

Blade Cluster





14 nodes DELL PowerEdge M600

2 QuadCore CPU (12MB cache, 3 Ghz)

16 GB RAM

Scheduler Sun Grid Engine

Queue priority

Projects



Blade Cluster

scheduler commands

qsub submit a batch job to Sun Grid Engine

-cwd launches job from current dir (output/error file)

-m b|e|a|s|n selects events for notification email

b (beginning)

e (end)

a (abort)

s (suspend)

n (never) #default

qstat show the status of Sun Grid Engine jobs and queues

-f shows full queue list

-j [num] shows full details of job #num



Blade Cluster

scheduler commands

qhost show the status of Sun Grid Engine hosts, queues, jobs

<no params> lists host list and their configuration

-q lists available queues on hosts

j lists running jobs on hosts

qdel delete Sun Grid Engine jobs from queues

-u <myaccount> deletes all my enqueued jobs

<job#list> deletes all jobs in the list



Shared Disk Space HDA dedicated volume

Course Root Dir /nfsd/hda

Datasets Root Dir /nfsd/hda/DATASETS

Personal Root Dir /nfsd/hda/<your_DEI_account>

Your personal root dir could be organized with many subfolders:

/nfsd/hda/<your_DEI_account>/my_prj_1

/nfsd/hda/<your_DEI_account>/my_prj_2

/nfsd/hda/<your_DEI_account>/my_group_prj

Respectively containing files of your personal project 1 and 2 and of your HDA final project.



git + sophisticated and powerful versioning system

+ distributed

+ main repository

git [local process] clone / add / delete / commit

(your daily work – could be offline)

[remote process] push / pull

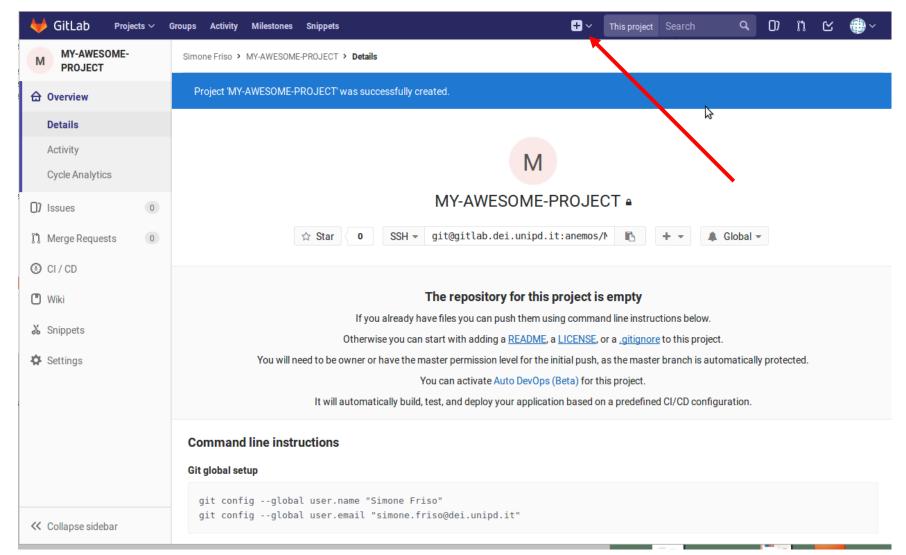
(toward the cloud, once you are online again)

collaborative work merge

Gitlab DEI git main repository



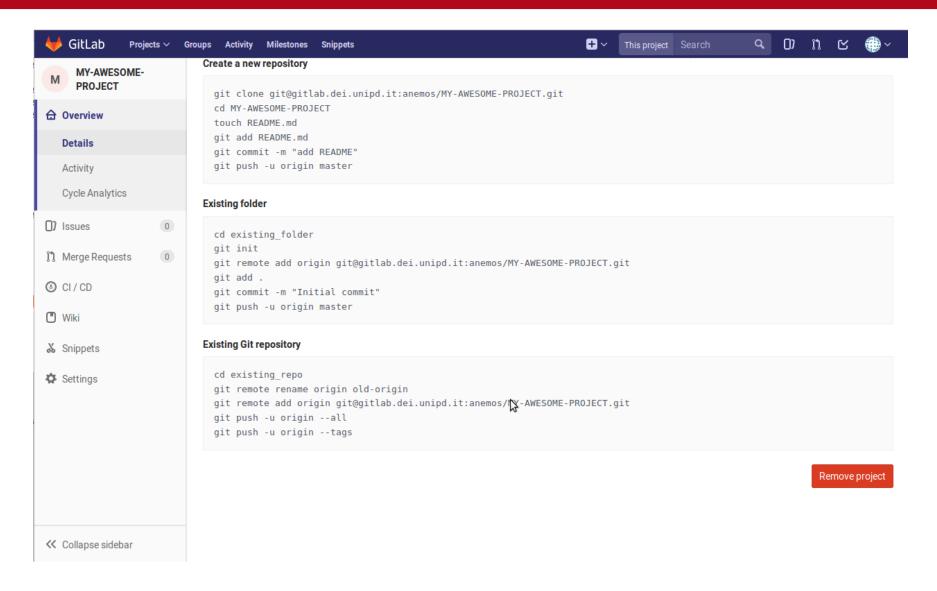
web based repo manager + issue tracker and wiki features





gitlab

a git cloud repository manager provided by DEI





Putting everything together

Login to Blade ssh <your_DEI_account>@login.dei.unipd.it

Your Dataset Root Dir /nfsd/hda/DATASETS/<my_dataset>

(NB Your dataset is possibly shared – i.e. used read only by many groups - saving disk space)

Project Root Dir /nfsd/hda/<your_DEI_account>/my_group_prj

In your project root dir you must have:

- 1) Main execution file of your code (.py) + dep
- 2) Script to submit your job to scheduler (.job)
- 3) git versioning (ignore for .job files)

Example 1 "Hello World!" from "dataset" file

Example 2 Tensorflow tutorial



Putting everything together

Example 1 Printing "Hello World!" to standard output reading

the string from a "dataset" file

Input files first.py # main Python file

firstsubmit.job # submission script

qsub job #1643278 submitted

Output files firstsubmit.job.e1643278 # error file (empty)

firstsubmit.job.o1643278 # output file (stdout dump)

anemos@login:/nfsd/hda/anemos\$ls -la total 20

drwxr-sr-x. 2 anemos hda 4096 May 23 16:47 ./

drwxrwsr-x. 44 root hda 4096 May 21 10:04 ../

-rw-r--r--. 1 anemos hda 96 May 18 14:15 **first.py**

-rw-r--r-. 1 anemos hda 490 May 18 14:17 firstsubmit.job

-rw-r--r-. 1 anemos hda 0 May 23 16:47 **firstsubmit.job.e1643278**

-rw-r--r-. 1 anemos hda 39 May 23 16:47 **firstsubmit.job.o1643278**