Predict-IT Administration Guide

Version 1.1

Copyright

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

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 Predict-IT and UCit products

World Wide Web page

You can find the latest information about UCit Predict-IT on its web site
https://www.ucit.fr/predict-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 Predict-IT Download

The latest Predict-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/

Table of Contents

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, when jobs are submitted on the cluster, some of them may not terminate within the specified time limit (walltime) and consequently no results are obtained. This way, to increase cluster profitability and production, it is necessary to enforce that the submitted computations will end up correctly.

That is why we have developed **Predict-IT**. By analyzing the job-submission historical data in the logs, Predict-IT creates a computational model of your cluster based on machine-learning techniques. Such model is then applied at the moment of job submission to **identify which jobs are likely to end-up in failure**; moreover, it also provides **recommendations for the walltime you need to specify to prevent job failure**, and the **amount of memory** (max RSS) a job will consume per node.

Configured specifically for your cluster, Predict-IT **steadily improves over time**: it adapts to your HPC environment by learning from your cluster logs and newly arrived jobs, each time becoming more and more accurate in its predictions. The more data you feed it, the higher the prediction accuracy can be.

How it works

Predict-IT **learns from the job scheduler logs** and **creates a model** of your cluster (HPC system) which will be used to perform predictions at the time of job submission. This is done in order to provide the user with predictions for the following jobs characteristics: the jobs finishing state (or simply **STATE**), the jobs maximum execution time (also called **EXECTIME** or **WALLTIME**), and the job maximum memory consumption, i.e., maximum resident set size (**maxRSS**) on a node.

The operation of Predict-IT is done in two phases. In the first one (**learning phase**), the algorithms embedded in Predict-IT devise a set of models for your cluster, based on the history logs of your system. Those models combined are able to forecast the correct STATE of a job, its WALLTIME, and its maxRSS up to a certain accuracy. In the second one (**prediction phase**), the trained models are queried by the user in order to obtain predictions for a particular job. Figure 1 depicts the basic flowchart for Predict-IT.

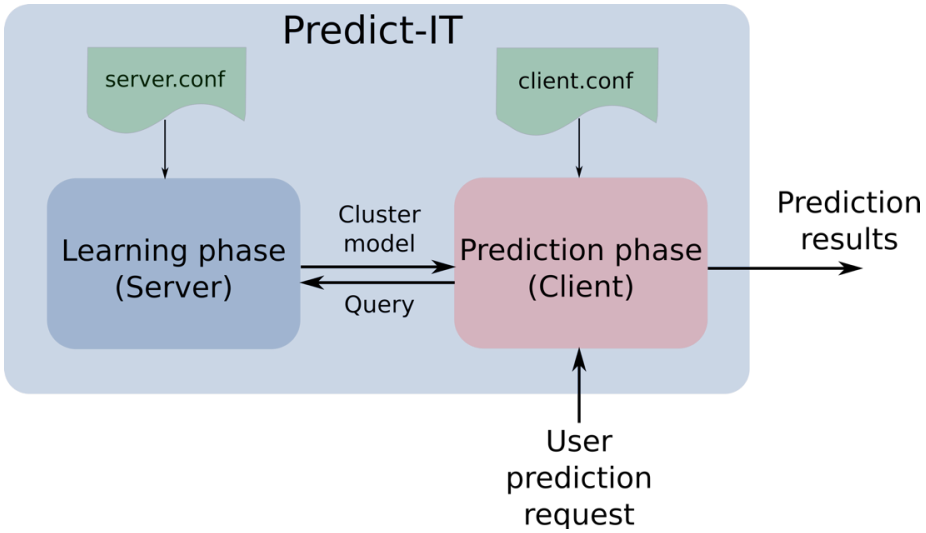

Figure 1: Overview of Predict-IT structure

At submission time, the user provides specific parameter values to the job scheduler. For example, one can set the number of requested CPUs, requested memory (REQMEM), partition etc. Additionally, each user has an identity (UID) and belongs to a specific group (GID) of the cluster. All those parameters are combined in the "user prediction request" (see Figure 1), which is then passed into Predict-IT to obtain predictions for that job (or group of jobs).
Predict-IT is composed by two components, namely the **server** and the **client**:

- Server: that's where the learning phase happens. This environment must be activated first, otherwise no cluster model will be created, and consequently no prediction will be possible.
- Client: after the learning phase is completed, the user can request predictions from the client side by passing the job submission parameters. The client queries the server which then uses the submission parameters together with the cluster model in order to forecast the STATE, WALLTIME, and MEMORY.

The configuration files `server.conf` and `client.conf` define parameters for the server side and the client side, respectively. Those parameters are user-customizable: their values depend on the characteristics of the dataset and the type of predictions the user wants to obtain.

Key Features

Predict-IT provides recommendations for the following jobs' characteristics:

- Job STATE: Detects the risk of a job finishing in timeout, i.e., reaching the walltime (maximum time allowed for running a job);
- Job EXECTIME (WALLTIME): Predicts the walltime that should be set by the user at the submission to prevent it resulting in timeout;
- Job MEMORY (maxRSS): Predicts the memory requirement that should be set by the user at the submission.

