

# TOTCSI

## Configuration for Simulation

Igor Jurkowski & Radosław Lazarz

September 12, 2012

## Contents

<b>1</b>	<b>How to start configuration</b>	<b>1</b>
<b>2</b>	<b>Parameter explanations</b>	<b>2</b>
2.1	General parameters . . . . .	2
2.1.1	config.task_type . . . . .	2
2.1.2	config.cmssw_dir . . . . .	2
2.1.3	config.map_output . . . . .	2
2.1.4	config.reduce_output . . . . .	2
2.1.5	config.save_using_pool . . . . .	2
2.2	TOTCSI workspace parameters . . . . .	3
2.2.1	config.workspace.create . . . . .	3
2.2.2	config.workspace.map_split_dir . . . . .	3
2.2.3	config.workspace.reduce_split_dir . . . . .	3
2.2.4	config.workspace.root_dir . . . . .	3
2.2.5	config.workspace.shell_template . . . . .	3
2.3	Input parameters . . . . .	4
2.3.1	config.input_config.map_path . . . . .	4
2.3.2	config.input_config.reduce_path . . . . .	4
2.4	Simulation parameters . . . . .	4
2.4.1	config.simulation.number_of_events . . . . .	4
2.4.2	config.simulation.number_of_jobs . . . . .	4
2.5	bsub parameters . . . . .	4
2.5.1	Default bsub parameters for map and reduce . . . . .	4
2.5.1.1	config.bsub.pool_size . . . . .	4
2.5.1.2	config.bsub.swap_size . . . . .	5
2.5.1.3	config.bsub.queue . . . . .	5
2.5.1.4	config.bsub.mail . . . . .	5
2.5.2	bsub parameters for map phase . . . . .	5
2.5.2.1	config.bsub_map.pool_size . . . . .	5
2.5.2.2	config.bsub_map.swap_size . . . . .	5
2.5.2.3	config.bsub_map.queue . . . . .	6
2.5.2.4	config.bsub_map.mail . . . . .	6
2.5.3	bsub parameters for reduce phase . . . . .	6
2.5.3.1	config.bsub_reduce.pool_size . . . . .	6
2.5.3.2	config.bsub_reduce.swap_size . . . . .	6
2.5.3.3	config.bsub_reduce.queue . . . . .	6
2.5.3.4	config.bsub_reduce.mail . . . . .	7
<b>3</b>	<b>Configuration example</b>	<b>7</b>

## 1 How to start configuration

Use the following lines at the beginning of your configuration file in order to import configuration class with all default options.

```
from totcsi.totcsi_configuration_logic import *  
config = TOTCSIConfiguration()
```

## 2 Parameter explanations

### 2.1 General parameters

#### 2.1.1 config.task\_type

Type of task to be run

Values:

- "Reconstruction"
- "Simulation"

*# Default:*

```
config.task_type = "Reconstruction"
```

#### 2.1.2 config.cmssw\_dir

CMSSW location (full path), beginning with "/"

Has to be given by user!

*# Example:*

```
config.cmssw_dir = "/afs/cern.ch/exp/totem/scratch... \  
...Release/raw_data_reco/rev5679/CMSSW_4_2_4"
```

#### 2.1.3 config.map\_output

Output directory for files (logs, results, configurations) for map phase.

Can be both CASTOR, EOS and AFS path.

It has to be set by user!

*# Example:*

```
config.map_output = "/castor/cern.ch/user/.../MapOutputFolder"
```

#### 2.1.4 config.reduce\_output

Output directory for files (logs, results, configurations) for reduce phase.

Can be both CASTOR, EOS and AFS path.

It has to be set by user!

*# Example:*

```
config.reduce_output = "/castor/cern.ch/user/.../ReduceOutputFolder"
```

#### 2.1.5 config.save\_using\_pool

Flag saying whether output should be saved first to pool and then copied to output location (might help with errors regarding CASTOR saving scripts) or directly to output location.

