PROCEEDINGS OF SCIENCE

Experiments support at INFN-T1

Daniele Lattanzio,[a,*] **Andrea Rendina,**[a] **Alessandro Pascolini,**[a] **Lucia Morganti,**[a] **Federico Fornari,**[a] **Carmelo Pellegrino**[a] **and Daniele Cesini**[a]

[a] *INFN-CNAF,*
Viale Berti Pichat 6/2, 40127 Bologna, Italy
E-mail: daniele.lattanzio@cnaf.infn.it, andrea.rendina@cnaf.infn.it, alessandro.pascolini@cnaf.infn.it, lucia.morganti@cnaf.infn.it, federico.fornari@cnaf.infn.it, carmelo.pellegrino@cnaf.infn.it, daniele.cesini@cnaf.infn.it

The Italian WLCG Tier-1 located in Bologna and managed by the Istituto Nazionale di Fisica Nucleare (INFN), provides computing resources to several research communities in the fields of High-Energy Physics, Astroparticle Physics, Gravitational Waves, Nuclear Physics and others. The facility is hosted at Centro Nazionale Analisi Fotogrammi (CNAF). Although the LHC experiments at CERN represent the main users of the Tier-1 resources, an increasing number of communities and experiments are also being supported in all of their computing activities. Due to this very heterogeneous user base, an efficient support system is needed in order to assure a smooth and appropriate exploitation of the computing infrastructure. In this framework, this role is played by the Tier-1 User Support group, which acts as a first level support and represents the entry point for services and support requests, and for problem reports. The group makes use of multiple systems to meet the different needs and specificities of the supported experiments. Moreover, the group continuously maintains a detailed knowledge based in the form of an online users guide. The communication channels are represented by ticketing systems and also by mailing lists used for a more direct communication with users, allowing to notify maintenance interventions, downtimes and more in general all the new features and services provided by the Datacentre. In this paper, the ticketing systems, tools, platforms and services that User Support offers, and the internal organisation of the department, will be described. Future workflow plans in view of the DATACLOUD project, which will require an increasing effort, will also be presented.

International Symposium on Grids and Clouds (ISGC) 2023,
19 - 31 March 2023
Academia Sinica Taipei, Taiwan

*Speaker

PoS(ISGC&HEPiX2023)013

1. Introduction: The INFN-CNAF Centre

The Worldwide LHC Computing Grid (WLCG) [3] is a collaboration involving around 170 computing centres in more than 40 countries. It provides computational resources to store, distribute and analyse the ~200 PB of data expected every year from the Large Hadron Collider (LHC) experiments at CERN. WLCG employes a *Tier-based* computing architecture[1], in which computing centres are organised in a hierarchical structure. The main Tier-0 Datacentre is hosted at CERN, physically near the detectors, and is directly connected via a high-performance network [2] to the 11 Tier-1s. A primary copy of the raw data is sent from the Tier-0 to the Tier-1s for storage preservation and analysis purposes. Thereafter, data are distributed to the over 160 Tier-2s of the GRID, for further analysis.
The Italian WLCG Tier-1 site is located in Bologna and it is hosted since 2003 at CNAF [4], the National Centre of INFN, the Italian National Institute for Nuclear Physics, dedicated to research and development on information and communication technologies. The Centre provides the resources, support and services needed for data storage and distribution, processing and analysis and Monte Carlo production to over 60 scientific communities.
CNAF supports either *local*[1], Grid and Cloud computing. The computing farm is composed of ~2000 computing nodes providing ~60.000 cores accessed via a batch system, and another partition running Cloud services that are accessible via either IaaS or PaaS solutions. Storage resources consist of ~70 PB of disk space served via a distributed filesystem (GPFS), and ~130 PB of tape space. The latter is mainly used as a long-term storage solution for archival data.
Although the LHC experiments are the main users of the Datacentre, CNAF represents the principal computing facility for many communities covering a broad scientific landscape, not only from the Physics field but also for Biomedical Health and Life science, requiring ISO-certified isolated computing zones.

In the next sections, the internal organisation and the main activities involving the INFN-T1 User Support unit, with a particular focus on the tools, platforms and services that the Group provides, will be showed. Future workflow plans in view of the DATACLOUD project and also of the imminent move of the Datacentre to a new location and its infrastructural expansion, will also be presented.

