

Distributed scientific computing ecosystem in France

Michel Jouvin – michel.jouvin@ijclab.in2p3.fr

Irfu workshop on computing strategy

Saclay, September 23, 2026

Outline

- High Througput Computing
 - The grid: a first attempt at a European/worldwide infrastructure
 - The cloud era
- The national and European HPC ecosystem
- EOSC
- MesoNET and MesoCloud
- Irfu : possible roles

Who am I?

- Joined LAL (now IJCLab) in 92... and almost retired!
 - Computing Department head: 2013-2022
- Involved in grid activities since 2001
 - GRIF site creation for LHC experiments in 2005: technical lead until 2023
 - France Grilles
- Leading role in UPSud/UPSaclay VirtualData creation (2011)
 - Collective effort around a shared facilities and shared resources
- Involved in cloud projects since 2010
 - First cloud at UPSud based on StratusLab then migrated to OpenStack (2016)
- Involved in UPSaclay mesocentre creation in 2020
 - VirtualData cloud + ENS/CS RUCHE + a global storage platform
- Involved in MesoNET project since 2021
 - Includes a participation to MesoCloud IR proposal... and now its steering group

The grid...

- First attempt to build a large-scale, data-intensive, multidisciplinary infrastructure
 - 98: computational grid concept launched by I. Foster & C. Kesselman with a toolkit: Globus
 - Supported by several European projects since 2020: EDG, EGEE, EGI
 - Similar projects in the US: PPDG followed by OSG projects since 2005
- High Throughput Computing (HTC) paradigm: a worldwide batch system
 - Goal: optimize the global job throughput rather than the individual job exec time
 - Unable to support efficiently massively parallel jobs: parallelism restricted to server cores
 - No efficient support of interactive jobs
- GRIF : a large grid site in Paris region built and managed by 4 labs
 - Irfu and 3 IN2P3 labs (LLR, LPNHE, LAL+IPNO/IJClab)
 - Significant amount of resources: among the largest in Europe
 - A rich network of expertise bringing together experts from different labs
 - More than 20 years and still alive!

... The grid

- WLCG : the LCG computing international collaboration
 - Based on a MoU signed by each site: resource commitments through *yearly pledges*
 - Bottom-up organization/collaboration: working groups, technical forum, management board
 - Importance of the human network: people know each other
- Several successes but failed to get a large adoption
 - Difficult to accommodate very different execution environments required by different communities

The cloud era

- The cloud concept emerges with Amazon Web Services in 2006
 - At the origin, completely disconnected from scientific computing
 - The core technology is the virtualization: environment seen by applications/users can be tailored as needed
 - Nothing like a batch system: either a machine that you need to configure or a “service”
- Several open-source projects started to provide a cloud middleware
 - 2 main ones: OpenNebula (2005, Europe) and OpenStack (2010, US)
- Academic community started look at clouds as a way of exposing site resources
 - **Conversely to the grid, not a technology to federate resources around the globe**
 - Worldwide commercial clouds are owned by one company...
 - CERN adopted the cloud technology (OpenStack) as the base infrastructure to deploy (almost) all its computing services in 2012
 - Kubernetes/DevOps emerging as the new approach to provision cloud (and other) resources

EGI / Frances Grilles

- European Grid Initiative (EGI, 2010) : a European infrastructure integrating all the national grids
 - Sharing the same grid middleware making the European infrastructure a global one
 - Extensions in Asia-Pacific
- Frances Grilles GIS started in 2010 as the French member of EGI
 - In fact started with the first European Datagrid project in 2001 and EGEE projects after
- ~2015: both extended to include a cloud-based infrastructure
 - A federation of clouds rather than one cloud but no real metacloud
 - Coexisted alongside the grid infrastructure, sharing the authentication technology/service and giving access to the grid storage
- Users: serving small/medium communities
 - So-called *long tail of science*: not necessarily well-structured communities
 - A lot of (grid) resources shared with WLCG but WLCG not part of EGI

Grid & clouds lessons learned

- WLCG
 - An agile governance is needed with a strong bottom-up component and the necessary top-down structures to be able to enforce decisions
 - Significant technological changes must be possible without a global disruption of the infrastructure
- France Grilles
 - Success: a rich network of expertise, collaboration between site administrators and (advanced) users/domain experts
 - Weakness: a top-down political governance based on research institutions, “disconnected” from the people running and using it
- Clouds
 - An interesting technology to expose site resources to users not requiring federated resources
 - Commercial clouds: access to (huge) resources for bursting or spiky workflows