Predict-IT keeps learning over time: the more data, the higher the accuracy.

Requirements

Hardware

Predict-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 the server (e.g., pituser)

Job scheduler

The following job schedulers are supported:

Name	Version	Notes
SLURM™	14.11.6 or later	SLURM™ binaries must be installed on the Predict-IT Server host, and in the PATH. The sacct command must be usable by the Predict-IT user (e.g., pituser). SLURM™ SlurmDBD server must be reachable from the Predict-IT Server host. 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 Predict-IT Server Host (default directory is `<PBS_HOME>/server_priv/accounting/`).
Torque	5.0 or later	The Torque logging directory must be readable from the Predict-IT Server Host (e.g., `<TORQUEROOT>/server_priv/accounting/`).
OAR	2.5.7 or later	The OAR database must be accessible from the Predict-IT Server Host.

The job scheduler must be accessible from Predict-IT Server host. If not, you will only be able to train Predict-IT with an input file; in this case the prediction model cannot be frequently updated automatically (see option `inputFile` in section "Server configuration file ").

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

Web browser

Predict-IT produces HTML which can be viewed with most popular browsers. Generally speaking, Predict-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 Predict-IT. If you do not have one yet, please contact helpdesk@ucit.fr or your Predict-IT reseller.

Installation

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

1. Save the `predictit-xx_rhel-yy.run` file in a folder of your choice, and execute it:
   `./predictit-xx_rhel-yy.run`
2. The files will be uncompressed, and the following will be printed:

```
################################################################################
#                    Welcome to Predict-IT installer
################################################################################

This installer will guide you during the installation of Predict-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/pituser/predictit`") and press "Enter":

```
Enter the path where Predict-IT must be installed.
################################################################################
Installation directory:
/home/pituser/predictit
```

5. Then specify which user will be running the Predict-IT service

```
Enter the user who will be executing Predict-IT server.
################################################################################
User [pituser]:
```

6. Optionally, install the service files

```
Do you want to install and register init.d scripts? (you need to be root)
################################################################################
Install init.d scripts (y or n):
y
```

7. The files will be installed in the chosen folder:

```
Installation of Predict-IT is now complete
Installation directory: /home/pituser/predictit
Log file: /tmp/Predict-IT-install-1.0-2018-02-02_12:44:31.log
################################################################################
Press "Enter" to exit
```

You will also need to copy your license file into `<INSTALLATION_PATH>/license/license.lic`. You can also place it in another directory, in this case you will need to edit the `<INSTALLATION_PATH>/conf/server.conf` file and set the `licenseFile` parameter with the path to the license file.

To start the Predict-IT service, just run
`service predictit start` on RHEL 6
`systemctl start predictit` on RHEL 7

Usage

Starting the Server

Before starting to predict the status and walltime for the submitted jobs, Predict-IT needs to extract information from the cluster logs and build a computational model of your HPC system. This is done as follows:

1. Go inside the Predict-IT installation folder (`cd <INSTALLATION_PATH>`). There you will find the folder `conf/` which stores the configuration files for both the server and client sides. All the parameters that control how Predict-IT learns from your cluster are defined in the `server.conf` file. Some examples of parameters that might be edited are:
   - `host` (default `127.0.01`): the host the server listens on;
   - `port` (default `9999`): the port server listens on;
   - `logLevel` (default is INFO/`20`): controls the amount of information that is displayed in the log file;
   - `licenseFile`: path to license file;
   - `driver`: the job scheduler driver to use. Predict-IT can automatically extract job scheduler's historical data on a periodic basis if this parameter is specified (along with the
   - `inputFile` (optional): manually extract the cluster logs by executing the provided job-scheduler data extraction scripts provided with Predict-IT (see section "
   - Extract job scheduler logs").

   The complete list of server-side parameters is provided in the section "Server configuration file ".

2. Once finished editing `server.conf`, save it and start the server:
   a. Either run it manually by running the `predictit-server` binary:
      `./bin/predictit-server`
   b. Or start the service
      `service predictit start` on RHEL 6
      `systemctl start predictit` on RHEL 7

   This starts the learning task. Some metrics will be displayed in the terminal prompt and/or saved in the Predict-IT log files. The learning phase is finished when the following line is displayed in the logs:

