Computing Cluster Usage and Cost Optimization

Melvyn Thomas[1], Saurabh Suman[2], Chaitali Pawar[3], Daniel Whitenack[4]

Purdue University Krannert School of Management

danie149@purdue.edu[1]; ssuman@purdue.edu[2]; cpawar@purdue.edu[3]; dwhitena@purdue.edu[4]

Abstract

This project aims to perform a detailed analysis on the Compute Cluster server logs by studying the past performance of clusters on a granular level and propose actions based on results. Every year, Purdue spends millions of dollars to build infrastructure to support its on-going research activities. Its ITaP Research Computing team delivered over 300 million computational hours to community cluster partners just in 2017. Our proposed method helps Purdue's research computing team maintain maximum utilization of its cluster capability and ensure a high resell value for the hardware after decommissioning. Our results motivate a method for all research teams to ensure that the supercomputing team maintains profitability by delivering services only for the resources they use. This eventually creates a win-win situation for both the buyers (researchers) and the seller (Purdue Supercomputing).

Introduction

Purdue's Professor in Physical Chemistry, Joseph Francisco has said,

"I wouldn't have been elected to the National Academy of Sciences without these clusters. Having the clusters, we were able to set a very high standard that led a lot of people around the world to use our work as a benchmark, which is the kind of thing that gets the attention of the National Academy."

- Computing clusters on campus are important for academic institutions in the US having excellent research faculty and industry tie-ups
- Purdue University has its own computing clusters with a lifespan of five years that support important research activities.
- Potential of these infrastructure can only be truly harnessed when their configuration and hardware of their clusters is at par with demand from the on-going research activities

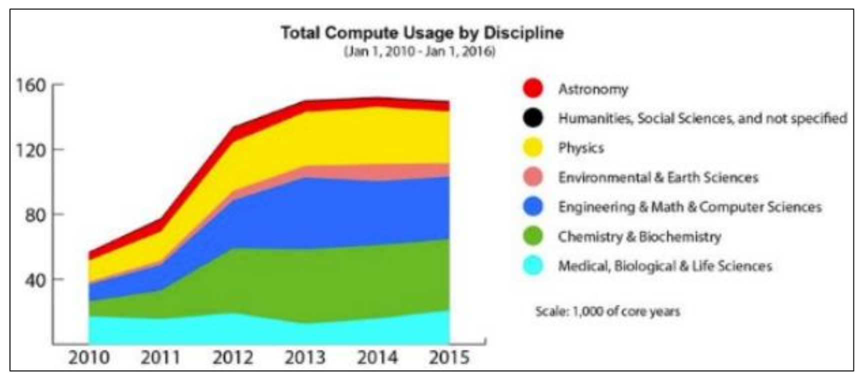

Figure 1. Compute usage of clusters by discipline

Applications running in the clusters are used carelessly rendering a very low percent utilization of clusters.

- Users are working on similar applications on different clusters, it is beneficial both for the user and the seller to choose the cluster that ensures maximum utilization with minimum costs
- Configuration, size and computational power of these clusters should align with the future research requirements for the length of time of clusters usage.
- Our analysis focuses on balancing the utilization of applications across the clusters as well as ensuring maximum cost advantage for users requiring a specific application requirement.

Literature Review

- In principle, there are several types of optimizations, like CPU optimization aims at maximizing the speed and efficiency of the operations being performed by the CPU, I/O optimization involves maximizing the efficiency of input/output operations whereas memory optimization focuses on maximizing the efficiency of a code's memory management. [1]
- Users with large job submissions generally have more HPC knowledge than users only with small job submissions. However, in general, the job runtime estimate is shown to be highly inaccurate. The super computer workload does not show stationary patterns over a month or a year. But the workload exhibits obvious patterns in weekly or daily cycles. [2]

