

# A Day in the Life of a Supercomputer

William Clem, Sean McCulloch, Tanzima Islam, Western Washington University  
Kathleen Shoga, Lawrence Livermore National Laboratory

We are developing an interactive dashboard for visualizing the statistics of how users interact with nodes of a supercomputer. The intended purpose of the dashboard is to periodically provide with the overview of the load and bandwidth utilization of the nodes of a large supercomputer and details of each node and job as selected by the user. The dashboard will be interactive which will enable users to zoom in on interesting parts of the visualizations to investigate in details.

## Goal

The goal of the visualizations is to graph data collected from the RZGenie supercomputer in such a way that the overall performance of RZGenie can be improved. The graphs help in comparing nodes against each other to determine if there are any “bad” nodes, if they are being underutilized, and if they are using too much of the shared resources among other things.

## Research Questions

- How do nodes perform relative to each other across different metrics?
- How much of the available bandwidth is used by the nodes throughout the day?
- How are jobs distributed across the nodes?
- Is there any obvious patters of the cluster usage with respect to the time of the day?

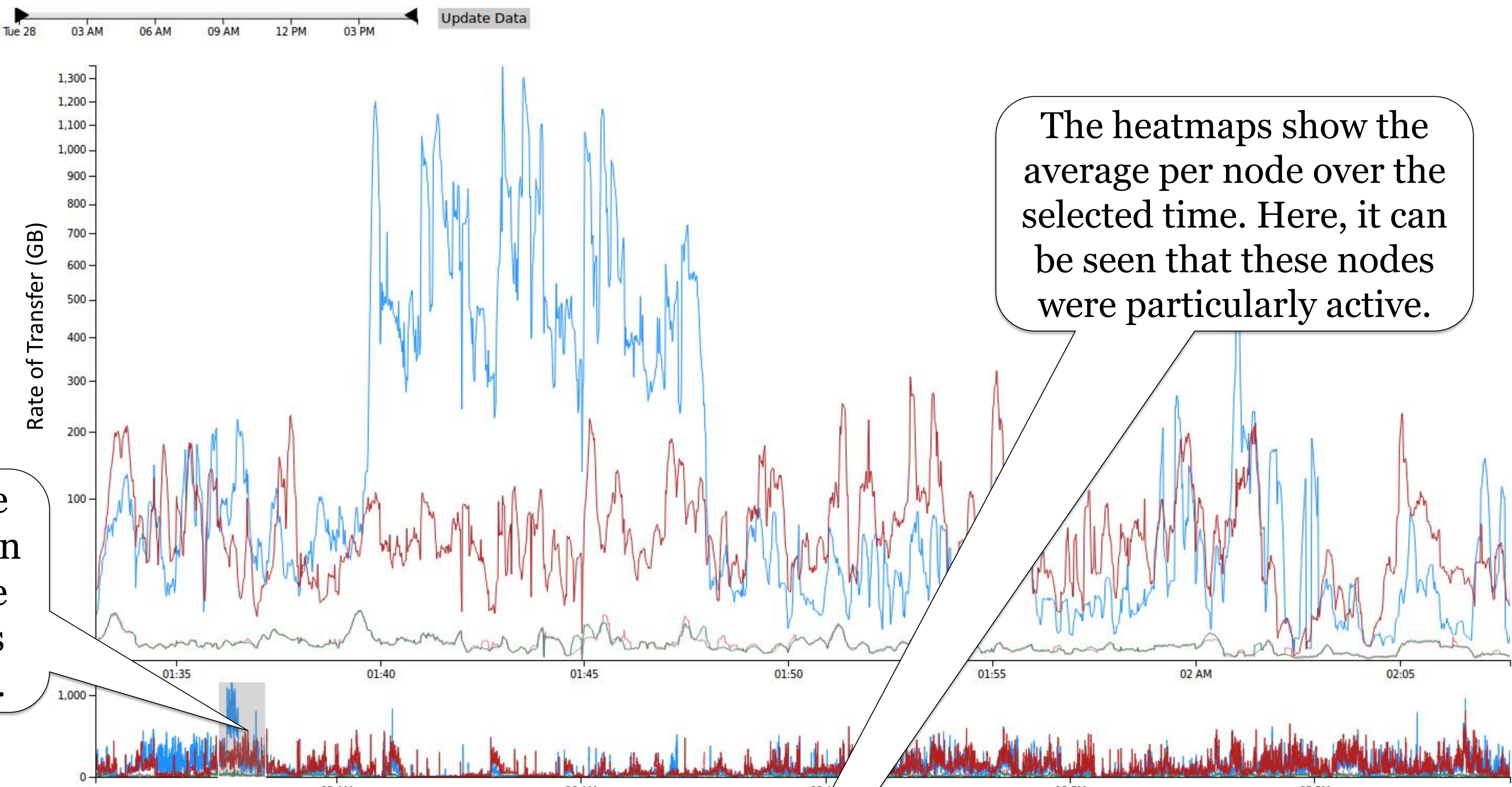
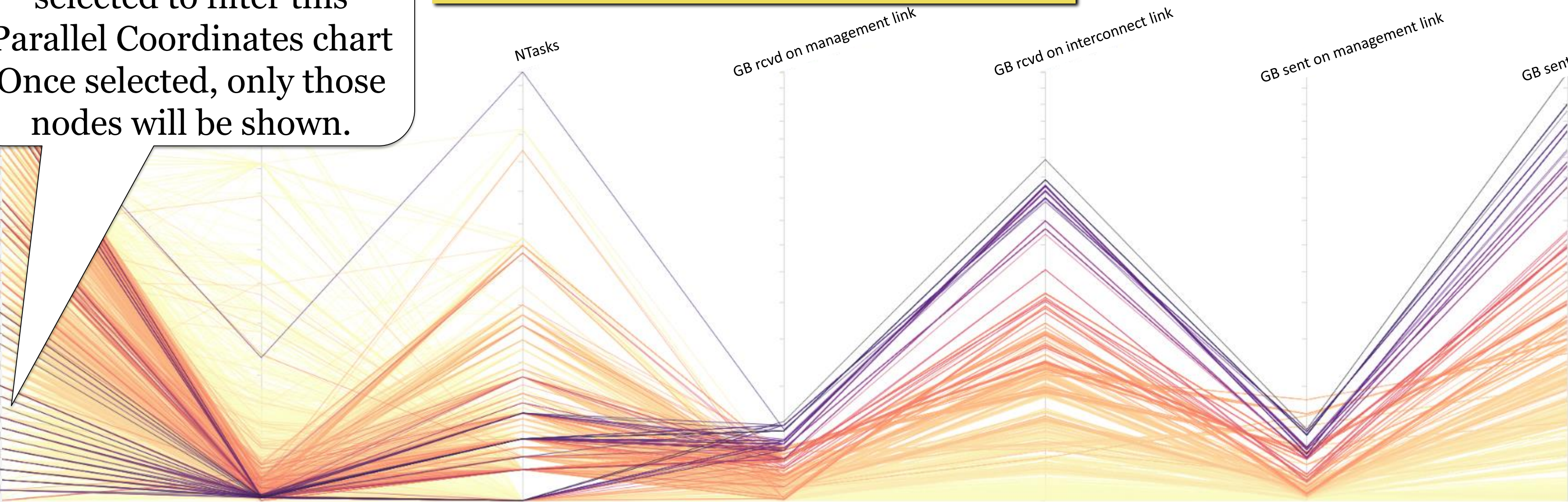
## Approach

