Supporting medium/small-sized experiments in the transition from X.509 to JWTs

Conference on Computing in High Energy and Nuclear Physics 2024, Krakow

C. Pellegrino, A. Shtimmerman, A. Pascolini, M. Barbetti, C. Giugliano, D. Lattanzio, L. Morganti, A. Rendina

INFN CNAF

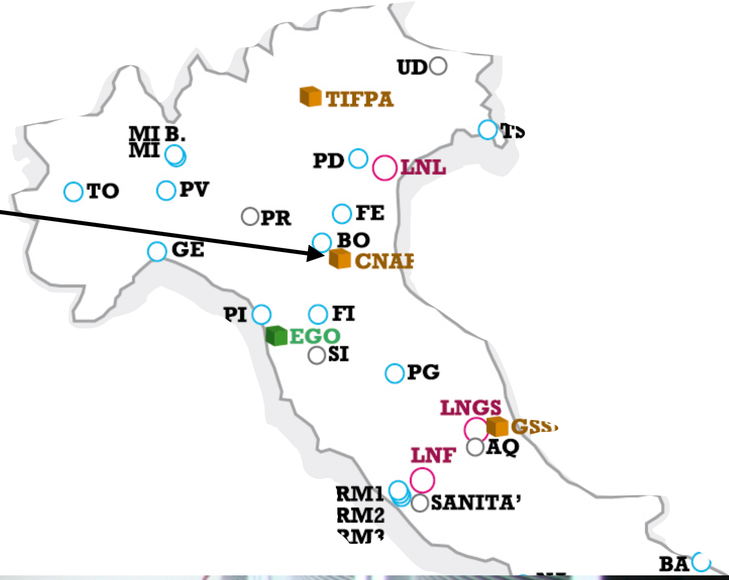
Outline

- The environment
  - INFN-CNAF
  - Our users
- What is a JWT
  - the problems it solves
  - the problems it introduces
- supporting in the transition
  - data transfer scripts
  - HTCondor mapping plugin
  - IAM management
  - mytoken

The environment

The Italian WLCG T1

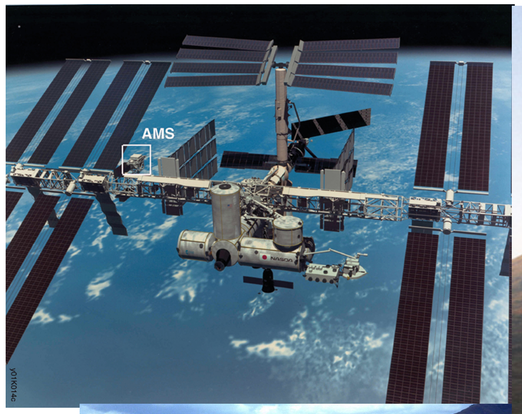
- The Italian WLCG Tier-1 is located in **Bologna** (**Emilia Romagna**)
  - managed by **INFN-CNAF** (https://www.cnaf.infn.it/)
- **~2.500 computing nodes** (physical and virtual machines)
  - **~60.000 core** managed by a batch system
- **~230 PB of disk**
- **~230 PB of tape** for long-term storage
- supports 60+ scientific communities
  - not only LHC and not only from the Physics field

