Valerio Ippolito
INFN Sezione di Roma

Computing in DarkSide

overview of 2026 requests and longer-term planning

INFN
Istituto Nazionale di Fisica Nucleare

What “DarkSide” means

- **DarkSide-50** (DS-50)
  - data taking 2013-2019
  - data analysis still ongoing (a few papers under submission)
- **DarkSide-20k** (DS-20k)
  - data taking ~2028-2038
  - simulation and software development started since 7+ years
    - support detector design choices
    - online / offline reconstruction
  - computing technical design report under INFN review
- **smaller scale setups** bridging/extending DS-50 and DS-20k
  - ReD (directionality, calibration for low-energy recoils)
  - Proto-0 (data taking ongoing)
  - photodetector characterisation facilities (PTF, NOA…)

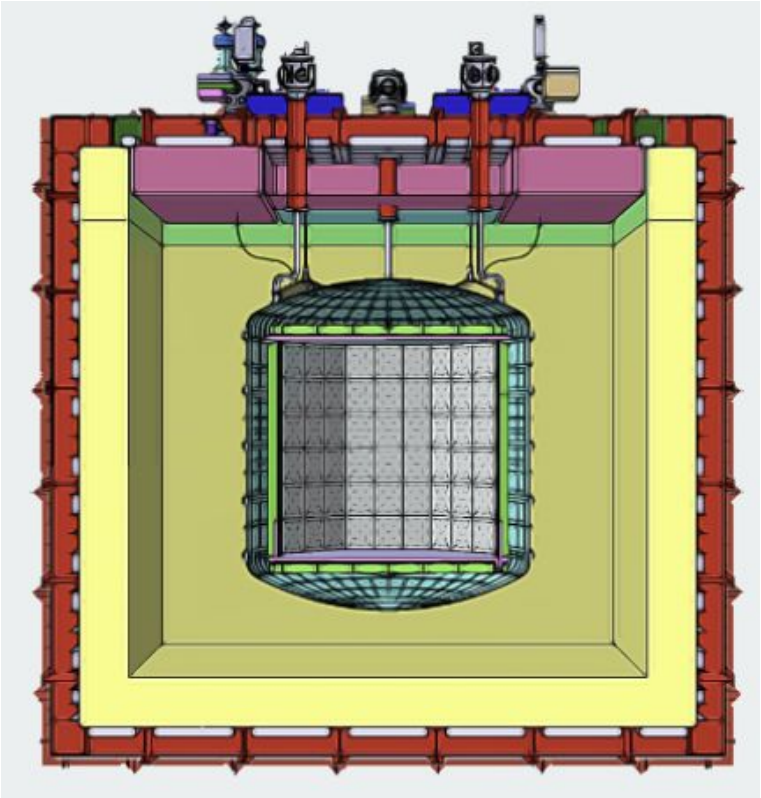
DS-20k: timeline and physics goals

- goal: discover dark matter (DM) and other new phenomena
  - WIMPs or other DM candidates, recoiling against nucleons/electrons
  - detector large enough to be sensitive to e.g. supernova neutrinos
- at LNGS: dual-phase liquid argon TPC + LAr inner veto (IV) + LAr outer veto (OV)
  - radio-pure materials and argon
  - O(100 Hz) background rate in TPC
  - 10y data taking, 2028-

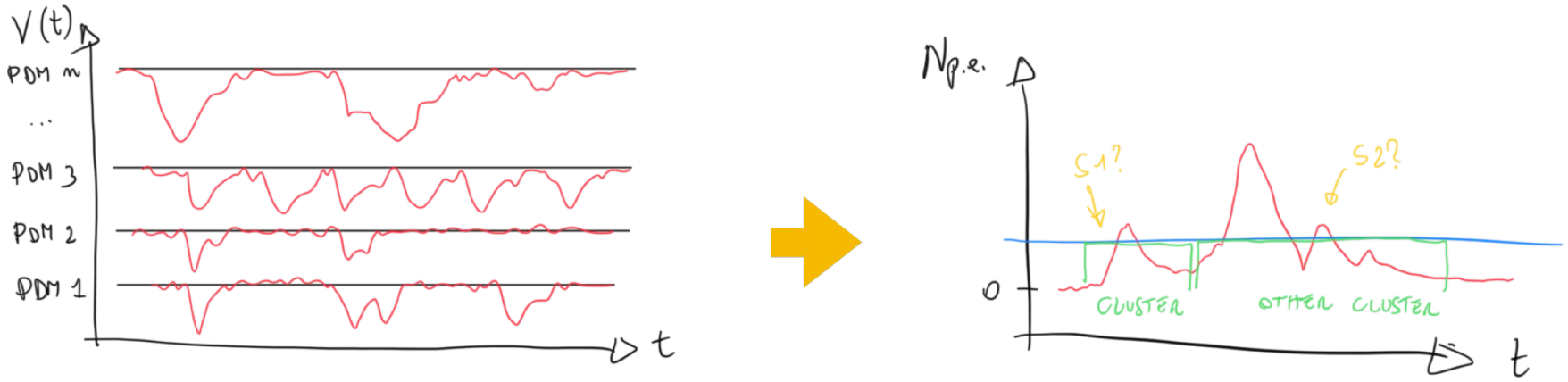
measure scintillation (S1) and ionisation (S2) signals

DS-20k as seen from DAQ/computing

- data acquisition: “trigger-less”
  - SiPM-based photodetectors, sensitive to single photo-electrons (PEs)
  - O(3000) channels (TPC+IV+OV), digitised at 125 MS/s (16 b/S)
  - max TPC drift time: 3.5 ms
- O(300) collaborators
  - ~10% of which are active in analysing data
    - expected to grow as construction phase reaches completion
  - Europe/North America/South America/Asia