Max. memory use (GiB/node) \ Job size (nodes)	(1,1]	(2,2]	(2,4]	(4,8]	(8,16]	(16,32]	(32,64]	(64,128]	(128,256]	(256,512]	(512,1024]	(1024,2048]	(2048,4096]	(4096,4290]
[0,12)	2.0%	1.4%	5.6%	10.4%	10.1%	9.4%	8.2%	5.4%	4.0%	2.3%	1.4%	0.5%	0.0%	0.0%
[12,24)	0.1%	0.6%	0.8%	1.9%	4.9%	3.3%	2.4%	2.0%	1.4%	0.6%	0.3%	0.2%	0.0%	-
[24,48)	0.2%	0.2%	0.6%	0.9%	1.3%	2.1%	1.5%	1.2%	1.5%	1.3%	0.2%	0.4%	0.0%	-
[48,96)	0.1%	0.2%	0.1%	0.3%	0.6%	1.1%	1.0%	2.1%	0.9%	0.3%	0.0%	0.2%	0.0%	0.0%
[96,128]	0.0%	0.0%	0.0%	0.0%	0.1%	0.1%	0.3%	0.1%	1.0%	0.2%	0.1%	-	0.0%	0.0%

Table 1. Memory Usage Vs Nodes (Job Size) [3]

- This study is novel because we compare the application level analysis in each cluster in contrast to all the research topics where cluster/node level analysis were being carried out.

Methodology

Figure 2 outlines the overall process showing how computing cluster infrastructure caters to user logs and how this log data is extracted, processed and combined with the cost data for nodes and are analyzed on tableau aiming for maximum utilization and cost optimization.