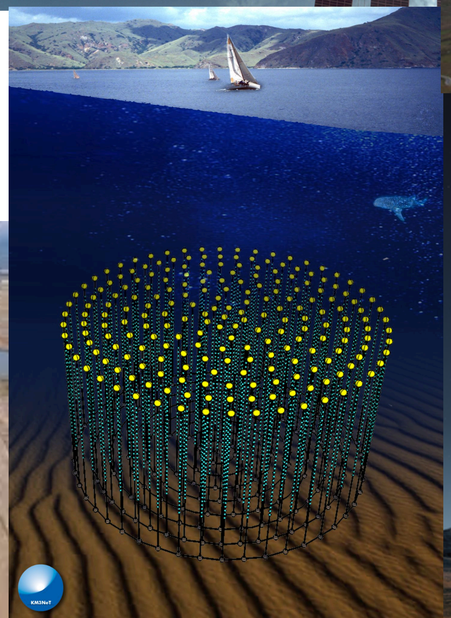


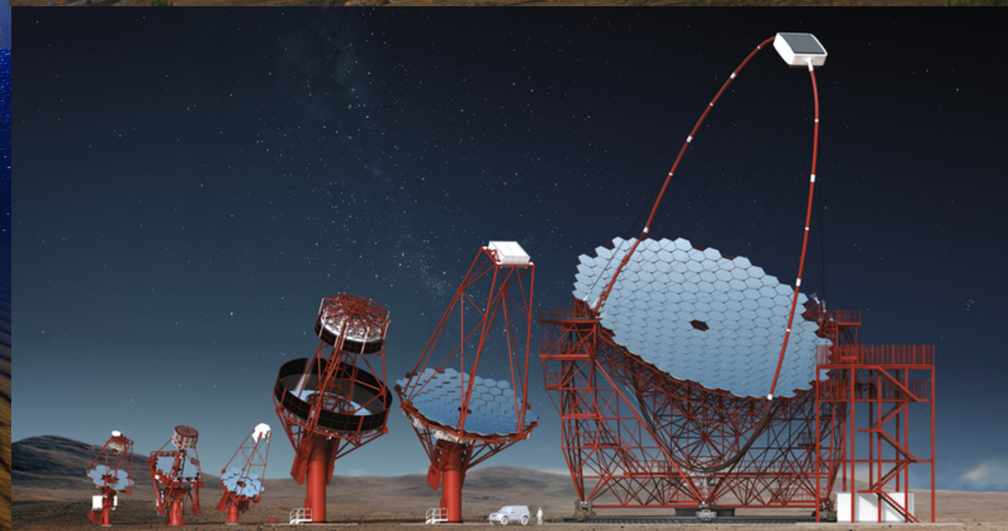
Not only WLCG

Other supported scientific communities

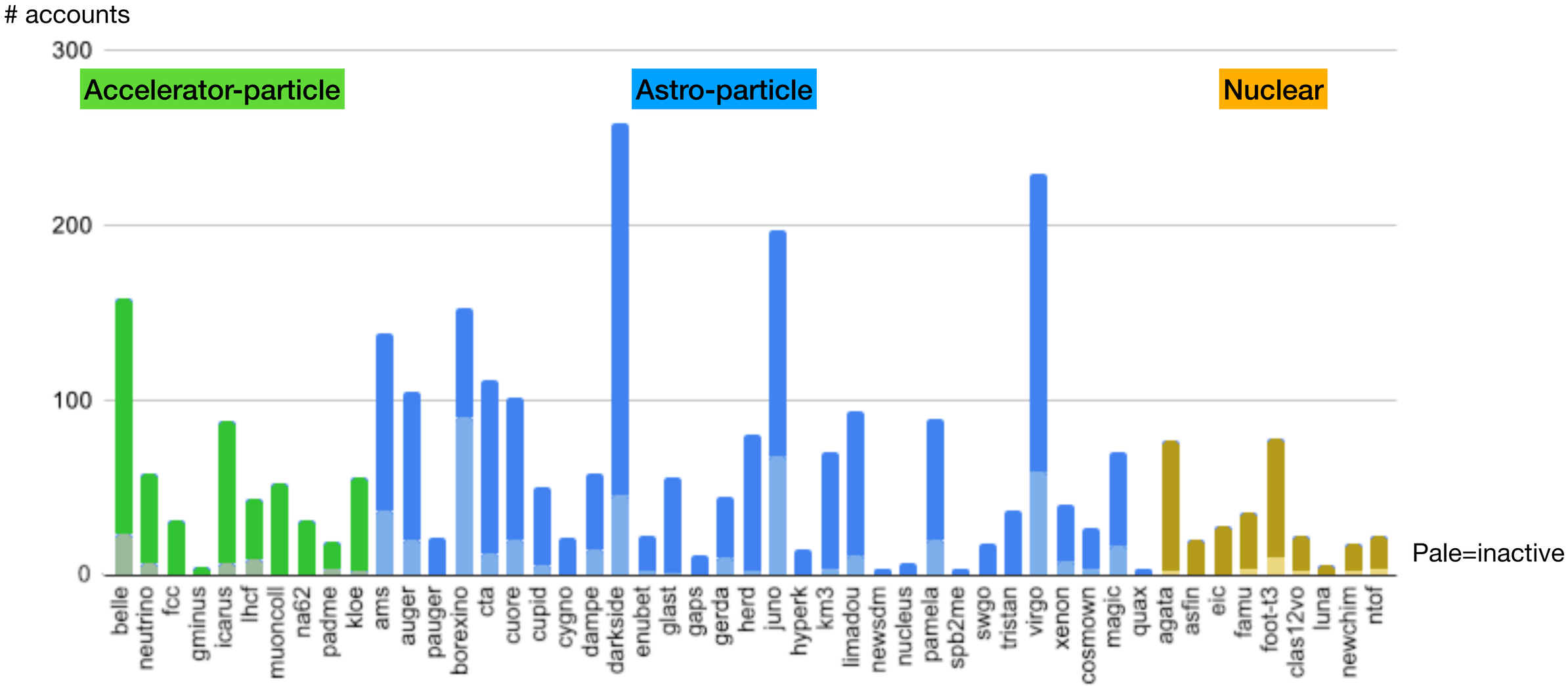
- High-Energy Physics: 8
- Astro-particle Physics: 18
- Gravitational Waves: 2
- Nuclear Physics: 16
- Dark Matter: 6
- others: 10







Medium/small-sized users per community



What is a JWT

JSON Web Token

TLDR: a (long) string

- eyJraW[...].**eyJzdWIiOiI4Yz[...]**.Hdpr__IbGkO[...]
- The **payload** is a JSON object containing a lot of information
  - User ID, groups, scopes, audence, expiry, etc...
- great for fine-grained authorisation!
- industrial standard
  - pillar for **OAuth2** and **OpenID Connect**
  - great interoperability with widely adopted tools


```
{
  "sub": "8c08e83b-5ebb-4f92-a362-c3a2a5d3ed2f",
  "iss": "https://iam-t1-computing.cloud.cnaf.infn.it/",
  "preferred_username": "budda",
  "client_id": "bcbfc16a-6e2a-442b-b8ce-a083270ce3d7",
  "wlcg.ver": "1.0",
  "aud": "https://wlcg.cern.ch/jwt/v1/any",
  "nbf": 1729512838,
  "scope": "compute.read compute.cancel compute.modify wlcg wlcg.
  "name": "Carmelo Pellegrino",
  "exp": 1729516438,
  "iat": 1729512838,
  "jti": "dea4753f-8555-49e1-badf-46dac3255c6d",
  "wlcg.groups": [
    "/darkside",
    "/dune",
    "/foot",
    "/jlab12",
    "/km3net",
    "/muone",
    "/ntof",
    "/padme",
    "/pauger",
    "/quax",
    "/swgo",
    "/user-support"
  ]
}
```

JSON Web Token

the problems it solves

- Enables interoperability with tools developed outside the HEP community
  - JupyterHub, Grafana, Gitlab and GitHub, S3, OpenStack, etc...
  - HTCondor-CE, ARC-CE, StoRM, dCache, XRootD...
  - ... avoiding things like the "http**g** nightmare"
- Improves IT security wrt x.509
  - allows implementing fine-grained authorisation via **scopes**, **audience**, and **groups**
  - **short-lived** => makes harder to tamper

