

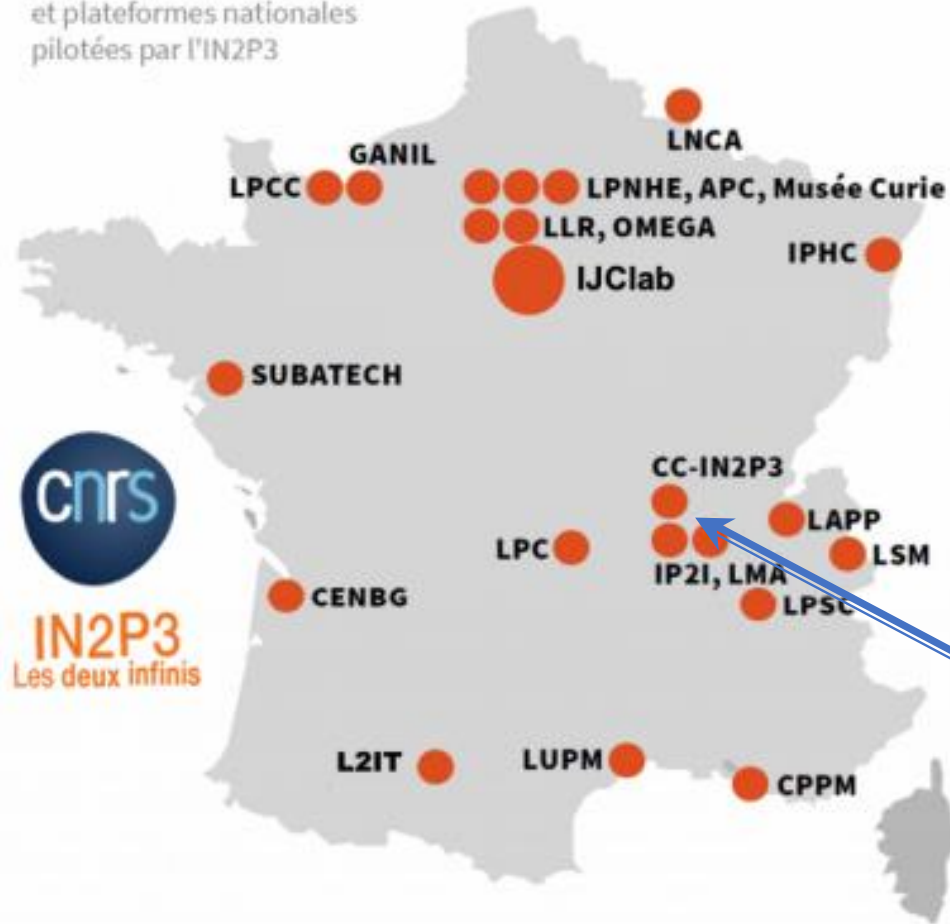
Centre de Calcul  
de l'Institut National de Physique Nucléaire  
et de Physique des Particules

# From SRB to iRODS: 20 years of data management at the petabyte scale

Jean-Yves Nief  
Yonny Cardenas  
Yvan Calas

# What is CC-IN2P3 ?

Carte de France des unités  
et plateformes nationales  
pilotees par l'IN2P3



## IN2P3:

- one of the 10 CNRS institutes.
- 19 labs dedicated to research in high energy physics, nuclear physics, astroparticles.

## CC-IN2P3:

- Computing resources provider for projects supported by IN2P3 (national and international collaborations).
- Resources opened both to french and foreign scientists.
- 2 computing rooms (2 x 850 m<sup>2</sup> or 9150 ft<sup>2</sup>):
  - 2,000 servers.
  - 800 virtual servers.





# Who is using CC-IN2P3 ?



# Why did we need a data management middleware ?

CC-IN2P3 participates to international collaborations:

- Other computing centers.
- Users spread worldwide.

