



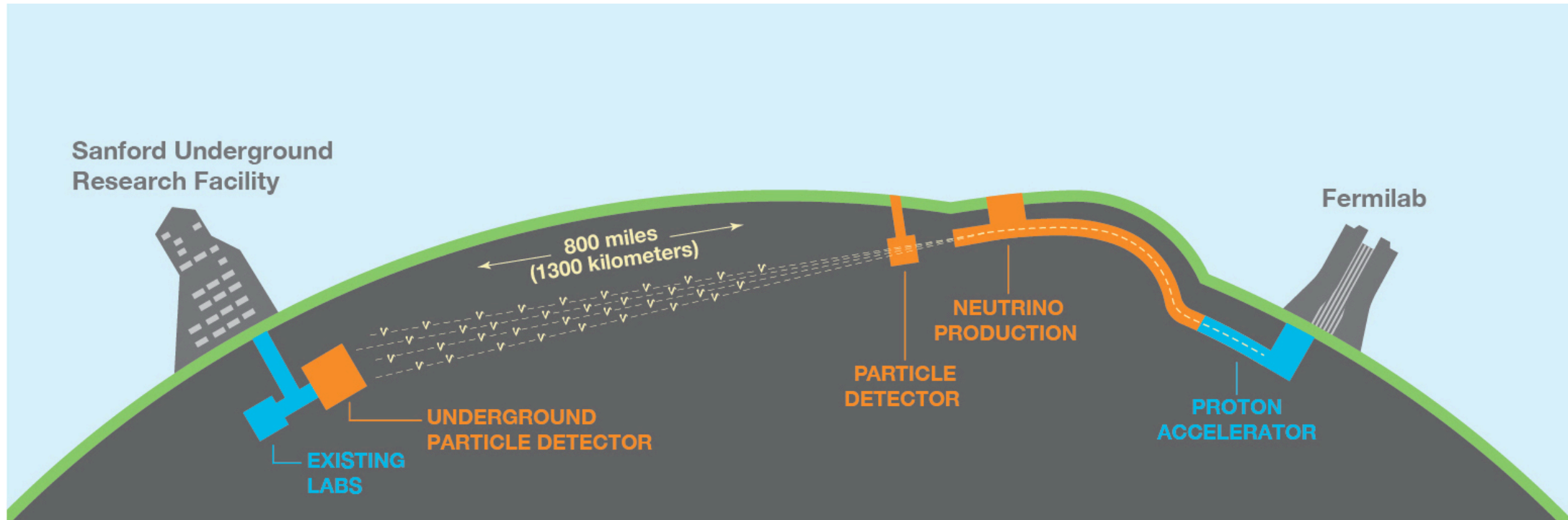
DUNE Plans for WLCG Data Challenge 2024

Mike Kirby on behalf of DUNE Computing Consortium

LHCOPN-LHCONE #50 - FZU (Prague, CZ)

19 April 2023

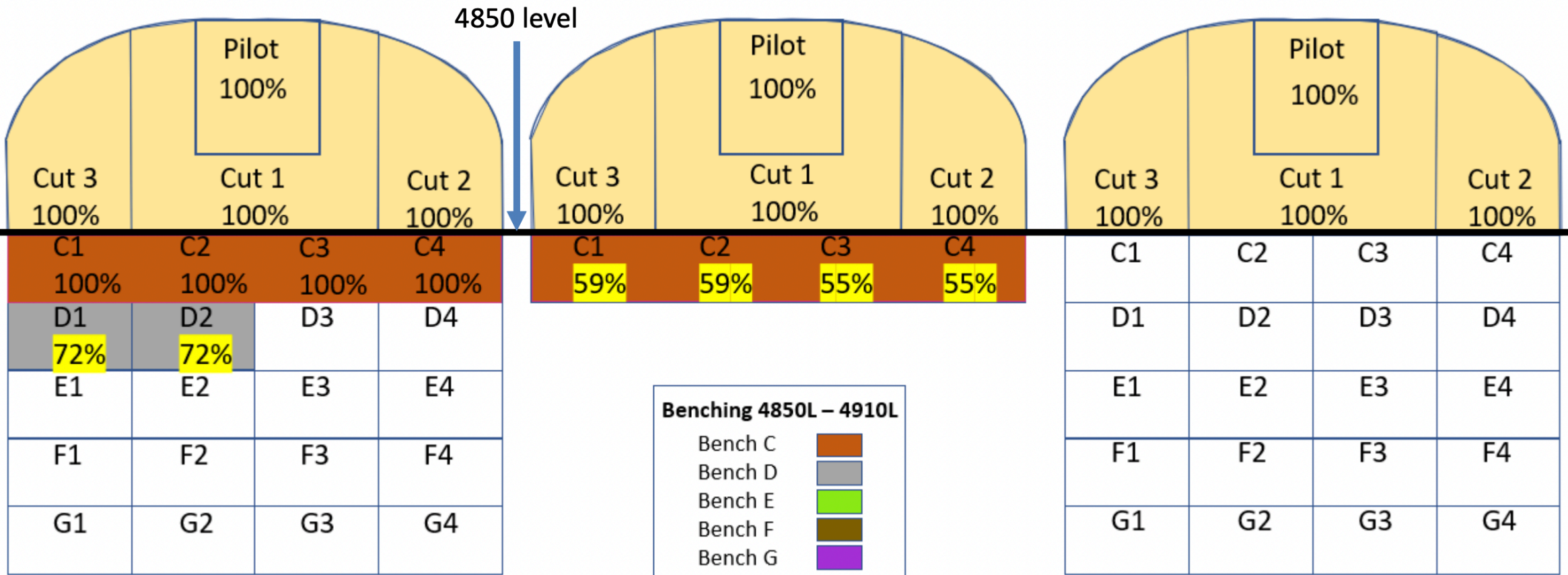
Quick reminder about DUNE



- neutrino experiment looking for neutrino oscillation parameter (mass ordering, matter vs antimatter asymmetry, unitarity), proton decay, supernova neutrinos, and more.
- four very large LAr TPC (17 kT) at 4850 ft underground in Lead, SD (Homestake Mine)
- Near Detector being designed (3 sub-detectors, two that move) at Fermilab
- Two prototypes at CERN - (ProtoDUNE II Horizontal Drift - ProtoDUNE II Vertical Drift)

Excavation continues...

Slide: C. Mossey



North Cavern

CUC Cavern

South Cavern

58.7% of in-situ rock volume removed as of 20 Mar 2023



North Detector Cavern – West End

Slide: C. Mossey



Photo by
Matt Kapust,
SDSTA; 19 Jan
2023

Drilling holes for blast charges for bench C (left) and removing muck (right) in North Detector Cavern (4850-33) west end

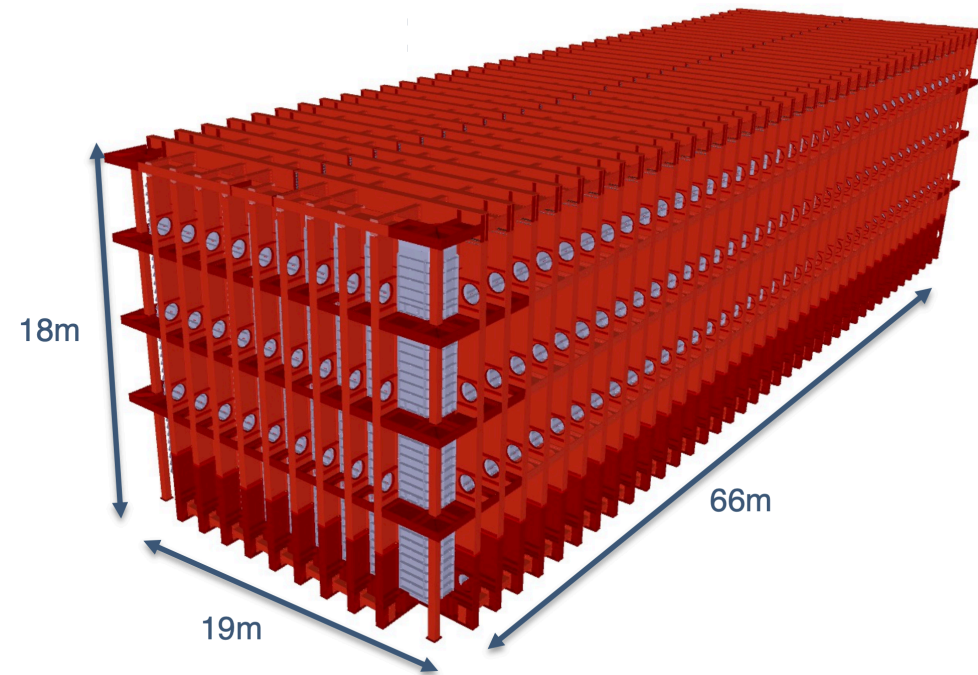
Central Utility Cavern



Data Taking Configurations

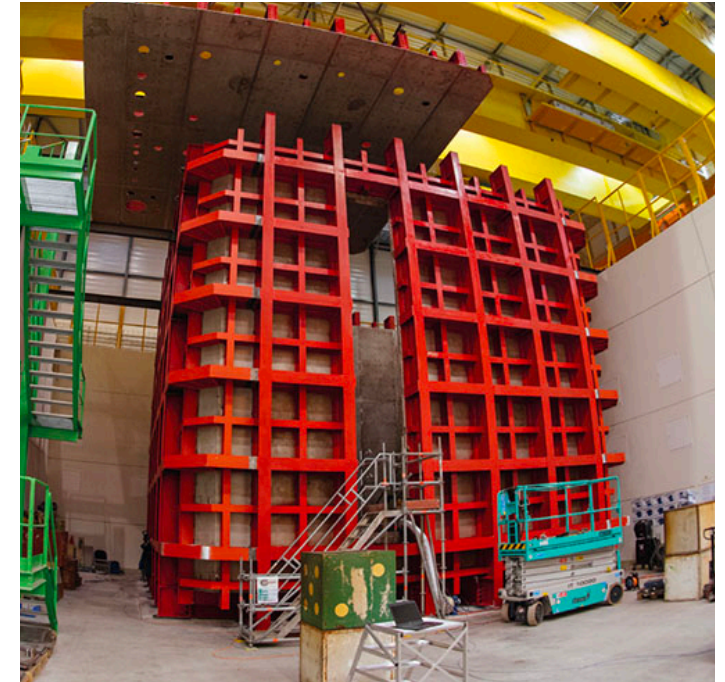
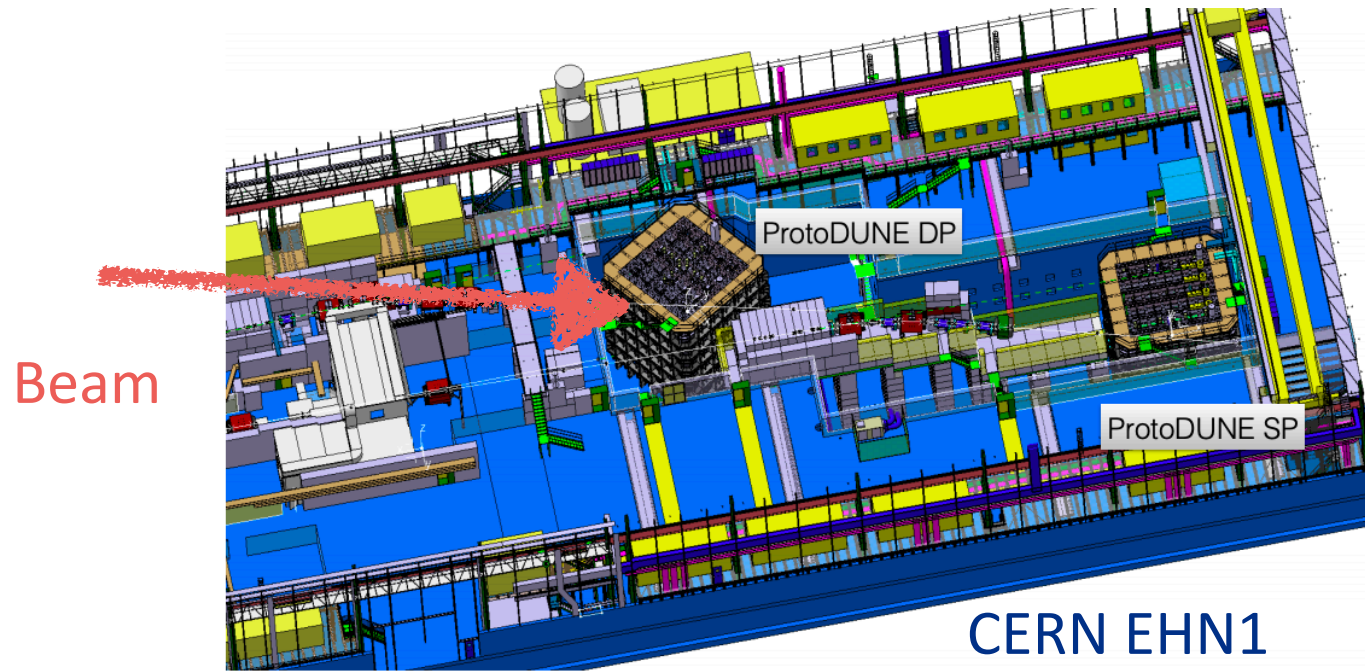
- beam coincidence events are extremely important, but of limited total volume - ~ 1 Hz beam rate
 - active online trigger in development
 - Region-of-Interest within module
 - online compression and zero-suppression being considered
- solar neutrino triggered events
- cosmic ray events and calibrations
- SuperNova readout events
 - ~ 150 TB in 100 seconds
 - transfer out in 4 hours and process in 4 hours to point for optical follow up
- DUNE requirement - less than 30 PB/year total to permanent storage from all active FDs

Far Detector Cryostat



Process	Rate/module	size/instance	size/module/year
Beam event	41/day	3.8 GB	30 TB/year
Cosmic rays	4,500/day	3.8 GB	6.2 PB/year
Supernova trigger	1/month	140 TB	1.7 PB/year
Solar neutrinos	10,000/year	35 TB/year	
Calibrations	2/year	750 TB	1.5 PB/year
Total			9.4 PB/year

