

Risorse di calcolo indivise per CSN2 analysis facility

G. Mazzitelli, C. Pellegrino et al. CSN2 Novembre 25



Shared pledge for the astro-particle community

CSN2 analysis facility

README.md



Shared pledge for the astro-particle community:

- One objective of the PSGE project—besides the generalization of the [Computing Model \(CM\)](#) through the CYGNO use case—is to standardize access to pledged resources and optimize their utilization for small- and medium-scale experiments, thereby enabling efficient resource allocation and amortizing potential surges in computational demand.
- To support this objective, an Identity and Access Management (IAM) service for authentication and authorization—currently limited to users with INFN platform credentials—has been created and is managed by the CSN2 Computing Working Group ([GdL Calcolo](#)): iam.psge.cloud.infn.it
- Through the [ui.psge.cloud.infn.it](#) interface, notebook, VS Code, remote desktop, and remote SSH login environments are available, enabling users to develop, run, and manage workloads directly on PSGE resources; see [UI howto](#)
- The user interface includes a minimal software stack for analysis and access to CVMFS storage. Jobs can be submitted via the HTCondor CE endpoint tier2-ce2.lnf.infn.it; see the [PSGE Guide to Job Submission](#) for step-by-step instructions.
- [rclone](#) is available to mount experiment storage

- info: <https://github.com/ICSC-Spoke2-repo/PSGE/tree/main/ASTROCM>

HOWTO login

possono accedere solo utenti riconosciti su AAI

- tutti gli utenti devono essere **riconosciti dall'INFN** (ospitalità virtuale LoA2) e aver quindi fatto il corso di sicurezza informatica
- prima di accedere all'user interface bisogna registrarsi sullo IAM di CSN2.
- E' essenziale dichiarare la/le **sigle** di appartenenza e il **referente per il calco/RN della sigla**.
- info: <https://github.com/ICSC-Spoke2-repo/PSGE/blob/main/ASTROCM/Guide2Services.md>

The screenshot shows a web browser window with three tabs: N-CSN2-QA, https://www.l, and https://portal. The main content area displays the CSN2 logo and the text "Welcome to PSGE". Below this, there is a "Sign in with" section featuring the INFN CCR-AAI logo and a "Local credentials" button. A "Register at PSGE" section follows, containing the text: "This is the PSGE registration page. You have been successfully authenticated, but your credentials are not yet linked to an PSGE account. To proceed with the registration please fill personal information below." It includes a link to abort the registration. The "Personal Information" form contains fields for Given name (Federico), Family name (Di Pierro), Email (federico.dipierro@ta.infn.it), Username (fdipier), and Notes (I work in the CTA experiment and the INFN CTA National Responsible is Massi Mariotti (mariotti@pd.infn.it)). At the bottom are "Register" and "Reset Form" buttons.

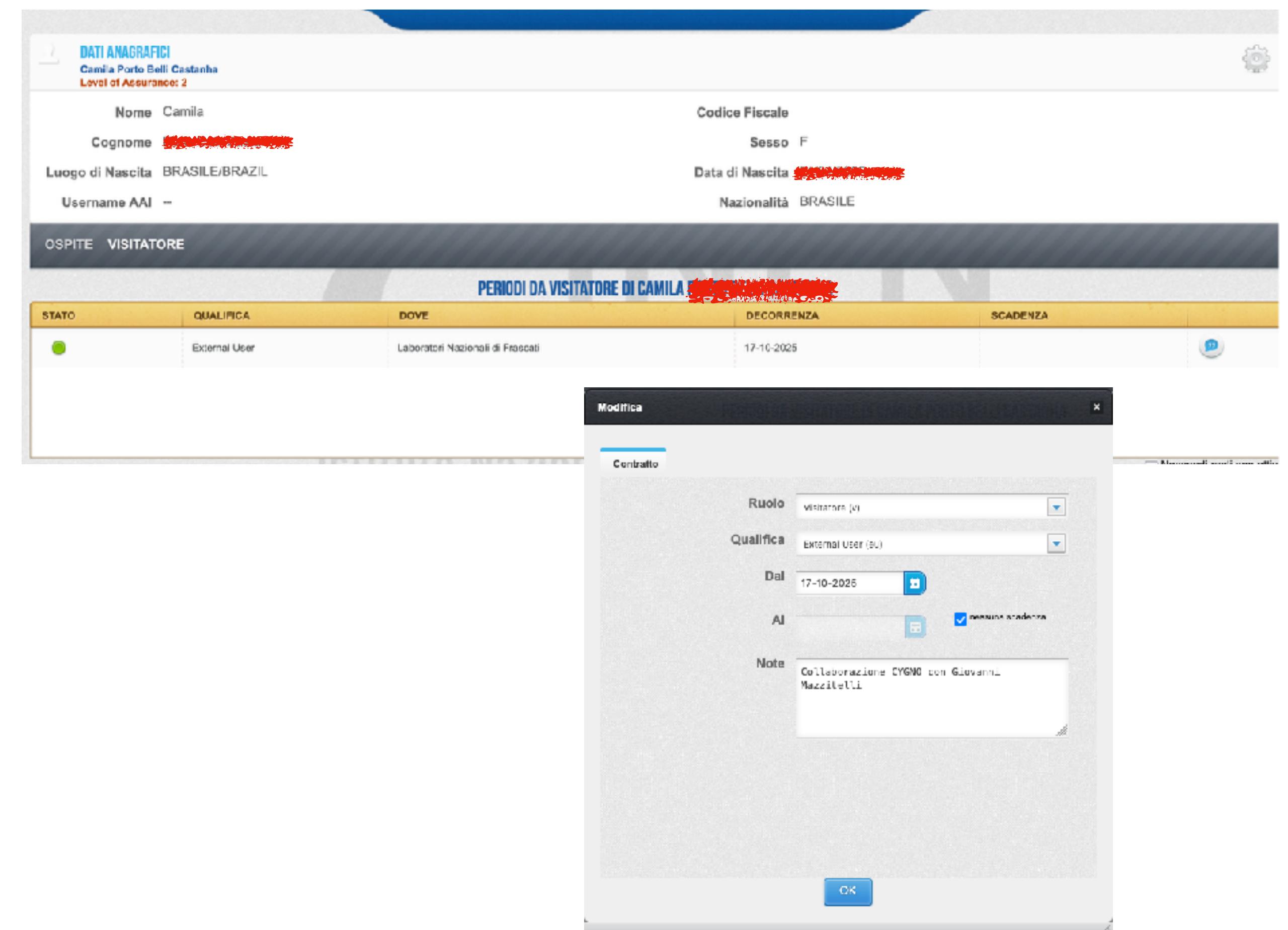
Request confirmed successfully

Your registration request has been confirmed successfully, and is now waiting for administrator approval. As soon as your request is approved you will receive a confirmation email.

