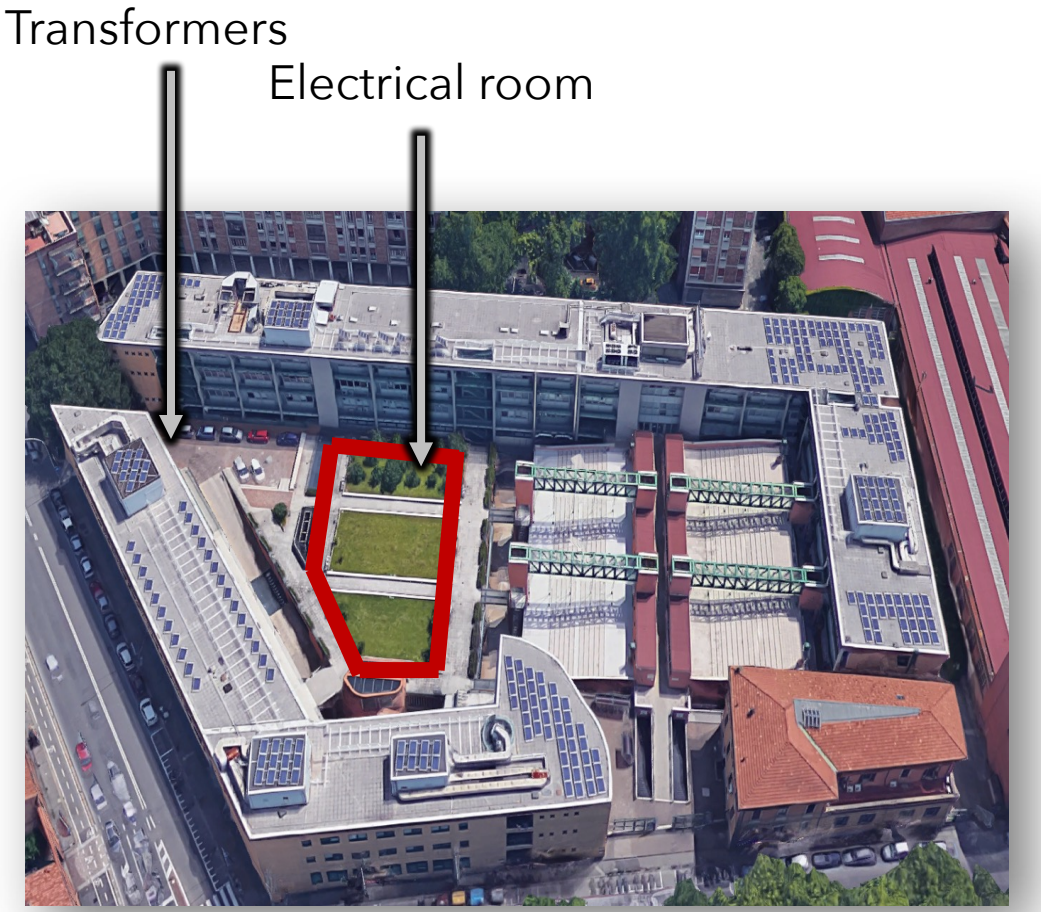
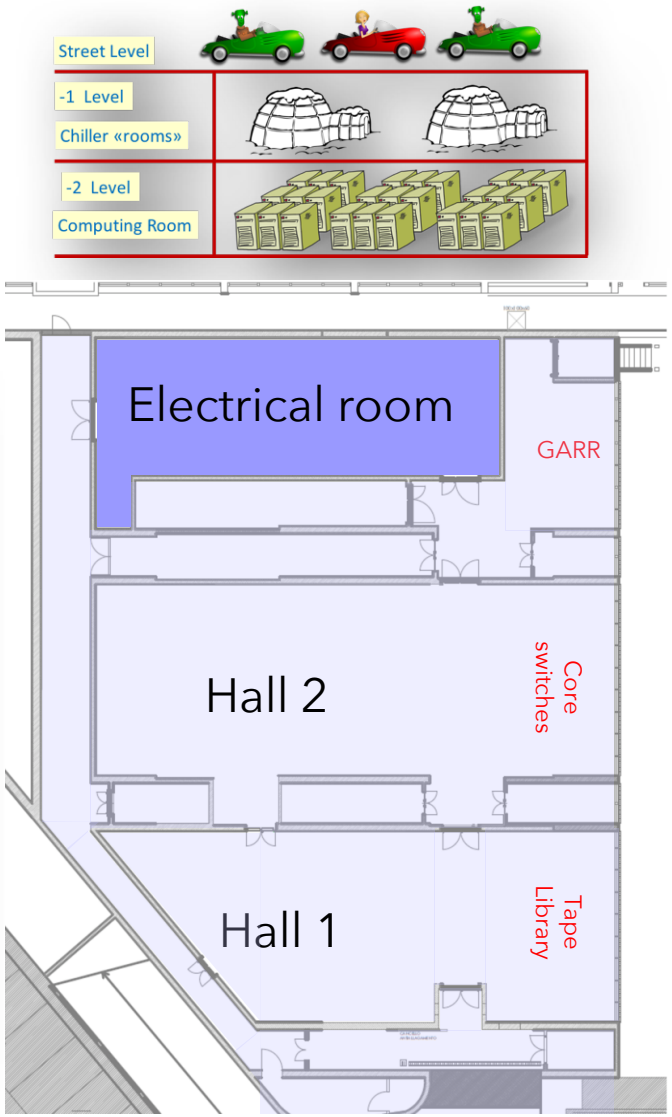
# INFN Tier-1 status

