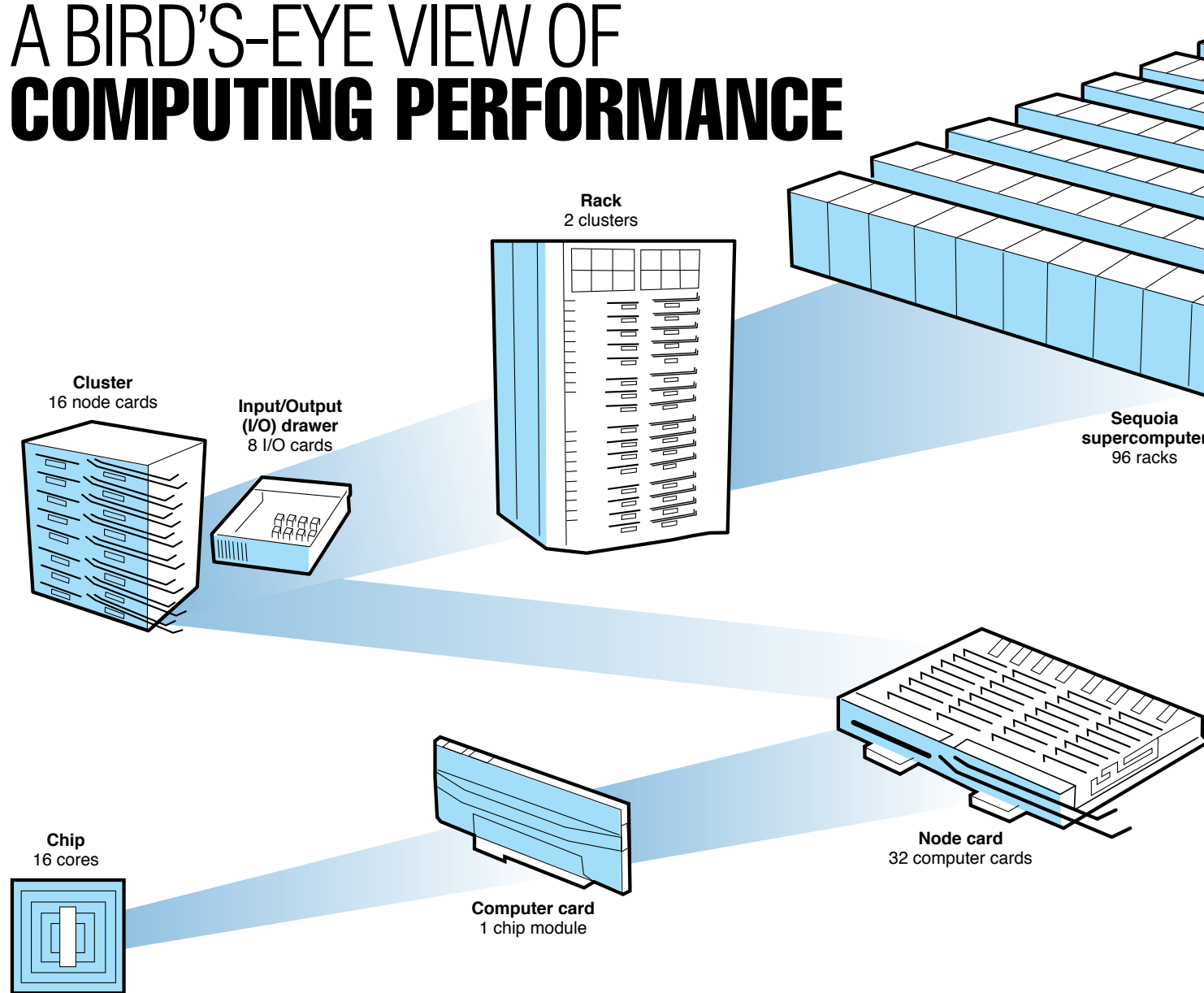


A BIRD'S-EYE VIEW OF COMPUTING PERFORMANCE



LAURENCE Livermore's high-performance computing (HPC) center runs 24 hours a day, 7 days a week. Inside the facility, the noisy hum of supercomputers mixes with the air currents of the building's sophisticated cooling system, providing evidence of the otherwise unseen high-powered processing that is required for a wide range of scientific applications. While the machines are hard at work, performance analysts collect data about every component within the HPC center to ensure smooth and efficient operations. Computer scientist Alfredo Giménez says, "To evaluate our facility's performance, we must make sense of the substantial amount of data the HPC systems generate."