1.1 INFN-T1 internal organisation

The Tier-1 operations are structured in 5 internal units (Figure 1): the batch and cloud resources are managed by the *Farming* group, the storage systems and the data transfer services are managed by the *Storage* group, the LAN and WAN connections of CNAF are managed by the *Network* group and the facilities of the Centre are managed by the *Infrastructure* group.
Working in close synergy with those, there is the *User Support* group, which represents the first level of support and acts as the entry point for users requests, problems, technical issues and more in general for all kind of communications between the experiments and the centre, thus maintaining a global visibility on its status.

[1] *local users* do connect to CNAF servers via SSH to manage their work either accessing the computing farm via a batch system and the storage via POSIX access to data.

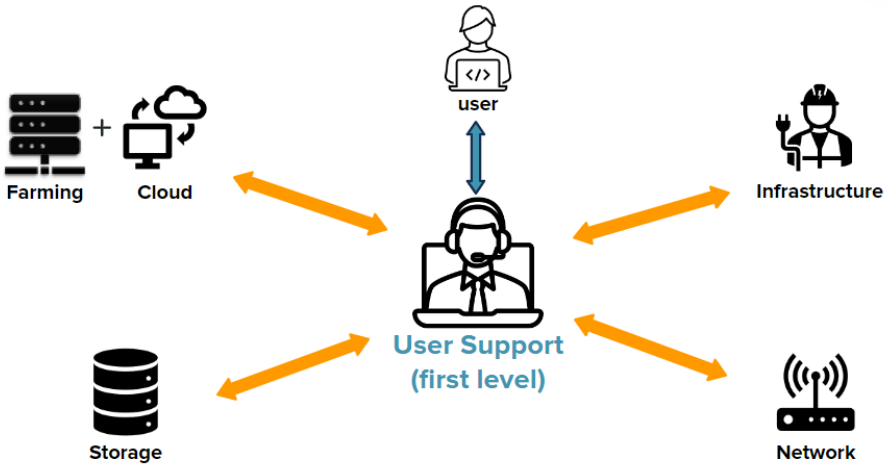

Figure 1: Sketch of the Tier-1 internal organisation and the relationship between the User Support unit and the others.

2. The User Support unit

By its definition, the duties of the INFN-T1 user support unit with respect to its users are:

- write and keep updated the official user documentation;
- act as their main contact point;
- assist them to ensure an efficient use of the computing resources;
- develop tools and procedures for the supported scientific collaborations;
- collaborate with the scientific communities in the definition of their computing models.

The User Support unit is currently composed by six people, all coming from a broad scientific experience, covering Astronomy, Astrophysics, Mathematics and Physics. Thanks to this fact, the group has the particular ability to understand the use-cases, problems and questions of its users by *speaking their same language*.

Furthermore, the User Support members' daily activities are often in overlap with that of other CNAF teams. Part of the second level support is also carried out in cooperation with them, which helps sharing information and technical competences.

2.1 Documentation

The group continuously maintains a detailed knowledge base available online [6] and organized in 14 chapters. It contains indications with simplified and practical examples on how to use the set of tools installed in the centre. The installation of this software is usually maintained by the sysadmins, although the installation of custom software by the users is allowed, to some extent, on dedicated shared filesystems.

Examples of centrally managed software deployments are *conda*[7], *apptainer*[8], *HTCondor*[9], *SLURM*[12], *oidc-agent*[10], *gfal2-util*[11] and many others.

The guide explains also all the procedures and best practices needed to access and efficiently use the Tier-1 resources: for instance explaining the account request procedure, how to access the user interfaces, obtain X.509 certificates or JSON Web Tokens for accessing the storage and computing elements via the Grid, compile software, use of container solutions interactively or via batch system, and more.

