

Activity report

2019



Centre de Calcul de l'Institut National
de Physique Nucléaire et de Physique des Particules



CC-IN2P3

2017-2019

The National Institute of Nuclear and Particle Physics (IN2P3) relies on the research infrastructure “IN2P3 Computing Centre” (**CC-IN2P3**) to accomplish its mission of setting up information systems allowing the storage, the processing and the valuation of all the scientific data concerned, as well as archiving it and making it available to the scientific community.

As a consequence of this, our activities range from operating a datacenter to providing resources for experiments in particle physics, astroparticles or nuclear physics. Among the major scientific

collaborations that use our services we can cite the four LHC experiments (ATLAS, CMS, ALICE, LHCb), GANIL / Spiral 2, LSST, CTA, KM3NeT. All appear on the national roadmap for research infrastructures as well as on the European ESFRI roadmap.

Operating an e-infrastructure designed to tackle the challenges coming from these communities requires a constant adaptation and evolution of our resources keeping in mind the next challenges that will come from the deluge of data of the forthcoming **High Luminosity LHC**



Summary

PREFACE	02
KEY NUMBERS	04
COMPUTING	06
STORAGE	08
SOFTWARE ENGINEERING	10
NETWORK	13
DATA CENTER	14
SUPPORT	16
RESEARCH	18
EUROPEAN PROJECT	20
COMMUNICATION	22
CONCLUSION	24

project at CERN, the new **SPIRAL 2** particle accelerator at GANIL, the **BELLE II** detector in Japan and several promising astrophysics projects such as the future **LSST** telescope, the **EUCLID** European satellite, and **CTA**.

This report does not aim to provide an exhaustive list or description of all the work made during the last two years. It rather highlights some of them in order to illustrate how CC-IN2P3 is continuously willing to improve its services to be at the best possible place.



KEY NUMBERS



HT. CORES*



SQUARE METERS
for 2 computing rooms



CURRENT MAXIMUM
STORAGE CAPACITY
100 PB of data stored



CARBON*
FOOTPRINT
(Estimated)



80



EXPERIMENTS

6.8M€



CNRS DOTATION
(Average on the last 3 years)

80



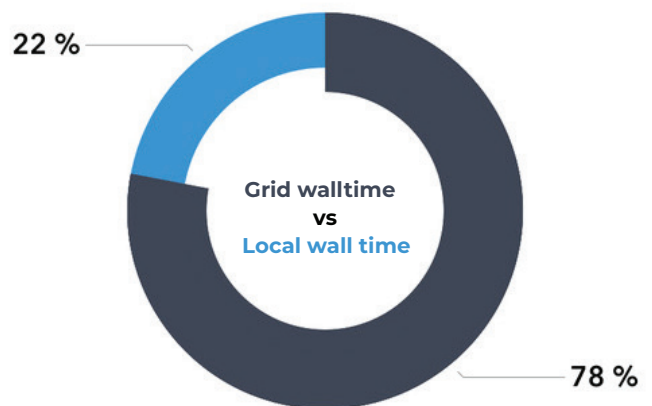
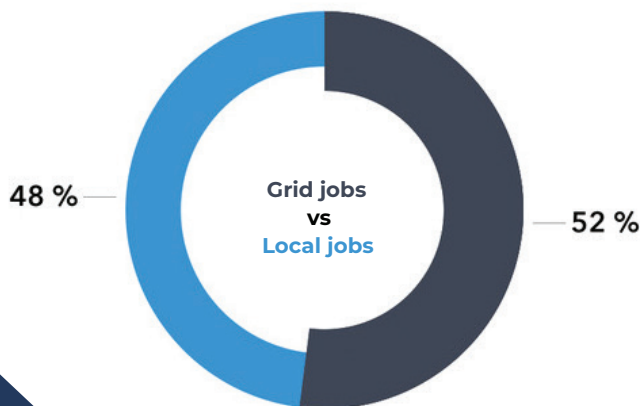
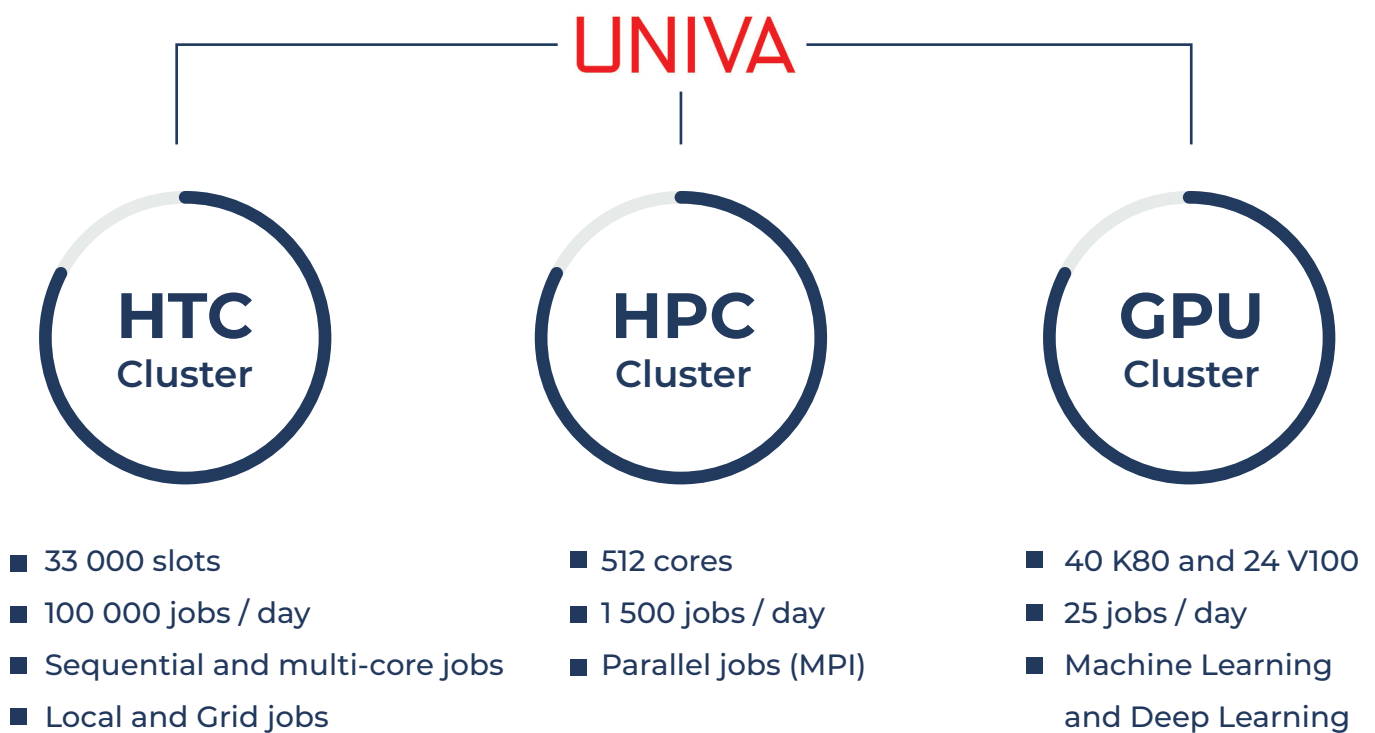
STAFF
16% short-term contract