Luca dell'Agnello, Gaetano Maron

Jan 16 2018

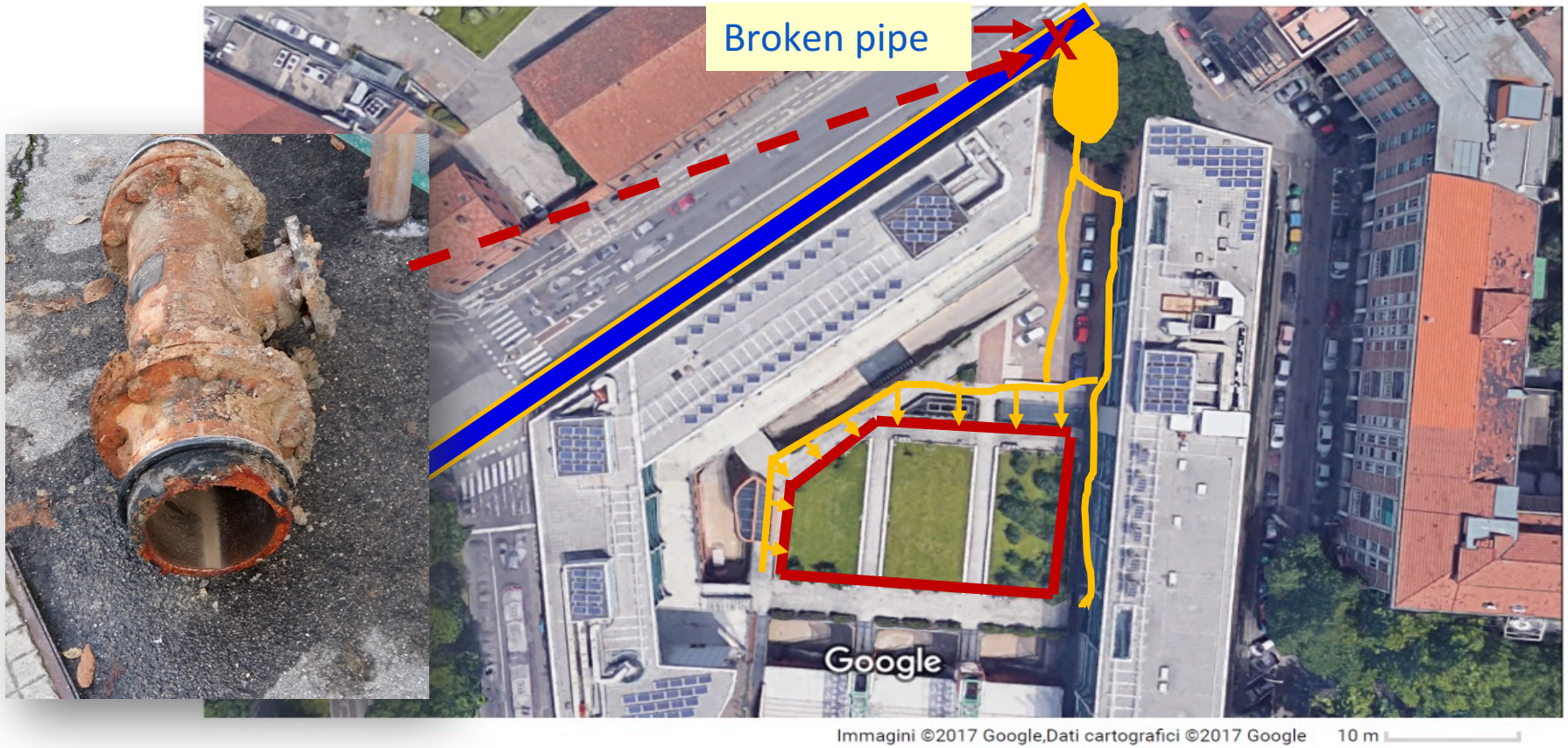
# The Tier-1 location



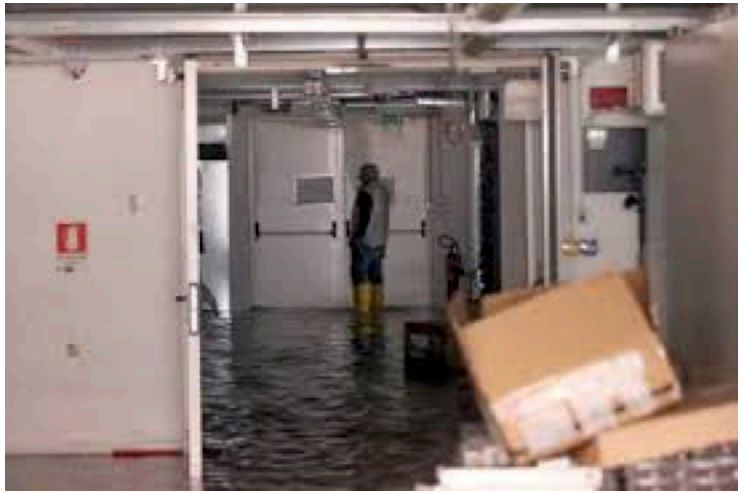
# 11/9: the flood

- The flood happened on November 9 early in the morning
  - Breaking of one of the main water pipelines in Bologna
  - Also the road near CNAF seriously damaged



# Outside the building…

# … and the Tier-1

All Tier-1 doors are watertight
Height of water outside: 50 cm
Height of water inside: 10 cm (on floating floor)
for a total volume of ~500 $m^3$

# First inspection

- Access to data center possible only in the afternoon
- Nearly all the electrical equipment in the electrical room damaged by the water
  - Both power lines compromised
- The two lower units of all racks in the IT halls submerged
  - Including the two lowest rows of tapes in the library