In addition to the user guide, the Group also provides some automatically-updated useful pages [14] as a form of *continuous documentation*. Such pages, automatically generated every night, aim to advertise specific information about the available services in a form that is easy to access and use. In particular:

- the list of *LCG environments* distributed via CernVM File System[13], CVMFS. As WLCG does not provide public documentation[2] on individual environments, it might take a long time for users to find one appropriate for their work. The page provides a table showing the compiler version, ROOT[15] version and Python support, together with a template of the shell command needed to load the environment;
- the list of *storage areas* for data transfer and management via Grid. As not all Virtual Organisation use the same transfer protocols, in this page all the storage endpoints are listed, with the exposed storage areas and their supported transfer protocols. For each storage area, the corresponding root path on the filesystem is also reported.

2.2 Communication

Communication represents a crucial part in the User Support daily activities and it is carried out in multiple forms:

- direct communication with users and experiments contact persons typically happens via e-mail;
- announcements regarding the Datacentre status, which are broadcasted via a mailing list;
- periodic meetings with the experiments. For instance, the monthly meeting "Comitato di Gestione" (*CdG* - "management committee") is an open meeting between the experiments' contact persons and the Datacentre, in which the most relevant issues and future plans are discussed;
- topical meetings with experiments' people to discuss issues or interventions that need a tighter cooperation, such as, as an example, the migration to different technologies.

Moreover, ticketing system are in place:

- GGUS [16], mainly used to interact with WLCG VOs.

[2]CERN login is required to access the documentation.

- an internal ticketing system based on Jira Service Desk[17], used to keep track of the internal activities involving other units of the Datacentre.

- an external ticketing system also based on Jira Service Desk, present as an a alternative communication channel to interact directly with users, although they seem to prefer a more direct approach in the form of an e-mail thread.

2.3 Procedures

The procedures carried out by the User Support members in their daily activities include:

- Handling account requests. This activity is very frequent, since almost every day new account requests are submitted, as shown in Fig. 3 below. In order to create a new account, a user must send an access authorisation form indicating what is the experiment or the community they need to access CNAF resources for, some personal details and a valid ID document number. Once the request is received, one of the User Support members proceeds to verify the new user's identity, by performing a *de-visu* recognition in case the new user is not already identified centrally by INFN. If the identity check is successful, the User Support member asks the corresponding experiment contact person for authorisation to proceed with the account creation. In case of positive verdict, the account request is forwarded to the CNAF system administrators who will eventually create it.

- Handling emails, tickets, and all the support requests, and trying to reproduce user's problems. This represents the core activity of the Group, which assists the user in finding a smooth and appropriate exploitation of the computing infrastructure.

- Security and performance monitoring. The Group has also the responsibility to ensure a correct usage of the resources from the security point of view. In particular, it is essential to monitor and supervise users, especially for violations of the Acceptable Usage Policy, e.g: personal account sharing, or running processes listening on network ports exposed to the Internet.

- Broadcast of announces. The User Support unit also takes care of sending announces on technical intervention that can reduce the overall Centre computing power, on scheduled down-times or on new features.

2.4 Workload

The figures below show respectively the number of local users per experiment registered between June 2017 and March 2023 and the local-users registration rate per month during the same period of time.

The number of CNAF users often reflects the number of researchers and students belonging to the scientific collaboration. This is especially true for small-medium collaborations, i.e. between 20 and 200 researchers. In larger communities - more than 500 people - Grid access is preferred and thus the number of local users is typically limited to a small group of computing experts.

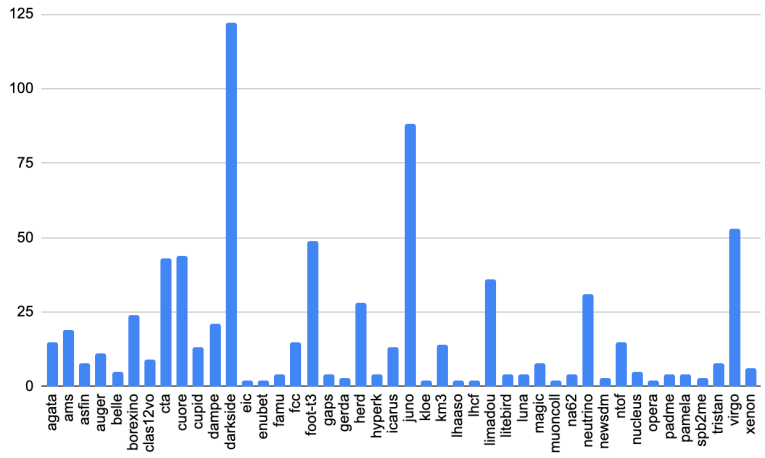

Figure 2: Local users per experiment cumulative since June 2017.