HT: hyper-threaded

CARBON: annual electricity consumption converted in CO2 tons
(source : <https://www.electricity-map.org/>)

COMPUTING

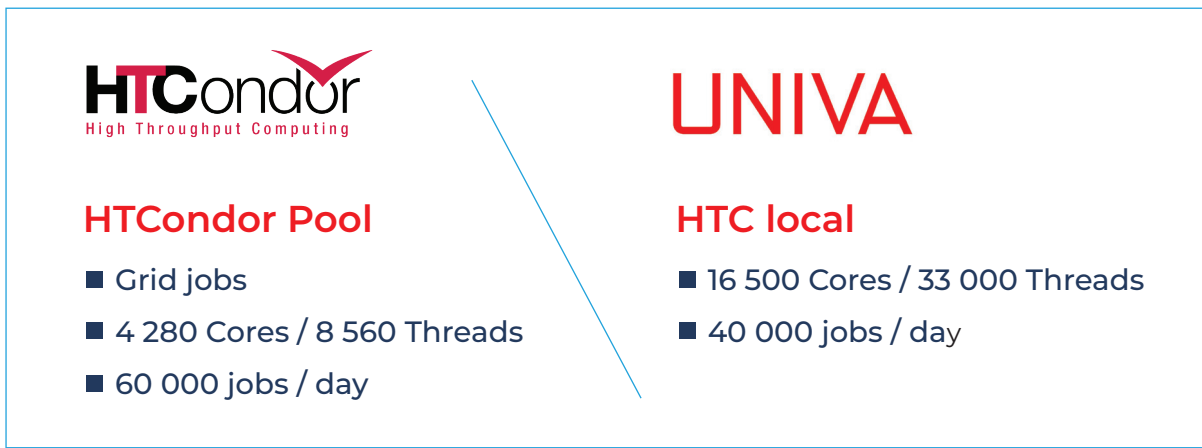
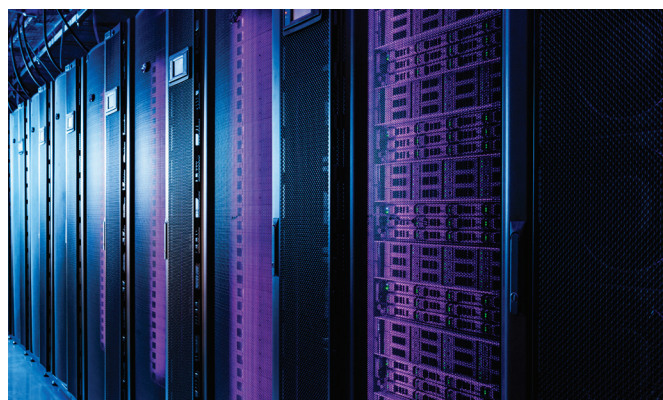
Our computing infrastructure is mainly operated using UNIVA Grid Engine as scheduling system. This infrastructure can be seen as three different logical clusters with well-defined tasks.



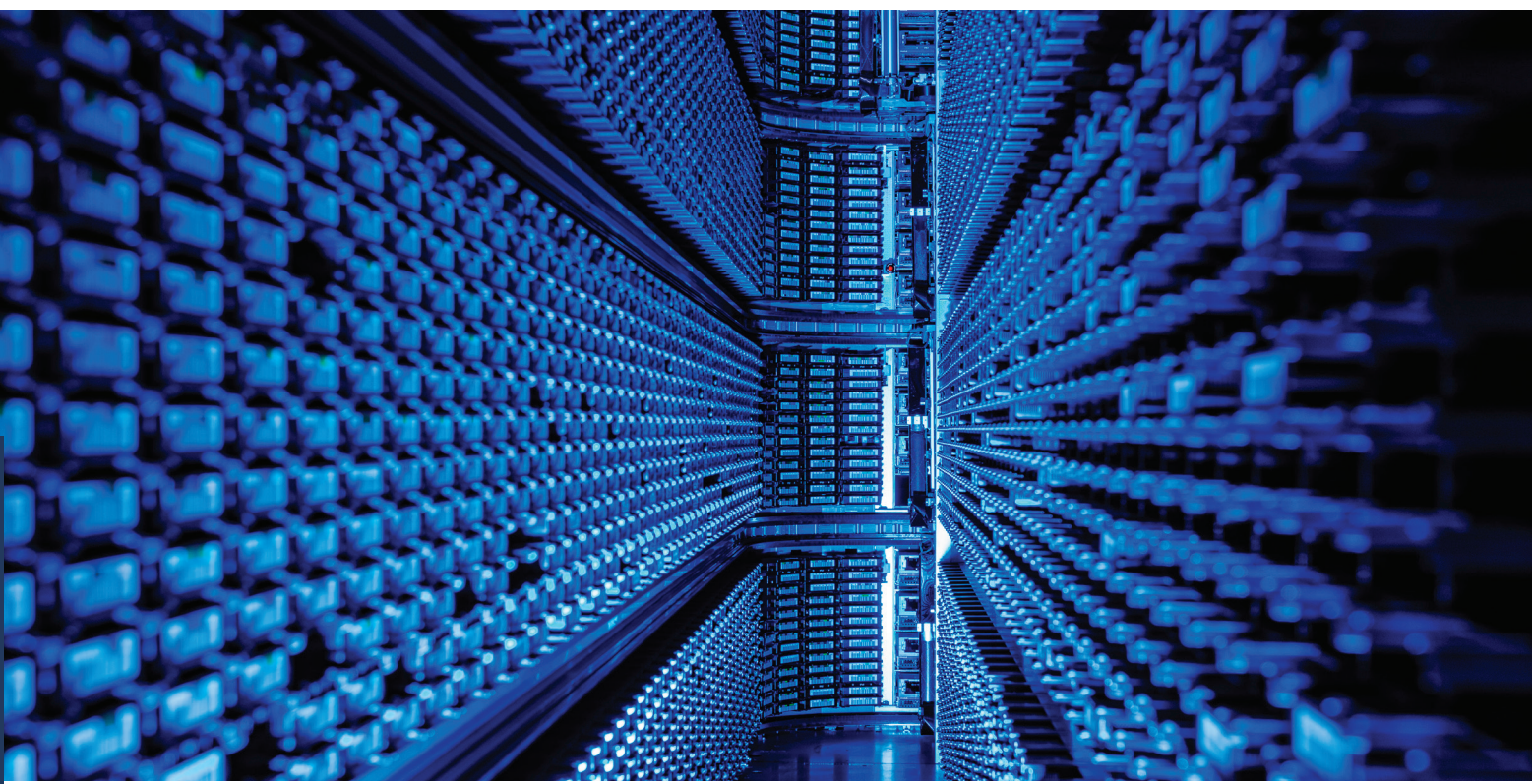
An evaluation of HTCondor as a alternative to Univa Grid Engine to manage a part of the computing infrastructure of the CC-IN2P3 has been conducted at the end of 2017. Taken in 2019, the decision to move to HTCondor was mainly motivated to contain the growing cost of the UGE license. Planed at first to be used for the GRID activities, it was also the opportunity to introduce a component widely deployed in the HEP community.

