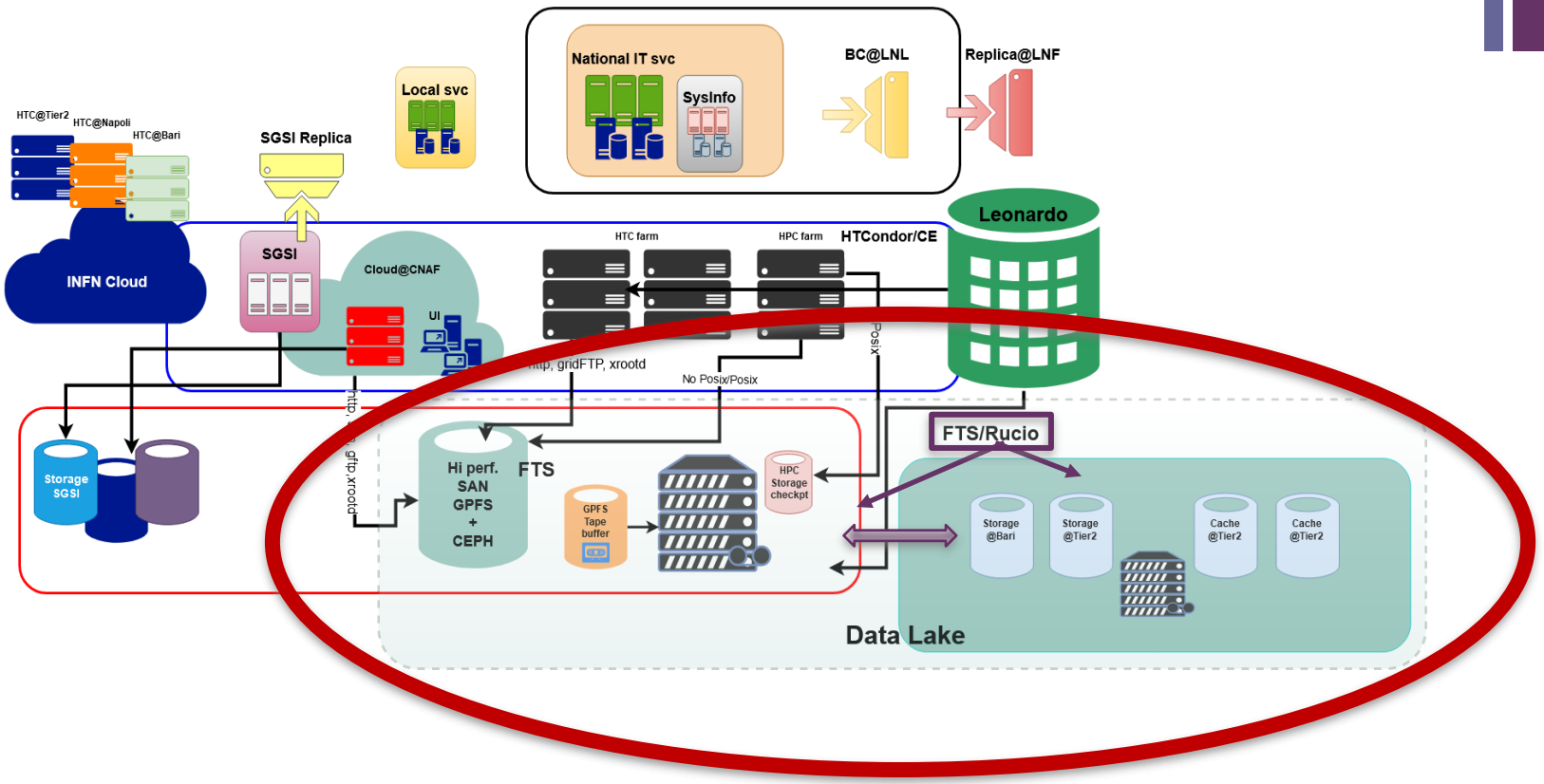
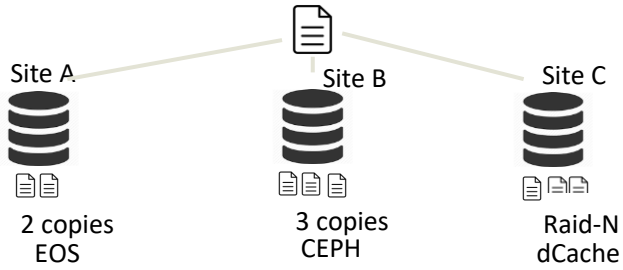
Piani per il Datalake e collaborazione con il CERN sul tema