```
======= End of test metrics =======

2018-01-10 20:18:54,969 : INFO : prediction.State : trainingTask._jobListener : Training job
 finished.
```

3. Now, access `http://host:port/metrics` (in case you are using the default values, `host=127.0.0.1` and `port=9999`) from your browser. This page displays the metrics for training and testing the computational model of the cluster (see Figure 2).

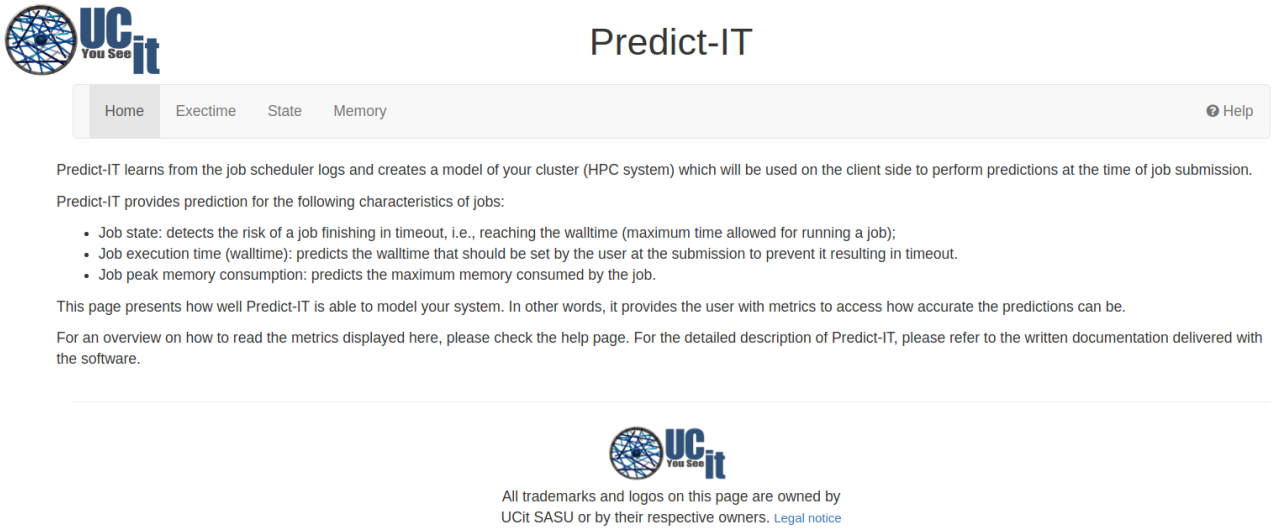

Figure 2. Metrics web page

4. There are different metrics for each target of interest (click on the links "State", "ExecutionTime", or "Memory"). Among them is the accuracy level that one should expect when requesting predictions for submitted jobs. See Figure 3 for an example of the State page.

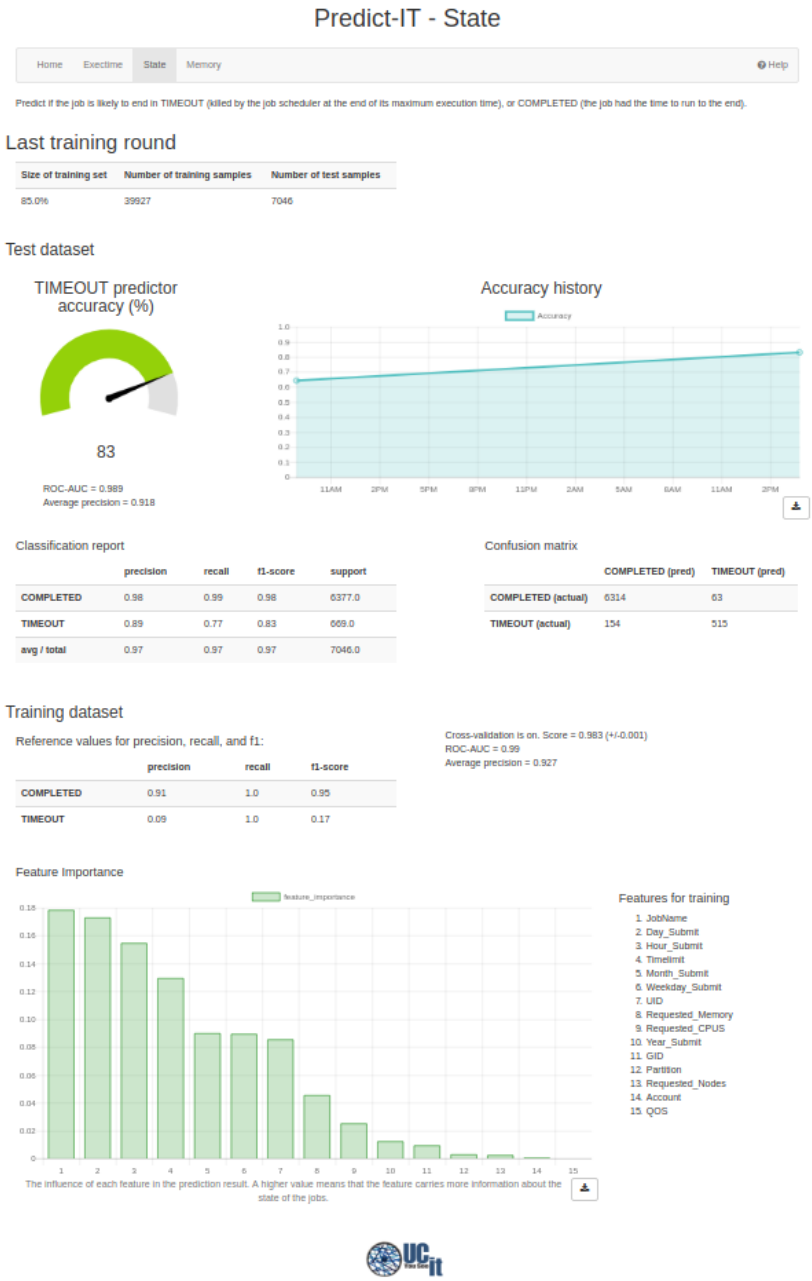

Figure 3. State metrics page example

For a complete description of the metrics page, including the meaning of each metric and how to interpret it, please refer to the help page accessible via the link "Help" on the upper right corner of the metric pages.

Client side

After the generation of a cluster model on the server side, the user can start requesting predictions for the state, maximum execution time (i.e. walltime), and memory of his jobs.

Command line client

The client comes with its own online help. To display it, just type `./client -h`