JSON Web Token

the problems it introduces

- rather different approach wrt X.509
  - users need to understand clients, tokens, refresh tokens, scopes, etc...
  - users need to adapt their workflows
- **short-lived**:
  - typical lifetime is 1h
  - deferred and long-lasting operation can be problematic
  - e.g.: recursive data copy of large number of files and writing job output to remote storage

Supporting in the transition

xfer-oidc

https://baltig.infn.it/cnaf-user-support/xfer-oidc

- A shell scripts to be used in extreme environments with minimal dependencies to register a client and get JWTs that can be used also to perform some basic data management operations via WebDAV
- Dependencies: curl, tar, jq, sed, sh (the POSIX shell)
- Use case:
  - a user is holding in a NAS powered by an immutable distribution of FreeBSD 15TB of experimental data to be transferred to CNAF

copy.sh (suggestions for a better name?)

https://baltig.infn.it/exp-supp/copy.sh

- A shell scripts to transfer large quantity of data (N >10^5 files, V > 1 TB), with parallel transfers
- Dependencies: gfal/curl, oidc-agent, GNU parallel[1], bash
- Use cases:
  - decommissioning of MUON-E storage on EOS at CERN
  - decommissioning of Auger storage at CC-Lyon

[1]: O. Tange (2018): GNU Parallel 2018, March 2018, https://doi.org/10.5281/zenodo.1146014

A catch-all INDIGO-IAM instance

https://iam-t1-computing.cloud.cnaf.infn.it

- Multi-VO IAM instance
  - 43 groups, O(100) users
- Gives access to storage and computing resources, also located outside CNAF
- Currently, CNAF accounts and IAM accounts are disjoint
  - we plan to integrate it with the new CNAF AAI system when the latter will be mature enough

HTCondor automatic mapping

https://baltig.infn.it/exp-supp/scitokens-mapping

- INFN-T1 is essentially an HTCondor-powered site
  - JWTs are called "*SCITOKEN*" in HTCondor jargon
- Pre-GSI phaseout there was support for GSI callouts
  - something similar introduced in version 10.5 in the form of *SCITOKEN plugins*
- A script to perform pool-account allocation and mapping to user identity
  - sqlite3 DB backend

HTCondor automatic mapping

https://baltig.infn.it/exp-supp/scitokens-mapping

- INFN-T1 is essentially an HTCondor-powered site

```
SEC_SCITOKENS_PLUGIN_A_COMMAND = $(LIBEXEC)/scitokens-mapping.sh
SEC_SCITOKENS_PLUGIN_B_COMMAND = $(LIBEXEC)/scitokens-mapping.sh SpecificVOName
```

  - JWTs are called "*SCITOKEN*" in HTCondor jargon

- Pre-GSI phaseout there was support for GSI callouts

  - somethi

```
# Multi-VO IAM instance
SCITOKENS /^https:\/\/iam-t1-computing\.cloud\.cnaf\.infn\.it\/,/ PLUGIN:A

# Plugin for an IAM with no groups defined
SCITOKENS /^https:\/\/iam-dedicated-instance\.example\.com\/,/ PLUGIN:B
```

    TOKEN plugins

- A script to perform pool-account allocation and mapping to user identity

  - sqlite3 DB backend

HTCondor automatic mapping

https://baltig.infn.it/exp-supp/scitokens-mapping

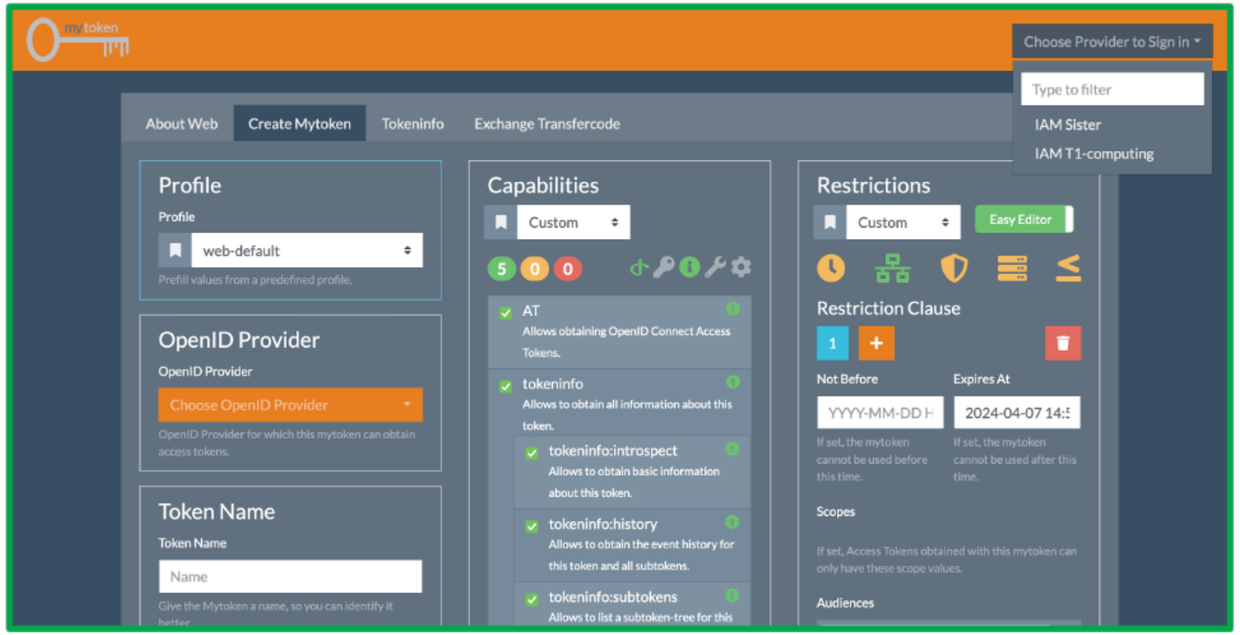
- On the user-side:
  - if the token issuer is dedicated, just create a token and do your condor_submit
  - if the token issuer is multi-VO (like iam-t1-computing), needs to bring your chosen VO at the top of the group list:
    - oidc-token wlcgt1comp -s 'wlcg.groups:/user-support'

