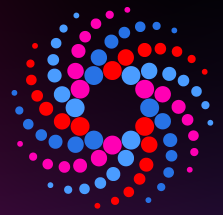
SAND2022-15595C



SC22
Dallas, TX | hpc accelerates.

AppSysFusion: CoMingling of appropriate data to drive Codesign of Applications, HPC Platforms, and Monitoring, Analysis, and Feedback Infrastructure

International Symposium on Quantitative Codesign of Supercomputers

Jim Brandt, Ann Gentile (Sandia National Laboratories)

11-13-2022

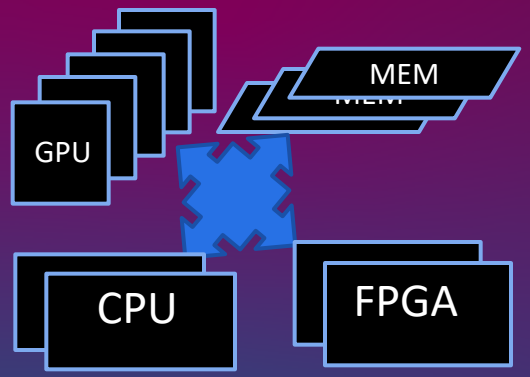
Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

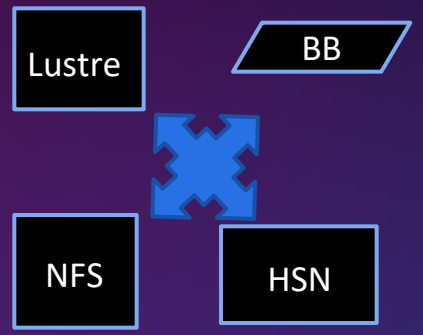
Motivation For Dynamic High Performance Computing Resource Management

Problem: Oversubscription of HPC systems results in **costly mission delays**: queue times can be weeks!

- Incorrect resource requests are **inefficient**, **waste resources,** and **exacerbate the problem**
  - **Users try to overestimate needs to minimize chances of early termination (OOM, timeout)**
- High run time variation due to **contention** and/or **component degradation ➔ inefficiency**
- **Wasted resources = wasted $$**
- ***Future:*** Increases in aggregate compute node capabilities may improve overall performance **BUT reduce overall efficiency**