[Back to Login Page](#)

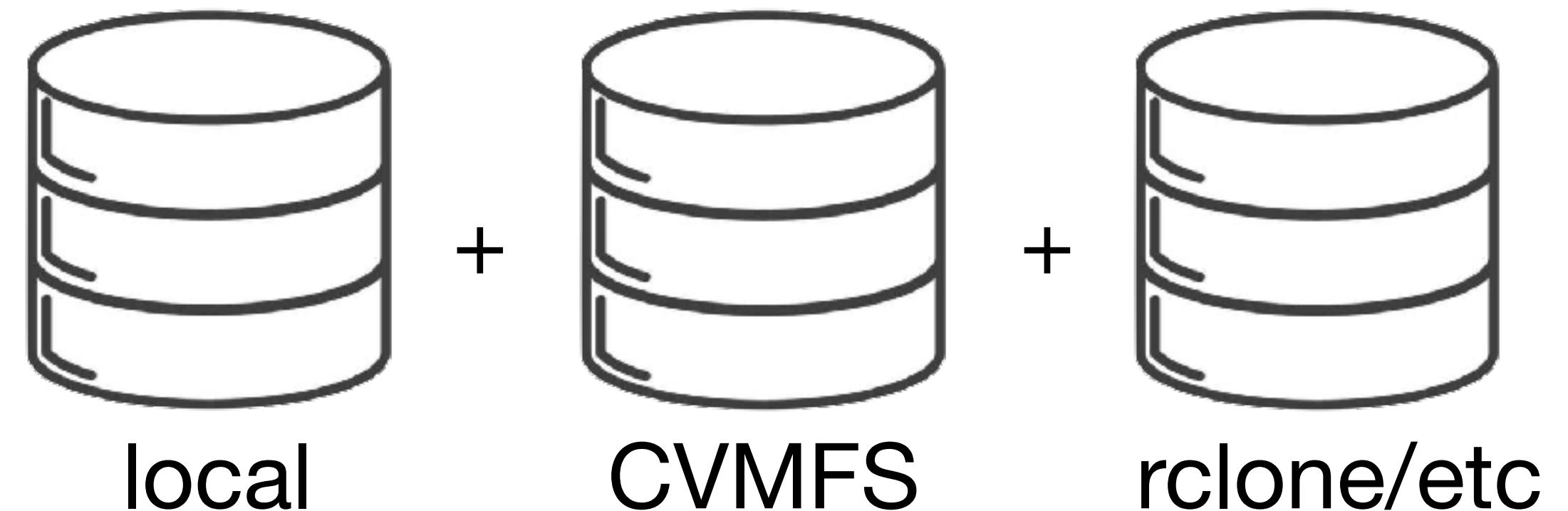
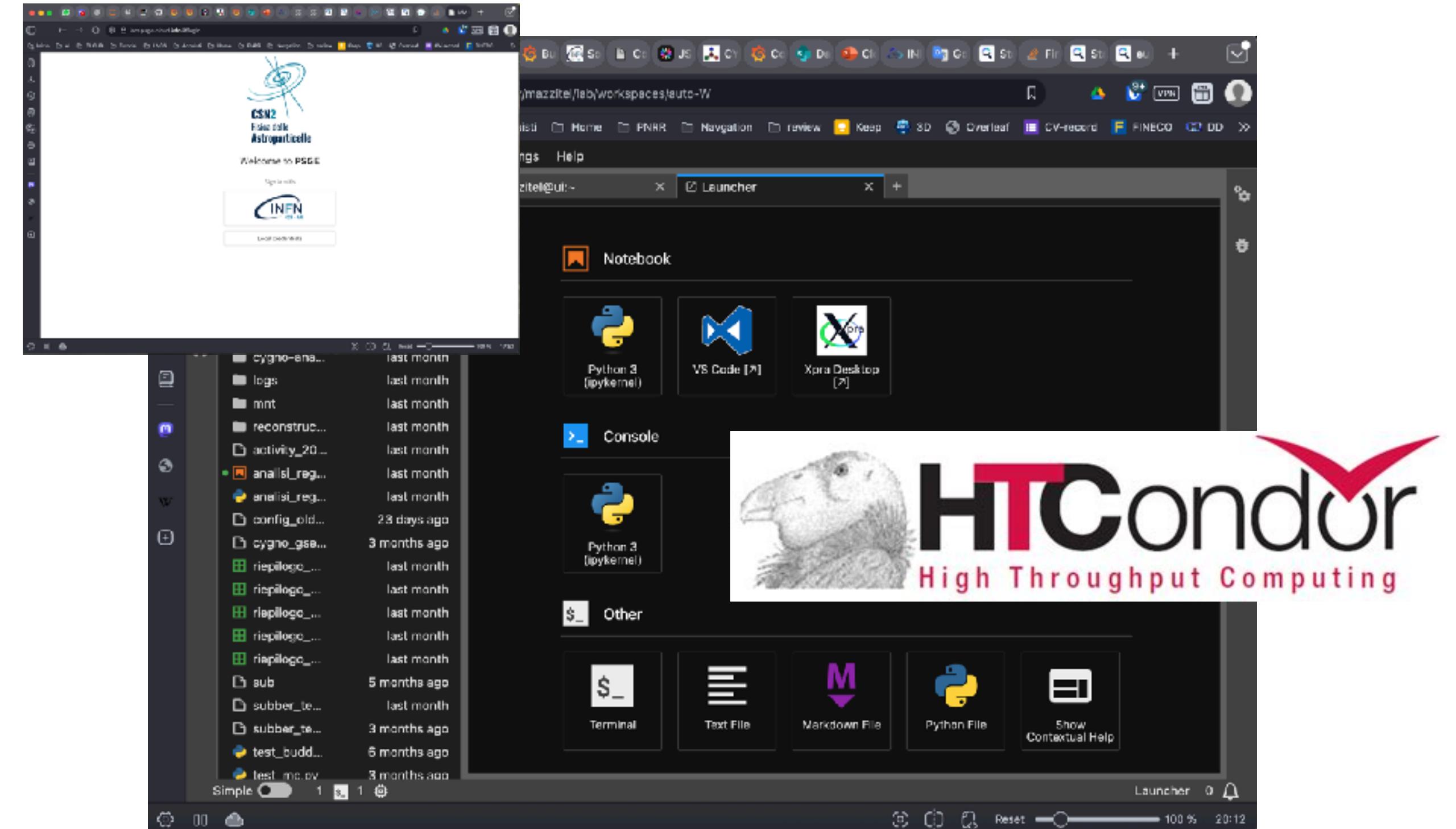
HOWTO eventuali utenti non in accesso AAI visitatori virtuali

