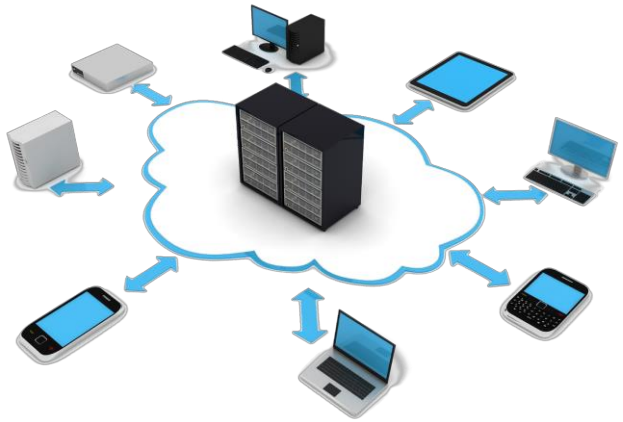


SILECS/SLICES

Super Infrastructure for
Large-Scale Experimental Computer Science

Christian Perez – LIP/Inria

Slides from F. Desprez – Inria & S. Fdida – Sorbonne University

Convergence of Computation and Communication

Computation

1950/1990 - Mainframes

1975 -Personal Computers

ENIAC
1946

1950 - Batchmode
1960 - Interactive

1967 - First virtualisation attempt

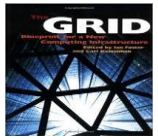
1970 - Terminals (clients/server concepts)

1999 - Salesforces SaaS Concept

2002 - Virtualised Infrastructure

2010 - Cloud democratisation

1999 - The Grid



2002- Amazon Initial Compute/Storage services

2006 - Amazon EC2 (IaaS)

2015
Fog / Edge
Software Defined XXX

Transistor
1947

micro processor
1971

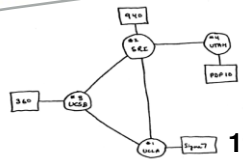
1838 - Telegraph

1876 - Telephone

1957 - satellite

1896 - Radio

1969 - ARPANET



1985 - TCP/IP Adoption
1973 - Ethernet

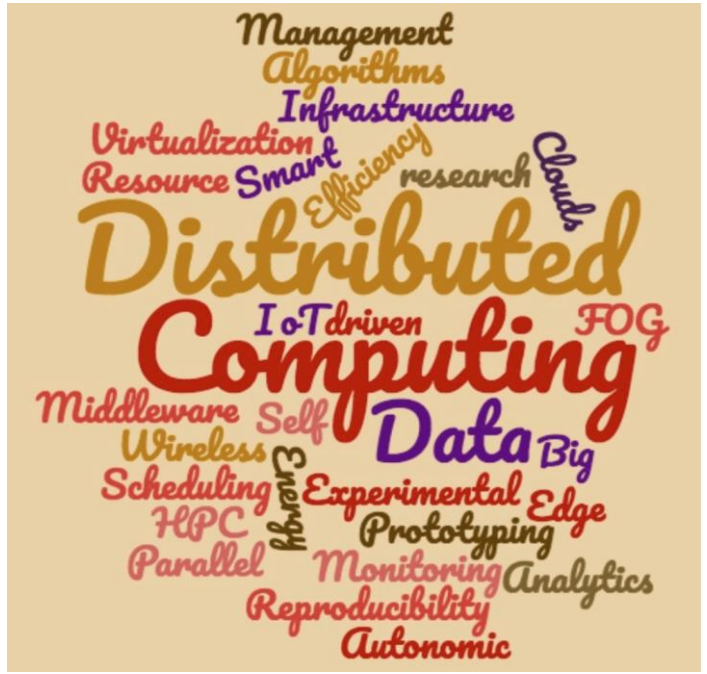
1995 - Commodity clusters

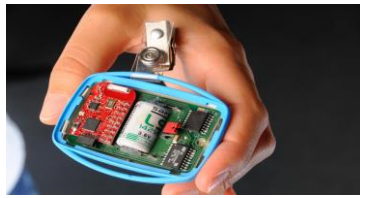
SmartPhones
2007

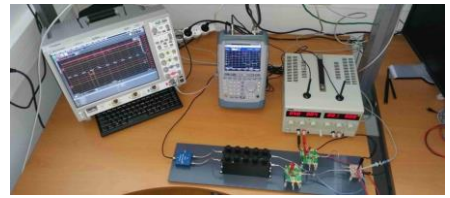
Communication

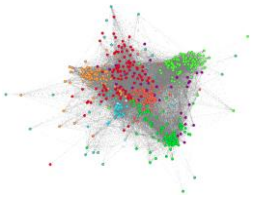
Motivation

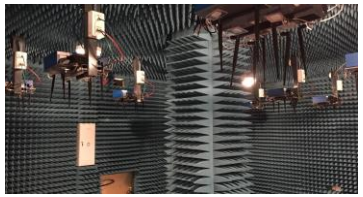
- Exponential improvement of
  - Electronics (energy consumption, size, cost)
  - Capacity of networks (WAN, wireless, new technologies)
- Exponential growth of applications near users
  - Smartphones, tablets, connected devices, sensors, …
  - Large variety of applications and large community
- Large number of Cloud facilities to cope with generated data
  - Many platforms and infrastructures available around the world
  - Several offers for IaaS, PaaS, and SaaS platforms
  - Public, private, community, and hybrid clouds
  - Going toward distributed Clouds (Fog, Edge, extreme Edge)



