



COMP SUPERSCALAR

COMPSS at BSC

MareNostrum 3 Manual

VERSION: 2.0.R.RC1702

February 21, 2017



**Barcelona
Supercomputing
Center**

Centro Nacional de Supercomputación

This manual only provides information about the COMPSs usage at MareNostrum. Specifically, it details the available COMPSs modules, how to load them and how to create and track COMPSs jobs.

If you want to install COMPSs on your local machine please refer to the *COMPSs Installation Manual* available at our webpage <http://compss.bsc.es>.

For further information about the application's execution please refer to the *COMPSs User Manual: Application execution guide* available at <http://compss.bsc.es> .

For further information about the application's development please refer to the *COMPSs User Manual: Application development guide* available at <http://compss.bsc.es/> .

For full COMPSs example application (codes, execution commands, results, logs, etc.) please refer to the *COMPSs Sample Applications* available at <http://compss.bsc.es/>

Contents

1	COMP Superscalar (COMPSs)	1
2	COMPSs Modules	2
2.1	Available modules	2
2.2	Configuration	2
3	COMPSs Jobs	4
3.1	Submitting COMPSs jobs	4
3.2	Tracking COMPSs jobs	6
4	Enabling COMPSs Monitor	8
4.1	Configuration	8
4.2	Execution	8

1 COMP Superscalar (COMPSs)

COMP Superscalar (COMPSs) is a programming model which aims to ease the development of applications for distributed infrastructures, such as Clusters, Grids and Clouds. COMP Superscalar also features a runtime system that exploits the inherent parallelism of applications at execution time.

For the sake of programming productivity, the COMPSs model has four key characteristics:

- **Sequential programming:** COMPSs programmers do not need to deal with the typical duties of parallelization and distribution, such as thread creation and synchronization, data distribution, messaging or fault tolerance. Instead, the model is based on sequential programming, which makes it appealing to users that either lack parallel programming expertise or are looking for better programmability.
- **Infrastructure unaware:** COMPSs offers a model that abstracts the application from the underlying distributed infrastructure. Hence, COMPSs programs do not include any detail that could tie them to a particular platform, like deployment or resource management. This makes applications portable between infrastructures with diverse characteristics.
- **Standard programming languages:** COMPSs is based on the popular programming language Java, but also offers language bindings for Python and C/C++ applications. This facilitates the learning of the model, since programmers can reuse most of their previous knowledge.
- **No APIs:** In the case of COMPSs applications in Java, the model does not require to use any special API call, pragma or construct in the application; everything is pure standard Java syntax and libraries. With regard the Python and C/C++ bindings, a small set of API calls should be used on the COMPSs applications.

2 COMPSs Modules

2.1 Available modules

COMPSs is configured in MareNostrum (MN3) as a Linux Module. Type *module available COMPSs* to list the available COMPSs modules through Linux Module configuration and *module load COMPSs/ < version >* to load it.

```
$ module available COMPSs
----- /apps/modules/modulefiles/tools -----
COMPSs/1.1.2_gpfs
COMPSs/1.1.2_scratch
COMPSs/1.2
COMPSs/1.3
COMPSs/1.4
COMPSs/2.0

COMPSs/release(default)
COMPSs/trunk

$ module load COMPSs/release
load java/1.8.0u66 (PATH, MANPATH, JAVA_HOME, JAVA_ROOT, JAVA_BINDIR,
                  SDK_HOME, JDK_HOME, JRE_HOME)
load MKL/11.0.1 (LD_LIBRARY_PATH)
load PYTHON/2.7.3 (PATH, MANPATH, LD_LIBRARY_PATH, C_INCLUDE_PATH)
load COMPSs/release (PATH, MANPATH, IT_HOME)
```

The following command can be run to check if the correct COMPSs version has been loaded:

```
$ runcompss --version
COMPSs version <version>
```

2.2 Configuration

The COMPSs module contains **all** the COMPSs dependencies, including Java, Python and MKL. Modifying any of these dependencies can cause execution failures and thus, we **do not** recomend to change them. Before running any COMPSs job please check your environment and, if needed, comment out any line inside the *.bashrc* file loading custom COMPSs, Java, Python and/or MKL modules.

