

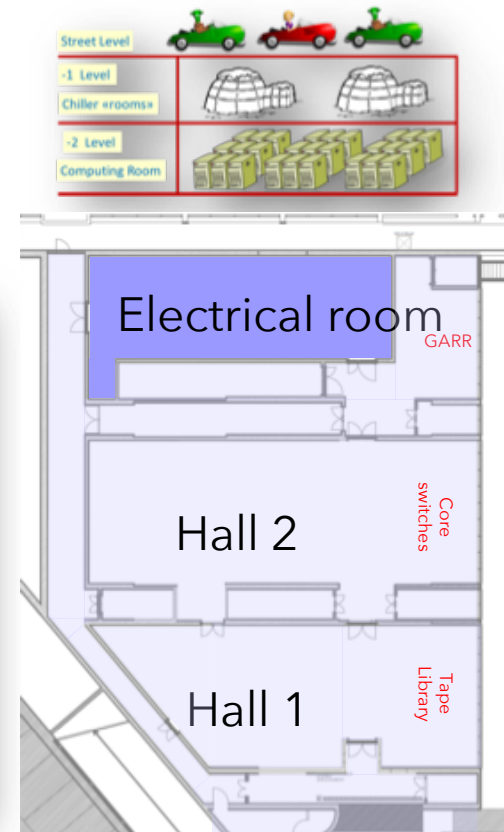
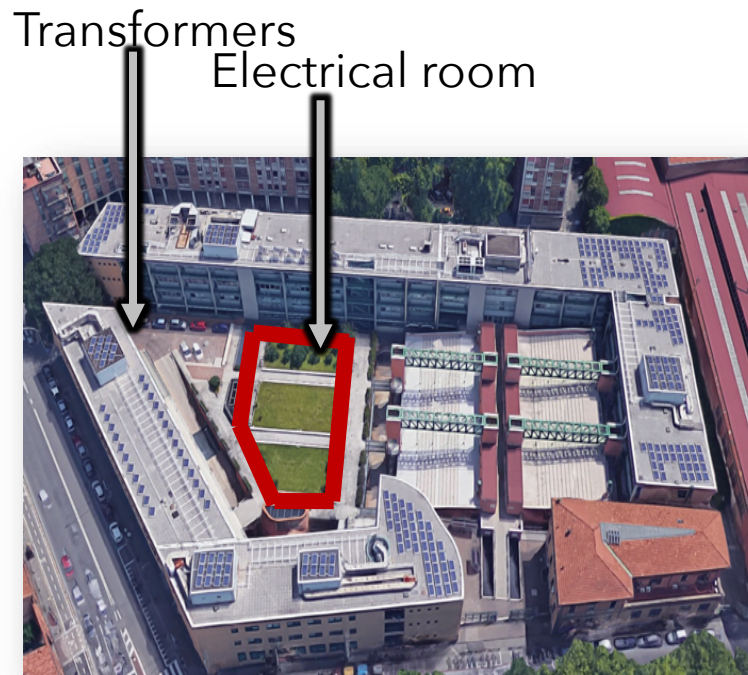