We focused on using graphic design principles along with processing the data provided to create visualizations that answered our research questions. We used a combination of Data-Driven Documents (D3.js), HTML, and CSS to make the visualizations.

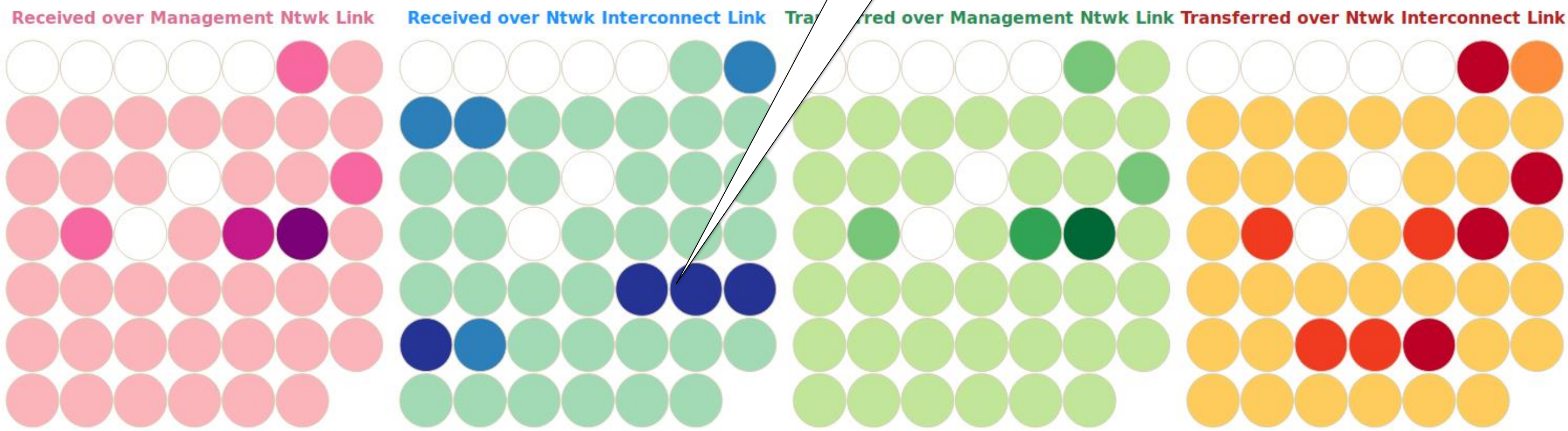
- The choice of these tools enabled us to:
- Create a portable web-based dashboard
  - Add interactivity and depth to the visualizations
  - Allow for continual development and refinements.