CERN and Neutrino Platform currently hosting ProtoDUNE

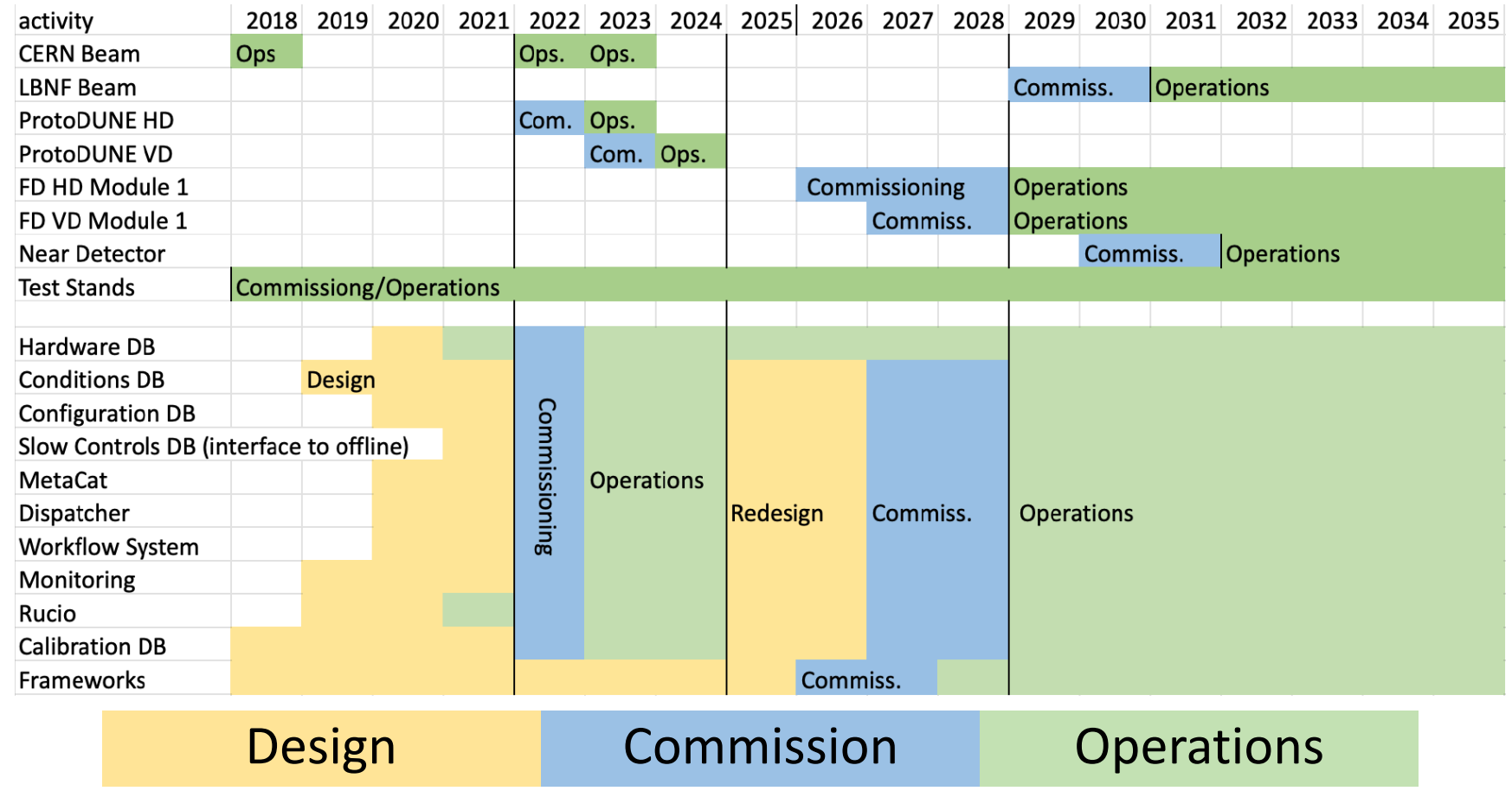


- ProtoDUNE Single Phase and Dual Phase
 - constructed and operated during the timeframe of 2018 - 2020
 - ProtoDUNE SP took 6 weeks of beam (~ 25 Hz)
 - invaluable information about performance, construction, and operations
- ProtoDUNE II currently under construction
 - Horizontal Drift and Vertical Drift
 - HD anticipate filling LAr 2023Q4
 - VD assembly will proceed in parallel and operate after HD operations
 - Beam operations in 2023/2024

tentative DUNE future timeline

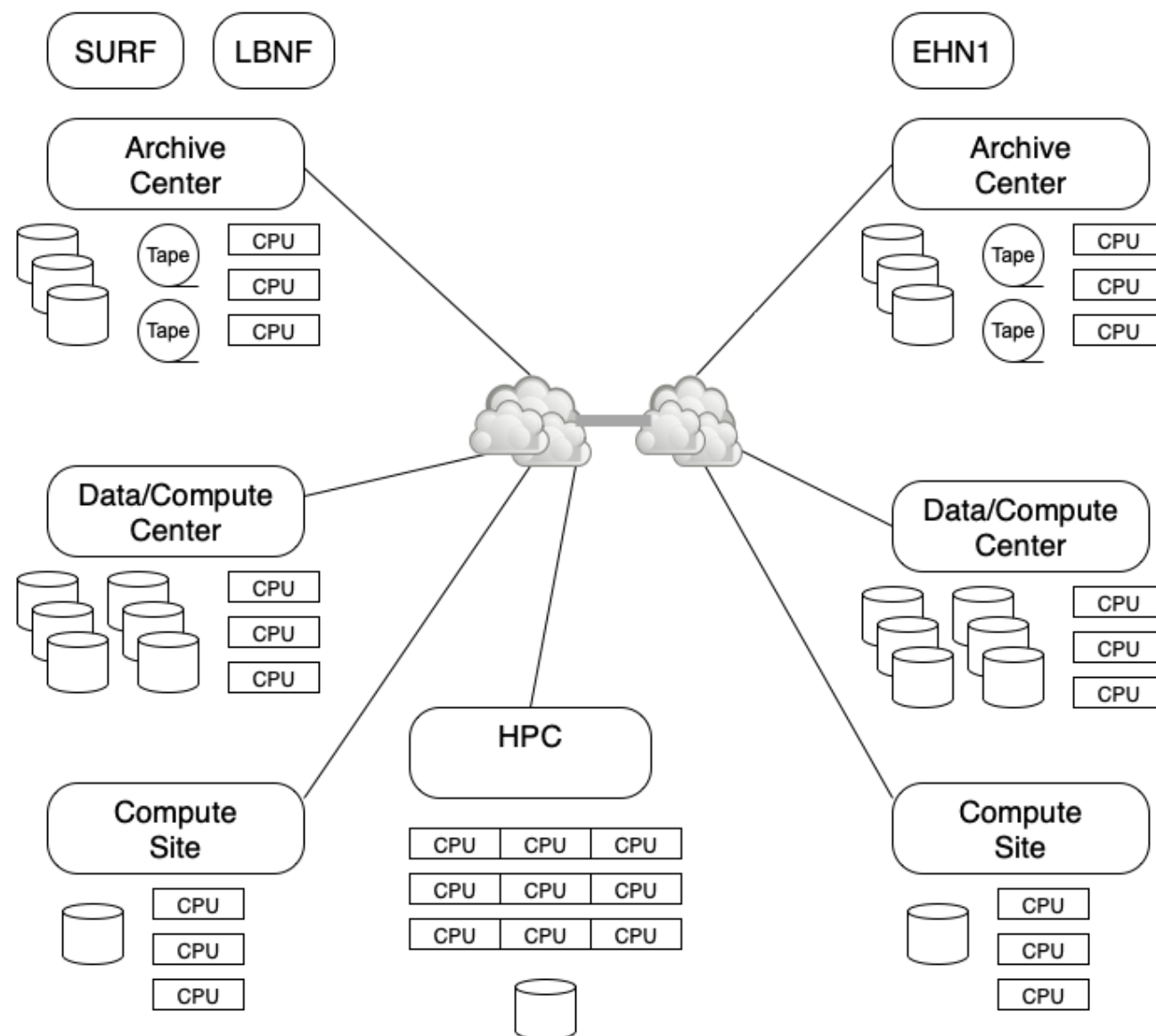
- **Fall 2023** - operations of ProtoDUNE HD
- **Spring 2024** - operations of ProtoDUNE VD
- **2023 - 2024** - DUNE computing operations at scale with PD II data
- **FD HD Module 1**
 - 2027 construction
 - 2028 commissioning
 - 2029 physics
 - FD VD Module - 1 year offset
- **2025-2027** - use this time for development addressing unique DUNE Challenges

Long Term Computing Project Schedule



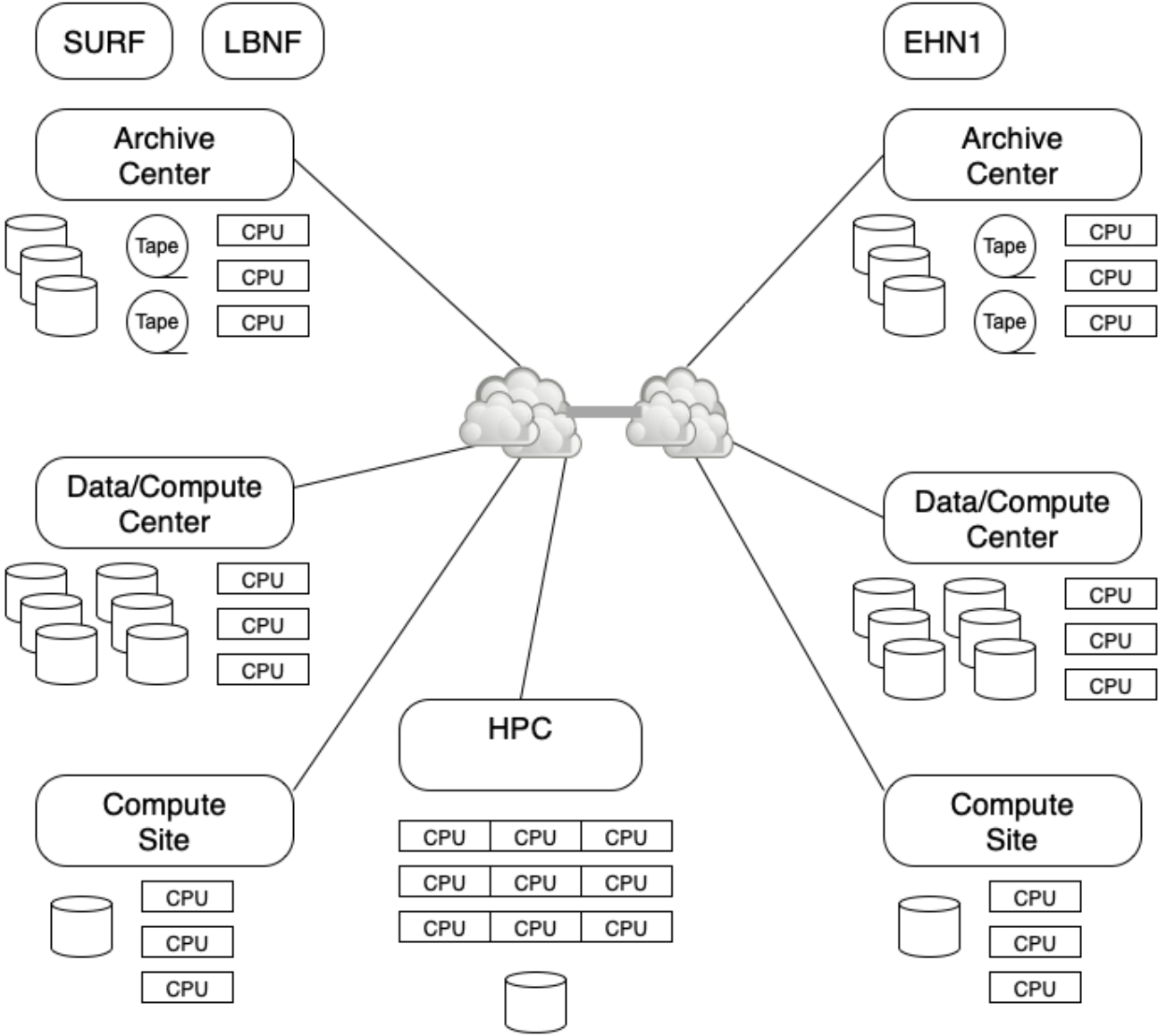
DUNE Computing Resource Model

- less “tiered” than current WLCG model → flatter model proposed by HSF DOMA working group
 - take advantage of existing WLCG sites that can add DUNE access
 - require a reasonable minimum size for storage elements
 - allow for CPU only sites with data streaming
- collaborating institutions (or groups of institutions) provide significant disk resources (~1PB chunks)
- plan to use common tools for most services
- participation in the HSF process important to provide and integrate new solutions

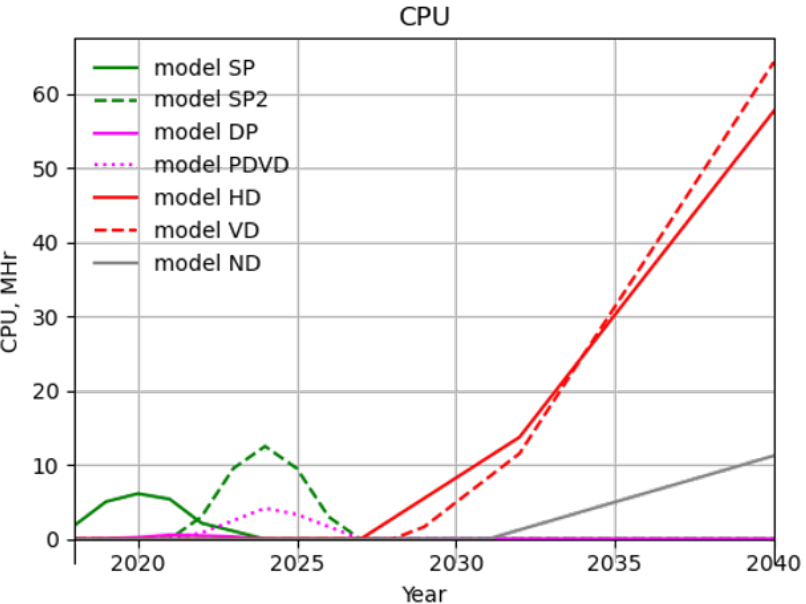
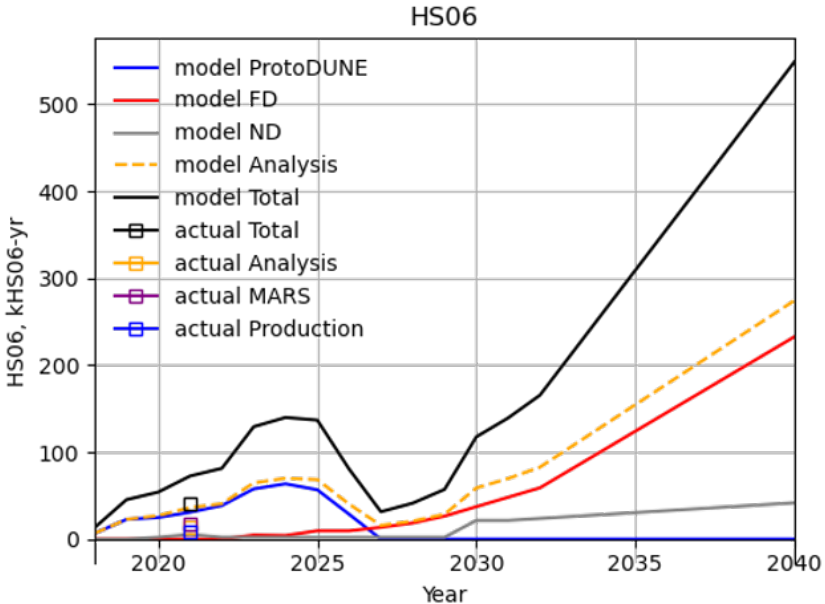
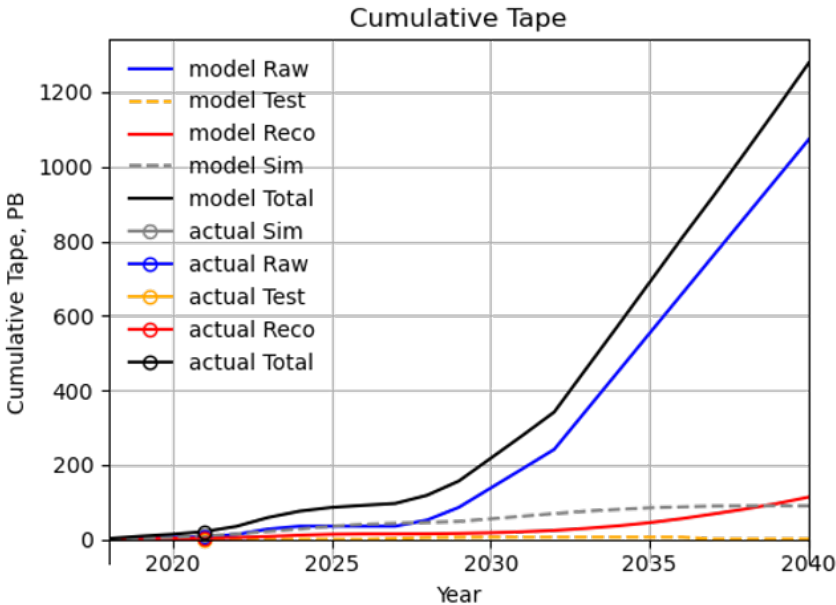


Data Placement Strategy

- 2 copies of raw data on tape (6 months on disk)
- 1 copy of “test” data stored for 6 months
- 1 copy of reco/sim on tape
 - Currently assume 1 reco pass over all data and 1 sim pass/year
 - Assume reco/sim resident on disk for 2 years
- Assume 2 disk copies of reco and sim
 - impose shorter lifetimes on tests and intermediate sim steps.



