Understanding Ganglia Resource Monitoring Data on SERC HPC systems.

09-02-2012

On the High Performance Computing systems (HPCs), a tool that can monitor and show real-time usage of resources is very useful. Such a tool can indicate to the potential users, the current state of resource availability so that necessary planning for job execution can be made. It also helps to analyse the running jobs in terms of how they are using the resources allocated. It is necessary that the data be presented in intuitive and visually understandable way so that quick inferences can be made. Ganglia is one such open source software that provides such monitoring with a visual interface and can be easily integrated into a variety of HPCs.

Ganglia software (http://ganglia.sourceforge.net/) provides a complete real-time monitoring and execution environment and is in use by hundreds of universities, private and government laboratories and commercial cluster implementers around the world. Ganglia was developed at the University of California, Berkeley Computer Science Division as way to link clusters across the Berkeley campus together in a logical way. Since it was developed at a university, it is completely open-source and has no proprietary components. All data is exchanged in well-defined XML and XDR to ensure maximum extensibility and portability.

Ganglia provides a wide variety of information. In SERC, Ganglia is installed on the basic HPCs, i.e., on tesla, fermi, dell and tyrone clusters. Ganglia ties up these systems as a single pool. At the top level, data regarding availability of nodes and cores and their average use is given (refer picture in Figure 1).

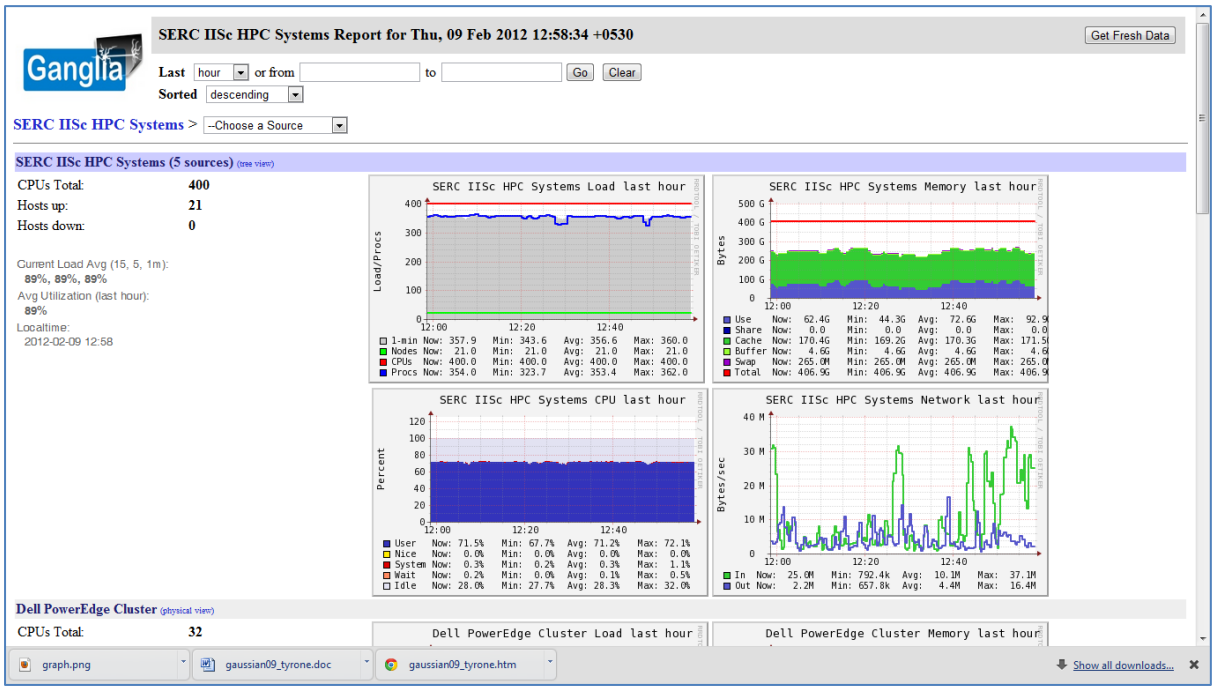

Figure 1 Top-Level Ganglia display for Basic HPC systems in SERC

The left hand side statistics display the global information quantitatively. For this picture, there are totally 400 CPUs that are being monitored from 5 systems, on which there are 21 hosts and all are up. The right side graphs display popular usage metrics pictorially. The information in these graphs display different metrics like system load averages (top left graph), different types of memory usage averages (top right graph), %CPU utilization averages (bottom left graph) and network bandwidth averages (bottom right graph). Each of these metrics is described below:

1. System Load averages: Indicates the average number of processes running on the systems. In this graph the average is around 89%. This means that around 356 processes (89% of 400) are currently running on these machines.
2. Memory usage averages: Indicates the average usage of memory, in the form of user process or shared memory areas, system cache, buffer or swap areas.

