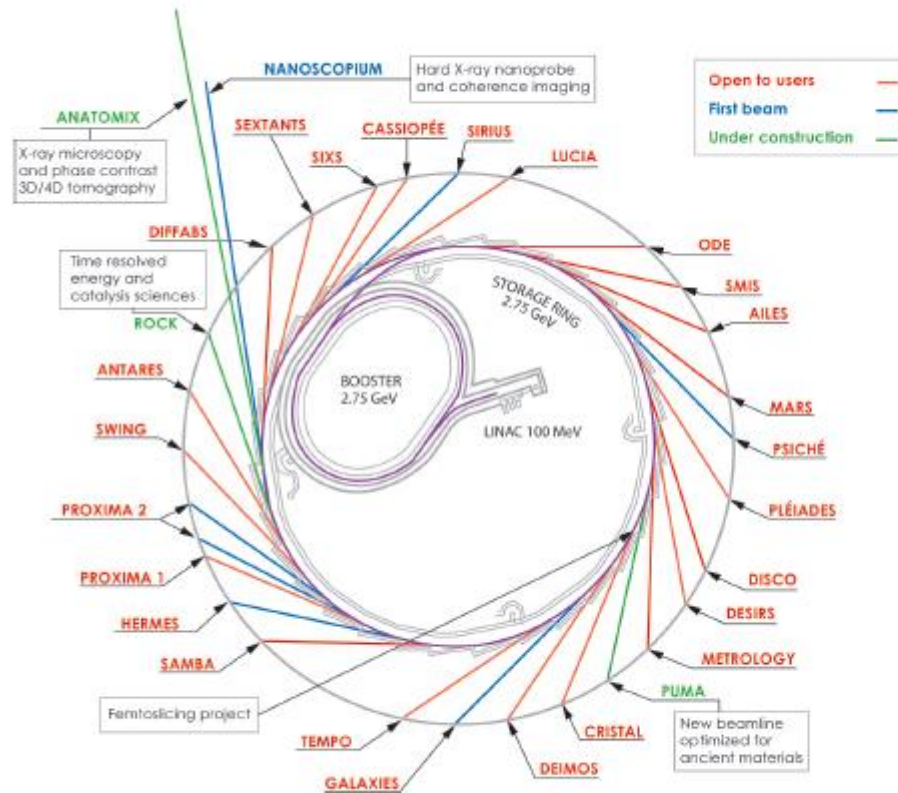


# IT infrastructure evolutions

*Emmanuel GIRARD, Philippe PIERROT*



# 26 beamlines open to users, 29 beamlines in 2017



- 26 beamlines open to users.
- 26 beamlines in the 2 phases of the SOLEIL program; all have seen the beam
- 3 additional beamlines (project funding).
- 24 on insertion devices; 5 on bending magnets.



# Scientific Strategy

**Spectroscopies**  
Time resolved (fs, ms)  
High resolution  
In situ

**Diffraction/Scattering**  
Automation, kinetics,  
coherence

**3D imaging**  
High resolution  
Phase contrast  
Multi-scale – multi-mode

**Chemistry – Physical Chemistry**  
Properties and reactivity of model  
and complex systems

**Biology - Health**  
Integrative approach from  
molecule to tissues

**Complex materials**  
Ancient materials, nanomaterials,  
extreme conditions

**Physics**  
Fundamental properties of matter from ideal to  
complex systems, nm to macroscopic

**Partnerships**  
Medium and long-term projects, instrumentation and support, R&D

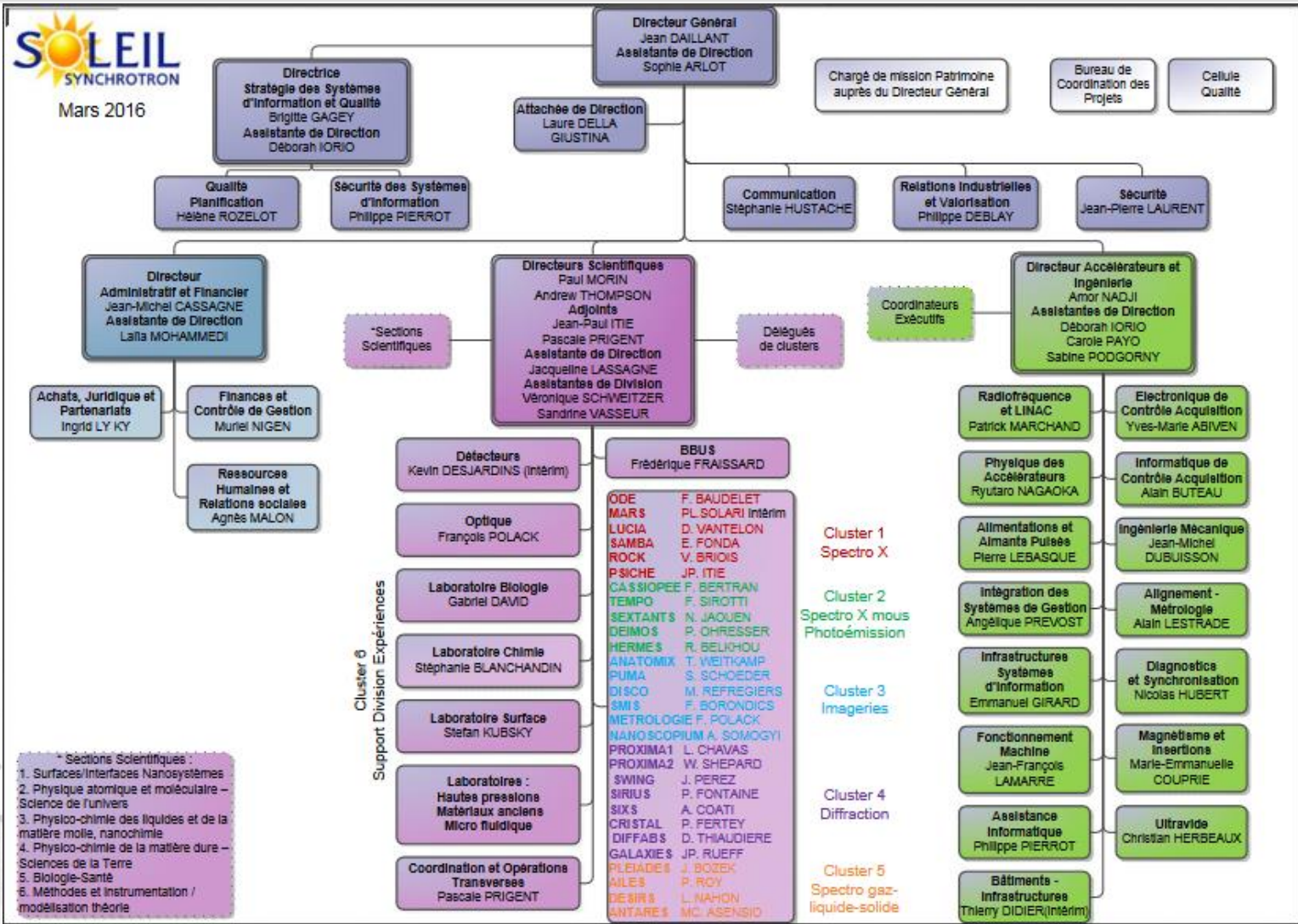
**INRA-SOLEIL**    **IPANEMA**    **MICASOL**    **COSMETOMICS**