The Sequoia supercomputer at Lawrence Livermore is shown here separated into its constituent parts—from the overall system to individual computer chips—to illustrate the scope of available performance data that are gathered at every level of this hardware hierarchy.

HPC performance data is collected from many different sources within the facility and includes metrics on network utilization, rack temperature and humidity, power consumption, application runtimes, and message routing. Such information is essential for understanding how efficiently the facility operates.

For example, if analysts can determine which applications generate the most heat during processing, they can reconfigure job scheduling policies to prevent resources from overheating. The challenge is that performance data quickly accumulates, making it more difficult to interpret. "As we explored various analytical approaches, we discovered that every data science problem created a data integration problem," says Giménez.

"Understanding data from many sources is more complex than one might initially imagine."

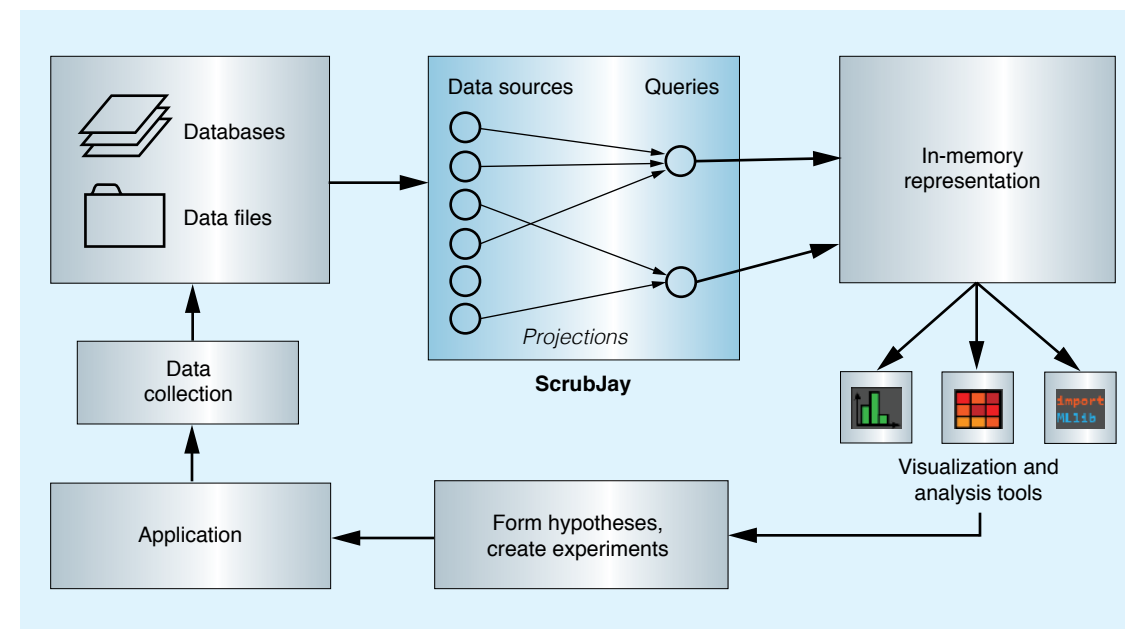
Giménez, along with colleagues in the Laboratory's Computation Directorate, investigates the information that can be extracted from data sources by defining the semantics (or meaning) of the data for better integration and analysis. Giménez incorporated this concept into an analysis tool called ScrubJay—named for the resourceful bird that memorizes scattered locations of food caches. ScrubJay automates and simplifies performance data analysis, helping analysts better understand interactions between applications and the complex systems within the HPC center.