# Tier-1 status

Luca dell'Agnello

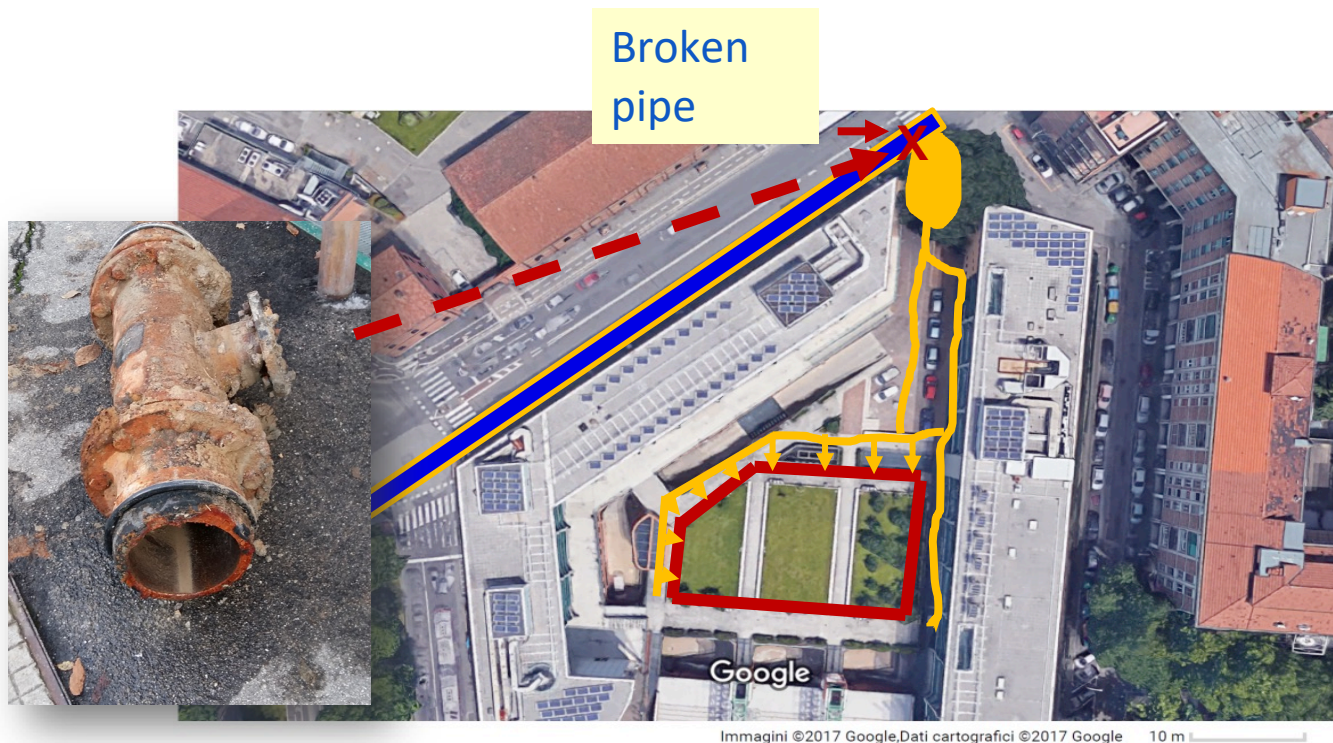
31/01/2018

# 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...





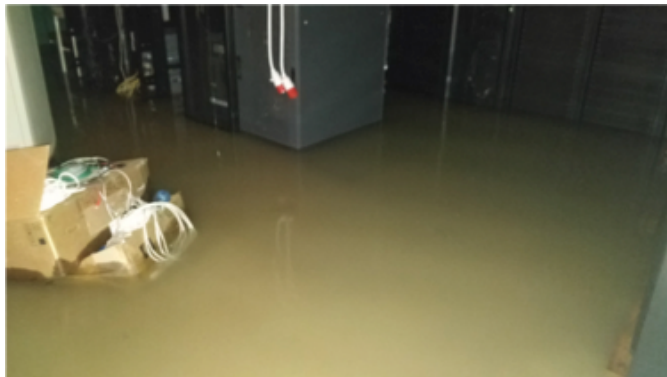
# ... and the Tier-1



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  $\sim 500 \text{ m}^3$

# 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



# Damage to IT equipment

## ■ 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 and other components damaged (*support intervention*)
- 4 TSM-HSM servers lost (*replacement servers ordered*)
- 136 tapes damaged (*being recovered in Oracle lab*)

