Analyze-IT User Guide

Version 1.0

UCit
2017

Summary

This document is the complete documentation of the product named Analyze-IT released and published by UCit.

It fully presents the software, it describes the key features, the installation requirements and the deployment procedure, and it explains how to use it.

UCit
2017

Copyright

Copyright © 2015-2018, UCit.
All rights reserved.

We Would Like to Hear from You

You can help us make this document better by telling us what you think of the content, organization, and usefulness of the information. If you find an error or just want to make a suggestion for improving this document, please send feedback to UCit Support.
Although the information in this document has been carefully reviewed, UCit does not warrant it to be free of errors or omissions. UCit reserves the right to make corrections, updates, revisions, or changes to the information in this document.

UNLESS OTHERWISE EXPRESSLY STATED BY UCit, THE PROGRAM DESCRIBED IN THIS DOCUMENT IS PROVIDED "AS IS" AND WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. IN NO EVENT WILL UCit BE LIABLE TO ANYONE FOR SPECIAL, COLLATERAL, INCIDENTAL, OR CONSEQUENTIAL DAMAGES, INCLUDING WITHOUT LIMITATION ANY LOST PROFITS, DATA, OR SAVINGS, ARISING OUT OF THE USE OF OR INABILITY TO USE THIS PROGRAM.

Document Redistribution and Translation

This document is protected by copyright and you may not redistribute or translate it into another language, in part or in whole, without the express written permission of UCit.

Trademarks

Linux is the registered trademark of Linus Torvalds in the United States and other countries.

Firefox® and Mozilla® are trademarks or registered trademarks of the Mozilla Foundation in the United States and/or other countries.

Apple®, Mac®, Mac® OS X® and Apple® Safari® are trademarks or registered trademarks of Apple, Inc. in the United States and other countries.

Altair® PBS Professional® is a trademark of Altair Engineering, Inc.

SLURM™ is a trademark of SchedMD LLC.

Google™ and Chrome™ are trademarks of Google Inc.

Red Hat® is a trademark of Red Hat, Inc.

Sun® and JavaScript® are registered trademarks of Oracle and/or its affiliates.

Learn about Analyze-IT and UCit products

World Wide Web page

You can find the latest information about UCit Analyze-IT on its web site
https://www.ucit.fr/analyze-it/
For more information about other UCit products and about the professional services provided by UCit you can refer to the company's web site https://www.ucit.fr/

UCit Analyze-IT Download

The latest Analyze-IT downloadable package and documentation are available at:
https://www.ucit.fr/download-products

UCit Support Contacts

Use one of the following to contact UCit technical support.

- Email: helpdesk@ucit.fr
- Helpdesk portal: https://helpdesk.ucit.fr/

UCit
2017

Table of Contents

Copyright 3
We Would Like to Hear from You 3
Document Redistribution and Translation 3
Trademarks 4
Learn about Analyze-IT and UCit products 4
  World Wide Web page 4
  UCit Analyze-IT Download 4
  UCit Support Contacts 4
Introduction 7
How it works 8
Key Features 8
Requirements 10
  Hardware 10
  System 10
  Job scheduler 10
  Web browser 11
License 11
Installation 12
Extract job scheduler logs 13
Analyses description 14
  Global information 14
  Consumers 14
  Concurrent users 15
  Job resubmission 15
  Cluster load 15
  Nodes load 15
  Jobs throughput 15
Filtering input data 16
  Specifying the period of time to consider for the analysis 16
  Filtering data based on the value of certain fields 16
  Specifying cluster information 16
Job scheduler parser 16
Customization 17
  Report rendering options 17
  Changing colors of graphs 18

UCit
2017

Example of usage 18

UCit
2017

Introduction

The job scheduler is the central point for your HPC infrastructure. It dispatches and monitors the jobs on the available resources, keeping track of the jobs allocation and their state. The generated data are stored in the job scheduler's logs, from which valuable information can be extracted in order to understand the cluster behavior.