```
Predict-IT client

positional arguments:
  Command         Select the command to run
    runtime       Predict job runtime
    state         Predict job end state (Timeout or Completed)
    memory        Predict job peak memory consumption
    version       Retrieve server version

optional arguments:
  -h, --help      show this help message and exit
  -v, --version   Print version and exit
```

As aforementioned, there are currently three types of predictions (targets):

- runtime, which predicts the maximum execution time (i.e. walltime) for the job
- state, which predicts the job final state (TIMEOUT or COMPLETED)
- memory, which predicts the peak memory consumption

Each target has its own help menu accessible by typing `./client <target> -h`. For example, for target Memory:

```
usage: prediction.py memory [-h] [-di | -i json | -f inputFile]
                            [-j {slurm,torque,oar,swf,pickleFile}]
                            [-s %Y-%m-%d %H:%M:%S] [-e %Y-%m-%d %H:%M:%S]

optional arguments:
  -h, --help            show this help message and exit
  -di, --describe       Describe service
  -i json, --input json
                        JSON input string
  -f inputFile, --file inputFile
                        Input file (JSON file by default, or job scheduler
                        logs if used in conjunction with the -j option)
  -j {slurm,torque,oar,swf,pickleFile}, --jobscheduler {slurm,torque,oar,swf,pickleFile}
                        Job scheduler parser to use for the input file (-f
                        option)
  -s %Y-%m-%d %H:%M:%S, --starttime %Y-%m-%d %H:%M:%S
                        Only keep jobs that where submitted after this date
  -e %Y-%m-%d %H:%M:%S, --endtime %Y-%m-%d %H:%M:%S
                        Only keep jobs that where submitted before this date
```

By default, if you call the client for any target without any other argument, you will enter an interactive mode. In this mode, the client first queries the server for the relevant fields you have to specify, and then asks you to fill them in.
Note: All fields are optional, the server sets default values for each field if not set manually by the user.

The interactive client also automatically fills in some of the fields:

- your user ID (UID, taken from the current user)
- your group ID (GID, taken from the current user)
- the submission time (supposed to be “now”)

Here is an example for state prediction:

```
Please provide the following parameter specific to the job you want to predict (all parameters are optional):
Account (string): lab1
JobName (string): ASAP
Partition (string):
QOS (string):
Requested_CPUS (integer, number): 2
Requested_Memory (integer, bytes): 62914560000
Timelimit (integer, seconds): 3600

Predictions for jobs' state:
+ Job 1 (ASAP) predicted state is COMPLETED (confidence 70%)
  Job details: {'Account': 'lab1', 'GID': '1001', 'JobName': 'ASAP', 'Requested_CPUS': '2', 'Requested_Memory': '62914560000',
'Timelimit': '3600', 'UID': '1001', 'confidence': 0.7, 'prediction': 'COMPLETED'}
```

Note that in the interactive mode you can only request the prediction for a single job, while with the `-i` and `-f` arguments you could query for multiple jobs: if you do not use the interactive mode, they should point respectively to the JSON input (provided directly in the command line) or the JSON file (which stores the JSON command).

Whether via command line or file, the JSON input should contain the fields (or a subset of them) described when typing `./client <target> -di`.
Example for runtime:

```
Description of maximum runtime prediction:
Predict the WALLTIME (maximum execution time) of jobs. Jobs are classified by bins (classes) of execution time, e.g., 1h < JOB
< 2h... 2h < JOB < 3h... 3h < JOB < 4h... and so on. This way, the prediction is not the exact execution time, but its upper an
d lower bounds.
This is the list of columns you can/should have in your input data:
 - Account
 - Day_Submit
 - GID
 - Hour_Submit
 - JobName
 - Month_Submit
 - Partition
 - QOS
 - Requested_CPUS
 - Requested_Memory
 - Timelimit
 - UID
 - Weekday_Submit
 - Year_Submit
```

Here is an example of a JSON input describing the submission values for two jobs:

```
{
    "jobInput": {
        "Account": [
            "default",
            "default"
        ],
        "GID": [
            "1200",
            "550"
        ],
        "JobName": [
            "rae_cruise",
            "acwssi1"
        ],
        "Partition": [
            "debug",
            "prod"
        ],
```