- ci deve essere un referente che lo **ospiti virtualmente** (l'RN o il Resp. Computing)
- registrarsi: <https://signup.app.infn.it/>
- seguire il corso di sicurezza informatica e accettare le security policy su <https://userportal.app.infn.it/>
- la segreteria riconosce la persona e assegna l'LoA2 come visitatore
- per il momento (il GdL AAI sta cercando di farlo in automatico) si deve assegnare una username chiedendolo ad esempio ad un CED locale
- chiedere l'accesso alle risorse di CSN2 <https://iam-psge.cloud.cnaf.infn.it/>



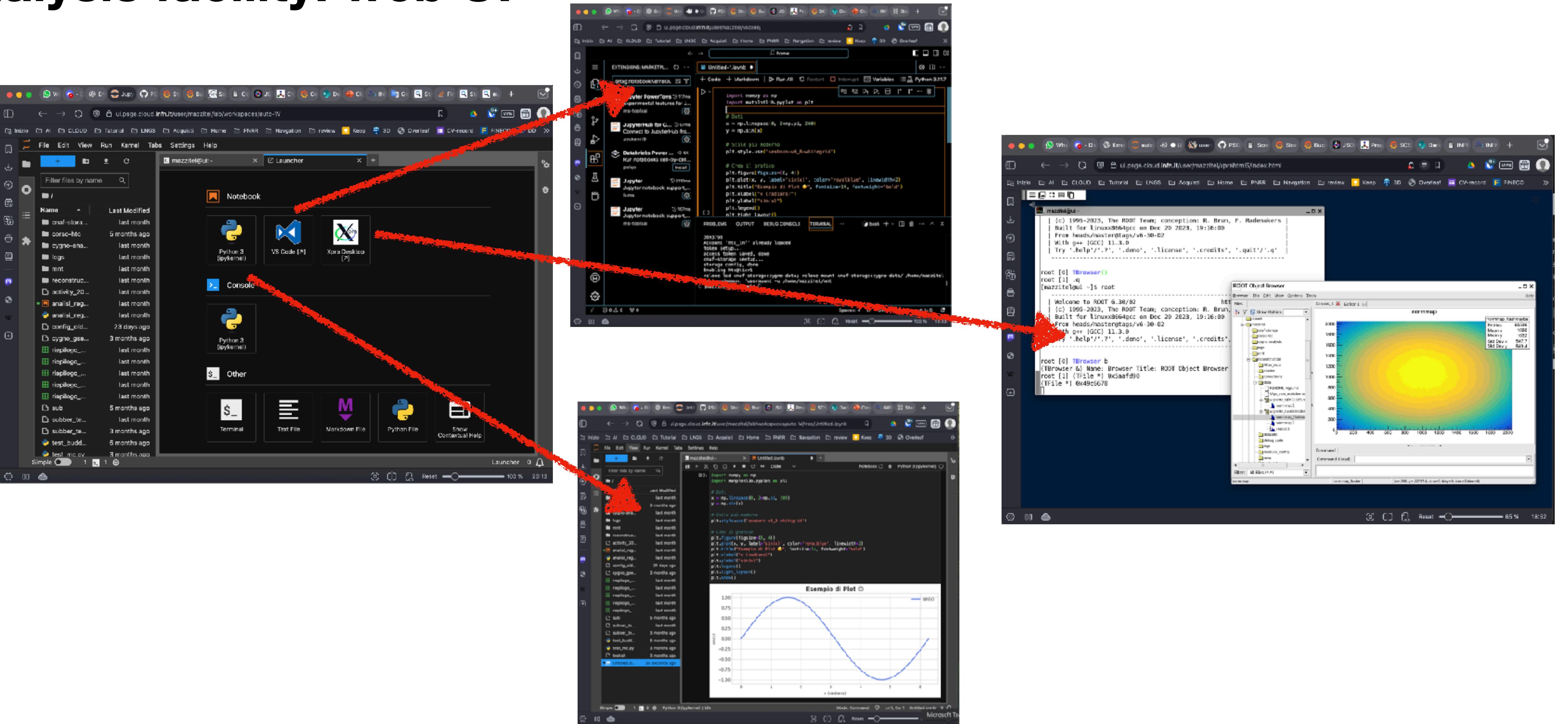
CSN2 analysis facility

- IAM PSGE per autenticazione e autorizzazione
- accesso code per il calcolo HTC - 8kHS06 - distribuite fra LNF e Catania, e...
- User Interface (web/X)/CVMFS
- storage spool di 9 TB, lifetime una settima (condor transfer data...)
- local quota di 10GB (1TB totale)
- batch system con autenticazione sul proprio storage per trasferire l'output dei job sottomessi



