Table of Contents

****
****

1. Tool general information

Profiling and Classification tool is used:
1. In order to acquire the footprints of an unknown application or one of its components, on a specific physical host. The footprints consist of average execution metrics concerning CPU, RAM, disk and network utilization.
2. In order to acquire the same type of footprints of a set of test-applications called benchmarks on the same physical host.
3. The results of both execution modes of the tool, will be served as an input to a classifier, which will decide upon the type of the application. Application types are defined according to the categorization of the benchmarks.

For an arbitrary application, the general approach is to run the profiling/classification process in three cases (low, medium, high workload case). In this way 3 categories are extracted, which represent the expected application behavior in the three different cases.

The outcome of the aforementioned process (categorization) can be used from the side of a Cloud Adopter in the process of selecting the best performing Cloud service (based on the specific category),

and also from the side of the Cloud Provider, in order to optimize its infrastructure resource allocation, based on the running application categories extracted.

****
****

2. Infrastructure and Software Requirements

For running the Profiling and Classification tool three machines will be needed.

1) The physical machine
2) A virtual machine running on (1) configured to host one or more components of your application
3) A virtual machine running on (1) configured to host benchmark tests

2.1 Physical machine requirements (1)

-Linux OS (Ubuntu/Mint/Debian distribution)
-Java installed
-SSH installed
-Check in the /etc/sudoers file that the line: "Default requiretty" does not exist, or comment it if exists

2.2 Configuring the Application VM (2)

No configurations apart from the ones indicated by the hosted application component/s.

2.3 Configuring the Benchmark VM (3)

-CentOS 64-bit V6.5 Server
-SSH installed
-Check in the /etc/sudoers file that the line: "Default requiretty" does not exist, or comment it if exists
-Follow the installation instructions of each benchmark.

****
****

3. Tool Installation

To install the Profiling and Classification tool, please run the sh script installer in the Ubuntu physical machine:

```
sudo sh Profiling_and_Classification-installer.sh
```

The script will request for a full installation path, which will be also used as the tool workspace:
"*Please provide an existing folder as the installation path/workspace (Linux full path)> "*
There the end user has to provide a full path of an existing folder (e.g. "/home/cloud/profiling-workspace"), in which he has access rights.

The script will create all needed folders and configuration files, and will also install the necessary monitoring tools (tshark, Pidstat) and download the tool latest version from the CloudPerfect repository. If everything went well you should get a success message:
"Profiler successfully installed at /home/cloud/profiling-workspace"
In the selected path, the folders "profiler" and "classifier" must exist, with various configuration files included.

In the "profiler" folder the file "benchmark_workloads.txt" is the workloads file containing the list of commands to run in the benchmarks VM, during the profiling of the benchmarks.

****
****

4. User Manual

4.1 Profiling of benchmarks (using the benchmarks VM)

4.1.1 Preparation

On the physical host:

a. Make sure that the benchmarks VM is up and running. This has all benchmarks and workloads already installed.
b. Find the PID of the benchmarks VM (Open a terminal on the physical machine and type: ps aux|grep "name of the vm")
c. Find the IP of the benchmarks VM
d. (**needed only for console mode execution**) Edit the installation-path/profiler/benchmark-info.txt configuration file so that it looks as follows:

```
installation-path
VM_PID
Host_IP,Host_interface
VM_IP,VM_User,Workloads_file,VM_User_pass,VM_Root_pass
```

Where:
installation-path: The path that the tool is installed (set by default)
VM_PID: PID of the benchmarks VM
Host_IP: Physical machine IP
Host_interface: Physical machine Network Interface
VM_IP: IP of the benchmarks VM
VM_User: Benchmark VM Username
Workloads_file: Full path of the workloads file (set by default)

VM_User_pass: Benchmark VM User password
VM_Root_pass: Benchmark VM root password

4.1.2 Execution

On the physical host open a terminal and navigate to the *installation-path.* Type:

```
sh profiler-run.sh
```

(a) To execute on console mode:

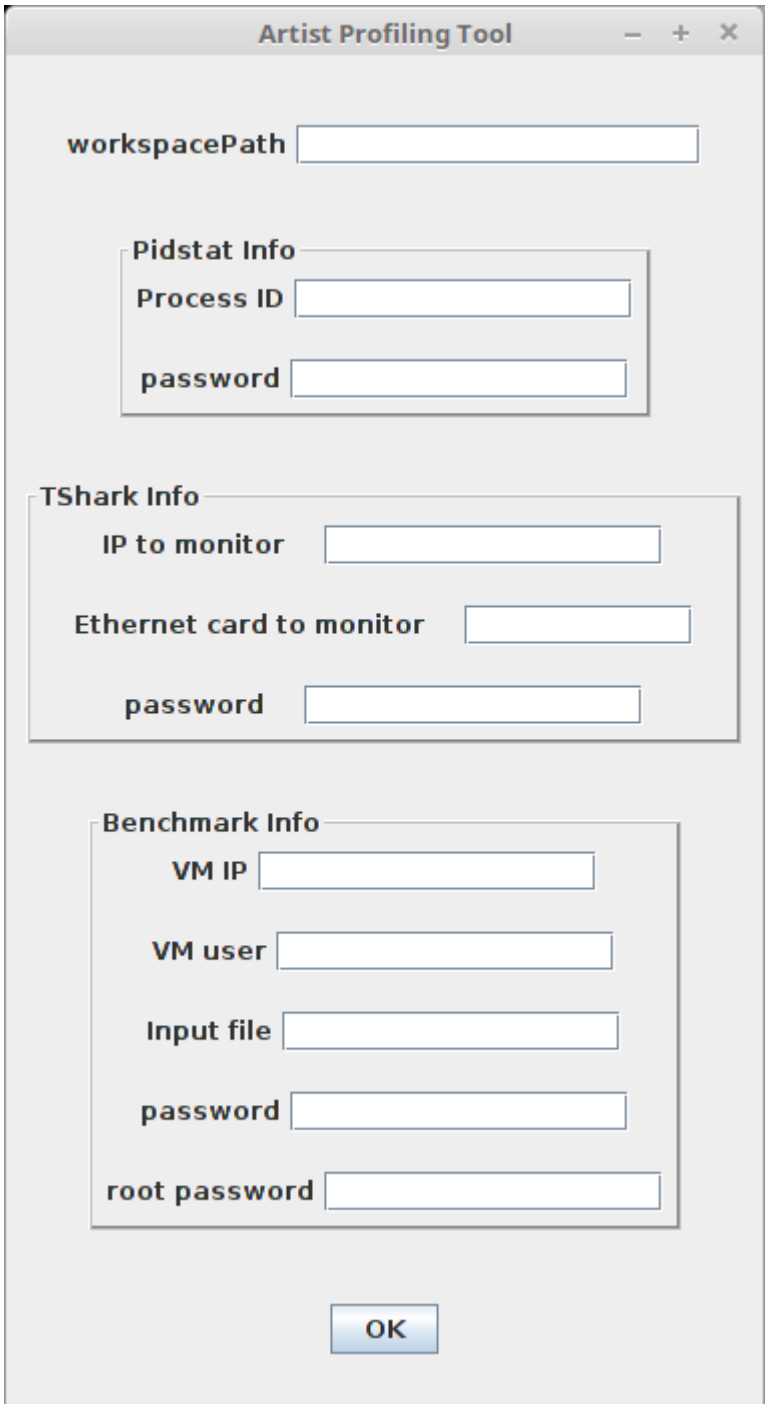
When the message “Choose your profiling target. Type [a] for application or [b] for benchmarks>” appears type “b”.
When the message “Please type [f] if you want to use the respective configuration txt file, else press enter (GUI will appear)> ” appears type “f” (this will take input values from the file defined in step (d) of preparation).
Provide the physical machine root password when asked.

Depending on the benchmarks and the workloads provided the execution may take more than an hour.

(b) To execute with graphical user interface mode:

When the message “Choose your profiling target. Type [a] for application or [b] for benchmarks>” appears type “b”.
When the message “Please type [f] if you want to use the respective configuration txt file, else press enter (GUI will appear)> ” appears press ENTER.