```
        "QOS": [
            "1",
            "1"
        ],
        "ReqCPUS": [
            "24",
            "120"
        ],
        "ReqNodes": [
            "1",
            "1"
        ],
        "Submit": [
            "2018-08-28 17:15:59",
            "2015-08-28 17:15:24"
        ],
        "Timelimit": [
            "7200",
            "21600"
        ],
        "UID": [
            "1201",
            "550"
        ]
    }
}
```


The snippet above can be read like so: taking for instance the first job ("`rae_cruise`"), it was submitted by UID (user ID) 1201, who belongs to GID (group ID) 1200, and requests 1 CPU, with a time limit (walltime) of 7200 seconds (2 hours).

The JSON above is stored in a file which will be used as the input to the prediction client (for the purposes of this documentation, let's call it `client.inputfile.json`). Two examples of predictions are provided below, one for STATE and the one for WALLTIME. The prediction for both targets is requested to the server via the following command:

```
./client <target> -f  client.inputfile.json
```

Calling that command with target 'state':

```
./client state -f  client.inputfile.json
```

The command returns the following results:

```
Predictions for jobs' state:
+ Job 1 (rae_cruise) predicted state is TIMEOUT (confidence 64.2%)
Job details: {'Account': 'default', 'Day_Submit': '28', 'GID': '1200', 'Hour_Submit': '17',
'JobName': 'rae_cruise', 'Minute_Submit': '15', 'Month_Submit': '8', 'Partition': 'debug', '
QOS': '1', 'ReqCPUS': '24', 'ReqNodes': '1', 'Second_Submit': '59', 'Submit': '2018-08-28 17
:15:59', 'Timelimit': '7200', 'UID': '1201', 'Weekday_Submit': '1', 'Year_Submit': '2018', '
confidence': 0.6424019607843137, 'prediction': 'TIMEOUT'}
+ Job 2 (acwssi1) predicted state is TIMEOUT (confidence 71.2%)
Job details: {'Account': 'default', 'Day_Submit': '28', 'GID': '550', 'Hour_Submit': '17', '
JobName': 'acwssi1', 'Minute_Submit': '15', 'Month_Submit': '8', 'Partition': 'prod', 'QOS':
 '1', 'ReqCPUS': '120', 'ReqNodes': '1', 'Second_Submit': '24', 'Submit': '2015-08-28 17:15:
24', 'Timelimit': '3600', 'UID': '550', 'Weekday_Submit': '4', 'Year_Submit': '2015', 'confi
dence': 0.7124019607843137, 'prediction': 'TIMEOUT'}
```

The first set of parameters (a dictionary) depicts the values for the first job in the JSON snippet. Similarly, the parameters for the second job are shown in the second dictionary. Notice that both dictionaries have extra parameters which were not previously selected by the user (e.g., `'Hour Submit', 'Month Submit', 'Requested_Memory_Per_CPU'` etc…). They are

automatically inserted by Predict-IT in order to have all the necessary information to query the model.

More importantly, notice the last parameter called `'prediction'`: *this shows the predicted state for the job if those parameters are selected*. For both jobs it is predicted that it will result in TIMEOUT.

This result is also reported on the server side by only displaying the amount of jobs in each prediction category:

```
127.0.0.1 - - [02/Feb/2018 22:29:34] "POST /1.0/state HTTP/1.1" 200 -
2018-02-02 22:30:55,592 : INFO : prediction : licenseVerifier.checkLice
nse : License is valid
2018-02-02 22:30:55,705 : INFO : prediction : stateResource.formatPred
iction : Prediction stats:
- 2/2 jobs have a TIMEOUT status
- 0/2 jobs have a COMPLETED status
```

Knowing that the jobs will probably result in TIMEOUT is valuable information to the user: changing the parameters here is a way to find the appropriate set of parameters which will provide the related job with a high probability of ending in COMPLETED state when really submitted on the cluster.

For instance, for the first job a possible action would be to increase the `"TimeLimit"` (WALLTIME) value. For the purposes of demonstration, if we increase `"TimeLimit"` from 7200.0 to 72000.0 (ten times more), Predict-IT forecasts that the first job will now be COMPLETED:

```
Predictions for jobs' state:
+ Job 1 (rae_cruise) predicted state is COMPLETED (confidence 54.9%)
  Job details: {'Account': 'default', 'Day_Submit': '28', 'GID': '1200', 'Hour_Submit': '17'
, 'JobName': 'rae_cruise', 'Minute_Submit': '15', 'Month_Submit': '8', 'Partition': 'debug',
 'QOS': '1', 'ReqCPUS': '24', 'ReqNodes': '1', 'Second_Submit': '59', 'Submit': '2018-08-28
17:15:59', 'Timelimit': '72000', 'UID': '1201', 'Weekday_Submit': '1', 'Year_Submit': '2018'
, 'confidence': 0.6424019607843137, 'prediction': 'COMPLETED'}
+ Job 2 (acwssi1) predicted state is TIMEOUT (confidence 71.2%)
  Job details: {'Account': 'default', 'Day_Submit': '28', 'GID': '550', 'Hour_Submit': '17',
 'JobName': 'acwssi1', 'Minute_Submit': '15', 'Month_Submit': '8', 'Partition': 'prod', 'QOS
': '1', 'ReqCPUS': '120', 'ReqNodes': '1', 'Second_Submit': '24', 'Submit': '2015-08-28 17:1
5:24', 'Timelimit': '3600', 'UID': '550', 'Weekday_Submit': '4', 'Year_Submit': '2015', 'con
fidence': 0.7124019607843137, 'prediction': 'TIMEOUT'}
```

Now, we request a prediction of the maximum execution time (runtime/walltime) using the same input file with the increased `"TimeLimit"` (from 7200.0 to 72000.0) for the first job:

```
./client runtime -f  client.inputfile.json
```

The results are:

```
Predictions for jobs' maximum runtime:
+ Job 1 (rae_cruise) predicted maximum runtime is 03:00:00 (confidence 46.2%)
  Job details: {'Account': 'default', 'Day_Submit': '28', 'GID': '1200', 'Hour_Submit': '17'
, 'JobName': 'rae_cruise', 'Minute_Submit': '15', 'Month_Submit': '8', 'Partition': 'debug',
 'QOS': '1', 'ReqCPUS': '24', 'ReqNodes': '1', 'Second_Submit': '59', 'Submit': '2018-08-28
17:15:59', 'Timelimit': '72000', 'UID': '1201', 'Weekday_Submit': '1', 'Year_Submit': '2018'
, 'confidence': 0.36233333333333334, 'prediction': '00:01:00'}
+ Job 2 (acwssi1) predicted maximum runtime is 10:00:00 (confidence 33.2%)
  Job details: {'Account': 'default', 'Day_Submit': '28', 'GID': '550', 'Hour_Submit': '17',
 'JobName': 'acwssi1', 'Minute_Submit': '15', 'Month_Submit': '8', 'Partition': 'prod', 'QOS
': '1', 'ReqCPUS': '120', 'ReqNodes': '1', 'Second_Submit': '24', 'Submit': '2015-08-28 17:1
5:24', 'Timelimit': '3600', 'UID': '550', 'Weekday_Submit': '4', 'Year_Submit': '2015', 'con
fidence': 0.33233333333333337, 'prediction': '00:01:00'}
```

This time, '*prediction' provides us with a forecast of the* WALLTIME *for each of the jobs*. Here we are especially interested in the first job since we know it was predicted to be COMPLETED. This way, it makes sense to focus only on it: we see that it was forecast to be completely executed within 3 hours.

REST API

The server can be queried through its REST APIs, accessible on its listening IP and port (e.g., http://localhost:9999/).

Server version

`GET srv-version` or `version`

Returns the current version of the server, for example:

```
$ curl -s -X GET "http://predictitserver:9999/version" | python -mjson.tool
{
    "version": "1.1"
}
```

All other methods must be queried with the server version prepended to the name of the prediction, i.e., http://HOSTNAME:PORT/VERSION/PREDICTION (with PREDICTION being either `runtime` or `state`).

Get prediction description

`GET` *`VERSION/PREDICTION`*

Returns a description of the prediction.

Example for `state`:

```
$ curl -s -X GET "http://predictitserver:9999/1.1/state" | python -mjson.tool
{
    "description": "Predict if the job is likely to end in TIMEOUT (killed by the job
 scheduler at the end of its maximum execution time), or COMPLETED (the job had the t
ime to run to the end).",
    "features": [
        "GID",
        "UID",
        "Month_Submit",
        "Requested_CPUS",
        "Hour_Submit",
        "Partition",
        "Year_Submit",
        "Weekday_Submit",
        "Timelimit",
        "JobName",
        "Day_Submit",
        "Requested_Nodes"
    ],
    "url": "state"
}
```

The `description` field is a human-readable description of what is predicted.
The list of `features` is the name of the parameters you can provide to the prediction algorithm as input to query a prediction. None of them are mandatory, but the more you provide, the better the prediction is likely to be.
The `url` field reminds the URL used to query the prediction.

Request a prediction

`POST VERSION/PREDICTION`

With `PREDICTION` one of `exectime`, `state`, or `memory`.

Data: a JSON structure containing the parameters related to the job submission. All parameters must be grouped inside the `jobInput` parameter. Then each parameter is a list, each element of this list referring to 1 job. Example:

```
{"jobInput":{"GID":["3000","3000"],"UID":["3177","3177"],"Requested_CPUS":["1
024","128"],"Requested_Memory":["10485760000","104857600"],"Requested_Nodes":
["1","1"],"Timelimit":["18000","7200"]}}
```

Returns a JSON structure containing the list of all requested jobs with the associated prediction in the `prediction` parameter. Each job has its own `prediction` parameter with the prediction value. (Note that the structure of the output is a bit different than from the input).