## ■ 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)
- 4 disk-servers (Alice)
- 3 FC switches

# What has been done (1)

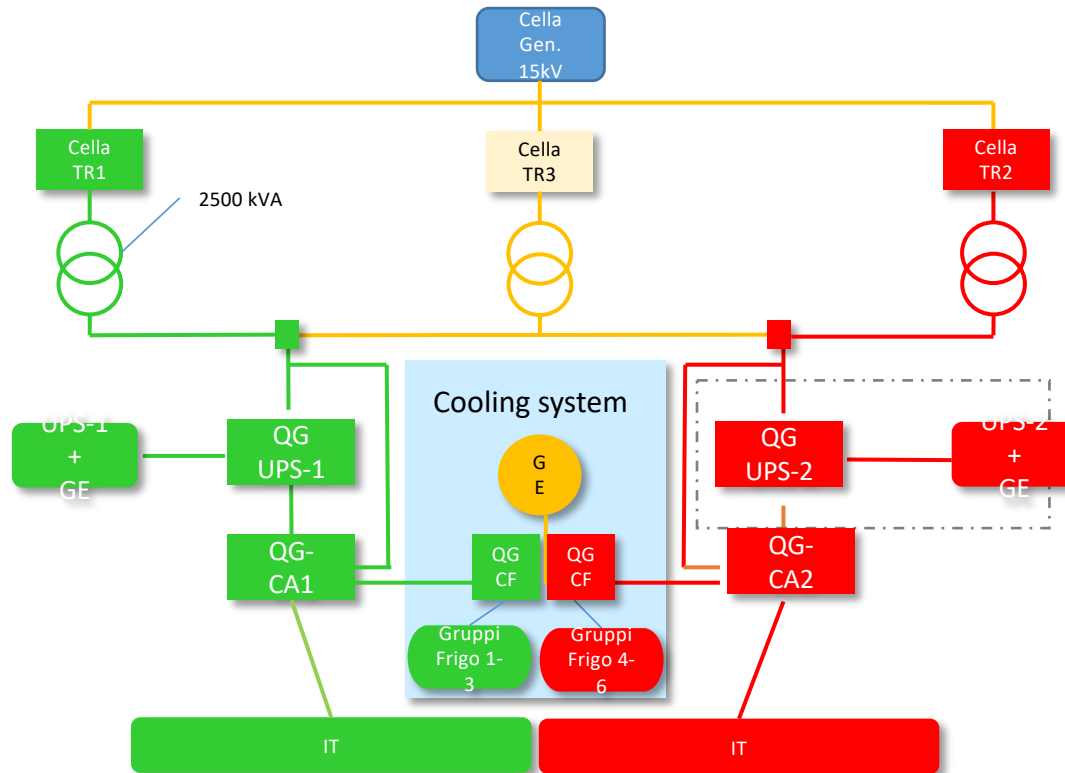
- IT services (non scientific computing) immediately moved outside CNAF (since November 10)
- Data center dried over the first week-end
- Activated a temporary power line (60 kW) after 1 weeks
  - Essential for GARR POP equipment
- The General IP connectivity restored few days after the flood
- Cleaning 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



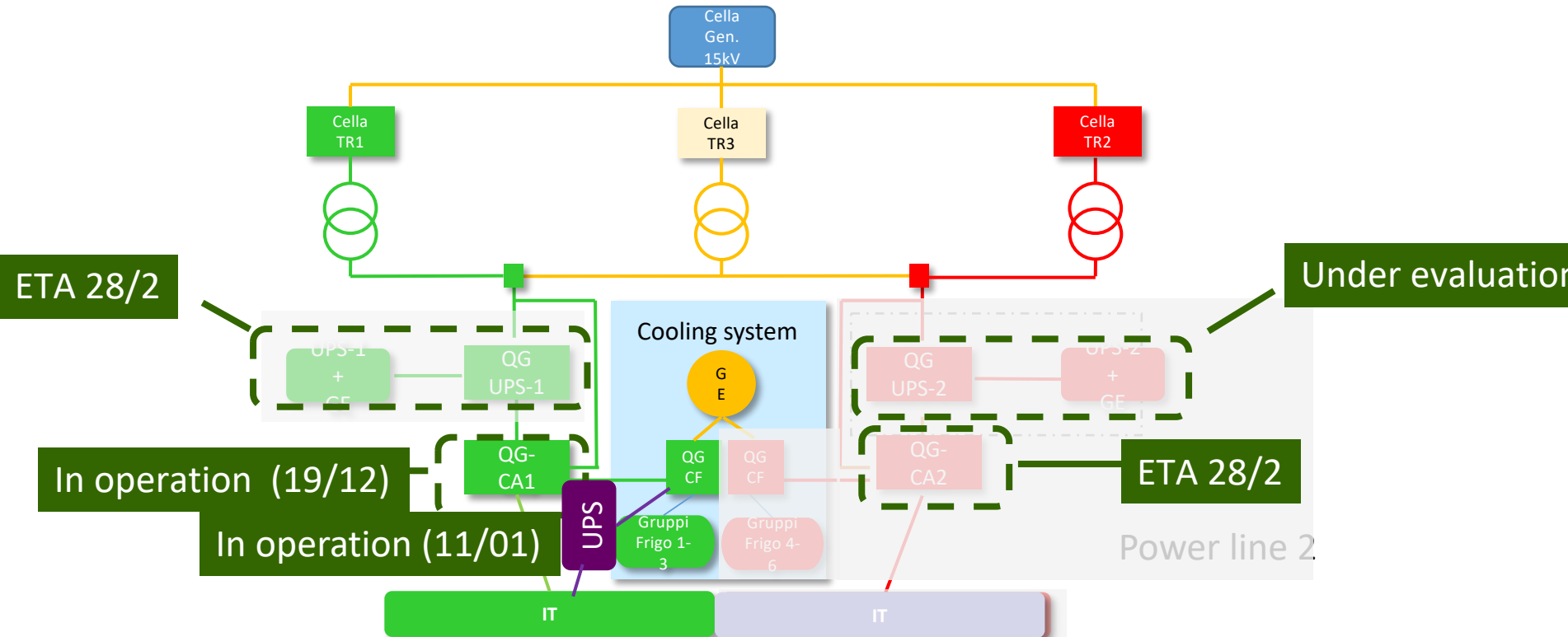
# What has been done (2)