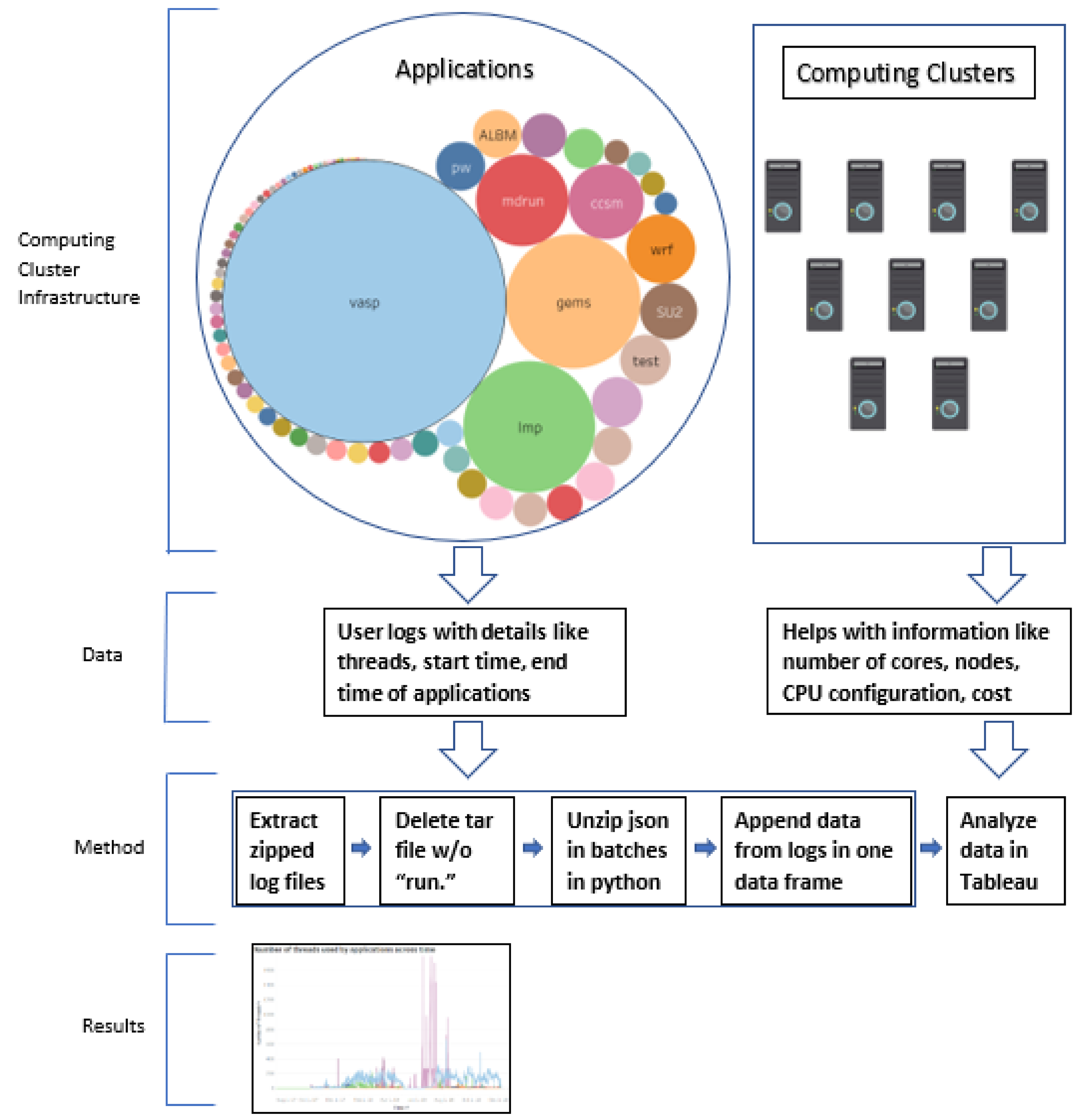

Figure 2. Complete Project Methodology

Data

- Log data was created in XALT by application usage on cluster nodes.
- We received JSONs from XALT directory using Globus application
- The per node cost and node data was obtained from the computing clusters.

Method

- Extract all the zipped log files in a local folder through shell scripting
- Run a shell script inside each tar.xz files that deletes the zip files that have no run*.json files as we are only interested in the running application logs.
- Run another application to create a directory that holds extracted files temporarily and clear all tar files not relevant for our analysis.
- It also extracts each json one by one and till it reaches 10000, append the rows in python dataframe and later clean data of form link*.json.
- This dataframe rows are appended to a master dataframe , which when reaches the maximum capacity of 40000 rows , generates an output in the form of csv file. The clean up is then done for all the processed files and next round is executed.

Business Performance Measures

- Runtime of applications across cores helps in evaluating the utilization of each application, this way we can find the $ value of resource utilized by each application. It will in turn help us get the most effective utilization pattern.
- Another important metric to analyze here is number of distinct users on clusters and number of distinct applications that are run on it.
- A better utilization of the cluster by different research groups aids in providing the most affordable way for research groups will be a win-win situation for both.

Results

The plot below shows how the CPU core usage of application varied with respect to time. We see that the last two year has been quite eventful in terms of resource utilization on supercomputing cluster.

