usage of each device (as specified by the manufacturer) that will be powered by the substation **or** (2) using at least one year's worth of operational data.
- The calculation from (1) showed that another mechanical substation would be needed at a cost of $2.5 million. However, using the OMNI data collect, we were able to analyze 2 years worth of online operational data of the facility with Cori & Edison. The analysis showed that the actual demand on the mechanical substation was much lower than the peak usage rating of each device and proved that our current mechanical substation could handle the new load.

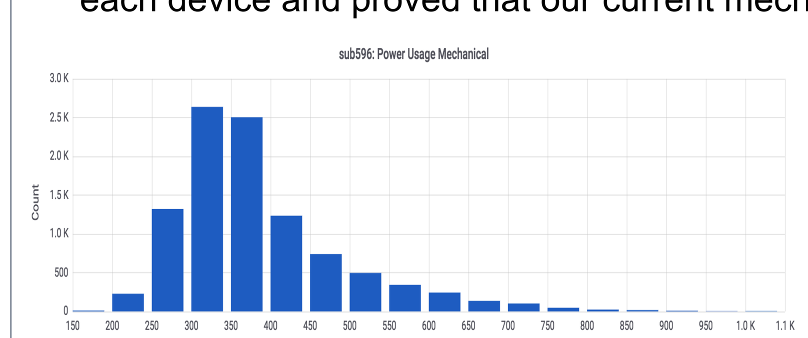


Figure 5, Substation 596 Power Consumption in 1-hr Increments. Substation 596 is the dedicated substation for mechanical units like air handlers, water pumps, etc.

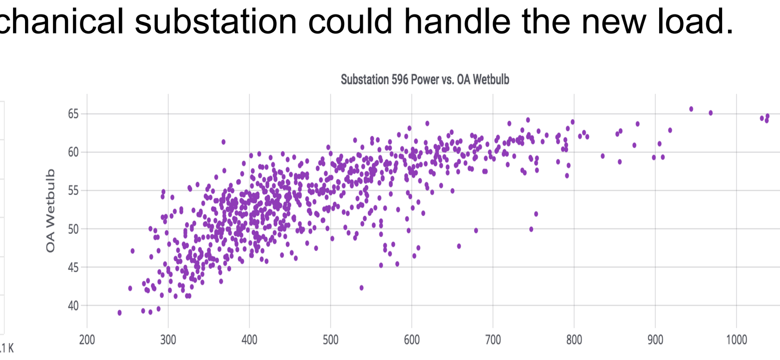


Figure 6, Substation 596 Power vs. Outside Air Wetbulb. In warmer temperatures with more humidity, the evaporative cooling functionality activates and draws more power.

Case 3: Tower Water Pump

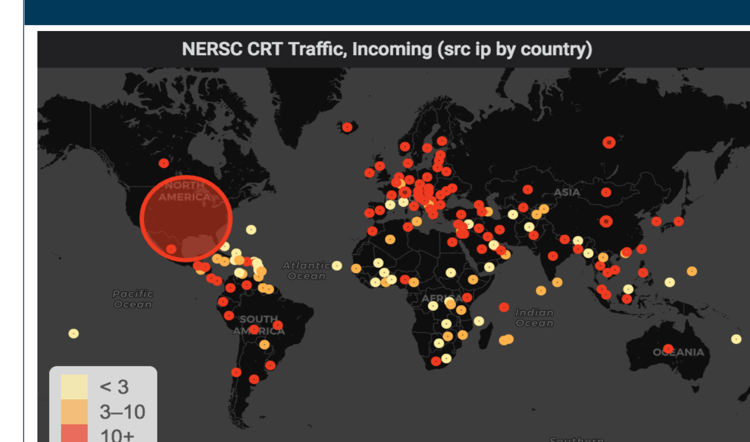
288 days

- Data showed Tower Water Pumps 1 & 2 were "maxing out," which can lead to reduced lifetime. The plots below illustrate the benefits of adding Pump 3 into the system.
- Before Pump 3 is added (left), the data shows that Pumps 1 & 2 are consistently hitting their theoretical max. After Pump 3 is added, the relative demand on Pumps 1 & 2 is lessened (right).