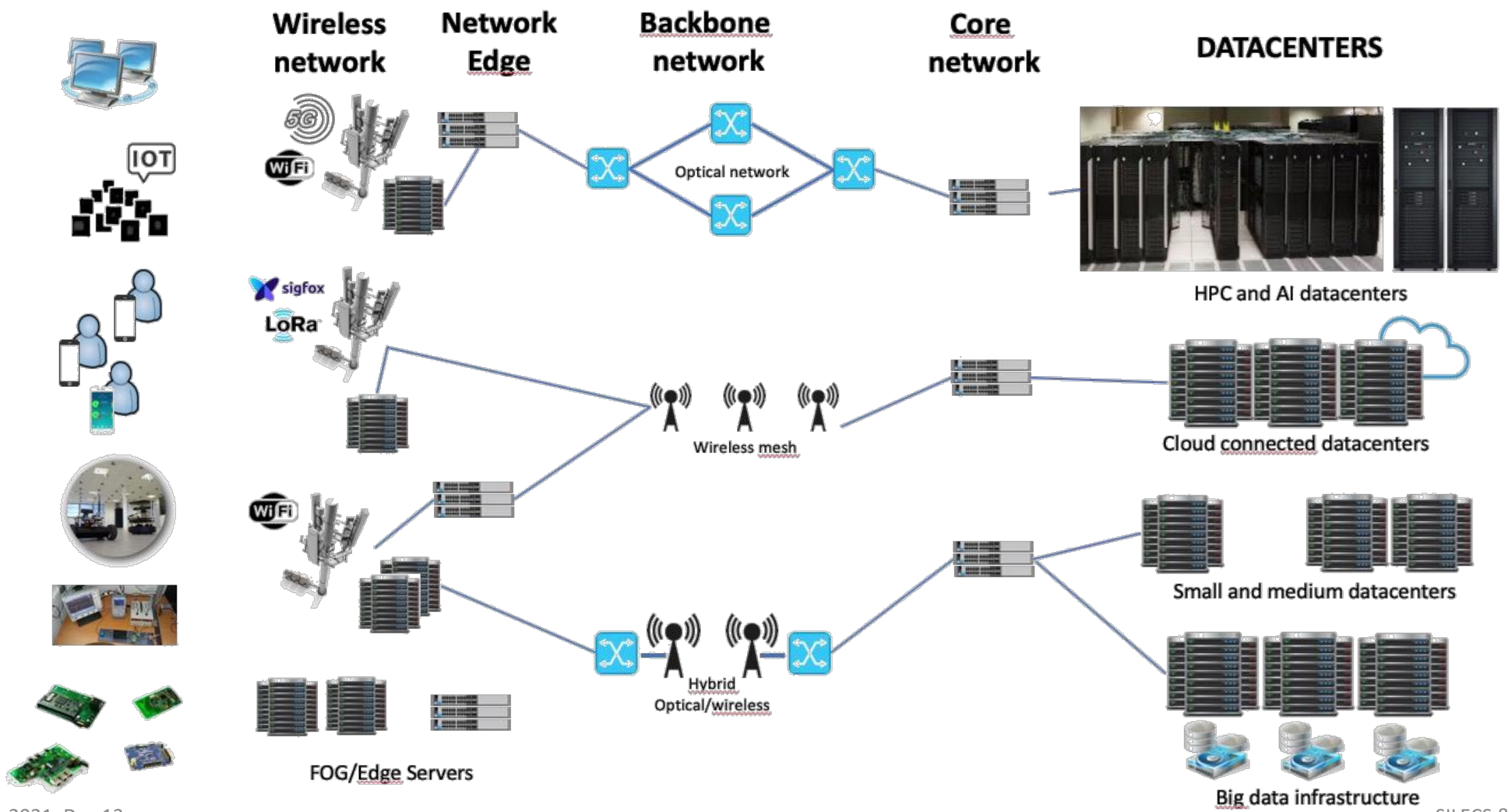








Digital Environment



The Discipline of Computing: An Experimental Science

The reality of computer science

- Information
- Computers, networks, algorithms, programs, etc.