- 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 and “urged” installation of tender 2017 storage
- Ordered replacement components for storage
- Temporary UPS (300 kW) + GE in place since Jan 12
  - Enough to switch on all the storage systems
- Recertification of tape library
- 3 chillers (out of 6) in operation since Jan 15
  - Possible to switch on also part of the farm (no continuity)
- Deep inspection of the data center to understand the flow of the water
- Other miscellaneous activities (e.g. floating floor replacement, recertification of fire alarm system etc...)

# Power Center configuration before the flood



# Present Power Center status



# Storage strategy

System	Strategy	Involved experiments
Huawei	Replacement of damaged components	All CSN2 and 3 experiments excepting AMS, Darkside e Virgo
Dell	Replacement of damaged components	Darkside and Virgo
DDN 1,2,...,6	Move data to new storage	ATLAS, Alice and LHCb
DDN 8	Move data to new storage	LHCb
DDN 9	Repaired using DDN8 disks	CMS
DDN 10, 11	Replacement of damaged components	ATLAS, Alice and AMS

Key element: availability of 2017 tender storage

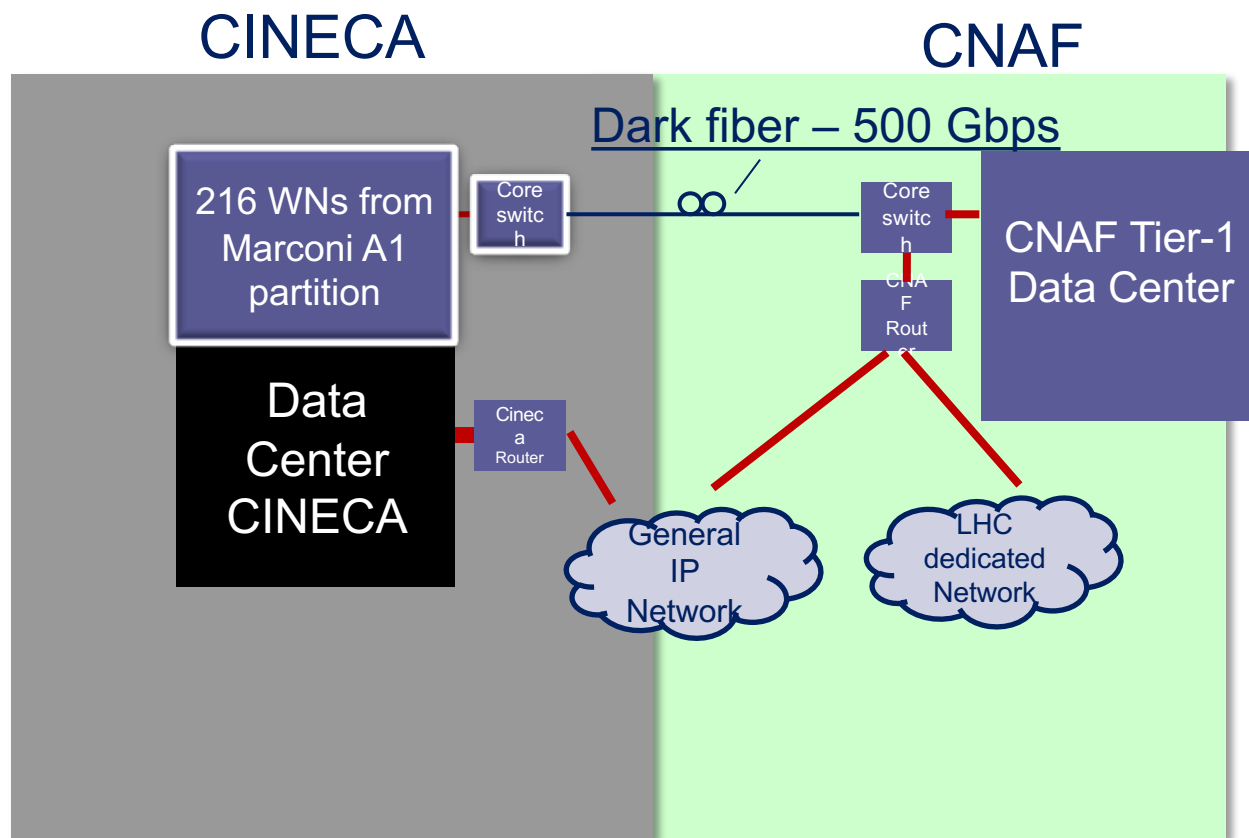
(\*) Replacement procedure: replacement of crates, switching on with old disks, verification of the integrity of the data and then replacement, one by one, of the disks