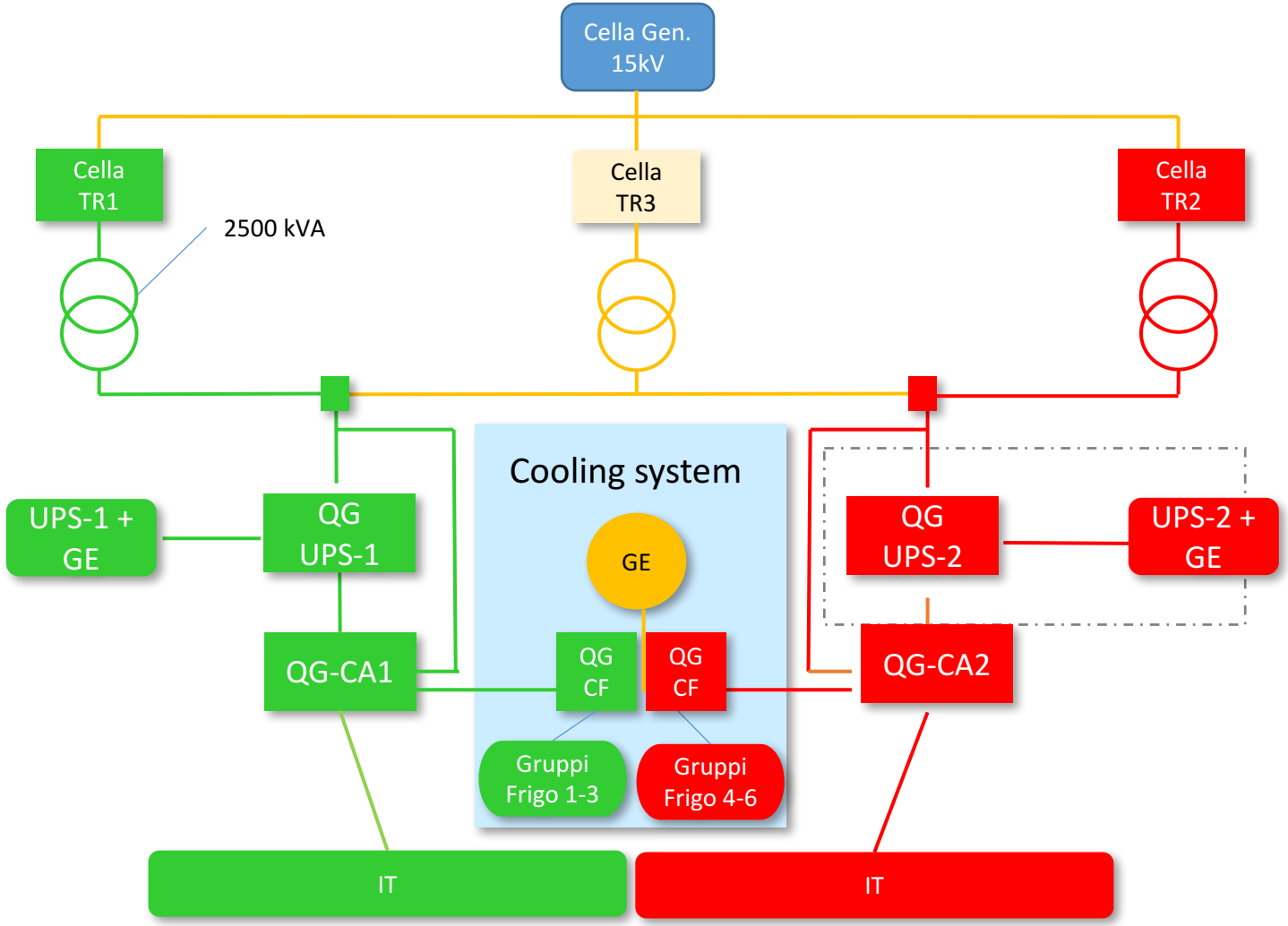
# What has been done (1)

- IT services (non scientific computing) immediately moved outside CNAF
- Data center dried over the first week-end
- Activated a temporary power line (60 kW) after 1 weeks
  - Essential for GARR POP equipment
- The General IP connectivity restored few days after the flood
- Cleaning from dust and mud completed during the first week of December
- Core switches tested and upgraded to 100 Gbit (Dec 15)
  - Needed to install new storage
- Recovery of first electrical line (1.4 MW) completed before Xmas

