SAND2020-1656C

Sandia National Laboratories

Advanced Simulation & Computing™

with ECP Exascale Computing Project

Attributing Performance Variation from Integrated Application and System Data

O. Aaziz, B. Allan, J. Brandt, J. Cook., K. Devine, J. Elliott, A. Gentile, S. Olivier, K. Pedretti, and T. Tucker

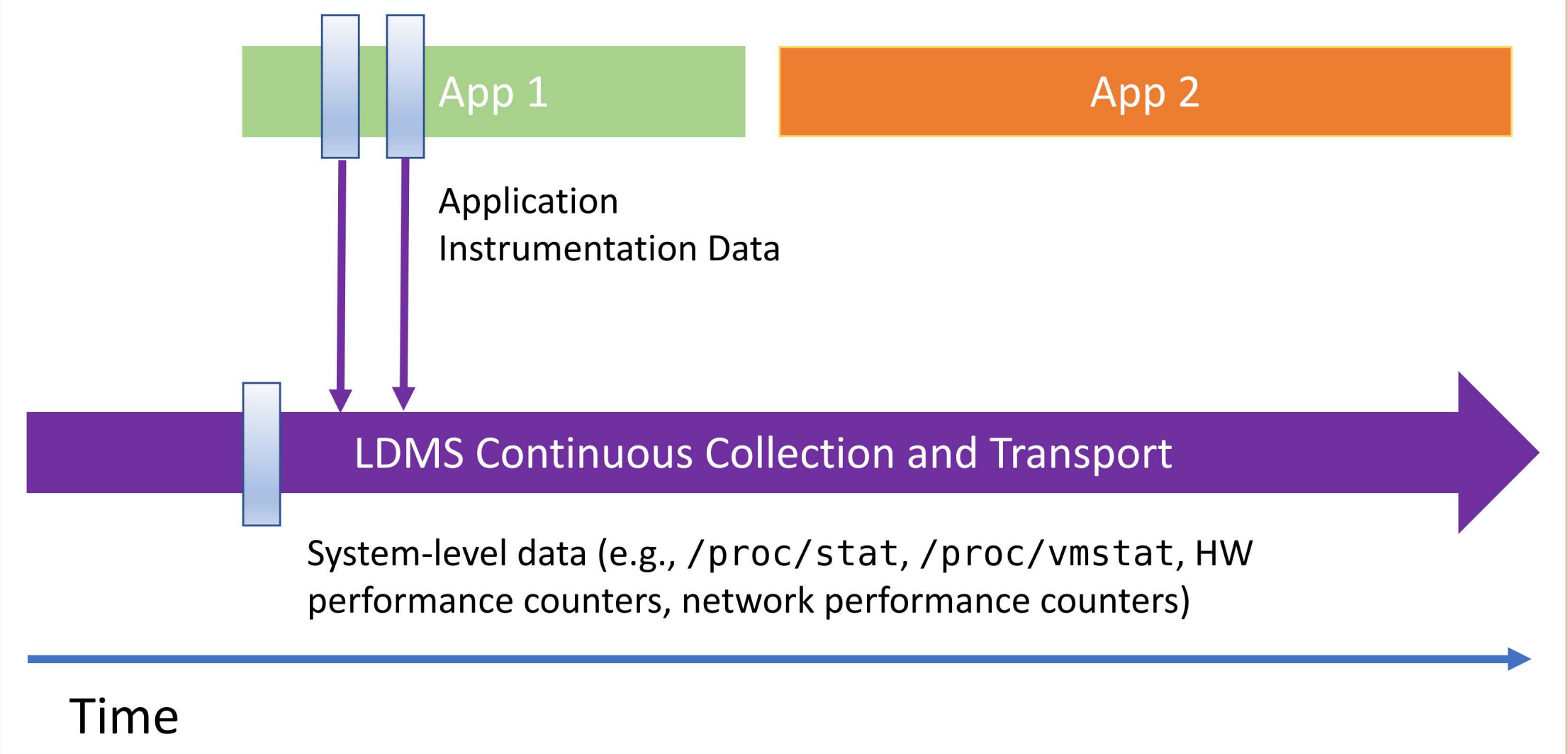
Objective

High fidelity integrated collection and analysis of application and system information to:

- Detect performance variation and diagnose root causes
- Assess effectiveness of code changes on use of architectural features (e.g., cache, memory, network) and runtime
- Detect inefficient resource usage
- Develop intelligent resource management techniques to improve system throughput

Coupling Application and System Data

- **Common representation for output from Kokkos profiling and Trilinos timers.** JSON key-value pairs as an application agnostic format
- **Minimize code modifications required to identify application phases and progress** by leveraging existing timers and built-in profiling tools
- **Export application metadata** to indicate comparable runs and capture execution environment
- **LDMS continuously collects whole system performance data**

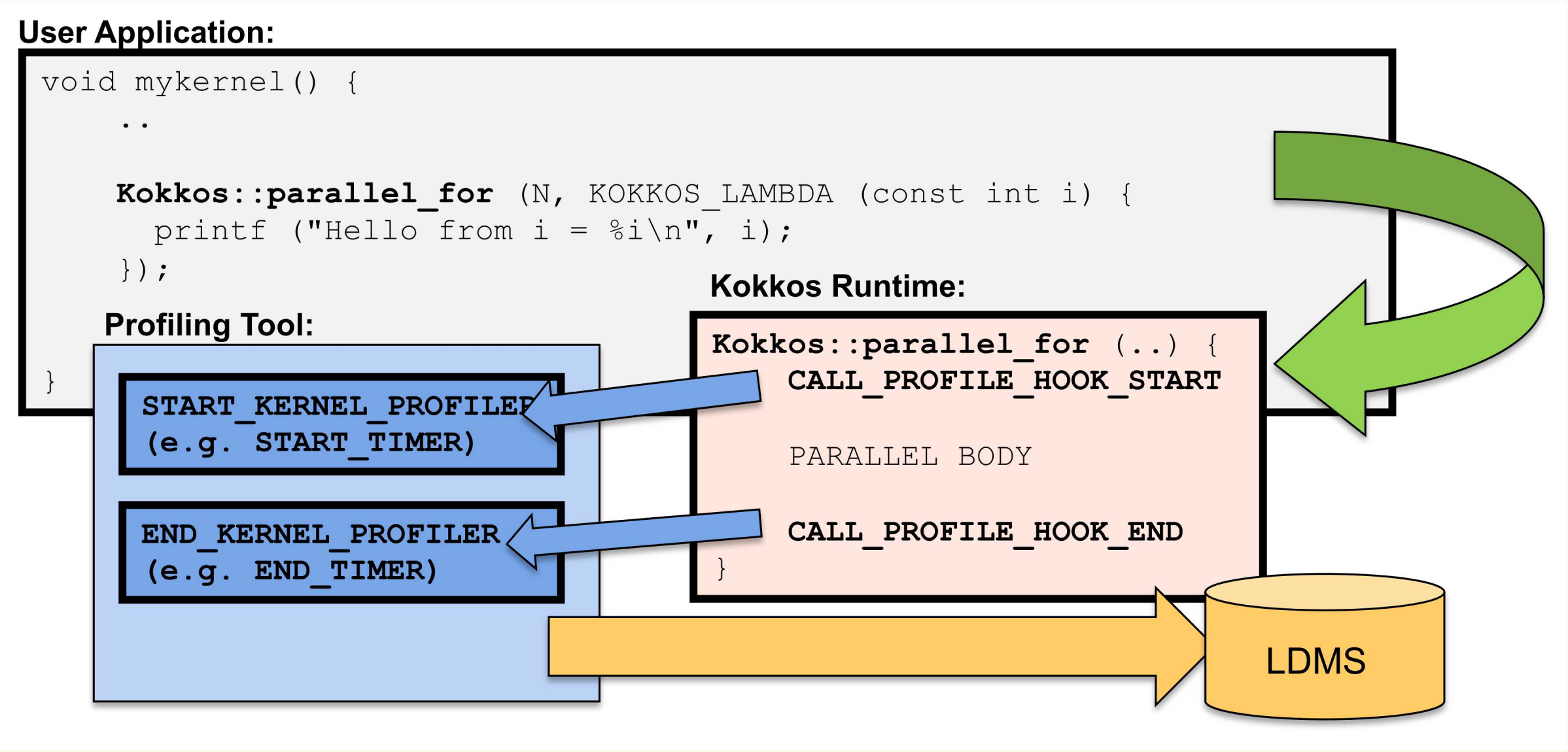