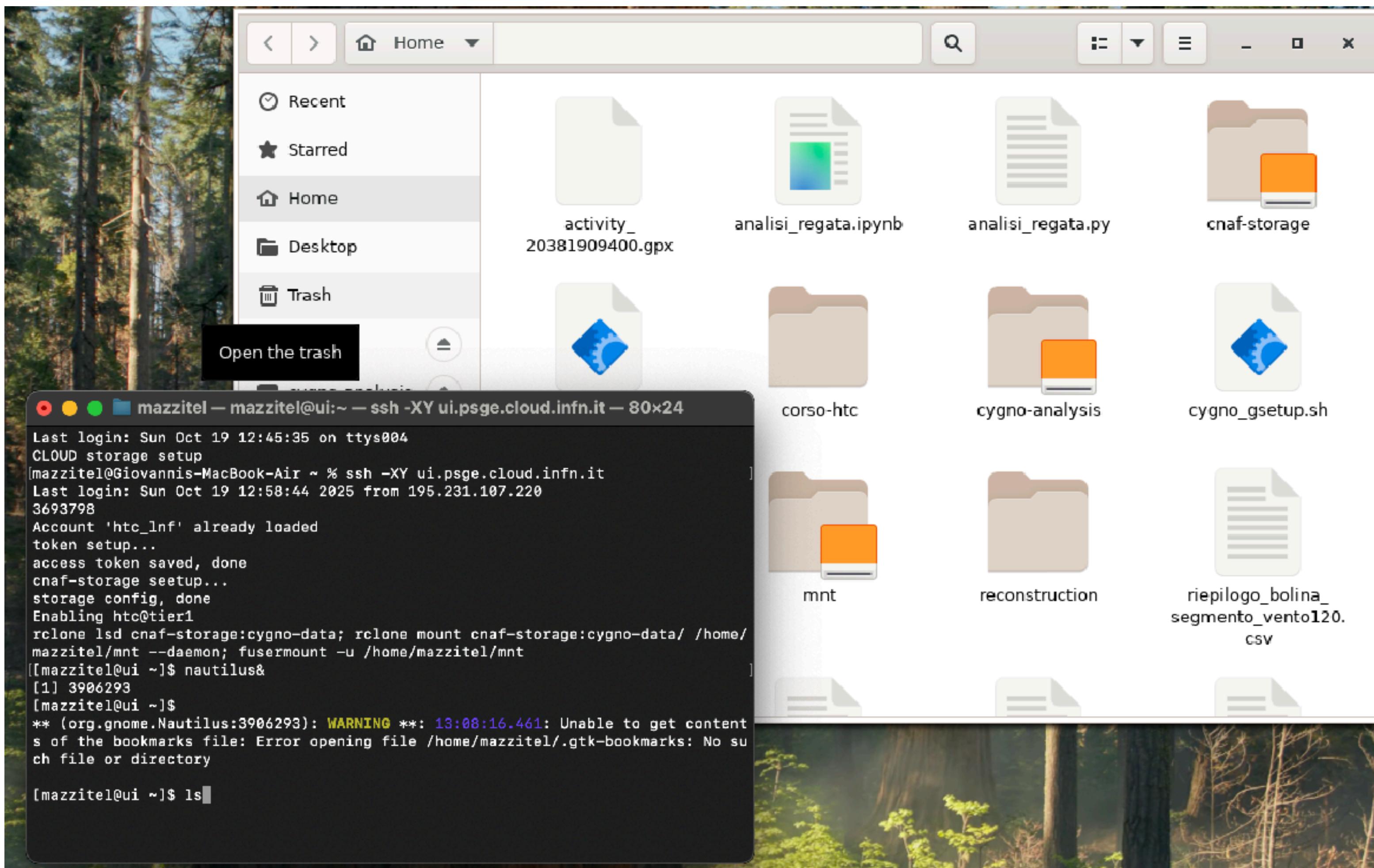
CSN2 analysis facility

analysis facility: web UI



CSN2 analysis facility

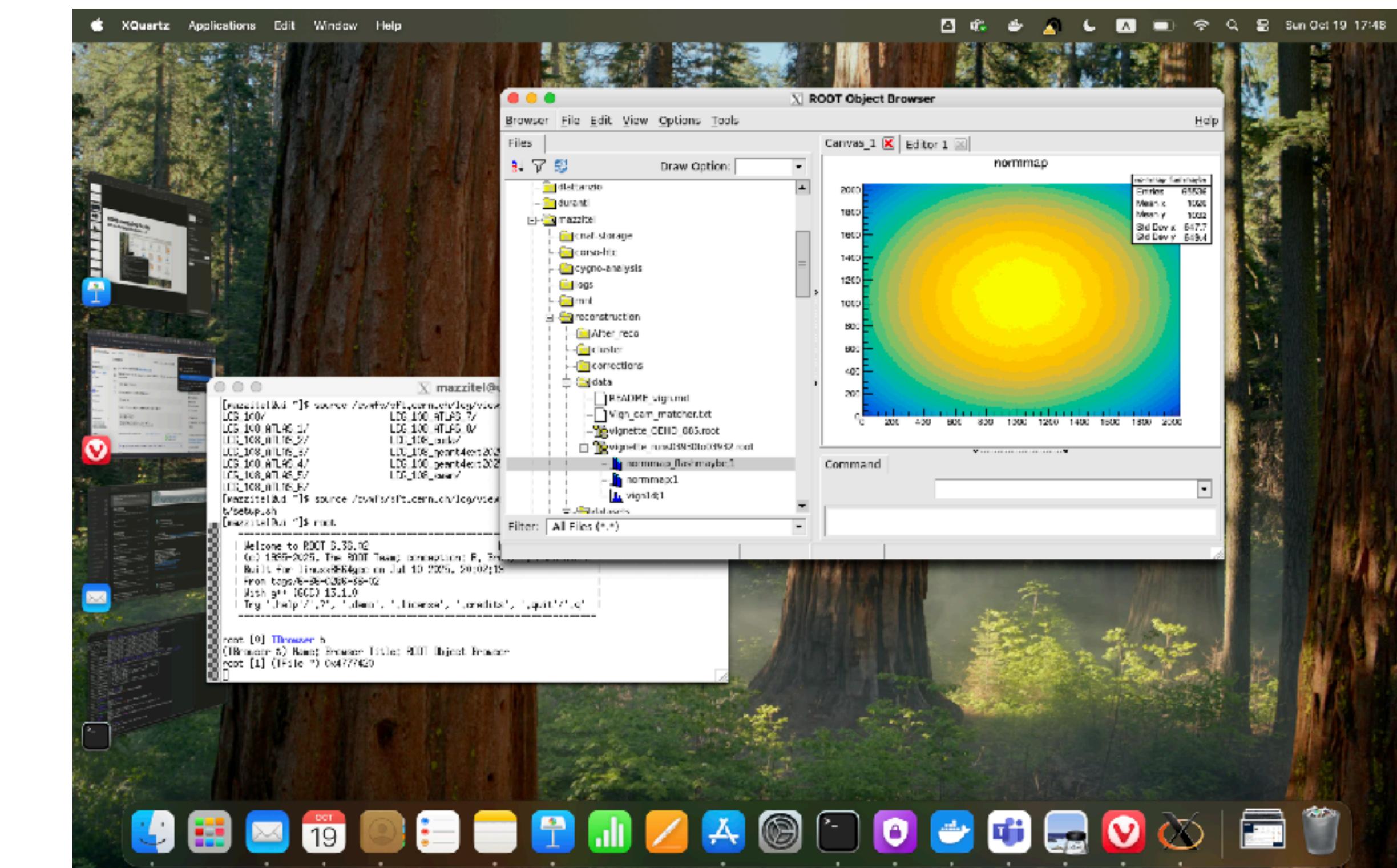
analysis facility: X UI - ssh user@ui.psge.cloud.infn.it



experiment software via CVMFS

analysis facility: read only file system per condividere il software

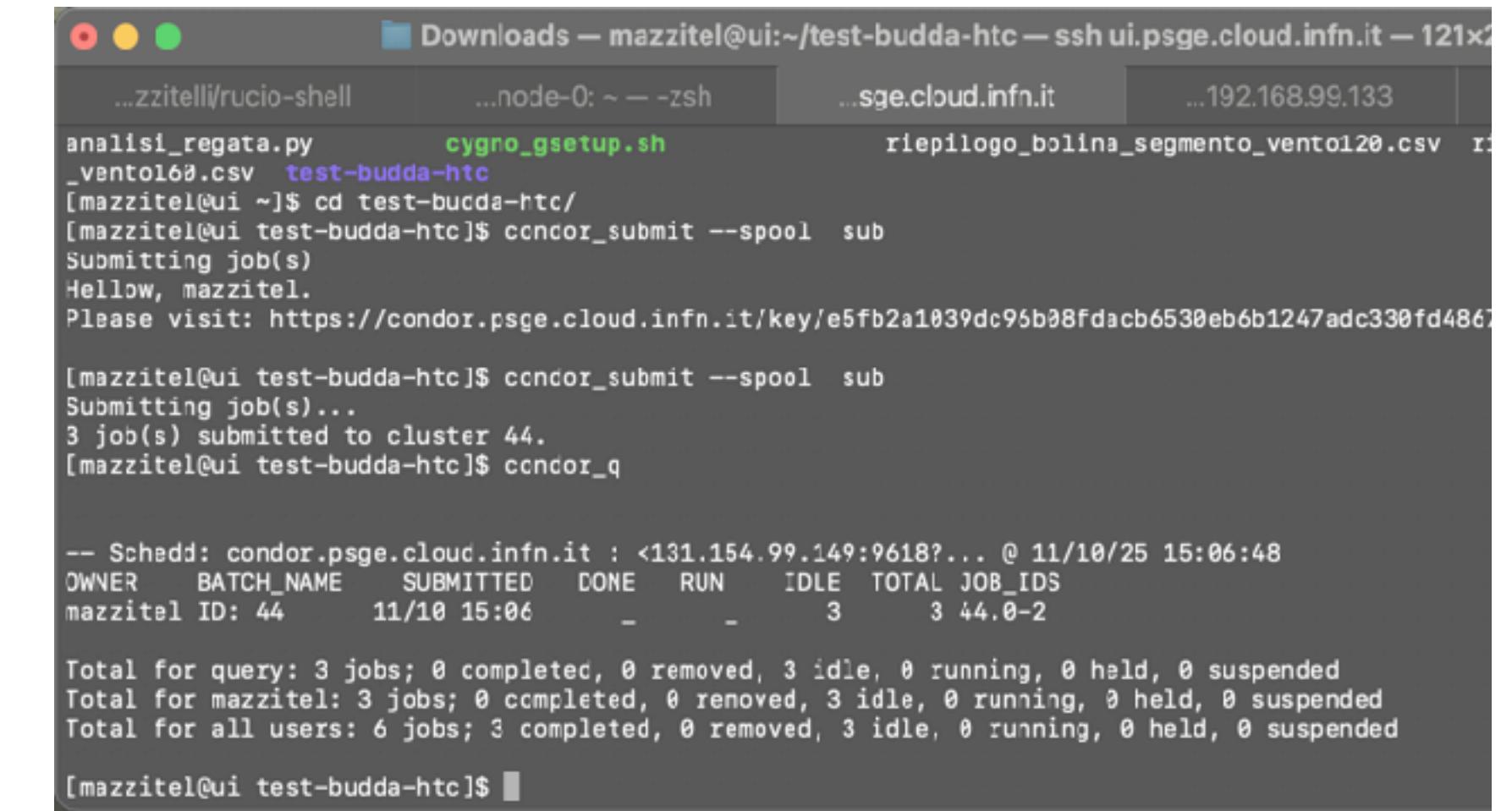
- si possono caricare i vari **software (ROOT/ GEANT, etc) dal CVMFS del CERN**
es. source /cvmfs/sft.cern.ch/lcg/views/LCG 105/x86 64-el9-gcc11-opt/setup.sh
info: <https://lcginfo.cern.ch>
- spazio e **personale** su DATA CLOUD - es:
/cvmfs/mazzitel-personalrepo.infn.it
- repository **dell'esperimento** - es:
/cvmfs/sft-cygnocern.it
- info: <https://confluence.infn.it/spaces/INFNCLOUD/pages/235864087/>
Software+Management+User+Guide