Tokens in deferred operations

- JWTs are usually short-lived (default: 1h)
- Typical deferred operation is to copy grid job output to grid/cloud storage
- you don't really want to attach your credentials or OIDC client to a remotely executed job

Working with mytoken

what is a mytoken?



- Extends on the concept of Refresh-Token
- JWT based
- Implements additional features:
  - Rotation
  - Restrictions (lifetime, connected IP, usage limit)
- It can be used in **oidc-agent** pretty like a refresh-token for a client registered on a **mytoken-server**
- https://mytoken-docs.data.kit.edu/
- Token-Transition-update-240327

Current activity

- Deployment of a self hosted mytoken-server
  - server configuration
  - setup CNAF profile (rotations, restrictions, capabilities, templates)
- connected to Tier-1 IAM instance and "iam-herd"
- Test phase with real use-case
  - see next slide

Current activity

- Collaboration with the HERD community
  - quite interesting computing model, given the size of the collaboration
  - several computing resources and technologies (HTCondor, MinIO, StoRM-WebDAV, K8s), distributed in various sites (INFN-T1, INFN-Cloud, INAF, ASI)
- Very promising test results during last summer:
  - interoperability among the various technologies has been reached:
    - execute anywhere, transfer to/from anywhere

Thank you for your attention

Backup

Working with mytoken

•Deployment of a self hosted **mytoken-server**

→ server configuration

→ setup CNAF profile

(rotations, restrictions, capabilities, templates)

→ connect to Tier-1 **IAM instance**

•Test phase on how to get and use **mytokens**

→ client choice (**mytoken-client**, **oidc-agent**)

→ submission tests to manage files with AT requested via **mytoken** flow

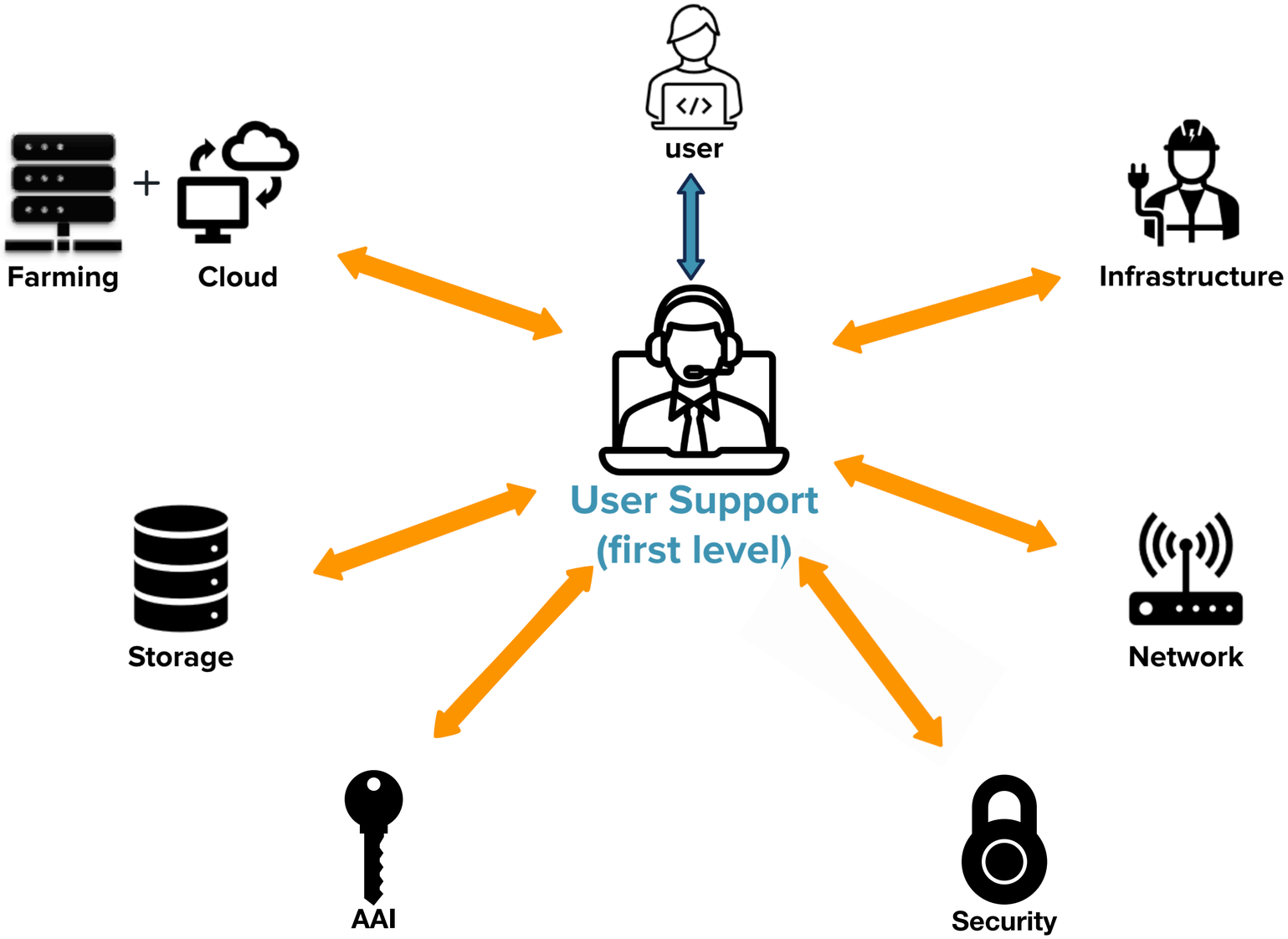
- INFN-T1 User Guide
- The group continuously maintains detailed knowledge base in the form of an online user guide
- The guide is public and organized in 14 chapters
- It contains suggestions with simplified and practical examples on how to use tools such as conda, singularity/apptainer, HTCondor, SLURM, oidc-agent, gfal2-util, and many others
- It explains also all the procedures and best practices needed to access and efficiently use the Tier-1 resources:
- How to request a new account, how to access the user interfaces, how to requests x509 certificates, how to obtain JWT tokens, etc…