A study made by the Operation team and the CCLab showed clear differences between jobs coming through the Grid and those directly submitted by “local” users, in terms of arrival rates, resource requests, and resource usages, which advocate for a separation of the scheduling process.

A HTCondor cluster is in production since July 2019 and serves the needs of the four LHC experiments, while the HTC jobs from local users, parallel, and GPU jobs are still handled by Univa Grid Engine.



The question of migrating the jobs submitted by local users to HTCondor remains open. It would drastically change the way users interact with the resources of the computing center. Impact of this migration thus has to be carefully estimated first. In the longer term, whether to keep or replace Univa Grid Engine for the parallel and GPU jobs would also have to be decided, leveraging the experience of other comparable computing centers sites.



STORAGE

During the last four years, the amount of data stored on our various storage systems has increased by a factor of 3 for disk and tape media, from a grand total of 35 PB in January 2015 to 100 PB as of today.

This increase could have an important impact on our disk based storage mainly in terms of physical space used (number of racks) and power consumption. This led to two axis of evolution:

- Promote the usage of higher density disk servers
- Increase the usage of our mass storage on tapes even more

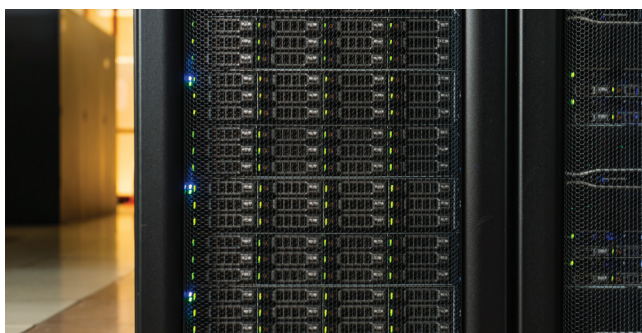
Relying more on our **tape** system implied to develop an efficient strategy to stage more quickly data from tape to disk. This has been achieved with the development of a tape request scheduler in order to sort the requests for reading tapes more efficiently. We have experienced some heavy activity, with up to 7000 tape mounts per day and a peak daily volume of 150 TB of data transferred from tape to disk.

Our tape libraries have been based on StorageTek / Oracle technology for more than 3 decades. However, the recent changes in Oracle strategy and their choice to give up their enterprise tape brand, forced us to define and choose a new direction for our future technical choices. Therefore, we have recently decided to purchase a SpectraLogic library using IBM enterprise tape drives. This library will gradually replace our Oracle libraries in the next few years.

The usage of filesystems on **disks** has evolved between 2015 and 2019: AFS which was used for the home of the users and software distribution has been phased out as its future was unclear. It has been replaced by NFSv4 served by a Network Attached Storage (NAS). The working space provided for short term living data was mostly based on Spectrum Scale[®] (IBM) in 2015. Following a change in the licensing policy, it has been decided to have this working space partially moved on a Dell EMC Isilon Scale-Out Network Attached Storage (NAS). This would be driven by the need of having better flexibility, scalability and ease of administration.

Cloud activity on OpenStack has increased in the last couple of years. A unified backend storage solution, both flexible and scalable, was needed in order to offer block storage, object storage and a filesystem interface. Ceph technology satisfies all these requirements and has been in production for more than a year.

Beside traditional relational **databases** service (RDBMS), new needs arise especially in the field of NoSQL database technology. This led to the creation of a MongoDB service and also some feasibility studies with other NoSQL technologies. Data mining being a key issue in data driven science, it has triggered an on-going effort on assessing Map-Reduce technologies such as Spark. A Spark testbed has been set up for our users and will be used in production for internal needs (fine grained storage statistics).



Overall, storing an ever growing amount of data and metadata for more than 100 scientific projects and thousands of users on more than 10 storage technologies implies questions that should be answered in order to improve the data life cycle management: describe data management policies, identify valuable data which should be archived and which is ephemeral, assess for how long data should be stored and kept safe, etc. Based on our experience and the realisation that a significant amount of data has become useless, we decided to build a **Data Management Plan** survey in order to make sure that the right data management policy is in place for every scientific project. This is still an on-going effort considering that all parties involved must be convinced that the DMP is a key point in the data life cycle. A daily data inventory helps in this activity. It is being improved using Map-Reduce technology in order to have fine grained information on all data stored at CC-IN2P3.

Archival, not just simple replication of data, has become a natural outcome of all this work on the data life cycle at CC-IN2P3. In order to satisfy this need which has been partially covered so far by our backup system, a mid-term preservation test bed implementing part of the OAIS standard is under assessment with real data.

SOFTWARE ENGINEERING

Web developments and infrastructure

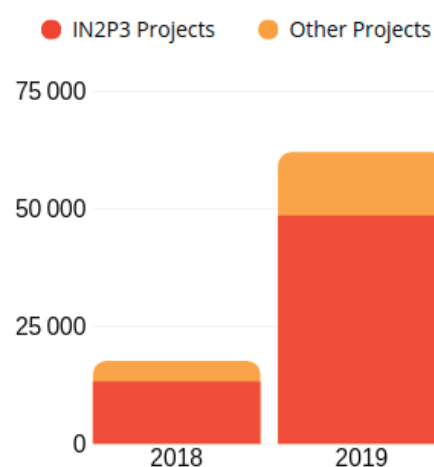
These last years have been marked by the outbreak of complex web developments.

Web development is a term that was previously used when people referred to the creation of several web pages and wikis for collaborations.

Today, it has a totally different meaning and is more likely to be used when people refer to web applications that rival the size and complexity of desktop applications.

This trend also applies to the **CC-IN2P3** and the software team has contributed to numerous complex web developments for internal needs, institutional needs and even European projects.

Number of Gitlab builds



Examples of significant projects:

- User Portal:** portals used to ease the daily operations of e-infrastructures developed in the context of European projects (EOSC-Hub, EGI)
- Operations Portal, ARGO:** portals used for the daily operations of european projects (EOSC-Hub, EGI)
- Hito:** web application initially dedicated to the follow-up of temporary funding contracts. Extended to the follow-up of permanent careers, it is now used by 3 others IN2P3 labs
- NSIP:** project management tool developed for IN2P3 collecting and exposing all the resources spent in the institutional projects

Following this trend, we are working on alternative web-based applications / services for users.

We are currently migrating our web hosting infrastructure to container orchestration technologies, .e.g Kubernetes. The aim is firstly to improve and automate the provisioning of web services for our end users but also to provide them on demand with services on an efficient cloud platform.

The main activity will be the migration of the existing web sites but also to provide and enrich the web services catalogue with multiple services from simple web static sites to CMS (Wordpress, Joomla...) and databases on demand.

This on-going activity will be the opportunity to replace the current web infrastructure but also a chance to increase our experience with container orchestration and envisage in the future to apply it to other areas like computing or storage.

Collaborative Tools

The Applications team selects and deploys software solutions to meet the needs of IN2P3. This particularly concerns collaborative tools. These tools allow members of IN2P3 projects to work together. They allow IN2P3 to maintain control over the information produced within the projects.

