

Improving the User Interface to OSG-LHC Network Metrics

High-Energy Physics relies upon the network to enable physicists to use the global set of resources available. Because of the importance of the network for distributed science, the Open Science Grid (OSG) has constructed a network measurement platform to gather detailed network metrics. This platform includes perfSONAR toolkit instances deployed in OSG and WLCG (Worldwide LHC Computing Grid) sites around the world and is managed and maintained by the OSG-LHC efforts in IRIS-HEP. The network metrics are centrally collected and pipelined into the ATLAS Analytics Platform at the University of Chicago. This Analytics Platform serves as a central repository for the complex and extensive network data. In principle, this rich data set is very useful for researchers and network engineers to understand and identify problems in the network and locate broken or mis-behaving parts of the network. In practice, when problems occur, it is difficult to identify the problem in the large, continuously expanding set of metrics. With large sets of metrics, network data holds information that may otherwise be completely overlooked or impossible to identify. We would like to be able to break down the complex information in ways that are user friendly, permitting rapid identifications of problems compared to an individual searching through the complex network data for an extended period of time. Such an automatic and quick system, therefore, would be very useful for users of our research and education networks, especially for high-energy physicists, allowing them to take full advantage of the entirety of the network data being collected.

My project would be to improve the user interface to the OSG-LHC network metrics to enhance the users' understanding of the entire global network, allowing speedier identification of problems and performance issues. This project will be supervised by Shawn McKee, OSG networking area coordinator and Director of the ATLAS Great Lakes Tier-2 (AGLT2). Similar to the AGLT2, the ATLAS Analytics Platform at the University of Chicago stores its data in Elasticsearch. My plan is to utilize Kibana to create relevant visualizations and organize them into dashboards for specific use-cases. There are already established dashboards that represent **specific metrics**, but my plan would be to create **new dashboards based upon important use-cases**. These dashboards will assemble different types of metrics already represented on the ATLAS Analytics Platform to provide new representations of data that can not only give a holistic overview of a problem, but also utilize the **cross-referencing of the metrics to help identify causes and triggers of the problems**. An example use-case would be for a site administrator wanting to understand if their site was working properly. The dashboard I would create for this use-case would utilize throughput, packet-loss and traceroute data, all with the specific site selected as either source or destination, depending upon the specific visualization. Having dashboards targeted at specific use-cases will make the data being gathered much **more useful** and **effective**. My plan includes identifying relevant use-cases and incorporating user feedback. I plan to engage with IRIS-HEP, SAND and the OSG community to identify the most relevant use-cases to support and then ask for feedback on my prototype implementations.

Week 1-2: Familiarize myself with the type of data stored on the ATLAS Analytics Platform and identify different use-cases that would be most beneficial to represent by asking users what would be the most beneficial to represent

Weeks 3-4: Create the dashboard(s) for the use-cases using both information based on the current metrics already being represented, as well as other metrics and information that I can identify that hasn't been created into a visualization.

Week 5: Work with the community who is the most likely to need the dashboards via email or presentation to identify areas that are working well and areas that need some more development.

Weeks 6-7: Going off the user feedback, continue to develop the use-cases, adding the information that was desired as well as any information that was discovered in the development, improving the capabilities of the use-case dashboards.

Week 8: Once again turn to the community for more feedback for more input on the current use-case dashboards created, but also querying about different use-cases that can also prove to be useful.

Weeks 9-10: Finish polishing the initial use-case dashboards and start to develop the newer use-case dashboards based on the community feedback on what is the most important to represent.

Week 11: Gather more community feedback on the final products of the first use-cases, as well as the developed second use-cases, identifying any areas that need to be better represented and working on them.

Week 12: Finalize and document the dashboards identified by working with the user community and incorporating their suggestions and feedback.