Studied objects are more and more complex

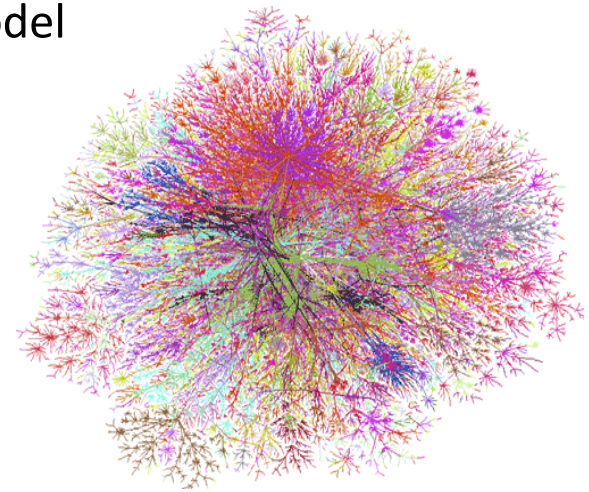
- Hardware, Systems, Networks, Programs, Protocols, Data, Algorithms, **…**

Experimental Validation: A good alternative to analytical validation

- Provides a comparison between algorithms and programs
- Provides a validation of the model or helps to define the validity domain of the model

Several methodologies

- Simulation (SimGrid, NS, …)
- Emulation (MicroGrid, Distem, …)
- Benchmarking (NAS, SPEC, LINPACK, ….)
- Real-scale (Grid'5000, FIT, FED4Fire, Chameleon, OpenCirrus, PlanetLab, …)



Good Experiments

A **good experiment** should fulfill the following properties

- **Reproducibility**: *must* give the same result with the same input
- **Extensibility:** *must* target possible comparisons with other works and extensions (more/other processors, larger data sets, different architectures)
- **Applicability:** *must* define realistic parameters and *must* allow for an easy calibration
- **"*Revisability*":** when an implementation does not perform as expected, *must* help to identify the reasons

ACM Artifact Review and Badging

https://www.acm.org/publications/policies/artifact-review-and-badging-current

SLICES and SILECS

Need of specific platforms to experiment

- To measure how programs behave and not only of the results they produce
- To (dynamically) change the execution environment (up to generate real faults)
- Tier 0,1,2 only enable to execute « *safe* » programs

European level: Slices

- The first European the first Research Infrastructure in computer science

French level: Silecs

- Based upon two existing infrastructures: Grid'5000 (HPC/cloud) and FIT (wireless/IoT)
- On the f*euille de route nationale des Infrastructures de recherche* since 2018
  - https://www.enseignementsup-recherche.gouv.fr/pid25366/acces-thematique.html?theme=317&subtheme=318

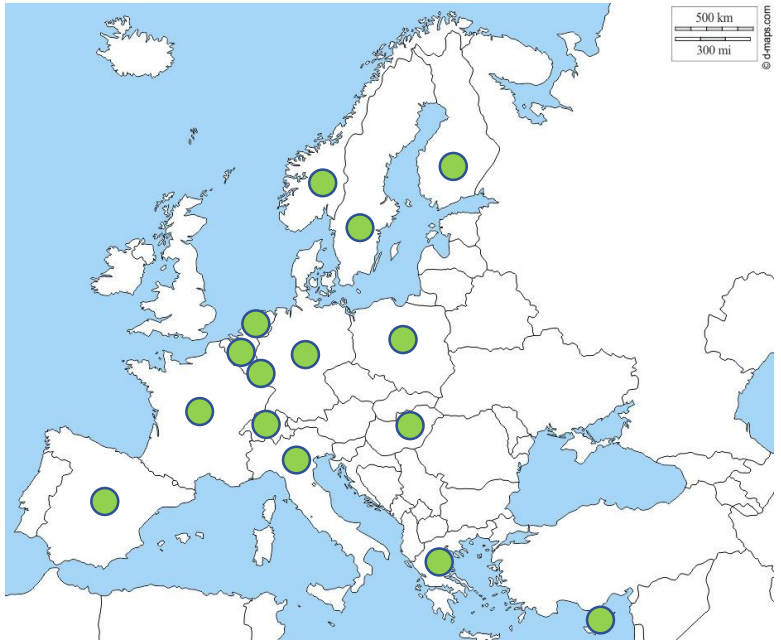
SLICES – ESFRI Project in 2021

Flexible platform designed to support
large-scale, experimental research focused on

- networking protocols
- radio technologies
- services
- data collection
- parallel and distributed computing

and in particular

- cloud and edge-based computing architectures and services.