3. CPU Utilization averages: Indicates % CPU utilization across all processes on all systems.
4. Network bandwidth usage averages: Indicates the average use of network bandwidth across the nodes.

The top bar located above the charts gives options for selection, with regard to the subscribing nodes of Ganglia monitoring setup. Using the options one can look for data about a specific system, in this case one of the clusters, fermi, tesla, dell, tyrone or altix350. One can also select information averages with respect to hour, day, week, month or year. A sample output for tyrone-cluster is illustrated below.

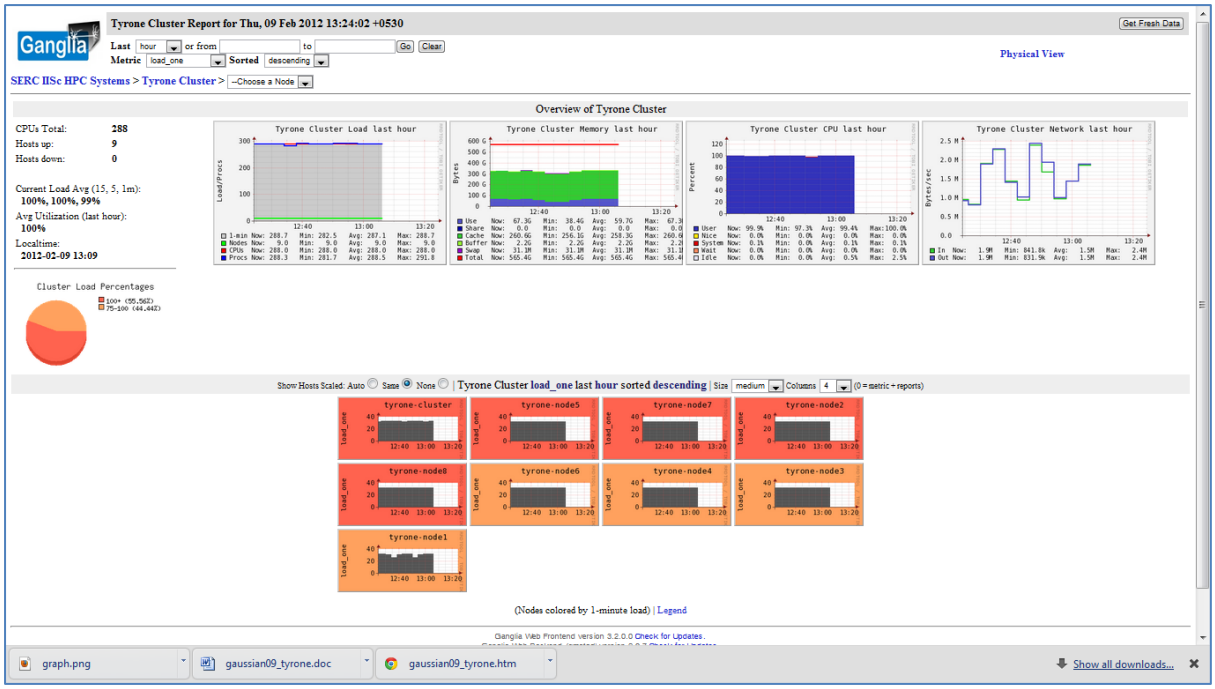

Figure 2: Tyrone-cluster usage report using Ganglia

The picture in Figure 2 gives a comprehensive view of resource usage on the different nodes of tyrone-cluster. Apart from the standard metrics, it also gives a color based utilization of each node. In this specific example, all nodes are seen to be having load averages of 75% (orange color) and above (red color). The top bar of the picture provides option for other metrics for the nodes of the cluster. This depiction helps a user to understand if his job running on a specific node is utilizing the allocated resources efficiently.