

## Working on TGCC

---

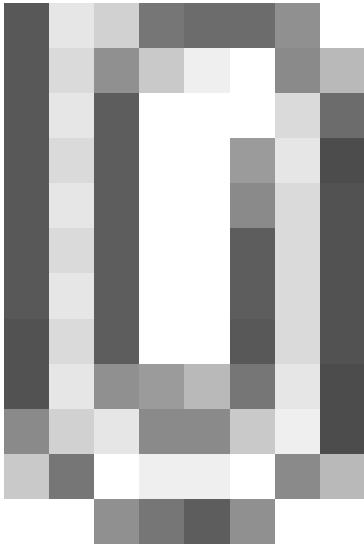
### Table of Content

<b>Working on TGCC</b>	<b>1</b>
<b>1. TGCC presentation</b>	<b>2</b>
<b>2. TGCC's machines and file systems</b>	<b>2</b>
<b>3. How to install your environment on TGCC</b>	<b>2</b>
<b>4. Project and computing needs</b>	<b>2</b>
<b>5. About file systems</b>	<b>3</b>
5.1. Quotas	3
5.2. CCCSCRATCHDIR	3
5.3. CCCWORKDIR	3
5.4. CCCSTOREDIR	3
5.5. ccc_home command to know directory complete pathname	3
5.6. Storage spaces available from ESGF/THREDDS	4
<b>6. Specific directories for projects</b>	<b>4</b>
<b>7. Specific file systems for CMIP6</b>	<b>4</b>
7.1. GENCMIP6_HOME	4
7.2. GENCMIP6_CCCWORKDIR	4
7.3. GENCMIP6_CCCSTOREDIR	5
7.4. GENCMIP6_SCRATCHDIR	5
<b>8. End-of-job messages</b>	<b>5</b>
<b>9. Simulation outputs</b>	<b>5</b>
<b>10. About password</b>	<b>5</b>
<b>11. The TGCC's machines</b>	<b>5</b>
11.1. Irene	5

## 1. TGCC presentation

■ <http://www-hpc.cea.fr/en/complexe/tgcc.htm>

## 2. TGCC's machines and file systems



## 3. How to install your environment on TGCC

- More information on the open-access website: ■ <http://www-hpc.cea.fr/en/complexe/tgcc.htm>  
Online access to the machines' users manual (you will need a TGCC login and password): ■ <https://www-tgcc.ccc.cea.fr/> .
- Online access to technical issues and news : ■ <https://www-tgcc.ccc.cea.fr/en/news/index.html>
- TGCC's machine is [Irene](#) (Bull Skylake)
- Note: the **\$HOME/.snapshot** directory contains hourly, daily, and weekly backups of your \$HOME files.

It is important to take the time to install a comfortable and efficient environment.

We suggest the user to use the igcmg environment (in bash) with a copy of the bashrc in his HOME.

```
ryyy999@irene: cp ~/igcmg/MachineEnvironment/irene/bashrc ~/.bashrc
```

Additionally, you can copy and complete the example of bashrc\_irene file to create your favorite environment (alias, module load ...). Don't forget to use it in .bashrc.

```
ryyy999@irene: cp ~/igcmg/MachineEnvironment/irene/bashrc_irene ~/.bashrc_irene
ryyy999@irene: vi ~/.bashrc # to point your own .bashrc_irene
```

WARNING : if you have a ~/.profile file, it's better to remove it to avoid any problem during the execution of a simulation with libGCM

In this environment is specified:

the path to the compiler tool `fcm` and to the `rebuild` tool which recombines output files from a parallel model:

```
export PATH=$(ccc_home -u igcmg)/Tools/fcm/bin:$(ccc_home -u igcmg)/Tools/irene/bin:$PATH
```

- the load of modules giving access to computing or post processing libraries and tools needed on our platform (done in `ccc_home -u igcmg/MachineEnvironment/irene/env_atlas_irene`).

## 4. Project and computing needs

- To find out the computing time used by the projects you are involved in (daily update):

```
ryyy999@irene: ccc_myproject
```

- When you will create a job you need to specify in the header the project from which you will use computing time:

```
#MSUB -A genxxx
```

## 5. About file systems

### 5.1. Quotas

To check the available and used storage capacities of HOME, CCCSCRATCHDIR, CCCWORKDIR and CCCSTOREDIR:

```
ryyy999@irene: ccc_quota
```

On the Irene machine this command will also return the space used by scratch (a specificity of the Irene machine).

This command has been improved and gives a lot of information : quotas and usage of shared space, type and duration of exception.

### 5.2. CCCSCRATCHDIR

The \$CCCSCRATCHDIR directory is often cleaned and only files that are less than 40 days are stored.

### 5.3. CCCWORKDIR

The \$CCCWORKDIR directory corresponds to the \$WORKDIR directory on Irene. It is large but its content is not backed up. Don't forget to do a backup (tar) for important directories.