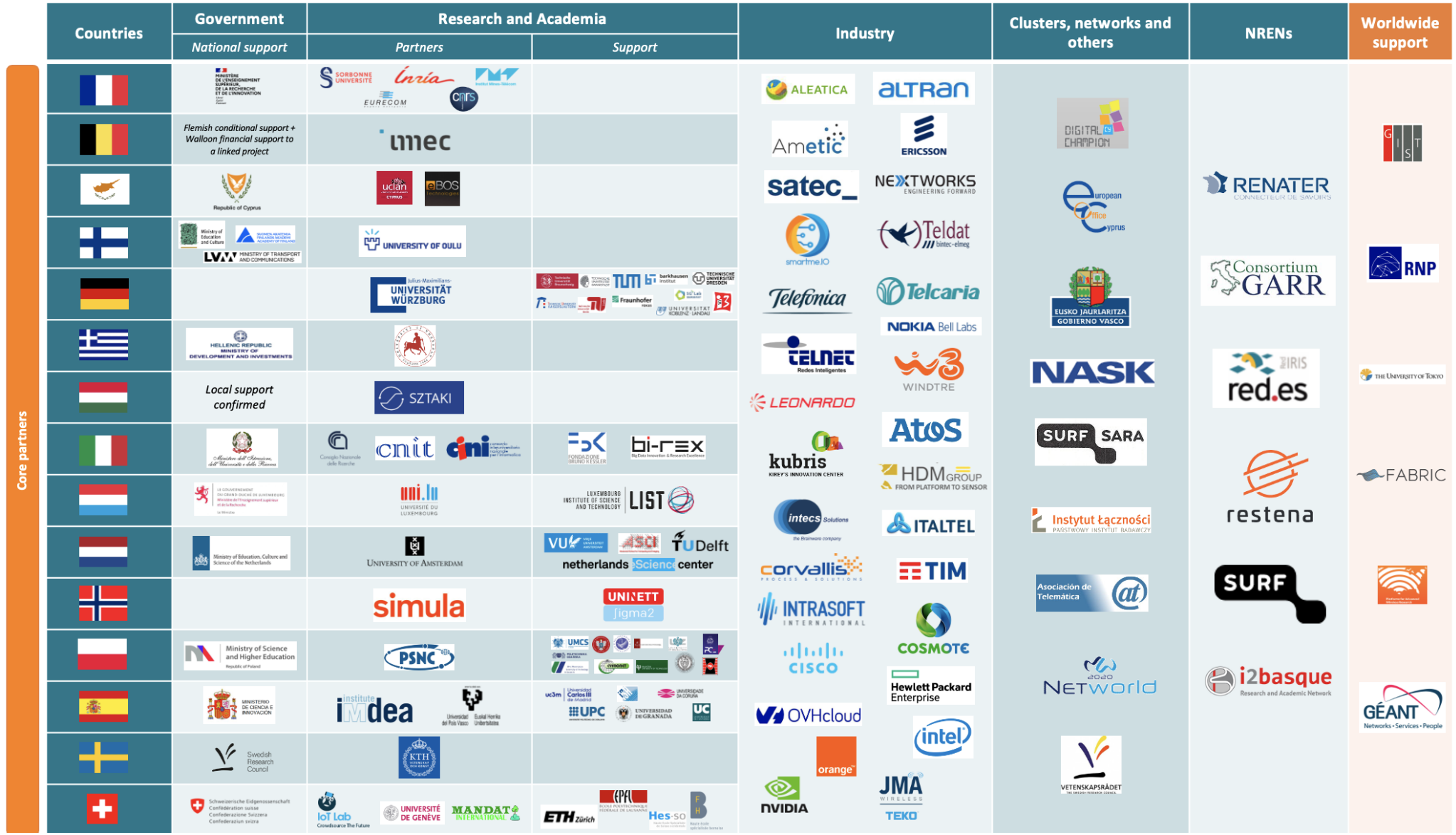
25 Participants from 15 countries

- Belgium
- Cyprus
- Finland
- **France (leader)**
- Germany
- Greece
- Hungary
- Italy
- Luxembourg
- The Netherlands
- Norway
- Poland
- Spain
- Sweden
- Switzerland