Datalake@Tecnopolo

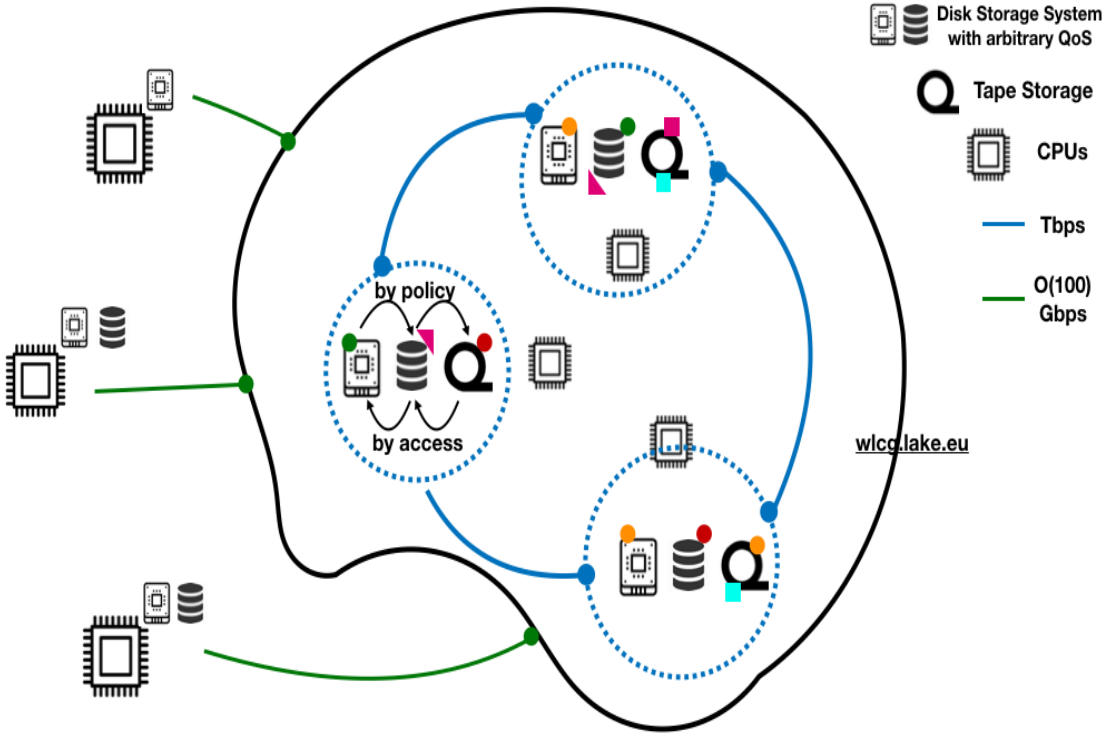


Datalake in WLCG – Storage Federation

- Reduce hardware cost: better exploiting the concept of QoS(Quality of Service)
  - Probably today we replicate more than we need
- Reduce Operational Cost: deploy fewer (larger) storage services maintaining high standards in availability and reliability
- Create large storage repositories that "look like one, but it is composed of many" ➔ **the Datalake**
- Co-location of Storage and CPU will not be guaranteed anymore
  - Need technologies for quasi-transparent data access from remote locations
    - Smart Caching