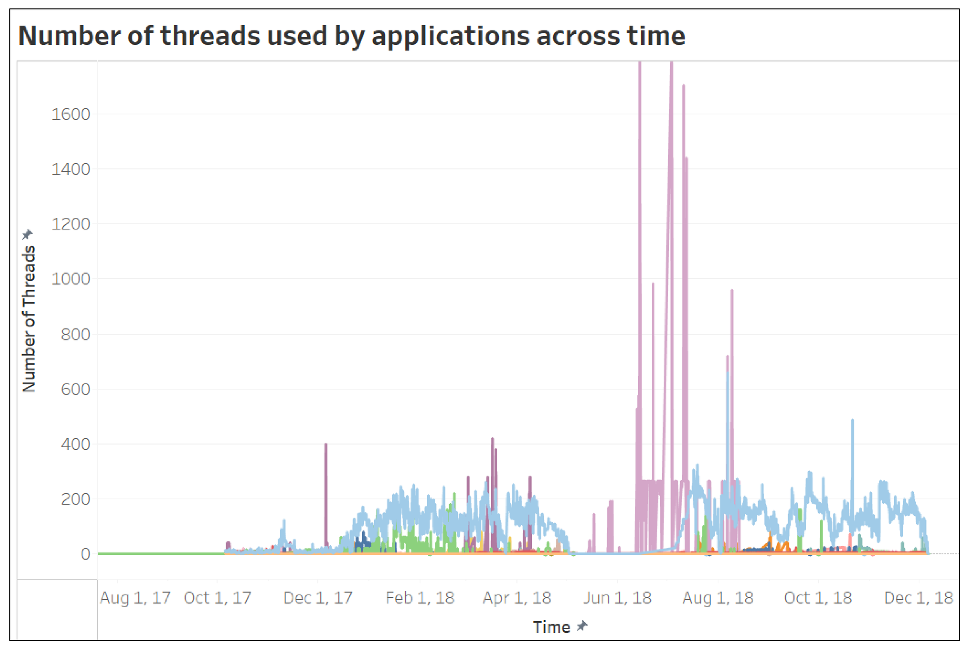

Figure 4. CPU cores(threads) used by applications across clusters

- The plot below shows that users of application that occupy the top three spots (vasp, gems, lmp), gives the following benefits in terms of dollars.
- The blue bar shows the savings occurred when the entire workload of given application is run on Rice and red bar similarly represents for Halstead. We see here that savings of Halstead is 5 times that of Rice.

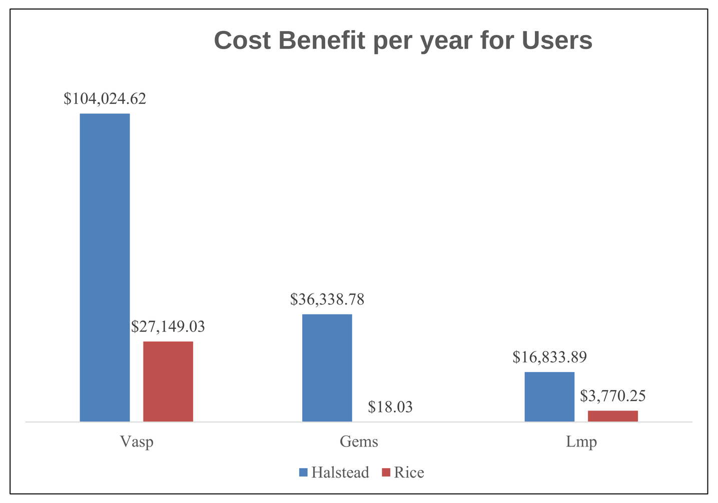

Figure 5. Cost benefit with shifting workloads

- The plot below indicates only the popular and adaptable clusters with ongoing campus research activities.
- For instance rice has most number of distinct users, and distinct application on its platform. It is followed by Halstead.

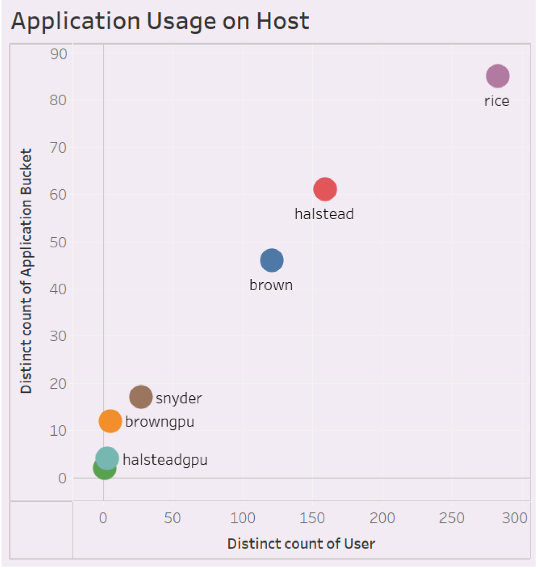

Figure 6. Cluster usage by application and user

Conclusions

- Millions of dollars are spent over building the supercomputer clusters and our analysis of server logs can help in running an application on the cluster that provides the cheapest rates for the user and **maximum utilization** for the Infrastructure provider
- Building compute resources that are similar to Rice , Halstead will ensure that utilization is maximum while ensuring the low cost of usage.
- It will eventually lead to considerable **cost cutting per cluster** which can also reflect in the selling price and eventually raise the demand for supercomputers.

Reference

1. C. D. Sudheer, A. Srinivasan, "Optimization of the hop-byte metric for effective topology aware mapping", High Performance Computing (HiPC) 2012.
2. 2. Todd Gamblin et al , "Mapping applications with collectives over sub communicators on torus networks", HiPC Networking Storage and Analysis (SC) 2012 Conference