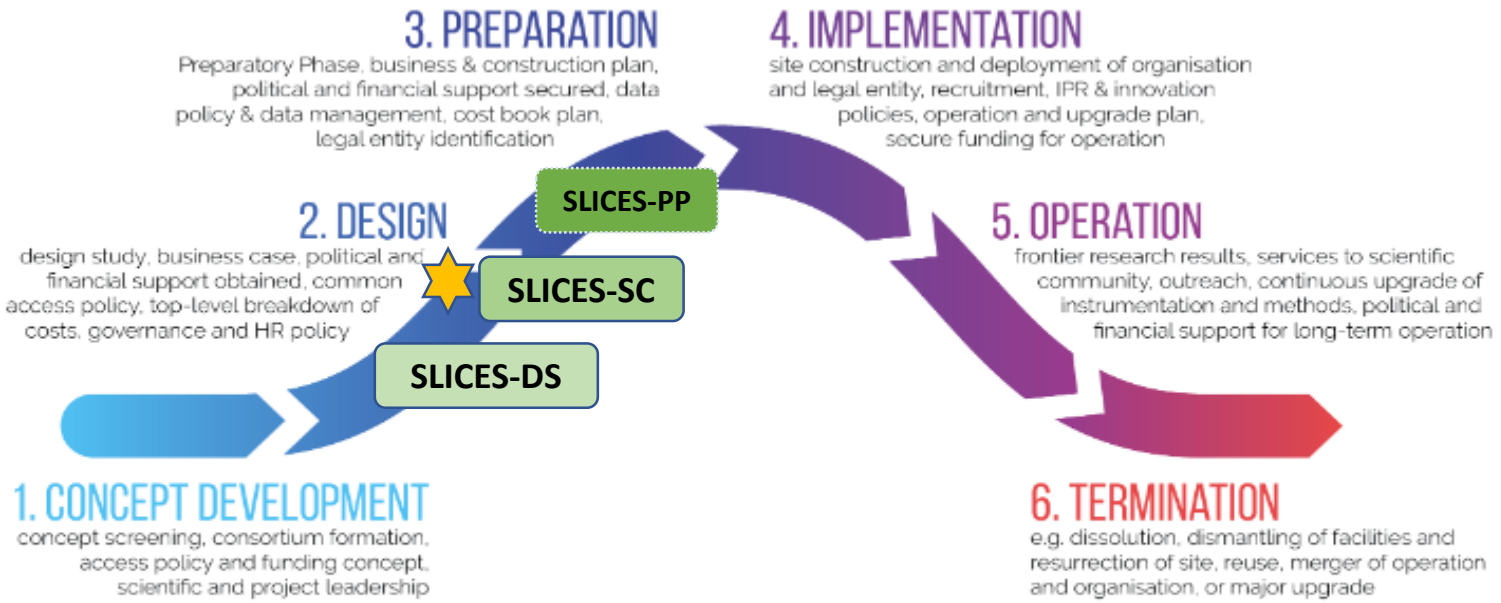
https://slices-ri.eu/

In cooperation with GIANT and national NRENs
Strong integration into the EOSC ecosystem

SLICES – Members and support



Lifecycle of an ESFRI Research Infrastructure



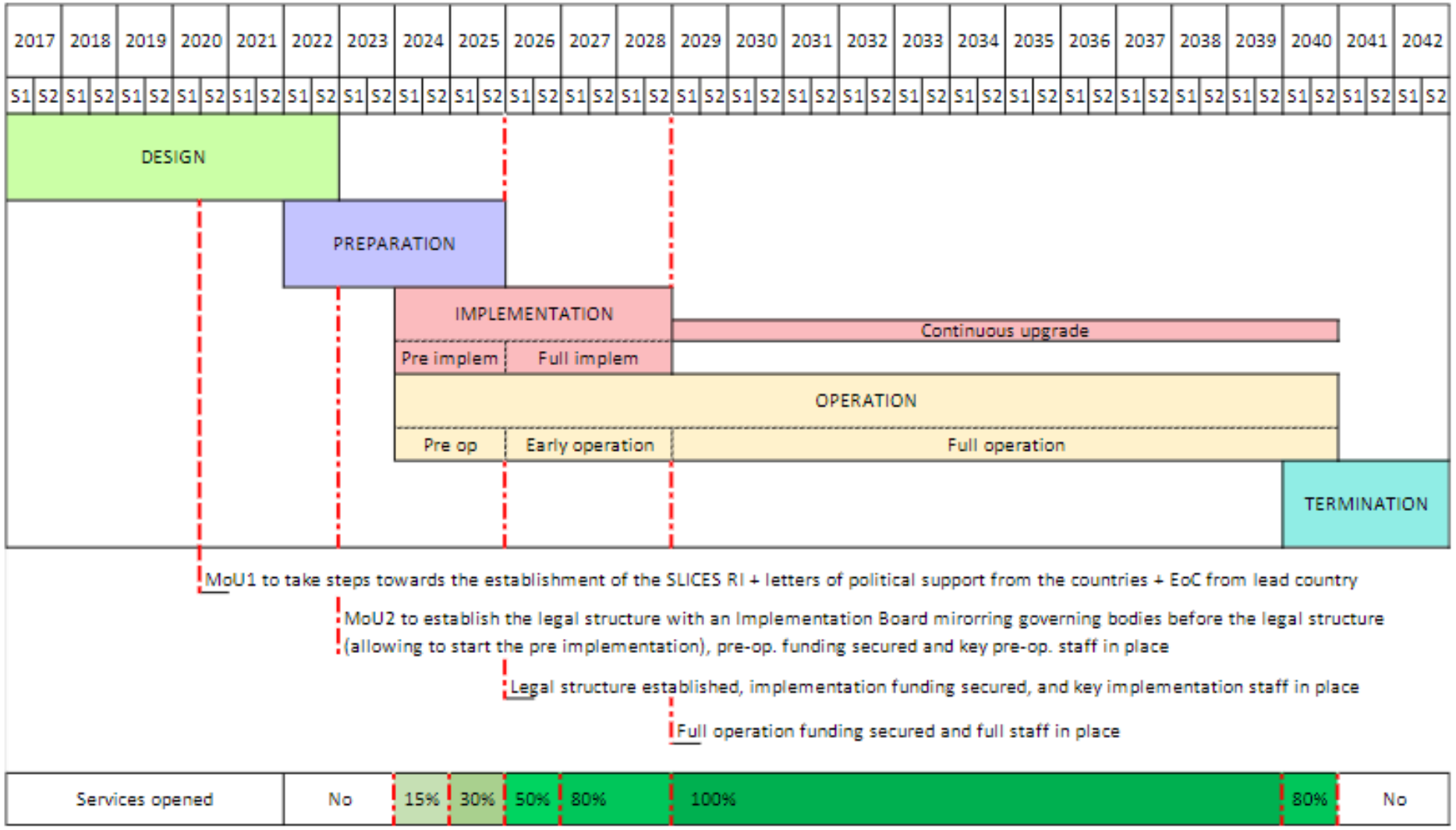
Supported by 2 projects started in 2020

- Slices Design Study (SLICES-DS)
- Slices Starting Community (SLICES-SC)