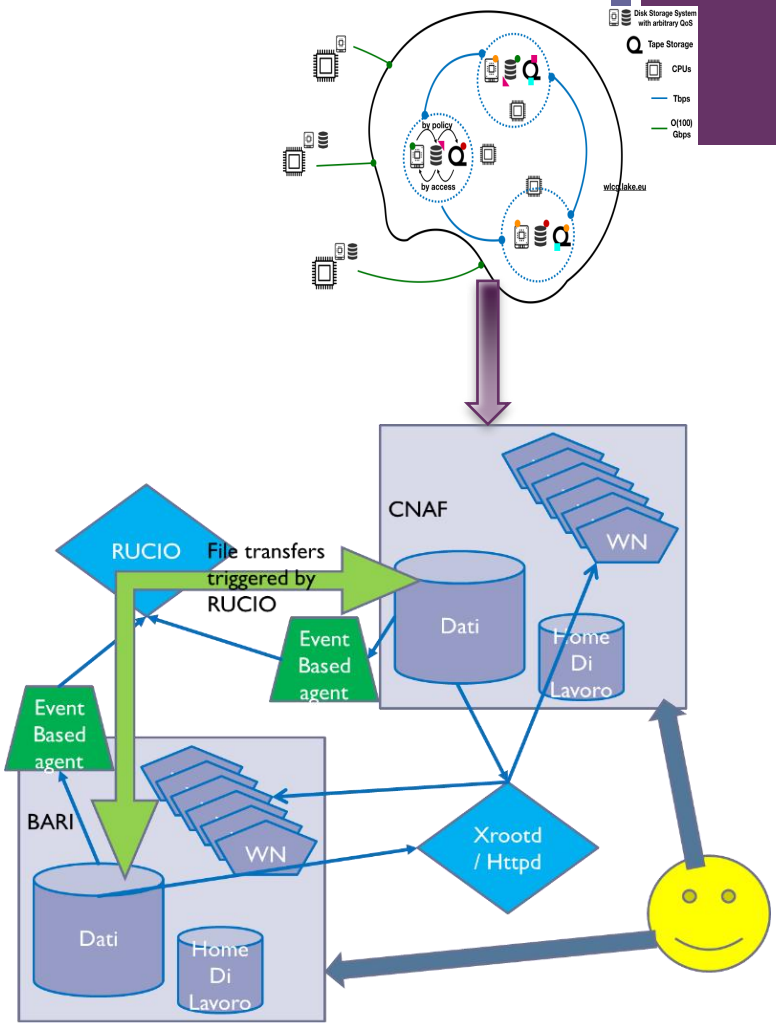


A stronger integration of sites could lead to a reduction of the number of copies



Datalake for non-LHC VOs

- Transparent creation of redundant federated storage resources
  - automatic replication in different, geographically distributed, availability zones
  - Replica performed at the infrastructure level, without the user intervention
  - To allow High Availability for VOs without a distributed computing model