App 1
App 2
Application Instrumentation Data
LDMS Continuous Collection and Transport
System-level data (e.g., /proc/stat, /proc/vmstat, HW performance counters, network performance counters)
Time

Combine Application-level data with System-level data, jointly transported via LDMS for integrated analysis

Data Representation

- **High Performance Object Store** to ingest 1000s of individual data types and 10s TB/day
- **Store complete data history** to analyze application performance progression and compare across platform generations
- **Python interface for analysis development**

Architecture


User Application:

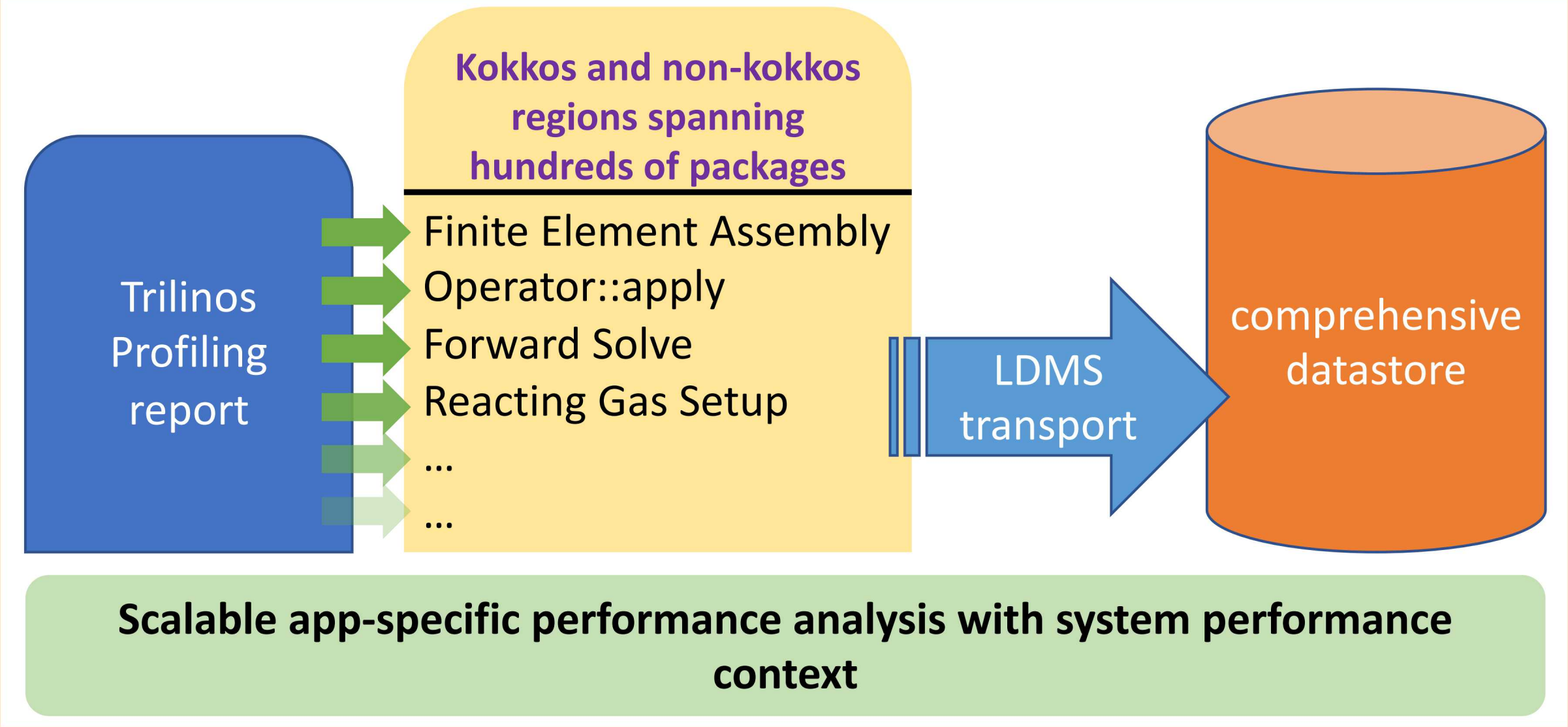
```
void mykernel() {
    ..

    Kokkos::parallel_for (N, KOKKOS_LAMBDA (const int i) {
      printf ("Hello from i = %i\n", i);
    });
}
```