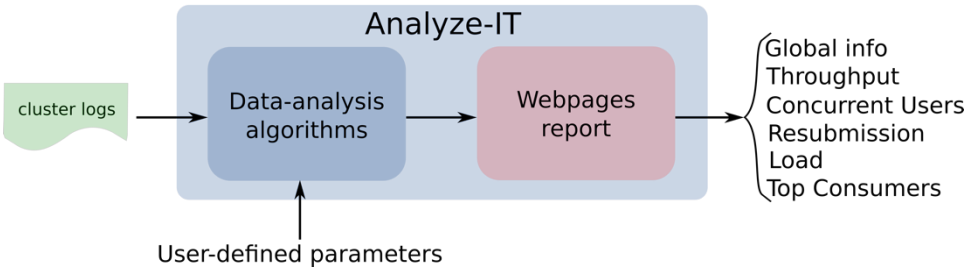
For example, the analysis of the logs makes it possible to identify practices and historical events that influence the overall cluster performance: which users are requesting more resources; which kind of jobs wait more time in the queue before starting; resources consumption per job state, and so on. Understanding the historical behavior of the cluster is crucial to choose the proper actions towards increasing production and profitability.

That is why we have developed **Analyze-IT**. By processing the job-submission historical data in the logs, Analyze-IT creates detailed analyses of the behavior of the jobs on your cluster. Moreover, it delivers a report containing indicators and **insights about your HPC infrastructure**.

UCit
2017

How it works

The first step is to extract the job submission logs of the client's cluster. This is done by the client himself, who runs the extraction scripts provided by UCit. This process generates three files with different extensions: `.jobs` (with information about the cluster jobs), `.partitions` (with information about the cluster partitions), and `.nodes` (information about the cluster nodes). They are automatically saved according to the following name format: `<cluster_name>_YYYY-MM-DDTHH:MM:SS.<extension>`.
Currently, Analyze-IT uses only the jobs file `<cluster_name>_YYYY-MM-DDTHH:MM:SS.jobs` to carry out the analysis. As depicted in Fig. 1, that file (represented by the cluster logs block) is fed to the data-analysis algorithms of Analyze-IT, which calculate a set of metrics defined by the user via input parameters. At the end of the data processing, a webpage report is generated: it shows the calculated metrics via tables and charts, providing the user with key information about the current state of the cluster and how it has been behaving over time.



Among the general metrics provided in the report are: global information about the cluster, throughput, concurrent users, resubmission, load, top consumers. Each of those has its own page in the report, where several other submetrics are provided.

Key Features

Analyze-IT provides two groups of metrics:

- Overview of the cluster usage: presented on the first two pages of the report, they provide quick information about the state of the cluster and their jobs.
- Detailed metrics: distributed in the remaining pages of the report. Among them are job-submission quality, cluster/nodes load, top 10 consumers, etc.

Below it is shown a list of all the metrics that can be presented in the report pages. Later in this document, it will be explained how the user can switch on/off the calculation & displaying of most of the following metrics:

- Home
  - General information
  - Cluster optimization index
- Global info
  - Percentage of jobs per job status
  - Percentage of core-hours per job status
  - Percentage of jobs per job status each month
  - Consumed resources repartition per job status each month
- Throughput
  - Inter-arrival
  - Slowdown
  - Amount of jobs arriving per 1 [day, week, month, year]
- Concurrent users

UCit
2017

    - Concurrent users who have submitted at least 1 job per 1hour
    - Concurrent users who have at least 1 job waiting or running per 1hour
    - Concurrent users who have submitted at least 1 job per 1day
    - Concurrent users who have at least 1 job waiting or running per 1day
    - Concurrent users who have submitted at least 1 job per 1month
    - Concurrent users who have at least 1 job waiting or running per 1month
- Resubmission
    - CANCELLED jobs that were resubmitted
    - TIMEOUT jobs that were resubmitted
- Load
    - Cluster load (cores allocated to jobs)
    - Number of jobs per node
- Top consumers by Account, GID, JobName, Partition, UID

UCit
2017

Requirements

Hardware

Analyze-IT should be installed on a virtual machine or physical server separated from the main cluster frontend node. Minimal hardware requirements are:

- CPU: 4 cores
- Memory: 16GB
- Disk space: At least 4GB of free disk space (for tool-generated files and installation files)

Note that the strongest requirement is the RAM: the more jobs you have in your logs, the more RAM you need.