The User Support unit

- Mission: solve most of the basic problems, and to write **documentation** to improve the usage of **solutions** and **standard tools** the Centre provides. Among them:
  - **HTCondor**, is the batch system for HTC, and **SLURM** for HPC
  - **gfal2-util**, is the tool for data transfer/management via Grid
  - **oidc-agent**, is the CLI tool to manage JWT tokens
  - **singularity/apptainer**, is the container solution
- Supporting the use of specific software:
  - personalised support on certain, specific, use cases. E.g.: user scripts, environment, etc...
  - different scientific communities need different software
- Composition: 5 people **coming from different scientific fields**, plus some effort from Storage and Farming

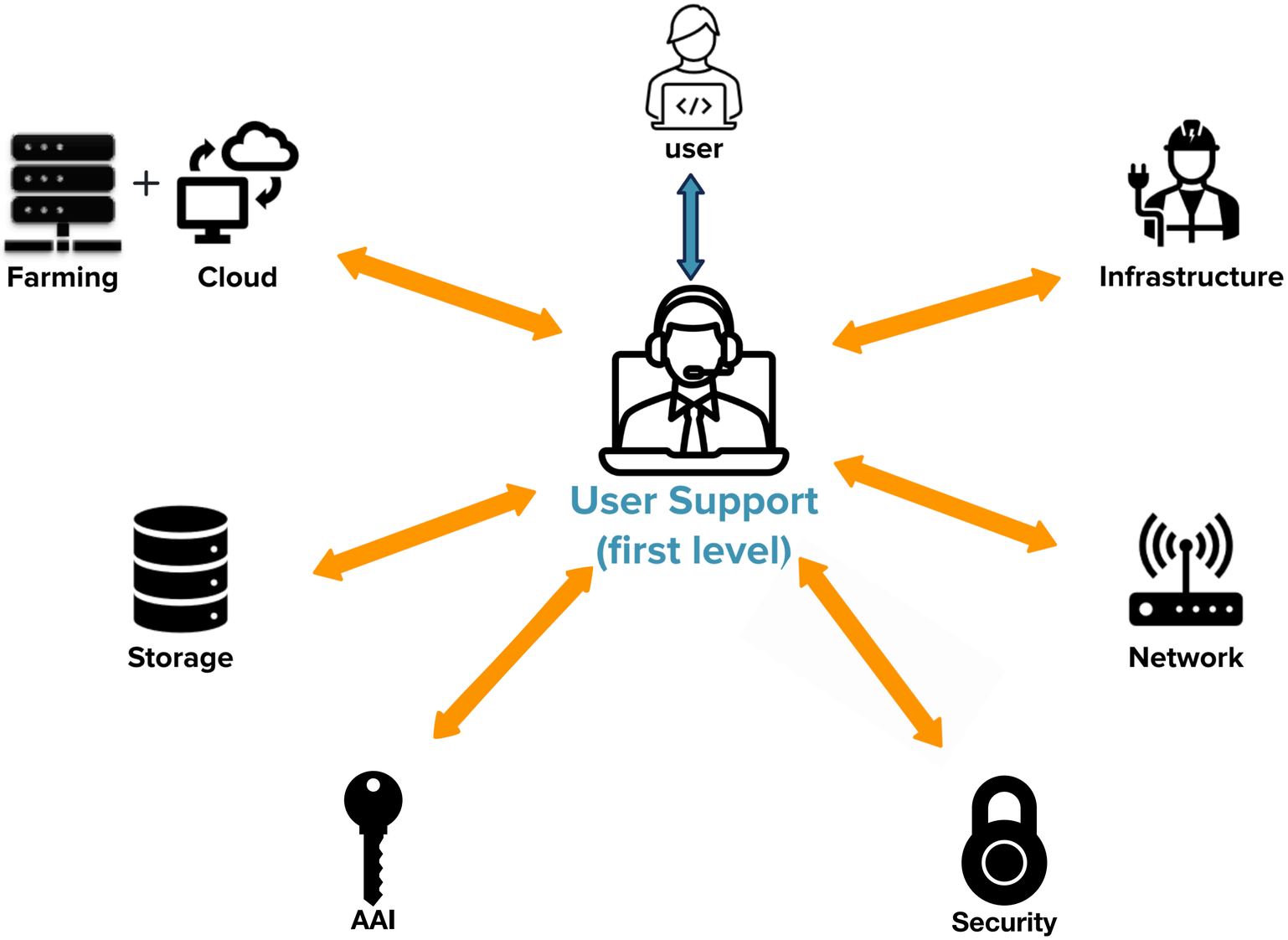
Support activities

- **On-boarding** of new **scientific communities** (projects, experiments, others)
- User **registration** procedure (recognition, authorisation, account creation)
- **Documentation** for users:
  - INFN-T1 **user guide** https://l.infn.it/t1guide
  - Automatically updated **useful pages** https://www.cnaf.infn.it/~usersupport/
- **Communication**:
  - Direct user communication (personal emails, chat)
  - Announces (mailing list, gocdb)
  - Periodic presentations (comitato di gestione (*CdG*), special events)
  - Dedicated meetings with experiments' people (on-boarding, special requests)


user
Farming
Cloud
Infrastructure
User Support
(first level)
Storage
Network
AAI
Security