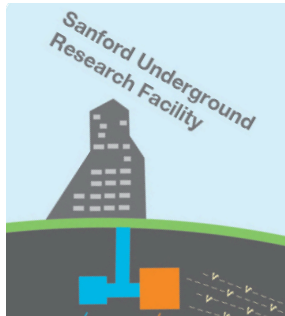
Data Volumes - through 2040



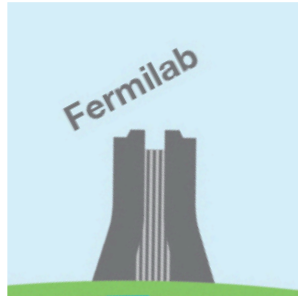
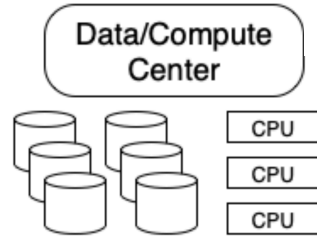
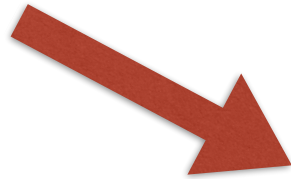
Longer term projections

FD Raw Data
archival storage

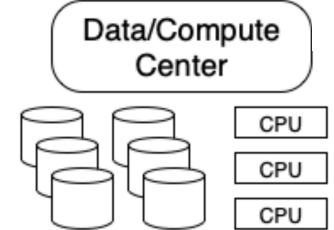
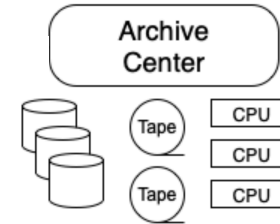
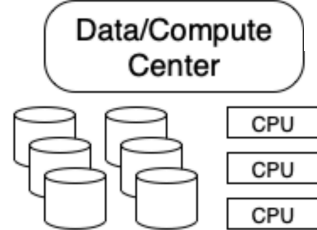
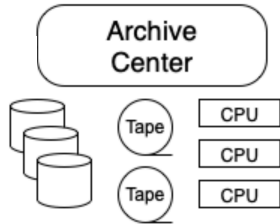
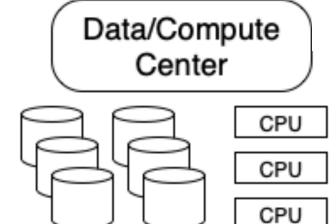
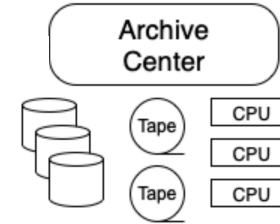
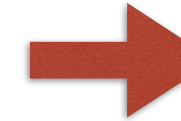
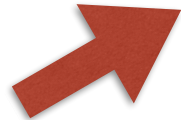
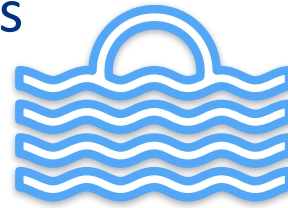
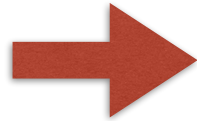
Cartoon version of data movement



8 Gbit/s



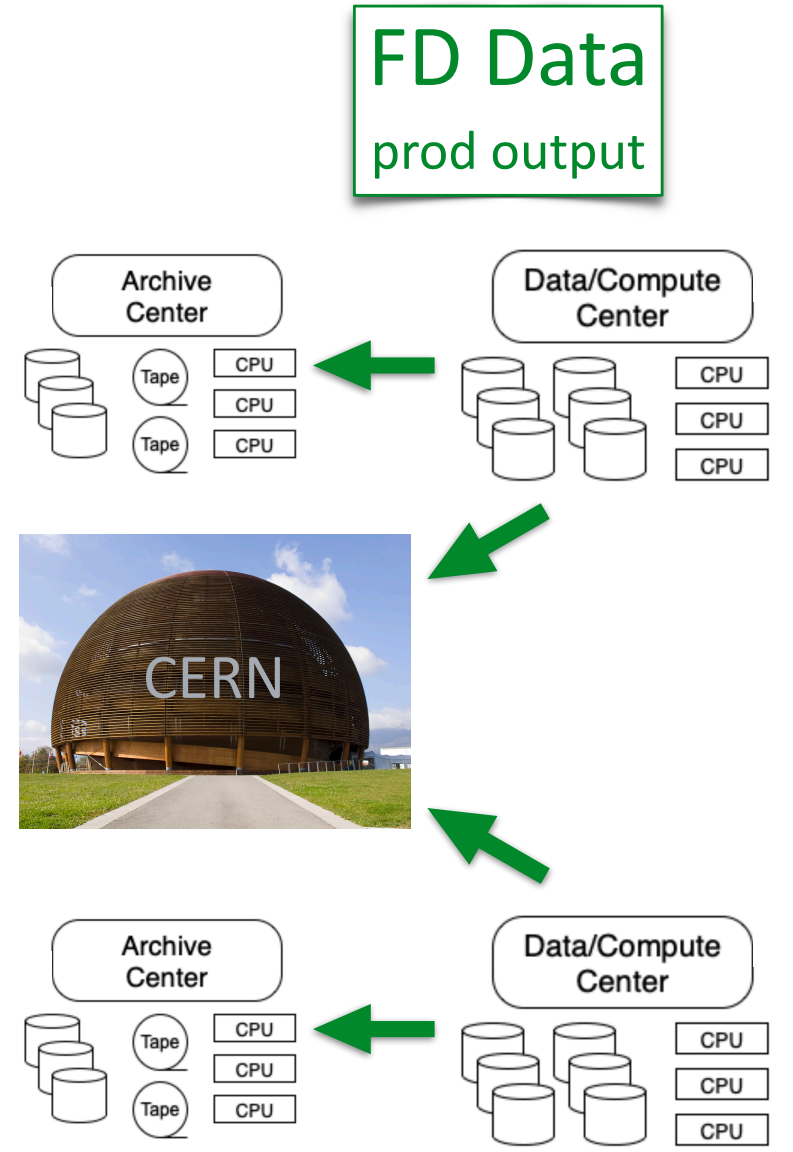
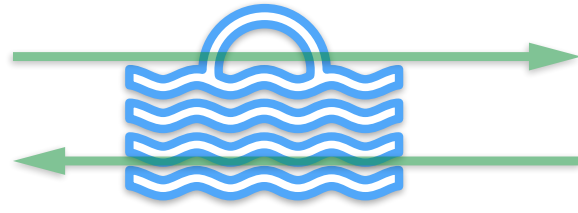
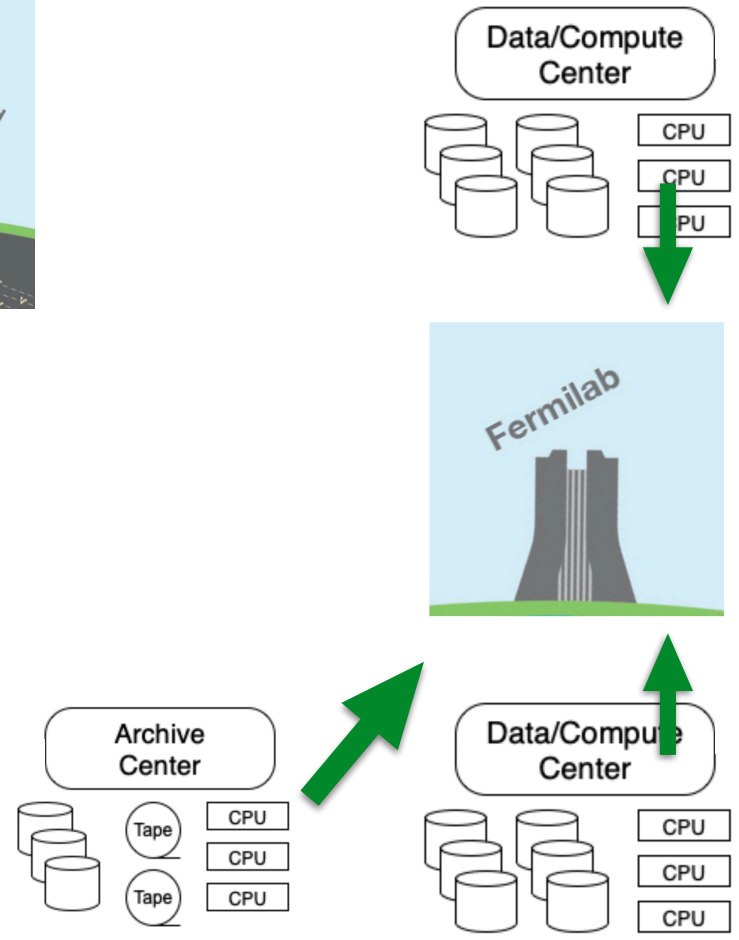
8 Gbit/s



not to scale, not a technical design
it's just a cartoon

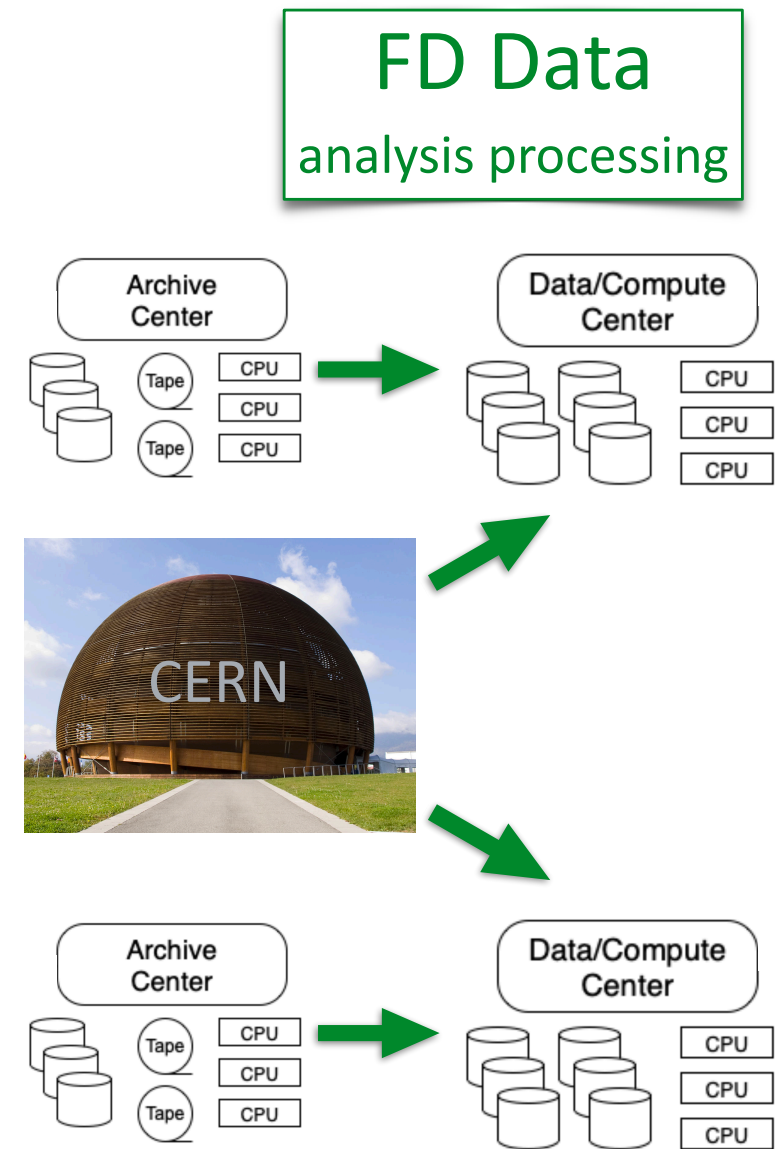
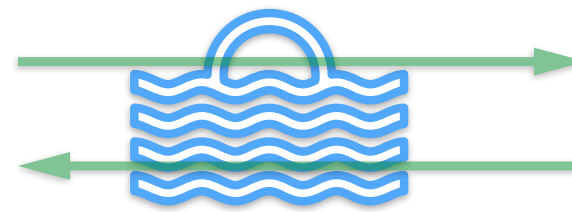
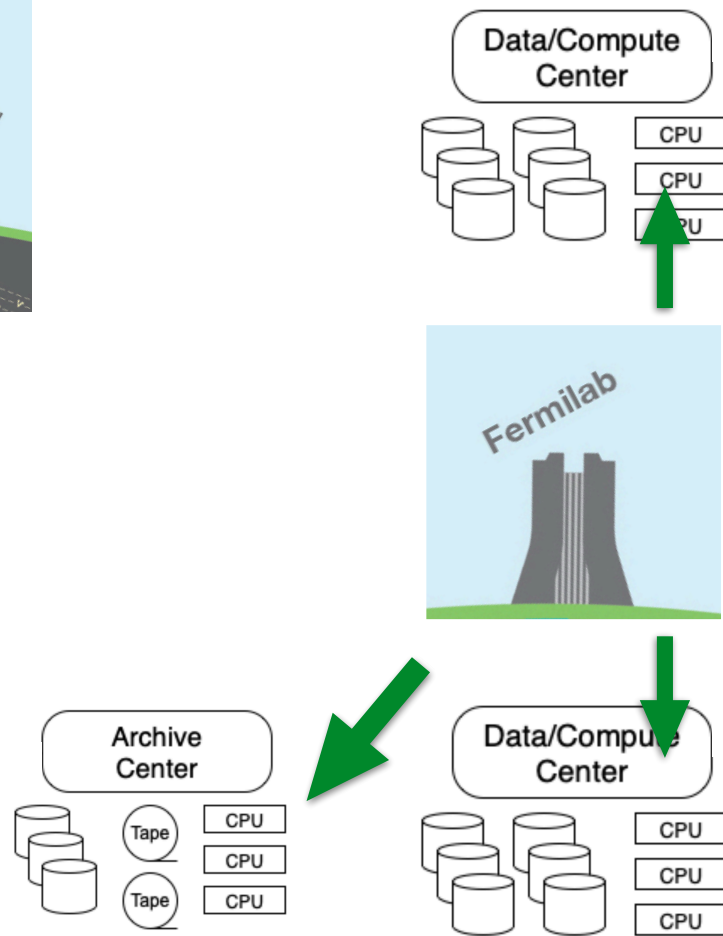
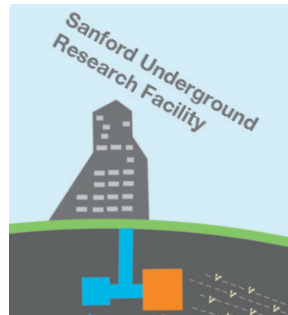
Use multiple archive sites
not in North America