The rate of local-users registration exhibits some specific features:

- an annual minimum during August, coinciding with the Italian period of vacation;
- two annual periods of higher intensity, in Spring and Fall, that can be put in correlation with the start of thesis and PhD start periods.

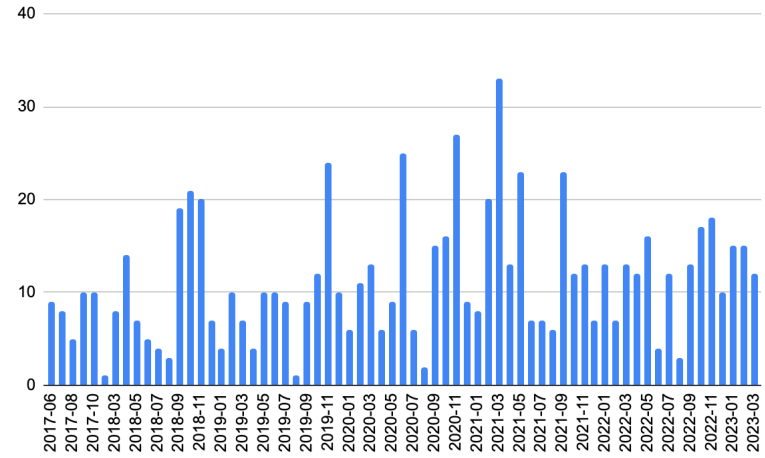

Figure 3: Local users registration per month since June 2017.

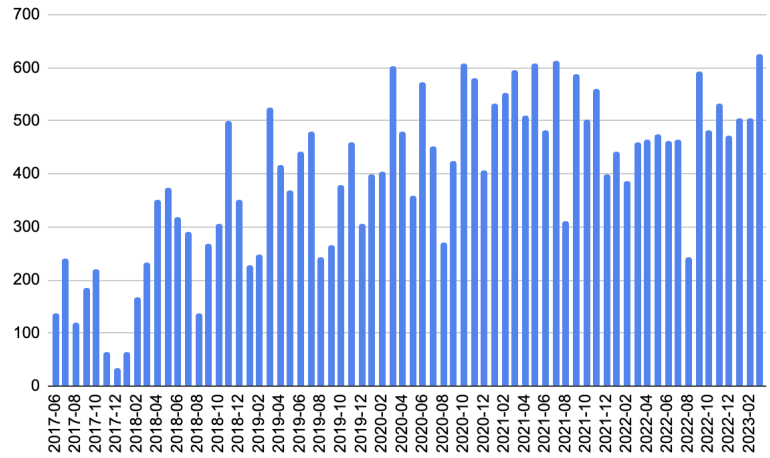

Figure 4: E-mail messages per month, both received and sent, since June 2017.

Thanks to these information, the effort on the user support group can be, to some extent, predicted. Indeed, Fig. 4 shows the number of e-mail messages, sent and received, since June 2017 and exhibits features similar to that of Fig. 3. The relatively low activity between November 2017 and March 2018 is due to the 2017 CNAF flooding event[18].

3. Future Plans

In this section the future developments involving the User Support activities, taking into account the increasing effort required in view of the new national INFN-supported project, and also considering the upcoming infrastructural changes CNAF is undergoing, will be discussed. Furthermore, in response to the continuously increasing amount of stored data and the expansion of the number of experiments demanding additional computing resources and technologies, a constant research for new tools and solutions, which can appropriately meet a larger number of uses cases, is needed.

3.1 Upcoming challenges

LHC experiments have recently entered the *RUN 3*[20], foreseeing a period of massive data production, requiring a very demanding computational power and storage capabilities, as shown in Fig. 5.