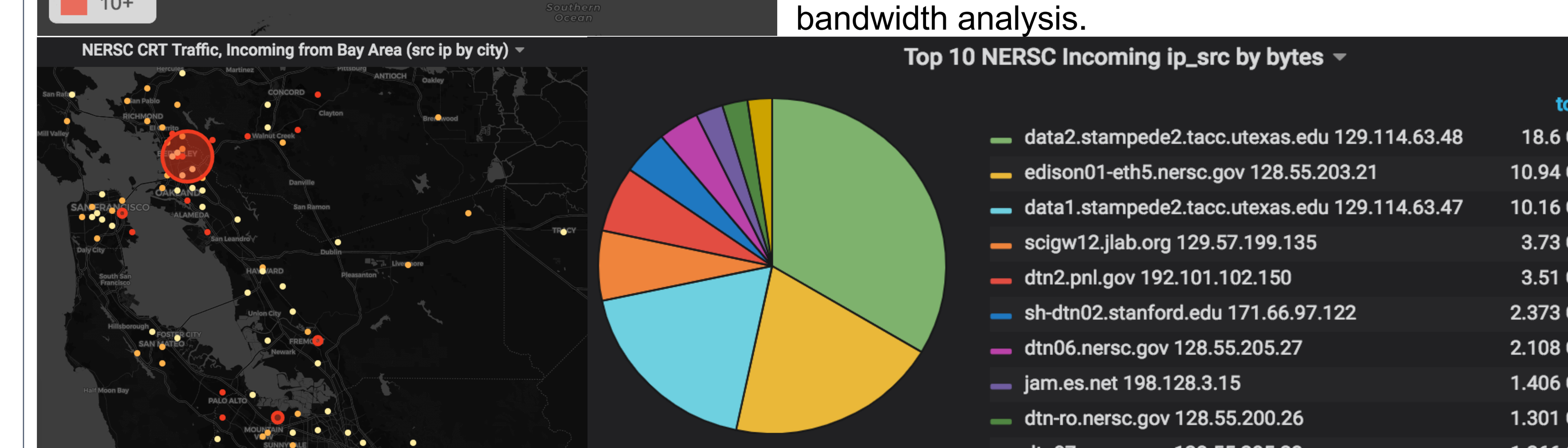
Figure 7, Benefits of HX Retrofit on Tower Water Pump Demand. The baseline draw for Pumps 1 & 2 is ~20kW while it is 0kW for Pump 3. Pumps 1 & 2 cannot be turned off because they are open enclosure systems. If turned off, outside elements (moisture, smoke particulates, etc.) have the chance to accumulate and can ultimately corrode the equipment. In contrast, Pump 3 is a closed enclosure system, thus it can selectively turn on only when it is needed.

Case 4: Network Traffic

6 hours



NERSC's Internet is serviced by the DOE's Energy Sciences network (ESnet). Once traffic arrives from ESnet to the NERSC border routers, it is routed to our internal systems, including Cori and Edison. In addition to SNMP data, we collect network samples (sFlow) at a rate of one in 512 packets. Using the sFlow data, we are able to create a snapshot of the source IP addresses connecting to NERSC, as well as examine the top source IPs for data transfer, which is useful in cybersecurity and network bandwidth analysis.



Future Work

- "Smart HPC Data Center": Connecting different data sources, applying machine learning techniques to spot trends or make predictions about failures, workload placement, and more.
- Power-aware scheduling. Work to better characterize node-level power usage on Cori and look for trends based on Darshan log workload characterization
- Publically available datasets coming end of Summer 2019. Check https://www.nersc.gov/news-publications/staff-blogs/otg/ for updates!

This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231.