Cartoon version of data movement



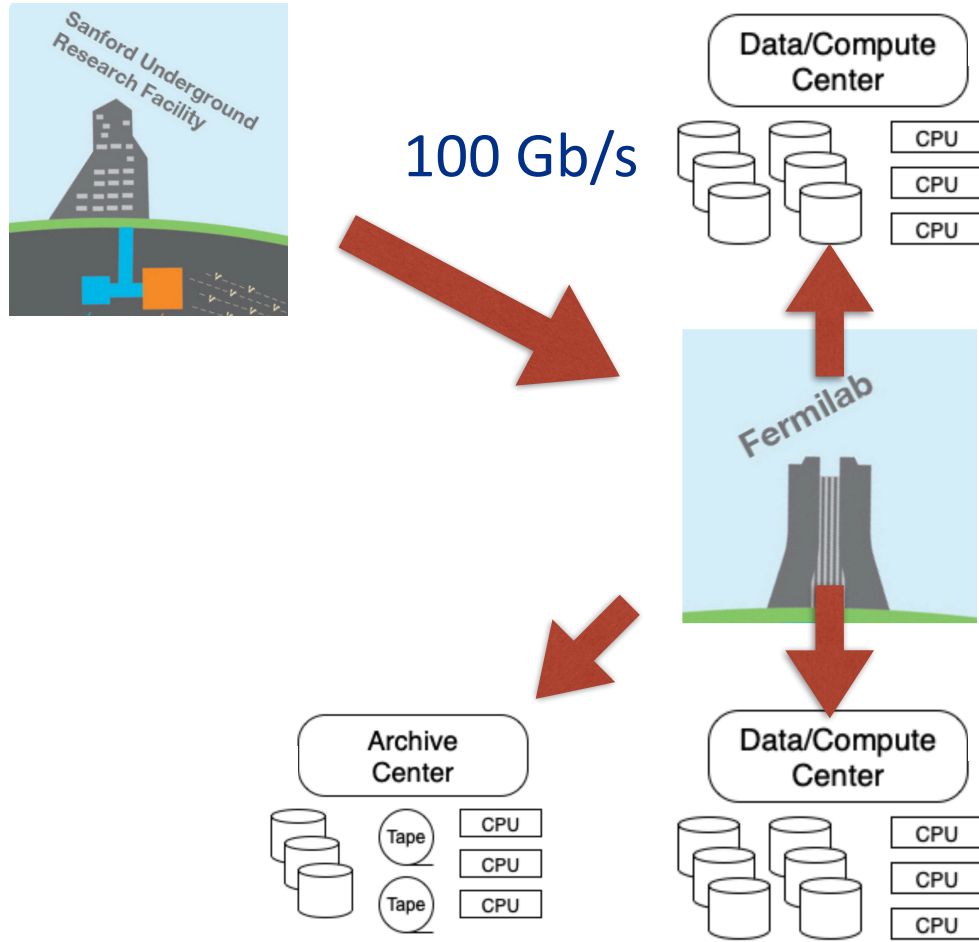
not to scale, not a technical design
it's just a cartoon

Cartoon version of data movement



not to scale, not a technical design
it's just a cartoon

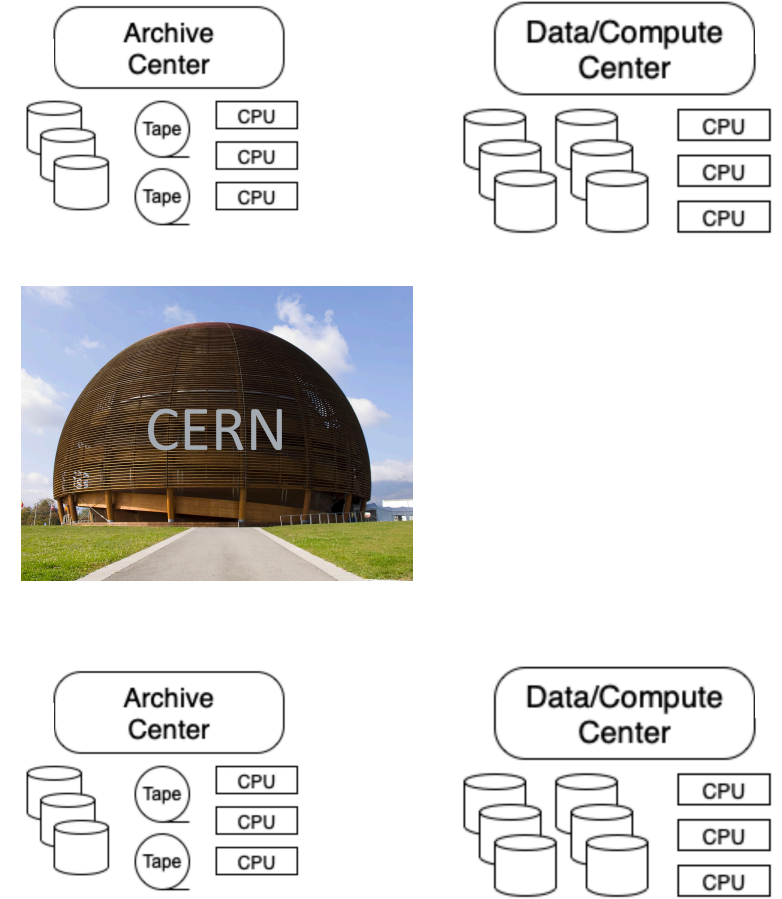
Cartoon version of data movement



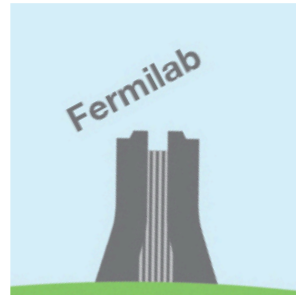
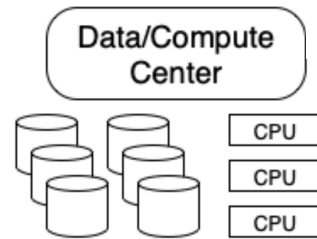
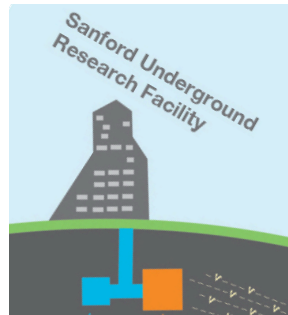
not to scale, not a technical design
it's just a cartoon

Not anticipated to be part of DC24

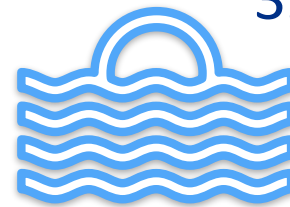
SuperNova Raw Data rapid transfer & processing



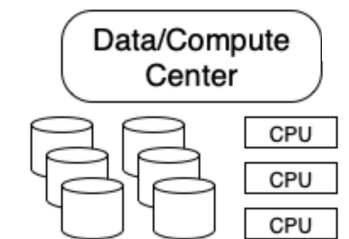
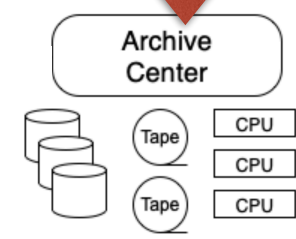
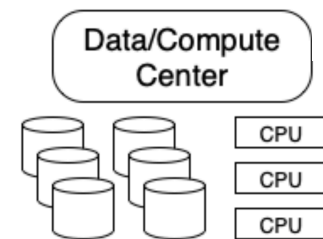
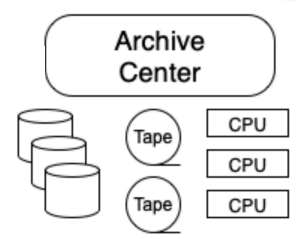
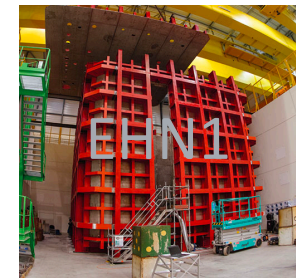
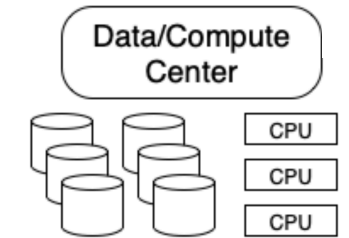
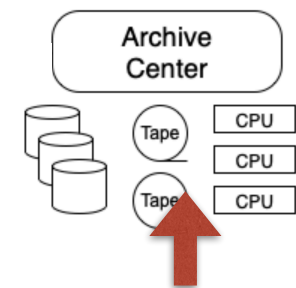
Cartoon version of data movement



32 Gbits/s

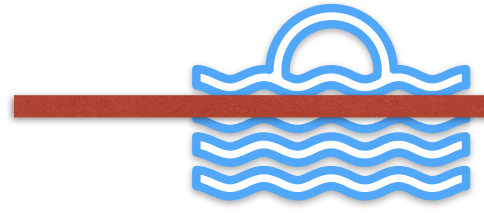
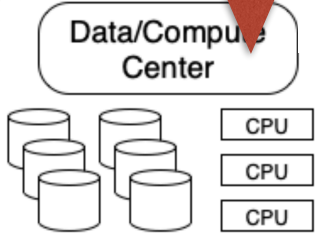
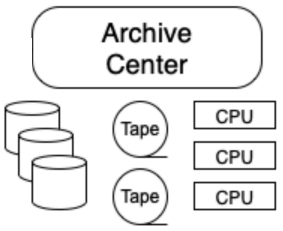
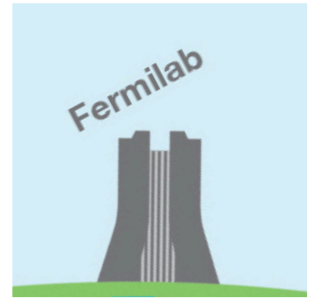
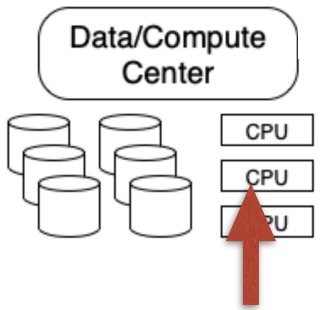


ProtoDUNE Raw Data
archival storage

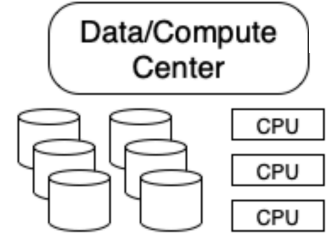
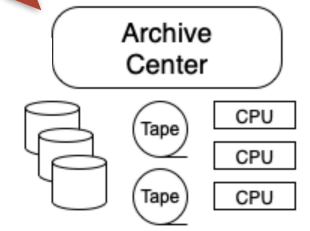
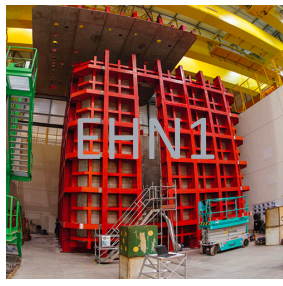
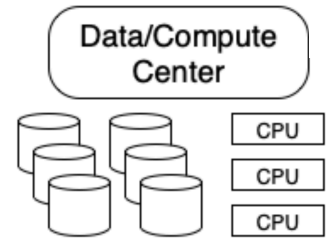
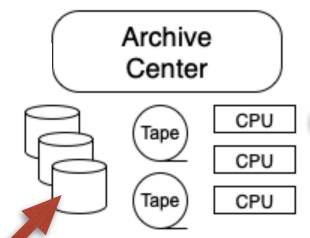