Needs:

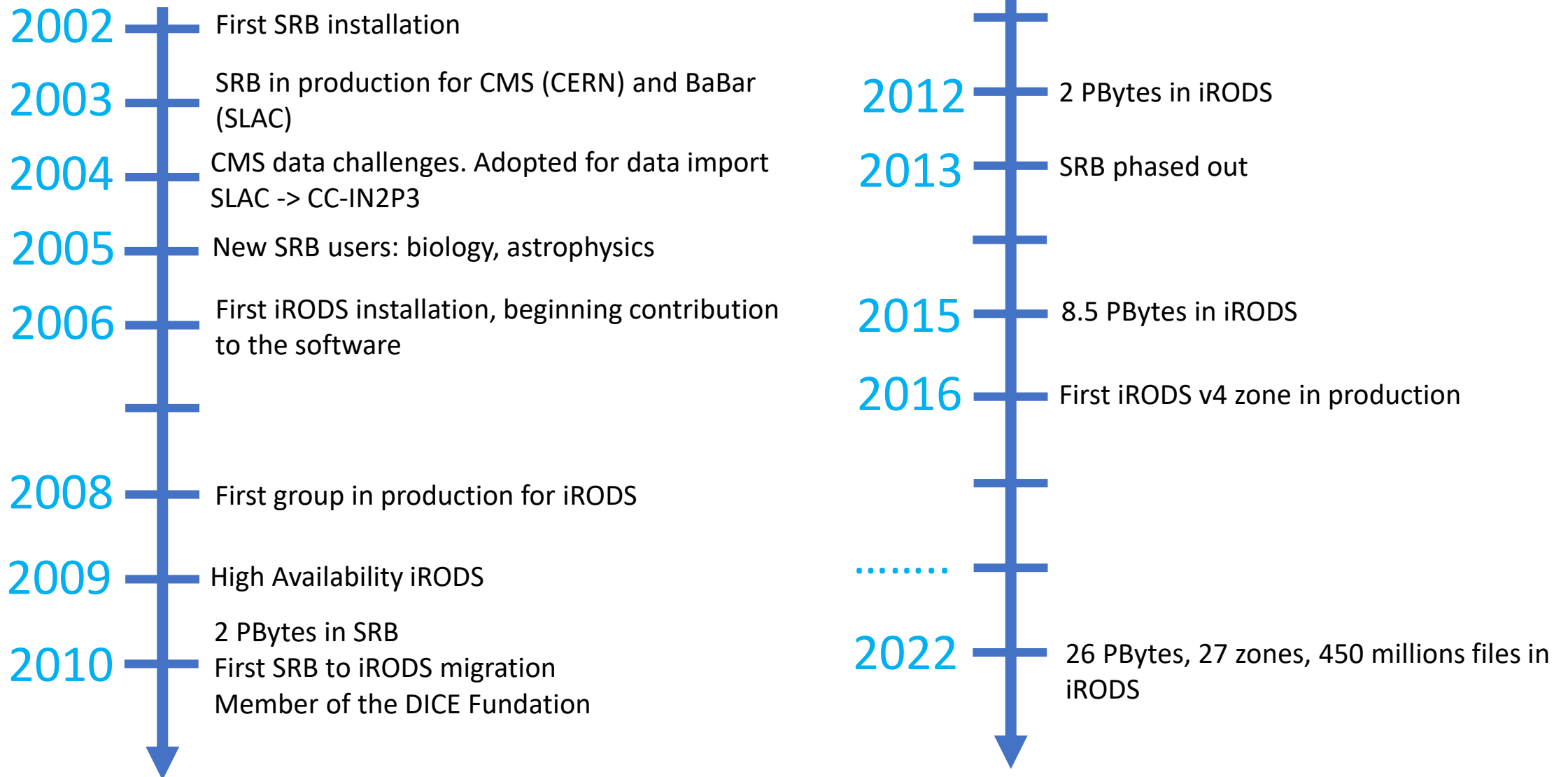
- Transfer, share and manage common sets of data.
- Identify and give different roles to the users.
- Heterogeneous technologies used:
  - OS, hardware, storage technologies, protocols.

Where do we come from ?

- In the 90s: data transfer using physical transport (van, plane etc...).
- Late 90s: start to build home made system to manage and transfer data between 2 sites.

Multiply home made systems for different projects / different sites ? No way!!!