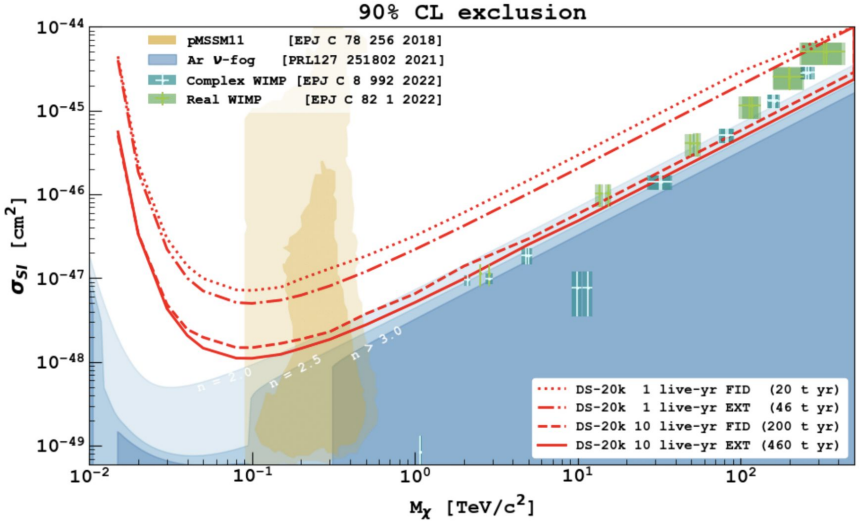


How TPC signals look like



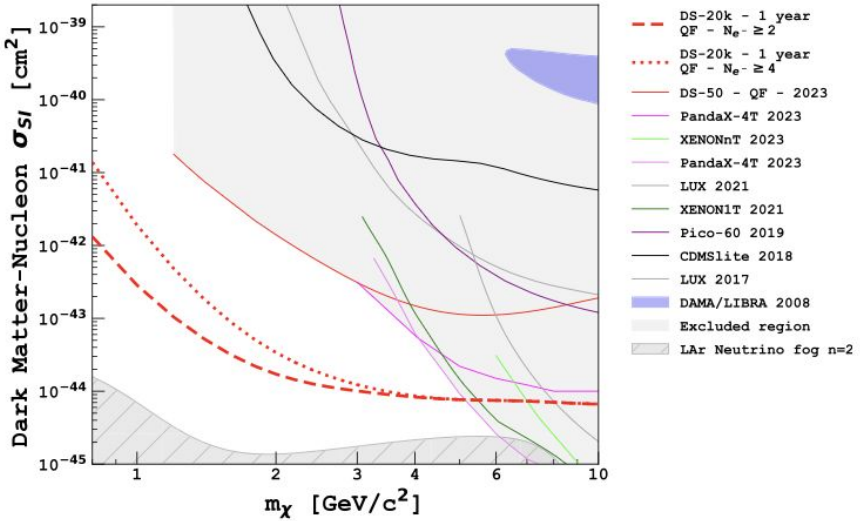
- *z* from t_{S2}-t_{S1}
- recoil type from S1 pulse shape
- *xy* from segmentation of light detection planes
- raw data = hits (filter) + waveform segments

Sensitivity to high- and low-mass WIMPs



S1+S2, ~10 years

https://doi.org/10.1038/s42005-024-01896-z



S2-only, ~1 year

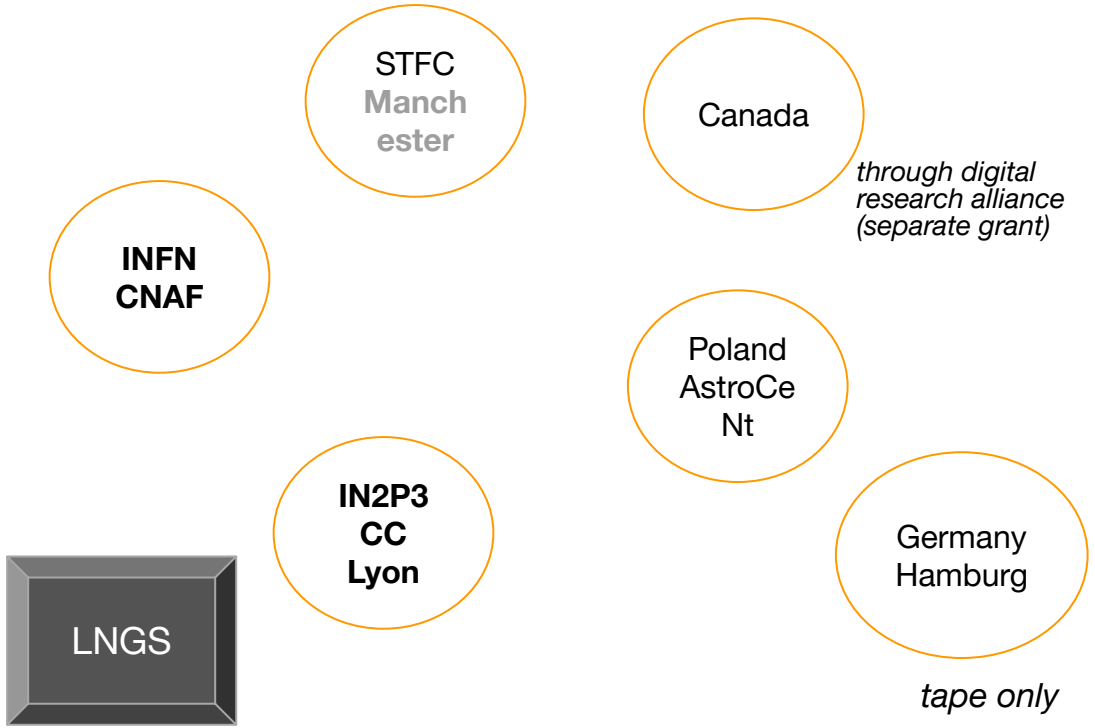
- computing technology should endure a ~15 years lifetime
- allocated resources should be enough to support fast turn-around of physics results