Project to be summited in January 2022

- H2020 Slices Preparatory Phase

SLICES – Timeline



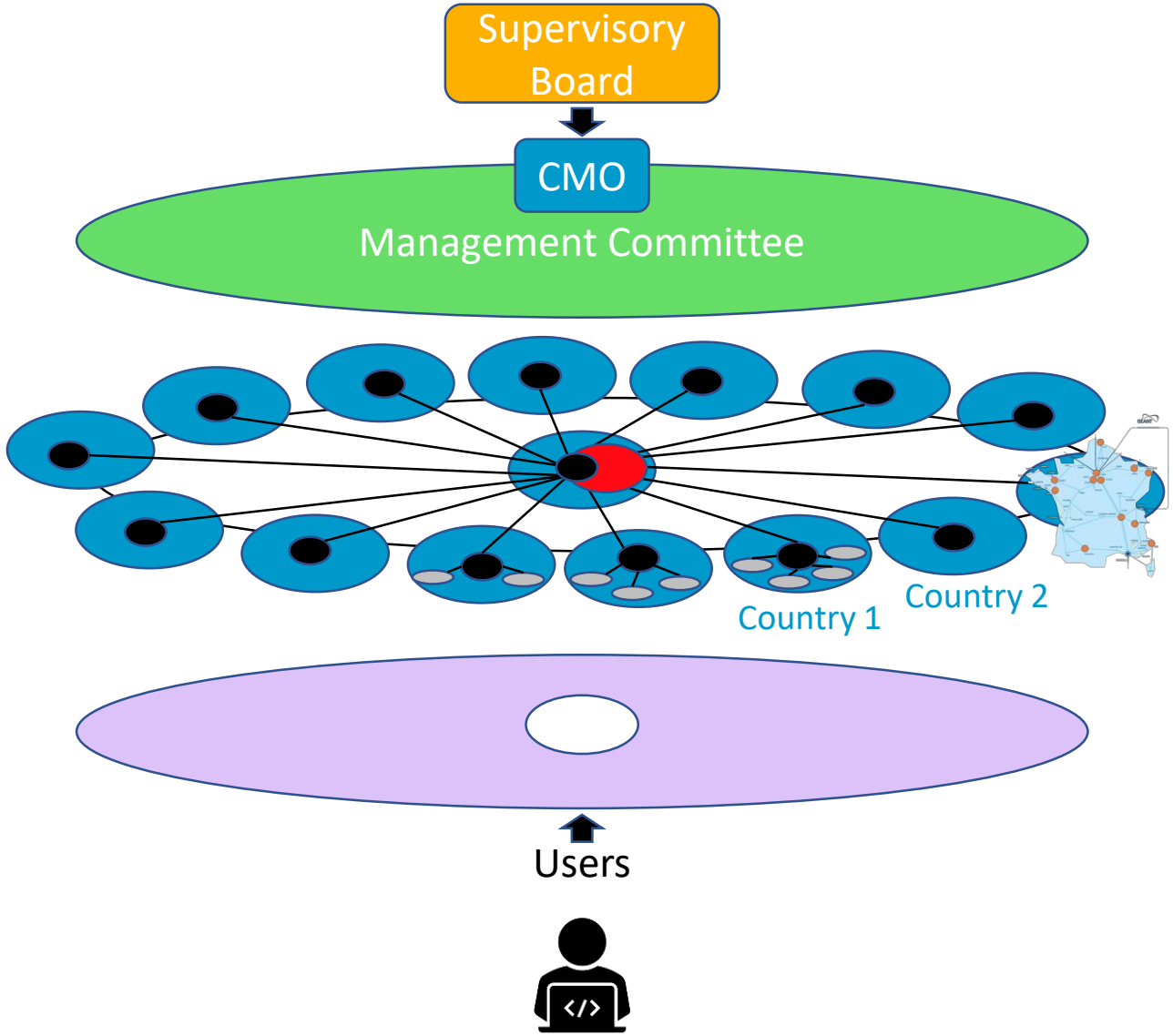
Q#4

SLICES is a distributed Research Infrastructure

Centralised governance

Distributed Infrastructure

Single entry point, single access policy



Joint investment strategy

Decisions on new nodes

Decisions on core functions and data centre

Optimize the distribution of resources according to needs and competences

SILCES: The French Node

Based on FIT and Grid'5000

SILECS: based upon two infrastructures

FIT

- Proving Internet players access to a variety of fixed and mobile technologies and services, thus accelerating the design of advanced technologies for the Future Internet