HOWTO htc

analysis facility: batch system condor

- il batch system e' accessibile via htcondor configurato di default sulla facility
- aprire in **terminale** sulla web UI o accedere via **ssh**
- **condor_q** mostra le risorse e la loro occupazione, monitorare il job.
- per sottomettere bisogna crearsi il proprio **sub** file
<https://codimd.infn.it/s/VD3RWisM6#Submitting-a-demo-job>
- sottomettere il job con **condor_submit --spool sub**
- se non si ha un proprio storage di esperimento al T1, si puo' usare la analysis facility come buffer dove accedere i dati, ci sono 9TB di disco buffer con una lifetime di una settimana, oppure possono essere ripresi attraverso **condor_transfer_data <job ID>** <https://htcondor.readthedocs.io/en/latest/users-manual/file-transfer.html> nella propria home (10GB)
- info <https://github.com/ICSC-Spoke2-repo/PSGE/blob/main/ASTROCM/GuideJobSubmission.md>

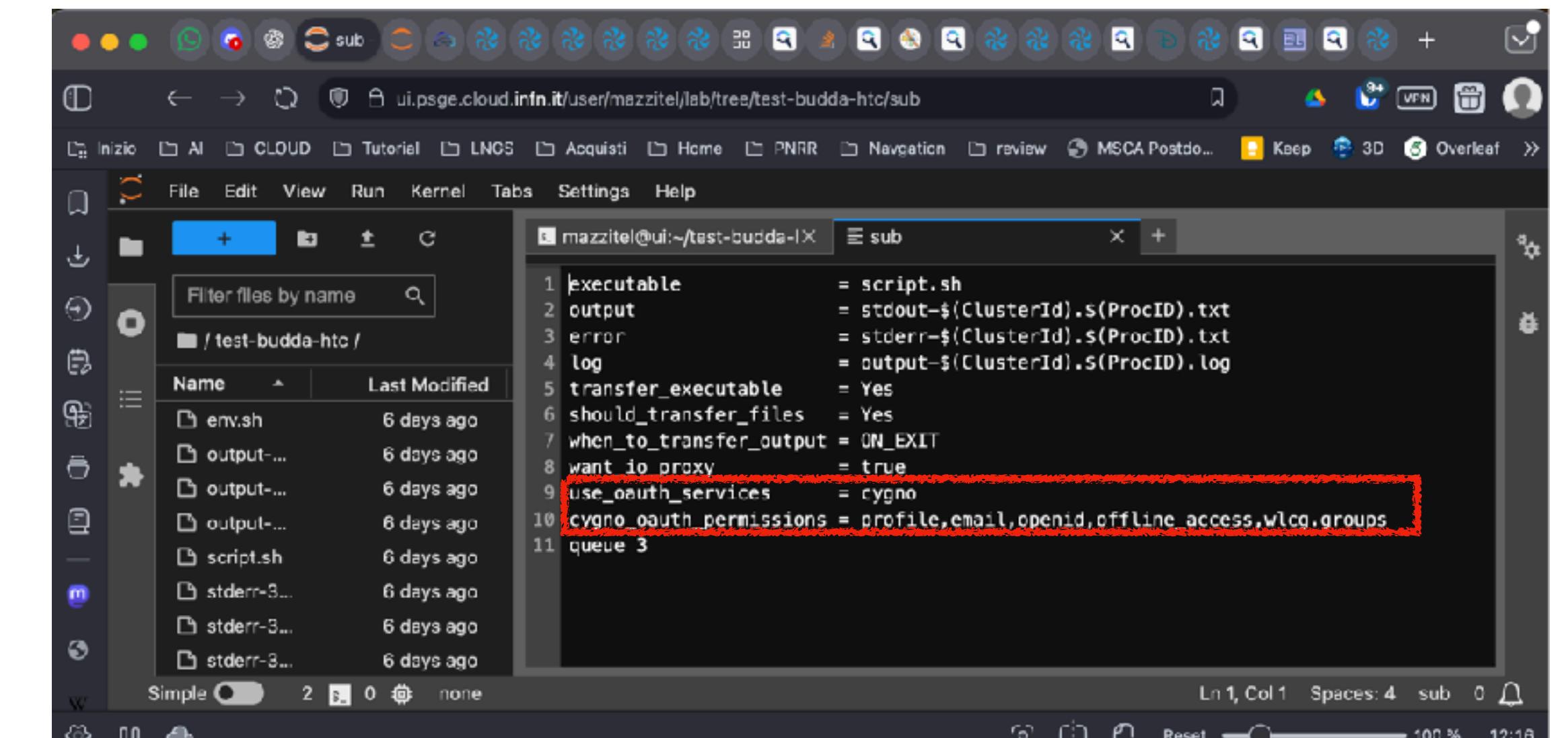


The screenshot shows a terminal window with several tabs open. The active tab displays the command `condor_submit --spool sub` being run, followed by a message indicating 3 job(s) submitted to cluster 44. Below this, the `condor_q` command is run, showing the queue status with columns for Schadd, OWNER, BATCH_NAME, SUBMITTED, DONE, RUN, IDLE, and TOTAL JOB_IDS. It lists one job for user mazzitel with ID 44, submitted at 11/10 15:06. The total for all users is 6 jobs, with 3 completed, 0 removed, 3 idle, 0 running, 0 held, and 0 suspended. The bottom of the window shows the prompt `[mazzitel@ui test-budda-htc]$`.