The COMPSs module needs to be loaded in all the nodes that will run COMPSs jobs. Consequently, the *module load* **must** be included in your *.bashrc* file. To do so, please run the following command with the corresponding COMPSs version:

```
$ cat "module load COMPSs/release" >> ~/.bashrc
```

Log out and back in again to check that the file has been correctly edited. The next listing shows an example of the output generated by well loaded COMPSs installation.

```
$ exit
$ ssh USER@mn1.bsc.es
load java/1.8.0u66 (PATH, MANPATH, JAVA_HOME, JAVA_ROOT, JAVA_BINDIR,
                  SDK_HOME, JDK_HOME, JRE_HOME)
load MKL/11.0.1 (LD_LIBRARY_PATH)
load PYTHON/2.7.3 (PATH, MANPATH, LD_LIBRARY_PATH, C_INCLUDE_PATH)
load COMPSs/release (PATH, MANPATH, IT_HOME)

$ runcompss --version
COMPSs version <version>
```

Please remember that COMPSs runs in several nodes and your current environment is not exported to them. Thus, all the needed environment variables **must** be loaded through the *.bashrc* file.

3 COMPSs Jobs

3.1 Submitting COMPSs jobs

COMPSs jobs can be easily submitted by running the **enqueue_compss** command. This command allows to configure any **runcompss** option and some particular queue options such as the queue system, the number of nodes, the wallclock time, the master working directory, the workers working directory and number of tasks per node.

Next, we provide detailed information about the *enqueue_compss* command:

```
$ enqueue_compss -h

Usage: enqueue_compss [queue_system_options] [COMPSs_options]
       application_name application_arguments

* Options:
  General:
    --help, -h                Print this help message

  Queue system configuration:
    --sc_cfg=<name>           SuperComputer configuration file to use.
                               Must exist inside queues/cfgs/
                               Default: default

  Submission configuration:
    --exec_time=<minutes>     Expected execution time of the application (in minutes)
                               Default: 10
    --num_nodes=<int>         Number of nodes to use
                               Default: 2
    --num_switches=<int>      Maximum number of different switches. Select 0 for no
                               restrictions.
                               Maximum nodes per switch: 18
                               Only available for at least 4 nodes.
                               Default: 0
    --gpus_per_node=<int>     Number of desired GPUs per node.
                               Leave this field empty if your application doesn't use GPUs.
                               Default:
    --queue=<name>            Queue name to submit the job. Depends on the queue system.
                               For example (MN3): bsc_cs | bsc_debug | debug | interactive
                               Default: default
    --reservation=<name>     Reservation to use when submitting the job.
                               Default: disabled
    --job_dependency=<jobID> Postpone job execution until the job dependency has ended.
                               Default: None
    --storage_home=<string>   Root installation dir of the storage implementation
                               Default: null
    --storage_props=<string> Absolute path of the storage properties file
                               Mandatory if storage_home is defined

  Launch configuration:
    --tasks_per_node=<int>    Maximum number of simultaneous tasks running on a node
                               Default: 16
    --node_memory=<MB>       Maximum node memory: disabled | <int> (MB)
                               Default: disabled
    --network=<name>         Communication network for transfers: default | ethernet
                               | infiniband | data.
                               Default: ethernet

    --prolog="<string>"      Task to execute before launching COMPSs (Notice the quotes)
                               If the task has arguments split them by "," rather than spaces.
                               This argument can appear multiple times for more than one
                               prolog action
                               Default: Empty
    --epilog="<string>"      Task to execute after executing the COMPSs application
                               (Notice the quotes)
```