# Some Figures

- Private company 72% CNRS, 28%CEA.
- Total construction cost (including salaries) 440 M€.
- Budget (2014): 65 M€ expenses, 51 M€ operating costs (33 M€ salaries and linked costs), 54 M€ from CEA and CNRS (Flat or slightly decreasing budget, increasing importance of targeted grants).
- Personnel:345 permanent contracts, 13 fixed term, 35 PhD students, 25 post-docs end 2014



# Organization



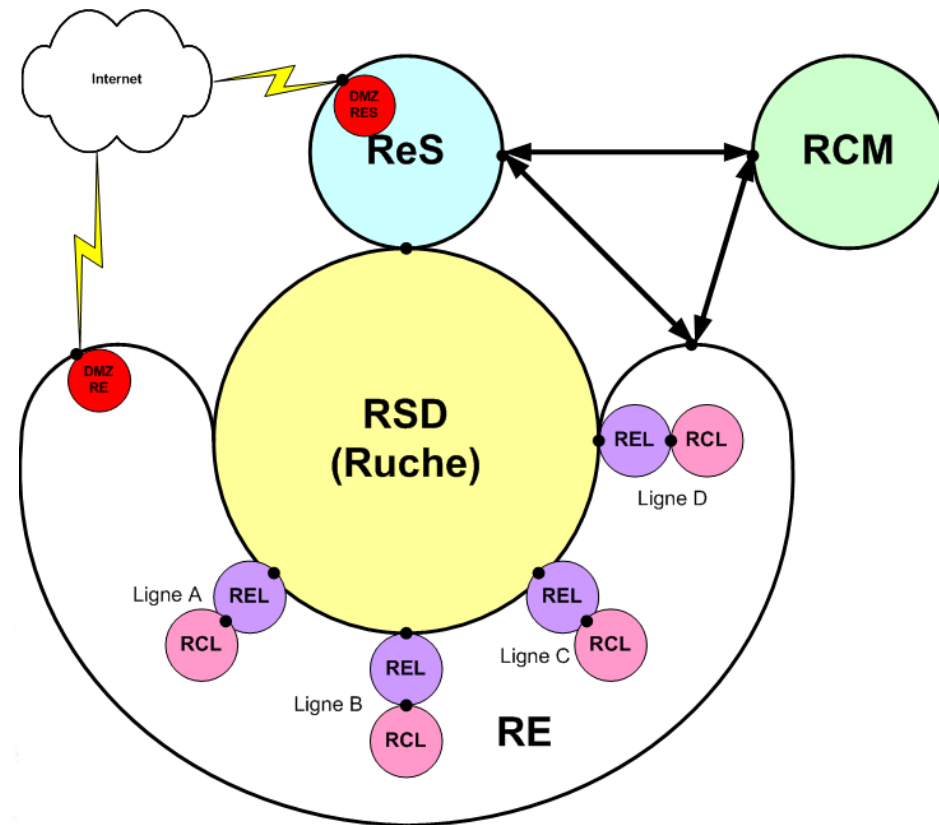
# Guidelines for infrastructure evolutions

- Evolutions are guided by:
  - Ensuring IT infrastructure reliability:
    - Replace items of equipment or software reaching end of service life, optimize operations, improve supervision and procedures...
  - Addressing emerging demands, anticipating them when possible:
    - 2D detectors, beamlines automation, remote access, image processing...
- Our ultimate goal is to contribute to SOLEIL excellence as a scientific instrument