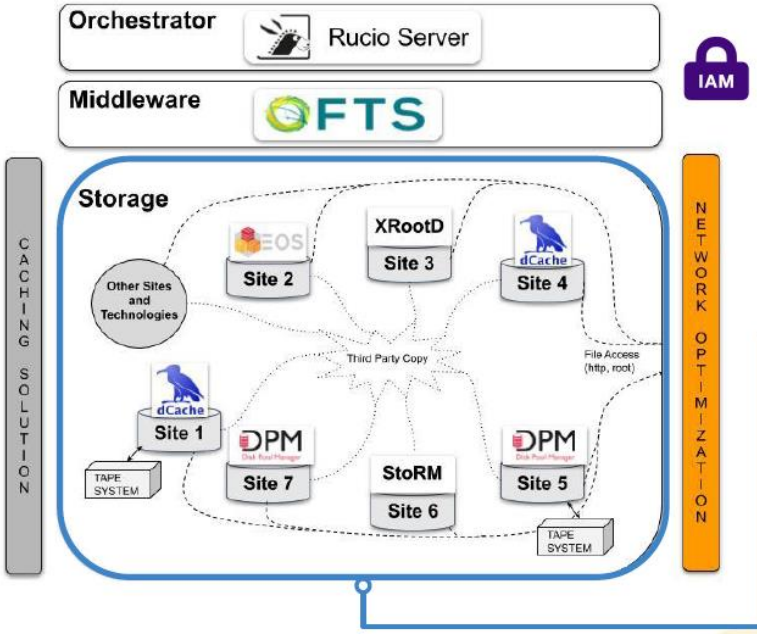


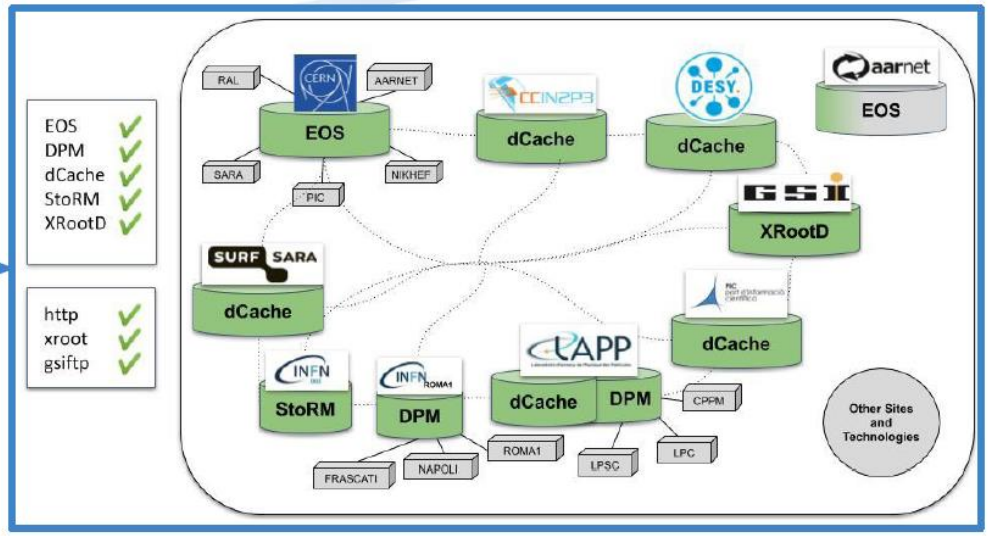
Participation to Activities on Datalakes

- **Networking solutions**
  - IDDLS
  - DCI CNAF/Tecnopolo ⇔ CERN (see Zani presentation later today)
- **Datalake Integration**
  - ESCAPE – WP2(DIOS) – creation of the Datalake for astroparticle and particle disciplines
    - https://wiki.escape2020.de/index.php/WP2_-_DIOS
    - INFN acts **as Resource Provider**….
      - T1 services: StoRM
      - NA,RM,LNF: DPM
    - …and **Technology Provider**
      - StoRM
      - IAM Integration and development (CNAF SD)
  - IDDLS
  - CNAF-CERN project T3.2/3.3
- **SW Development**
  - WLCG-DOMA Working Groups
  - XDC (R.I.P.)

ESCAPE

ESCAPE DataLake





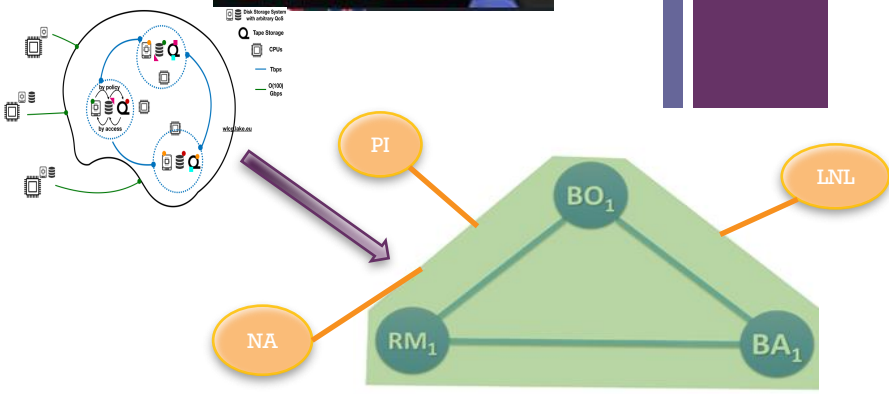
- Hiding complexity and providing transparent access to data.
- Heterogeneous federated storage and operations model.
- Some centers joining even if not funded by ESCAPE.