	<p>If the task has arguments split them by <code>" , "</code> rather than spaces. This argument can appear multiple <code>times</code> for more than one epilog action</p> <p>Default: Empty</p>
<code>--master_working_dir=<path></code>	<p>Working directory of the application</p> <p>Default: .</p>
<code>--worker_working_dir=<name path></code>	<p>Worker directory. Use: <code>scratch</code> <code>gpfs</code> <code><path></code></p> <p>Default: <code>scratch</code></p>
<code>--worker_in_master_tasks=<int></code>	<p>Maximum number of tasks that the master node can run as worker. Cannot exceed <code>tasks_per_node</code>.</p> <p>Default: 0</p>
<code>--worker_in_master_memory=<int> MB</code>	<p>Maximum memory in master node assigned to the worker. Cannot exceed the <code>node_memory</code>.</p> <p>Mandatory <code>if</code> <code>worker_in_master_tasks</code> is specified.</p> <p>Default: disabled</p>
<code>--jvm_worker_in_master_opts="<string>"</code>	<p>Extra options for the JVM of the COMPSs Worker in the Master Node.</p> <p>Each option separated by <code>" , "</code> and without blank spaces (Notice the quotes)</p> <p>Default: <code>" "</code></p>
Runcompss configuration:	
Tools enablers:	
<code>--graph=<bool>, --graph, -g</code>	<p>Generation of the <code>complete</code> graph (<code>true/false</code>)</p> <p>When no value is provided it is <code>set</code> to <code>true</code></p> <p>Default: <code>false</code></p>
<code>--tracing=<level>, --tracing, -t</code>	<p>Set generation of traces and/or tracing level (<code>[true basic] advanced false</code>)</p> <p>True and basic levels will produce the same traces.</p> <p>When no value is provided it is <code>set</code> to <code>true</code></p> <p>Default: <code>false</code></p>
<code>--monitoring=<int>, --monitoring, -m</code>	<p>Period between monitoring samples (milliseconds)</p> <p>When no value is provided it is <code>set</code> to 2000</p> <p>Default: 0</p>
<code>--external_debugger=<int>, --external_debugger</code>	<p>Enables external debugger connection on the specified port (or 9999 <code>if</code> empty)</p> <p>Default: <code>false</code></p>
Runtime configuration options:	
<code>--task_execution=<compss storage></code>	<p>Task execution under COMPSs or Storage.</p> <p>Default: <code>compss</code></p>
<code>--storage_conf=<path></code>	<p>Path to the storage configuration file</p> <p>Default: None</p>
<code>--project=<path></code>	<p>Path to the project XML file</p> <p>Default: <code>/opt/COMPSs/Runtime/configuration/xml/projects/default_project.xml</code></p>
<code>--resources=<path></code>	<p>Path to the resources XML file</p> <p>Default: <code>/opt/COMPSs/Runtime/configuration/xml/resources/default_resources.xml</code></p>
<code>--lang=<name></code>	<p>Language of the application (java/c/python)</p> <p>Default: <code>java</code></p>
<code>--summary</code>	<p>Displays a task execution summary at the end of the application execution</p> <p>Default: <code>false</code></p>
<code>--log_level=<level>, --debug, -d</code>	<p>Set the debug level: <code>off</code> <code>info</code> <code>debug</code></p> <p>Default: <code>off</code></p>
Advanced options:	
<code>--extrae_config_file=<path></code>	<p>Sets a custom extrae config file. Must be in a shared disk between all COMPSs workers.</p> <p>Default: <code>null</code></p>
<code>--comm=<path></code>	<p>Class that implements the adaptor for communications</p> <p>Supported adaptors: <code>integratedtoolkit.nio.master.NIOAdaptor</code> <code>integratedtoolkit.gat.master.GATAdaptor</code></p> <p>Default: <code>integratedtoolkit.nio.master.NIOAdaptor</code></p>
<code>--conn=<path></code>	<p>Path of the connector jar/s that should be loaded.</p>