Multiplicity of **compute node** resources ➔ suboptimal use

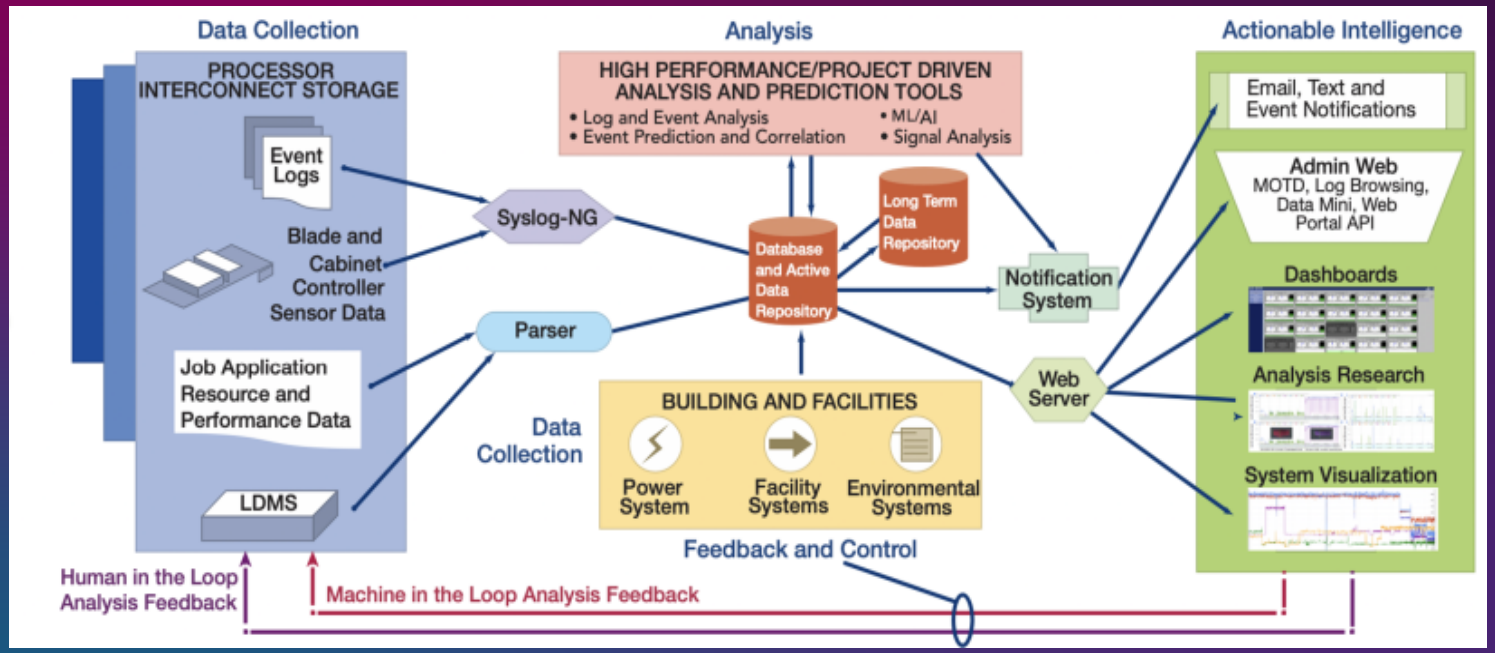


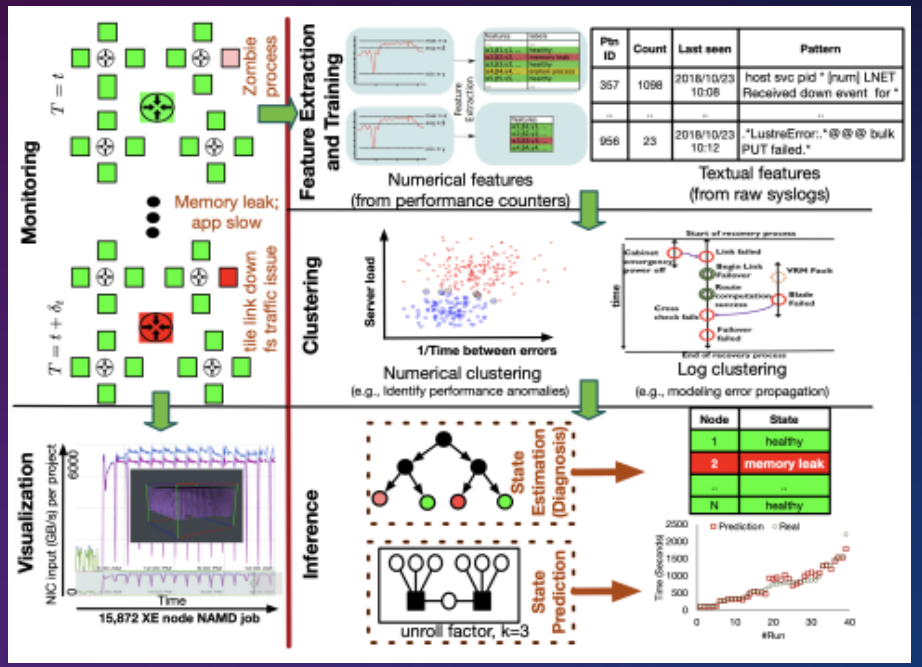
Multiplicity of shared **system** resources ➔ prone to oversubscription

DOE Heterogeneity report: Heterogeneity of current and future architectures does not lend itself to efficient utilization when the **characteristics of available resources** and the **low-level needs of user applications**, which may **change over execution time**, are **not known by the scheduling and resource management software**.