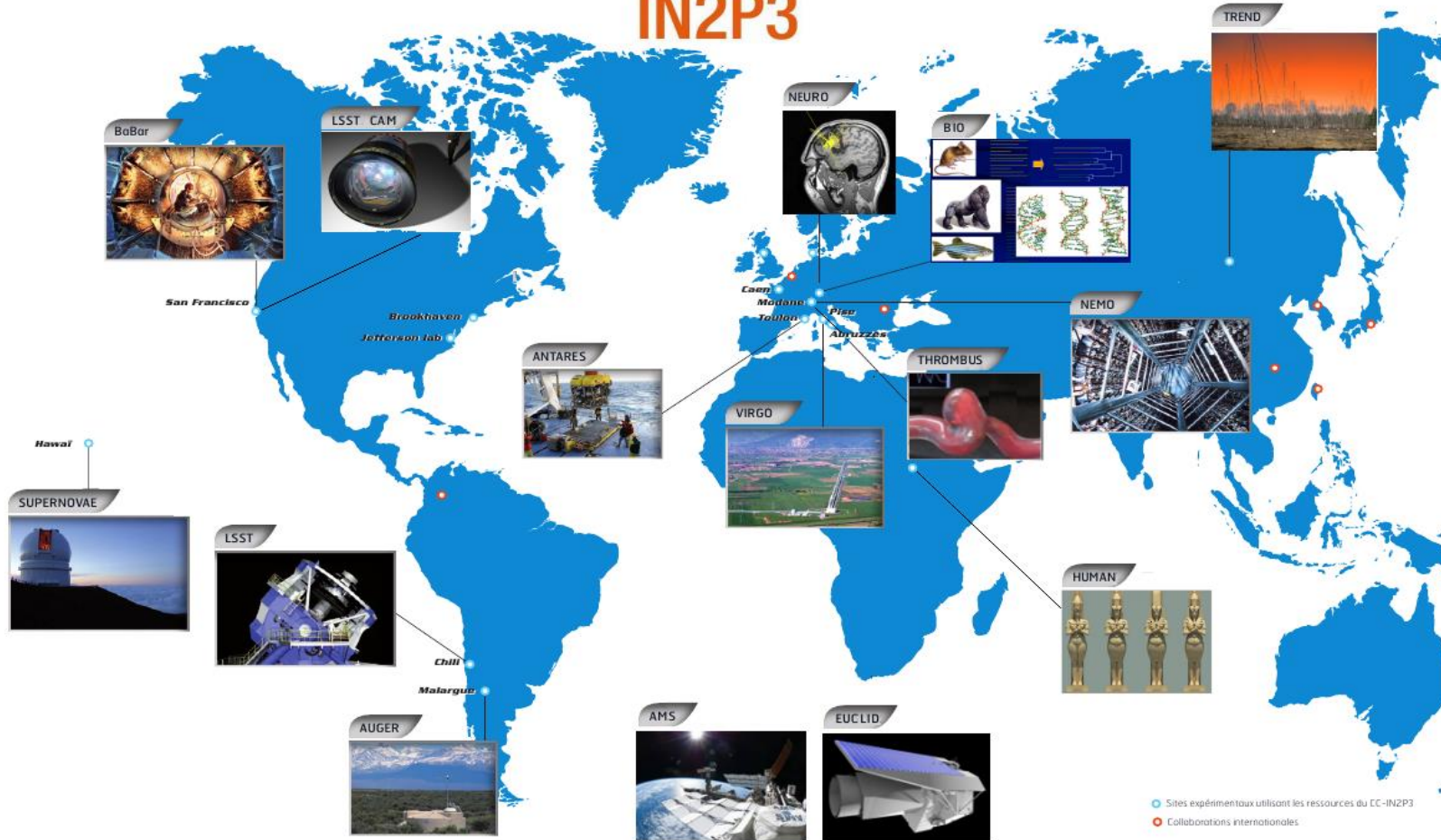
# Some historical background: from the Storage Resource Broker to iRODS





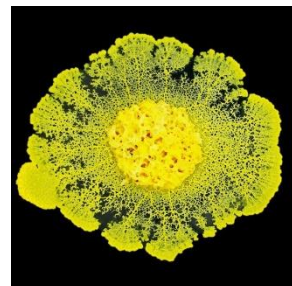
# Who are (were) the iRODS (SRB) users ?

## iRODS IN2P3



### Scientific fields:

- Particle physics.
- Nuclear physics.
- Astroparticle and astrophysics.
- Biology, biomedical studies.
- Arts and Humanities.



# Evolution of users' needs (part I)

## Needs:

- Remotely share, manage data from wide-spread collaborations.
- Upload/download files on local sites or CC-IN2P3 for data analysis through other technologies (NFS like, mass storage system, xrootd etc...).
- Manage workflows for distributed computing resources.

## Users profile:

- Developers:
  - provide applications hiding the iRODS usage.
- End users:
  - Direct iRODS usage for skilled users (through scripts and tools for various workflows)
  - Indirectly using iRODS through Web or other applications.

## How to interact with iRODS ?:

- iCommands (mainly), APIs (Java, C).
- Web applications, visualization apps.
- webDav becoming more and more popular.

## Main pitfall:

- Lack of control on amount of parallel connections (can come from anywhere).

