Global Experiences with HPC Operational Data Measurement, Collection and Analysis

Michael Ott
michael.ott@lrz.de
Leibniz Supercomputing Center

Norman Bourassa
njbourassa@lbl.gov
Lawrence Berkeley National Laboratory

Melissa Romanus AbdelBaky
mromanus@lbl.gov
Lawrence Berkeley National Laboratory

Woong Shin
shinw@ornl.gov
Oak Ridge National Laboratory

Natalie Bates
Natalie.jean.bates@gmail.com
Energy Efficient HPC Working Group

Torsten Wilde
wilde@hpe.com
Hewlett Packard Enterprise
Energy Efficient HPC Working Group

Mission

- First early adaptor sites are starting to deploy Operational Data Analysis (ODA) systems for collecting and analysing HPC platform and facility data together
- As a team, we want to learn and understand:
  - What tools are the sites using for this?
  - How are they using the collected data?
  - What are the lessons learned?
- Provide guidance for the community to deploy similar systems

Methodology

- **Initial survey to identify interest**
  - Precursory questions sent to 24 sites
  - Participation in in-depth questionnaire and interview
    - Yes (7 sites), No (2 sites), Unclear/No response (5 sites)
- **In-depth survey to understand more**
  - Participation from 8 sites (RIKEN, LRZ, CINECA, ORNL, NCAR, NREL, LBNL, LLNL)
  - Detailed questionnaire, teleconference interviews with the ODA team & site individuals
- **Synthesis**
  - How sites are really using their data collection systems with specific use case scenarios.
  - To what extent are data streams from HPC platform and infrastructure integrated into a common platform.
  - Implementation issues; such as system-to-system interface.
  - Known or potential scalability constraints.
- **Report**
  - Initial synthesis results presented at the 11th Workshop for the Energy Efficient HPC Working Group @SC19
    https://eehpcwg.llnl.gov/assets/sc19_11_425_525_operational_data_analytics_ott_bourassa.pdf
    - Use-cases & impact
    - Architectural pattern reported from sites.
    - Common tools used throughout the sites
  - ODA team pursuing SC20 paper for in-depth analysis

Survey Details

- **Initial Survey**
  - Has your site implemented data collection, aggregation and analysis for operational management of both facilities and HPC systems (including energy and power management) in a production environment on at least one large-scale system (Top500 sized system) with integration (or plans for integration) that extends from the HPC data center down through the platform to the CPU?
  - Willingness to extend to in-depth survey
- **In-depth Survey**
  - Scope & degree of data aggregation implemented for the site
  - Value added actions utilizing the aggregation system
  - Architecture, tools & systems in use
  - Data Management Policy
  - Lessons learned & future plans

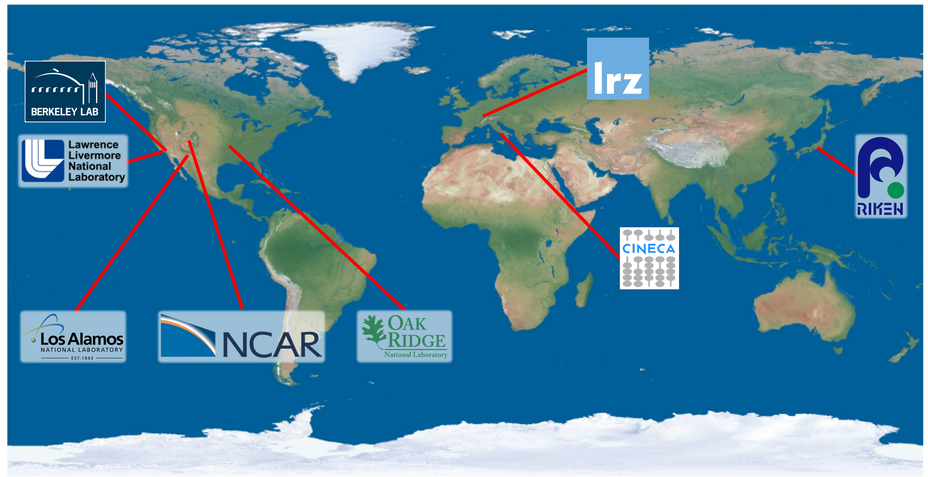

Figure 1 - Methodology: Initial survey to 24 sites to identify interest followed by an in-depth survey with the participation of eight sites.