Computing today

how we have been working so far

Computing sites

- DS-50: CNAF + FNAL
  - plus in-kind resources at IN2P3-CC (mainly simulation)
- **today: CNAF is the main computing center**
  - + IN2P3-CC (as in DS-50)
  - + opportunistic resources in Manchester
- DS-20k: requested funding to extend beyond CNAF
  - grid-like model



Current pledges

- one user interface
  - ui-darks.cr.cnaf.infn.it
- CNAF HTC: 6500 HS06
  - used via batch system (ad-hoc scripts)
- CNAF disk: 3150 TB
  - simulation + user (for all DS projects)
- CNAF tape: 1920 TB
  - datasets from all DS projects (dominated by DS-50 raw data)
- INFN Cloud: 1200 HS06, 200 TB
  - for current- and DS-20k-services

	Y0
CPU [HS06]	7700
Disco [TB]	3350
Tape [TB]	1920
HPC [Mcore*hours]	0

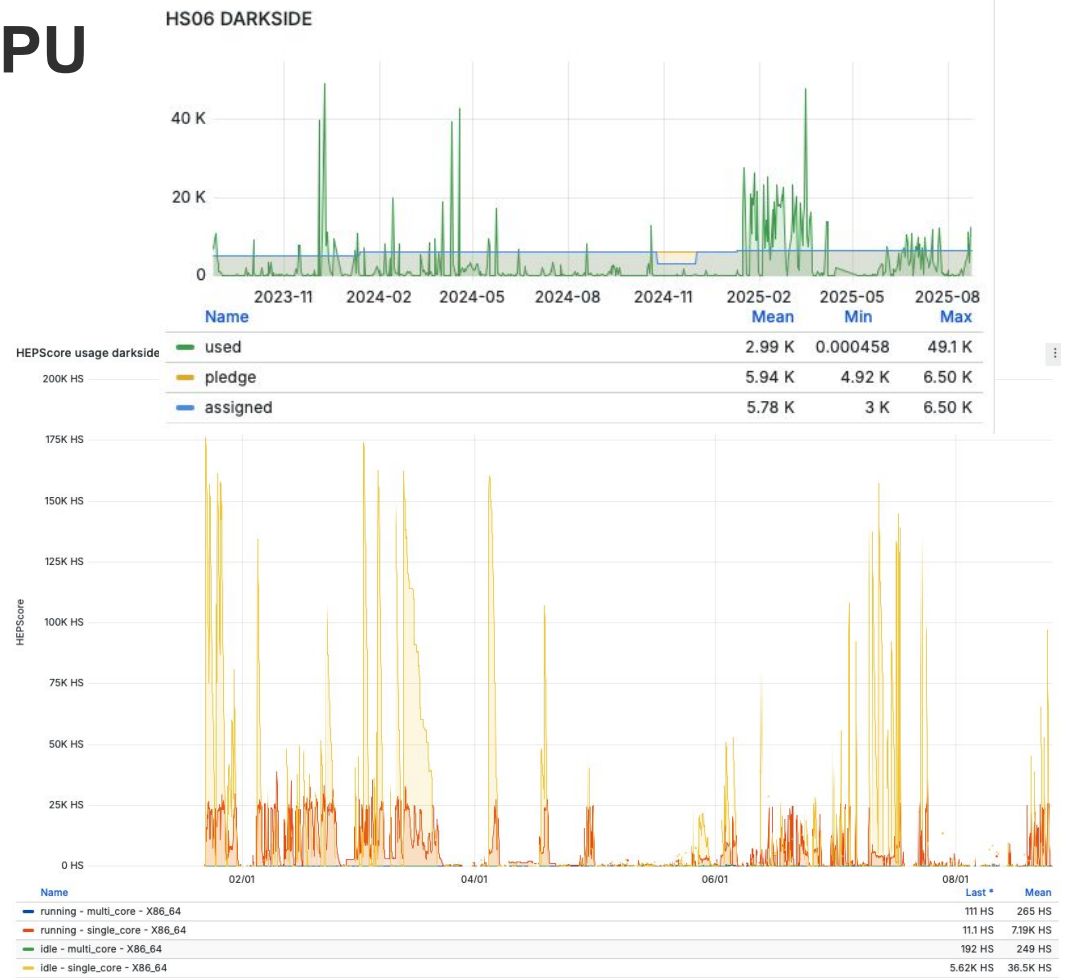
- IN2P3 provides in-kind about 300 HTC cores, used mainly for simulation
- pilot allocation in Manchester (from Oct 2025): 90 HTC cores (4 GB), 400 TB disk, 230 TB tape
- similar funding requests placed in other countries