```
$ curl -s -X POST -d '{"jobInput":{"GID":["3000","3000"],"UID":["3177","3177"],"Requested_CPUS":["1024",
"128"],"Requested_Memory":["10485760000","104857600"],"Requested_Nodes":["1","1"],"Timelimit":["18000","
7200"]}}' "http://predictitserver:9999/1.1/exectime"
{
  "prediction": [
    {
      "GID": "3000",
      "Requested_CPUS": "1024",
      "Requested_Memory": "10485760000",
      "Requested_Nodes": "1",
      "Timelimit": "18000",
      "UID": "3177",
      "confidence": 0.7395,
      "prediction": "04:00:00"
    },
    {
      "GID": "3000",
      "Requested_CPUS": "128",
      "Requested_Memory": "104857600",
      "Requested_Nodes": "1",
      "Timelimit": "7200",
      "UID": "3177",
      "confidence": 0.6295,
      "prediction": "03:00:00"
    }
  ]
}
```

Errors

Here are some of the errors you can get, and the associated explanation.

If the URL does not exist:

```
{
    "error": "Not found"
}
```

If the server has not finished preparing its prediction algorithms:

```
{
    "action": "Retry later",
    "message": "No prediction algorithm available",
    "status": 400,
    "sub_code": 1
}
```

If there are no input parameters:

```
$ curl -s -X POST -H 'Content-Type: application/json, accept: application/json' -d '{}'
"http://localhost:5900/1.0/runtime"| python -mjson.tool
{
    "action": "Update your input parameters",
    "message": "Needed input parameters: data",
    "status": 400,
    "sub_code": 2
}
```

If the input parameters have the wrong format:

```
$ curl -s -X POST -H 'Content-Type: application/json, accept: application/json' -d '{"jobIn
put":{"UID":"12"}}' "http://localhost:5900/1.0/runtime"| python -mjson.tool
{
    "action": "Check your input data (If using all scalar values, you must pass an index)",
    "message": "Problem while parsing your input data",
    "status": 500,
    "sub_code": 1
}
```

Extract job scheduler logs

If you do not want Predict-IT to automatically and periodically extract new data from the job scheduler, you can use the scripts provided 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.
- `extractPBSTorqueData.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 Predict-IT, only the first file is currently relevant: `HOSTNAME_DATE.jobs`, you need to specify the path to this file in the `server.conf` with the `inputFile` in the `jobScheduler` section.

Administration

Here we describe all the parameters in the configuration files `server.conf` and `client.conf`. Those can be customized depending on the user dataset and what target is supposed to be predicted.

Environment Variables

Configuration files location can be configured through environment variables. It is a good practice to preserve the original (default) server and client configuration files the way it is. If it is necessary to modify any parameter, we recommend creating an editable copy and set the system variable to let Predict-IT know where it is:

`PIT_SERVER_CONF`. specifies the location of the server configuration file.
Example: `export PIT_SERVER_CONF=/home/usename/server.conf`

`PIT_CLIENT_CONF`: specifies the location of the client configuration file.
Example: `export PIT_CLIENT_CONF=/home/usename/client.conf`

Server configuration file

The `server.conf` file is organized in the following sections.

Section	Parameters
`[server]` General parameters	**host: hostname or IP address to listen on (default is 127.0.0.1). Use 0.0.0.0 to listen on all interfaces** `host = 127.0.0.1` **port: listening port (default is 9999)** `port = 9999` **Path to license file. If not specified, the license file will be searched in license/license.lic**. `licenseFile = PATH_TO_FILE` **Log file path (default is /tmp/prediction.log)** `logFile = /tmp/prediction.log` **Log Level (default is INFO/20)** *Level Numeric value* *CRITICAL 50* *ERROR 40* *WARNING 30* *INFO 20* *DEBUG 10* *NOTSET 0* `logLevel = 10` **Separate logs for the different targets? (default False)** *If true, then for each target, a specific file will be created for its logs.* *The files will be named with <logFile>.<target>* `separateLogs=False` **Select the target(s) for the prediction** *Currently supported targets: ExecutionTimeBins, State (default is ExecutionTimeBins,State)* `target=ExecutionTimeBins,State`

`[jobScheduler]` Reading job-scheduler logs	**Which job scheduler should we use? (this parameter is mandatory)** *- slurm* *- torque* *- pbs* *- oar* *- swf (not a real job scheduler, just reads an swf file)* *- pickleFile (not a real job scheduler, just reads a pickle file)* `driver=slurm` **Optional argument to load data from a file instead of from really calling the job scheduler** `inputFile=PATH_TO_FILE` **Optional Start time to get the historical data (YYYY-MM-DD HH:MM:SS)** `starttime=2015-01-01 00:00:00` **Check job scheduler (True/False) (default True)** *If you use an input file, you should set that to False* `check=false` **Additional parameters for the job scheduler plugin.** *All additional parameters must start with 'param_'.* *+ **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* *+ PBS:* *- acct_dir: path to accounting directory (default="/var/spool/pbs/server_priv/accounting/")* *+ **Slurm**:* *- - 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, it allows to get the MaxRSS values properly.* *+ **SWF**: None* *+ **Torque**:* *- - acct_dir: path to accounting directory (default=/var/spool/torque/server_priv/accounting/)* `param_acct_dir=/var/spool/torque/server_priv/accounting/` `param_separator=slurmdefault`

`[XXX_algorithms]` General parameters for the algorithms that predict `XXX` (with `XXX` being `ExecutionTimeBins`, `State`, or `MaxRSSBins`. See parameter `target` in section `server`)	**Comma separated name of algorithms that must be used.** Then, you must create one section in this configuration file for each algorithm Supported algorithms: randomForest, decisionTree, extraTrees, gradientBoosting `usedAlgorithms=randomForest` **Currently only the voting metaAlgorithm is supported. Do not change!** `metaAlgorithm=voting` **Train size should be a float value within (0, 1).** *For example, 0.999. Default = None* `trainSize=0.95` **Remove features that do not have a significant number of samples (default 0.5)** `significant_n_samples = 0.5` **refreshModel=YEAR MONTH DAY HOUR MINUTE SECOND** *The following table lists all the available expressions for use in the fields from year to second. Multiple expression can be given in a single field, separated by commas.* *Expression Field Description* ** any Fire on every value* **/ any anyFire every a values, starting from the minimum* *a- any anyFire on any value within the a-b range (a must be smaller than b)* *a-b/ any anyFire every c values within the a-b range* *x,y,z any Fire on any matching expression; can combine any number of any of the above expressions"""* *example to refresh every 5 hours:* `refreshModel=* * * */5 * *` **trainOnce (True,False) (default False)** *train the model only once, and then pauses the training task* `trainOnce=True` **Switch on/off (True/False) data scaling (normalizing) (default True)** `scale=False` **Switch on/off (True/False) data Balancing (default True)** `balance=False` **Filenames for saving metrics** `file_train_parameters = /tmp/STATE_training_parameters.p` `file_train_metrics = /tmp/STATE_training_metrics.p` `file_test_metrics = /tmp/STATE_test_metrics.p`

	`file_trainingTask = /tmp/STATE_trainingTask.p` **Parameters for cross validation.** *cv={True,False}: to switch on/off cross validation. (default True)* *Might be useful to reduce overall calculation time.* *fold_strategy={kfold,stratified} (default kfold)* *scoring multiclass={f1 micro, f1 macro} (default f1 micro)* *n_splits: integer, number of folds for kfolds (default 3)* `cv=True` `fold_strategy=kfold` `scoring_multiclass=f1_micro` `n_splits=4`
`[XXX_randomForest]` Specific parameters for the initialization of random forest algorithm for `XXX` predictions (with `XXX` being `ExecutionTimeBins`, `State`, or `MaxRSSBins`. See parameter `target` in section `server`)	**Parameters for the initialization of the algorithm.** *All these parameters must start with 'param_'.* *Parameters and default values:* *max_depth = None* *max_features=auto* *min_samples_split = 2* *random_state = 0* *n_estimators = 10* *criterion = gini* *class_weight=None* *param_oob_score=True*
`[voting]`	*Parameters for the initialization of the voting algorithm.* *All these parameters must start with 'param_'.* *Parameters and default values:* *voting_case={hard,soft} (default hard)* *weight_algorithms = None (array-like, shape = [n_classifiers]) (default None)* *n_jobs: integer, number of threads to use (-1 means no restrictions, use all cores) (default -1)*
Other sections: `[decisionTree]` `[extraTrees]` `[gradientBoosting]`	Templates for other algorithms. See `server.conf` for a complete list of parameters.

Client configuration file

The `client.conf` file contains only one section described below:

Section	Parameters
`[server]`	URL used by the client side to contact the server. *Format is* `http://IP_OR_HOSTNAME:PORT/` `baseURL=http://127.0.0.1:9999/` Warning: the IP address indicated in the baseURL variable must be the IP indicated in server.conf server-side.

Troubleshooting

Here we provide a list with known issues and the actions that you can take (if necessary) to circumvent them.

Value error
The following lines are printed on the terminal prompt during the execution of the tool:

```
ValueError: Buffer dtype mismatch, expected 'Python object' but got 'long'
Exception ignored in: 'pandas._libs.lib.is_bool_array'
ValueError: Buffer dtype mismatch, expected 'Python object' but got 'long'
ValueError: Buffer dtype mismatch, expected 'Python object' but got 'long'
Exception ignored in: 'pandas._libs.lib.is_bool_array'
ValueError: Buffer dtype mismatch, expected 'Python object' but got 'long'.
```

Such “error” is actually a warning about the data type that is being used in the conversion of the log data (usually strings and numbers) into categorical data (used by the embedded algorithm that performs the prediction). This does not affect prediction performance and should be removed in a coming release.

Selection of parameter class_weight in `[XXX_randomForest]`
The parameter class_weight allows for equalizing the relevance of the classes. It is especially useful when in State prediction the number of TIMEOUT jobs is much lower than the number of COMPLETED, providing the user with another way of balancing the dataset. Four options are available for class_weight: 'None', 'balanced', 'balanced_subsample', and '{0: weight_class_0, 1: weight_class_1, ...}', where weight_class_j for j=0,1,… can have any value greater or equal to zero.

In State prediction, if the user chooses class_weight = {0: weight_class_0, 1: weight_class_1} while keeping cv=True (calculation of cross validation in [State_algorithms]), the execution breaks and no prediction model is generated. This way, if using a dictionary like the one above to define custom weights for each class, make sure to turn off the calculation of cross validation. This issue will be fixed in the next release.