System

- Operating Systems: RHEL 6/7
- A user account, different from the root account to run Analyze-IT (e.g., `aituser`)

Job scheduler

The following job schedulers are supported:

Name	Version	Notes
SLURM™	14.11.6 or later	SLURM™ binaries must be installed on the server where you will run the log extraction script, and in the PATH. The sacct command must be usable by the user (e.g., `aituser`) who runs the script. SLURM™ SlurmDBD server must be reachable from the server where you will run the log extraction script. Accounting must be turned on with SlurmDBD: `AccountingStorageType` must be set to `accounting_storage/slurmdbd`, `JobAcctGatherType` must be set to `jobacct_gather/linux`[1] (see https://slurm.schedmd.com/accounting.html for more details)
Altair® PBS Professional®	13.0 or later	The PBS Professional® logging directory must be readable from the server where you will run the log extraction script (default directory is `<PBS_HOME>/server_priv/accounting/`).
Torque	5.0 or later	The Torque logging directory must be readable from the server where you will run the log extraction script (e.g., `<TORQUEROOT>/server_priv/accounting/`).
OAR	2.5.7 or later	The OAR database must be accessible from the server where you will run the log extraction script.

[1] `jobacct_gather/cgroup` does not gather correctly memory usage in SLURM™ version lower than 17.02

UCit
2017

Web browser

Analyze-IT produces HTML which can be viewed with most popular browsers. Generally speaking, Analyze -IT supports the latest versions of each major platform's (Linux®, Apple® Mac® OS X®, and Microsoft® Windows®) default browsers (Google™ Chrome™, Mozilla Firefox®, Apple® Safari®, Microsoft® Edge®).
JavaScript® must be enabled on browsers.

License

You need a valid license to install and run Analyze -IT. If you do not have one yet, please contact helpdesk@ucit.fr or your Analyze -IT reseller.

UCit
2017

Installation

Analyze -IT is installed via the execution of the file `analyzeit-xx_rhel-yy.run` (with `xx` the version of Analyze -IT, and `yy` the version of RHEL)

1. Save the `analyzeit-xx_rhel-yy.run` file in a folder of your choice, and execute it:
   `./analyzeit-xx_rhel-yy.run`
2. The files will be uncompressed, and the following will be printed:
   ```
   ###############################################################################
   #                  Welcome to Analyze-IT installer
   ###############################################################################

   This installer will guide you during the installation of Analyze-IT.
   ###############################################################################
   Press "Enter" to continue or press "q" to quit:
   ```
   Press "Enter" to continue.
3. The license agreement is shown. To continue with the installation, you should type "`accept`", otherwise type "`quit`" to decline (the latter will terminate the installation).
4. Choose the installation path (in this example the installation path is "`/home/aituser/analyzeit`") and press "Enter":
   ```
   Enter the path where Analyze-IT must be installed.
   ###############################################################################
   Installation directory:
   /home/aituser/analyzeit
   ```
5. The files will be installed in the chosen folder:
   ```
   Installation of Analyze-IT is now complete
   Installation directory: /home/aituser/analyzeit
   Log file: /tmp/Analyze-IT-install-1.0-2018-01-22_14:23:29.log
   ###############################################################################
   Press "Enter" to exit
   ```
6. To check that the files exist, navigate to the installation folder and inspect it with "`ls -lth`".