We have installed the following software solutions



Electronic document management based on Nuxeo software called "Atrium".



A project management tool based on Redmine + EasyRedmine software.



A collaborative software development platform based on GitLab software.



A collaborative notetaking tool based on the Etherpad software.



A survey organization tool based on the LimeSurvey software.



Zimbra is an enterprise-class email, calendar and collaboration solution.

We are finalizing deployments of 2 new tools:



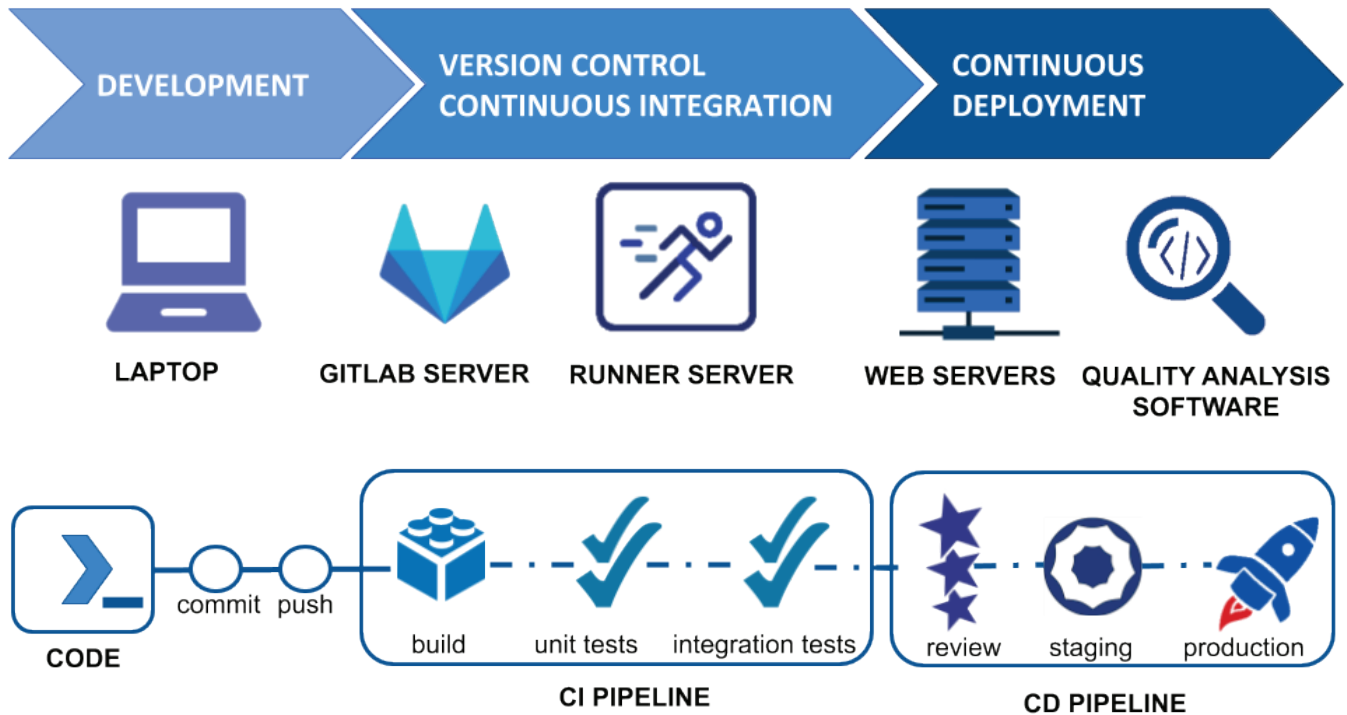
A collaborative Word, Excel and Powerpoint document editing solution based on nextCloud + OnlyOffice software.



An instant messaging solution based on Rocket-Chat software.

SOFTWARE ENGINEERING

DevOps philosophy



Cloud computing, mobile devices, efficient network are driving increased demand for code. That demand has compelled many organizations to transform their software development practices significantly and follows the DevOps philosophy. **CC-IN2P3** is also following this orientation.

A huge effort has been made to add automation: Continuous Integration, Continuous Testing, Continuous Deployment, and Continuous Delivery. This schema is illustrating how a local development could be deployed automatically after several steps of testing and checks.

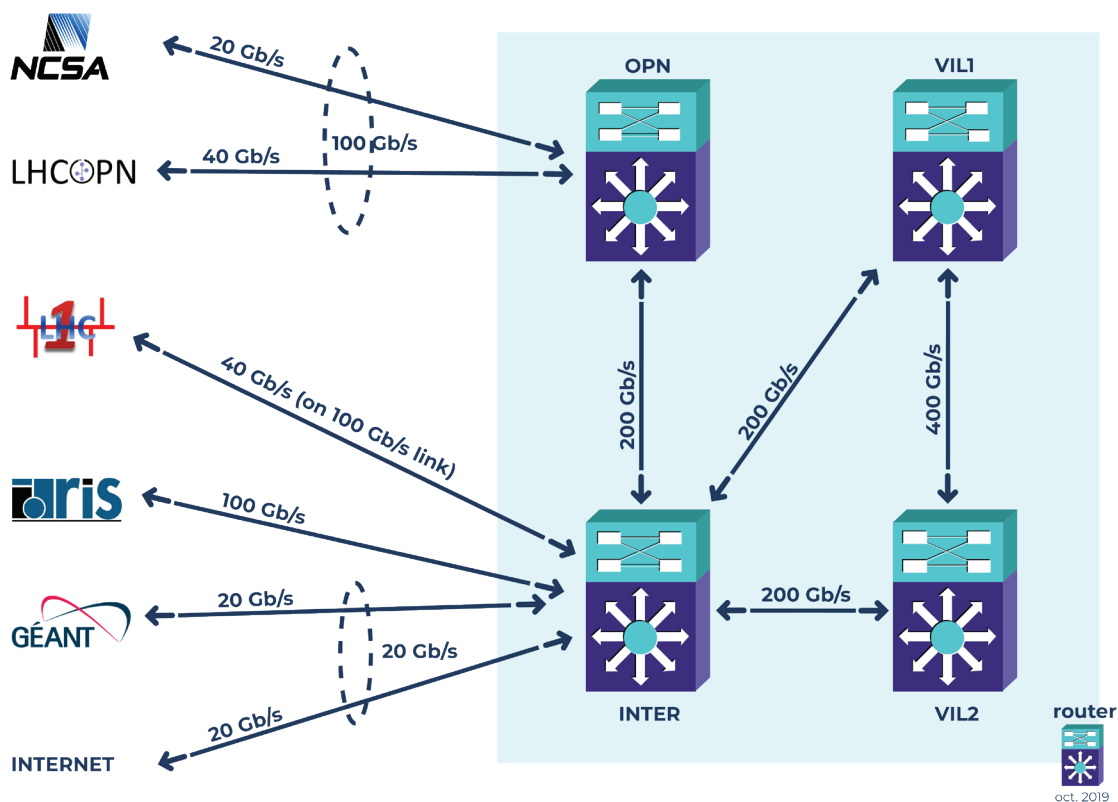
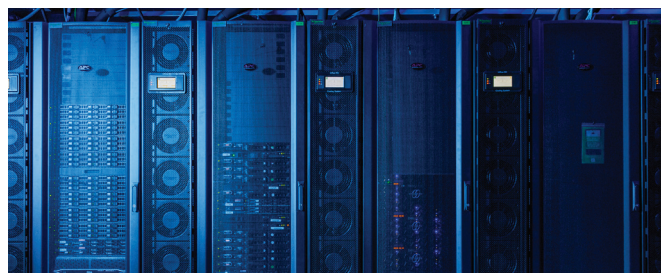
Automation is the ultimate need for DevOps practice and 'Automate everything' is the key principle of DevOps. In DevOps, automation kick starts from the code generation on Developers machine until the code is pushed to production and even after that to monitor the application and system in production. This philosophy is illustrated with the extensive use of some tools like GitLab (see graph below) and will be reinforced by the use of Kubernetes in the future. These changes are also reflected by a close collaboration with the system administration team.