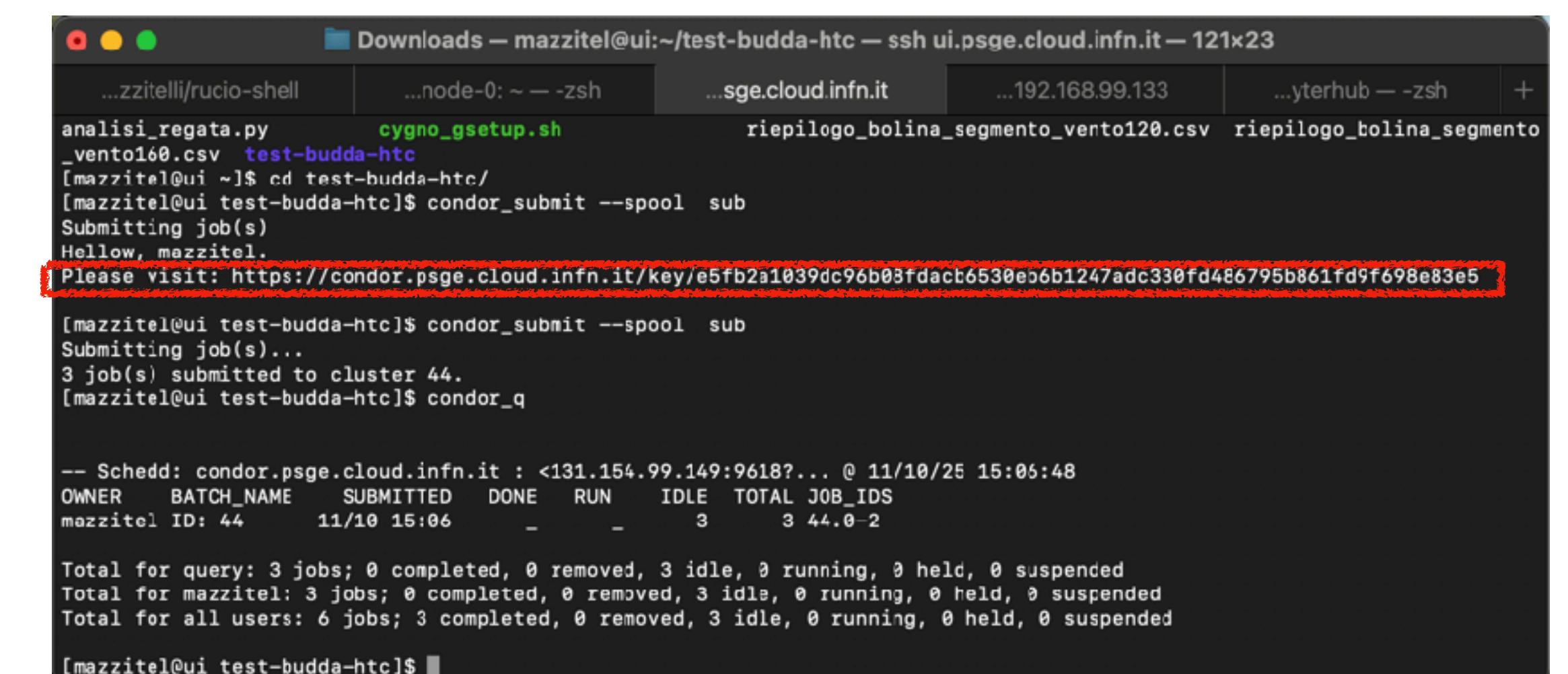
HOWTO accesso allo storage via htcondor analysis facility: batch system con token per lo storage

- creare il proprio **sub file** (vedi esempio)
- le righe importanti sono **use_oauth_services** e **cygno_oauth_permissions**, queste due variabili devono stare necessariamente prima del comando **queue**
- **use_oauth_services** lega il vostro token per lo storage al servizio e quindi dal job potete direttamente caricare il vostro output sul vostro storage di esperimento
- sottomettere una prima volta **condor_submit --spool sub** e seguire il link per autenticarsi sullo storage (questa prima sottomissione va fatta una tantum)
- sottomettere il vero job **condor_submit --spool sub**
- da dentro il vostro script potete fare ad esempio:
export TOKEN=\$(jq -r .access_token \${CONDOR_CREDS}/cygno.use) per usare il vostro token