French HPC : GENCI et mésocentres

- French HPC ecosystem started to develop in mid-2000
 - 3 big facilities quickly established: CEA/TGCC, CNRS/IDRIS, University/CINES (HPC T1)
 - Several universities started to buy small/medium HPC clusters (HPC T2)
- 2006: launch of the mésocentre concept
 - A not well-defined name to designate various resources for scientific computing
 - **Strong focus: local user community, support and machine experts close to users**
 - Mainly HPC resources but a grid site like GRIF was labelled mésocentre in 2006
 - During many years, a loose collaboration to share experiences with 1 meeting/year
- Increased incentive to have a federation of resources at regional level
 - Initially effort put on regional datacenters (with some “big/dense” regions having 2 or 3)
 - Mésocentre focused on scientific computing: not necessarily coordinated with emerging regional datacenters
- Recently, several mésocentres developed cloud-like resources for data analysis

European HPC : Prace / EuroHPC

- European collaborations developed in the HPC world
 - Similar to what happened in the HTC world
 - In parallel to the HTC projects and infrastructures: very little links, no global strategy
- Mid-2000 : PRACE collaboration
 - Partnership for Advanced Computing in Europe
 - Bottom-up, science-driven, European collaboration building a European network of national big facilities
 - 2016 : registered on ESFRI roadmap
- 2018 : EuroHPC Joint Unit (legal entity)
 - Co-finance acquisition of leading HPC machines (HPC T0) in Europe, deployed in a different country for each generation (Alice Recoque, 2027, at TGCC is one of them)
 - Strong collaboration with PRACE to organize the user community around these facilities

European Open Science Cloud (EOSC)

- A major European initiative to integrate all the digital infrastructures for scientific computing
 - Computing : EGI and PRACE
 - Data : EUDAT
 - Network: GEANT
 - Open science : OpenAIRE
- Promote interoperability and ease of access to all infrastructures
- Strong focus on open data and FAIRness
 - FAIR: Findable, Accessible, Interoperable, Reproducible
- A marketplace where users can find the resources they need
- In fact, existing infrastructures still live quite independently..
 - EOSC has been very top-down so far, despite the French EOSC association...
 - EOSC node concept being pushed to provide more integrated resources at the national level

MesoNET: the project

- An Equipex+ project (2021-2027) to build a national federation of mésocentres
 - ~20 partners: 1 per region + a few exceptions + GENCI
 - Rationale: impossible/non optimal that all regions have the resources/technologies (HPC, HTC, HPDA, AI, Quantum, ...) for all use cases and the expertise around them
 - Support the training of students (master) and research staff to develop the required expertise
- Budget : 30 M€ half-funded by ANR
 - Procurement of ~10 machines, including future architecture prototypes (*code-formation*)
 - A global storage infrastructure usable by all resources and distributed on ~14 sites
 - **~10 CDDs located with a commitment from partners to make them permanent positions**
- Prepare the future beyond the project
 - Technical R&D: computing architectures/technologies, storage (Rucio, S3)...
 - Design and propose a permanent infrastructure beyond the project

2 Université des Antilles

3 CINAURA (Université
Grenoble Alpes, FLMSN)

4 GIP numérique de Bretagne

5 Université de Bourgogne
Franche Comté (UBFC)

6 Université d'Orléans
Fédération CaSciModOT

7 Université de Tours
Fédération CaSciModOT

8 Université de Corse
Pascal Paoli

9 Université de Reims
Champagne-Ardenne

10 Université de Strasbourg

11 Université de Lille

12 Université Paris Saclay

13 Centrale Supélec

14 ENS Paris Saclay

15 Paris sciences et lettres (dont
Observatoire de Paris)

16 CRIANN (Rouen)

17 Université de Bordeaux

18 Université de Toulouse
(Cal mip)

19 Université de Montpellier
(Meso@LR)

20 Aix-Marseille Université

21 Université Côte d'Azur

22 Centrale Nantes

23 Lyon 1

MesoNET: achievements

- **A great achievement: build a national network of expertise**
 - The right project at the right time: previous projects created links between some people
 - GSM : Groupe de Support Mutualisé
 - Several technical working groups: access portal, authentication, storage
 - **Commitment to give permanent positions** to the experts recruited and trained by MesoNET
- Significant deployment of new resources in ~10 regions
 - Not an infinite lifetime but most of them deployed in the last 2 years...
- More and more university courses using resources not necessarily local
- A portal developed allowing a central management of projects and affiliated users
 - Used by project to request access to MesoNET resources: get a yearly allocation on them
 - Developed as an asset for the future beyond the project