Ongoing software/computing activities

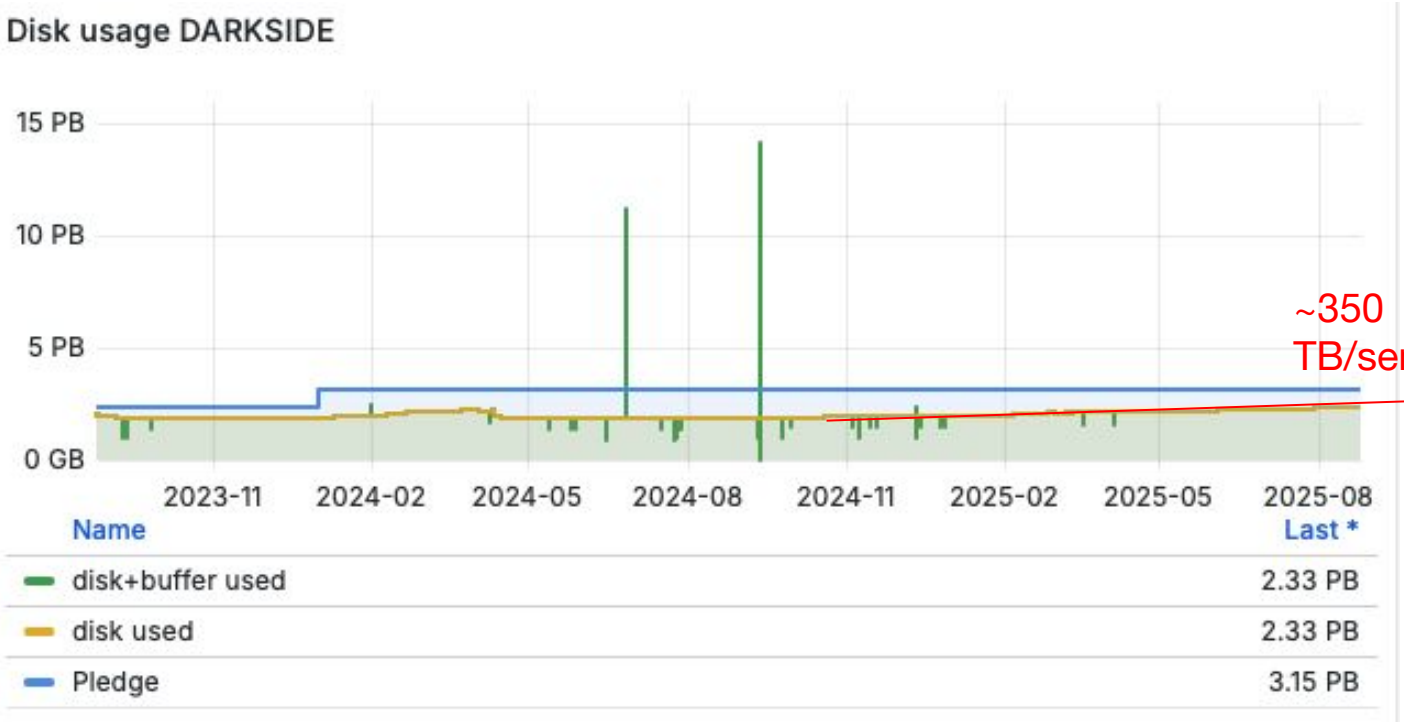
Project	Activity	Where
detector design	simulation	CNAF
core software	development	CNAF
DS-20k	Monte Carlo	CNAF, IN2P3
	reconstruction	CNAF
	data management, workload management	INFN-Cloud
	databases	INFN-Cloud
DS-50	storage	CNAF (, FNAL)
	data analysis	CNAF
ReD	storage, data analysis	CNAF
DS-Proto0	storage, data analysis	CNAF
photodetector characterisation	storage, data analysis	CNAF
	databases	CNAF

Last year's experience: CPU

- until 2024, CPU usage mostly in "bursts"
  - which correspond to different simulation campaigns
- more intensive since early 2025
  - higher precision needed for physics perspective studies
  - stress-test of reconstruction software for computing TDR ("data challenges")
  - analysis of DS-Proto0 data
  - increase of user base

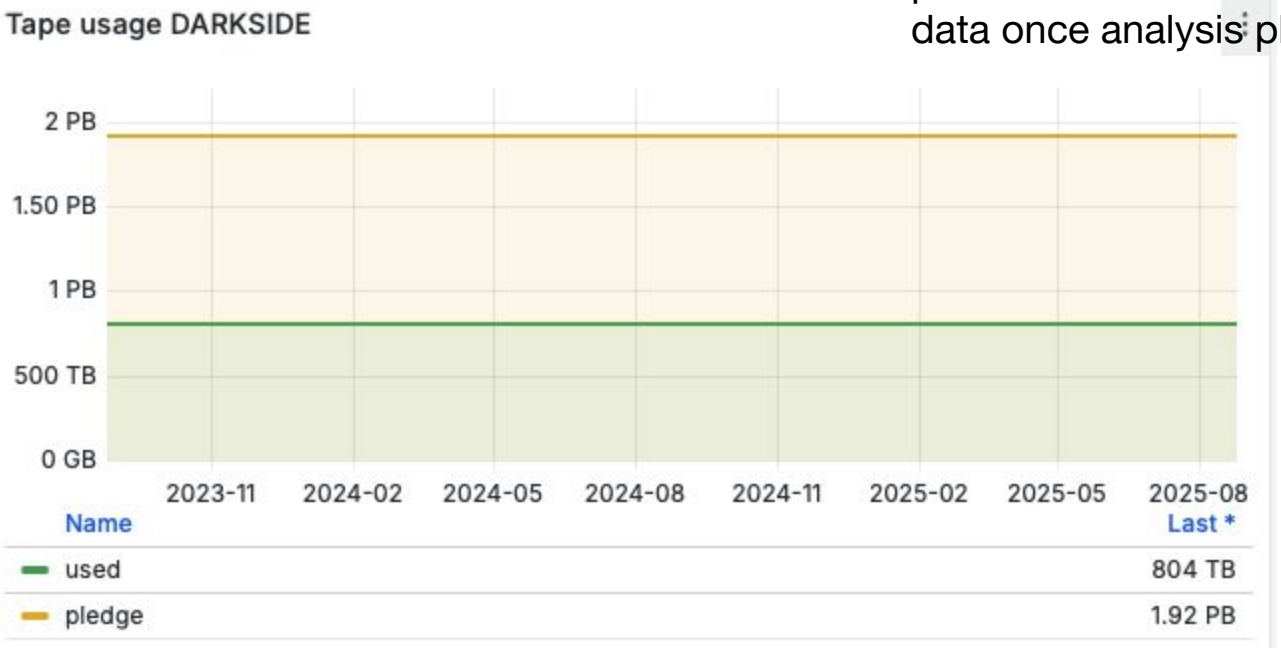


Last year’s experience: disk



Last year’s experience: tape

- dominated by DS-50 data
- will move there ReD / Proto-0 / photodetector characterisation data once analysis phase is over



Last year's experience: INFN-Cloud

- used for test deployments of grid middleware for DS-20k
- 1200 HS06 in 2025
  - kubernetes cluster for pandaWMS instance (highest resource usage)
  - software release build (docker/apptainer) and CVMFS publisher node
    - NB: actual CVMFS deployment is on WP1 DataCloud (i.e. not accounted for in requests)
  - prototype calibration database (in usage for DS-Proto0)
  - data management "user interface"
    - NB: actual resources (rucio) deployed on the WP6 kubernetes cluster (i.e. not accounted for in requests)
      - same for WP1's for the rucio database instance

Next year?