IAM-Instance	name of the oauth service in PSGE	Description
iam-t1-computing.cloud.cnaf.infn.it	t1	multi-VO
iam-cygno.cloud.cnaf.infn.it	cygno	CYGNNO-VO
iam.cloud.infn.it	infncloud	multi-VO



```
executable = script.sh
output = stdout-$(ClusterId).$(ProcID).txt
error = stderr-$(ClusterId).$(ProcID).txt
log = output-$(ClusterId).$(ProcID).log
transfer_executable = Yes
should_transfer_files = Yes
when_to_transfer_output = ON_EXIT
want_io_proxy = true
use_oauth_services = cygno
cygno_oauth_permissions = profile,email,openid,offline_access,wlcg.groups
queue 3
```



```
...zzitelli/rucio-shell ...node-0: ~ -- zsh ...sge.cloud.infn.it ...192.168.99.133 ...yterhub -- zsh
analisi_regata.py cygno_gsetup.sh riepilogo_bolina_segmento_vento120.csv riepilogo_bolina_segmento
_vento160.csv test-budda-htc
[mazzitel@ui ~]$ cd test-budda-htc/
[mazzitel@ui test-budda-htc]$ condor_submit --spool sub
Submitting job(s)
Hellow, mazzitel.
Please visit: https://condor.psge.cloud.infn.it/key/e5fb2a1039dc96b03fdact6530e36b1247adc330fd486795b861fd9f698e83e5

[mazzitel@ui test-budda-htc]$ condor_submit --spool sub
Submitting job(s)...
3 job(s) submitted to cluster 44.
[mazzitel@ui test-budda-htc]$ condor_q

-- Schedd: condor.psge.cloud.infn.it : <131.154.99.149:9618?... @ 11/10/26 15:06:48
OWNER BATCH_NAME SUBMITTED DONE RUN IDLE TOTAL JOB_IDS
mazzitel ID: 44 11/10 15:06 - - 3 3 44.0-2

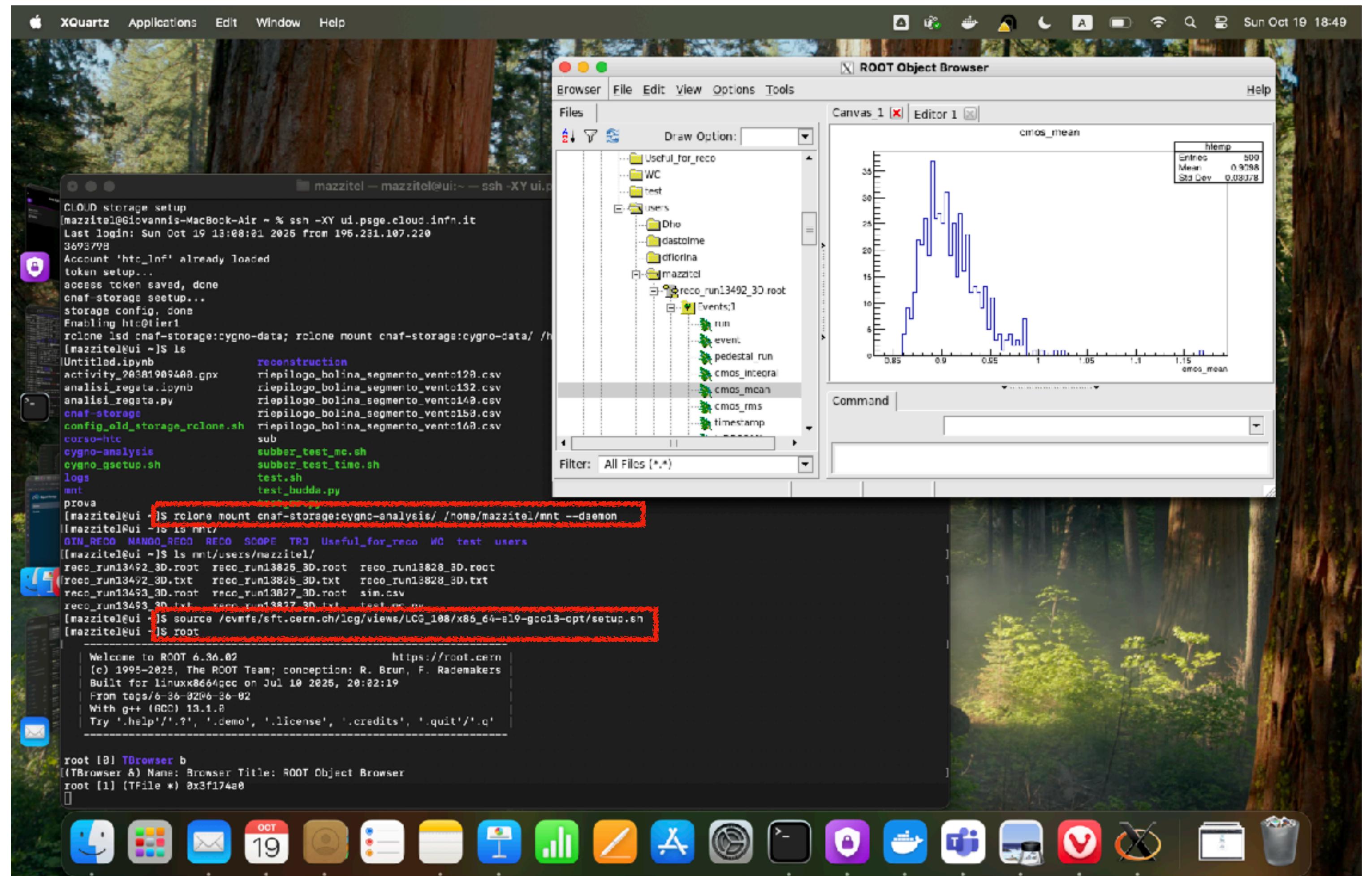
Total for query: 3 jobs; 0 completed, 0 removed, 3 idle, 0 running, 0 held, 0 suspended
Total for mazzitel: 3 jobs; 0 completed, 0 removed, 3 idle, 0 running, 0 held, 0 suspended
Total for all users: 6 jobs; 3 completed, 0 removed, 3 idle, 0 running, 0 held, 0 suspended

[mazzitel@ui test-budda-htc]$
```

HOWTO accesso storage remoto

rclone, etc per l'accesso allo storage d'esperimento o personale

- gfal, curl, boto(*), etc
- con **rclone** potete montare localmente qualunque tipo di storage (es gDrive)
- se **configurato** basta una comando del tipo: rclone mount drive:folder localdrive –daemon
- info: <https://rclone.org/docs/>
- (*) py lib per l'accesso a s3:
<https://boto3.amazonaws.com/v1/documentation/api/latest/index.html>



CSN2 analysis facility

conclusioni

- grazie alle risorse del PNRR, gli sviluppi di DATACLOUD e l'implementazione di C. Pellegrino, e il supporto del GdL calcolo nasce la **analysis facility di CSN2**
- questa facility permette di dare **risorse immediate** a tutti gli esperimenti di commissione che lo richiedano ed eventualmente **assorbire momentanei picchi estemporanei**
- **L'accesso via ssh e web UI**, e l'integrazione con i più comuni editor e software, permettono anche sviluppo del software e/o la visualizzazione di dati i modo accessibile
- l'integrazione di un batch system basato su htcondor, con una notevole quantità di risorse, permettere a tutti di fare calcolo intensivo, se necessario.
- sono in **fase di sviluppo** i tool di monitoraggio e accounting, la spool accessibile per il retrieve dei file, e alcune piccole implementazioni per allungare la lifetime dei token
- abbiamo in programma di fare un **corso di formazione INFN** aperto a tutti gli interessati.