Further info: https://wiki.escape2020.de/index.php/WP2_-_DIOS#Datalake_Status

IDDLS

Italian Distributed Datalake for Science

- INFN-GARR collaboration to realize a prototype of an Italian Datalake exploiting:
  - Last generation networking technologies provided by GARR
    - DCI (Data Center Interconnection) equipment
    - 3 years 2019-2021

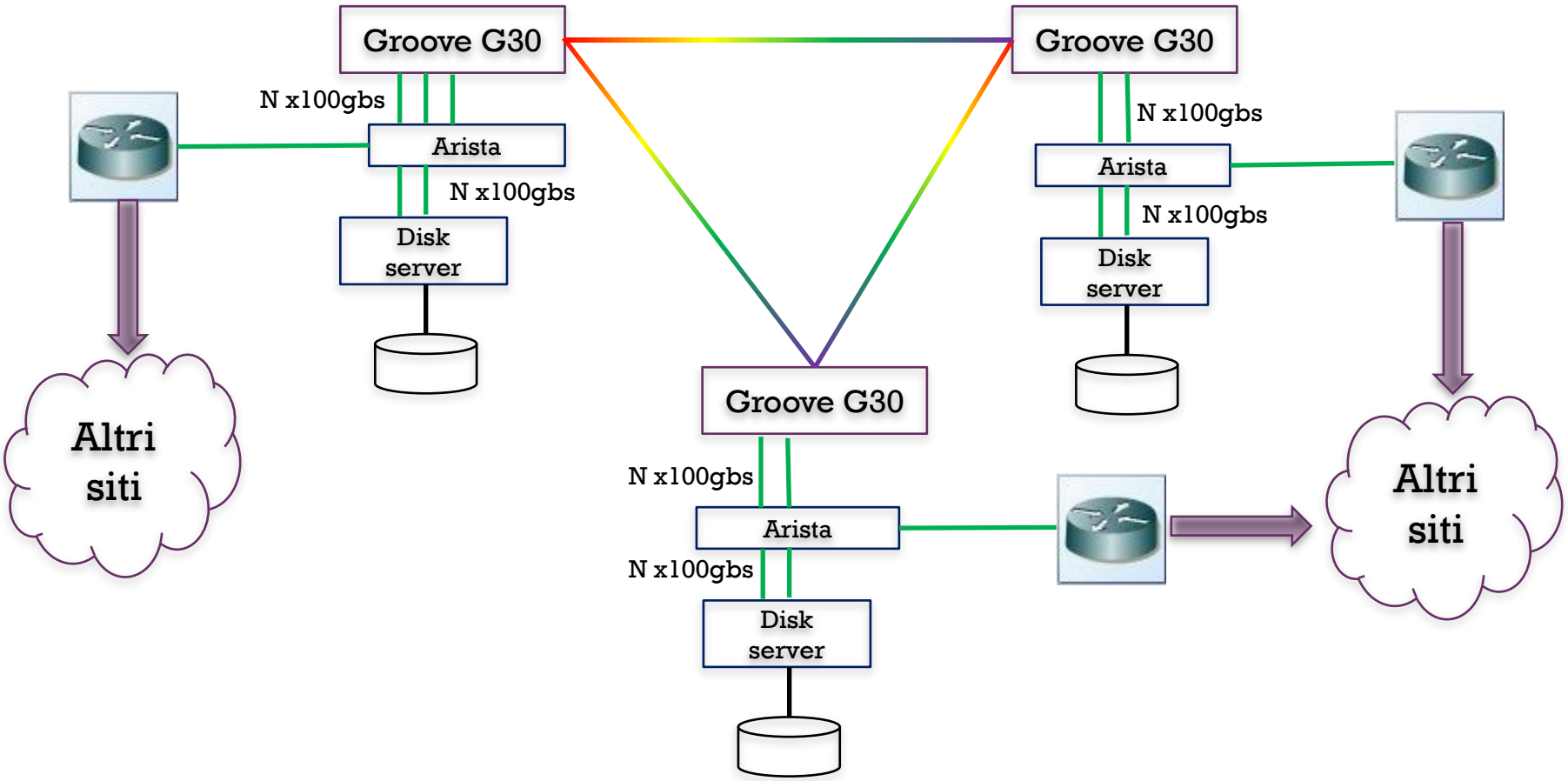


Topology of the GARR Network with DCI for the Datalake

- Status
  - Networking hw acquired after evaluation of 3 solutions from different vendors
    - Transponders (GARR)
    - Switches and Servers (INFN)
  - Very preliminary tests
  - **Stuck during the last year due to Covid**