UCit
2017

Extract job scheduler logs

Analyze-IT provides a set of scripts to extract the historical data from the job scheduler. You can find them in `<INSTALLATION_PATH>/jobSchedulers/extractXXXXData.sh`, with `XXXX` the name of the job scheduler:

- `extractSlurmData.sh` for SLURM™. It relies only on `sacct` and `scontrol` to extract raw data from SLURM.
- `extractTorqueData.sh` for Torque, Altair® PBS Professional®. It only extracts raw data from accounting files, and from pbsnodes and qstat.
- `extractOARData.sh` for OAR. It uses `psql` (PostgreSQL client) to extract data from OAR database.

Each script creates 3 text files in the current directory:

- **historical data on jobs**: `HOSTNAME_DATE.jobs`
  - From a given date up to now
  - Retrieve all data gathered by the job scheduler (job id, requested/obtained resources, node list, submit/start/run times…)
  - For confidentiality reasons, the output is anonymized: by default, usernames and groups are not retrieved, only UID and GID are
- current **list of nodes** and their description: `HOSTNAME_DATE.nodes`
  - Node name
  - Number of cores, number of cores per socket, memory
  - Specific resources (features described in the job scheduler)
- current **list of partitions/queues** and their description: `HOSTNAME_DATE.partitions`

The execution of the script is very lightweight. No processing of the data is done.
Each script comes with its own online help, run `./extractXXXXData.sh -h` to see this help.

For Analyze-IT, only the first file is currently relevant: `HOSTNAME_DATE.jobs`.

UCit
2017

Analyses description

Navigate to the folder where Analyze-IT was installed
`cd <INSTALLATION_PATH>`

Analyze-IT is run via the execution of the file `./Analyze-IT`. If you call it with the help option `-h` or `--help`, it lists all possible analyses
`./Analyze-IT --help`

By default, Analyze-IT runs all the available analyses it offers, but you can also specify which ones you would like to see in your final report (with the `-a` option). Keep in mind that some metrics shown in the final report (e.g., the "Cluster Optimization Index") rely on several of these analysis. Thus, if you run only a subset of them, the index value will truly not reflect your cluster's behavior.

We describe below the analysis you have access to, with the associated options you can use to customize the analysis to your needs. On top of all these options, you also have two general options for the running Analyze-IT:

- `-o OUTPUT, --output OUTPUT` : Output directory for the report (default: report)
- `-mnp MAXNBPROCESSES, --maxnbprocesses MAXNBPROCESSES`: Maximum number of processes to launch for the analysis (default: None)

Global information

This analysis is accessible through the `-a ginfo` option.

Associated options:

- `-tc TOTALNBCORES, --totalnbcores TOTALNBCORES`: Total number of cores on the cluster per period. Format is the following: a string "`nbcores[,YYYY-MM-DD:nbcores...]`" (the date represents the start date at which point the number of cores change and take a new value). (default: None)

Consumers

This analysis is accessible through the `-a group` option.

Associated options:

- `-g {JobName,UID,GID,Partition,Account} [{JobName,UID,GID,Partition,Account} ...], --group {JobName,UID,GID,Partition,Account} [{JobName,UID,GID,Partition,Account} ...]`: Name of the columns to group the data with to analyze consumption. (default: `'JobName', 'UID', 'GID', 'Partition', 'Account'`)
- `-gj, --groupjobnames`: Group similar job names using fuzzy string matching: similar job names will be replaced by a single job name that is considered as the "main" name. For example, if Analyze-IT has detected that the following job names are similar ["job1", "job2", "job3"…] and that "job1" is the most significant name, all jobs in the group will have their job name replaced by "job1".
- `-gds GROUPDETAILNBSAMPLES, --groupdetailnbsamples GROUPDETAILNBSAMPLES`: Maximum number of points (jobs) on the graphs on the details analysis pages. If None, all jobs are displayed. (default: None)

UCit
2017

- `-t number, --top number`: Number of jobs to be highlighted in each of the consumer analysis (default: 10)

Concurrent users

This analysis is accessible through the `-a concusers` option.

