Arbiter: Dynamically Limiting Resource Consumption on Login Nodes

Dylan Gardner, Robben Migacz, and Brian Haymore
University of Utah Center for High Performance Computing

Center for HIGH PERFORMANCE COMPUTING
THE UNIVERSITY OF UTAH

Arbiter prevents the abuse of shared resources on cluster login nodes by tracking usage and applying progressively stricter and longer throttling on the resources available to users with excessive consumption.

Motivation

Login nodes are the access point for all users of a high-performance computing cluster. High usage on such nodes can result in service interruptions and degraded performance for users trying to access computational resources or perform lightweight tasks like compiling software, scripting, and staging data.

Arbiter is a service that protects login nodes from users who use shared resources immoderately by:

- Limiting access to resources for all users to protect nodes from high usage and ensure reliability
- Throttling users with excessive usage to encourage running resource-heavy tasks on computational resources
- Sending an email with usage information, as in Figure 1, to the user and staff when throttling occurs

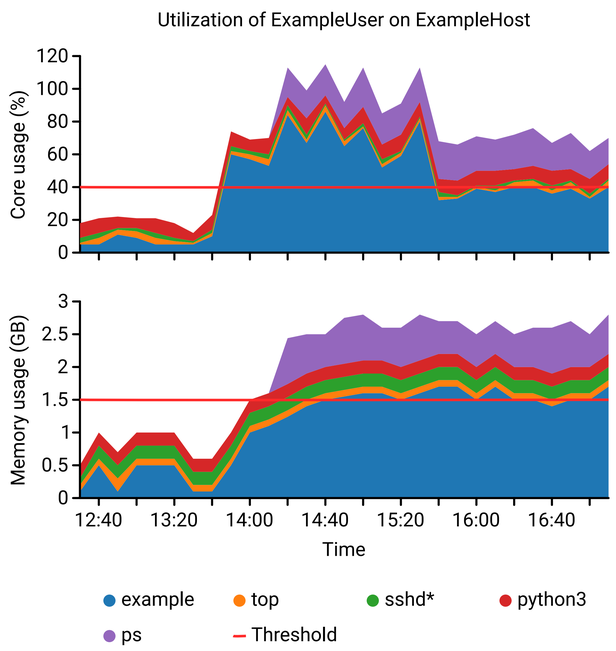

Figure 1. An example of a usage plot include in an email to a user. The approximate usage of each process is stacked (and ordered by impact). A table of recent averages of usage metrics is also included in emails. Verbose descriptions of usage help users understand which processes have the biggest impact on the login node and are therefore most likely to affect others' workflows.

Implementation

Arbiter, written in Python 3, uses *cgroups*, a feature of the Linux kernel, to monitor and limit groups of processes. A slice is associated with the user and used to read usage information and set quotas.

Tracking and Notifying Users

The Arbiter service collects usage information periodically and uses average values to determine a user's impact on the node.

- A *badness score*, depicted in Figure 2, is used to track impact over time
  - This system allows brief tests using all resources above their thresholds
  - Scores decay slowly with time if usage is below the threshold
- When the badness score is larger than a threshold, an email is sent to staff and the user (who is subsequently throttled)
  - Repeat offenders receive lower limits for longer times
- The total usage on the node is also tracked; emails about high overall usage are sent to system administrators

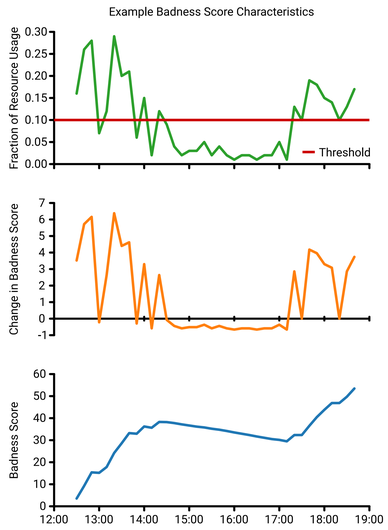

Figure 2. An example of a user's badness score for a given resource usage pattern. The user's score increases when usage is above the threshold and decreases when it is below.

Applying Limits on Resources

The resources available to a given user are reduced when the badness score reaches a threshold. Default quotas are also applied when users log in. An example of the effects is shown in Figure 3.

Memory is stepped down continuously if the usage exceeds the quotas associated with throttling.

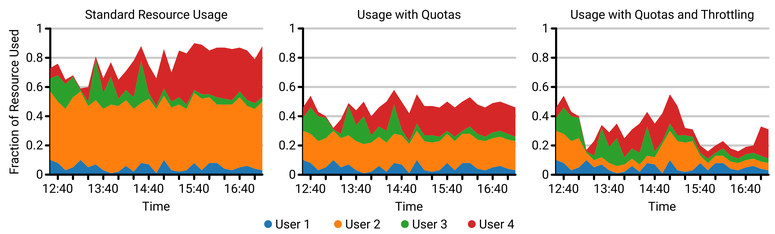

Figure 3. An example of how Arbiter reduces the load (for a given resource) on a server. Left: usage with no limits. Center: usage with standard quotas applied; overall load on the server is reduced. Right: usage with quotas and throttling of users with high usage. The total system load is further reduced.

Impacts

Arbiter has been running on CHPC resources for only a short time, but it has already yielded results on login nodes.

Reducing Load on Login Nodes

Arbiter reduces the impact any one user can have on a given login node and helps ensure the resources remain available and responsive to others. High overall usage is now possible only where multiple users are responsible.

There has been a marked decrease in the number of high usage emails since throttling was enabled, as shown in Figure 4. Before this, Arbiter was running but not throttling usage and was sending messages only to CHPC staff.

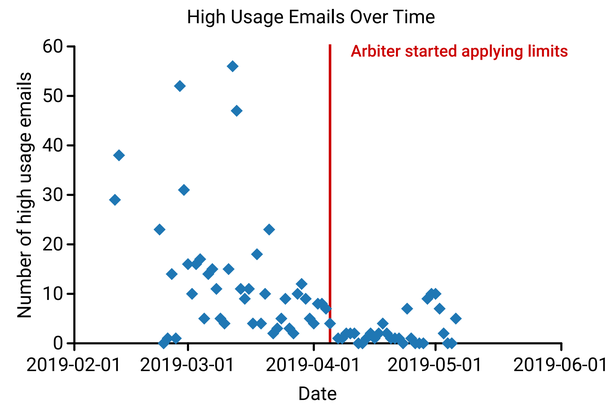

Figure 4. The number of overall high usage emails received on a given date (across all login nodes on all clusters). There is a marked decrease in the number of emails when Arbiter's throttling features are enabled.

Arbiter throttles users an average of 10.7 times a day across all login nodes on all clusters deployed at CHPC.

Improving Communication with Users

Messages sent by Arbiter describe the recent usage on the node with easy-to-digest plots and tables of processes.

Policy information is also included in email messages to help users better understand acceptable usage on the login nodes and provide alternatives for running programs that might impact other users.

Reports included in emails can be re-created from log files and easily searched, allowing administrators to identify trends and adjust Arbiter to better meet their needs.

Acknowledgments

We would like to thank Anita Orendt and the user services team at CHPC for their continued support. We would also like to thank Paul Fischer for his contributions and suggestions and Albert Lund for creating the predecessor of Arbiter.