- Storage facilities
- High Performance Computing
- Backup system
- Network
- Servers
- Security

# Quick reminder: Current network architecture at SOLEIL

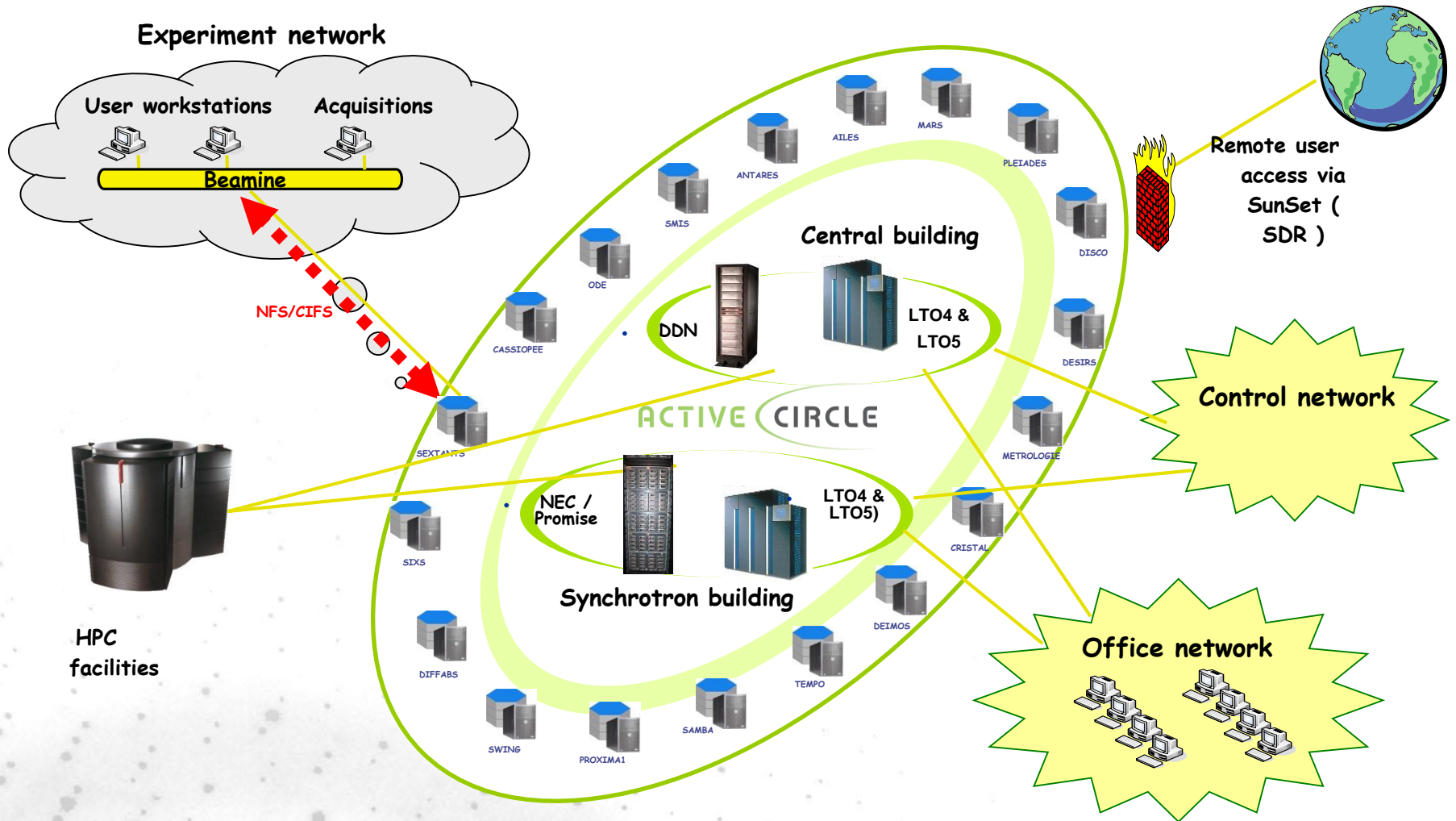
- Two Computer rooms
- 3 interconnected networks (RES, RE, RCM)
- Clustered Core router in each network
- Each beamline has two subnetworks of RE: REL and RCL
- 1 additional network (RSD) dedicated to experiments data storage