NETWORK

Over the last few years, the ever growing need for network bandwidth inside the computing centre and with our partners led the evolution of our local and wide area networks.

In 2019, we prepared for the extension of our second computer room. In order to provide resources for the WLCG and other experiments, the number of storage servers with 10 Gb/s connections is continuously increasing. We also saw the installation of the first rounds of 10 Gb/s computing nodes. The connection between our two computer rooms was upgraded to multiple 100 Gb/s links.

For our external network links, we are working with RENATER on 100 Gb/s long distance links, to connect to IDRIS, to LHCOPN (CERN), to LHCONE (WLCG Tiers 1 and Tiers 2), and to NCSA (for the LSST).



DATA CENTER

INFRASTRUCTURE

Since 2017, we managed to cope with our IT needs while **improving hosting** and **upgrading our power and cooling infrastructure**.

Vil-2 became the preferred location for new equipment installation due to its better PUE. Skipping hardware installation in Vil-1 allowed us to free raised floor for structure service. An asbestos removal program preceded this long due maintenance. We lead it during summer 2019 without impact on IT services. We reconsidered the chilled water provisioning system of Vil-1. We replaced a 30 years old chiller with a new one (340 kW) with direct exchange with groundwater, providing heat for the rest of the building. Power consumption of the most recent unit has also been trimmed down through installation of speed regulation: cooling costs go down with drills slowing down.

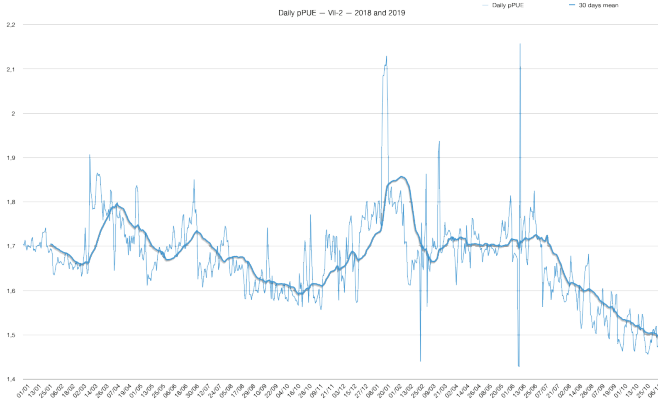
IT rack procurement and Vil-1 organization have been reviewed. Server cooling is evolving from raised floor technology to a more efficient hot aisle

corridor system. Demands for hosting increased, going up to 72 racks. As less power hungry systems of 7 kW racks with redundant PDU (ATS in option).

Concerning the newest building, our power and cooling upgrades have been driven by IT load. We improved its infrastructure in two steps: electrical upgrade in 2015-2016 (allowing 1.2 MW IT power) then cooling in 2019 (enabling 2.2 MW cooling). IT racks addition has been phased with our needs, after proper recycling of decommissioned servers.

We installed in 2015-2016 a second secured electrical chain maxing IT load up to 1.2 MW. We also introduced a redundant cooling network in 2019 on two hot aisle corridors. Hence, we now provide one hot aisle with a Tier II equivalent redundancy (up to 460 kW IT) and two hot aisles with a Tier III





Vil-2 Daily / Monthly Partial PUE



Vil-2 available space (racks and U)

equivalent (power and cooling fully redundant up to 1.2 MW IT). We had to twist our plans from 2011 but managed to maintain a final target of 3.2 MW IT. During december 2019, we installed 43 IT racks in a new hot aisle corridor in this fully redundant set.

We also improved energy efficiency for Vil-2. Cooling setpoints went up for cooling production and usage. Chillers went from 8°C-13°C to 15°C-20°C while server input moved around 27°C, without impact on reliability (the only concern left is for GPU servers, which rely upon lower input temperature). Warmer temperatures lessen energy bill.

We set up High Pressure control on every chiller of Vil-2. This reduced their consumption. Heat recovery has been in function for 2 years. We provide

calories from DC to a nearby chemistry building. While this is good for environment, it adds an over-consumption to our bill during winter.

Our last cooling group procurement is based on HFO fluid, which has a lower impact on global warming. One of those new chillers is also using freecooling: 100% freecooling below external temperature of 7°C. It's ROI is under 4 years.

Finally, although Vil-1 retrofit allows its rebirth as a privilege hosting target, its limitations on cooling and power distribution must be taken into account.



- VIL1: First computing room
- VIL2: Second computing room built in 2011
- PUE: Power usage effectiveness
- ATS: Automatic Transfert Switch
- PDU: Power Distribution Unit

SUPPORT

The “**User Support**” team helps interface business with IT, through various activities including training, service documentation, resource demand management, and the daily user support tasks (communication and service desk activities).



Training

One of the most significant changes that we introduced is the organization of trainings, which has considerably ramped up since early 2019 (nine trainings proposed this year). These trainings target end users and cover various topics, including programming languages, use of GPU devices, databases, storage, etc. The usual audience represents about 25 attendees, and often the demand cannot be fully covered. We usually send a survey after each training, in order to get feedback from the attendees and new ideas for future trainings. Such trainings appear to be really appreciated by users, and we should keep on proposing them.

Single entry point

Another strategic choice was to start building a unique entry point where users can connect and get access to most of the information related to their own activities and resource usage, as well as pointers to recent news, documentation, collaborative tools etc. This is addressed by the User Portal: we started the project in early 2017 and delivered a first version in the spring of 2018. Since then we implemented and deployed several improvements, but the project is not done yet as several new features still need to be provided. Over the last year, the portal has been used in a rather steady way, with about 50 user connections per day. As of today, the User Portal only provides information to connected users, in a read-only mode. In parallel, two services are under development in the Application team: the self-service project for account

management and, in the longer term, the PAIR project which aims to provide an fully-interactive service to access the resources. We believe the resulting products will have to merge with the User Portal at some point, to provide a more complete entry point for users. Effort needs to be engaged on that respect.