# Farm status

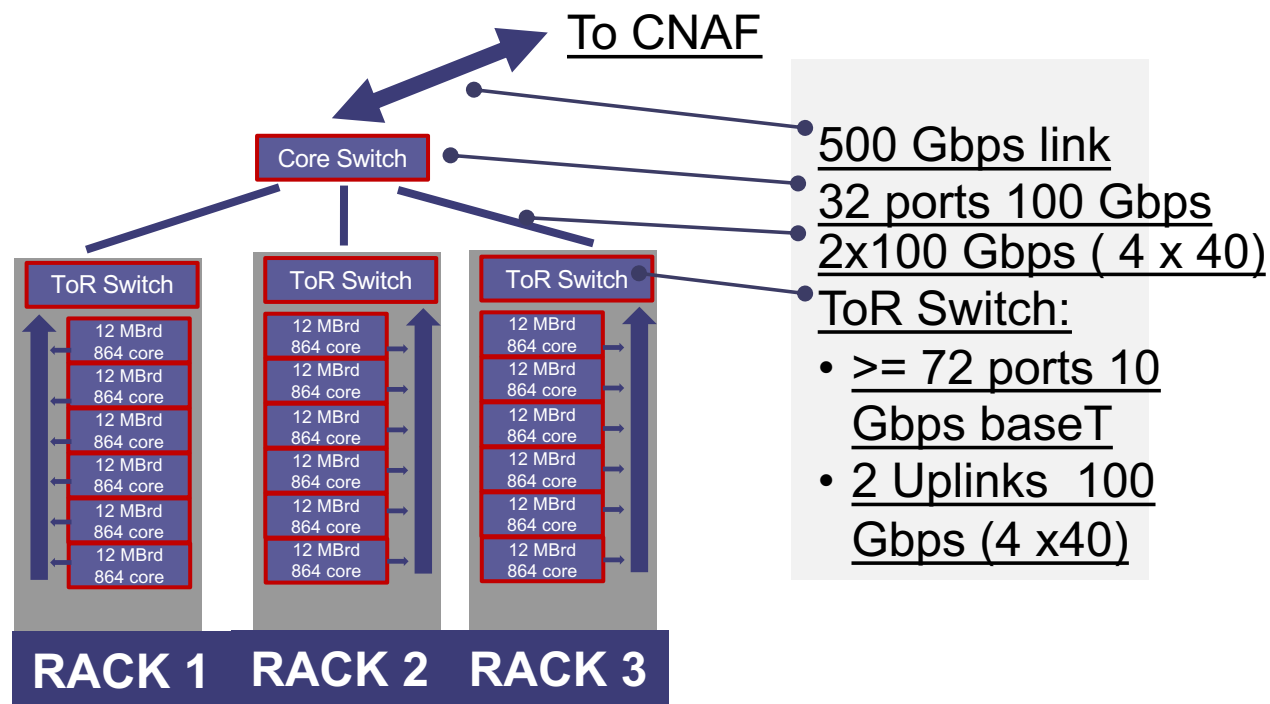
- Farm and auxiliary systems (e.g. provisioning, accounting monitoring, DOCET...) recovered
- Only part of the local farm available (~50 kHS06)
  - only 3 chillers in production and w/o continuity
- VPN CNAF-BARI ok
- Remote farm partition in Bari-RECAS (~24 kHS06)
- CNAF-CINECA extension farm (~ 170 kHS06) should be ready from mid February
  - VPN to be set and tested
  - WNs to be installed