Conclusions and perspectives

- Challenges for the User Support:
  - **keep its central role** between scientific communities and the INFN computing ones
  - support over multiple infrastructures => increase in workload driven by the DataCloud project (see poster 27 on Thursday)
    - an increasing adoption of automation techniques
    - getting more people involved to keep a sustainable personal effort
- Future plans:
  - **Harmonisation** of the INFN-Cloud and T1 **documentations**
  - Gain good visibility of on both **cloud** and **T1** usage.
- Fostering the creation of **a community of users** who provide **mutual support** on common computing topics


user
Farming
Cloud
Infrastructure
User Support
(first level)
Storage
Network
AAI
Security

The INFN Tier-1 User Guide

https://l.infn.it/t1guide

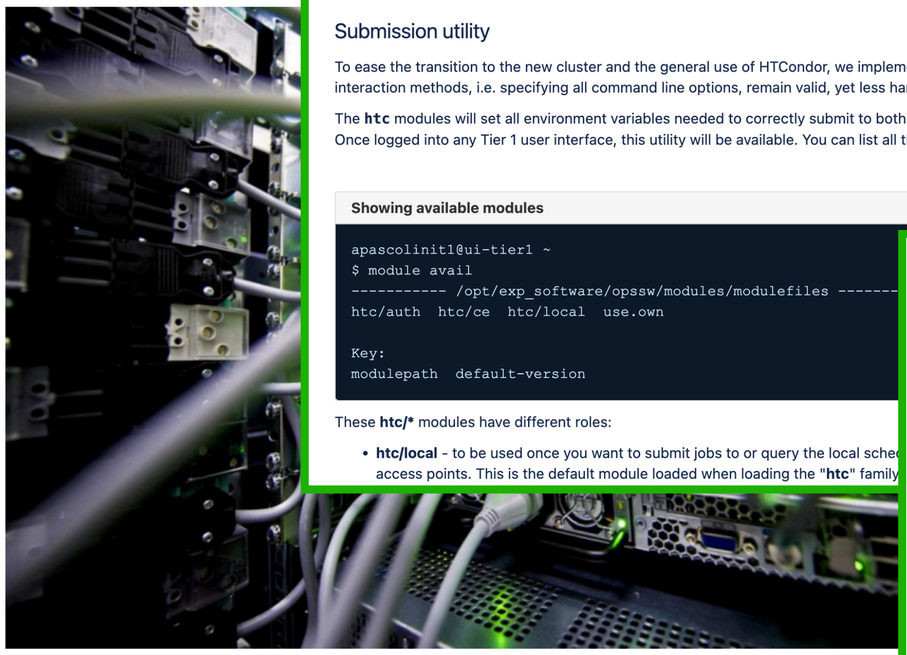
Tier1 - Documentation

PAGE TREE

- INFN-CNAF Tier-1 User Guide
  - 1 - CNAF
  - 2 - Tier-1
  - 3 - Bastion & user interfaces
  - 4 - Farming
  - 5 - Storage
  - 6 - The HPC cluster
  - 7 - Cloud @ CNAF
  - 8 - Digital Personal Certificate
  - 9 - Job submission
  - 10 - Data Transfers
  - 11 - Monitoring
  - 12 - Helpful information and ti
  - 13 - Support
  - 14 - Problem report
  - Appendix A - Submit Descripti
  - Appendix B - Helpful links
  - Bibliography
- Monitoring

Pages / Tier1 - Documentation

INFN-CNAF Tier



INFN-CNAF Tier-1 user guide
Summary

1. CNAF
2. Tier-1
3. Bastion & user interfaces

Submission to the new cluster HTC23

- Submission utility
- Local Submission
- Grid Submission
  - Token submission
  - SSL submission

Submission utility

To ease the transition to the new cluster and the general use of HTCondor, we implemented a solution based on
interaction methods, i.e. specifying all command line options, remain valid, yet less handy and more verbose.

The **htc** modules will set all environment variables needed to correctly submit to both the old and the new HTC
Once logged into any Tier 1 user interface, this utility will be available. You can list all the available modules using

Showing available modules

```
apascolinit1@ui-tier1 ~
$ module avail
---------- /opt/exp_software/opssw/modules/modulefiles -------
htc/auth  htc/ce  htc/local  use.own

Key:
modulepath  default-version
```

These **htc/*** modules have different roles:

- **htc/local** - to be used once you want to submit jobs to or query the local sche
  access points. This is the default module loaded when loading the "**htc**" family

Pages / ... / 10 - Data Transfers

- **Removing a file**

```
[arendina@ui-tier1 ~]$ gfal-rm davs://xfer-archive.cr.cnaf.infn.it:8443/j
davs://xfer-archive.cr.cnaf.infn.it:8443/juno/test0107  DELETED
```

Third-party-copies

In order to properly perform a third-party-copy between two endpoints which support the http protocol
macaroon.
Indeed, this token is used to authenticate the user always to the second endpoint. For this reason, the se
copy is in pull or push mode.