### 5.4. CCCSTOREDIR

To manipulate the files in /ccc/store a few commands are useful:

```
# Demigrate a list of files on CCCSTOREDIR, see also "ccc_hsm -h"
ccc_hsm get $CCCSTOREDIR/FILE1 $CCCSTOREDIR/FILE2 ...

# Demigrate recursively the files from a CCCSTOREDIR directory, see also "ccc_hsm -h"
ccc_hsm get -r $CCCSTOREDIR/DIRECTORY

# Find out the used space on CCCSTOREDIR
cd $CCCSTOREDIR ; find . -printf "%y %s %p \n" | \
    awk '{ SUM+=$2 } END {print "SUM " SUM/1000000 " Mo " SUM/1000000000 " Go" }'

# or use --apparent-size with du :
du -sh --apparent-size
```

### 5.5. ccc\_home command to know directory complete pathname

ccc\_home could help you to find directory complete pathname for an other user or for you .

```
>ccc_home -h
ccc_home: Print the path of a user directory (default: home directory).
usage: ccc_home [-H | -s | -t | -W | -x | -A | -a | -n] [-u user] [-d datadir]
              [-h, --help]

-H, --home          : (default) print the home directory path ($HOME)
-s, -t, --cccscratch : print the CCC scratch directory path ($CCCSCRATCHDIR)
```

```

-X, --ccchome      : print the CCC nfs directory path ($CCCHOMEDIR)
-W, --cccwork     : print the CCC work directory path ($CCCWORKDIR)
-A, --cccstore    : print the CCC store directory path ($CCCSTOREDIR)
-a, --all         : print all paths
-u user           : show paths for the specified user instead of the current user
-d datadir        : show paths for the specified datadir
-n, --no-env      : do not load user env to report paths
-h, --help        : display this help and exit

> ccc_home -A -u ryyy999
/ccc/store/cont003/genXXX/ryyy999

```

## 5.6. Storage spaces available from ESGF/THREDDS

To store a file for the first time on esgf/thredds, you must ask for esgf/thredds write access by mail to the TGCC hotline access : [hotline.tgcc@cea.fr](mailto:hotline.tgcc@cea.fr). On Irene, files available on \$CCCWORKDIR are candidates to be available from ESGF/THREDDS :

- use `thredds_cpcommand`
- files will be hardlinked here : `/ccc/work/cont003/thredds/login`

From a server web : [https://vesg.ipsl.upmc.fr/thredds/catalog/work\\_thredds/catalog.html](https://vesg.ipsl.upmc.fr/thredds/catalog/work_thredds/catalog.html)

More information about output data available from ESGF/THREDDS [here](#).

## 6. Specific directories for projects

3 different spaces are available for each project : HOMEDIR, WORKDIR, SCRATCHDIR, STOREDIR.

## 7. Specific file systems for CMIP6

For `gencmip6` project, and only for it, 3 more file systems and 4 more directories are available. Phase 1 have been installed in april 2016. Phase 2 and Phase 3 will come later in 2017 and 2018.

To use them, in interactive mode, you have to do : `module load datadir/gencmip6`.

Since `libGCM_v2.8.1`, if you set your project to `gencmip6/devcmip6`, they are automatically used in place of usual `HOME`, `CCCWORKDIR`, `CCCSTOREDIR` and `CCCSCRATCHDIR` : `module switch dfldatadir dfldatadir/gencmip6` called from `libGCM`.

### 7.1. GENCMIP6\_HOME

- 50 TB
- `gencmip6` group quota
- dedicated to sources and scripts
- strongly recommended for CMIP6 sources and simulations scripts
- regular snapshot are taken by the system. See `$GENCMIP6_HOME/.snapshot` Attention : you need an interactive connexion on a compute node :

```

> ccc_mprun -s -p standard -A devcmip6 -T 1800 -Q test
> cd
> . .bash_login
> cd .snapshot
> ls -l
total 44
drwxr-sr-x. 13 xxx gencmip6 4096 Dec 17 09:47 daily.2017-02-07_0010
drwxr-sr-x. 13 xxx gencmip6 4096 Dec 17 09:47 daily.2017-02-08_0010
...

```

### 7.2. GENCMIP6\_CCCWORKDIR

- 2.5 PB in phase 1, 5 PB in phase 2
- gencmip6 group quota
- dedicated to small output files (ATLAS, MONITORING)
- available through <https://esgf.extra.cea.fr> following work\_thredds
- no backup

### 7.3. GENCMIP6\_CCCSTOREDIR

- 2.5 PB in phase 1, 5 PB in phase 2 and 14 PB on tape in phase 3
- gencmip6 group quota
- dedicated to large (more than 1GB) output files (Output, Analyse)
- available through <https://esgf.extra.cea.fr> following store\_thredds
- linked with HSM (tapes)

### 7.4. GENCMIP6\_SCRATCHDIR

- same file system as GENCMIP6\_CCCWORKDIR
- used during batch execution (RUN\_DIR) and erased at the end of the execution
- regular cleaning after 40 days

## 8. End-of-job messages

To receive the end-of-job messages sent by the job itself: end of simulation, error,... you must specify your address in the \$HOME/.forward file.

News in June 2018 : On Irene you have to duplicate a .forward for each project HOME.

## 9. Simulation outputs

Final simulation outputs are stored in \$CCCSTOREDIR/IGCM\_OUT and on \$CCCWORKDIR/IGCM\_OUT regarding the ATLAS and MONITORING directories.

The esgf/thredd server on TGCC are available via:

- <https://vesg.ipsl.upmc.fr/thredds/catalog/catalog.html>, click on store, click on your login or directly : <https://vesg.ipsl.upmc.fr/thredds/catalog/store/YOURLOGIN/catalog.html> and ATM (or an other component) for Analyse files (TS or SE)
- <https://vesg.ipsl.upmc.fr/thredds/catalog/catalog.html>, click on work, click on your login or directly : <https://vesg.ipsl.upmc.fr/thredds/catalog/work/YOURLOGIN/catalog.html> for ATLAS and MONITORING

## 10. About password

ccc\_password\_expiration helps you to know expiration date of your password. Currently password have to be changed one time per year.

```
> ccc_password_expiration
Password for xxxxx@USERS-CCRT.CCC.CEA.FR: Pppppppppp
Your password will expire in 70 days on Fri Nov 22 08:42:59 2013
> ccc_password_expiration -h
Usage: ccc_password_expiration [username[@realm]]
```

## 11. The TGCC's machines

### 11.1. [Irene](#)

See the documentation for [Irene](#).