# Evolution of users' needs (part II)

Some users came from far away:

- Using external drives to handle data...

Metadata:

- AVUs metadata been there for almost 20 years:
  - Could be seen as a gizmo by some people.
  - Data description within the collections name.
- Mind set has changed a lot in the last couple of years:
  - Users naturally asking for this kind of feature.
- But key/value pairs do not necessarily make a good match (for comple metadata).



## Storage virtualization:

- Ease and help a lot handling wide spread heterogeneous storage systems with a single interface.
- Danger: might make life too easy (and not fully aware of the limits on the systems underneath).

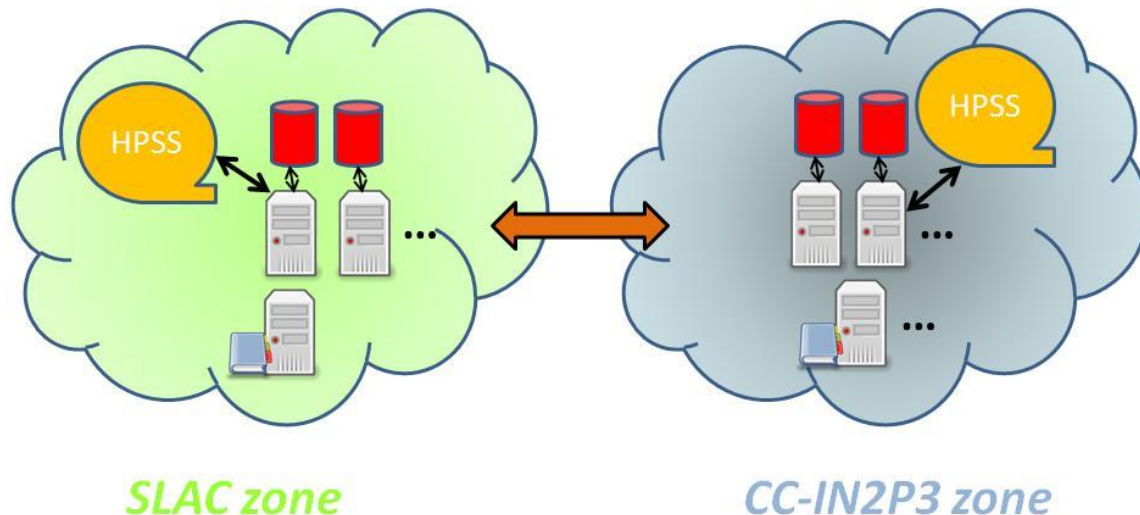
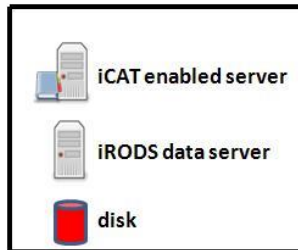
## Policy virtualization:

- An evidence for a storage system admin or a data manager.
- Allow excellent adaptation to a lot of pre existing scenarios.
- Not obvious for the end users:
  - Might not understand what it means.
  - Might not see the interest for it.