Actually, if both the endpoints are able to release a macaroon and the used gfal version is greater or equ
Otherwise, if only one of the two endpoints can release a BEARER_TOKEN, or equivantly just one endpoi
macaroon to that endpoint.

Two easy examples follow below.

Pull-copy

ProxyJump is a feature of SSH clients used to facilitate access to a remote server through one or more intermediary server ase IHEP, and th
bastion.cnaf.infn.it is a jump host. s happens becau

How ProxyJump works

When using `ProxyJump`, the client establishes an SSH connection to the first server (the jump host) and then, through this usercert.pem
connection to the target server. This process can be extended to multiple intermediary servers if needed.

Configuring ProxyJump for SSH into CNAF User Interfaces

It is possible to configure the `ProxyJump` by configuring the SSH client of **your PC**. The `~/.ssh/config` file can be used

Example Configuration in the '~/.ssh/config' File:

```
Host bastion
    hostname bastion.cnaf.infn.it
    User <username>
Host t1
    hostname ui-tier1.cr.cnaf.infn.it
    User <username>
    ProxyJump bastion
```

In the `Host` field, you can specify the name that you want to use to identify the target-server that you want to connect to.
Once this example file is written, it will be possible to SSH into ui-tier1 by just typing the following command:

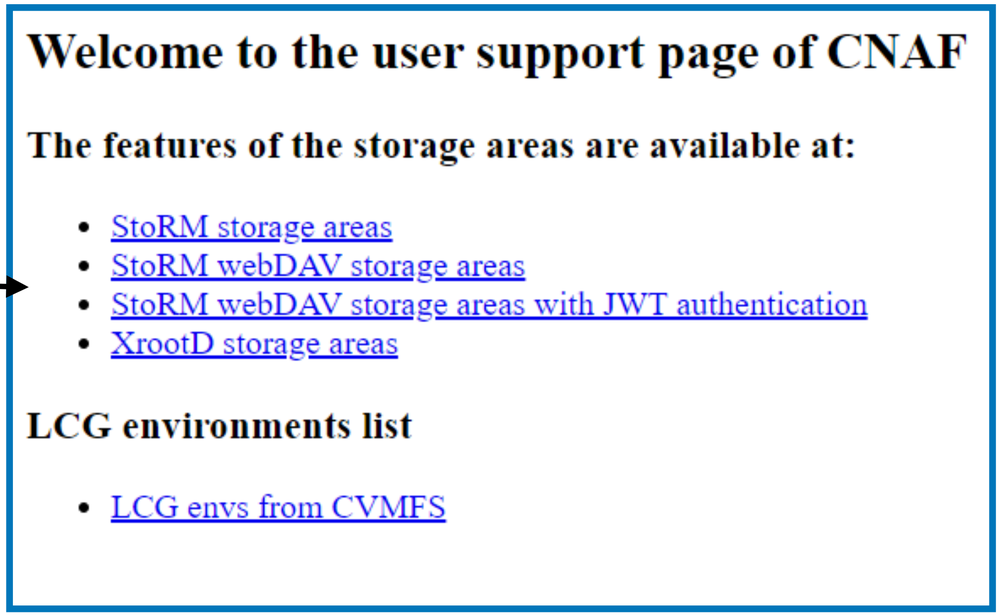
```
ssh t1
```

Handy links to useful pages 1/2

- Automatically updated useful pages every night
- To advertise specific information about the services available to the communities in a form that is easy to access and use:
  - https://www.cnaf.infn.it/~usersupport/

Storage Areas per service and experiment

LCG envs via CVMFS



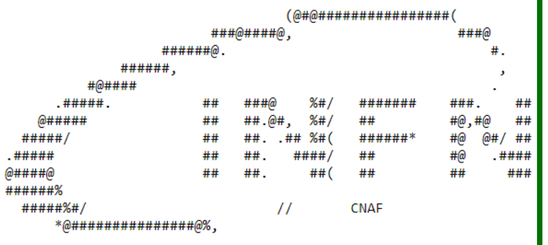
Handy links to useful pages 2/2

LCG Environments navigator

In the table below you find the updated list of LCG environments available through CVMFS.
Pick one of your choice from the list below, depending on the compiler version, root version etc..
and then run the following command on a user interface:

```
source /cvmfs/sft.cern.ch/lcg/views/<env>/<env_version>/setup.sh
```

env	env_version	compiler	root_version	python_version	python2_version	python3_version	cpp_version
LCG_97apython3_LHCB_4	x86_64-centos7-gcc9-opt	g++	6.22/04	2.7.16	2.7.16		cxx17
LCG_99	x86_64-ubuntu2004-gcc9-opt	c++	6.22/06	3.8.6		3.8.6	cxx17
LCG_99	x86_64-centos7-gcc8-opt	g++	6.22/06	3.8.6		3.8.6	cxx17
LCG_99	x86_64-centos7-gcc10-opt	g++	6.22/06	3.8.6		3.8.6	cxx17
LCG_99	x86_64-centos7-clang10-opt	clang++	6.22/06	3.8.6		3.8.6	cxx17
LCG_99	x86_64-centos8-gcc10-opt	g++	6.22/06	3.8.6		3.8.6	cxx17
LCG_99cuda	x86_64-centos7-gcc8-opt	g++	6.22/06	3.7.6		3.7.6	cxx17
LCG_geant4ext20210118	x86_64-centos8-gcc10-opt	g++	6.22/06	3.8.6		3.8.6	cxx17
LCG_geant4ext20210118	x86_64-centos7-gcc10-opt	g++	6.22/06	3.8.6		3.8.6	cxx17