Associated options:

- `-cr {10min,1hour,1day,1week,1month,1year} [{10min,1hour,1day,1week,1month,1year} ...], --concusersres {10min,1hour,1day,1week,1month,1year} [{10min,1hour,1day,1week,1month,1year} ...]`: Time resolution for analysis of concurrent users. (default: `'1hour', '1day', '1month'`)

Job resubmission

This analysis is accessible through the `-a resubmit` option.

Associated options:

- `-msr MAXSECONDSRESUBMIT, --maxsecondsresubmit MAXSECONDSRESUBMIT`: Maximum of seconds allowed between two jobs to consider a job resubmission (default: 172800, i.e., 48 hours)

Cluster load

This analysis is accessible through the `-a clusterload` option.

Associated options:

- `-tc TOTALNBCORES, --totalnbcores TOTALNBCORES`: Total number of cores on the cluster per period. Format is the following: a string "`nbcores[,YYYY-MM-DD:nbcores...]`" (the date represents the start date at which point the number of cores change and take a new value). (default: None)

Nodes load

This analysis is accessible through the `-a nodesload` option.

Jobs throughput

This analysis is accessible through the `-a throughput` option.

Associated options:

- `-tr {10min,1hour,1day,1week,1month,1year} [{10min,1hour,1day,1week,1month,1year} ...], --throughputres {10min,1hour,1day,1week,1month,1year} [{10min,1hour,1day,1week,1month,1year} ...]`: Time resolution for analysis of job throughput. (default: `'10min', '1hour', '1day', '1week', '1month', '1year'`)

UCit
2017

Filtering input data

Specifying the period of time to consider for the analysis

The following options allow you to consider only jobs in a given time frame, both can be used at the same time:

- `-s %Y-%m-%d %H:%M:%S, --starttime %Y-%m-%d %H:%M:%S`: Consider jobs that where submitted after this date (default: 1970-01-01 00:00:00)
- `-e %Y-%m-%d %H:%M:%S, --endtime %Y-%m-%d %H:%M:%S`: Consider jobs that where submitted before this date (default is the current date)

Filtering data based on the value of certain fields

When analyzing your job scheduler data, you may want to analyze only a subset of the data depending on the value of certain fields, for example you may want to analyze the behavior of certain groups of users, and thus only keep jobs submitted by a list of GIDs for the analysis and drop all the other jobs. Analyze-IT allows you to filter the input data to really target what you want to analyze.

`-fc FILTERCOLUMN, --filtercolumn FILTERCOLUMN`: Filter columns to keep only certain jobs.

Format is the following: `columnName:value` or `columName:value1,value2,...`

Multiple filters can be provided by separating them with ";" Example:
`UID:1234;JobName:Abaqus,Fluent...`

Available columns (depends on the input data) are: `Account, Allocated_CPUS, Allocated_Nodes, _Memory, Cluster, Comment, Eligible, End, Exit_Code, GID, Group, JobID, JobName, MaxRSS, NodeList, Partition, QOS, Requested_CPUS, _Nodes, Requested_Memory, Requested_Memory_Per_CPU, Start, State, Submit, , UID, User` (default: None)

Specifying cluster information

Sometimes the input data do not contain all the necessary information to really compute the actual values of some analysis. The following options allow you to set default values when they are missing in the input data (you will usually see 'nan' showing up in the result in the analysis when some values are missing):

- `-c CORESPERNODE, --corespernode CORESPERNODE`: Number of cores per node. Used to compute the number of cores used by a job when only the allocated number of nodes is known in the logs, and vice-versa. (default: 1)
- `-dq DEFAULTQUEUE, --defaultqueue DEFAULTQUEUE`: Default queue name to use when queue is absent in the input data. (default: None)

Job scheduler parser

Analyze-IT does not detect automatically which job scheduler you took your data from. You must specify the right parser Analyze-IT needs to use. The following options allow you to configure this:

- `-j {slurm,torque,oar,swf}, --jobscheduler {slurm,torque,oar,swf}`: Job scheduler parser to use (default: `slurm`). Use Torque if you have logs from Altair® PBS