# SCIENTIFIC COMPUTING INFRASTRUCTURE

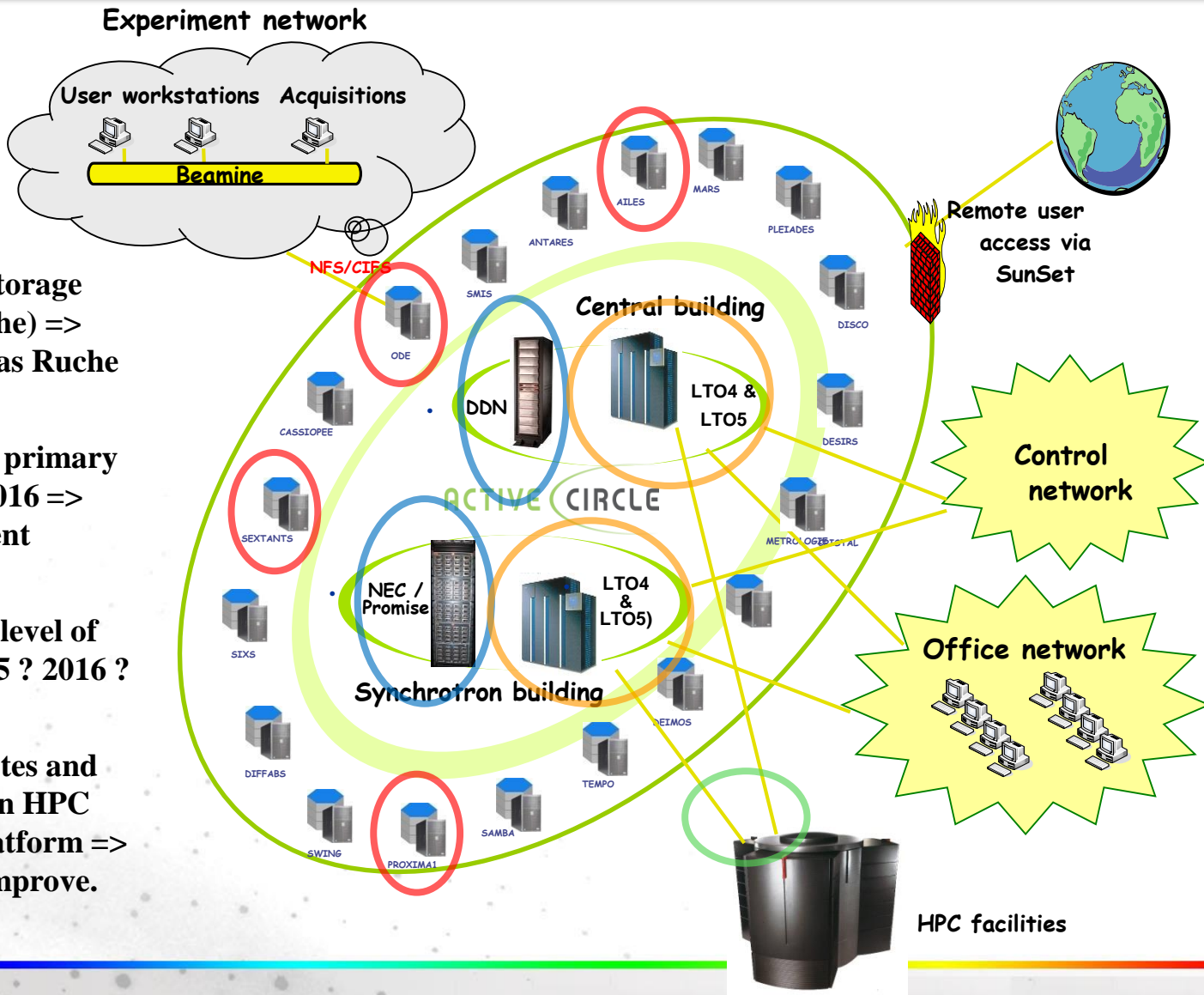


# Storage facilities : experiment data architecture



# Storage facilities : experiment data architecture evolutions

- Replacement of oldest storage nodes (10 GE, large cache) => ongoing project known as Ruche V2
- Planned replacement of primary storage nodes in 2015/2016 => ongoing related to current strategy
- Evolution of the second level of HSM to the cloud ? 2015 ? 2016 ? => still on going studies
- Improvement of data rates and high availability between HPC facilities and storage platform => use of flash storage to improve.



→ subject to budget !

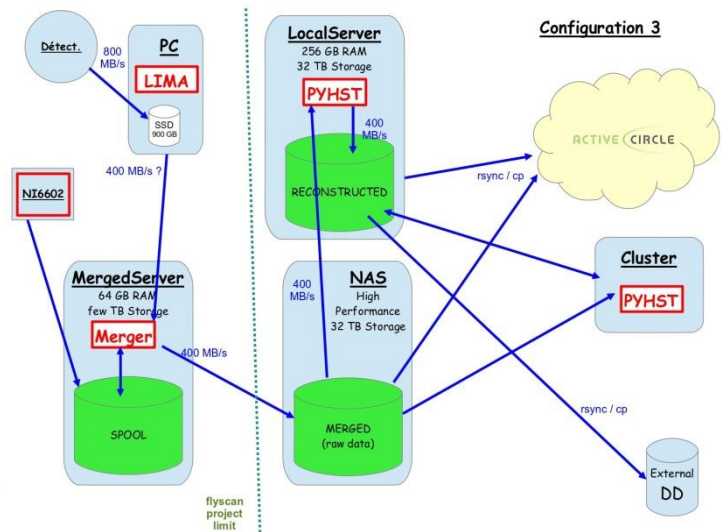
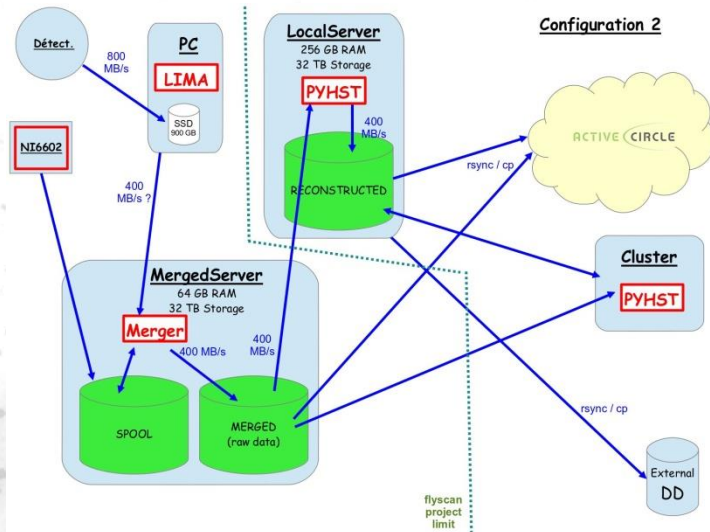
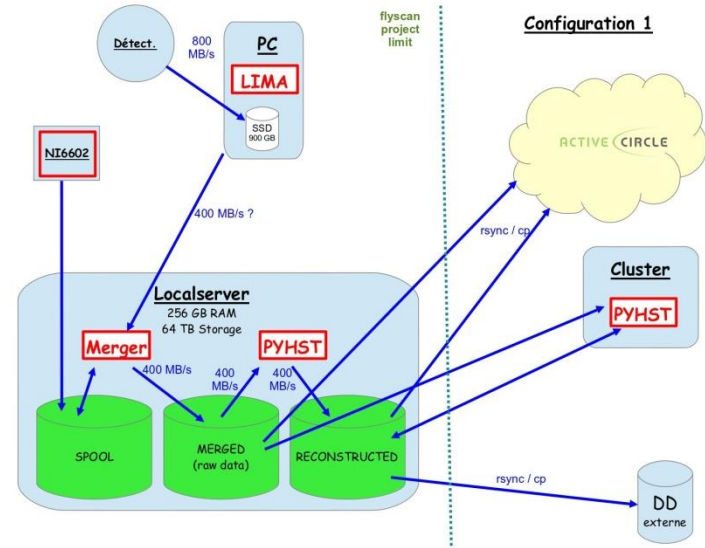
# Fast local storage for beamlines