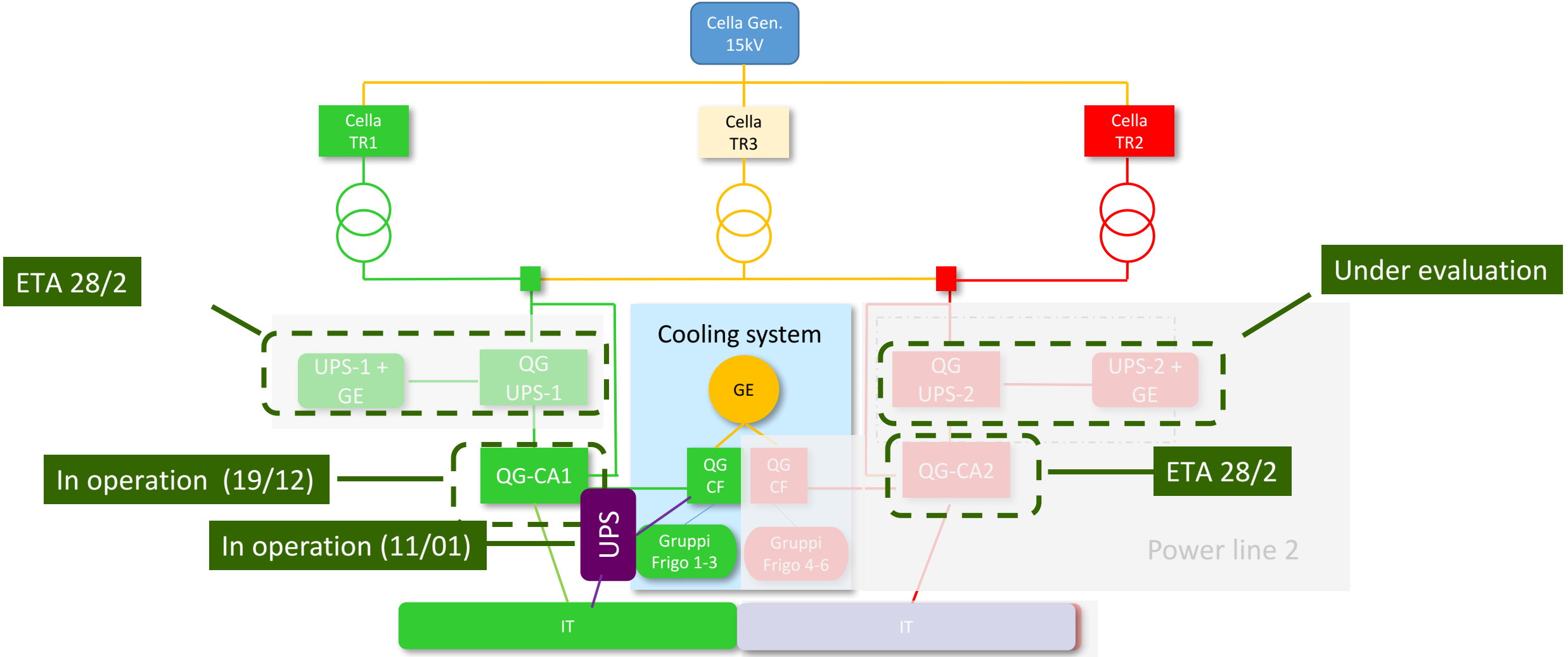
# What has been done (2)

- In the meanwhile activity to recover wet IT equipment
  - Cleaned and dried (using oven when appropriate) disks, servers, switches,….
- IT components to be replaced have been ordered
- Temporary UPS (300 kW) + GE in place since Jan 12
  - Enough to switch on all the storage systems
- 3 chillers (out of 6) in operation since Jan 15
  - Possible to switch on also part of the farm (no continuity)
- Other miscellaneous activities (e.g. floating floor replacement, fire alarm system recertified etc…)

# Power Center configuration before the flood



# Present Power Center status



# In the meanwhile….

- Deep inspection of the data center to understand the flow of the water
- Recovery of wet IT equipment
  - Cleaned and dried (using oven when appropriate) disks, servers, switches,….
  - IT components to be replaced have been ordered
- Installation of tender 2017 storage on going
  - Completed for mid February

# Damage to IT equipment (1)

- Computing farm
  - ~34 kHS06 are now lost (~14% of the total capacity)
  - No special action taken