A graphical interface appears which introduces a form with empty text fields.



Specification of the fields (see step (d) of preparation) :

-Workspace path: The installation path

Pidstat Info Group:

-Process ID: Process ID of the application VM.

-password: Password for executing commands under root privileges.

TShark Info Group:

-IP to monitor: IP of the VM to be monitored.

-Ethernet card to monitor: The ethernet card on which tshark will capture traffic (the eth card of the physical host)

-password: Password for executing commands under root privileges.

Benchmark Info Group:

-VM IP: The IP of the Benchmark VM (used for the ssh command)

-VM user: The user for the SSH command

-Input File: The input file containing the commands for the workload execution.

-password: Password for the user used for the SSH command

-root password: Password for running commands under root privileges (in the Benchmark VM)

Fill the fields of the form and press the OK button. Then wait until the whole process is finished.

Note: Execution can be monitored via the terminal. When the whole process has finished make sure that every executed command has been terminated with exit value 0.

4.1.3 Results

Output profiling results will be written in the “BenchmarkProfiles-trainingFile.txt”, which will be used as the training file for the classifier.

The aforementioned results file will contain some labels (benchmark-workload name) and numeric values (profile vector) in each line.

The profile vector is consisted of 23 values which represent the following pidstat and tshark output metrics (in the same order as provided in the results):
%user, %system, %guest, %CPU, CPU, kB_rd/s, kB_wr/s, kB_ccwr/s, minflt/s, majflt/s, VSZ Virtual Size, RSS, %MEM, cswch/s, nvcswch/s, number of packets, data byte rate (bytes/s), average packet size (bytes), average packet rate (packets/s).
Last four values of network metrics are provided both for the monitored IP acting as source as well as acting as destination.

Detailed information about profile metrics can be found in deliverable D2.3, sect. 5.1.1: Monitored Metrics per VM.

****

4.2 Profiling and Categorizing the Application VM

4.2.1 Preparation

On the physical host:

a. Make sure that the application VM is up and running.

b. Find the PID of the application VM (Open a terminal on the physical machine and type: ps aux|grep "name of the vm")

c. Find the IP of the application VM

d. (**needed only for console mode execution**) Edit the installation-path/profiler/application-info.txt configuration file so that it looks as follows:

```
installation-path
VM_PID,Exec_Time
Host_IP,Host_interface
```

Where:
installation-path: The path that the tool is installed (set by default)
VM_PID: PID of the benchmarks VM
Exec_Time: The time that the profiling of the application will run (in seconds)
Host_IP: IP of the VM to be monitored
Host_interface: Physical machine Network Interface

4.2.2 Execution

On the physical host open a terminal and navigate to the *installation-path.* Type:

```
sh profiler-run.sh
```

(a) To execute on console mode:

When the message “Choose your profiling target. Type [a] for application or [b] for benchmarks>” appears type “a”.

When the message “Please type [f] if you want to use the respective configuration txt file, else press enter (GUI will appear)> ” appears type “f” (this will take input values from the file defined in step (d) of preparation).

Provide the physical machine root password when asked.