- Working with Experiments Division and « flyscan » project to define a reference architecture to be implemented on the beamlines – different types of architectures have been tested => Ruche V2 was born.
- Tests done on PSICHE and PROXIMA 2 end of October
- From 500 MB/s (Write) to 900 MB/s (Read) on 10 GE network via NFS
- The new servers have SSD drives to improve data acquisition and the server is also an Active Circle Cell thus allowing:
  - Computing to be done on the beamline
  - data to be archived on the Ruche.



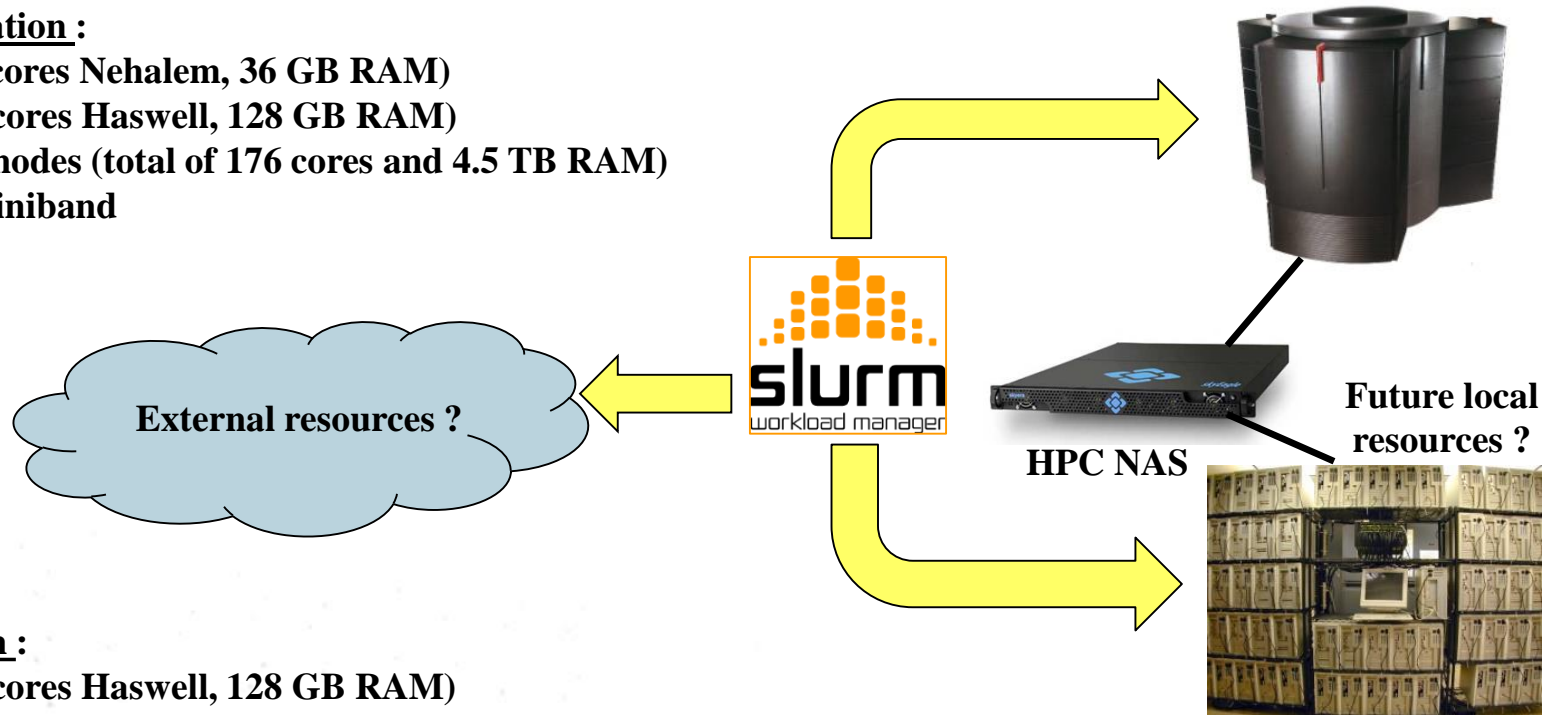
# Fast local storage for beamlines

## ➤ Configurations tested



## Current configuration :

- 112 nodes (8 cores Nehalem, 36 GB RAM)
- 16 nodes (20 cores Haswell, 128 GB RAM)
- 2 interactive nodes (total of 176 cores and 4.5 TB RAM)
- Lustre on Infiniband