Moving Towards Autonomous HPC Facilities

- **Currently:** HPC facility operation and rule-based orchestration is based on human experience, inference, and reaction times.
- **Goal (short term):** AI provides data-driven **Decision Support** for **humans** to support better resource utilization and higher resource availability (equivalent to blind spot or lane warning in auto)
- **Goal (long term):** AI used to **Orchestrate and Optimize resources** and **operations** subject to physical constraints such as power and priority with humans providing support in terms of repair and iteration on next generation capabilities (equivalent to fully autonomous in auto)





Key Points

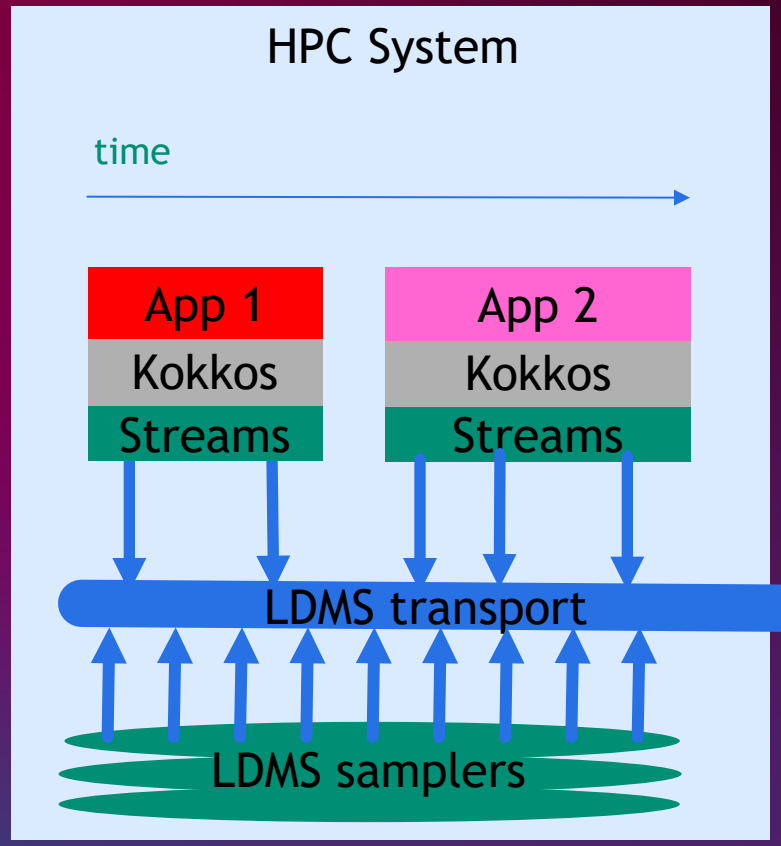
- Need to move from considering individual application performance to considering the **efficiency of the multi-dimensional HPC ecosystem** (i.e., optimize utilization of resources and performance while honoring constraints such as power and priority)
- The dynamic and complex nature of HPC workloads requires **continuous orchestration** of overall HPC ecosystem which **relies on continuous insight** into all dimensions

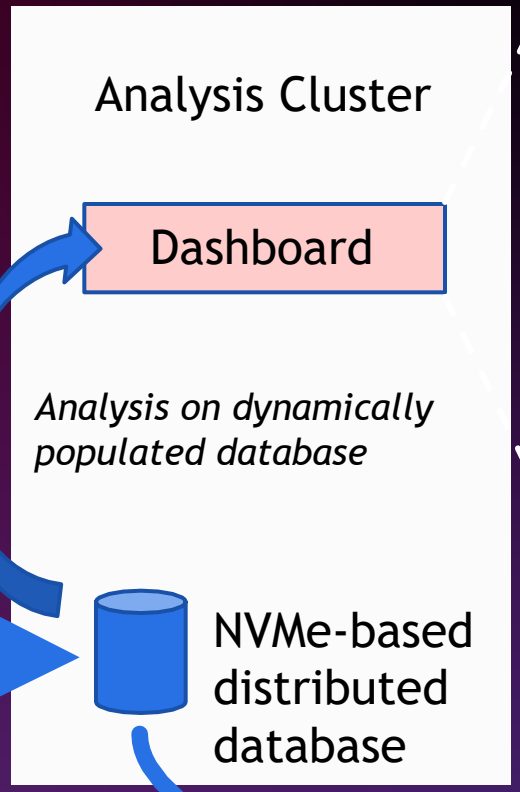
Approach: Re-imagine **HPC resources as autonomous peer components** that can negotiate among themselves, and with applications, to **optimize global efficiency**