Profiling Tool:

```
START_KERNEL_PROFILER
(e.g. START_TIMER)

END_KERNEL_PROFILER
(e.g. END_TIMER)
```

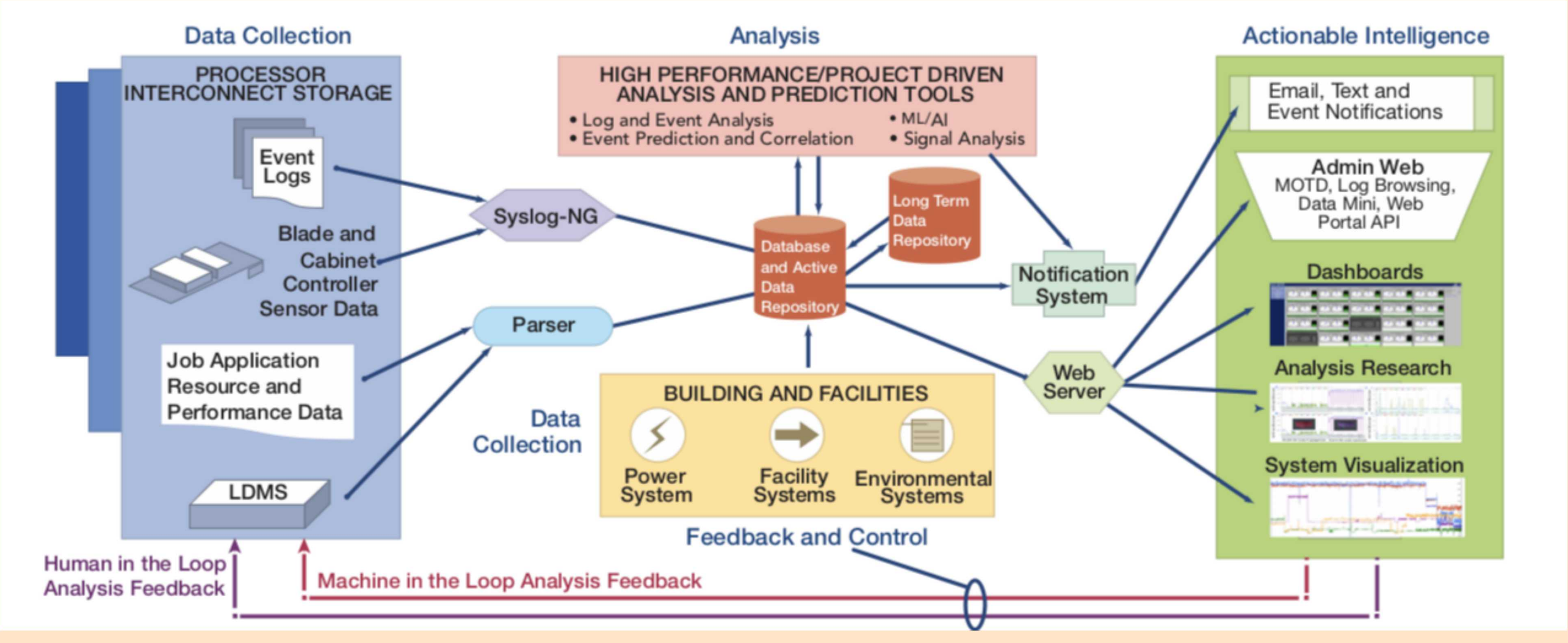
Kokkos Runtime:

```
Kokkos::parallel_for (..) {
    CALL_PROFILE_HOOK_START

    PARALLEL BODY

    CALL_PROFILE_HOOK_END
}
```