## New configuration :

- 13 nodes (24 cores Haswell, 128 GB RAM)
- 1 interactive node (total of 64 cores and 4.0 TB RAM)
- 52 GPU on K80 cards for tomography
- Lustre on Infiniband ( 128 TB )
- Total Teraflops : 88

## Evolutions

- **Adding a high-performance flash storage NAS on /work to improve data access. 24TB.**



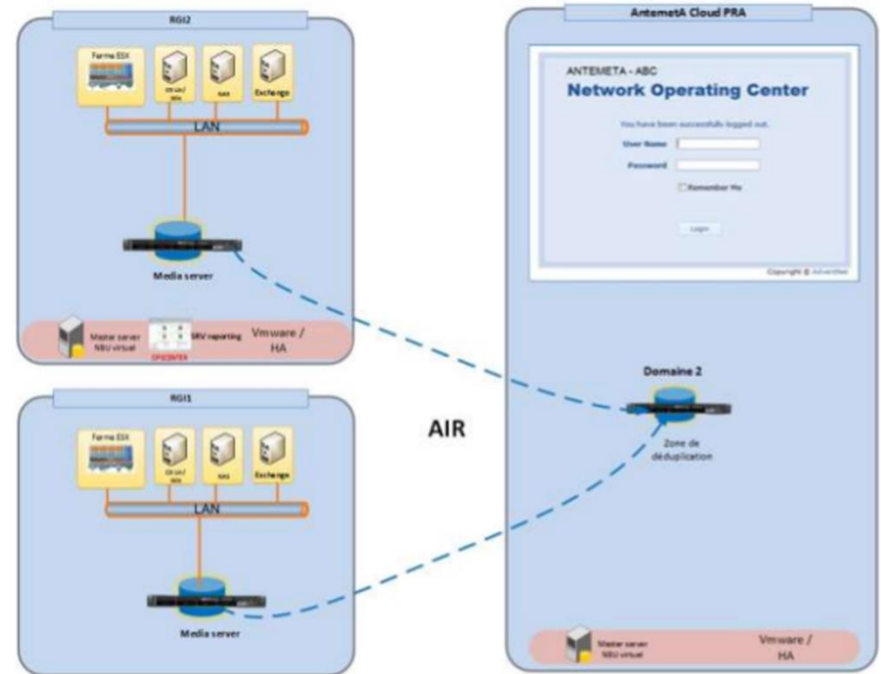
- **Strong relation with CCRT HPC facility for 400k core in 3 years beginning September '16. up to 19,6 Tflops**
  - <http://www-ccrt cea.fr/>