AppSysFusion **First Steps**: Characterization of Application and System State/Performance

Applications dynamically and irregularly inject data into the LDMS transport

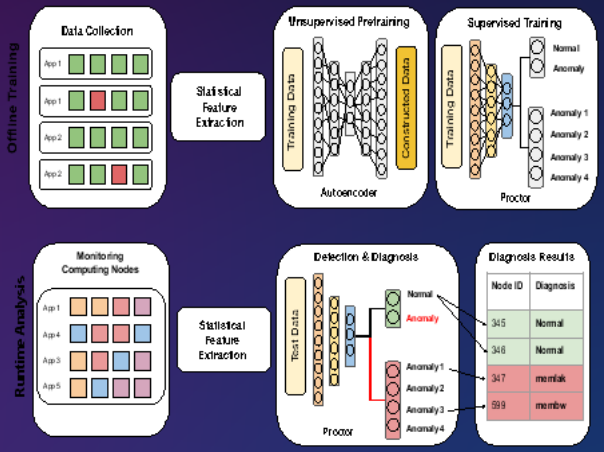
LDMS continuously and regularly collects and transports full system data





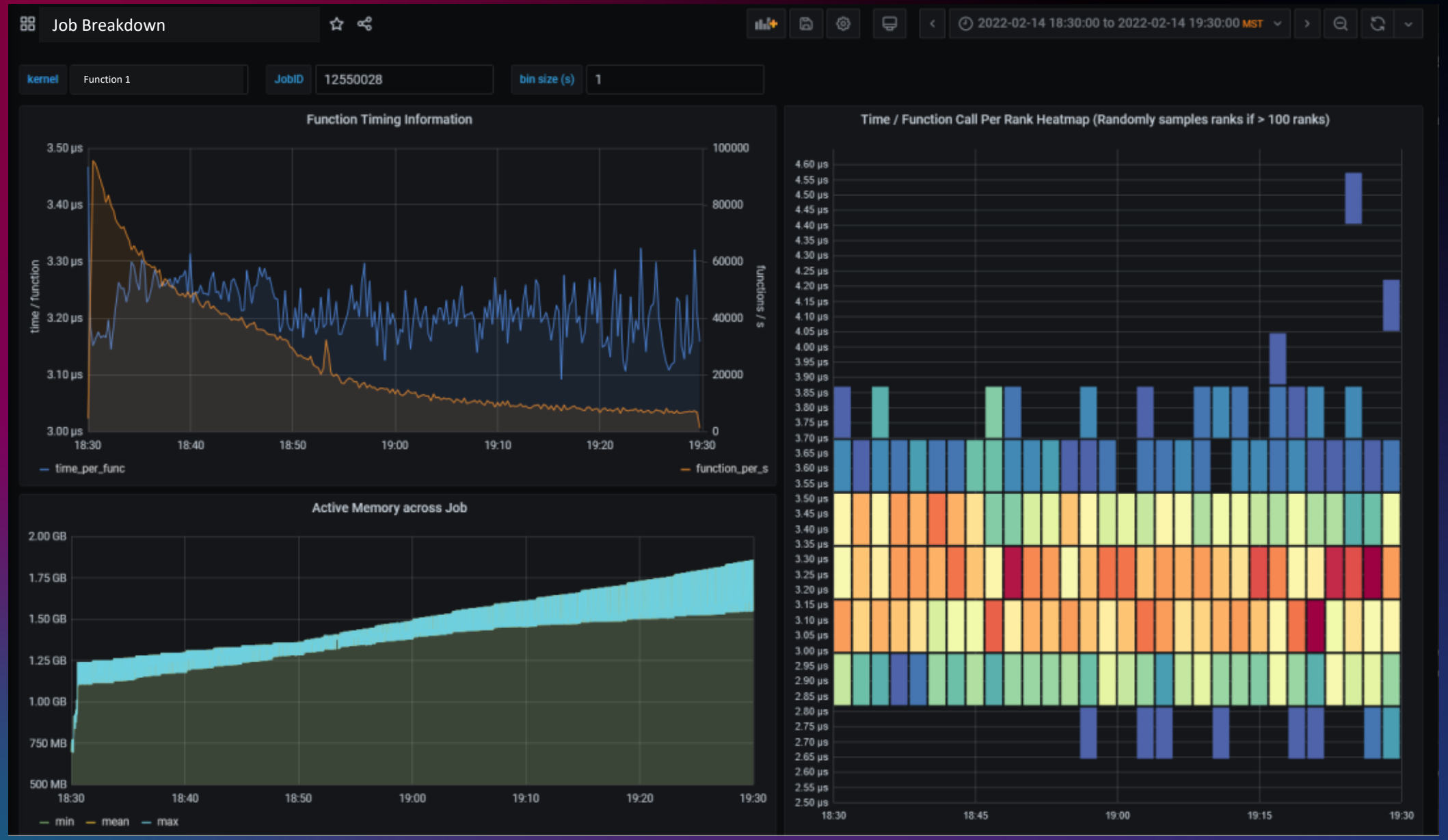
Application Progress

System Metric

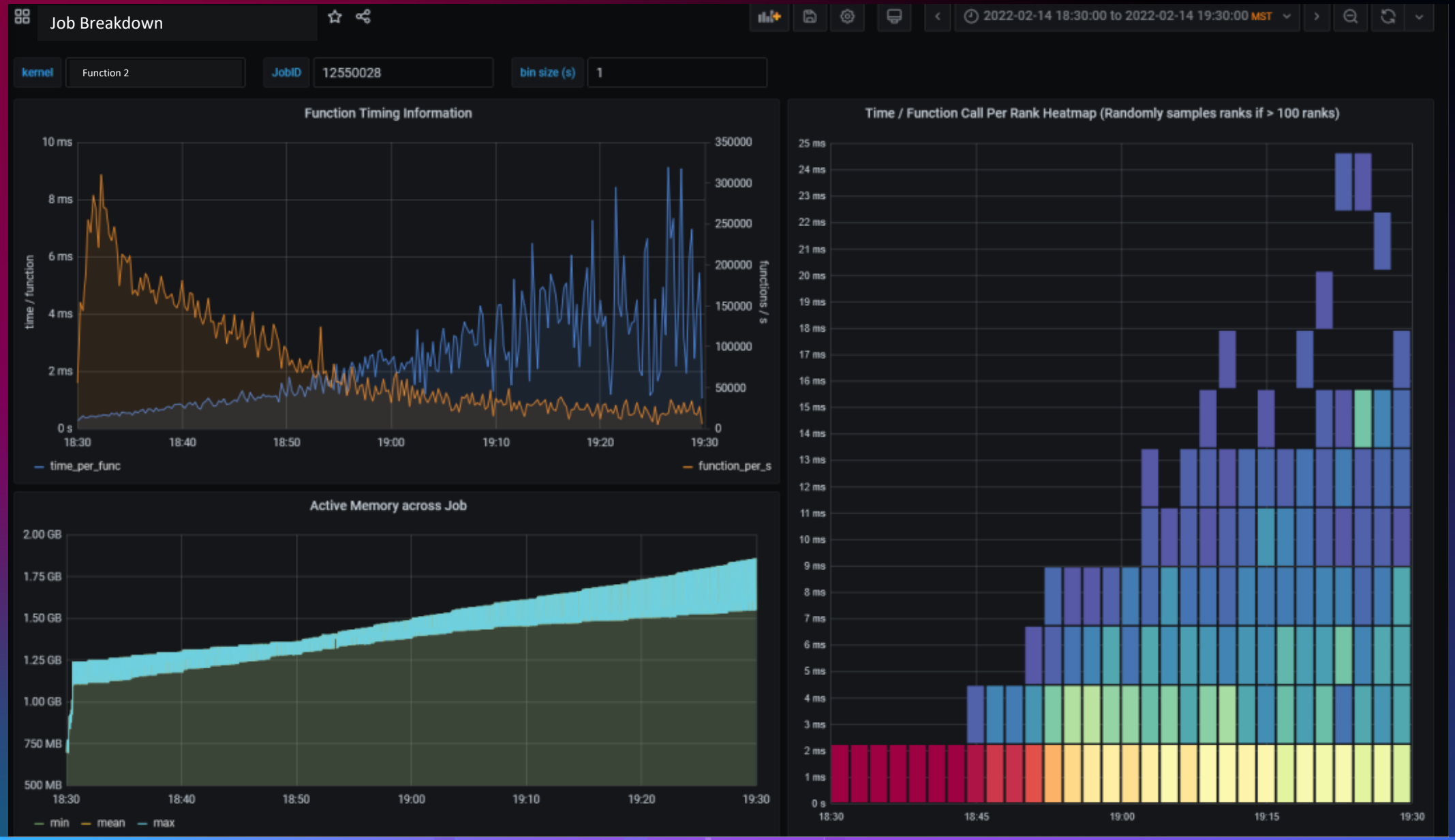


ML-based Decision Support