# Interconnection CINECA – CNAF



# FARM INFN-Tier1 at CINECA



# Summary

- Farm open for Alice, Darkside and Virgo
  - Alice disk-less
- Still working on recovery storage for other experiments
- Library ready but HSM services

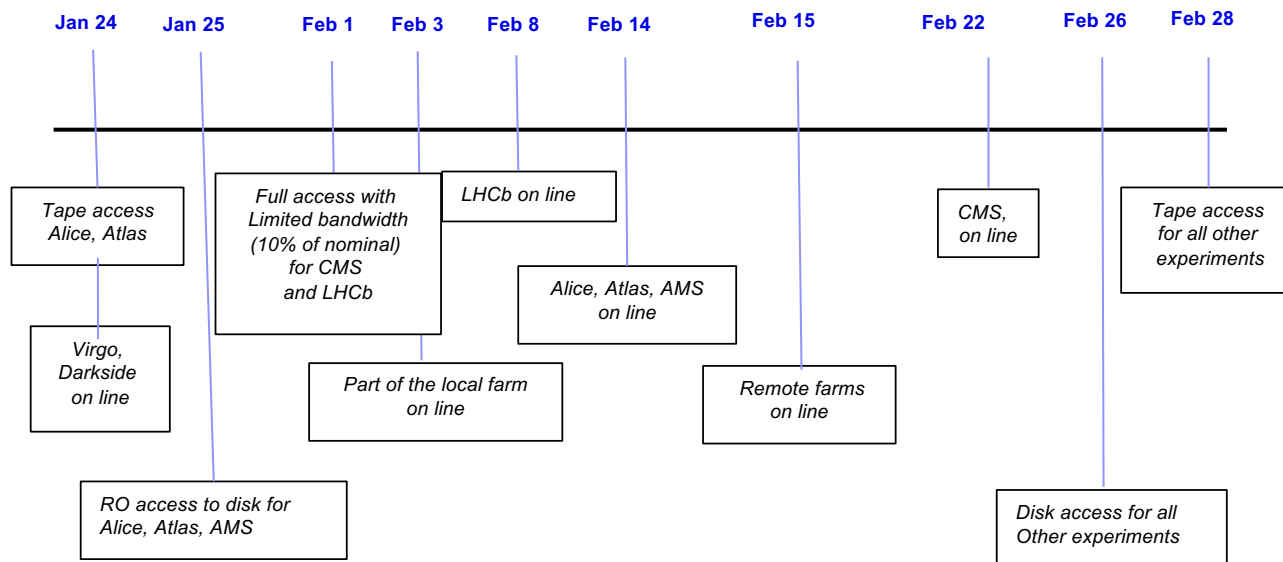
# BACKUP

# Storage status

System	PB	Activity	Involved experiments	Status	Readiness
Huawei	3.4	TBD (26/1)	All CSN2 and 3 experiments excepting AMS, Darkside e Virgo + Archive	Critical	
md1	2.2	Replacing disks	Archive, Darkside and Virgo	1 <sup>st</sup> fs check OK	fs OK Virgo and
md2			Archive, Darkside and Virgo	1 <sup>st</sup> fs check OK	Archive
md3		Replacing disks	Archive, Darkside and Virgo	1 <sup>st</sup> fs check OK	need also library
md4			CMS	LUN OK	
Md5,6		Disks replaced	Metadata exp LHC	To be recovered	After DDN1
DDN 1	0.9		Buffer ATLAS, Alice	LUN OK	library
DDN 2,4	0.9	TBD	LHCb		
DDN 3			Backup db TSM	LUN OK	
DDN 5,6	0.9	TBD	Atlas		
DDN 8	2.7	TBD	LHCb	February	
DDN 9	3.8	TBD	CMS	After DDN8	
DDN 10, 11	10	TBD (>29/1)	ATLAS, Alice and AMS	ETA: 29/1	
Total	23.9				