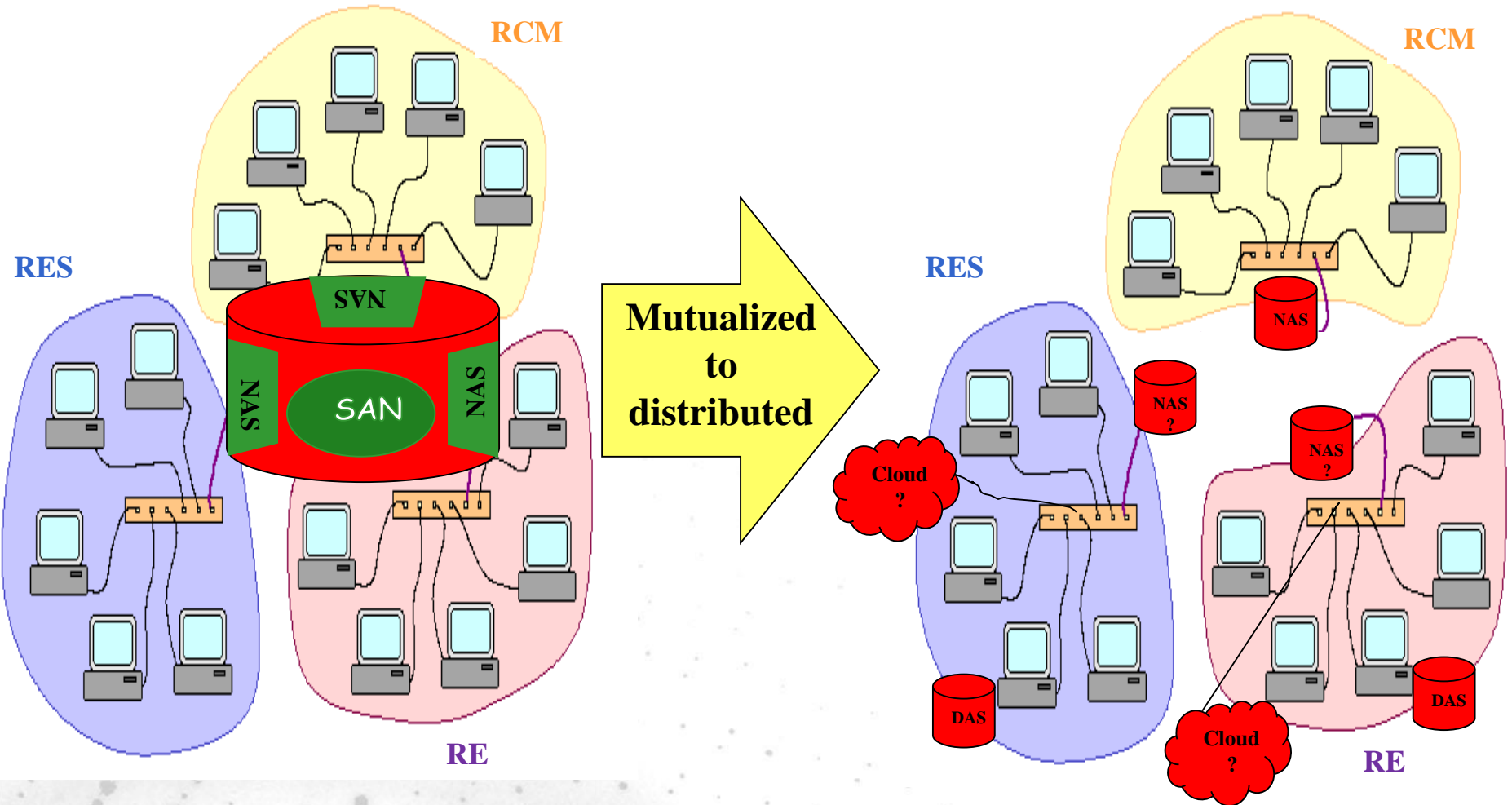
# OFFICE IT EVOLUTIONS

# Backup system architecture

- Reminder: not for experiments data (cf Active Circle configuration)
- Objectives :
  - Replacement of the current solution installed in 2008 (NetVault)
  - Reduce backup duration, especially on weekends
  - Extension of the retention to 6 months at least (a few weeks before)
  - Service availability in case of computer room shutdown
  - To be ready for 10G evolution
- AntemetA proposal :
  - 2 local appliances (17TB)
  - NetBackup software (Symantec)
  - Replication of metadata and deduplicated backups in a AntemetA's datacenter
  - Restore possible from the backup site
  - Reporting : dashboard and daily report
  - Solution under supervision 24/7, while allowing us to assure ourselves what we want
  - Pay per use, more flexible for future evolutions



# Storage facilities : office filer evolution





# Storage Facility

## ➤ Findings:

- Mutualized storage was used primarily to leverage budget but:
  - Mutualized storage constraints are more important than evaluated
  - It is almost impossible to make changes without impacting storage on all networks ( RES/RE/RCM )
- Split storage facility in different needs:
  - Machine control on RCM => done as January 2016
  - Office automation on RES/RE
  - Experience data storage is always on Active Circle.

## ➤ Improvements:

- Installation of NexentaStor on RCM: Zfs based solution – not so easy as it seems – lack of compatibility with Windows 2k/XP environment for legacy control chassis – DONE
- Tender to change SAN/NAS storage on office automation, based on needs rather than location/facility => ability to cloudify a part of our storage.

# Office automation virtual infrastructure

- Current infrastructure: 14 HP blades in 4 chassis
- 91 VMs
- 15 blades are used as physical servers.
  
- Migration to Vsphere 5.5 (done).
- licensing fees by VMWare changed and increased significantly in version 6.x.
- Planned evolutions:
  - Chassis has to be renewed – many servers have to be renewed – the only SAN storage remaining will be for VMWare datastore only.
  - Maybe It's time to redesign ( or rethink ) our VMWare architecture
    - Should we choose Hyperconvergence ( simplivity, nutanix, ... ) ?
    - Should we improve the current design ?

# MACHINE CONTROL IMPROVEMENTS



# Dedicated storage for machine control

- Previous architecture : mutualized EMC<sup>2</sup> 2 x CX4-240 with 4 NAS NS42g
- New architecture: Nexenta Stor with a cluster on RGI1 and a backup appliance on RGI2
- No transparent failover of storage from RGI1 to RGI2 but with cabling improvement we can have a ( almost ) no downtime architecture.