Finding Common Ground

Although many open-source performance analysis tools exist, ScrubJay is unique in that it provides a way to integrate information from various data sources. For example, the batch system that runs parallel applications on Livermore's machines records which nodes ran an application, while the network monitoring system captures the amount of data written on each node. To optimize network utilization, ScrubJay can help analysts cross-reference these two data sets, and others, to determine how different applications affect network communication. Computer scientist Todd Gamblin states, "If we can identify the cause of performance problems, we can schedule jobs to avoid saturating the network or other resources."

Furthermore, analysts must contend with variations in formats, databases, and data types. One data collection tool may record comma-separated values in a plain-text file, and another may store facility information in a third-party database. In addition, time stamps may vary between data sets. ScrubJay eliminates the need for manual integration of such heterogeneous data by first annotating raw data sets with semantics that describe the types of information—for example, whether a table column contains time or temperature recordings. When analysts query ScrubJay, data sets are converted into common in-memory representations via parsing functions known as data wrappers.

ScrubJay's semantics are generalizable to many kinds of numerical and categorical data. Values in a data set can be defined as ordered or unordered, continuous or discrete, and so on. An analyst can also input customized semantics, such as



ScrubJay helps performance analysts evaluate how applications interact with complex high-performance computing systems and enables them to focus on high-level analysis queries rather than detailed data-processing operations.

particle lookup functions in a particle-transport code. ScrubJay’s “derivation engine,” named after inference engines proposed for artificial intelligence in the 1980s, uses the semantics to determine which information can be obtained from one or more available data sets. For example, a new data set may arise from the conversion of units or when a range of values is separated into discrete values. Alternatively, two data sets may merge when their values are mapped to each other. In both cases, ScrubJay integrates the data points through a common semantic dictionary and in a shared format. In addition, ScrubJay contains logic for overlapping data. For one-to-many relationships, the derivation engine aggregates values with the same semantics. This process may result in averaged values, such as a single point in time mapped to multiple temperature readings.

A Collaborative Ecosystem

ScrubJay arose from a Livermore project called PAVE (Performance Analysis and Visualization at Exascale), which established a framework for analyzing the performance of parallel scientific codes. Giménez cites Boxfish, a Livermore-developed platform for performance data visualization, as another important influence in ScrubJay’s development. He states, “Boxfish exposed many research problems identified by PAVE that we subsequently

targeted in ScrubJay, such as how to relate different parts of the computing center to each other.”

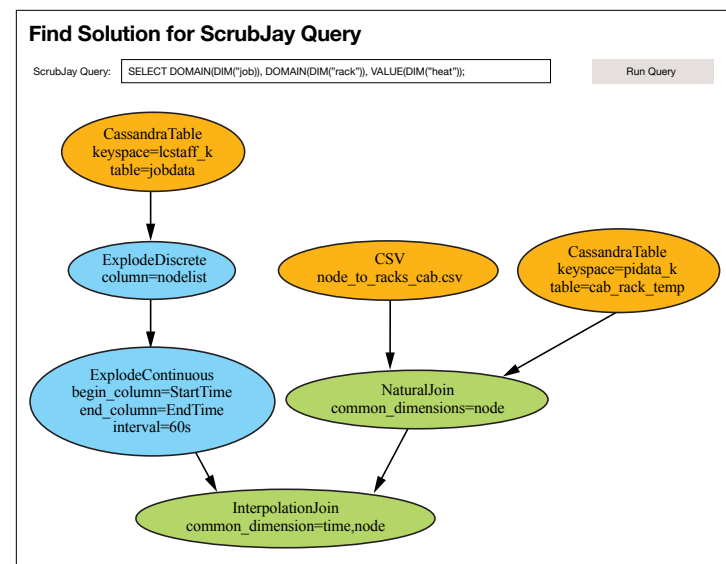
Now funded by multiple programs across the Laboratory, ScrubJay plays an important role in the larger HPC ecosystem. The HPC center’s sheer volume of performance data—tens of gigabytes per day—is monitored, collected, stored, and processed by an infrastructure called Sonar. The system consists of a dedicated computing cluster, a distributed database, and monitoring tools that continuously and securely manage performance data without affecting computing operations. Gamblin, who manages Sonar, explains, “No single person has the knowledge of code experts, data analysts, and system administrators to pull together data from millions of processors. Sonar brings all of the information into one place where ScrubJay can integrate it.”