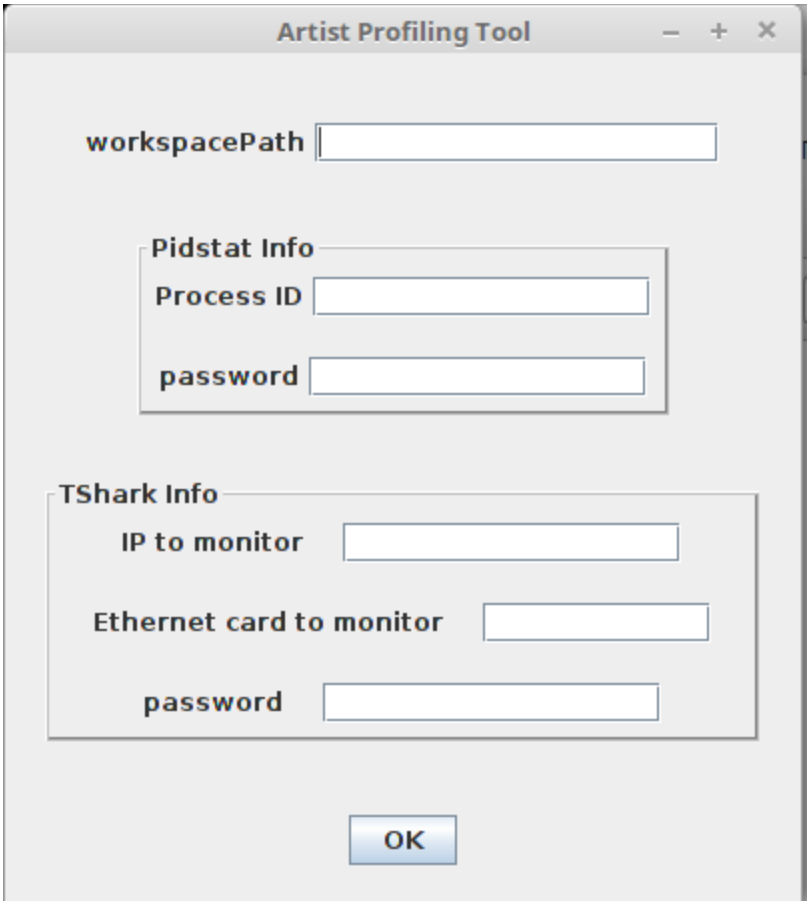
The profiling will run based on the *Exec_Time* parameter provided in the configuration file and will then generate two profile result files in the profiler folder (Pidstat.txt and Tshark.txt).
Those files will be automatically provided to the classifier, along with the “BenchmarkProfiles-trainingFile.txt” (profiling of benchmarks must have already ran), which will finally return the application Category in the console.

(b) To execute on graphical user interface mode:

When the message “Choose your profiling target. Type [a] for application or [b] for benchmarks>” appears type “a”.

When the message “Please type [f] if you want to use the respective configuration txt file, else press enter (GUI will appear)> ” appears type ENTER.

Then a graphical interface appears which introduces a form with empty text fields.



Specification of the fields:

-Workspace path: The workspace directory (see step (d) of preparation)

Pidstat Info Group:

-Process ID: Process ID of the application VM.

-password: Password for executing commands under root privileges.

TShark Info Group:

-IP to monitor: IP of the VM to be monitored.

-Ethernet card to monitor: The ethernet card on which tshark will capture traffic (the eth card of the physical host)

-password: Password for executing commands under root privileges.

Fill the fields of the form and press the OK button.