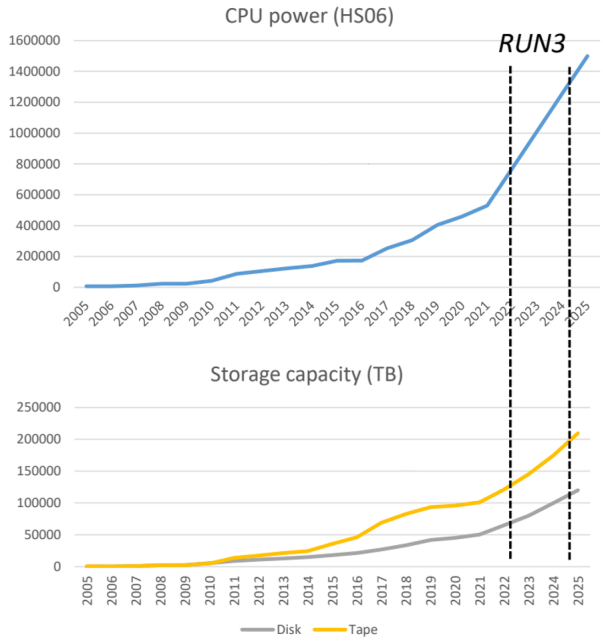

Figure 5: The expected increasing need for resources usage in the next years, both for CPU and storage.

In order to be able to support the foreseen increase of computing demand, the datacentre will be physically relocated to the "Data Valley Hub", a site located few kilometres north of the Bologna city centre, together with two larger datacentres, the European Centre form Medium-Range Weather Forecasts[21], ECMFW, and the Consorzio Interuniversitario dell'Italia Nord Est per il Calcolo Automatico[22], CINECA.

The new site, also named *tecnopolo*, will be hosting $\sim$ 80% of the total computing capacity in Italy and it represents the Italian Government and Emilia-Romagna Region largest investment in Big Data, Supercomputing and Research Infrastructure[23].

The presence of the three datacentres represents an opportunity to cooperate on large projects, and indeed several Italian national projects[24] will take benefit.

This will result in a number of challenges for the User Support group:

- large increase of the user base, both for the already supported experiments and also in response to the start of the new national projects;
- multiple infrastructures to support, due to the vast variety of hardware that will be employed, such as, High Throughput Computing systems, High Performance computing systems, either with and without accelerators (GPUs, FPGAs), Cloud-based systems;
- furthermore, it will be even more necessary to keep the central role in interfacing between the scientific communities and the computing ones.

To cope with aforementioned points, an increase of the staff is foreseen.

4. Conclusion

The INFN-T1 User Support Unit plays a central role in favouring the interaction of researchers with the computing services provided by the centre. It is involved in the definition of the computing model of several Physics experiments and provides guides and tools to allow the users to efficiently make use of the resources provided by the INFN-T1 datacentre. The increase of the workload in response to the new compelling INFN-supported national projects has been discussed.

References

[1] https://cerncourier.com/a/the-lhcs-worldwide-computer/

[2] https://twiki.cern.ch/twiki/bin/view/LHCOPN/WebHome

[3] https://wlcg.web.cern.ch/

[4] https://www.cnaf.infn.it/

[5] https://indigo-iam.github.io/v/current/

[6] https://l.infn.it/t1guide

[7] https://conda.io/

[8] https://apptainer.org/

[9] https://htcondor.org/

[10] https://indigo-dc.gitbook.io/oidc-agent/

[11] https://dmc-docs.web.cern.ch/dmc-docs/

[12] https://slurm.schedmd.com/

[13] https://cernvm.cern.ch/fs/

[14] https://www.cnaf.infn.it/ usersupport/

[15] https://root.cern/

[16] https://ggus.eu/

[17] https://www.atlassian.com/software/jira/service-management/features/service-desk

[18] https://www.cnaf.infn.it/en/datacenter-temporarely-down/

[19] https://www.cloud.infn.it/

[20] https://home.cern/events/launch-lhc-run-3

[21] https://www.ecmwf.int/en/computing/our-facilities/supercomputer-facility

[22] https://www.cineca.it/

[23] Legge regionale 17 giugno 2019, n. 7, https://demetra.regione.emilia-romagna.it/al/articolo?urn=er:assemblealegislativa:legge:2019;7

[24] https://home.infn.it/en/pnrr/infn-pnrr-presentation