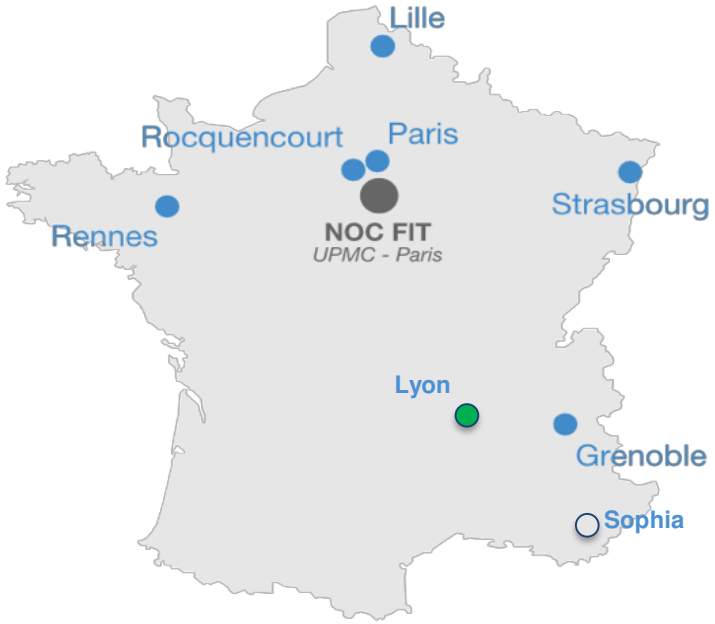
Grid’5000

- A scientific instrument for experimental research on large future infrastructures: Clouds, datacenters, HPC exascale, Big Data infrastructures, networks, etc.

SILECS/FIT

Providing Internet players access to a variety of fixed and mobile technologies and services, thus accelerating the design of advanced technologies for the Future Internet

- **FIT-R2Lab:** WiFi mesh testbed (DIANA)
- **FIT-CorteXlab:** Cognitive Radio Testbed 40 Software Defined Radio Nodes (SOCRATE)



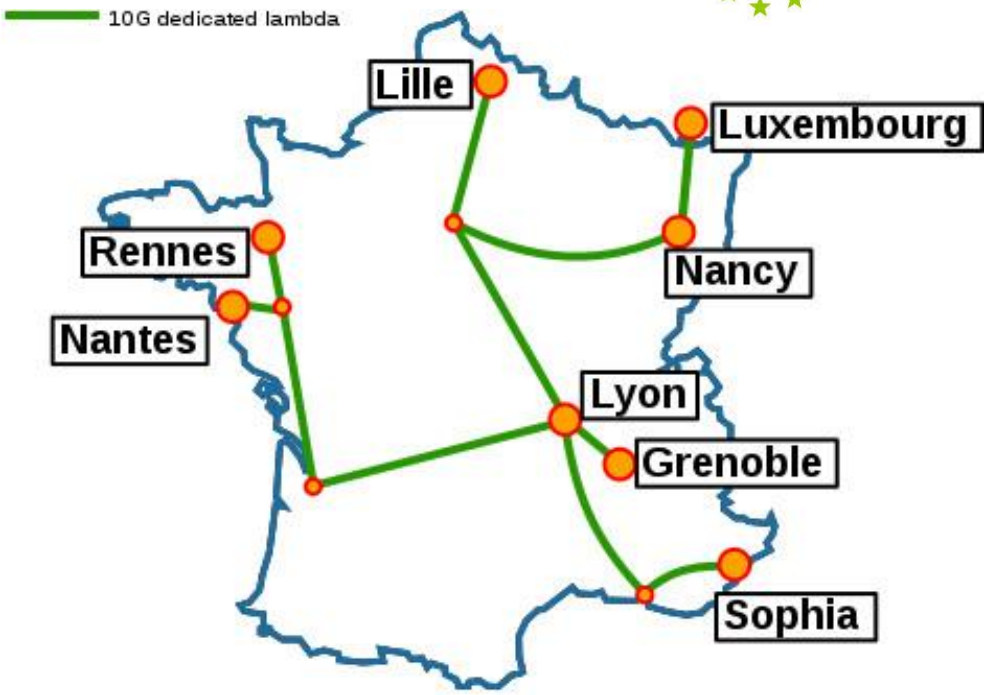
- **FIT-IoT-LAB**
  - 2700 wireless sensor nodes spread across six different sites in France
  - Nodes are either fixed or mobile and can be allocated in various topologies throughout all sites