- Next Steps - 2021
  - Deployment of the networking systems on the sites
  - Creation of the Storage federation + QoS management
    - RUCIO + FTS
  - Tests with Experiments

IDDLS Network Architecture

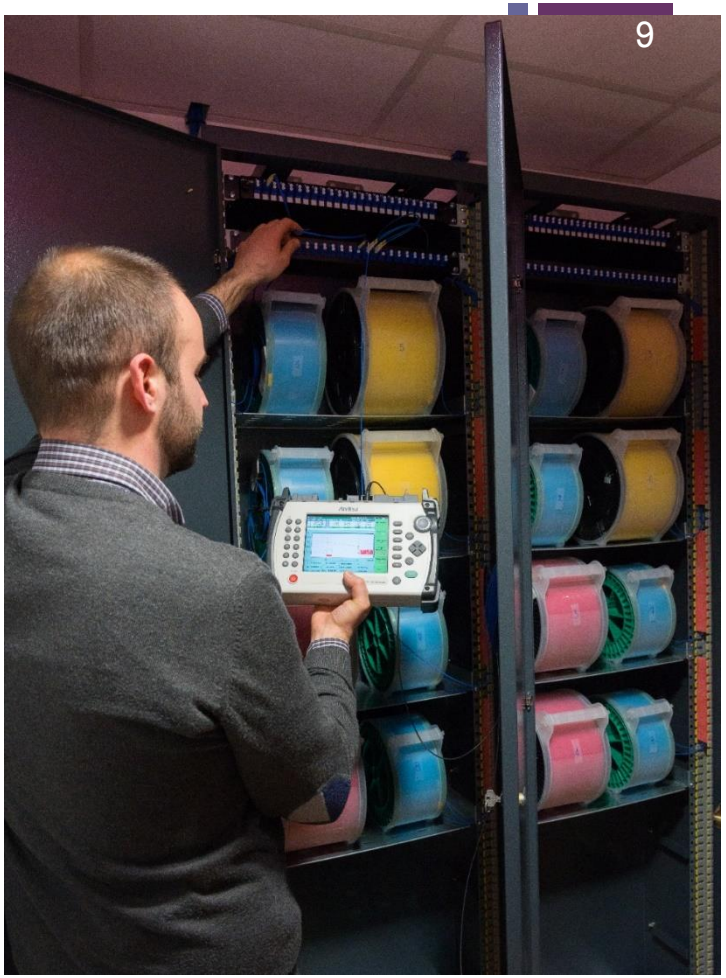


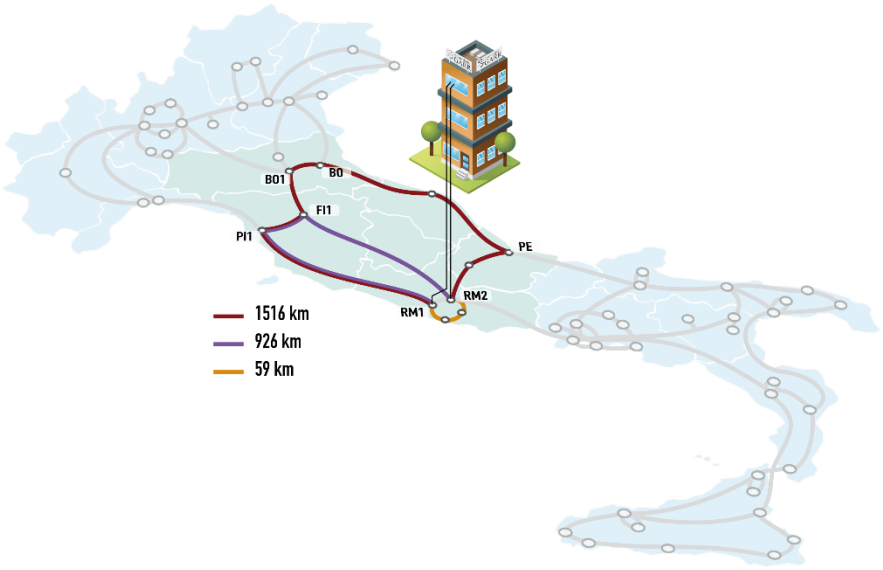
Fiber spools

- 16 spools of single G.652d fiber
- vendor: Fujikura
  - 8 spools 50km -> 4 pairs
  - 8 spools 25 km -> 4 pairs

  In total 600km single fiber

  The **flexible topology** changes through a patch panel on the rack top





CNAF-CERN Project: WP3 Task3.2

- Implementation of **automatic replication** in different, geographically distributed, availability zones
  - Performed at the infrastructure level
  - Quasi-real-time replication
  - Different QoS for the replicas
- Activities:
  - Evaluation of the adoption of a community-independent, storage management **orchestrator**
    - based on RUCIO (+FTS)
  - Evaluation of **different mechanisms** to implement the Datalake
  - Evaluation of the **storage events notifications** approach for quasi-real-time replication through different storage solutions.
    - Evaluation of the replication performed at the filesystem level without an orchestrator
  - Implementation of the **storage federation** with the sites belonging to the Datalake

Evaluation of Storage Events

- Intercept relevant events on the filesystem and receive corresponding notification
  - File Creation
  - File Write
  - Directory access
- **Inject the notifications into the policy manager system** to trigger data movements
  - Within the site or the Lake, including QoS change
    - QoS change without scanning the buffer
- Technologies tested:
  - GPFS
    - an internal utility that needs further licenses
  - CEPH
    - script developed ad hoc to parse log files
- Preliminary test successfully completed
  - Mainly on write operations
    - 0.8Hz (gpfs)
    - 2.8 Hz (ceph)

CNAF-CERN Project WP3 Task3.3

- Targets the non-LHC experiments and the **"long tail of science"**
  - To provide a lightweight point-to-point data transfer service.
  - To providing an **enriched data management Web service with sync-and-share capabilities**
  - A free-for-use **competitor of Globus Online**.