Trilinos Profiling report
Kokkos and non-kokkos regions spanning hundreds of packages
Finite Element Assembly
Operator::apply
Forward Solve
Reacting Gas Setup
…
…
LDMS transport
comprehensive datastore
Scalable app-specific performance analysis with system performance context

Existing, extensive Trilinos profiling interface provides app specific timings at a coarser granularity than Kokkos

Application data injected into LDMS transport using LDMS Streams interface


Data Collection
PROCESSOR INTERCONNECT STORAGE
Event Logs
Blade and Cabinet Controller Sensor Data
Job Application Resource and Performance Data
LDMS
Analysis
HIGH PERFORMANCE/PROJECT DRIVEN ANALYSIS AND PREDICTION TOOLS
• Log and Event Analysis
• Event Prediction and Correlation
• ML/AI
• Signal Analysis
Syslog-NG
Parser
Database and Active Data Repository
Long Term Data Repository
Notification System
Web Server
Data Collection
BUILDING AND FACILITIES
Power System
Facility Systems
Environmental Systems
Feedback and Control
Actionable Intelligence
Email, Text and Event Notifications
Admin Web
MOTD, Log Browsing, Data Mini, Web Portal API
Dashboards
Analysis Research
System Visualization
Human in the Loop Analysis Feedback
Machine in the Loop Analysis Feedback

Runtime analytics enable insights and operational decision-making while applications are running

Analysis

Machine Learning and Statistical analysis:

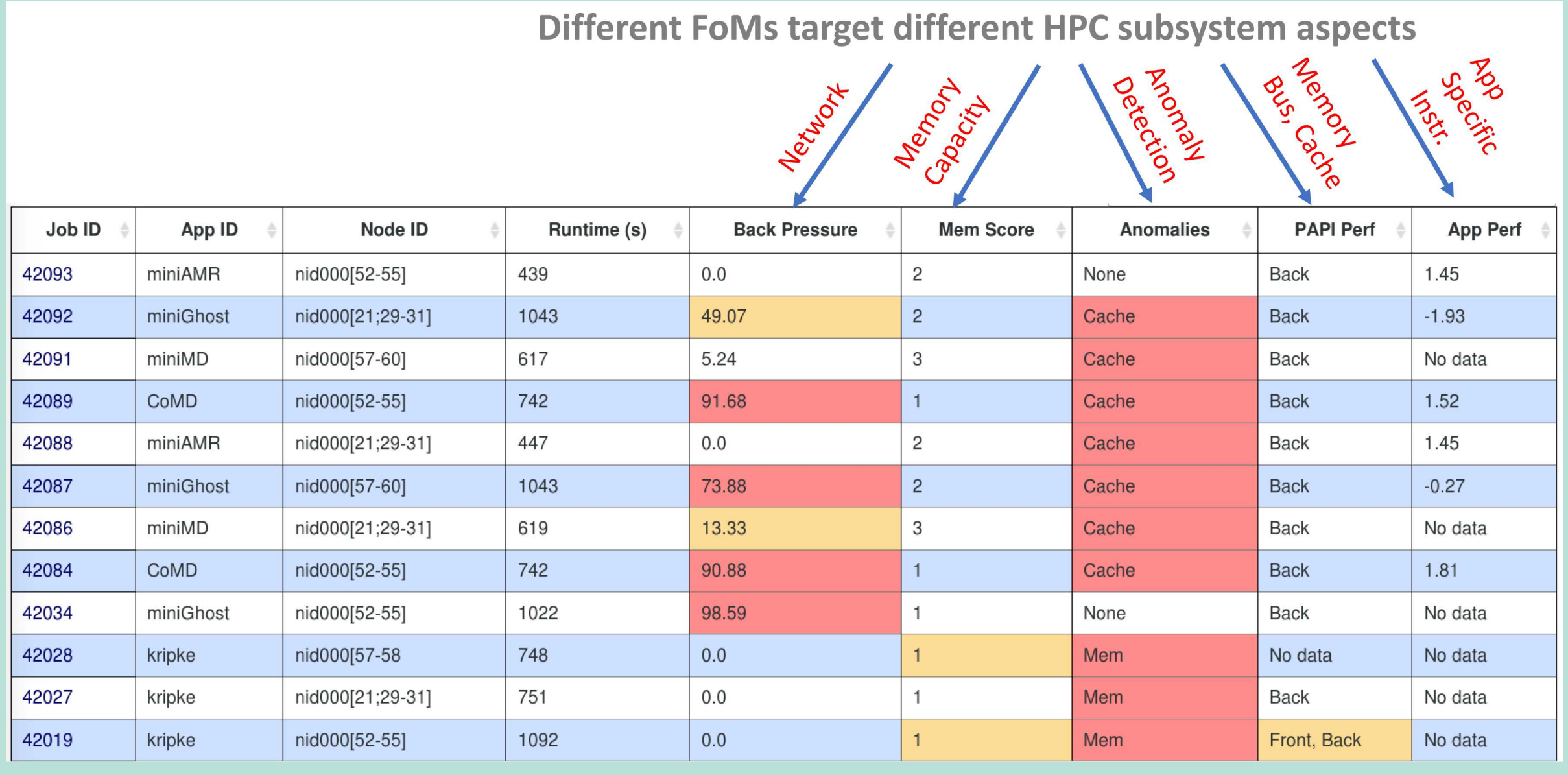
- Detect anomalous application performance
- Determine most important data features
- Quantify relationships between ensembles of data values and application performance
- Root cause attribution of performance variation
- **Incorporate Architecture and Application relevant data features**