MesoCloud: the project

- Proposed as an **Infrastructure de Recherche** during the 2025/2026 update
 - Well received and accepted last Spring, now in the building phase
- MesoNET experience used as a basis but extended to include France Grilles
 - France Grilles as a structure stopped end of 2025 (prolonged several times...)
 - France Grilles has a very close collaboration culture despite technology differences (grid, clouds) and a different governance model compared to MesoNET
 - A few sites/people belonging to both infrastructures
 - Work in progress since last Spring to identify the concrete technical topics that could be used to make the merge concrete
- Governance model keeping the strong bottom-up component of MesoNET
 - Led by CNRS as a multi-disciplinary organization used to drive global initiative
 - Members are sites providing resources (their institution) and not user communities
 - Supported by a CNRS UAR (Unité d'Appui à la Recherche) to be created and providing a legal entity

MesoCloud: status

- Goal: MesoCloud started before MesoNET end (October 2027)
 - Smooth/adiabatic transition planned
- Elaboration of the Consortium Agreement
 - Spring 2025: ~20 partners expressed their intent to join the consortium
 - A first draft of the principle circulated last Spring, v2 sent beginning of September
 - Expected to last a long time... always difficult to sort out last details, will start with a lighter proto-agreement
 - Next milestone: Cargèse workshop, November 11-12
- CNRS UAR setup: hope to get it ready by the end of 2027
 - Tight schedule as CNRS top management changed last summer...
- Technical discussion started with France Grilles (FG) to identify issues that could be worked on together
 - Authentication, storage (FG and MesoNET run an iRods infrastructure)...

Possible CEA/Irfu role in this landscape

- CEA and CEA/Irfu have been important actors in this history
 - CEA/Irfu, involved in all the CERN/LHC experiments, is one of the GRIF site founders and is hosting ~35-40% of the resources
 - CEA has been one of the France Grilles members (national research organisations)
 - CEA is a major actor in the French HPC ecosystem
- CEA and CEA/Irfu also had difficulties in some of the projects and infrastructures
 - No significant involvement in cloud operations, despite a good expertise on how to use them (e.g. how to use efficiently Kubernetes for workflow deployments) as demonstrated in SVOM and SKA
 - Absent from MesoNET project as it is not involved in a mésocentre
- Momentum to link with these new efforts in the MesoCloud context
 - Expertise in how to use distributed computing infrastructure is an important asset: CEA/Irfu has developed a quite unique one for large projects with SVOM and now SKA
 - CEA being a member of Paris Saclay, could join the national effort by a presence in the UPSaclay mesocentre on this particular expertise, something at the core of the mesocentre

Conclusions: ecosystem

- A diverse but federated scientific computing ecosystem has been developed in France for 25 years
 - Provide access to the various kind of existing computing technologies
 - Allow users to store their data in a global storage federation usable from all resources
 - Global authentication service
- Historic divide in France between HTC and HPC is over
 - MesoNET project has been a game changer, MesoCloud paves the road to the future
 - Researchers need both, several sites diversified their resources
 - AI to play a major role in future scientific workflows: new hardware requirements that are neither pure HPC or pure HTC
- New scientific projects requiring access to federated resources, e.g. SKA
 - CERN/LHC is no longer unique

Conclusions : Irfu

- Irfu developed an expertise at the interface of infrastructures and scientific domains, e.g. SVOM and SKA
 - Often the missing bit at sites...
 - Critical for big communities and complex workflows: cannot expect end users to tackle of the issues related to an efficient use of distributed resources
- UPSaclay is a dense ecosystem with a big mésocentre
 - A unique opportunity to provide both significant resources in the global ecosystem and a rich network of expertise around it (currently: Maison de la Simulation, UPSaclay DSI)
 - Suggestion: connect more officially the work done at CEA/Irfu/DEDIP to the UPSaclay mésocentre, as part of this network of expertise
 - Could include some of the computing infrastructure administrators to participate to some aspects of the UPSaclay/VirtualData cloud administration