# Tentative timeline



RO = read only

Esperimenti	Stato dischi	Stato Buffer HSM	Stato complessivo	Note
VIRGO	file system (fs) check ok su tutti i dischi	<ul style="list-style-type: none"> <li>FS ok</li> <li>attesa library</li> </ul>	Ready to go ON-LINE	Tape library on-line 31/1
DARKSIDE	fs check ok su tutti i dischi	<ul style="list-style-type: none"> <li>da verificare (il buffer è condiviso con gli altri)</li> </ul>	ON-LINE	
ALTRI ESP CSN2 E CSN3	NUOVI enclosure e DISCHI CONSEGNATI	<ul style="list-style-type: none"> <li>da verificare</li> <li>attesa library</li> </ul>	<ul style="list-style-type: none"> <li>INSTALLAZIONE IL 26/1</li> </ul>	Tape library on-line 31/1
AMS	In attesa di consegna dei dischi e enclosure (29/01)	<ul style="list-style-type: none"> <li>da verificare (il buffer è condiviso con gli altri)</li> </ul>	<ul style="list-style-type: none"> <li>CONSEGNA IL 29/1</li> <li>ON-LINE 12/2</li> </ul>	
ATLAS	<ul style="list-style-type: none"> <li>DDN 10 E 11; DISCHI DA CONSEGNARE</li> <li>DDN 5,6: LUN OK</li> </ul>	<ul style="list-style-type: none"> <li>FS ok</li> <li>attesa library</li> </ul>	<ul style="list-style-type: none"> <li>CONSEGNA IL 29/1</li> <li>ON-LINE 14/2</li> <li>Tape buffer ready to go</li> </ul>	Tape library on-line 31/1
ALICE	<ul style="list-style-type: none"> <li>DDN 10 E 11; MATERIALE DA CONSEGNARE</li> </ul>	<ul style="list-style-type: none"> <li>FS ok</li> <li>attesa library</li> </ul>	<ul style="list-style-type: none"> <li>CONSEGNA IL 29/1</li> <li>ON-LINE 14/2</li> <li>Tape buffer ready to go</li> </ul>	Tape library on-line 31/1
CMS	<ul style="list-style-type: none"> <li>Dell MD4 OK</li> <li>DDN 9: INIZIO RECUPERO 12/2</li> </ul>	<ul style="list-style-type: none"> <li>FS unico da verificare</li> <li>attesa library</li> </ul>	<ul style="list-style-type: none"> <li>IN ATTESA HUAWEI</li> <li>INSTALLAZIONE 29/1</li> <li>ON-LINE 22/2</li> </ul>	<ul style="list-style-type: none"> <li>Il recupero inizia dopo che il nuovo HUAWEI è online</li> <li>Tape library on-line 31/1</li> </ul>
LHCB	<ul style="list-style-type: none"> <li>DDN2 DDN4 ok</li> <li>DDN 8: INIZIO RECUPERO 12/2</li> </ul>	<ul style="list-style-type: none"> <li>FS unico da verificare</li> <li>attesa library</li> </ul>	<ul style="list-style-type: none"> <li>IN ATTESA HUAWEI</li> <li>INSTALLAZIONE 29/1</li> <li>ON-LINE 22/2</li> </ul>	DDN8: Il recupero inizia dopo che il nuovo HUAWEI è online

# Storage recovery strategy

- Recertification of tape library and recovery of wet tapes
- Disk storage systems from 3 brands (DDN, Dell, Huawei)
- Replacement components ordered only for systems still under support in 2018 (DDN9, 10,11, DELL, Huawei)
  - DDN1, 2,3,4,5,6: damaged components replaced in house with spare parts
- 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
- Replacement parts for DDN9, 10 and 11 (yet) to be delivered
- Dell systems recovered
- Problems with Huawei replacement parts
- Key element: installation of 2017 tender storage
  - Still on-going ☹