Abstract
Understanding and improving the performance and efficiency of HPC centers requires detailed analysis of running systems. To this end, modern HPC facilities provide extensive capabilities for collecting performance-related data for analysis. However, these data sources are most often disparate from one another, measuring different components in different domains. It is not clear, for example, how to correlate per-rack temperature readings with mesh input sizes recorded for a particular physics simulation.

We are developing a performance analysis system that combines disparate data sources into a centralized database and automatically performs complex transformations on the data to yield indirect relationships between them.

HPC Performance Data Sources
...in the hardware domain:
- Network Data
- Node Data
- Rack Data
- Facility Level Data
- Cluster Level Data
- System-level Data
- Performance data
- Enclave-level data
- Application-level data

...in the software domain:
- Libraries
- Applications
- Environments
- Annotations

Merging Disparate Data
Disparate data sources often require more advanced merging than a simple SQL JOIN operation.

Case: No one-to-one mapping
<table>
<thead>
<tr>
<th>Time</th>
<th>Temperature</th>
<th>FLOP Counter</th>
</tr>
</thead>
<tbody>
<tr>
<td>10:00</td>
<td>24453</td>
<td>6453</td>
</tr>
<tr>
<td>10:01</td>
<td>24453</td>
<td>6453</td>
</tr>
<tr>
<td>10:02</td>
<td>786</td>
<td>786</td>
</tr>
<tr>
<td>10:03</td>
<td>24453</td>
<td>6453</td>
</tr>
</tbody>
</table>

Case: Same domain, different units
<table>
<thead>
<tr>
<th>Time</th>
<th>Temperature</th>
<th>FLOP Counter</th>
</tr>
</thead>
<tbody>
<tr>
<td>10:01</td>
<td>64.6</td>
<td>24453/24</td>
</tr>
<tr>
<td>10:02</td>
<td>55.6</td>
<td>24453/786</td>
</tr>
</tbody>
</table>

Solution: Semantic Table

<table>
<thead>
<tr>
<th>Column</th>
<th>Units</th>
<th>Aggregator</th>
<th>Conversions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>cycles</td>
<td>N/A</td>
<td>[T] =&gt; [time]</td>
</tr>
<tr>
<td>Temperature</td>
<td>Celsius</td>
<td>Average</td>
<td>[T] =&gt; Fahrenheit</td>
</tr>
<tr>
<td>FLOP Count</td>
<td>count</td>
<td>sum</td>
<td>none</td>
</tr>
</tbody>
</table>

This tells us:
1) If two data sources may be merged, and
2) how to merge them.
3) Possible datasets that may be produced by different sequences of merges (see below)

Preliminary Results
Dedicated Access Time (DAT) for 2 days on Cab (1296 nodes)
Data collected:
1. Job queue information (slurm)
2. Facility temperature (9 sensors per rack; 23 racks)
3. Facility layout (assignment of nodes to racks)
How much heat is generated by different jobs?

The SONAR Data Cluster
Below:
- Heat over time for rack 17
- Generated heat = temperature difference between hot and cold aisles

SONAR:
- newly deployed data cluster
  - 13 nodes, SSDs, data software stack
  - Apache Cassandra distributed database
  - Apache Spark distributed data-local processing (used here)