    - Cleaning and drying where possible or order new systems
- Library and HSM system
  - 1 drive damaged
  - Cleaning needed (completed)
  - Recertification on going (ETA: Jan 19)
- Tapes
  - 136 tapes damaged
  - **“wet” tapes are being recovered in Oracle lab**
    - ETA: end of February

# Damage to IT equipment (2)

- Nearly all storage disk systems involved
  - 11 DDN JBODs (LHC, AMS)
    - RAID parity affected
  - 2 Huawei JBODs (all non-LHC experiments excepting AMS, Darkside, Virgo)
  - 2 Dell JBODs including controllers (Darkside and **Virgo**)
    - Most critical - 2 trays out of 5 went underwater. **RECOVERED**
  - 4 disk-servers (4 Alice) + 4 TSM-HSM servers

| System | PB | JBODs | Disks | Involved experiments |
|---|---|---|---|---|
| **Huawei** | 3.4 | 2 | 150 x 6 TB | All CSN2 and 3 experiments excepting AMS, Darkside e Virgo |
| **Dell** | **2.2** | **2** | **120 (48) x 4 TB** | Darkside and **Virgo** |
| **DDN 1,2** | 1.8 | *4* | | ATLAS, Alice and LHCb |
| **DDN 8** | 2.7 | *2* | | LHCb |
| **DDN 9** | 3.8 | 2 | | CMS |
| **DDN 10, 11** | 10 | 3+2 | 252 x 8 TB | ATLAS, Alice and AMS |
| **Total** | **23.9** | **9** | **~4 PBytes** | |