    - As an evolution of the CERN FTS

- The following activities are proposed:
  - Evaluation of the available data management solutions to be offered to not-experienced user communities and selection of the most suitable one.
  - Integration and/or development of a common sync-and-share solution for end-users
  - Evaluation of the proposed solution and testing of the integrated services
  - Feedback loop among users and developers
  - Integration with RUCIO and FTS as backend services to steer data movements based on policies in a transparent way for the end-user.
  - The project could potentially attract EU funding that should be invested both in the development and commissioning efforts.

CNAF-CERN Project WP3 Task3.3

- Targets the non-LHC experiments and the "long tail of science"
  - To provide a lightweight point-to-point data transfer service.
  - To providing an enriched data management Web service with sync-and-share capabilities
  - A free-for-use competitor of Globus Online.
    - As an evolution of the CERN FTS
- The following activities are proposed:

Since the very first meetings it was clear that the implementation is possible only in case of dedicated projects/initiatives to fund effort

Status - Milestones - Effort

- **IDDLS**
  - **Stuck during the last year** due to Covid
  - Ready to implement the infrastructure, but no results at the moment
  - The project will end in Dec. 2021
- **CNAF-CERN T3.2/3.3**
  - We now have experience on **RUCIO** administration/configuration
    - Worked on the containerization of both systems having in mind a future k8s deployment
  - Working a brand-new **FTS** instance (for testing)
  - We have a prototype of a system to collect **Storage Events**
    - But the integration with RUCIO is missing
  - **We need to consolidate the effort to speedup the activities and test at scale**
    - **At least 2 FTE needed** (currently about 0.5 from SDDS/T1) to have by the end of the year:
      - a working prototype that spans at least throughout the IDDLS sites
      - Support for QoS and possibly storage events
      - Tests in real life with selected experiments
- DCI CNAF ⇔ CERN
  - Stefano talk