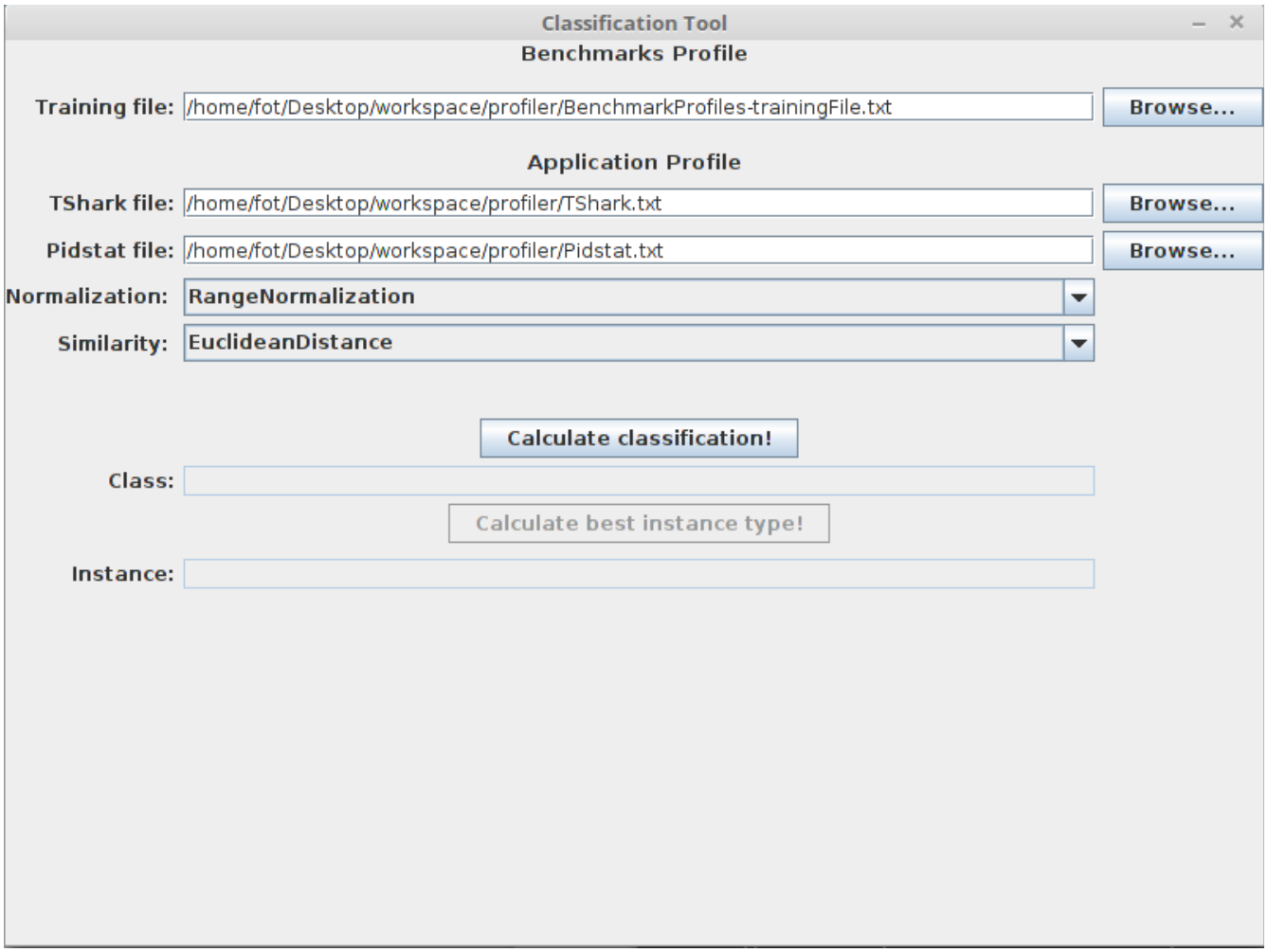
Two buttons (Start/Stop) will appear on the screen.

Press the Start button, wait until enough time has passed (for the application to execute the defined workload) and then press the Stop button to save the profiles.

Finally, in the Linux terminal, run the classifier script:

sh ../classifier/classifier-run.sh

The following GUI will appear:



Please provide the training file (BenchmarkProfiles-trainingFile.txt) along with the profiling result files (Pidstat.txt and Tshark.txt) to calculate the application Category.

Note: Execution can be monitored via the terminal. When the whole process has finished make sure that every executed command has been terminated with exit value 0. For other errors see the errors section.

****
****

5. Monitor Execution and Cases of Misuse

This section describes possible reasons for failure of the profiling process, together with some monitoring advice for solving them.

Note: Every execution detail is printed on the terminal (System commands, actions). Monitor the exit values of the system commands.
Execution was successful only if every command terminated with exit value: 0.

Possible error messages:

-Invalid workspace! Exiting... : Error meaning that the workspace path provided by the user does not exist.

-Wrong password (Pidstat or Tshark)! Exiting... : Error meaning that the password the user provided for running the corresponding
commands under root privileges was wrong.

-Wrong Process ID! Exiting... : Error meaning that the PID the user provided does not correspond to a valid process of the OS.

-Wrong network device! Exiting... : Error meaning that the name of the ethernet card was not valid.

-Error in command. Exiting ... : Error meaning that one or more of the system commands executed was not
terminated with 0 exit value (Unexpected error, the profiling process must be repeated)

-Please enter argument: application/benchmarks : Error meaning that the command for executing the jar file was incorrect. User
must retype the command correctly.

-Wrong number of arguments! : Error meaning that the number of arguments provided with the java -jar commands was
more than 2. User must retype the command correctly.