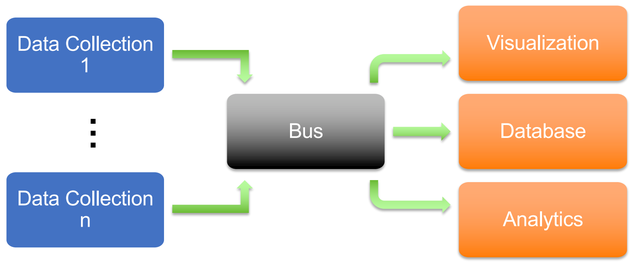

Figure 2 - Schematic representation of common architecture components: As we analyzed the data of the survey and conducted interviews with individual sites it became obvious that many sites are in fact building similar architectures for their operational data analytics efforts.

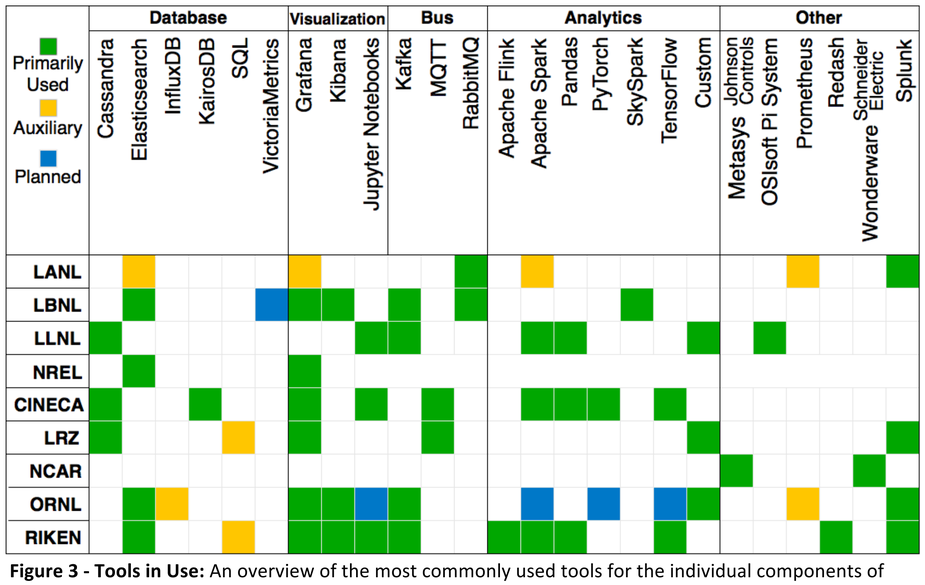

Figure 3 - Tools in Use: An overview of the most commonly used tools for the individual components of such ODA architectures at the different sites that participated in the survey

Use-cases

- **NCAR: Infrastructure Optimization**
  - Correlation of CDU monitoring data with power quality meter showed brown out condition in the same timeframe
  - Move CDU pumps to UPS-backed power
- **LRZ: Application Tuning**
  - Memory bound application can run at lower frequency without hurting performance but lowering their energy-to-solution
  - Use performance counters to assess application characteristics and DVFS to underclock memory bound applications
- **ORNL: Cooling optimization**
  - Correlate water temperatures & flow rates with CPU/GPU temperatures
  - Optimize temperature and flow for energy efficiency, maintain safe & un-throttled operational conditions of HPC system
- **RIKEN: Power Management**
  - Monitor datacenter power draw, compare with utility rate limit and HPC system application job schedule
  - Kill high-power jobs to reduce total power draw
- **LBNL: Cooling Optimization**
  - Use actual HPC air demand to control AHU instead of rule of thumb
  - Aggregate individual server air demand (air temperatures, fan speed) using Prometheus client and feed this into AHU controls

Conclusion

- Sites are collecting and integrating data from HPC platform and infrastructure for analysis and visualization purposes.
- Scalable open source tools are becoming more available for implementing such pipelines.
- Analysis currently is mostly manual by visual inspection, use of AI/ML is still an active research activity.
- ODA can pay for itself by lowering OPEX.

Further Questions

- If you had to do it over again, what would you do differently?
- What did you want to achieve when you first began?
- What have been some of your biggest challenges or obstacles?
- Do you have any lessons learned?
- What questions do you have for your peers?

The Energy Efficient HPC Working Group (EEHPC-WG) invites other supercomputing sites to participate in enhancing this survey. It welcomes questions, feedback, and comments from the entire HPC community.

This QR code alongside, points to the EEHPC-WG webpage (https://eehpcwg.llnl.gov/) that contains additional links to the white paper related to this poster, the feedback form, and other information on ways you can participate in this era of Energy and Power-aware computing.