	You can use multiple by splitting with ':'
	Supported connectors: compss-{CONN}-connector.jar (where {CONN} can be: "jclouds", "amazon", "docker", "one", "rocci", "vmm", etc.)
--scheduler=<path>	Class that implements the Scheduler for COMPSs Supported schedulers: integratedtoolkit.components.impl.TaskScheduler integratedtoolkit.scheduler.readyscheduler.ReadyScheduler Default: integratedtoolkit.scheduler.readyscheduler.ReadyScheduler
--library_path=<path>	Non-standard directories to search for libraries (e.g. Java JVM library, Python library, C binding library) Default: Working Directory
--classpath=<path>	Path for the application classes / modules Default: Working Directory
--appdir=<path>	Path for the application class folder. Default: /home/cramonco
--base_log_dir=<path>	Base directory to store COMPSs log files (a .COMPSs/ folder will be created inside this location) Default: User home
--specific_log_dir=<path>	Use a specific directory to store COMPSs log files (the folder MUST exist and no sandbox is created) Warning: Overwrites --base_log_dir option Default: Disabled
--uuid=<int>	Preset an application UUID Default: Automatic random generation
--master_name=<string>	Hostname of the node to run the COMPSs master Default:
--master_port=<int>	Port to run the COMPSs master communications. Only for NIO adaptor Default: [43000,44000]
--jvm_master_opts="<string>"	Extra options for the COMPSs Master JVM. Each option separated by "," and without blank spaces (Notice the quotes) Default:
--jvm_workers_opts="<string>"	Extra options for the COMPSs Workers JVMs. Each option separated by "," and without blank spaces (Notice the quotes) Default: -Xms1024m,-Xmx1024m,-Xmn400m
--task_count=<int>	Only for C/Python Bindings. Maximum number of different functions/methods, invoked from the application, that have been selected as tasks Default: 50
--pythonpath=<path>	Additional folders or paths to add to the PYTHONPATH Default: /home/cramonco
--PyObject_serialize=<bool>	Only for Python Binding. Enable the object serialization to string when possible (true/false). Default: false
* Application name:	
For Java applications:	Fully qualified name of the application
For C applications:	Path to the master binary
For Python applications:	Path to the .py file containing the main program
* Application arguments:	
Command line arguments to pass to the application. Can be empty.	

3.2 Tracking COMPSs jobs

When submitting a COMPSs job a temporal file will be created storing the job information. For example:

```
$ enqueue_compss \
--exec_time=15 \
--num_nodes=3 \
--tasks_per_node=16 \
```

```

--master_working_dir=. \
--worker_working_dir=gpfs \
--lang=python \
--log_level=debug \
<APP> <APP_PARAMETERS>

SC Configuration:      default.cfg
Queue:                 default
Reservation:           disabled
Num Nodes:             3
Num Switches:          0
GPUs per node:         0
Job dependency:        None
Exec-Time:             00:15
Storage Home:          null
Storage Properties:    null
Other:
    --sc_cfg=default.cfg
    --tasks_per_node=16
    --master_working_dir=.
    --worker_working_dir=gpfs
    --lang=python
    --classpath=.
    --library_path=.
    --comm=integratedtoolkit.nio.master.NIOAdaptor
    --tracing=false
    --graph=false
    --pythonpath=.
    <APP> <APP_PARAMETERS>
Temp submit script is: /scratch/tmp/tmp.pBG5yfFxEO

$ cat /scratch/tmp/tmp.pBG5yfFxEO
#!/bin/bash
#
#BSUB -J COMPSs
#BSUB -cwd .
#BSUB -oo compss-%J.out
#BSUB -eo compss-%J.err
#BSUB -n 3
#BSUB -R "span[ptile=1]"
#BSUB -W 00:15
...

```

In order to trac the jobs state users can run the following command:

```

$ bjobs
JOBID  USER  STAT  QUEUE  FROM_HOST  EXEC_HOST  JOB_NAME  SUBMIT_TIME
XXXX   bscXX  PEND  XX     login1     XX         COMPSs    Month Day Hour

```

The specific COMPSs logs are stored under the `./COMPSs/` folder; saved as a local *runcompss* execution. For further details please check *COMPSs User Manual: Application Execution* available at our webpage <http://compss.bsc.es> .

4 Enabling COMPSs Monitor

4.1 Configuration

As MareNostrum nodes are connection restricted, the better way to enable the *COMPSs Monitor* is from the users local machine. To do so please install the following packages:

- COMPSs Runtime
- COMPSs Monitor
- sshfs

For further details about the COMPSs packages installation and configuration please refer to the *COMPSs Installation Manual* available at our webpage <http://compss.bsc.es> . If you are not willing to install COMPSs in your local machine please consider to download our Virtual Machine available at our webpage.