Particle Migration Simulation



Particle Migration Simulation Continued



Automatic, Detailed Resource Forecasting

- An agent for each host will track resource availability and forecast future availability using the current model for each running application
- The host agents will then seek to maximize utilization of their resources by advertising a resource vector of their projected availability
- Host agents use detailed local resource and workload information in conjunction with hierarchical organization to achieve efficient identification of a target resource collective that maximizes the system utility function without need for global detailed information-sharing.

Conceptualization of multi-tenant, multi-resource allocation for a compute host over time

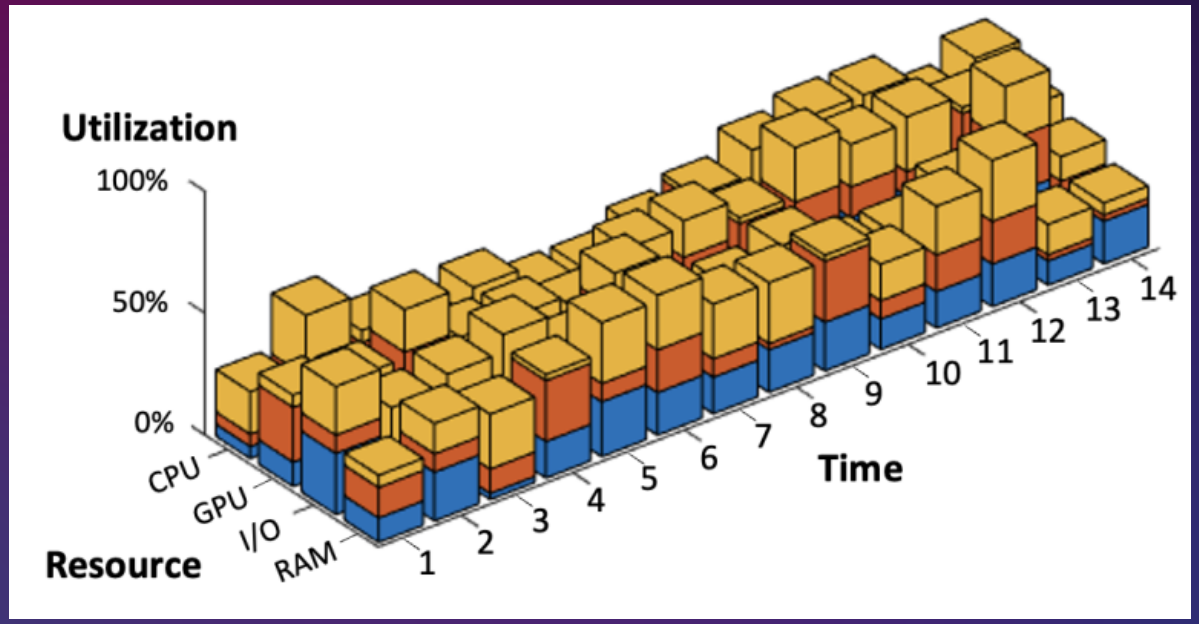


Figure by Robert Abbot (SNL)