Analysts access ScrubJay via a web-based dashboard. Instead of querying specific database tables, analysts enter the desired data sources and measurements. The tool automatically stores derivation sequences that can be re-run later or modified. “Analysts can examine a saved ‘recipe’ to see how ScrubJay derived the resulting data set. This reproducibility saves time for both analysts and the machines,” says Giménez. The dashboard interface also allows analysts to create visualizations of the data.

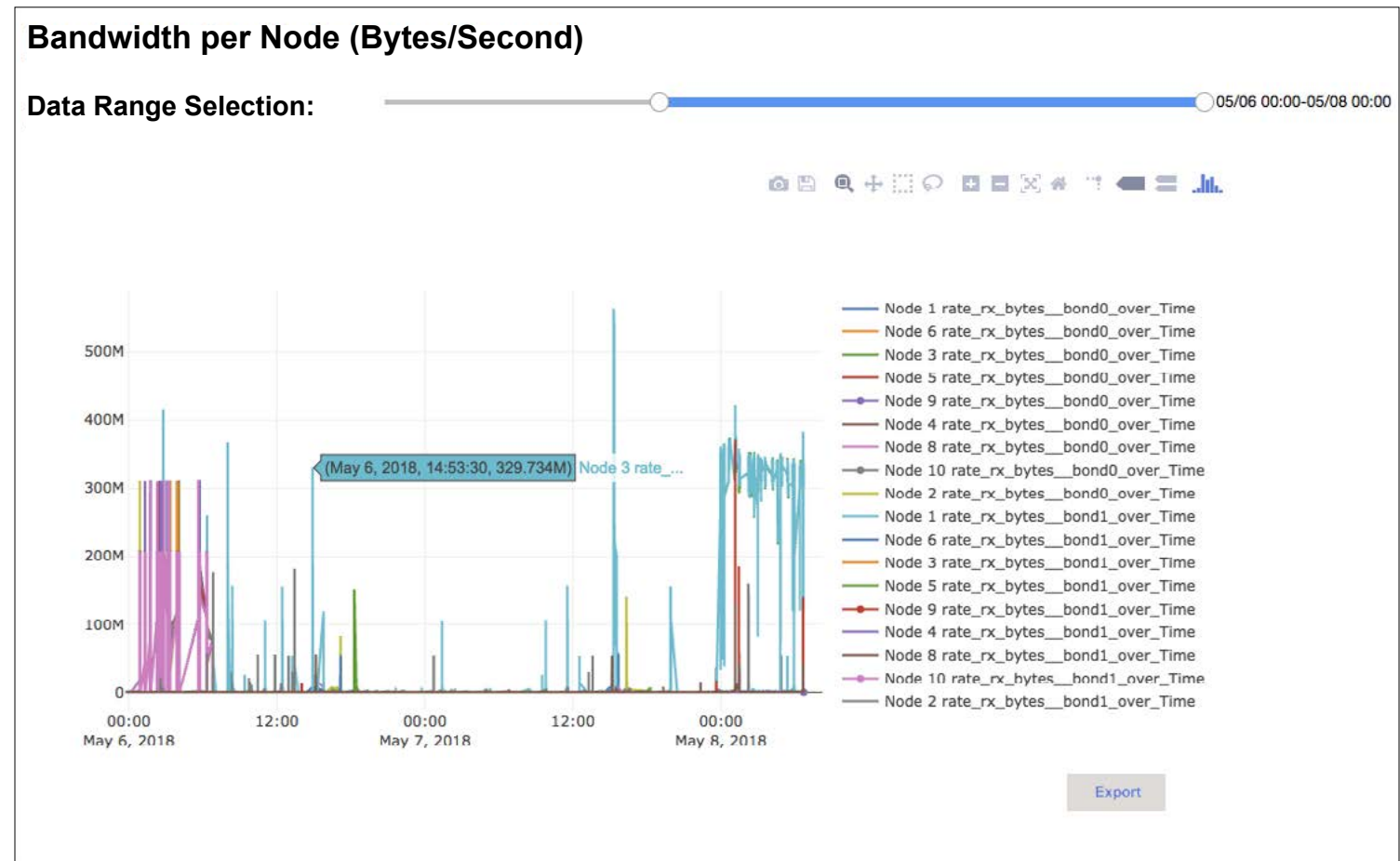
Proof Positive

The ScrubJay team presented two case studies at the international SC17 supercomputing conference to illustrate the tool’s accuracy in integrating facility-wide information. In the first study, Giménez’s team collected job queue logs; rack temperature, humidity, and power usage; and nodes-per-rack layout for one of Livermore’s production clusters. “We wanted to analyze correlations between heterogenous workloads on the cluster, ultimately discerning which applications generated the most heat within the facility,” states Giménez. The team incorporated the resulting data wrappers and semantics into ScrubJay’s knowledge base for future analysis of the cluster’s performance.

The second ScrubJay study added motherboard status, central processing unit (CPU) counters, and other performance variables to the query. The team sought to evaluate the effect of variable CPU frequency on node power consumption. The cluster’s temperature readings were recorded every two minutes, while raw CPU, motherboard, and node data were collected at one- to three-second intervals. Manual analysis would have required tedious mapping of time stamps as well as raw data conversion to derive performance rates of different components on the motherboard. ScrubJay’s automated transformations enabled the team to create visualizations of the derived information to answer their questions. Giménez notes, “This information is useful for understanding how architectural designs affect the way power is consumed from different computing resources.”



Analysts query ScrubJay through a dashboard that displays the derivation sequence and interactively renders the resulting data. In this case, the analyst requested a data set describing the heat generated from different jobs and their corresponding racks. ScrubJay produced a solution involving three original data sets (orange), two transformations (blue), and two combinations (green).