not to scale, not a technical design
it's just a cartoon

Cartoon version of data movement



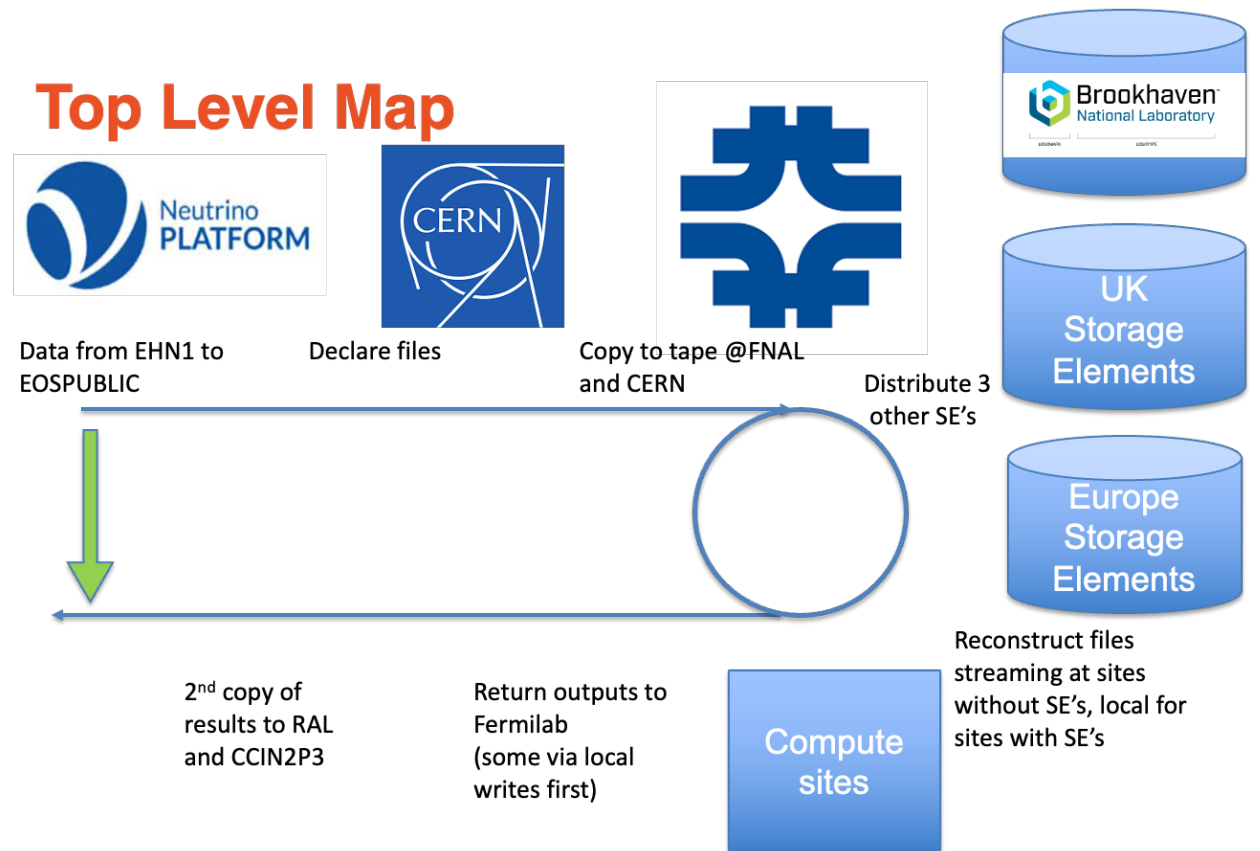
ProtoDUNE Raw Data raw data processing



not to scale, not a technical design
it's just a cartoon

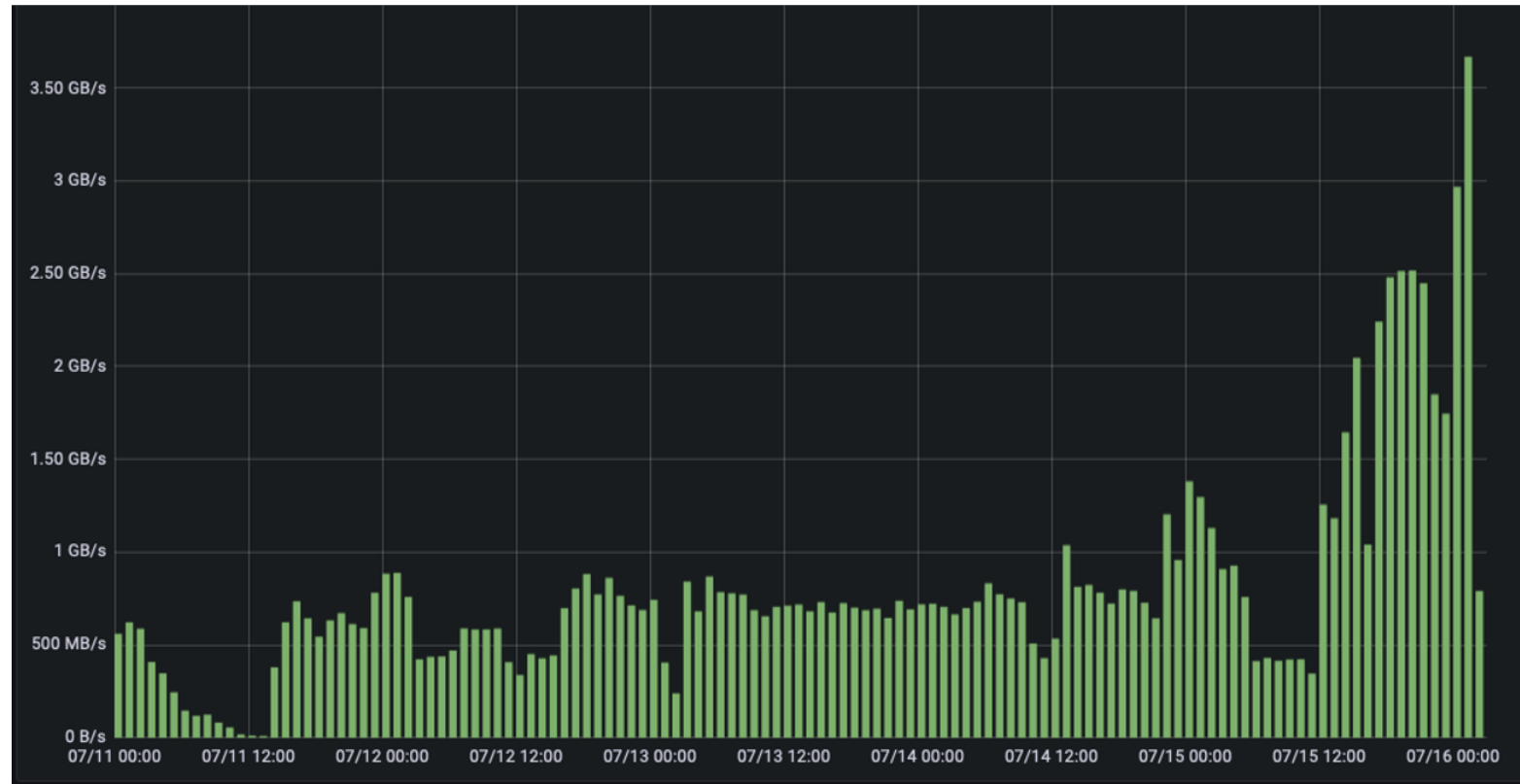
Summer 2022 Data Challenge 4 - ProtoDUNE

- Goals of the Data Challenge 4 were to test all the services and procedures that will be used in the forthcoming beam runs of ProtoDUNE-HD and ProtoDUNE-VD
 - These tests were done and were successful in transferring and processing data at the necessary scale
- recap of Phase 1 of the data challenge (Data Pipeline)
 - (data pipeline EHN1->CERN->FNAL)
- Then description of Phase 2 which tested two processing models
 - The Workflow System (now “justIN”)
 - The Data Dispatcher



Data Challenge 4 - Phase 1 - Data Pipeline for ProtoDUNE

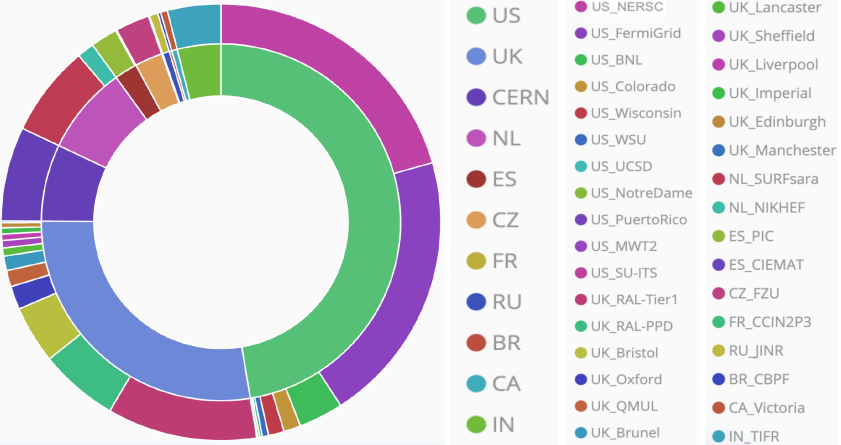
- Ran 5 days, July 11-15, Midnight UTC to Midnight UTC
- Generated ~500TB of data total
- Near end of the challenge reached a peak of 3.6GB/s = 28.8Gbit/s
- utilized FTS-3 for initial ingest, Rucio for replication, https and 3rd party transfers
- helped to identify issues that have to be fixed before beam run begins



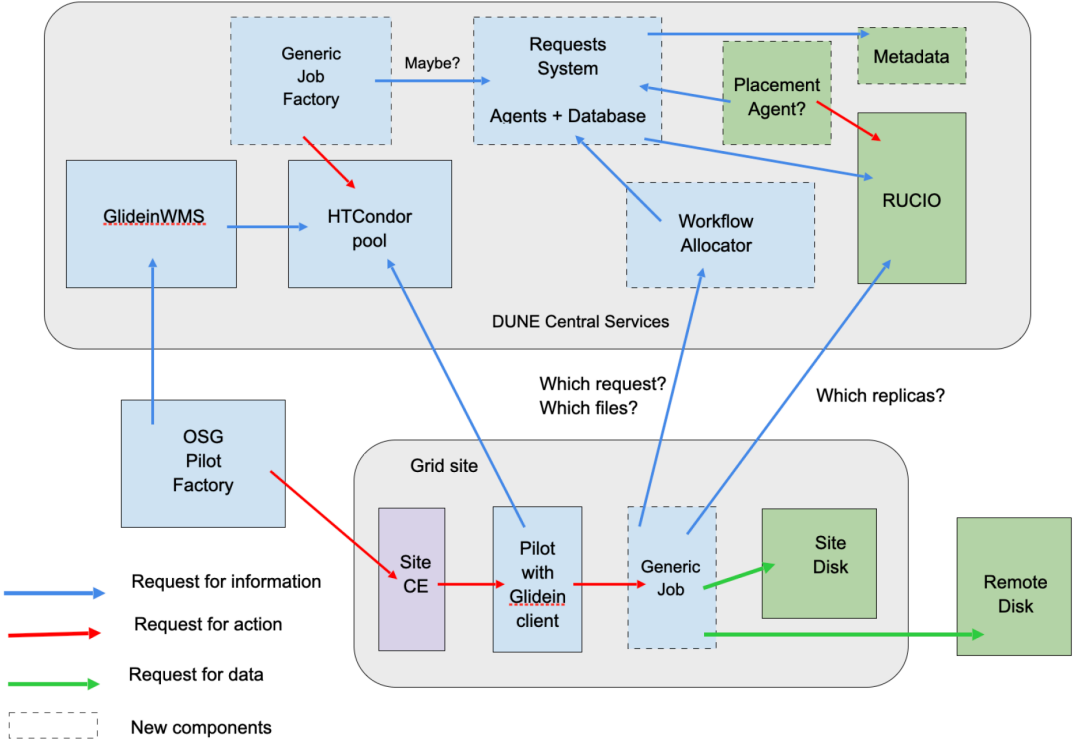
DC-4 Phase 2 - Data Processing

- Goal was to sustain 5000 concurrent jobs for keep up processing
- studies have shown significant drop in performance for jobs where large input data files not located “near” job node
 - new design addresses this problem
 - Manchester leading effort on workflow management development and testing
- need to actively manage data placements and job workflows taking the “landscape” into account

DUNE Production Job Wall Time



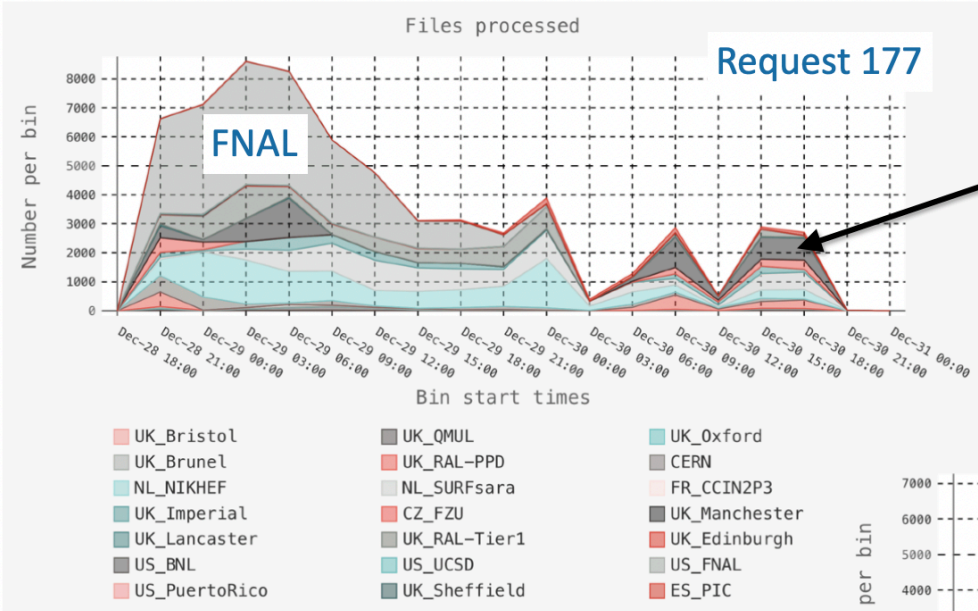
Workflow Management Dev Design



DC-4 Phase 2 - Data Processing

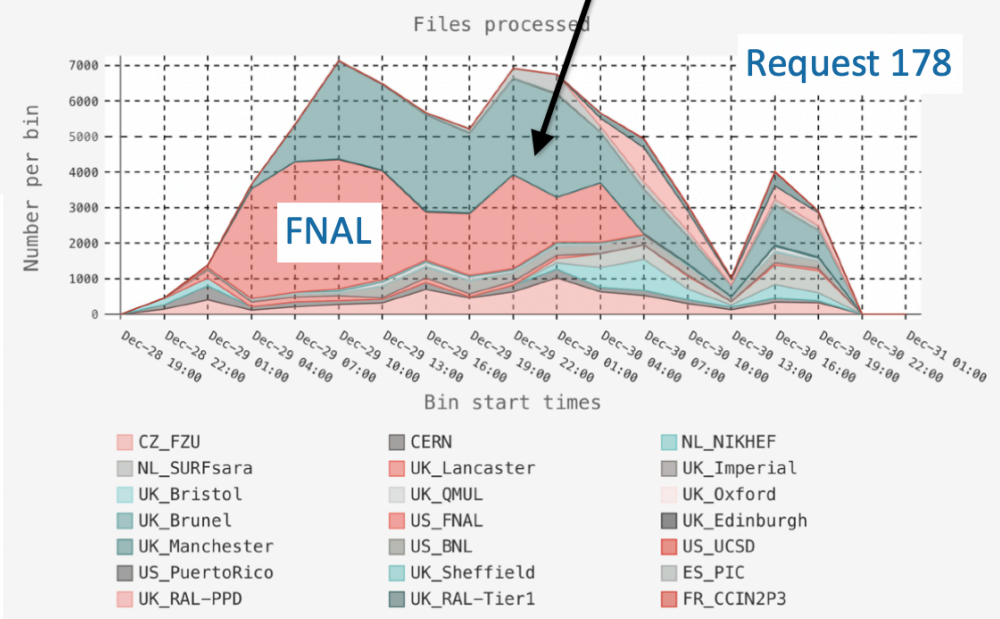
- Successfully shown ability to easily sustain 5000 jobs
- more importantly, automatic matching between the job locality and data locality
- looking for improved efficiency in the “flatter” computing model

(Charts from WFS Dashboard)



Not much Manchester

Lots of Manchester!



Allocations follow the data to suitable sites, using replica locations originally from Rucio

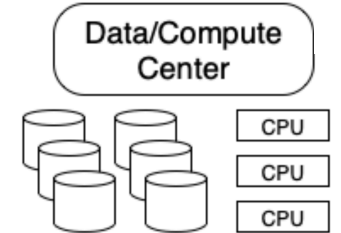
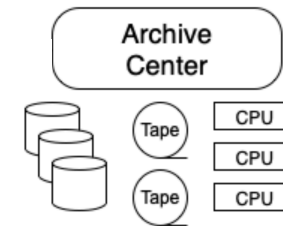
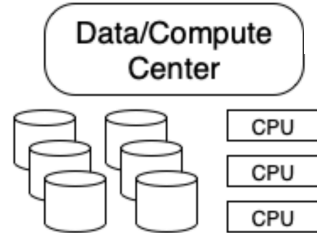
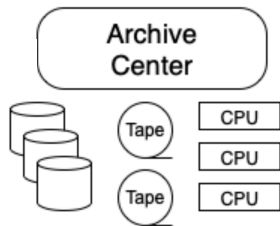
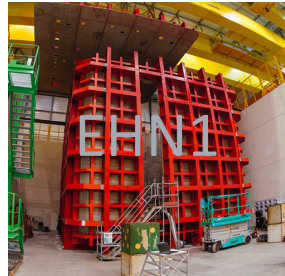
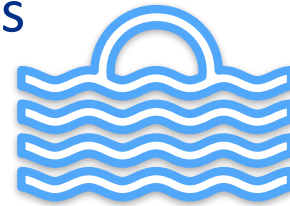
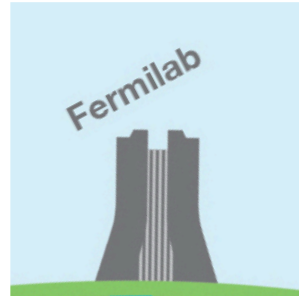
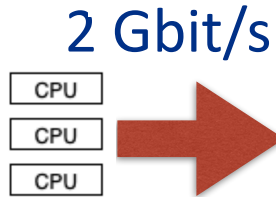
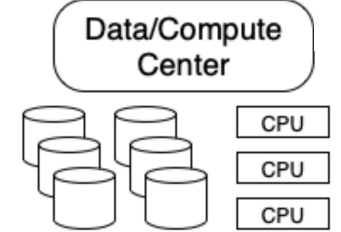
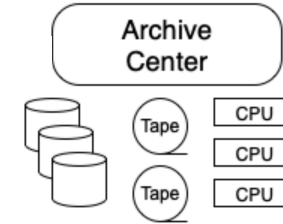
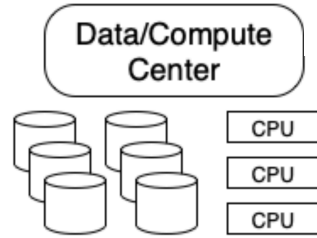
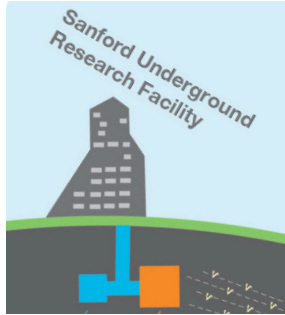