- **1000 TB** disk increase at CNAF
  - extrapolating the usage increase observed in 2025 (simulation, reconstruction, data challenge) - **+700 TB**
  - **+500 TB** data from DS-Proto0 and photodetector characterisation facilities
  - **-200 TB** from moving first DS-Proto0 runs and older simulation to tape
- **3500 HS06** HTC increase at CNAF
  - +50%, based on observed usage and effective HS06 <-> cores conversion
    (which brings to updated extrapolations)
- tape at CNAF unvaried

	Y0	Y0+1
CPU [HS06]	7700	15000
Disco [TB]	3350	4450
Tape [TB]	1920	1920
HPC [Mcore*hours]	0	

	CPU			DISCO		TAPE		Infrastruttura (Tier1, Cloud,...)
	HS06	kEuro	Range temporale mesi	TB	kEuro	TB	kEuro	
INCREMENTO	+3500	+35		+1000	+100	0	+0	CNAF Tier-1
PLEDGE RICHIESTO (attuale + incremento)	10000	100	12	4150	415	1920	19.2	CNAF Tier-1
INCREMENTO	+3800	+38	12	+100	+10			INFN-Cloud
PLEDGE RICHIESTO (attuale + incremento)	5000	50	12	300	30			INFN-Cloud
INCREMENTO				+100	+10			LNGS
PLEDGE RICHIESTO (attuale + incremento)				100	10			LNGS

Next year?

- **+3800 HS06** and +**200 TB** on INFN Cloud
  - consolidation of panda deployment and its integration with rucio
  - development of DS-20k database (started mid-2025)
- **funded elsewhere** (DS-20k costbook)**:** 100 TB
  disk buffer at LNGS and data transfer server
  - in the LNGS external labs, integrated with rucio

	Y0	Y0+1
CPU [HS06]	7700	15000
Disco [TB]	3350	4450
Tape [TB]	1920	1920
HPC [Mcore*hours]	0	

	CPU			DISCO		TAPE		Infrastruttura (Tier1, Cloud,...)
	HS06	kEuro	Range temporale mesi	TB	kEuro	TB	kEuro	
INCREMENTO	+3500	+35		+1000	+100	0	+0	CNAF Tier-1
PLEDGE RICHIESTO (attuale + incremento)	10000	100	12	4150	415	1920	19.2	CNAF Tier-1
INCREMENTO	+3800	+38	12	+100	+10			INFN-Cloud
PLEDGE RICHIESTO (attuale + incremento)	5000	50	12	300	30			INFN-Cloud
INCREMENTO				+100	+10			LNGS
PLEDGE RICHIESTO (attuale + incremento)				100	10			LNGS

Issues for the short term

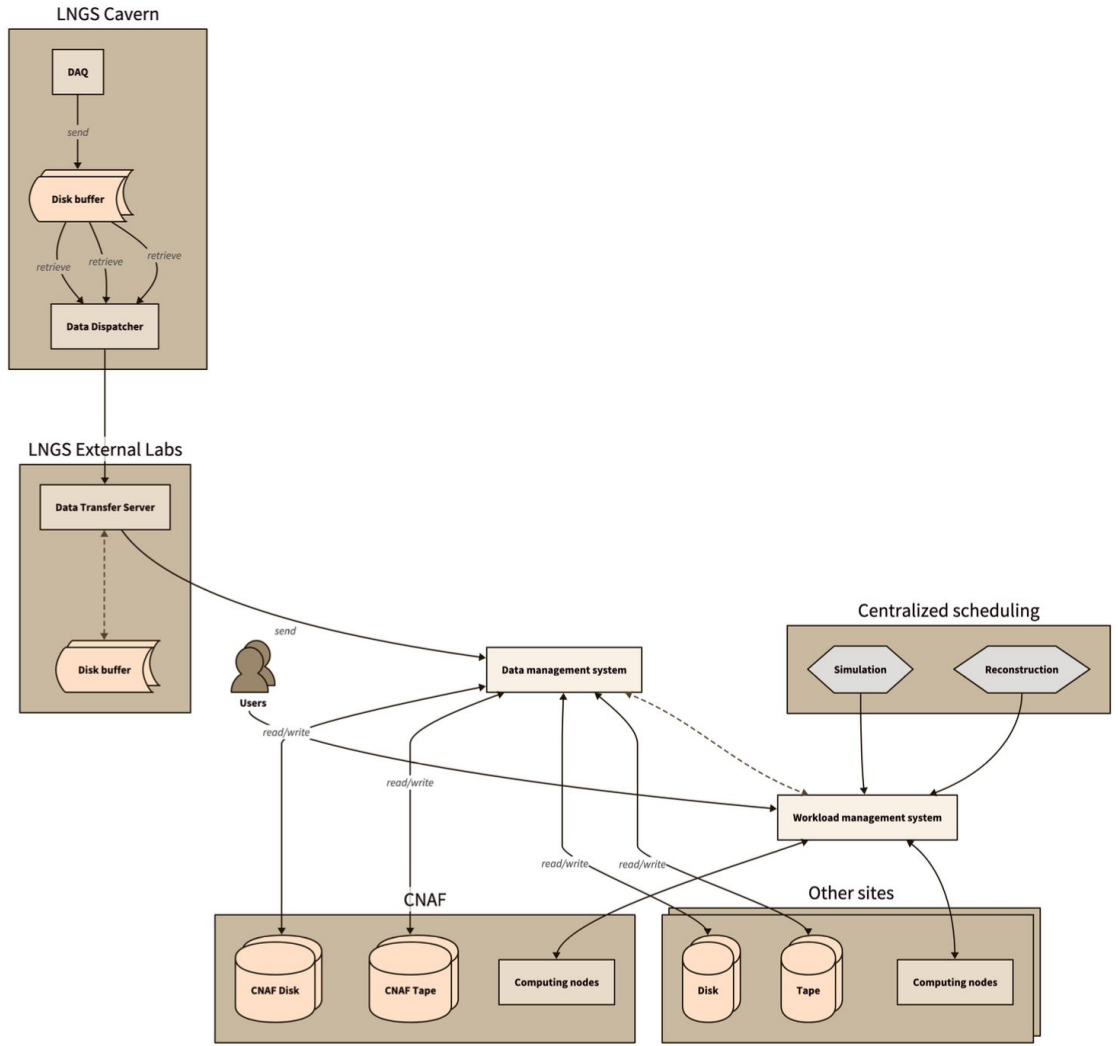
- user interface (UI) starts being insufficient
  - 4 processors, 8 GB
  - too many concurrent users [e.g. 5 users on August 25]
    - P(one runs interactively and eats up all resources) is high
    - a few instances in 2025 where ssh access was broken
  - younger people reasonably like vscode - which halts the machine soon
  - solutions (besides discouraging interactive usage): load-balancing? larger UI? autokill?
- is testing tape carousel a possibility?
  - requested in TDR review, to reduce foreseen disk allocation for DS-20k
    - the challenge here is to measure the amount of the 20 PB total we could park on tape while still being able to reprocess them ~once a year
  - in parallel, we are exploring the possibility to do this at IN2P3 CC