Nodes of interest can be selected to filter this Parallel Coordinates chart. Once selected, only those nodes will be shown.

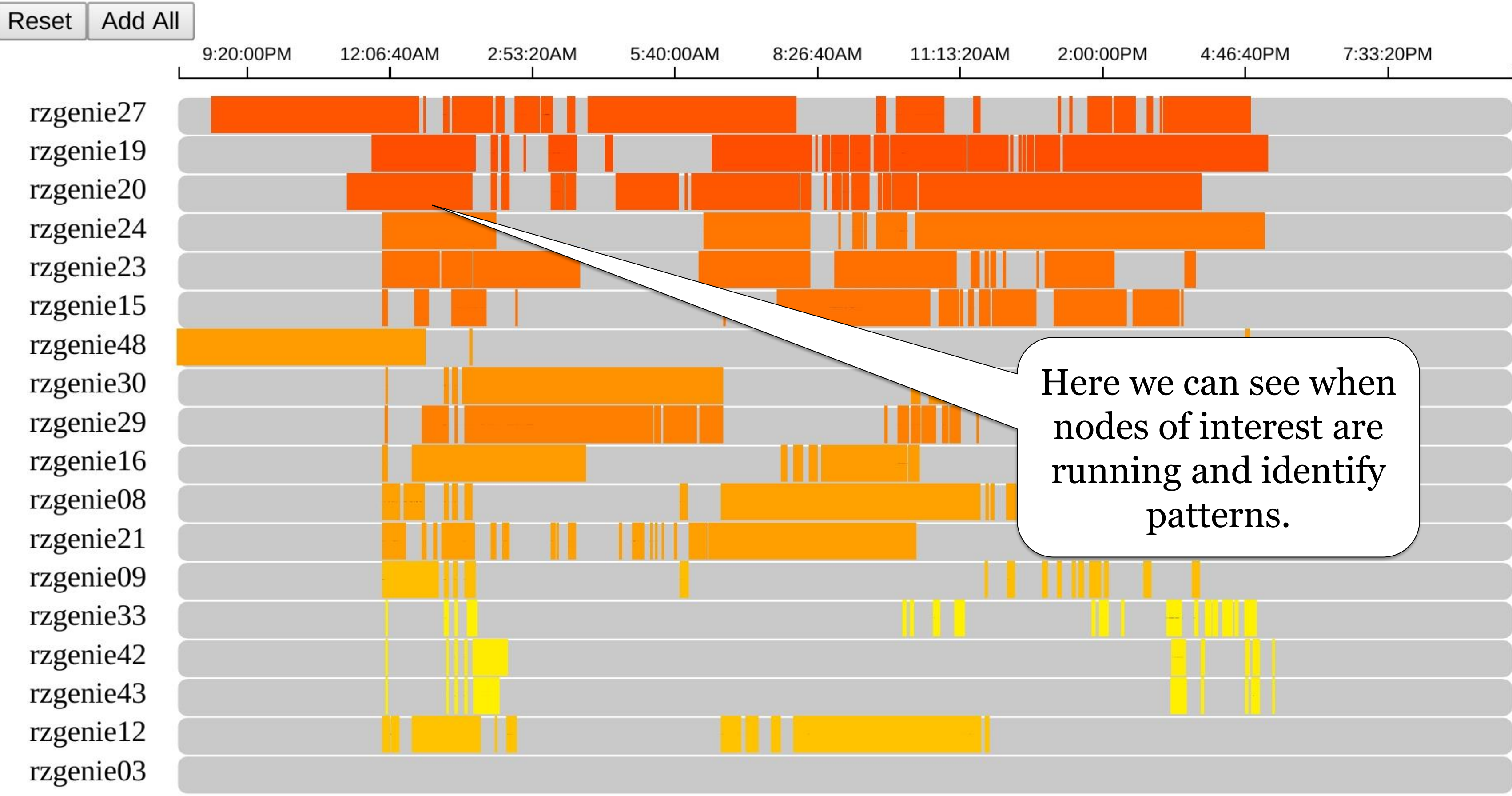
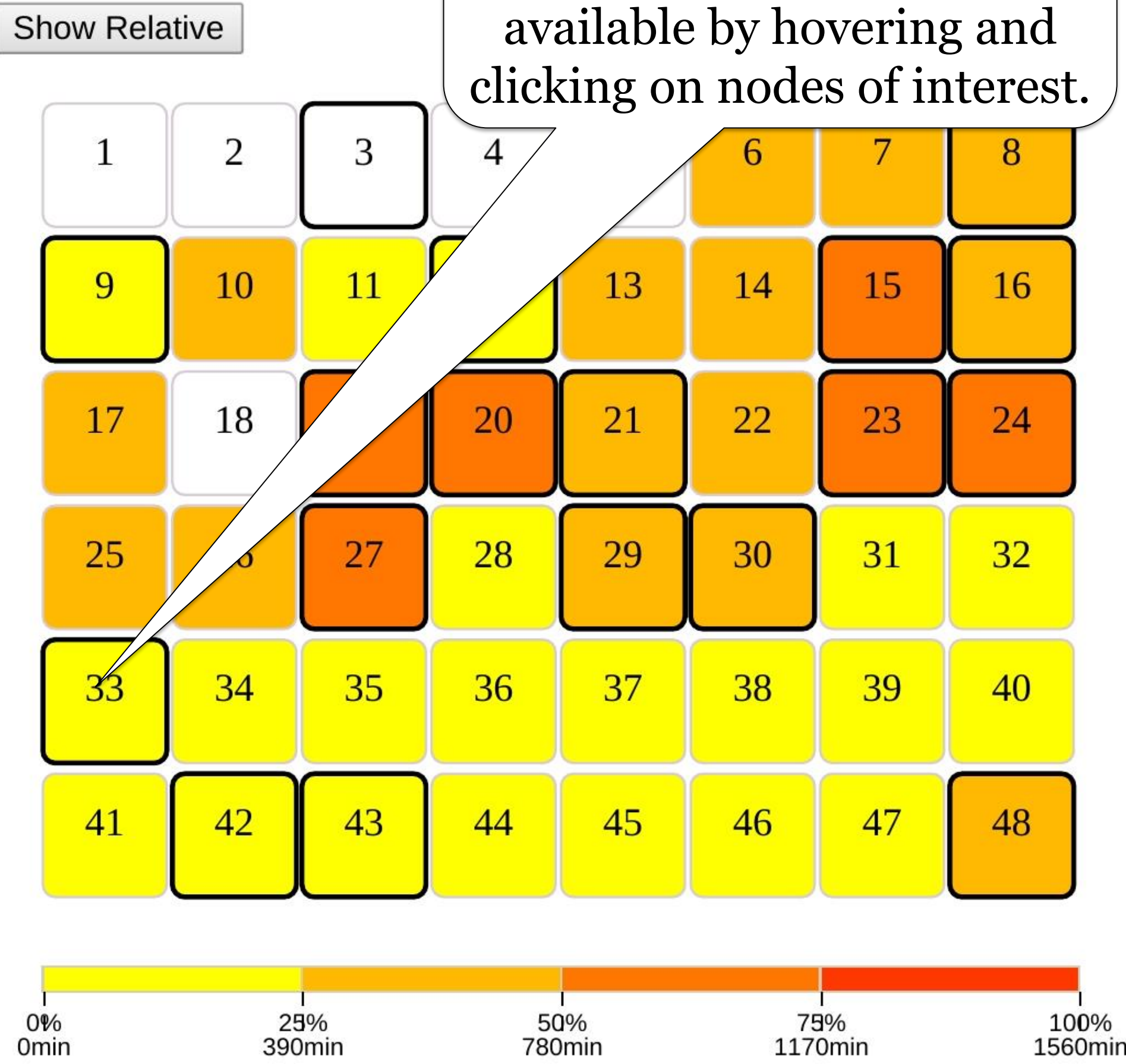
## HPC Job Statistics



The heatmaps show the average per node over the selected time. Here, it can be seen that these nodes were particularly active.



Total node activity can be viewed here, more info is available by hovering and clicking on nodes of interest.



Here we can see when nodes of interest are running and identify patterns.

## Future Work

- Answer more of the presented research questions
- Put all the graphs into one dashboard
- Improve behind-the-scenes performance