# Federation capability

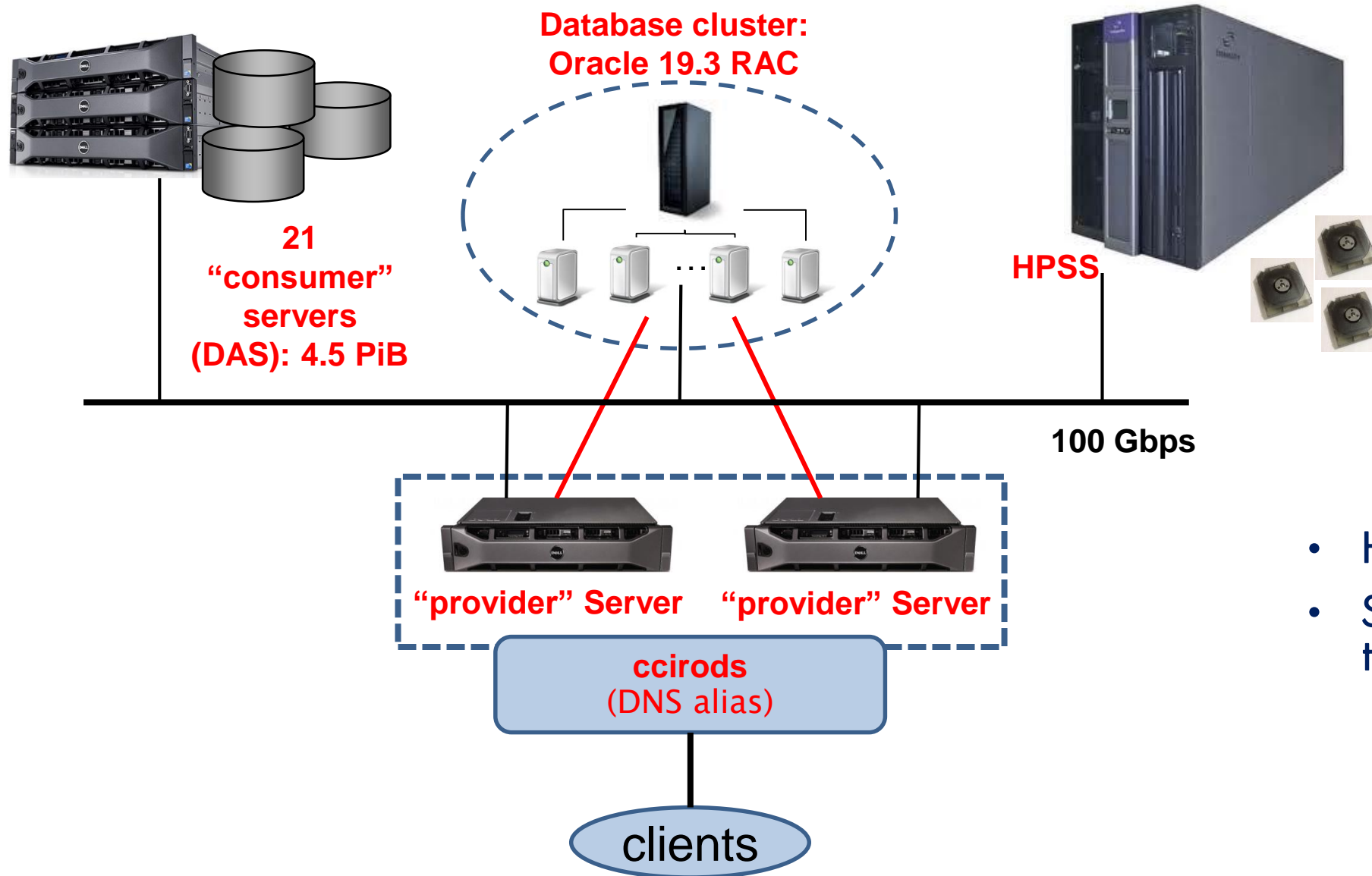
A couple of zone federations (2 zones):

- On the campus.
- National.
- Transatlantic (example below).



- **archival** in Lyon of the entire BaBar data set (total of **2 PBs**).
- **automatic transfer** from tape to tape: 3 TBs/day (no limitation).
- **automatic recovery** of faulty transfers.
- ability for a SLAC admin to recover files directly from the CC-IN2P3 zone if data lost at SLAC.

# CC-IN2P3 architecture



And also:

- 5 webDav portals.
- 1 gridFTP portal.

- High availability
- Shared infrastructure for the 27 iRODS instances.



- 27 zones.
- 110 groups.
- 1236 users:
  - Maximum of 1M connections per day.
  - Maximum of 8M connections per month.
- 450 millions of files.
- 26 PB:
  - Disks: 1.8 PiB.
  - Tape: 23.3 PiB.
  - Up to +100 TiB daily growing rate.

## OS:

- Started on Solaris 8 / Sparc to CentOS7.
- Starting to switch to RockyLinux next year.
- Migrations not a big deal thanks to SRB/iRODS portability.

## Hardware:

- Physical migration has always been run on the fly.
- PBs have been migrated: not a big deal, but small files migration is always an issue (true also outside the SRB/iRODS world).

## Technology:

- SRB -> iRODS: disruptive, made step by step in 3 years.
- iRODS v3 -> v4: still ongoing.
- Not that easy: need to adapt to our HA system and multi installations per server.

A lot of talk on FAIR data every where.....

Try to stay in the real world and make real things (work in progress):

1. Data Management Plan:

- Make our DMP machine actionnable: check if the data policy described in the DMP in agreement with what is really happening.

2. Long term data archival: OAIS based (using RODA) + iRODS.

Access mode:

- REST APIs, web access.
- Authentication: OpenID.
- Mechanism to throttle incoming wide opened connections



Thanks to my colleagues at CC-IN2P3:

- Yonny Cardenas.
- Yvan Calas.

## Big thanks to the iRODS teams through the age!