DUNE involvement in DC24

- Simulate the archival of 25% of the raw data rate from the Far Detector
 - translates to 2 Gbit/s from SURF to FNAL
 - replicate that “FD” raw data to archival storage facilities around the world
 - replicate the “FD” raw data to disk storage elements around the world for prompt access from compute elements
- Maintain continuous processing workload at distributed sites commensurate with 25% “FD” raw data rate
 - utilize compute elements across the WLCG and OSG
 - match the locality of jobs with locality of data at nearby RSEs
- sites will need to opt-in to participate in order to not interfere with WLCG DC24 goals (e.g. Tier2s that might have bandwidth limitations)
- RSE to RSE w/ token authentication/authorization

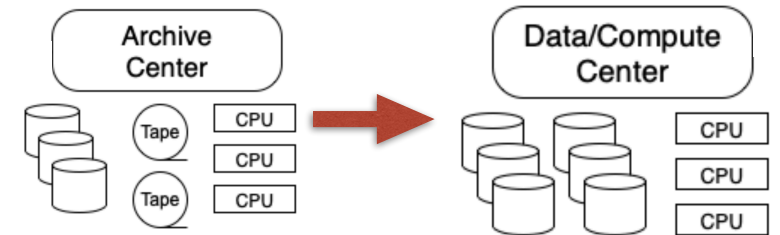
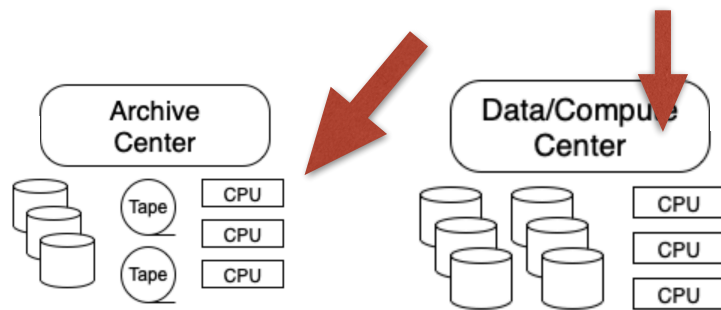
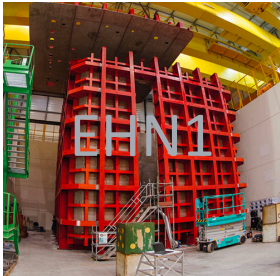
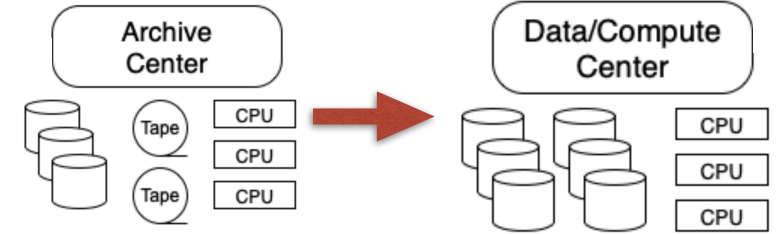
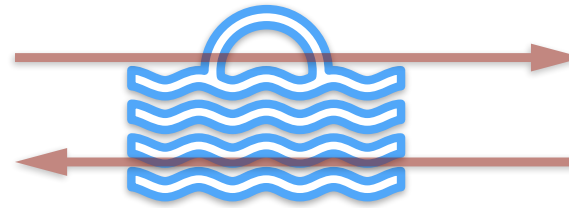
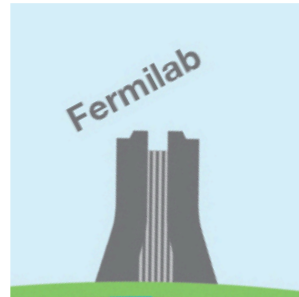
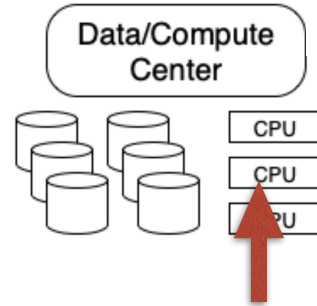
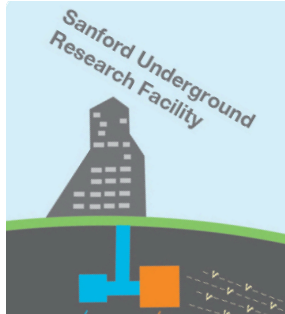
“FD” Raw Data archival storage

DUNE Involvement in DC24



**not to scale, not a technical design
it's just a cartoon**

Cartoon version of data movement

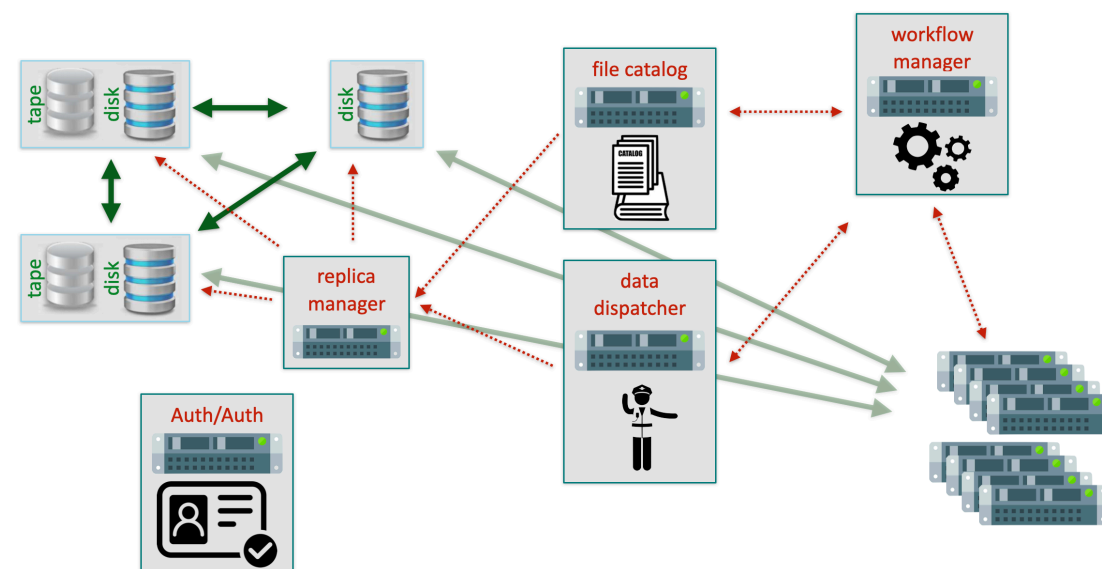


“FD” Raw Data
raw processing

not to scale, not a technical design
it's just a cartoon

Summary

- DUNE successfully utilizing WLCG resources
- Looking forward to addressing challenges
 - more complete Rucio integration
 - developing new workflows and workflow management - including access to HPC
 - integrate GPU software and hardware for processing - data prep especially
- looking forward to improved understanding that will come from involvement in DC24 and ProtoDUNE II operations
- exploring ideas for analysis centers
- improved projections for resource needs

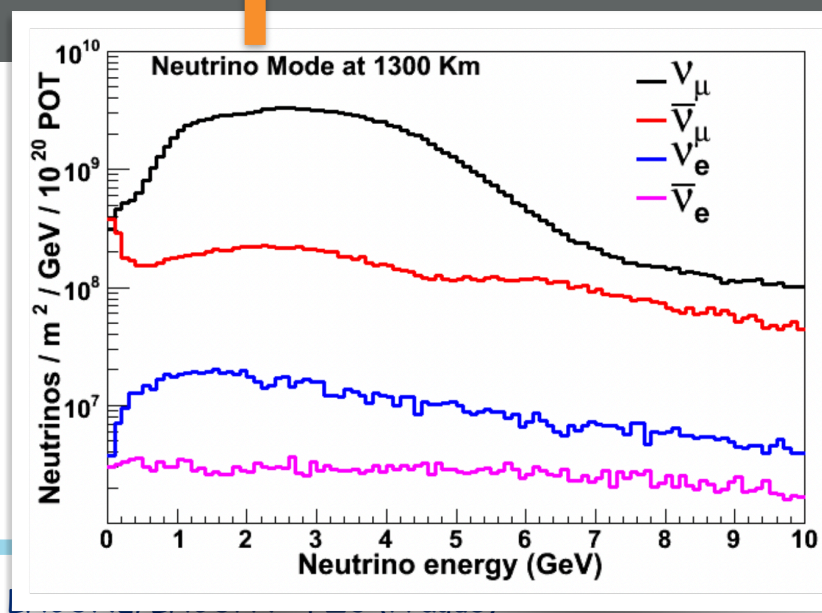
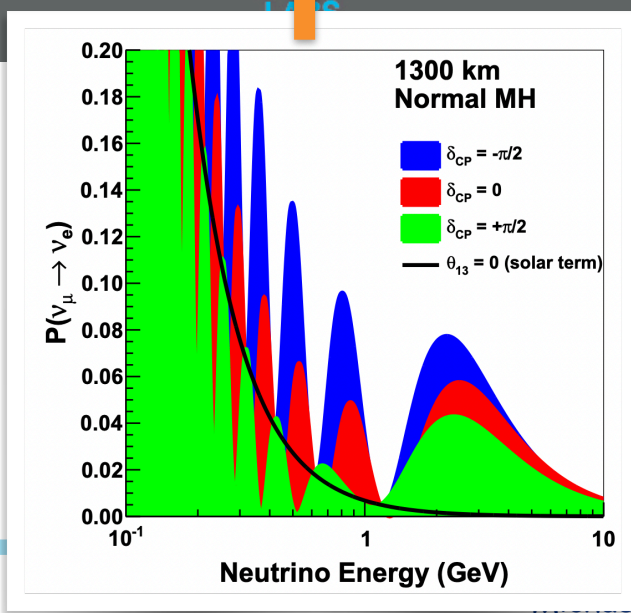
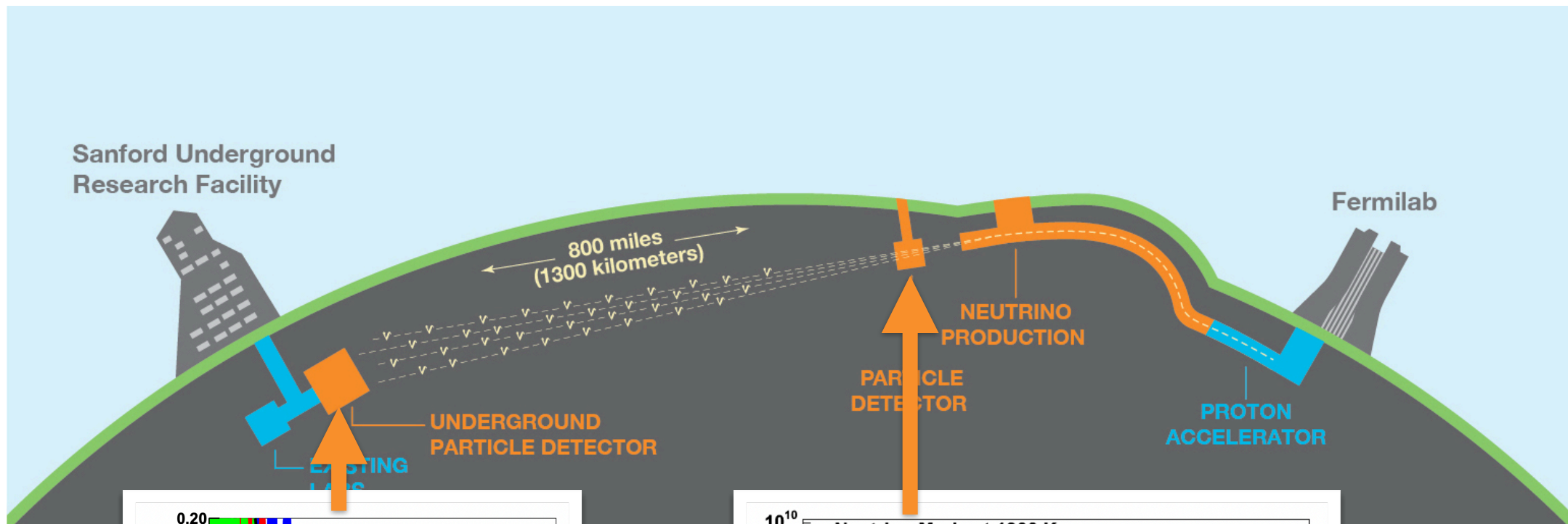


Thank you to the organizing committee



Backup

Optimized location of the Far Detector

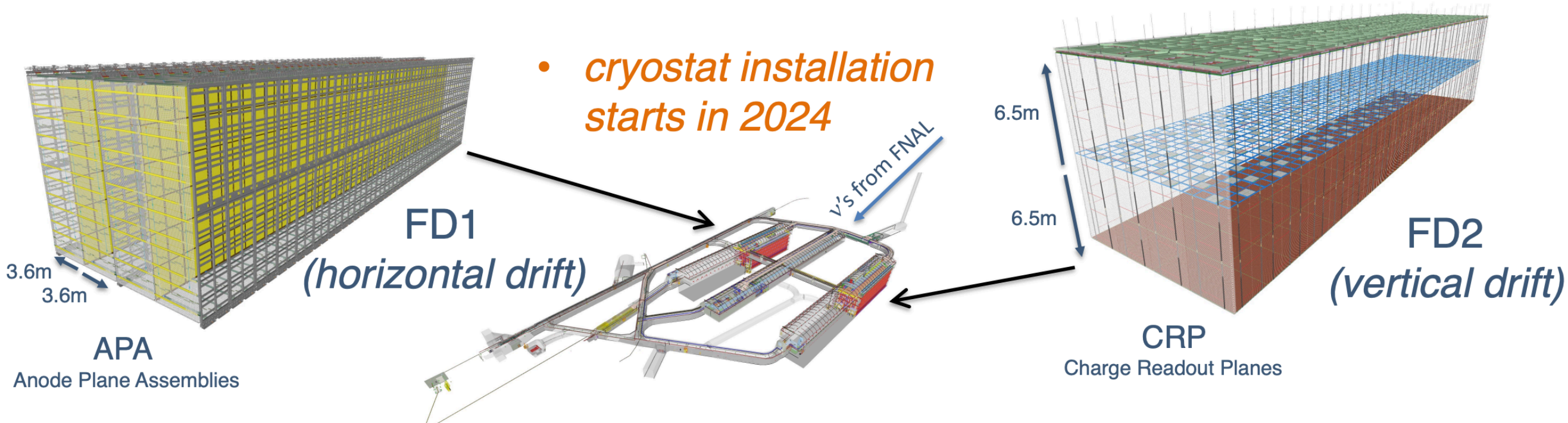


Kirby, WLCG

DUNE Far Detectors (Phase I)

- Phase 1 will include caverns for 4 detector modules in South Dakota and 2 far detector modules, each 17 kton of LAr, the largest LAr TPCs ever constructed.
 - FD1: horizontal drift (ala ICARUS, MicroBooNE)
 - FD2: vertical drift (capitalizing on protoDUNEs)