Values:

- True - save with a use of pool (Default)
- False - try to directly save to output location

*# Default:*

```
config.save_using_pool = True
```

## 2.2 TOTCSI workspace parameters

### 2.2.1 config.workspace.create

Flag specifying whether workspace directories should be created if they do not exist.

```
# Default:  
config.workspace.create = False
```

### 2.2.2 config.workspace.map\_split\_dir

Folder with map splitted configurations

Relative to root\_dir

It will contain the map configurations generated by TOTCSI.

```
# Default:  
config.workspace.map_split_dir = "map_split"
```

### 2.2.3 config.workspace.reduce\_split\_dir

Folder with reduce splitted configurations

Relative to root\_dir

It will contain the reduce configurations generated by TOTCSI.

```
# Default:  
config.workspace.reduce_split_dir = "reduce_split"
```

### 2.2.4 config.workspace.root\_dir

Base directory for all configuration files of TOTCSI

Can be both relative and full path.

```
# Default:  
config.workspace.root_dir = "."
```

### 2.2.5 config.workspace.shell\_template

Path to shell template. It is filled by TOTCSI and then submitted to computing cluster as job.

```
# Default:  
config.workspace.shell_template = "totcsi/templates/job_template.sh"
```

Default template looks like:

```
#!/bin/sh  
export RFIO_USE_CASTOR_V2=YES  
export STAGE_HOST={{stage_host}}  
export STAGE_SVCCLASS=default  
export SCRAM_ARCH=slc5_amd64_gcc434  
source /afs/cern.ch/cms/cmsset_default.sh  
cd {{CMSSW_path}}  
eval `scram runtime -sh`  
cd -
```

```
{{rm_type}} {{file_to_remove}}
rm -f {{log_to_remove}}
{{mkdir_type}} {{dir_to_create}}
cmsRun {{python_conf_file}}
{{optional_copy_command}}
```

## 2.3 Input parameters

### 2.3.1 config.input\_config.map\_path

Path to map template (production)  
Has to be given by user!

```
# Usage:
config.input_config.map_path = "/afs/.../MapItTemplate.py"
```

### 2.3.2 config.input\_config.reduce\_path

Path to reduce template (validation)  
Has to be given by user!

```
# Usage:
config.input_config.reduce_path = "/afs/.../ReduceItTemplate.py"
```

## 2.4 Simulation parameters

### 2.4.1 config.simulation.number\_of\_events

Number of events to be generated (integer number) per job  
It has to be set by user!

```
# Example:
config.simulation.number_of_events = 1500
```

### 2.4.2 config.simulation.number\_of\_jobs

Number of jobs generating the simulation data (integer number)  
Each job generates number of events specified by config.simulation.number\_of\_events  
It has to be set by user!

```
# Example:
config.simulation.number_of_jobs = 10
```

## 2.5 bsub parameters

### 2.5.1 Default bsub parameters for map and reduce

#### 2.5.1.1 config.bsub.pool\_size

Pool size for the job (integer)

```
# Default:
config.bsub.pool_size = 30000
```

### 2.5.1.2 `config.bsub.swap_size`

Swap size for the job (integer)

```
# Default:  
config.bsub.swap_size = 2000
```

### 2.5.1.3 `config.bsub.queue`

Type of the queue to be used

Values (sorted by decreasing priority) ::

- 8nm - 8 minutes of CPU time available on cluster
- 1nh - 1 hour of CPU time available on cluster
- 8nh - 8 hours of CPU time available on cluster
- 1nd - 1 day of CPU time available on cluster
- 2nd - 2 days of CPU time available on cluster
- 1nw - 1 week of CPU time available on cluster
- 2nw - 2 weeks of CPU time available on cluster

For more information run the `bqueues` command on `lxplus`.

```
# Default:  
config.bsub.queue = "1nh"
```

### 2.5.1.4 `config.bsub.mail`

Should the mail be sent after the job finishes

```
# Default:  
config.bsub.mail = False
```