The new documentation system

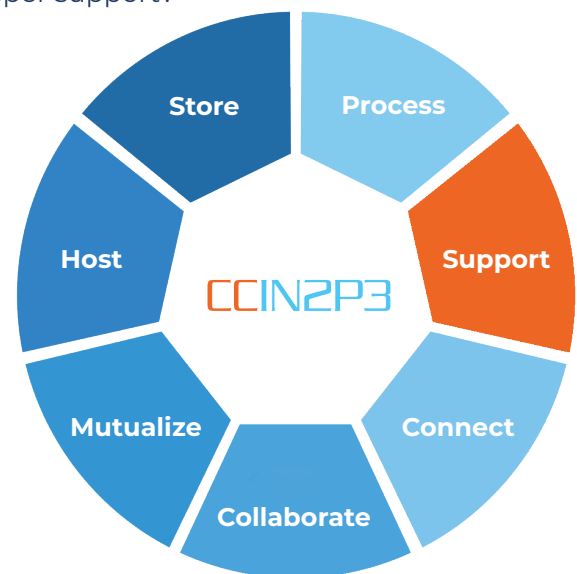
One of the main projects currently carried out by the User Support is the new documentation system. We delivered significant effort since early 2019 to provide a more modern interface for user-dedicated documentation. As of now, most of the current documentation content has been ported to the new system, and some work still remains to finalize the project. We expect the new system to be ready for production by early 2020.

Another important project carried out by the Support Team is the implementation of a set of tools that provide an aggregated view of user activities and data. The latter part is especially important from an account management point of view, to take appropriate decisions related to user (i.e. scientific) data when deleting accounts: that helps speed-up a heavy procedure. The product was delivered about three years ago and has been maintained since then.

Last, because of the rising demand in virtualization functionalities, the User Support has worked on the provisioning of container-related services for end users. Singularity was integrated directly to the batch service, and we put in place a catalogue of Singularity images that cover most of the demand.

To conclude and anticipate for the near and mid-term future, we would like to emphasize one of the recent growing trends in demand for GPU devices. CC-IN2P3 currently provides compute servers equipped with nVidia graphic cards, and

these servers are quite solicited. We expect such demand to keep on increasing, due to the rising needs coming from astro-particle experiments. One of the key questions that arise now is whether User Support should invest time in training and getting expertise on programming software on such devices, in order to be able to provide a deeper support?



Finally, we will end the yearly “experiments day” where users and experiments representatives would come to CC-IN2P3 to hear about our services and expose their use cases. Instead, members of the User Support team will visit the IN2P3 laboratories in their own premises. This way, we expect those visits to be an opportunity to target more audience and better publicize the services provided by CC-IN2P3. We will certainly meet more users and have more time to discuss their needs. Such visits should not only help improving the interaction between the Center and its end users, but also show that we are here to help. This will of course require time and effort from our staff. We wish to start this new approach in early 2020.

RESEARCH

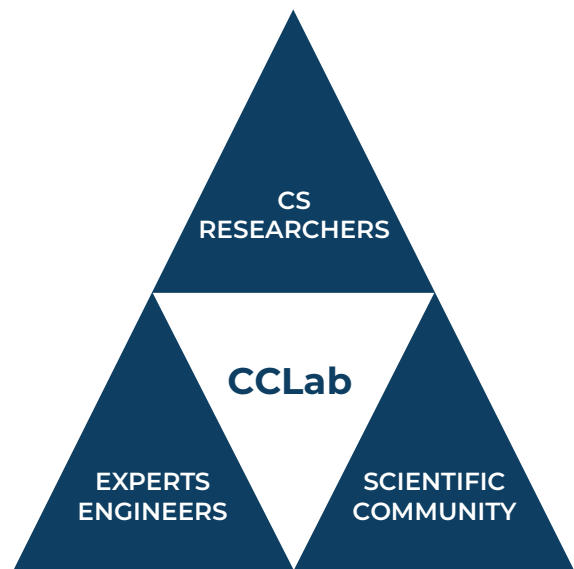
A computer Science research lab in a Computing Center

The computer science of the CCLab reflects the traditional setup of a research laboratory with a mix of different profiles and skills focused on a common goal.

The objective of the CCLab is to identify specific and concrete research actions for which a dedicated task force is composed.

Examples:

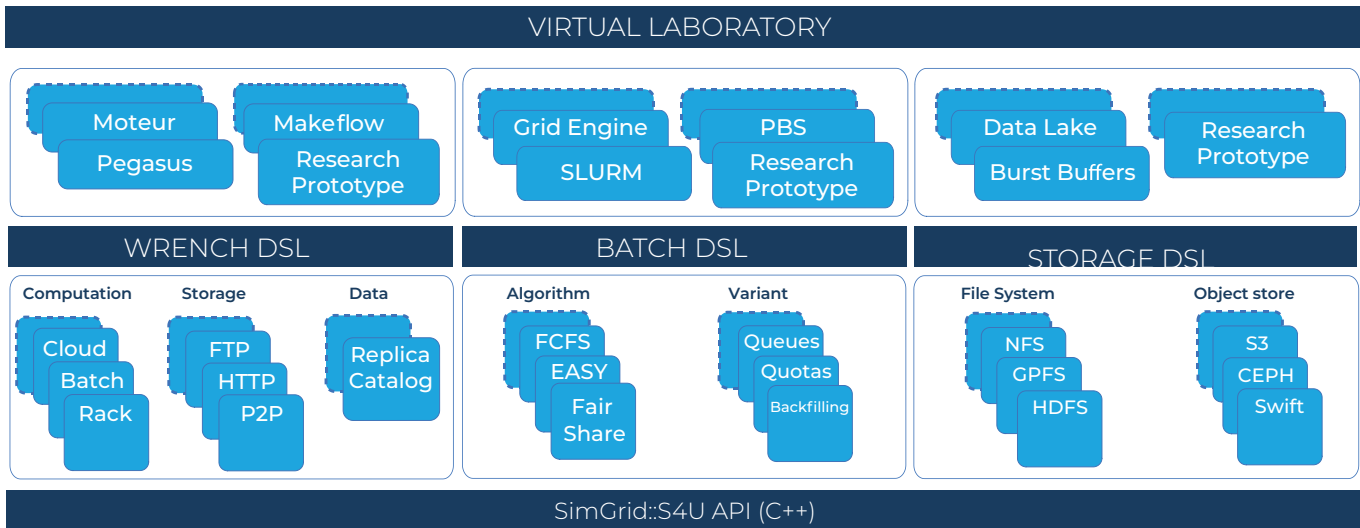
- Improvement of data management for medical imaging workflows
- Analysis of the workload processed at CC-IN2P3 and optimization of the batch scheduling system



A virtual laboratory for data intensive sciences performance evaluation

Over the next few years, the IN2P3 Computing center will have to increase its capacities by at least an order of magnitude to answer the combined requests for computing and storage resources made by the experiments it serves. This increase will come along with a diversification of the needs and usages. Building on the success of the SimGrid and Wrench projects, the proposed virtual laboratory will help the CC-IN2P3,

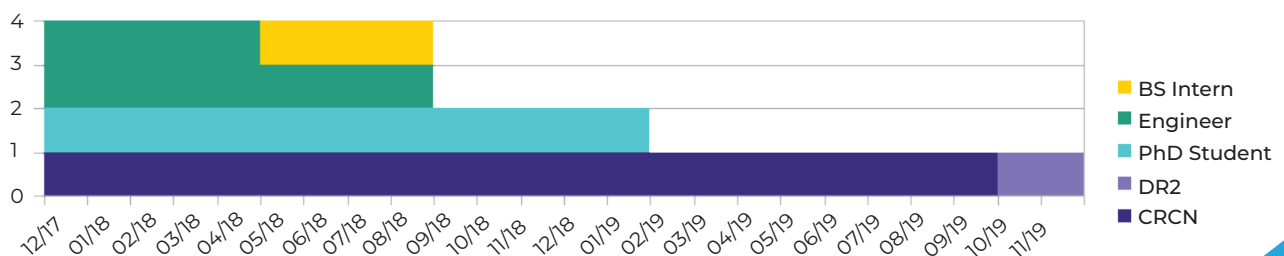
regional datacenters, and the IN2P3 laboratories to evaluate in an objective way the cost-performance ratio of a software or hardware upgrade of their infrastructure, hence optimizing their investments. It will also help the user communities to identify and evaluate different workflow management systems and make an educated decision based on objective indicators.