UCit
2017

Professional®

- `-jp JSPARAMS, --jsparams JSPARAMS`: Job scheduler additional parameter. Format is "`key=value,key=value...`". (default: None).
  - OAR:
    - `dbhost` (mandatory): hostname of the PGSQL server hosting OAR database
    - `username` (mandatory): username to connect to the database
    - `password` (mandatory): password to connect to the database
    - `database` (mandatory): database name
  - Slurm:
    - `separator`: the only possible value is `'slurmdefault'`. Analyze-IT uses by default '@|@' as a field separator when retrieving slurm historical data. If your input file uses slurm's default separator ('|'), then you must specify `separator=slurmdefault`
    - `noalloc`: Gather details about each job step or not? If `noalloc` is True (default), then we gather the details for all steps of the jobs, allows to get the MaxRSS values properly.
  - Torque:
    - `acct_dir`: path to accounting directory (default=/var/spool/torque/server_priv/accounting/)
- `-rn, --realnames`: Use real user and group names if available instead of UID/GID. (default: False)
- `-sd, --savedata`: Save data once parsed into a `.pickle` file. Allows for faster further analyses. (default: False). The file will be saved in the same directory as the input file, with the same name, and a `.pickle` extension. You can then use directly provide the `.pickle` file to Analyze-IT as an input file.

Customization

Report rendering options

The following options allow you to customize the report (add a logo, change its name, modify the graphs…):

- `-n name, --name name`: Name of the analysis. By default, this will be the name of the input file. (default: None)
- `-l LOGO, --logo LOGO`: Path to PNG logo to display on the top of the pages (recommended size: height="80" width="160") (default: None)
- `-fm, --fullmenu`: Generate the full menu with all the links even if only a subset of the analysis is selected. Beware that the cluster optimization index might be wrong if you do not run a global analysis. (default: False)
- `-op OPACITY, --opacity OPACITY`: Opacity of RGB color for ploting. Value should be a real number in [0, 1] (default: 0.2)
- `-cls CLOADNBSAMPLES, --cloadnbsamples CLOADNBSAMPLES`: Maximum number of points (jobs) on the graphs on the cluster load analysis page. (default: 10000)
- `-ni, --noindex`: Do not generate index (useful if you only refresh some pages, and do not want to mess with the cluster optimization index computation). (default: True)

UCit
2017

Changing colors of graphs

Update the `rgb_colors.csv` file in the `template` directory.

Make sure that the following 'colors' are defined: `CANCELLED`, `COMPLETED`, `FAILED`, `NODE_FAIL`, `PREEMPTED`, `TIMEOUT`, `BOOT_FAIL`, `REQUEUED`, `Unused resources`.

Also make sure that you have at least 6 more colors (hence a total of 15 colors).

Colors are taken in the order they are defined in the file to generate the graphs.

Example of usage

Here it is shown an example of how to run Analyze-IT with its default settings, i.e., performing all the available analysis, and provide an overview of the resulting report.