Feedback and Response

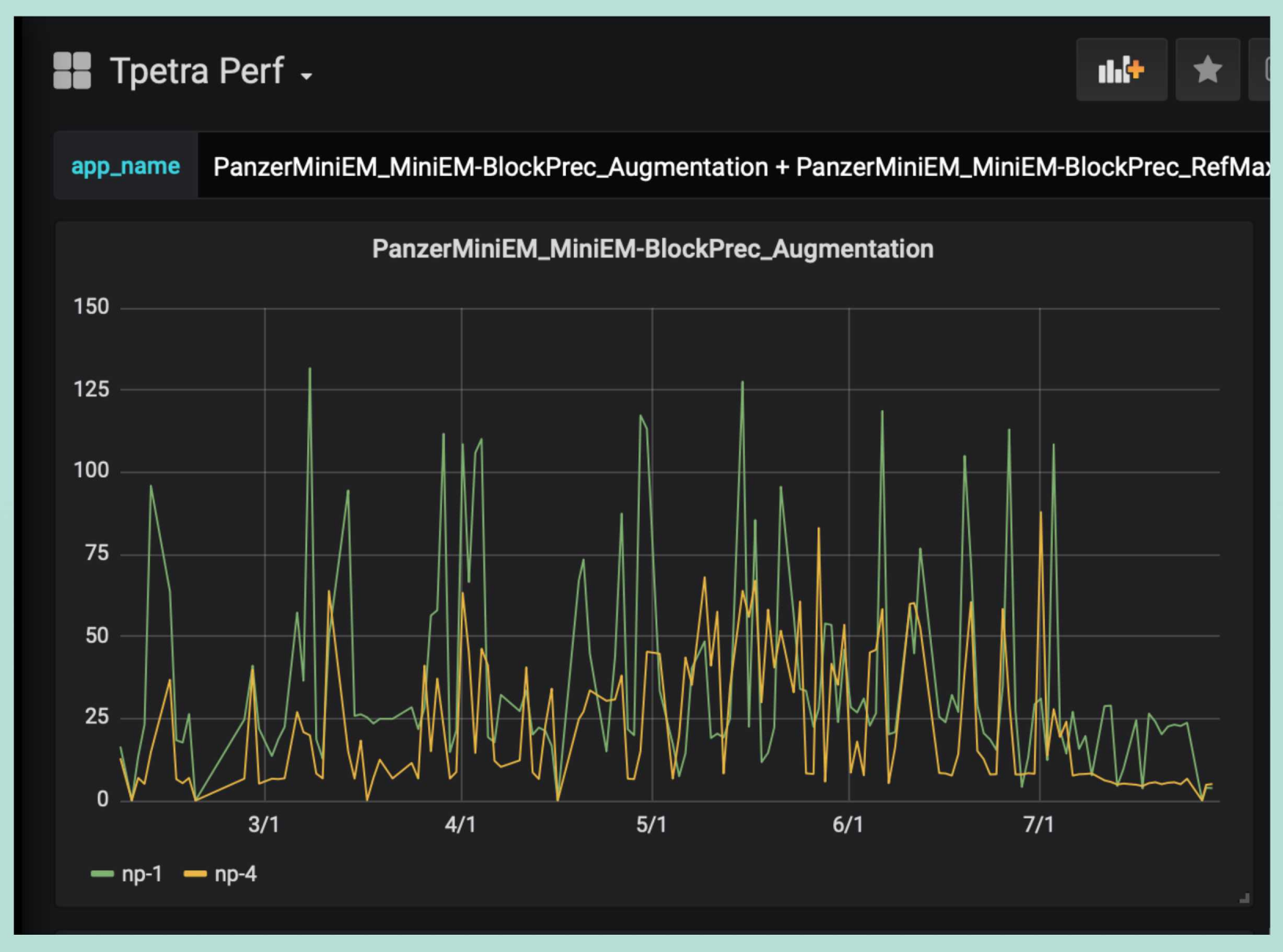
- **Runtime feedback of analysis results** to applications and system software to enable better application-to-resource mapping and co-scheduling decisions
- Enable app teams and library developers to quickly identify **and investigate performance regressions and runtime issues**
- Data collection can occur from existing *independent* tool chains. e.g., **nightly regression testing infrastructure feeds continuous performance data** for improved developer R&D
- **Low-latency Autonomous response can improve HPC operations**

Visualization

User and SysAdmin dashboard associates overall application performance with subsystem Figures of Merit to guide diagnosis


Different FoMs target different HPC subsystem aspects

Job ID	App ID	Node ID	Runtime (s)	Back Pressure	Mem Score	Anomalies	PAPI Perf	App Perf
42093	miniAMR	nid000[52-55]	439	0.0	2	None	Back	1.45
42092	miniGhost	nid000[21;29-31]	1043	49.07	2	Cache	Back	-1.93
42091	miniMD	nid000[57-60]	617	5.24	3	Cache	Back	No data
42089	CoMD	nid000[52-55]	742	91.68	1	Cache	Back	1.52
42088	miniAMR	nid000[21;29-31]	447	0.0	2	Cache	Back	1.45
42087	miniGhost	nid000[57-60]	1043	73.88	2	Cache	Back	-0.27
42086	miniMD	nid000[21;29-31]	619	13.33	3	Cache	Back	No data
42084	CoMD	nid000[52-55]	742	90.88	1	Cache	Back	1.81
42034	miniGhost	nid000[52-55]	1022	98.59	1	None	Back	No data
42028	kripke	nid000[57-58	748	0.0	1	Mem	No data	No data
42027	kripke	nid000[21;29-31]	751	0.0	1	Mem	Back	No data
42019	kripke	nid000[52-55]	1092	0.0	1	Mem	Front, Back	No data



Application metadata enables tracking of historical performance variations

App-specific names and timers *are selectable via Grafana variables (mapping to queries)*

Let us know!

- What information or analyses might you find useful?
- How would you like to get this feedback?

U.S. DEPARTMENT OF ENERGY

NNSA National Nuclear Security Administration

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.
SAND No. ___________

Sandia National Laboratories