# Storage recovery roadmap

- Dell systems recovered
  - Replacement of damaged parts only (i.e. crates)
    - Compromised disks can be replaced during normal operations
    - LUNs show optimal conditions
  - Switched on Jan 15
  - Servers under test
  - File-system to be verified
- Replacement parts for DDN10 and DDN11 to be delivered on 25/1
- Huawei replacement parts to be delivered not before 27/1
- DDN1, DDN2: damaged components replaced with spare parts: to be tested!
- Data on DDN8 (out of maintenance) will be moved onto new storage (after acceptance test: mid of February?)
- Disks of DDN8 will be used to replace wet disks of DDN9

# Storage recovery status

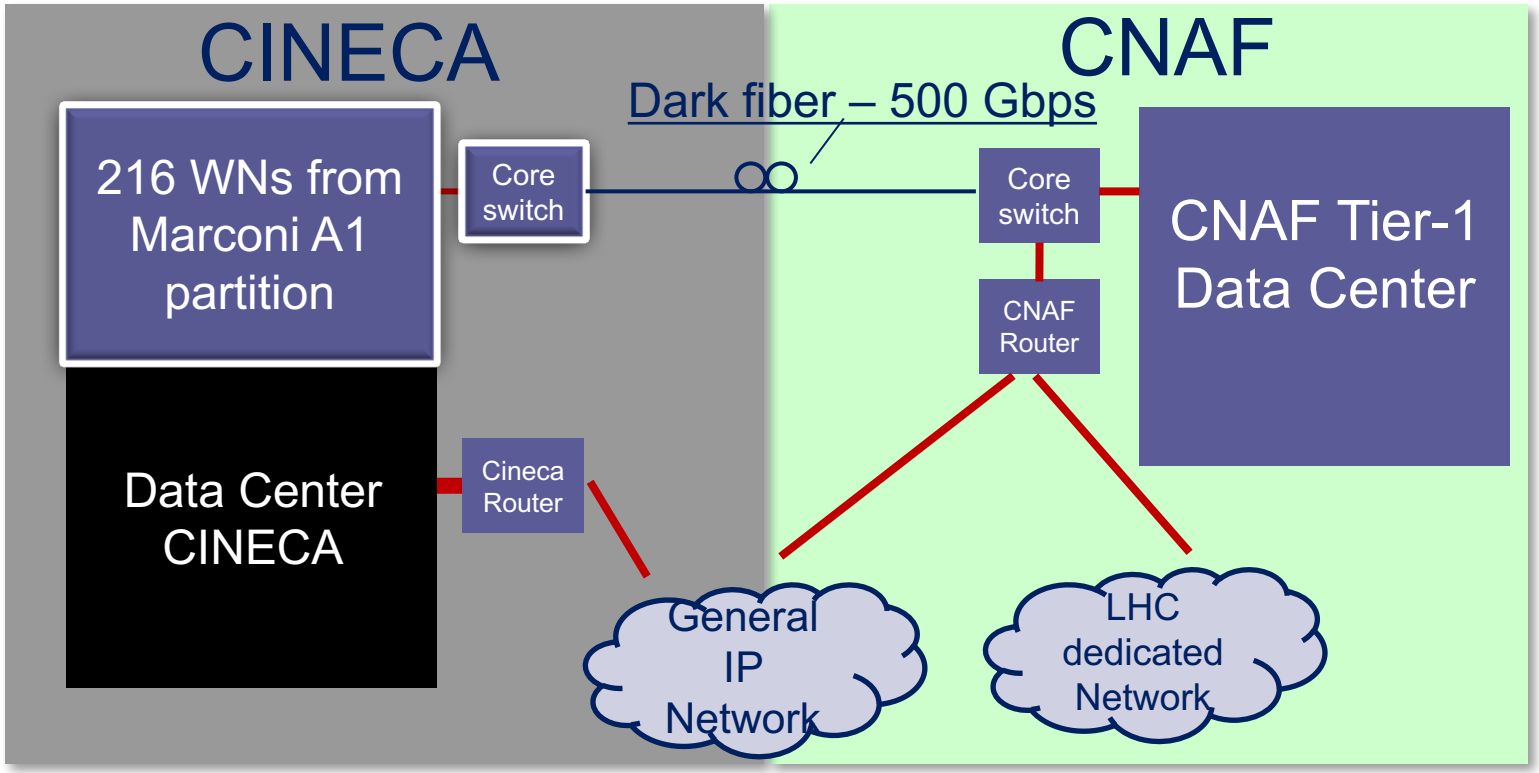
| System | PB | Strategy | Involved experiments | Status |
|---|---|---|---|---|
| **Huawei** | 3.4 | Replacement of damaged components | All CSN2 and 3 experiments excepting AMS, Darkside e Virgo | ETA> 27/1 |
| **Dell** | **2.2** | **Replacement of damaged components** | **Darkside and Virgo** | **OK** |
| **DDN 1,2** | 1.8 | Move data to new storage | ATLAS, Alice and LHCb | To be tested |
| **DDN 8** | 2.7 | Move data to new storage | LHCb | |
| **DDN 9** | 3.8 | Repaired using DDN8 disks | CMS | |
| **DDN 10, 11** | 10 | Replacement of damaged components | ATLAS, Alice and AMS | ETA: 25/1 |
| **Total** | **23.9** | | | |