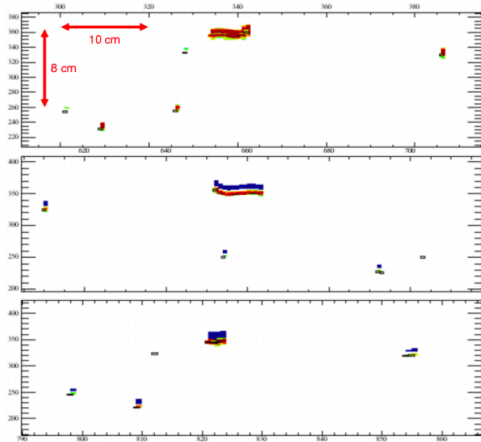
• cryostat installation starts in 2024



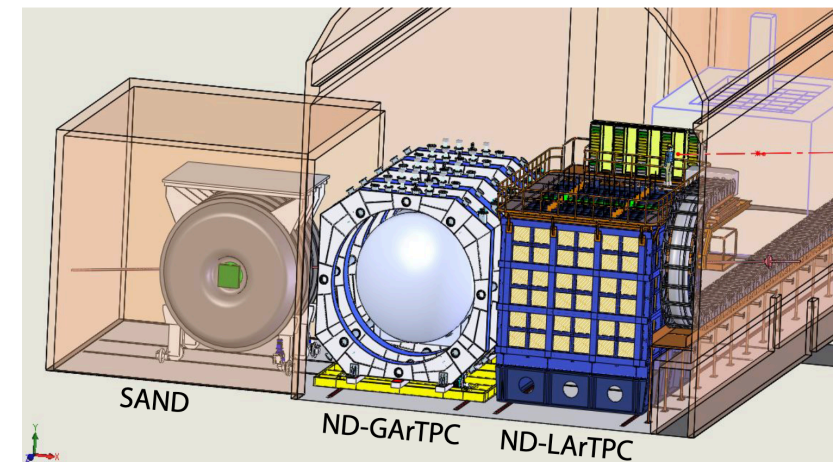
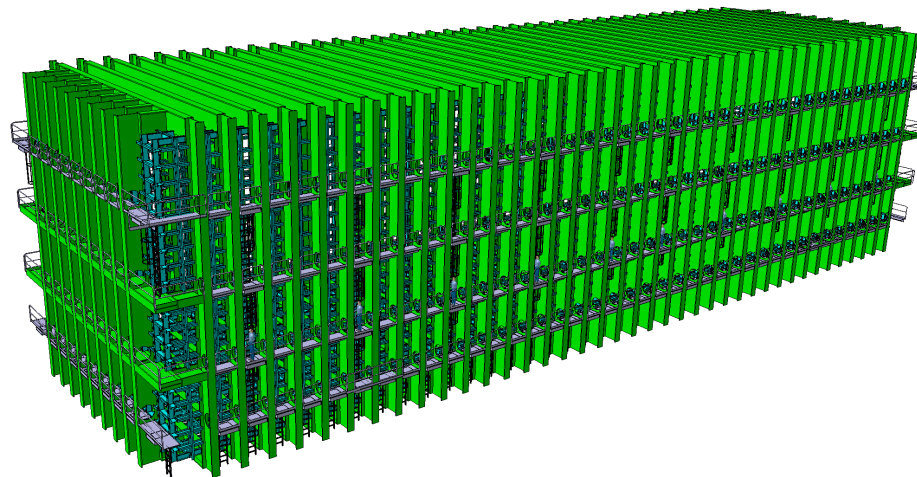
- Order of magnitude more mass than has been deployed up to now from all LAr TPCs

Recently published DUNE Computing CDR - <https://arxiv.org/abs/2210.15665>

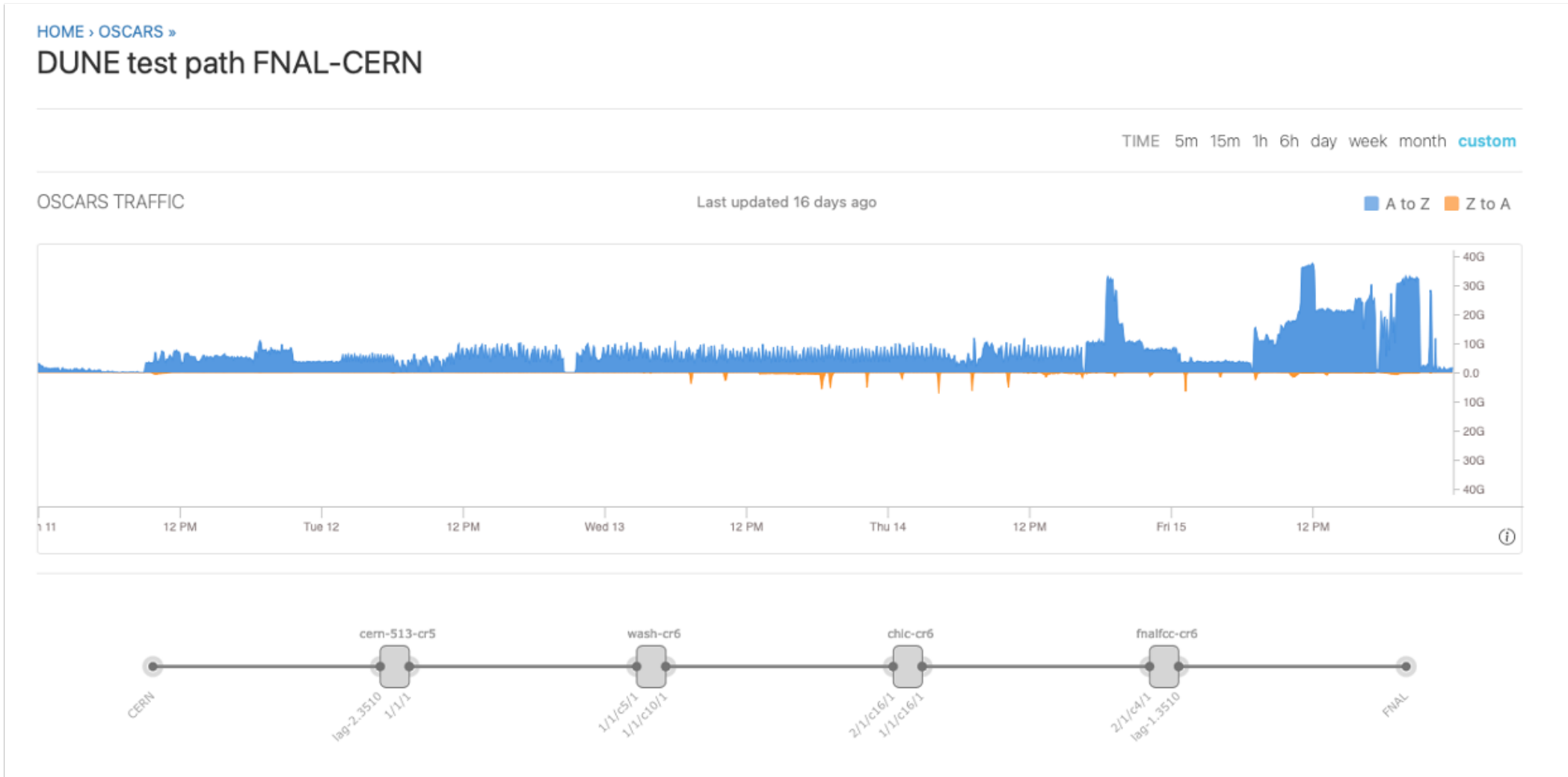
Simulated SN neutrino interaction



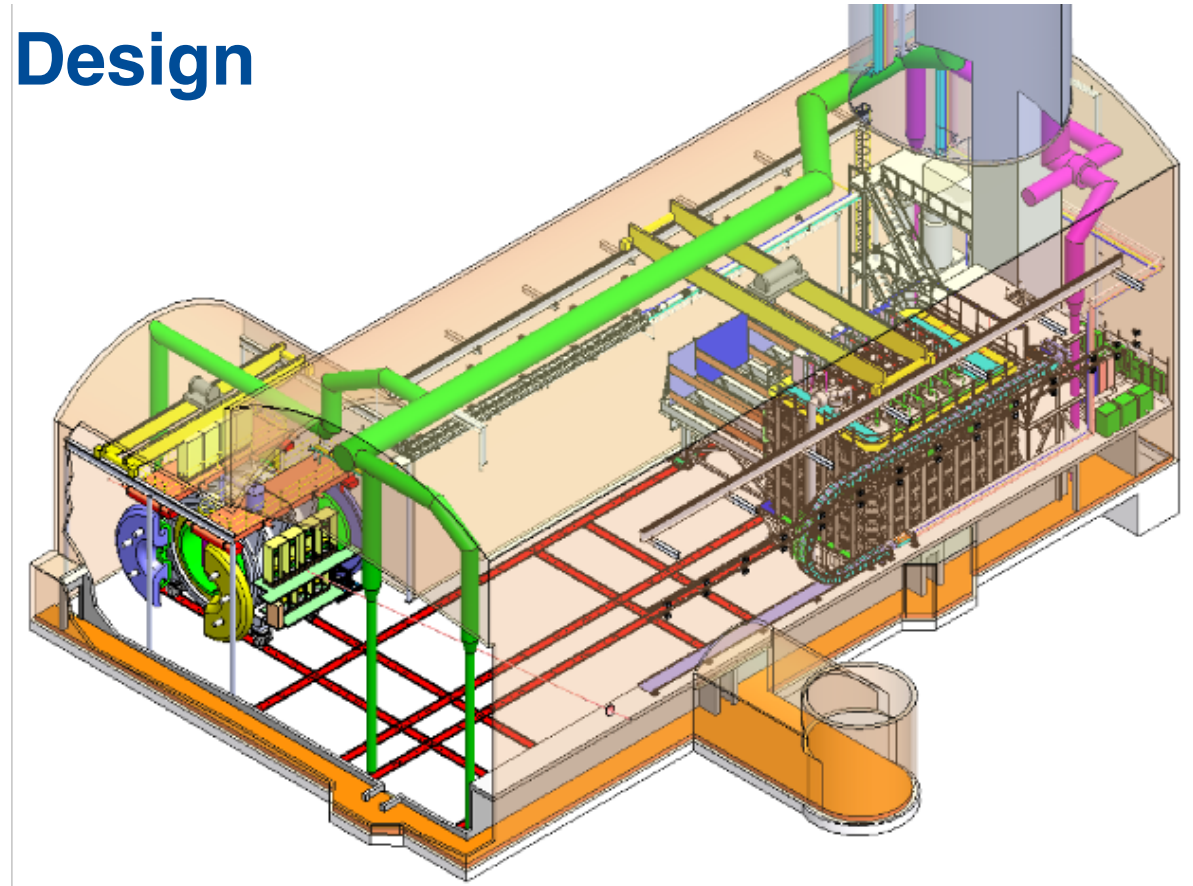
- starting from the DUNE physics goals, and the design of the Near and Far Detector define the offline computing challenges facing the Collaboration
- define the scope of the Offline Computing Consortium responsibilities
 - Code Management vs Algorithm Development
 - Running Processing Jobs vs Develop/Definition of a MC sample configuration
- list the requirements for computing services necessary to address those challenges
- layout tasks necessary to development/integrate solutions that address those requirements
- hopefully serve as a roadmap for contributions from collaboration members and computing experts going forward



CERN-FNAL traffic



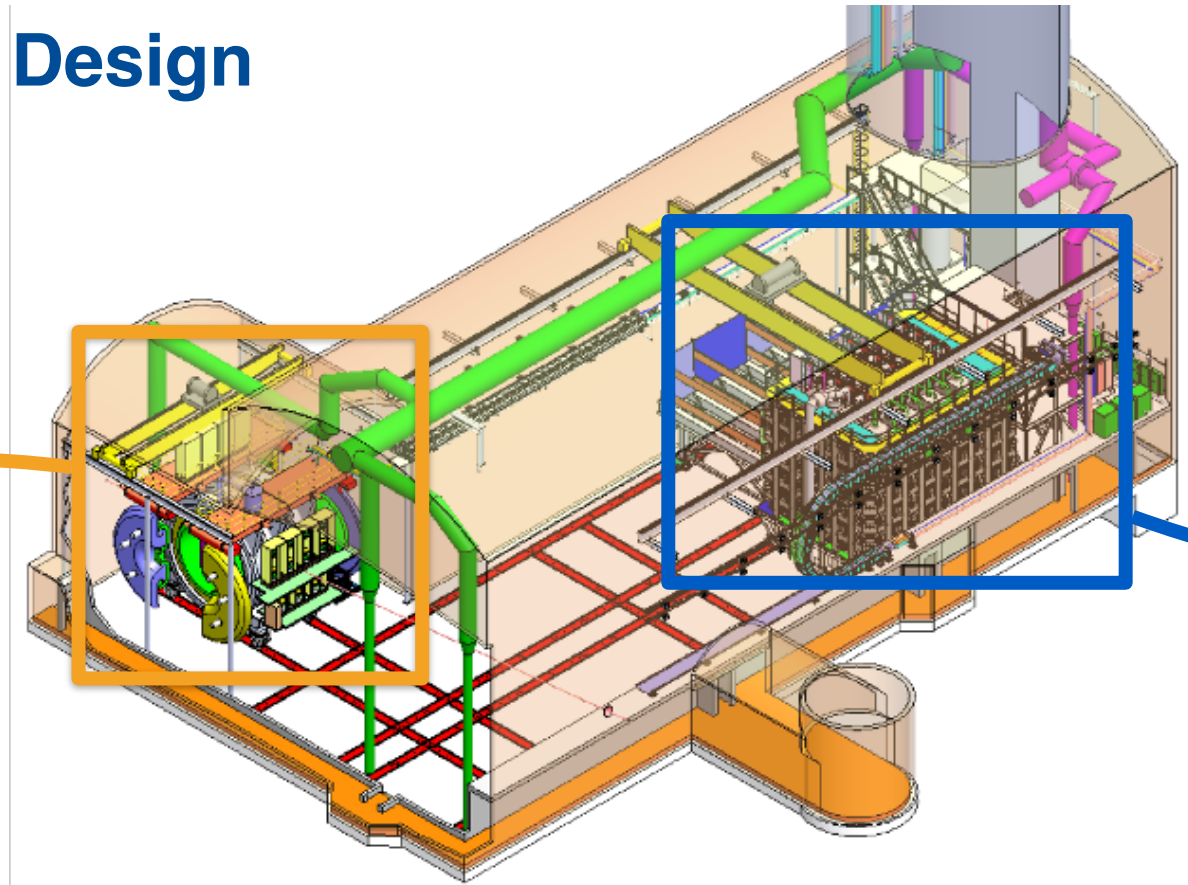
DUNE Near Detector Design



- Near Detector - three sub-detectors serving different purposes
- ND-LArTPC: Highly segmented Liquid Argon Time Projection Chamber
- TMS - Muon Spectrometer (Phase 1) / ND-GArTPC - Gaseous Ar TPC (Phase 2)
- SAND - scintillator-based tracking and active argon target for on-axis beam monitoring
- DUNE-PRISM: Movement of LAr + TMS/GArTPC transverse to the beam, sampling beam profile, energy spectrum, and ν interactions to reduce systematic uncertainties