Computing tomorrow

a glimpse at the DS-20k TDR

A map for DS-20k

- main challenge: significant scaling up w.r.t. DS-50
  - quantitative and qualitative
- storage buffer @ LNGS
  - underground + CED
  - ~1 week (100 TB) each
  - mitigate risk of network outage
- grid-like model for storage and processing
  - CNAF Tier-1 + other sites
    - 1 full INFN replica on disk and tape
    - another distributed elsewhere
  - central reconstruction/simulation
  - users analyse data via workload-management system (WMS)
  - users access data via data-management system (DMS)



Main requirements for DS-20k computing

- support analysis for ~5 years after experiment completion
- storage redundancy (disk vs tape, geographical)
- sustain 60 MB/s raw data output, peaks of 200 MB/s
  - 2 PB/year for 10 years
- redundant buffers against network outages
  - main critical items: underground-to-LNGS, LNGS-to-internet
- support data access control at all stages
  - and blinding!
- multi-site computing, at CNAF and elsewhere, as transparent for the user as possible
  - middleware should support integration of computing resources as they become available
- system sized to allow few months of turnaround for delivery of physics results
  - centralize production and “event skimming” to reduce costs

The current analysis model

- RAW data are hits and consolidated information from waveform segments from the O(10k) SiPM channels
  - organised in **time slices** corresponding to a given time interval (~1 s)
- reconstruction first identifies **pulses**, and then combines pulses into **events**
  - e.g.: WIMP should give one S1 pulse followed by one S2 pulse within t_{drift}
  - this association is often delicate
- current model assumes hit information is not propagated to higher-level formats
  - e.g. DATA_EVENT is a TTree organised per event, with pulse information but no info on what was seen by channel #4553
  - this makes reconstructed data lighter, but requires access to RAW data for reprocessings
- to minimize CPU usage, users access DATA_EVENT - reprocessings are only performed centrally
  - but still they need to be performed (new algorithms, bug fixes…)
  - our sizing of necessary resources is based on a target turn-around
    - few days in the commissioning phase, few months in year 10+

Expected longer-term impact on resources at CNAF

- in the TDR we propose to extend computing beyond CNAF
  - interest in other countries at providing computing resources (hardware & peopleware)
  - CNAF would still provide core resources (1:1 = CNAF:elsewhere)
- CPU need can be projected to be almost flat after commissioning
- disk need increases 2 PB/year
  - tape carousel being explored as an alternative for older data

in the TDR we foresee this allocation to be reviewed by RRB every few years

	Y0	Y0+1	Y0+2	Y0+3	Y0+4
CPU [HS06]	7700	15000	20000	30000	30000
Disco [TB]	3350	4450	6000	9350	11350
Tape [TB]	1920	1920	6000	7920	9920
HPC [Mcore*hours]	0				

Main items and challenges / 1

- join LHCOne to exploit higher bandwidth from LNGS to CNAF
  - upgrade of connection from 10 Gbps to 100 Gbps foreseen in the next few years
- use WLCG tools to ease (*en-passant*-) integration of additional resources
  - **rucio** for data management
    - successful deployment thanks to WP6 team
    - challenge: make it a stable solution for the next ~15 years
  - workload management
    - no one-size-fits-all solution: considered **pandaWMS** and DIRAC
      - adapted panda to DS-20k use case with support from core developers
        - mostly successful, few items to be ironed out
      - clearly a challenge - “WMS-as-a-service” would be the best option…
  - software distribution via **CVMFS**
    - successful deployment on DataCloud CVMFS instance
  - IAM based authN/authZ
    - experiment IAM deployed by CNAF

Main items and challenges / 2

- deployment of services on INFN-Cloud
  - extremely valuable asset from our point of view (thanks!!!)
    - **XXX-as-a-service** helps experiments (like DS-20k) where computing person power and expertise are naturally more limited than at LHC
  - we think support from INFN on **transition from "pilot" to "production"** stages would be precious for the success of the experiment
    - example: deployment of database on cloud resources would require **high-availability**
    - accounting (resources, technical support…) should reflect this
  - brings a strategy question for INFN!