Developing the CCLab

Since the last ESC, the number of members of the CCLab has slowly decreased to be reduced to F. Suter, who has been promoted on a «Directeur de Recherche» position. This mainly corresponds to the disengagement of European projects too disconnected from the research activities mentioned during the last ESC.

A new position of Junior Researcher will be opened in 2020. Hiring another permanent researcher with a different background and area of expertise will be an opportunity to broaden the scope of answers the CCLab can provide. Two research topics are considered: Data Science and management of large amounts of data. New Master and PhD students will also be hired in 2020 to increase the manpower of the CCLab.



EUROPEAN PROJECTS

Why participating?

For many years, Europe has devoted significant effort and budget to shape an ambitious computing landscape. The purpose is to propose a large range of distributed resources addressing multidisciplinary scientific fields with maximal interoperability and minimal access penalty, building the so-called EOSC, European Open Science Cloud.

HIGH ENERGY PHYSICS

ASTRO-PARTICLE

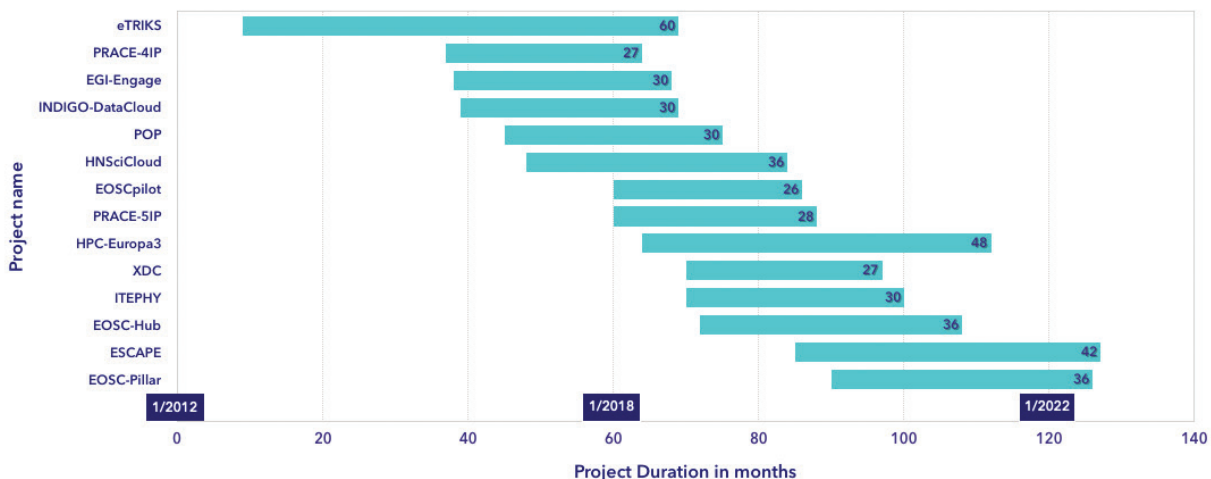
NUCLEAR PHYSICS EXPERIMENTS

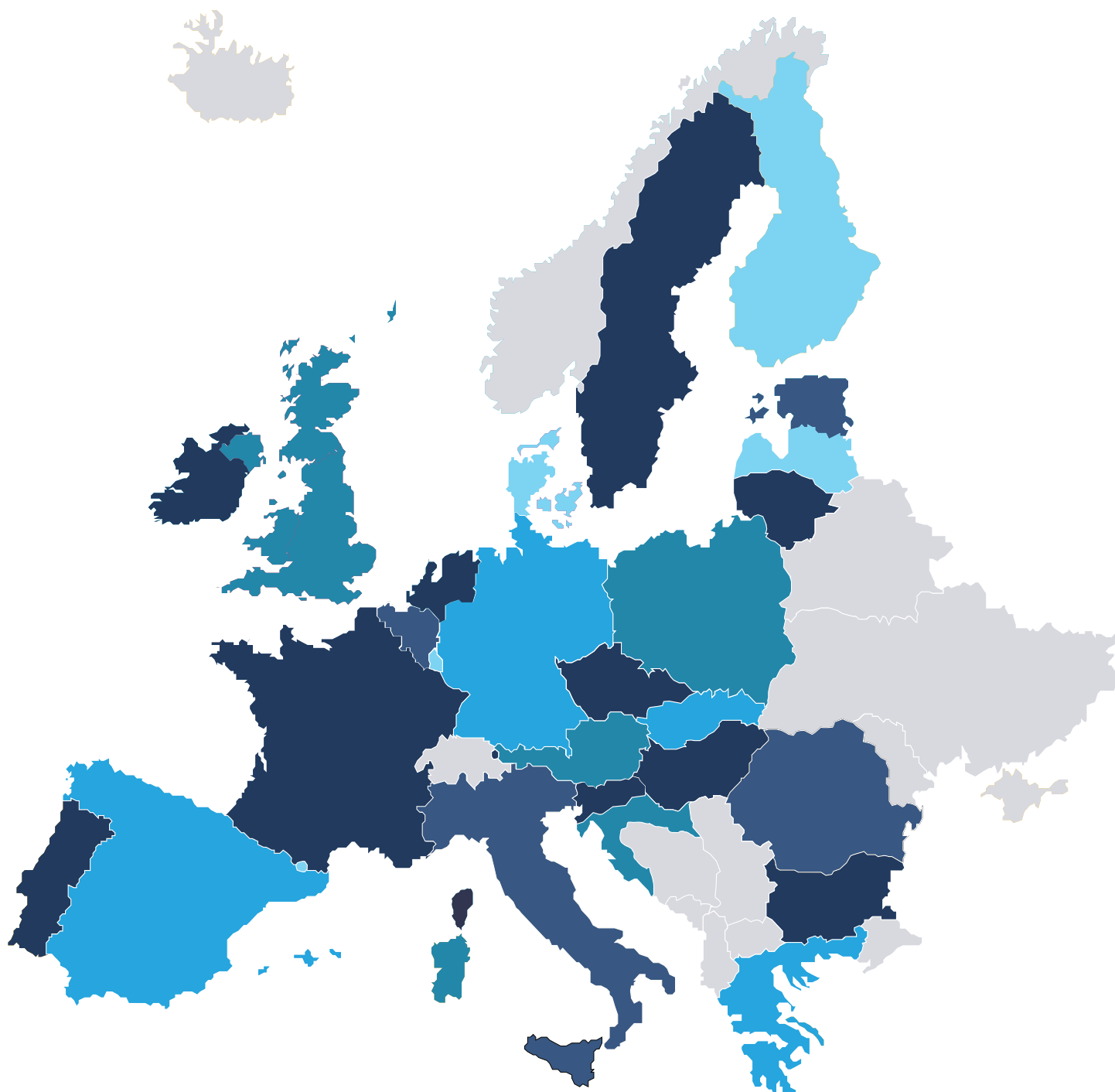
CC-IN2P3, as a major national infrastructure and a host to large high energy, astro-particle and nuclear physics experiments is primarily concerned by this issue and in the best position to contribute.