(*) Replacement procedure: replacement of crates, switching on with old disks, verification of the integrity of the data and then replacement, one by one, of the disks

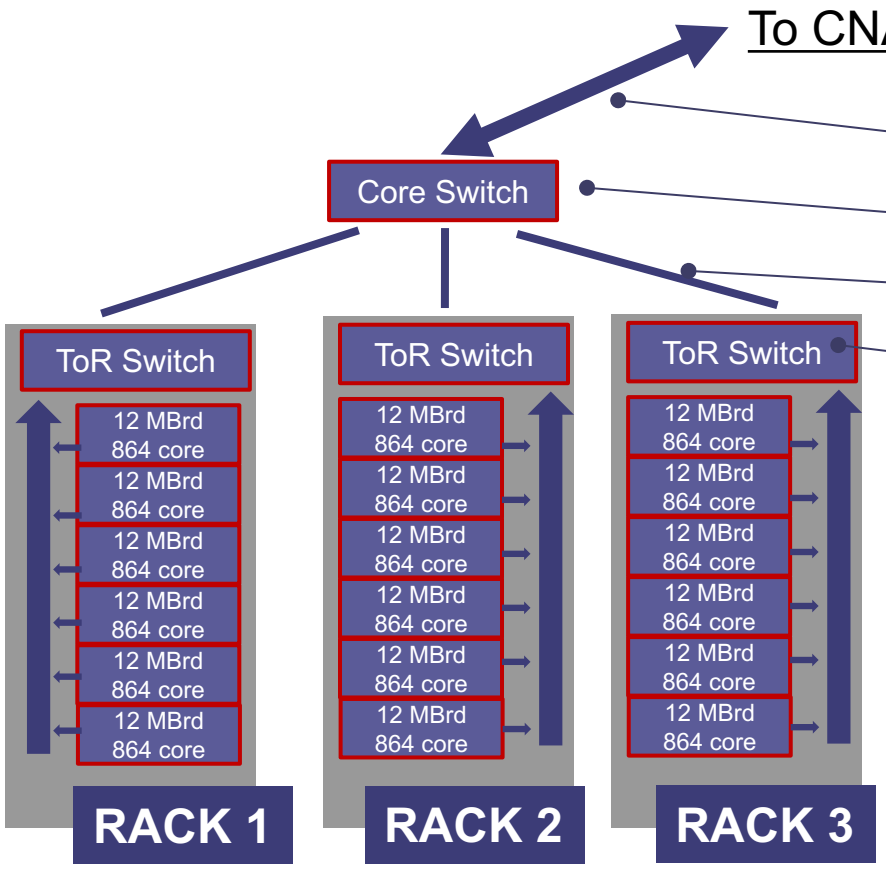
# Farm recovery

- Test and recovery of farm services starting this week
  - LSF masters, CEs, squids etc…
- Then upgrade of WNs
  - Middleware, security patches (i.e. meltdown etc..)
- Only part of the local farm can be powered on (only 3 chillers in production) and w/o continuity
  - 30-50 kHS06 (to be verified)
- Exploiting the CNAF farm elastic extension to provide more computing power
  - Restart of remote farm partition in Bari-RECAS (~24 kHS06)
  - Install the CNAF-CINECA extension farm (~ 170 kHS06)