The team has also overcome efficiency challenges. ScrubJay is configured to search a large number of possible data derivations, then execute them on multiple large data sets. ScrubJay also distributes the data-processing operations over the entire Sonar cluster. Giménez explains, “We’re improving ScrubJay’s efficiency in a few ways, including caching intermediate results for reuse.”

New Ideas Take Flight

HPC performance analysis must keep pace with evolving technology. As Livermore’s Sierra supercomputer comes online with ever-faster processing speed, the Laboratory’s HPC teams are adapting complex computing architectures to meet powerful performance requirements. (See *S&TR*, September 2016, pp. 4–11; and March 2017, pp. 4–11.) Optimized performance leads to better scalability, so monitoring and analysis tools such as ScrubJay are crucial for ensuring the Laboratory’s HPC center functions at its best.

Giménez is working toward releasing ScrubJay as open-source software (<https://github.com/LLNL/ScrubJay>), and he looks forward to learning about external researchers’ experiences with the tool at other HPC facilities. In the meantime, he plans to

An interactive plot of a ScrubJay-derived data set shows active bandwidth on different networking interfaces for each node in a cluster. The raw data set contains the number of bytes transferred over time, from which ScrubJay obtains bandwidth per node. This information helps analysts determine how much available bandwidth is being used over prolonged periods of time.

incorporate additional data transformations and apply machine learning to explore conditions that affect a code’s efficiency. Giménez is also adapting ScrubJay for integration with other scientific and simulation data, including physical simulations from Livermore’s National Ignition Facility. He states, “Data integration is a big challenge for all data scientists. We hope our solutions can help a variety of scientific fields.”

—Holly Auten

Key Words: data analysis, data integration, data science, derivation engine, high-performance computing (HPC), ScrubJay, Sonar.

For further information contact Alfredo Giménez (925) 422-0431 (gimenez1@llnl.gov).