Challenges

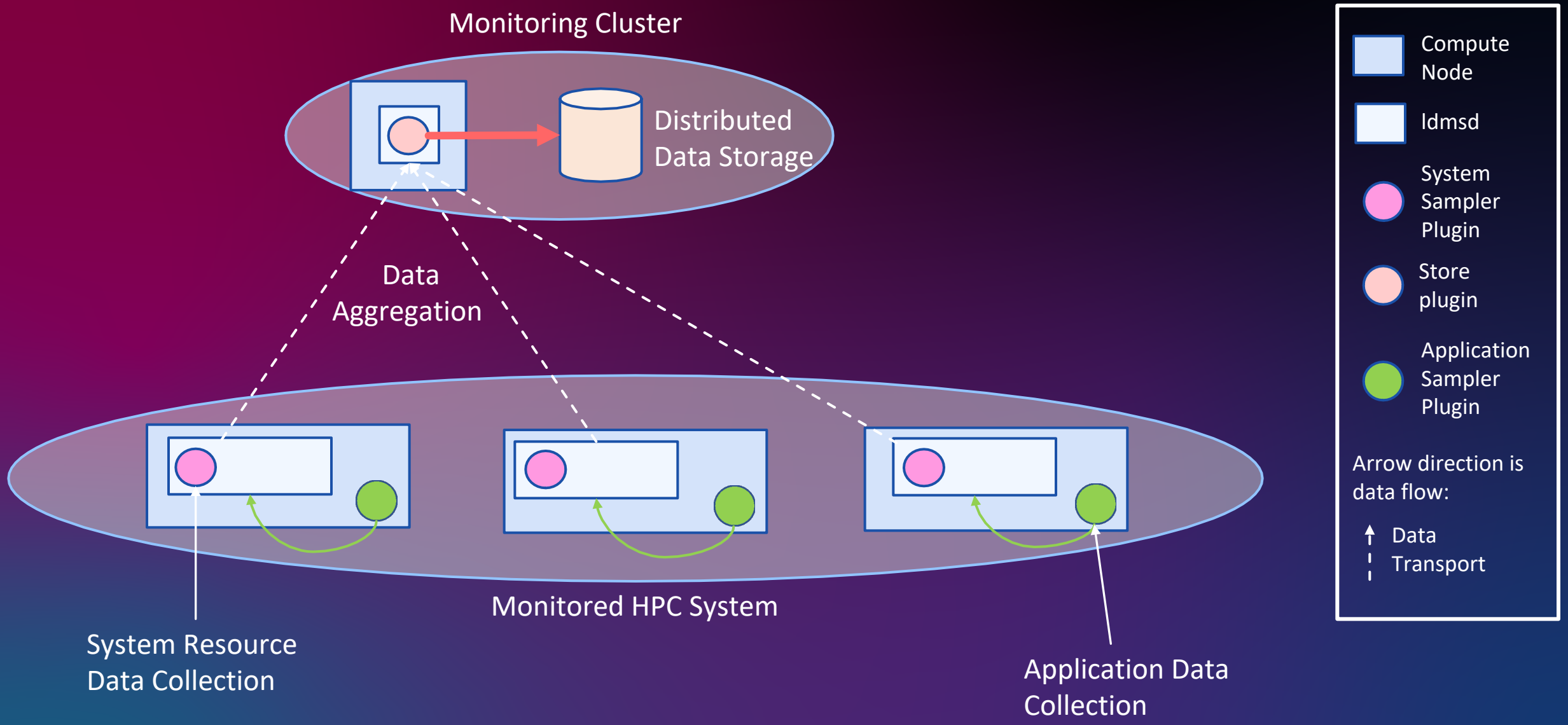
Major challenges to taking this approach are development and testing of:

1) Software and hardware to enable scalable distributed low-latency data-driven decision-making without negatively impacting resources executing user workflow performance
2) A communication protocol that is lightweight and can accommodate required communication and latency bounds
3) System software for birth-to-death management of workflows including addition of self-aware workflow components that can interact in a run time customer-provider relationship with distributed intelligent resource components.