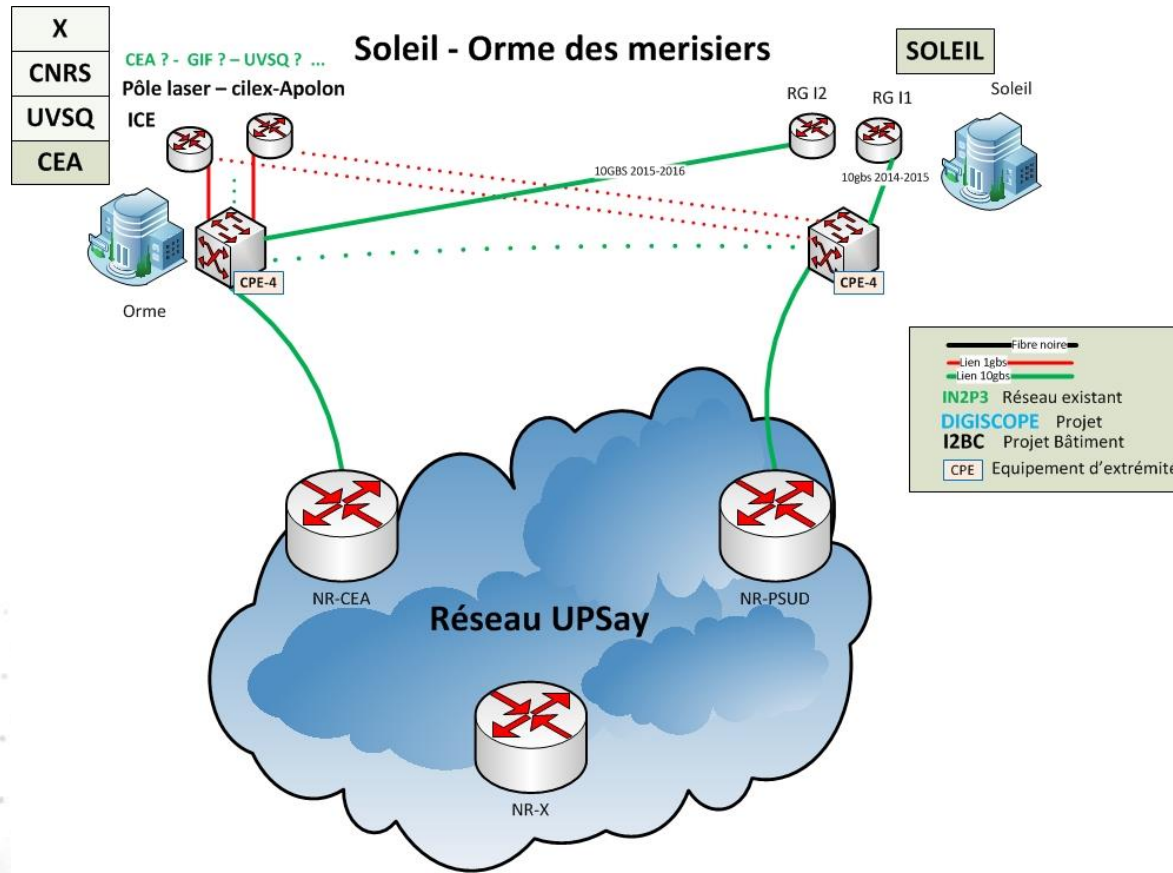
# NETWORK UPGRADE

# Network: Switch upgrade

- Beamlines (RE)
  - 10Gb/s switches (24 ports) installed on 10 beamlines
    - Allows faster data transfers between detectors, control server, local storage and central storage – DELL Switches
    - Currently tested on PSICHE beamline, in Flyscan configuration
  - Replacement of old 3750 series switches by new 3750X series on all beamlines (75% achieved, 60 switches replaced) – Cisco switches
  
- Machine (RCM)
  - Replacement of all 3750 series switches by new 3750X series (completed, 32 switches replaced) - Cisco switches
    - Hot-plug redundant power supplies (improved availability)
    - 10Gb/S uplinks
  
- Office automation (RES)
  - Replacement of all 3750 series switches by new 3750X series in all buildings (completed, 86 switches replaced)

# Network: Internet link

- Our own access to RENATER was shutdown last summer because of TCSP worksite  
Currently using one of CEA backup link at 1Gb/s
- Dual 10Gb/s link to RENATER planned for the end of this year using FCS fiber loop
- Revamping of our DMZ done:
  - High availability (RGI1 – RGI2)
  - Improved performances
  - Separate path for internal usage and beamlines needs
  - Improved monitoring (web stats, spam handling, etc.)
- Full 10Gb/s path to the Internet after Firewall replacement (Planned this year)





# Supervision -

- Our old NAGIOS3.X – Centreon 2.X monitoring platform is mostly out of date. We plan to move to ZABBIX end of 2016.
- We experiment an ELK architecture to store and analyze logs. Currently testing our RES DMZ

# IT SECURITY

# IT Security

- IT security audit points several aspects to be corrected, mostly regarding our organization.

## FIRST STEPS:

- → Appointment of an IT Security Officer:
  - Establish and maintain IT security policy.
  - Add security concerns early in projects development.
- Rewriting of our IT charter (done)
- Document management policy (done).
- Strengthening of our password policy (done)
- Migration or replacement of XP computers (some are still remaining)
- Survey about Remote Access usage for SOLEIL employees done.

## COMING SOON:

- Global IT Security Policy to be defined in the next months ( PSSI du Gouvernement Français )

**Security  
projects  
progress  
reported  
to SOLEIL  
Board**

# CONCLUSION



# Conclusion

- Summary: Infrastructure next moves
  - Fast local storage on beamlines
  - Outsourcing (Full or partial, IaaS, SaaS,...)
  - Faster path for experiment data processing and retrieval
  - Datacenter and supervision improvements.
  - Improve user experience.
  
- And still maintaining efforts to keep things running at their best.

# Our Concern ( 1 )

## ➤ Related to experiments

- What kind of storage architecture do you have for experiences – what kind of archiving infrastructure do you have ?
- How do your users get their data back home ?
- HPC on premise or on cloud. What would be the criterias for cloudification ?
- Remote accès to beamline – acces control – acces security ?
- Experiences automation

## Our Concern ( 2 )

### ➤ Related to IT :

- Network Management and separation.
- cabling infrastructure – preventing failures due to age – test procedure if any ?
- Datacenter design – supervision – UPS.
- ITSM portal for employees and users.
- Users authentication - employees authentication – Which directory for which usage ? – Id federation.
- what kind of function would you like to offsite/cloudify ?
- Data privacy – protection system – access management - cyphering