StoRM WebDAV storage areas with JWT authentication

aa.wp6

StoRM WebDAV endpoint	Access point	Root path
xfer.cr.cnaf.infn.it	/DataCloud-TB	/storage/gpfs_escape/datacloud-tb

belle

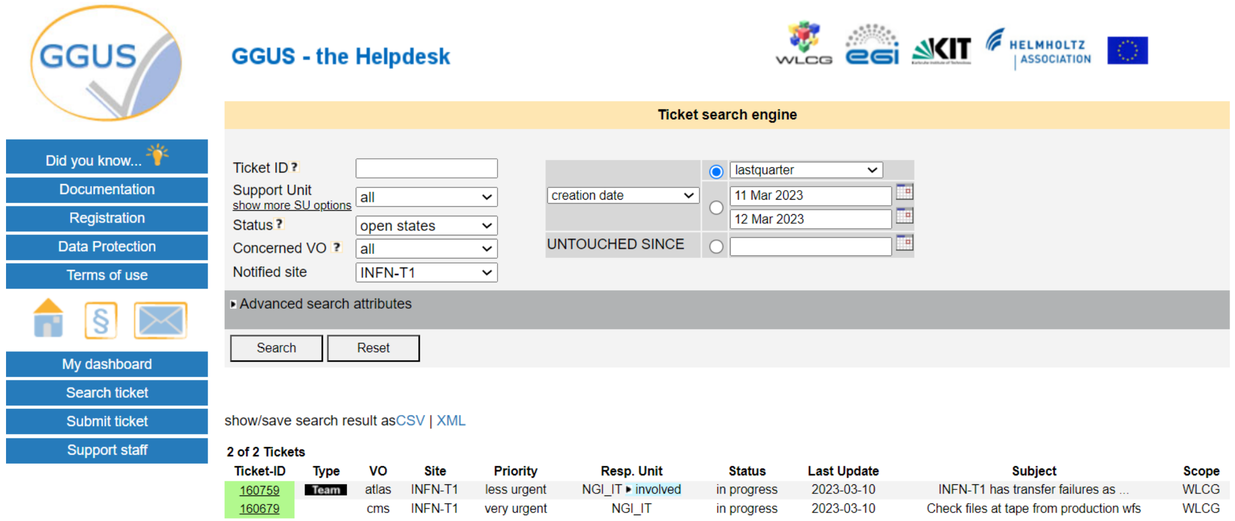
StoRM WebDAV endpoint	Access point	Root path
xfer-archive.cr.cnaf.infn.it	/belle	/storage/gpfs_data/belle

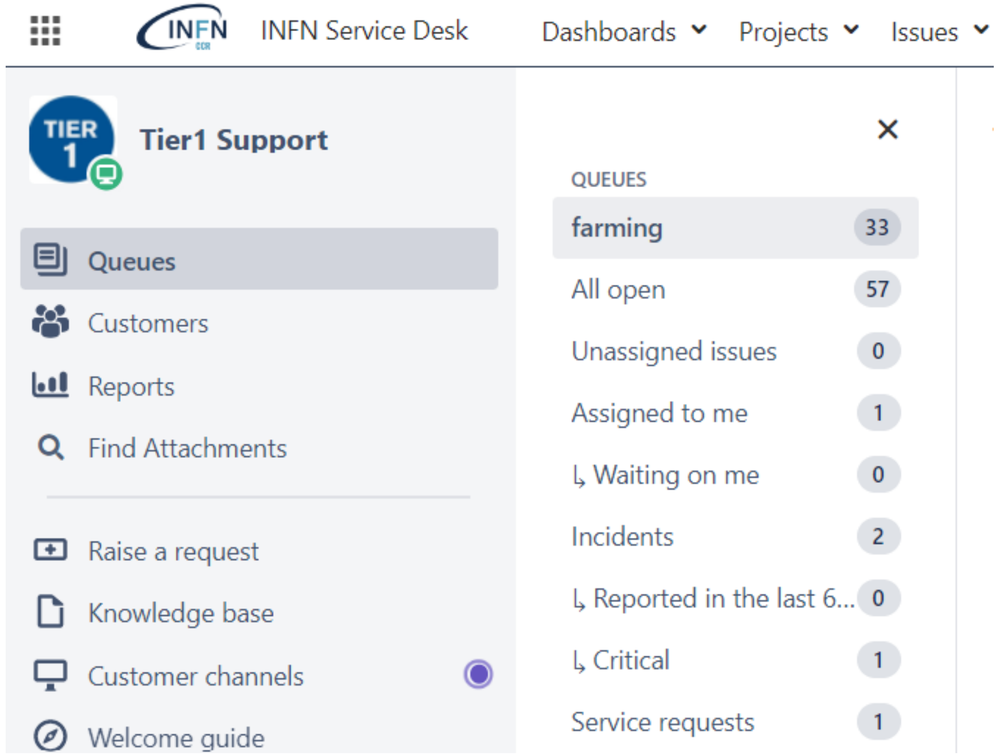
cta-lst

StoRM WebDAV endpoint	Access point	Root path
xfer-archive.cr.cnaf.infn.it	/cta-lst	/storage/gpfs_data/ctadisk/cta-lst

Communication channels

- Mailing lists to reach the users regarding the datacentre status
- Ticketing systems:
  - GGUS, mainly for WLCG VOs
  - Ticketing system for internals
  - Ticketing system for users (in development)





Typical issues

- **First level** support
  - disk quota exceeded
  - issues with batch jobs (not running, getting killed, etc…)
  - explanations/documentation requests
- **Second level** support (usually escalated to other CNAF teams)
  - installation of software
  - filesystem access management (SA configuration, POSIX permissions)
  - network problems
- Due to the overlap with other units, part of the second level support is also carried out in cooperation with the User Support team