## 2.5.2 `bsub` parameters for map phase

### 2.5.2.1 `config.bsub_map.pool_size`

Pool size for the job (integer) for map phase (overrides `config.bsub` parameter)

```
# Default:  
config.bsub_map.pool_size = 30000
```

### 2.5.2.2 `config.bsub_map.swap_size`

Swap size for the job (integer) for map phase (overrides `config.bsub` parameter)

```
# Default:  
config.bsub_map.swap_size = 2000
```

### 2.5.2.3 config.bsub\_map.queue

Type of the queue to be used for map phase (overrides config.bsub parameter)

Values (sorted by decreasing priority) ::

- 8nm - 8 minutes of CPU time available on cluster
- 1nh - 1 hour of CPU time available on cluster
- 8nh - 8 hours of CPU time available on cluster
- 1nd - 1 day of CPU time available on cluster
- 2nd - 2 days of CPU time available on cluster
- 1nw - 1 week of CPU time available on cluster
- 2nw - 2 weeks of CPU time available on cluster

For more information run the bqueues command on lxplus.

```
# Default:  
config.bsub_map.queue = "1nh"
```

### 2.5.2.4 config.bsub\_map.mail

Should the mail be sent after the job finishes in map phase (overrides config.bsub parameter)

```
# Default:  
config.bsub_map.mail = False
```

## 2.5.3 bsub parameters for reduce phase

### 2.5.3.1 config.bsub\_reduce.pool\_size

Pool size for the job (integer) for reduce phase (overrides config.bsub parameter)

```
# Default:  
config.bsub_reduce.pool_size = 30000
```

### 2.5.3.2 config.bsub\_reduce.swap\_size

Swap size for the job (integer) for reduce phase (overrides config.bsub parameter)

```
# Default:  
config.bsub_reduce.swap_size = 2000
```

### 2.5.3.3 config.bsub\_reduce.queue

Type of the queue to be used for reduce phase (overrides config.bsub parameter)

Values (sorted by decreasing priority) ::

- 8nm - 8 minutes of CPU time available on cluster
- 1nh - 1 hour of CPU time available on cluster
- 8nh - 8 hours of CPU time available on cluster

- 1nd - 1 day of CPU time available on cluster
- 2nd - 2 days of CPU time available on cluster
- 1nw - 1 week of CPU time available on cluster
- 2nw - 2 weeks of CPU time available on cluster

For more information run the `bqueues` command on `lxplus`.

```
# Default:
config.bsub_reduce.queue = "1nh"
```

### 2.5.3.4 config.bsub\_reduce.mail

Should the mail be sent after the job finishes in reduce phase (overrides `config.bsub` parameter)

```
# Default:
config.bsub_reduce.mail = False
```

## 3 Configuration example

```
from totcsi.totcsi_configuration_logic import *ex

config = TOTCSIConfiguration()

config.task_type = "Simulation"
config.cmssw_dir = "/afs/cern.ch/i/ijurkows/cmssw/CMSSW_4_2_4"

config.map_output = "/castor/cern.ch/user/i/ijurkows/output"

config.save_using_pool = False #save right away

#setting the splitter workspace
config.workspace.create = True
config.workspace.root_dir = "/afs/cern.ch/i/ijurkows/private/MyTOTCSIWorkspace"
config.workspace.map_split_dir = "SplittedMaps"
config.workspace.reduce_split_dir = "SplittedReduces"

config.input_config.map_path = "/afs/cern.ch/user/i/ijurkows/public/\
    configurations_for_map/RPElasticbeta90energy3p5TeV2200ctober2011a3_cfg.py"
config.input_config.reduce_path = "/afs/cern.ch/user/i/ijurkows/public/\
    configurations_for_reduce/valRPElasticbeta90energy3p5TeV2200ctober2011a3_cfg.py"

config.bsub.pool_size = 20000
config.bsub_map.pool_size = 26000

config.simulation.number_of_events = 2000
config.simulation.number_of_jobs = 100
```