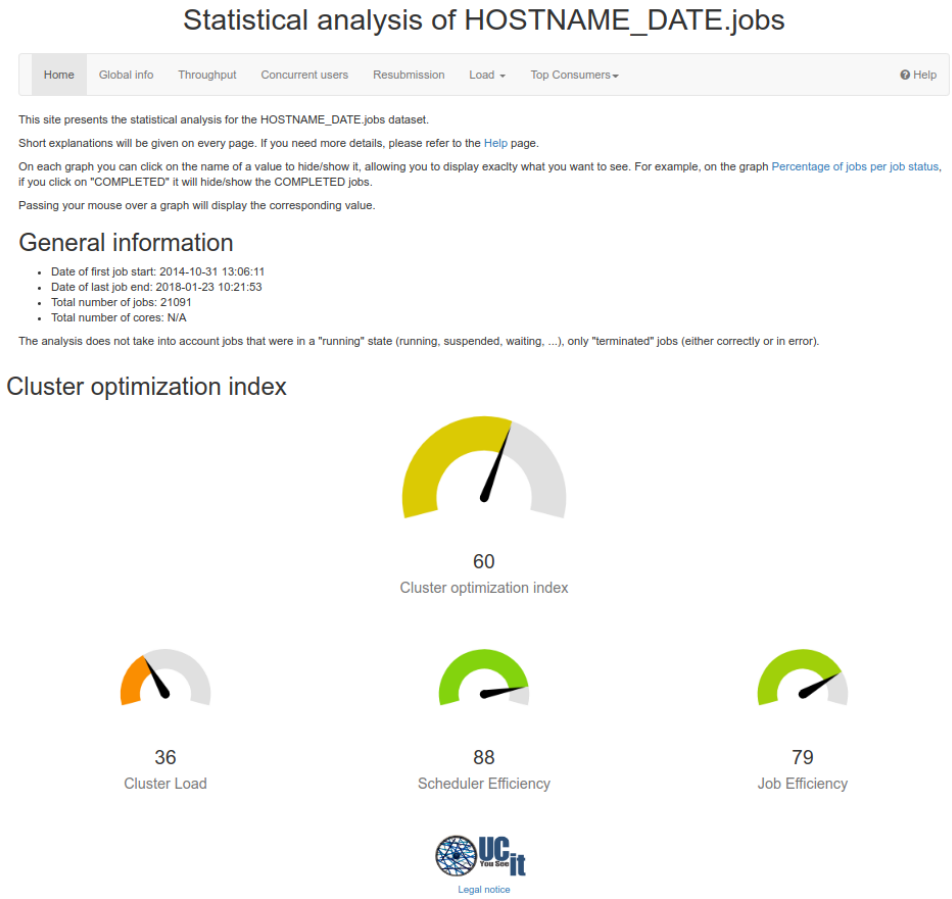
Navigate to the folder where Analyze-IT was installed
```
cd <INSTALLATION_PATH>
```

To start the analysis, type the following command (It is assumed that the job-scheduler logs are already extracted – check Section "Extract job scheduler logs")
```
./Analyze-IT  HOSTNAME_DATE.jobs
```

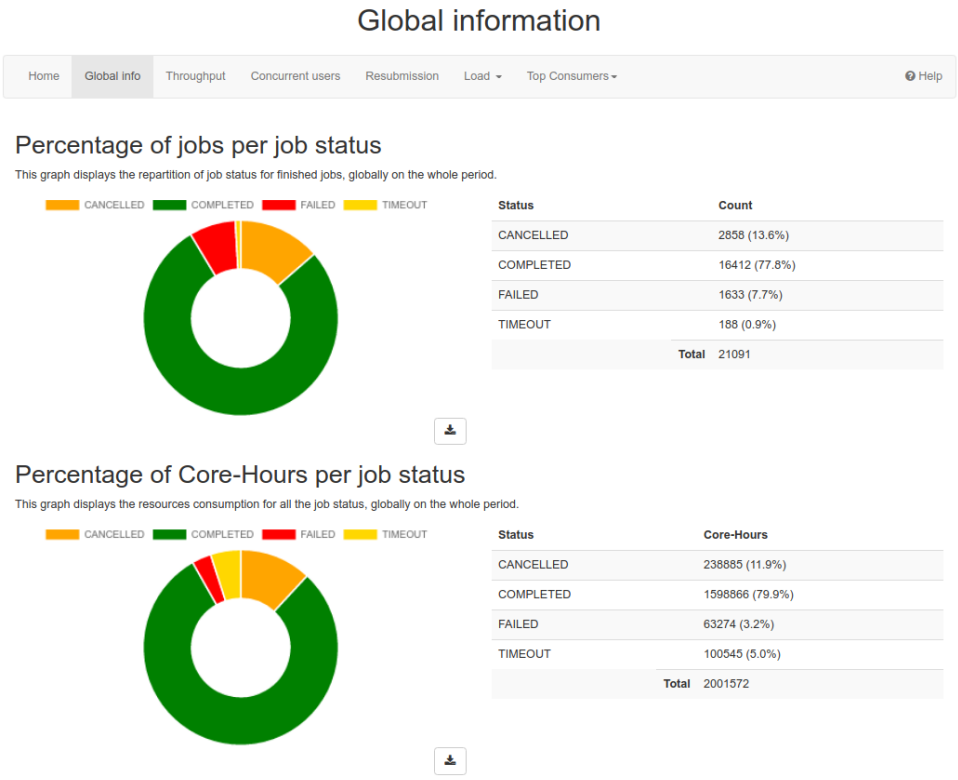
This will run all the available analysis, with their default values, for the file `HOSTNAME_DATE.jobs`. A web-based report will be generated and saved in the folder `report` (automatically created inside `<INSTALLATION_PATH>`). Go inside it and click on index.html, which will be opened in your default web browser.