Once the packages have been installed and configured, users need to mount the sshfs directory as follows (*MN_USER* stands for your MareNostrum user and the *TARGET_LOCAL_FOLDER* to the local folder where you wish to deploy the MareNostrum files):

```
compss@bsc:~$ scp $HOME/.ssh/id_dsa.pub ${MN_USER}@mn1.bsc.es:~/id_dsa_local.pub
compss@bsc:~$ ssh MN_USER@mn1.bsc.es
"cat ~/id_dsa_local.pub >> ~/.ssh/authorized_keys;
rm ~/id_dsa_local.pub"
compss@bsc:~$ mkdir -p TARGET_LOCAL_FOLDER/.COMPSs
compss@bsc:~$ sshfs -o IdentityFile=$HOME/.ssh/id_dsa -o allow_other
MN_USER@mn1.bsc.es:~/COMPSs
TARGET_LOCAL_FOLDER/.COMPSs
```

Whenever you wish to unmount the sshfs directory please run:

```
compss@bsc:~$ sudo umount TARGET_LOCAL_FOLDER/.COMPSs
```

4.2 Execution

Access the COMPSs Monitor through its webpage (<http://localhost:8080/compss-monitor> by default) and log in with the *TARGET_LOCAL_FOLDER* to enable the COMPSs Monitor for MareNostrum.

Please remember that to enable **all** the COMPSs Monitor features applications must be ran with the *-m* flag. For further information please check the *COMPSs User Manual: Application Execution* available at our webpage <http://compss.bsc.es>.

Figure 1 illustrates how to login and Figure 2 shows the COMPSs Monitor main page for a MareNostrum application.

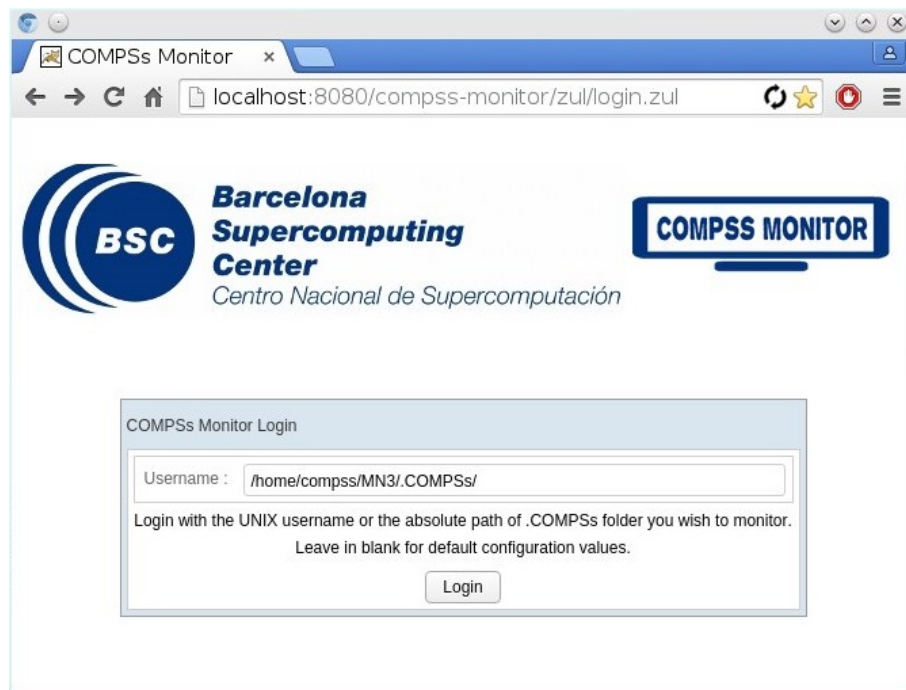


Figure 1: COMPSs Monitor login for MareNostrum

Applications

2435288

Refresh

Resources information

Tasks information

Current tasks graph

Complete tasks graph

Load chart

Runtime log

Execution Information

Statistics

Status	Resource Name	CPU Computing Units	GPU Computing Units	FPGA Computing Units	OTHER Computing Units	Memory Size	Disk Size	Provider	Image	Running Actions
	s04/2b83	16	-	-	-	28.0 GB	-GB	-	-	195 203 387 395 355 443 411 403 379 324 363 427 435 371 419 348
	s04/2b80	16	-	-	-	28.0 GB	-GB	-	-	226 210 218 275 315 299 339 307 283 259 267 331 243 251 235 292

Figure 2: COMPSs Monitor main page for a test application at MareNostrum

Please find more details on the COMPSs framework at
`http://compss.bsc.es`