# Interconnection CINECA – CNAF

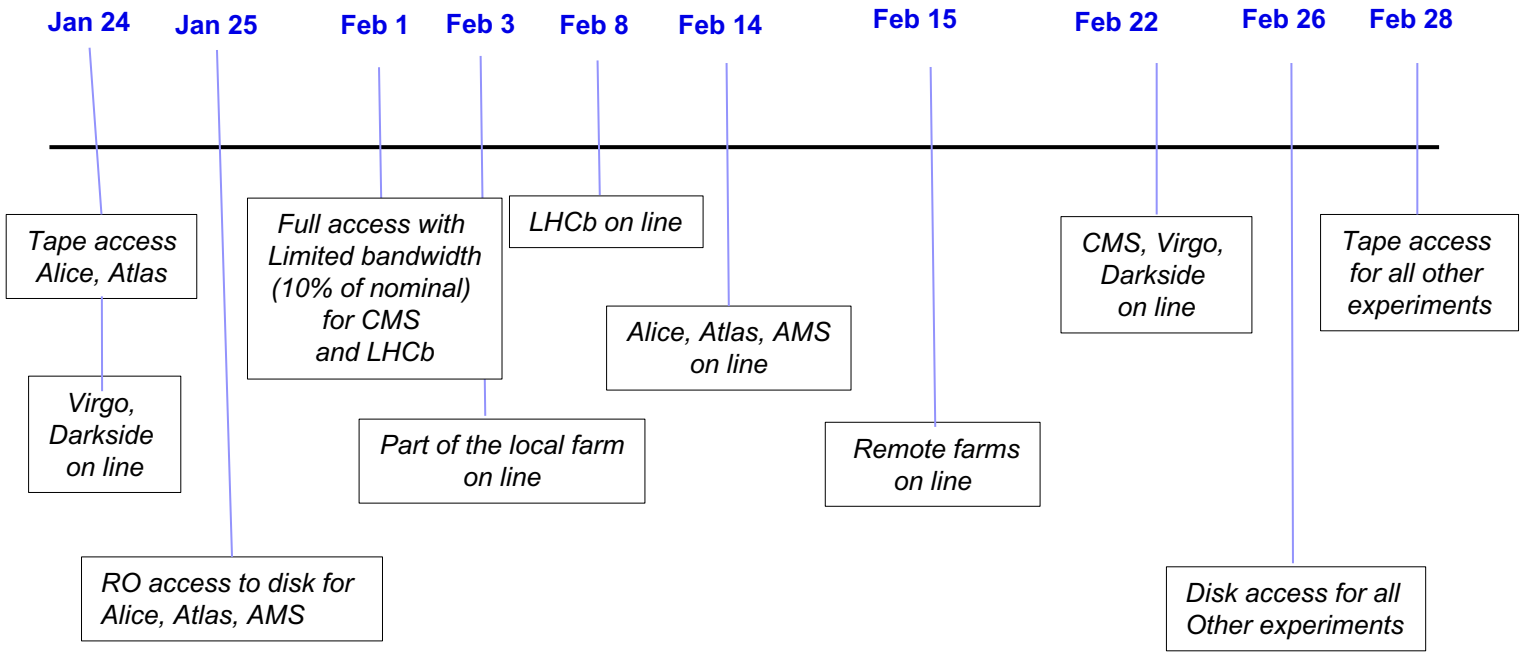


# FARM INFN-Tier1 at CINECA




- 500 Gbps link
- 32 ports 100 Gbps
- 2x100 Gbps ( 4 x 40)
- ToR Switch:
  - >= 72 ports 10 Gbps baseT
  - 2 Uplinks 100 Gbps (4 x40)

# Tentative timeline




*RO = read only*

# Summary

- Virgo data on disk apparently ok
  - File system to be verified
- Library and HSM back in operation before end of January
- Part of the farm in operation in the first half of February