https://www.iot-lab.info/hardware/

https://fit-equipex.fr/

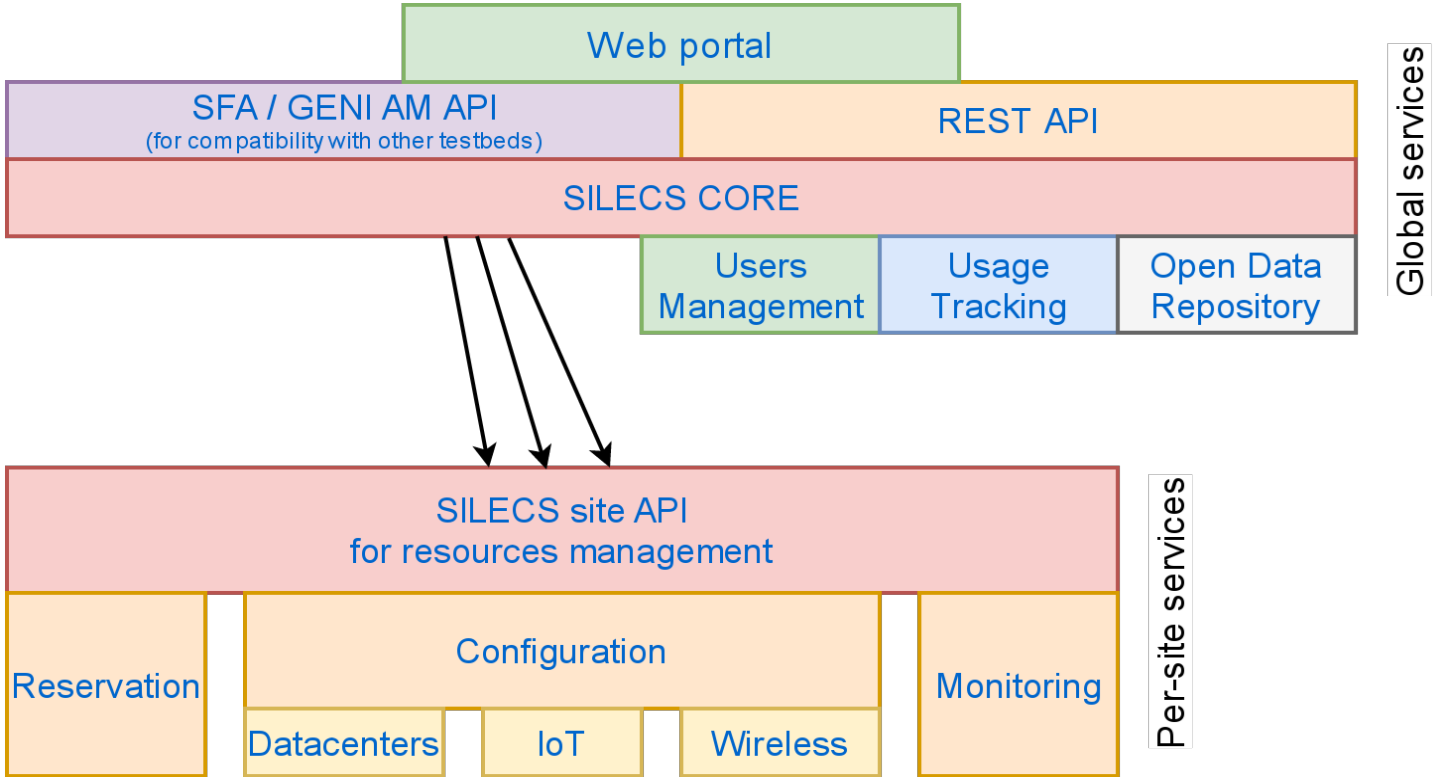
SILECS/GRID'5000

- **Testbed for research on distributed systems**
  - Born in 2003 from the observation that we need a better and larger testbed
  - HPC, Grids, P2P, and now Cloud computing, and BigData systems
  - A complete access to the nodes' hardware in an exclusive mode (from one node to the whole infrastructure)
  - Dedicated network (RENATER)
  - Reconfigurable: nodes with Kadeploy and network with KaVLAN
- **Current status**
  - 8 sites, 38 clusters, 763 nodes, 15768 CPU cores, 314 GPU
    - Memory: ~100 TiB RAM + 6.0 TiB PMEM
    - Storage: 1.42 PB (1515 SSDs and 953 HDDs on nodes)
    - 617.0 TFLOPS (excluding GPUs)
  - Diverse technologies/resources (Intel, AMD, Myrinet, Infiniband, two GPU clusters, energy probes)
- **Some Experiments examples**
  - In Situ analytics
  - Big Data Management
  - HPC Programming approaches
  - Network modeling and simulation
  - Energy consumption evaluation
  - Batch scheduler optimization
  - Large virtual machines deployments



https://www.grid5000.fr/

Services & Software Stack



Built from already functional solutions

Conclusions

- **SLICES**: ESFRI Research infrastructure for experimental computer science and future services in Europe
- **SILECS**: Research infrastructure in France based on two existing instruments (FIT and Grid'5000)
- **Challenges**
  - Design a software stack that will allow experiments mixing both kinds of resources while keeping reproducibility level high
  - Keep the existing infrastructures up while designing and deploying the new one
- **Keep the aim of previous platforms** (their core scientific issues addressed)
  - Scalability issues, energy management, …
  - IoT, wireless networks, future Internet
  - HPC, big data, clouds, virtualization, deep learning, ...
- **Address new challenges**
  - IoT and Clouds
  - New generation Cloud platforms and software stacks (Edge, FOG)
  - Data streaming applications
  - Big data management and analysis from sensors to the (distributed) cloud
  - Mobility
  - Next generation wireless
  - …
- **Next steps**
  - Continue preparing Slices (H2020 Slices-PP) and Silecs (PEPR)