Current Approach

- Create a container-based emulator
- Develop, deploy, and explore interaction of the above-described software components on large-scale emulated system (thousands of emulated compute nodes)
- Utilize labeled data traces from production systems as ground-truth for workflow component behavior and resource requirements/utilization

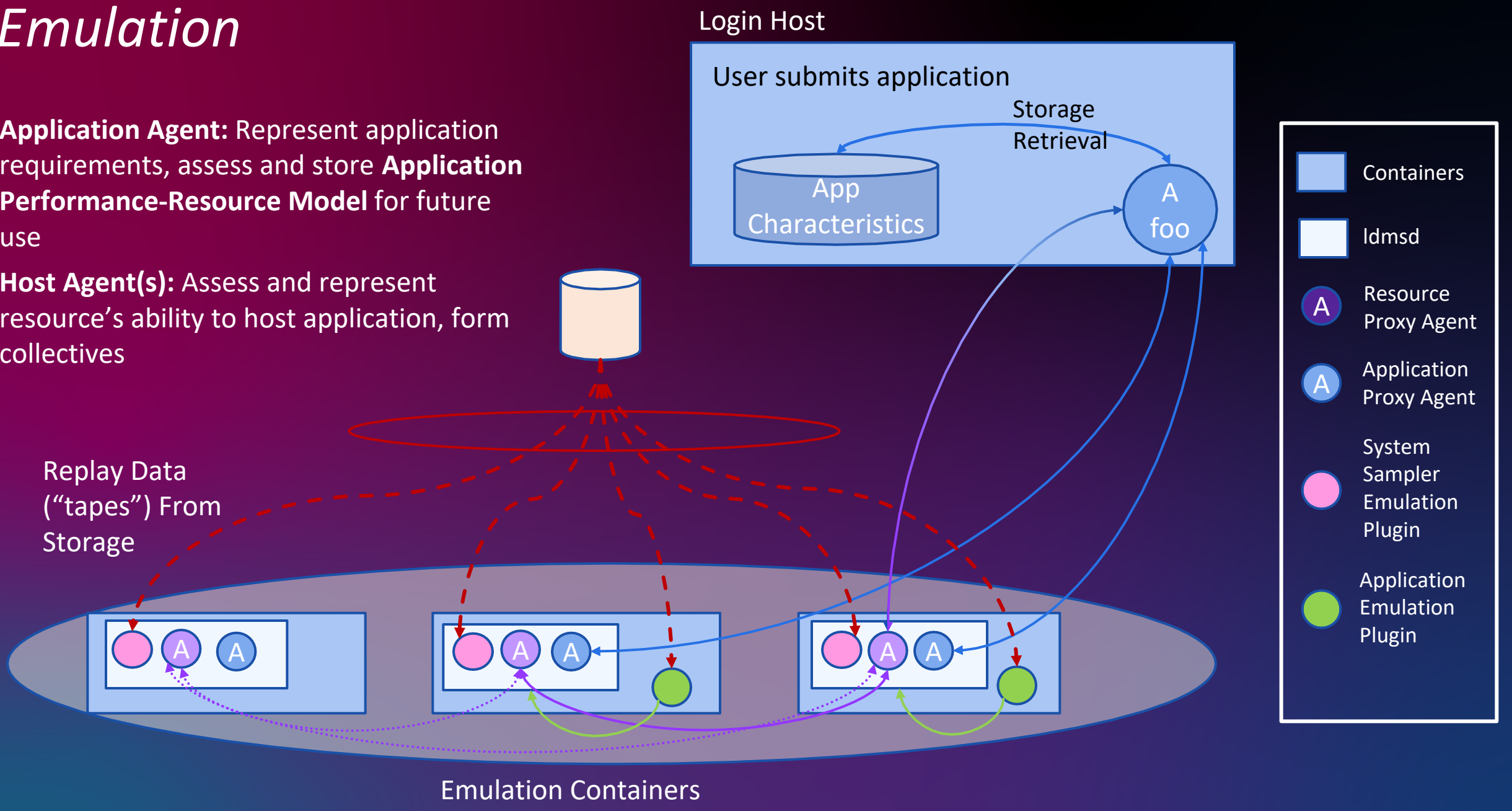
Sandia's AppSysFusion Is An Enabling Technology



Emulation

Application Agent: Represent application requirements, assess and store **Application Performance-Resource Model** for future use

Host Agent(s): Assess and represent resource's ability to host application, form collectives



Questions?