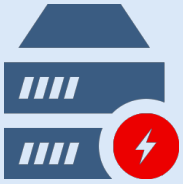
Analyze IT

Understand
Advise
Optimize



Overloaded Clusters ?

HPC requires capacity and performance…

But are there cheapest ways than a hardware upgrade to improve performance?

Unoptimized Scheduling?

Scheduling policies are always compromises…

But are there ways to reduce job wait-time or means to prevent they'd be killed before being completed?

Happy Users?

Probably not the ones who speak the most…

But can we provide them with more autonomy and improve the quality of IT service?

Analyze-IT packages our combined expertise in data analysis and HPC job schedulers to provide you instantly with key performance indicators you need. Based on your own data, we deliver a customized report containing advices for optimizing your HPC infrastructure.

84
Cluster optimization index

97
Cluster Load

70
Scheduler Efficiency

72
Job Efficiency

Analyze IT

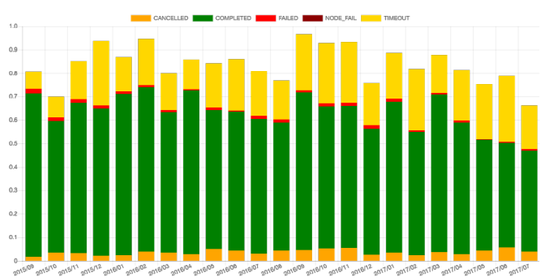
Understand
Advise
Optimize

The **job scheduler** is the central point of your HPC infrastructure. Besides allocating jobs on the available resources, it also keeps track of those allocations and the jobs performance, storing the collected data in log files. Such logs contain valuable information which is processed by us in order to deliver indicators and **insights about your HPC infrastructure**.

Overall Cluster Usage

You paid for all those nodes, but how well are they used? **This is revealed by assessing the following metrics during the analysis of your logs**: the ratio of jobs per completion status; the resources consumption (number of cpu-hours) per job status; cluster load (how well the job scheduler maximizes cluster resource usage); nodes load (how well the jobs are spread on all the nodes of the cluster); job interarrival statistics (measures cluster throughput and efficiency); job slowdown ratio (the impact of cluster load on result retrieval times)…

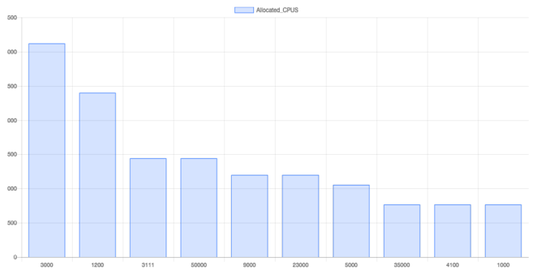
Additionally, we highlight **potential waste of resources** (e.g., power, compute, memory), **problematic user behavior** (e.g., wrong submission parameters), and **job scheduler configuration vs. usage unsuitability**.



Consumed resources repartition per job status each month

Specific Cluster Usage

How well your cluster is used, and by whom? We assess the cluster utilization by each user, group of users, and account: the number of jobs, the associated resources consumption (number of allocated cores and allocated memory), and jobs runtimes...



The maximum number of allocated CPUs by group of users

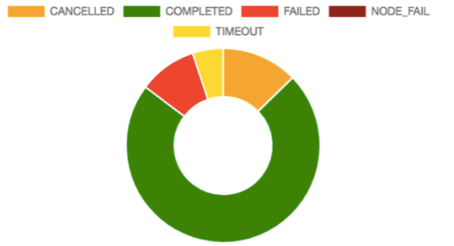
Our analysis also tries to **detect patterns** in job names to see which applications are the most widely used, and which ones use the cluster the most. This can be done ***globally on the whole cluster***, or can be **tailored to a particular interest** (e.g., analyze the behavior of a laboratory, select only those jobs that use more than 2048 cores).

Analyze IT

Understand
Advise
Optimize

Quality of Job Submission

Jobs are waiting for too long? This might be a symptom of an inefficient choice of submission parameters. There is a **set of actions we can recommend** by analyzing the logs: update scheduling policies, modify user priorities, reorganize queues, define default walltimes and propose template scripts for your applications.

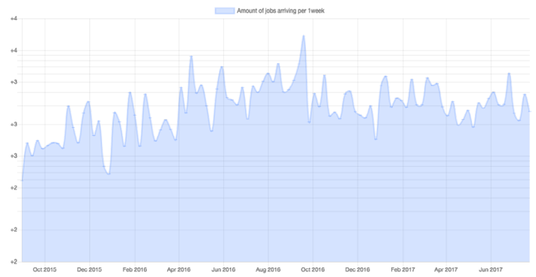


Percentage of jobs per job status

Infrastructure Upgrade

You intend to upgrade or replace your cluster? Our analysis can also help you define **your future computing requirements** (resources), either on premises or in the cloud. We will provide you with solid information to back up your decisions. Knowing how to optimize your cluster enables you to better understand your infrastructure and plan accordingly.

The above metrics and analyses are then gathered in a detailed report providing an overall score for your cluster and improvement directions. Moreover, the report helps you understand if the cluster is “correctly shared” between users/groups, and can also serve as a basis for periodic reporting towards your end-users or stakeholders. Based on that, we can suggest you a number of corrective actions such as user education or training, a specific configuration of your job scheduler (adapt its behavior to measured usage), or recommend decision-making tools to help your end-users select the best job-submission parameters. As a result, there are **several potential benefits to be gained**, e.g., increased cluster efficiency, cost savings, and the improvement of end-user satisfaction.



Amount of jobs arriving per 1week

Cycle of an analysis

1. **Collection of Anonymized Data**
2. **Statistical Data Analysis**
3. **Raw-Data Web Site Generation**
4. **Final Report Presentation**

Legal Mentions

© Copyright UCit 2017
UCit
5 rue des cigales
34470 Perols - France

Produced in France
November 2017

This document is current as of the initial date of publication and may be changed by UCit at any time. The performance data discussed herein is presented as derived under specific operating conditions. Actual results may vary. It is the user's responsibility to evaluate and verify the operation of any other products or programs with UCit products and programs.The information in this document is provided « as is » without any warranty, express or implied. UCit products are warranted according to the terms and conditions of the agreements under which they are provided.Statements regarding UCit future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Please Recycle

CC BY-NC-ND UCit 2017