We have been involved since early 2000 in the projects that shaped the computing on the grid (DATAGRID, EGEE). In a strategy to open the centre to multidisciplinary sciences, we designed cloud platform for translational research (eTRIKS project) and enriched it with all innovative technologies developed for our mainstream projects.

Many other projects followed where we built new competences and tackled new technologies to enrich the catalogue of services offered to our scientific community: from parallel and GPU computing with the PRACE and HPC-Europa3 projects to the use of commercial clouds in HNSciCloud and recently the research and development towards future data-lakes with XDC and ESCAPE.

EUROPEAN PROJECTS GANTT CHART





European projects are a unique opportunity to build strong collaborations with other European data centres, participate together into providing the ambitious amount of resources and services requested in the coming thirty years by the HEP and Astro-particles communities and including all the long-tail of sciences into this process.



COMMUNICATION

Nowadays, with information constantly flowing from all sides, communication is an essential activity for a CNRS research infrastructure such as CC-IN2P3. Explaining its role, demonstrating its dynamism, and presenting its expertise are today essential to be a known and recognized actor within its scientific community, but also by other institutions and general public.

Goals of the external communication:

- Promote the technological infrastructure and expertise of the teams
- Demonstrate the role of CC-IN2P3 in large physics experiments
- Play a noticeable role in the local community of research and general public

Goals of internal communication:

- Support the CC-IN2P3 strategy and its explanation inside the unit
- Facilitate exchanges between members and teams
- Help its members to move together towards the same objective



As the website is relatively up-to-date, online communication is currently mainly done through the twitter account. In addition to targeted communications on events, news about training, technical changes, appointments, and open positions are posted.

Based on these objectives, the communication at CC-IN2P3 is organized around key recurring events:

The **Festival Particule.com** is a scientific mediation event supported by CC-IN2P3, organized once every two years in collaboration with local laboratories during the Science Fair. It is an opportunity to propose innovative mediation supports (theatre play, virtual reality, games, etc.).

The purpose of the **Journées du CC** is to bring all staff together to conduct a collective reflection on the situation of the CC-IN2P3, its functioning, the achievement of the various challenges and future commitments, while improving interpersonal relations.

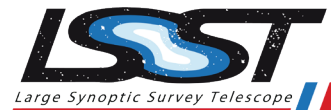
Each year, CC-IN2P3 leads a communication project at **Supercomputing**, the largest conference in the field of high performance computing, networking, storage, and analysis. It sets up a stand on the exhibition floor with IDRIS and GENCI. The aim is to promote the CC-IN2P3 to this international audience, bringing together researchers, industrials, French and foreign partners.

Obviously, apart from these key events, other actions are taken according to the actuality. 2019 was an opportunity to celebrate the 30th anniversary of the Web and to recall the role of CC-IN2P3 in the development of the first French web server. This important date was celebrated at two separate events: one for the media and the general public, which provided an interesting press coverage, and the other for a school audience, with an action to raise awareness on the use of the web today.

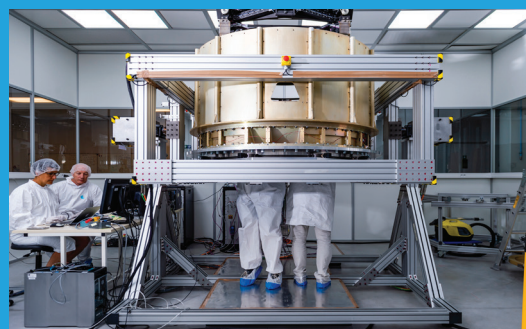
Over the last two years, a lot of work has been done on image. CC-IN2P3 now has many pictures that can be used on different media, especially social networks. A reorganization of the communication unit made it possible to recruit a person on a fixed-term contract with a multimedia skills. The use of video and motion design animations is for example becoming widespread at CC-IN2P3 and allows for new communication opportunities.

The communication team is continuing its partnership with the Dell EMC company which, in addition to providing the CC-IN2P3 with a technological platform, allows us to benefit from relatively recurrent sponsorship of events.

Throughout the year, the communication team coordinates internal seminars, visits of the computer rooms and the CC-IN2P3 museum, requested by partners, students, or CC-IN2P3's users (several hundred people per year).



CC-IN2P3 dedicates a part of its staff to the communication of the **LSST experiment**. Over the last two years, several major events (LSST@Europe3, LSST DESC conference) have been organized by the head of the CC-IN2P3 communication team. An objective in 2019 was also to promote the realization of the filter exchange system built by the IN2P3 laboratories, and now transferred to SLAC to be integrated with the other elements of the camera. The filter exchange system has been photographed at each stage of the construction. These pictures were used for a storytelling on Twitter and a press action.



CONCLUSION

In this report we choose to present some of the most visible part of all the works and activities that took place at CC-IN2P3 during the last two years. This doesn't mean that others works were not as important: they are also contributing to our approach to continuously improve our services. For CC-IN2P3, as production data processing center, improving or offering new services requires some skills and adaptability that, if we're used to, may always be seen as a too long process for our users. Our guideline is to the propose as best as possible experience to our users, which at the first order is synonym of reliability.

We also dedicate part of our work time to study and test new technologies that may allow the experiments to handle, in the future, the unique data coming from the detectors they conceived in the framework of the international collaborations they are part of. Therefore, the relevance of such R&D can only be reached by cooperating with comparable computing centers all over the world and listening carefully to the researchers needs.

1, 3, 9 / *Journées du CC 2019* including a visit of the CMS experiment

2/ 30th anniversary of the Web

4/ *Journées du CC 2017*

5/ Visit of the technical infrastructure after the electrical and cooling works

6/ International Women's Day 2019

7/ Visit of the computing room during the Science Fair

8/ Theatre play at *Festival Particule.com 2017*

10/ Science Fair 2019

11/ Picture of maintenance work taken by CNRS Images

12/ Webcast team during the finale of *Ma thèse en 180 secondes*

13/ Supercomputing 2019 (Denver, US)



Depuis 80 ans, nos connaissances
bâtissent de nouveaux mondes

Centre de Calcul de l'IN2P3 / CNRS

21 Avenue Pierre de Coubertin
CS70202
69627 VILLEURBANNE cedex
FRANCE
Tél. : +33 (0) 478 93 08 80
Fax : +33 (0) 472 69 41 70
<http://cc.in2p3.fr>