Backup slides

which I will probably remove before converting to PDF

Data acquisition (DAQ)

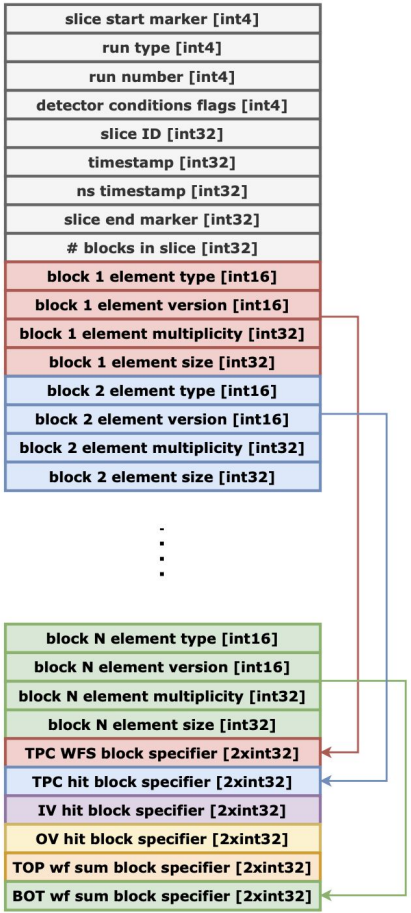
- cannot sustain rate in scope mode: SiPM waveforms are preprocessed by DAQ
  - uncommon for direct detection searches
- raw data := hits + waveform segments (WFS)
- design rate in “physics mode”: 60 MB/s
  - up to 200 MB/s during commissioning/calibration

data taking mode	**rate**	**notes**
physics	100 Hz in the TPC	DM search
laser calibration	100 Hz external trigger	~daily
random trigger	100 Hz external trigger	unbiased measurement of noise and hit-finding efficiency
calibration	100 Hz + source rate	radioactive sources (in calibration system or dissolved in LAr)
random waveform	<10 Hz	unprocessed waveform, to monitor preprocessing performance

Raw data format

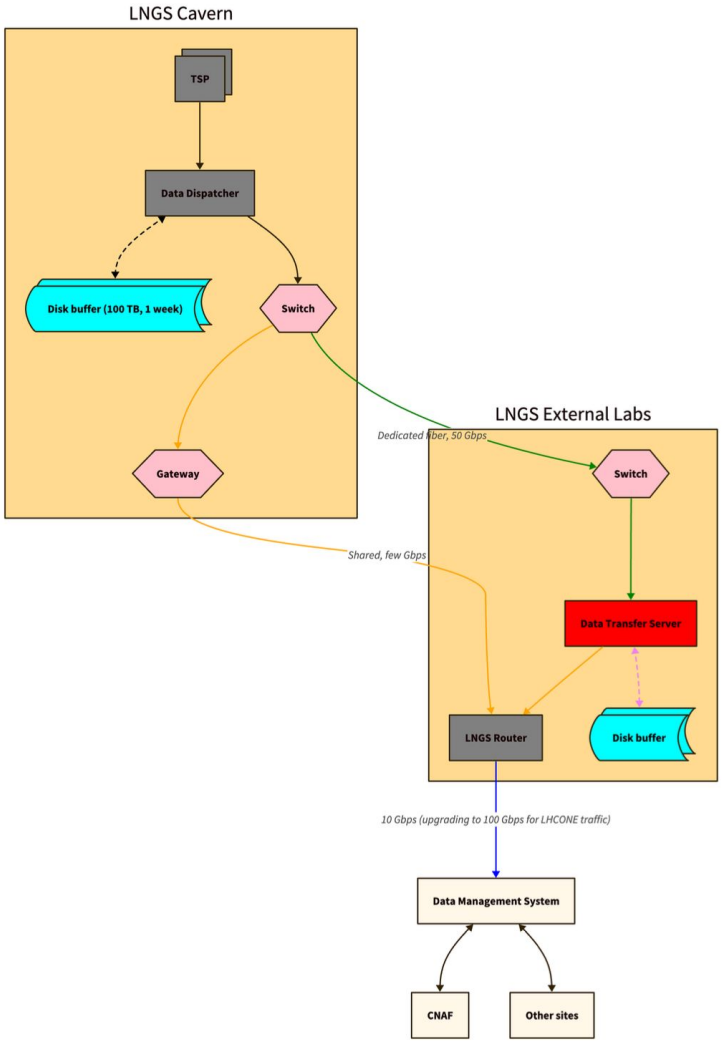
- waveform digitisers (WFDs) “prepare” raw data
  - hits and WFS
  - from groups of channels
- time-slice processors (TSPs) merge channel information together
  - 1 s slices, overlapping in time
  - follow expandable event data model (EDM), custom binary
- data merger sorts them and writes RAW data files on disk
- monitoring and possibly fast online analysis also happen
  - at different stages (merger, TSPs…)

this is when offline computing starts



Data flow and data management / 1

- two lines connecting underground DAQ room (in LNGS Hall-C) to the world
  - 50 Gbps optical fiber for experiment data, dedicated to DS-20k
  - few Gbps shared fiber for service tasks (e.g. software updates, remote shifts)
- data transfer server bridges experiment network and LNGS network
  - ultimately connected to internet via GARR
- funding (O(50k€)) to be requested to INFN this year
  - Hall-C fiber + transceivers + switch + data transfer server
- in this picture, bottleneck is the LNGS-GARR connection
  - 10 Gbps, to be upgraded to 100 Gbps
  - in parallel: associating DS-20k to LHCOne (as e.g. Xenon) to exploit dedicated bandwidth



Data flow and data management / 2

- 1 run (~6h) = 1.2 TB
  - ~15’000 physics runs total over 10 y
- 130 files (MIDAS’ “subruns”) per run
  - 10 GB each
    - determined (and configurable) at the level of the data merger
    - actual file size to be optimised based on data management pilots
      - stress-test rucio+FTS
        - first tests to start soon involving LNGS-CNAF and CNAF-Manchester transfer
- RAW files are organised per time slice
- promptly processed (and reprocessed) to reconstruct *events*