Below is an example of the report homepage:

UCit
2017



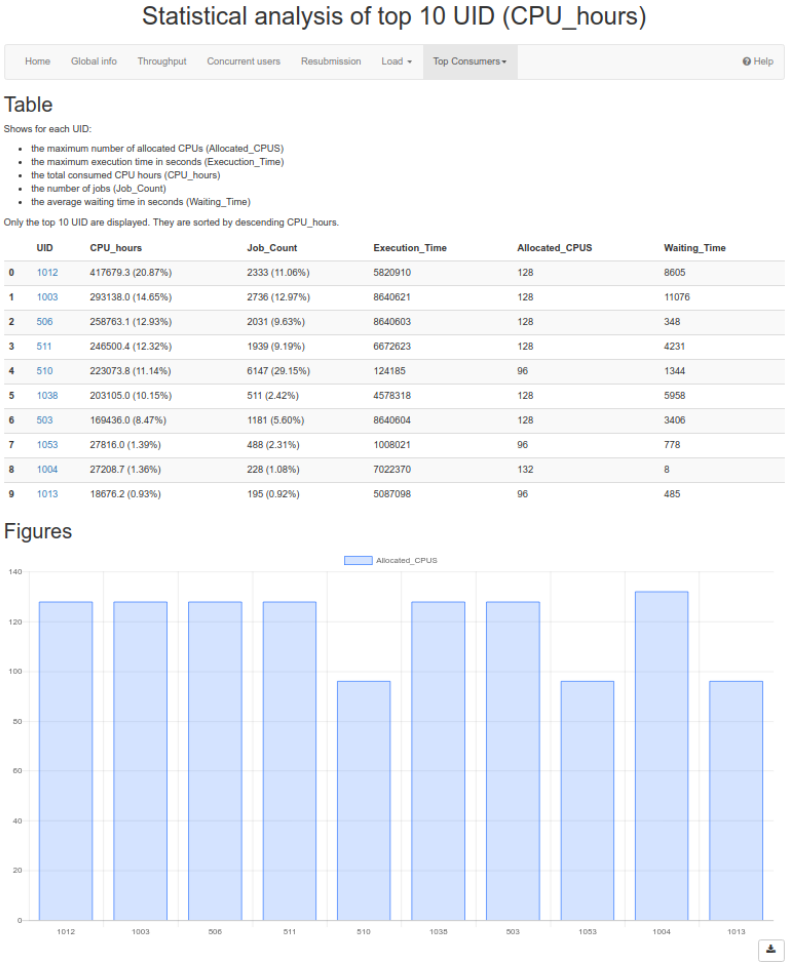
Notice that for each analysis type a navigation link is created. For example, a section of the "Global info" page is shown below, with metrics like the "percentage of jobs per job status" and "percentage of core-hours per job status".



Another example, the "Top Consumers" pages shows the top GIDs, JobNames, Partitions, and UIDs in

UCit
2017

terms of Allocated_CPUs, CPU_hours, Execution_Time, Job_Count, and Waiting_Time. The snapshot below lists the 10 UIDs that consume more CPU_hours.



UCit
2017