Analysis model

- reconstruction takes *time slices*, identifies *pulses* and reconstructs *events*
  - current implementation allows splitting into intermediate steps
- key difference w.r.t. LHC: hit information is kept in RAW data alone, and must be accessed for event reconstruction
  - i.e. intermediate (“PULSE”) data formats are not an augmentation of the RAW data format
    - the equivalent of ATLAS’ ESD would be larger than the RAW data
- this means RAW data should be available on disk (which is the cost driver)
  - TAPE acts only as a backup
- reconstruction scheduled centrally, includes skimming to pre-select events
  - to be used for blinding and for end-user analysis
    - exact application of blinding requires more thought and extensive testing
  - end-user on Y10 will analyse a fraction (~10%) of the 70 TB of reconstructed data
- use same computing resources for reconstruction and user analysis
  - priority to central scheduling
- see T. Hessel’s presentation

The problem of end-user analysis

- full 10Y dataset in EVENT format: ~70 TB
  - skimming may reduce them to ~7 TB of interesting data, and users run on them in batch via the WMS
    - it should be possible to go back to RAW data information (e.g. hits) from this “skimmed” data format
- this brings the challenge of supporting end-user analysis
  - basic option: “Tier-3 equivalent” (access to user interface at CNAF or local institution clusters)
    - recent tests of Jupyter over CNAF batch system
      - direct and dask-based access to datasets
    - does not scale well
      - today, single darkside-ui@cnaf with $N_{cores} < N_{users}$
  - HSF efforts may bring better solutions on the market
- architecture / pilots to be developed in the coming future

Authentication and authorisation

- ideal model: IAM-based authentication and authorization
  - DS-50: legacy DS VOMS, used mainly for data transfers by few users
  - DS-20k: Indigo-IAM instance setup by CNAF, to be used by “everyone” for “everything”
    - integrated with EduGAIN
    - three different groups (enabling user/production/blind data access)
      - more on this in the data management part
- current status: IAM for computing, but “collaboration tools” still require separate accounts
  - LNGS account for critical resources within LNGS VPN
  - INFN-AAI account for access to agendas, CNAF, and other services
    - problem: the INFN policy does not allow all collaborators to get an account
      - in contact with INFN responsibles to sort these issues out, but will take time

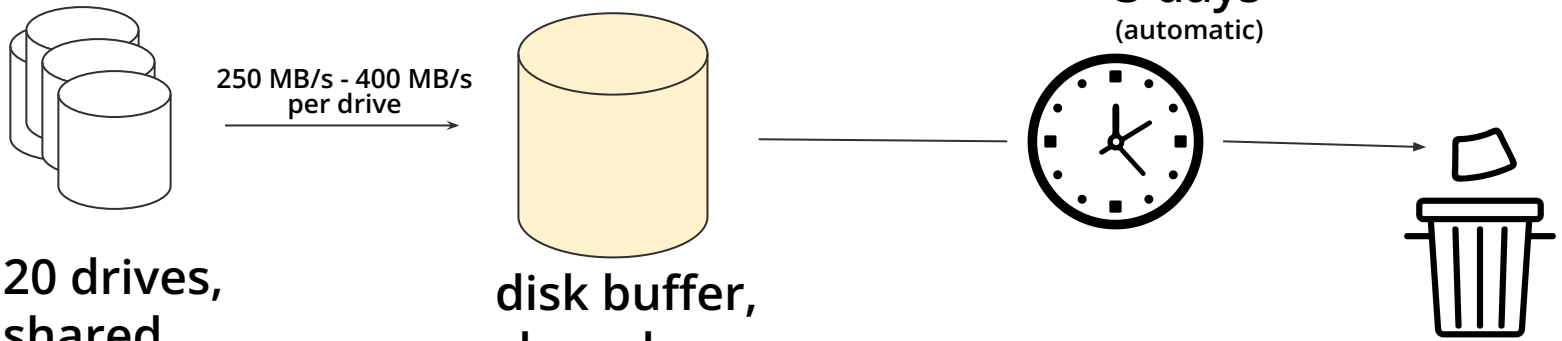
Data management system

- specific requirement: DMS should enable blinding data
  - “embargo”
- we chose rucio as our DMS
  - pilot deployed successfully with INFN Cloud support
- embargo is not supported directly: delegated to storage configuration
  - IAM releases token with requested storage scope when authorised
    - implemented via IAM scope policy, tested successfully
  - dataset-level permission restrictions within rucio will be supported in a future version of rucio
    - would require syncronised mapping of IAM users
    - in contact with rucio developers
- additional layer of complication brought by how panda uses rucio
  - user token vs service account

Beyond standard CPUs

- software development uses state-of-the-art languages and tools
  - heavily python-based, using e.g. numpy, scipy
- this should simplify the extension to new architectures
  - ARM, RISC-V…
  - containerisation should help building dedicated releases
    - panda already used in integration of ARM resources at ATLAS
- GPUs and other accelerators to improve reconstruction?
  - testing started for xyzE, underlying python tools make it easy
- GPUs for accelerating machine learning (ML)
  - ongoing development for ML at trigger level (supernovae) and reconstruction level (pulse finding / event building)
  - explorative USA/Canada funding requests placed for a dedicated machine in DAQ farm
    - will use next years to exploit possibility to extend requests also for core computing
- bottom line: potential for development, but not in the baseline strategy yet

Tape @ CNAF (VERY preliminary)



e.g. N-3 years of RAW data

~2 weeks of RAW data

~1 day to process them (or move them elsewhere)

more CPUs (700 if reco speed is 50 lives-s)

likely limitations:

- takes 1-2 months to retrieve 1Y data
- same for writing them on tape!
- limits possibility of full data reprocessing

⇨ need **cost-benefit analysis**

Valerio Ippolito
INFN Sezione di Roma