DUNE Near Detector Design

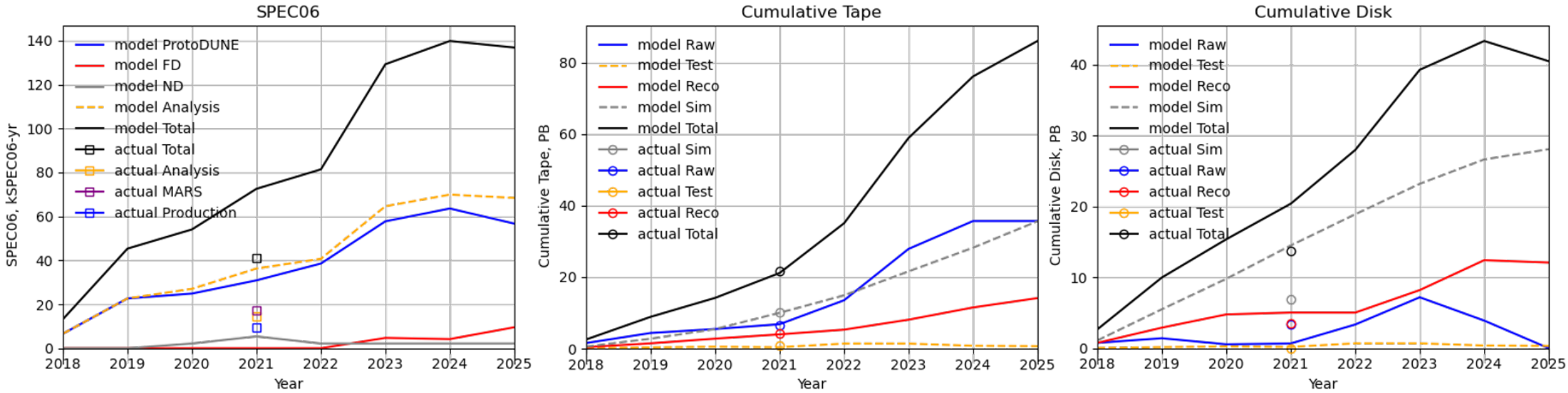
on-axis neutrino detector for beam monitor



ND that is functionally identical to FD - but moveable

- Near Detector - three sub-detectors serving different purposes
- ND-LArTPC: Highly segmented Liquid Argon Time Projection Chamber
- TMS - Muon Spectrometer (Phase 1) / ND-GArTPC - Gaseous Ar TPC (Phase 2)
- SAND - scintillator-based tracking and active argon target for on-axis beam monitoring
- DUNE-PRISM: Movement of LAr + TMS/GArTPC transverse to the beam, sampling beam profile, energy spectrum, and ν interactions to reduce systematic uncertainties

Data Volumes - through 2025



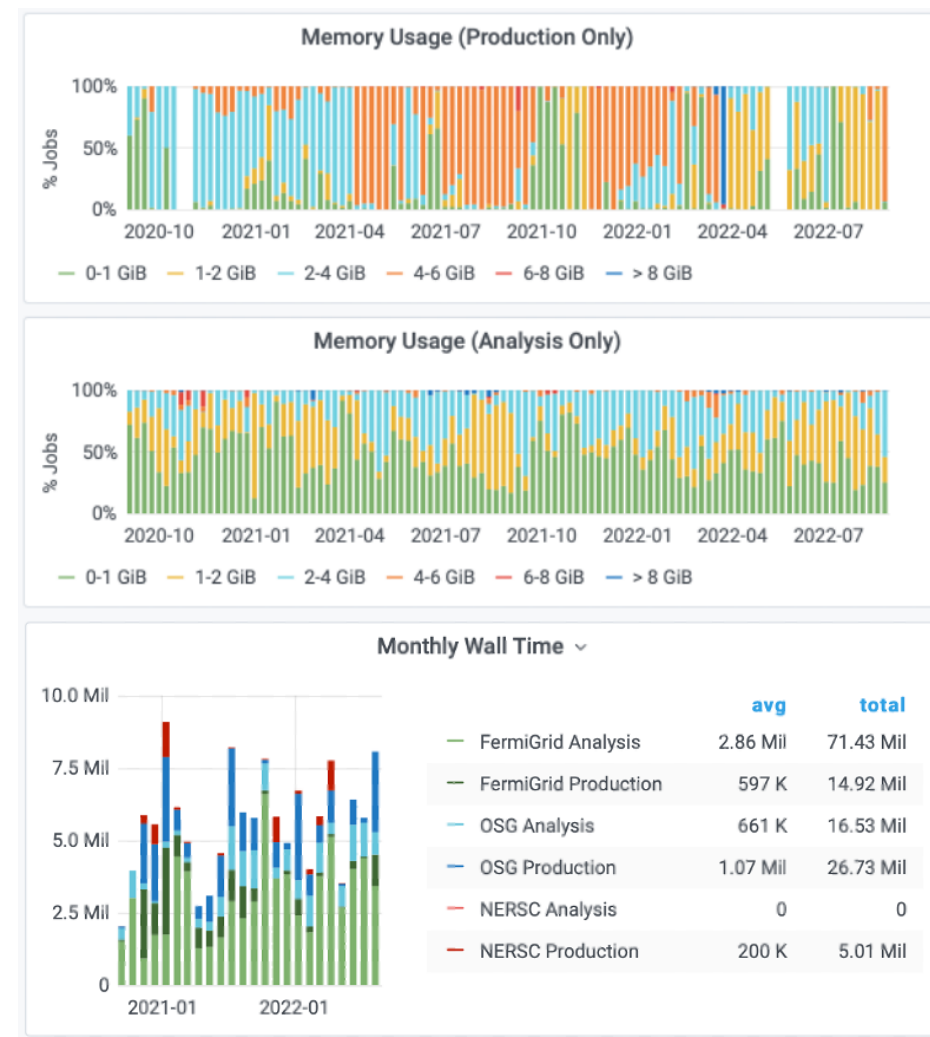
CDR Resource Estimates through 2025

Defining the Computing Challenges

Offline reconstruction experience and impact on workflow and resources

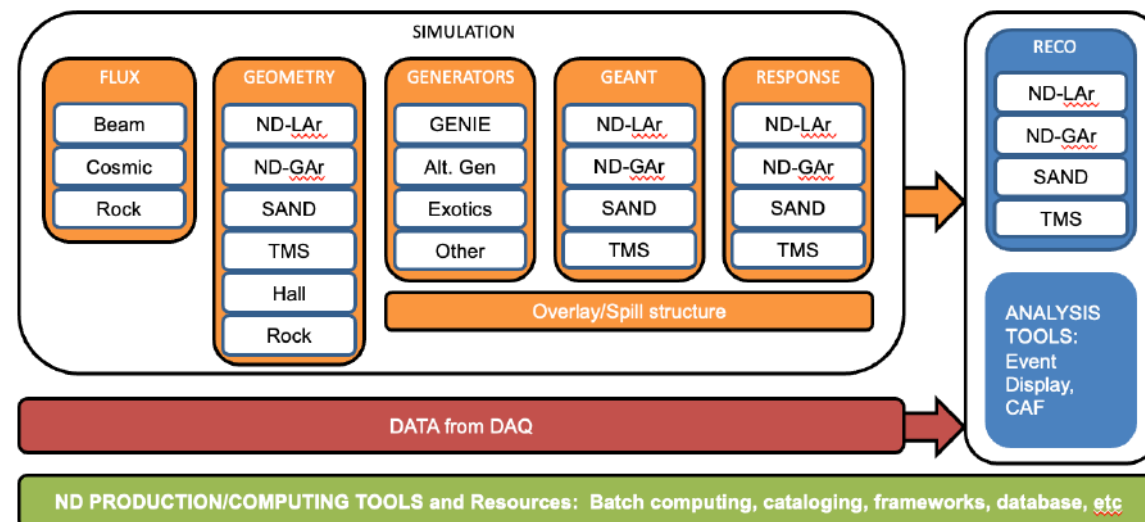
- ProtoDUNE offline reconstruction
 - signal processing memory “pinch point”
 - adapt workflow & software to fit site constraints
 - adjust to future expectation of localized readout
 - process time extended trigger records
- Near Detector simulation and reconstruction
 - strong algorithm development and important achievements helping to define ND design
 - need for integration into a framework
 - reco simulation samples for resource estimates
 - advances in ML/AI show need for an improved standard development environment

Last 2 years Grid job stats



Offline reconstruction experience and impact on workflow and resources

- ProtoDUNE offline reconstruction
 - signal processing memory “pinch point”
 - adapt workflow & software to fit site constraints
 - adjust to future expectation of localized readout
 - process time extended trigger records
- Near Detector simulation and reconstruction
 - strong algorithm development and important achievements helping to define ND design
 - need for integration into a framework
 - reco simulation samples for resource estimates
 - advances in ML/AI show need for an improved standard development environment



GARSoft run time stats

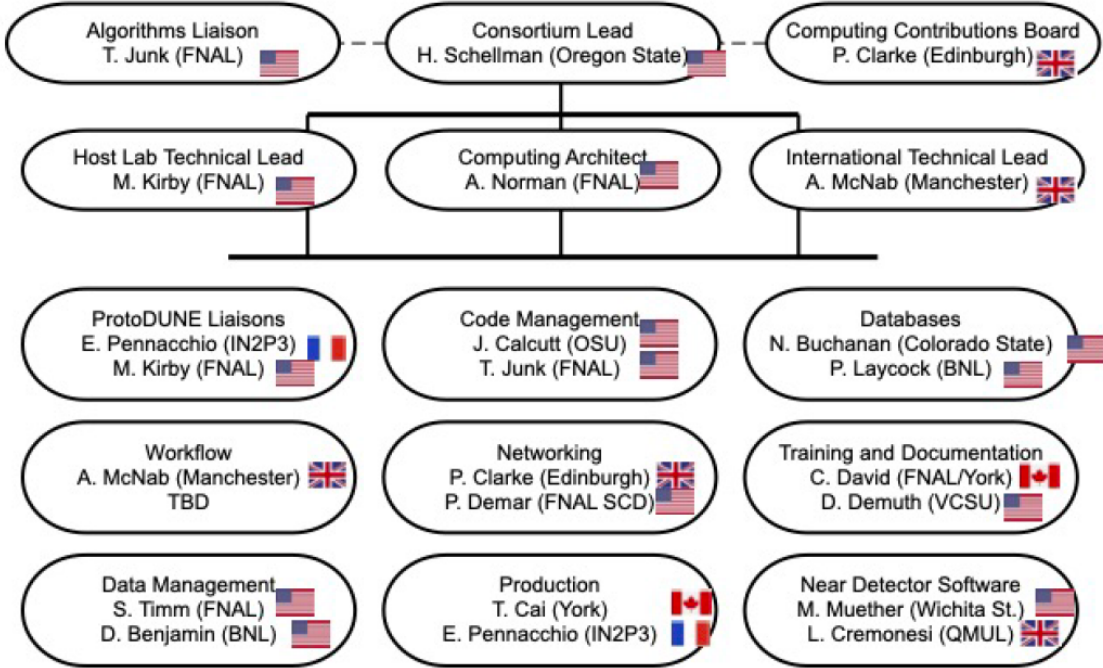
Module Label	time/event (sec)
RootInput(read)	0.00
init:EventInit	1.31275e-05
hit:CompressedHitFinder	0.00488308
tpclusterpass1:TPCHitCluster	0.0091922
vechit:tpcvechitfinder2	0.0103787
patrec:tpcpatrec2	0.0130245
trackpass1:tpctrackfit2	0.014211
vertexpass1:vertexfinder1	0.000851081
tpcluster:tpccathodestitch	0.0269436
track:tpctrackfit2	0.0135842
vertex:vertexfinder1	6.19847e-05
veefinder1:veefinder1	9.96417e-05

Characterization of the Large Scale Processing Challenge

Use case	memory	input file size	output file size	CPU time	input	cores/job
units	GB	MB	MB	sec	MB/s	
Simulation+reco	6	100	2000	27000	0.00	1
data reco	4	8000	4000	60000	0.13	1
tuple creation	3	4000	1	100	40.00	1
calibration	3	4000	1	100	40.00	1
Parameter estimation	1	400	1	600000	0.00	68000

Overview of the Computing Consortium

- organization of the working groups and management
- define source of resources
 - CCB - computing resources
 - division of effort for development
- layout interfaces with other consortia



Consortium

- Project management
- Core software development
- Core operations
- Interfaces to DUNE consortia
- Interfaces to other projects
- Training
- User support

Sites

- Storage
- CPU
- Networking

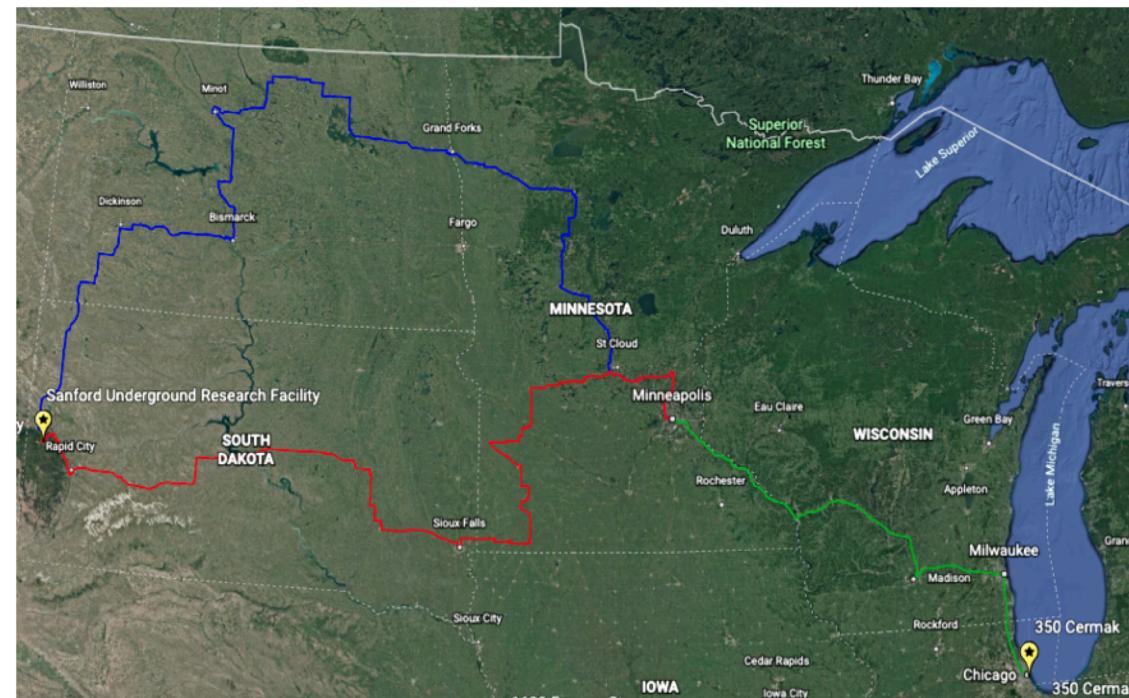
• Collaboration

- Algorithms
- Operations help
- User support
- Production group
- Validation
- Calibrations
- New ideas

Listing the Unique DUNE Requirements

Networking for DUNE

- important requirements for operations
 - potential natural disasters combine to make multiple paths important
- interface between online and offline resources
- unique challenges with SNB data transfer and opportunity for fast processing
- wide area network:
 - DUNE is very well served by national science networks
 - ESnet, Geant, Janet, Surfnet, Renater....
- DUNE relies on these and thanks them.
 - fully engaged and participate in the global LHCOPN/LHCONE meetings
 - NRENS are ~1 TB backbones w/ 100 GB to many sites
 - discussed with ESnet potential/need for virtualised networks - not foreseen at this time



DUNE FD WAN Bandwidth Timeline Projections:

Date	Stage of the experiment	Primary Path	Secondary Path	Tertiary Path
Now	Cavern excavation	10GE	< 1GE via SURF	none
2025	Detector construction	10GE	< 1GE via SURF	none
2027	Computing/DAQ deployment	100GE	10GE	< 1GE via SURF
2028	Cryo deployment completed	100GE	10Gb/s+	10GE
2029	Start of science	100GE	10Gb/s+	10GE

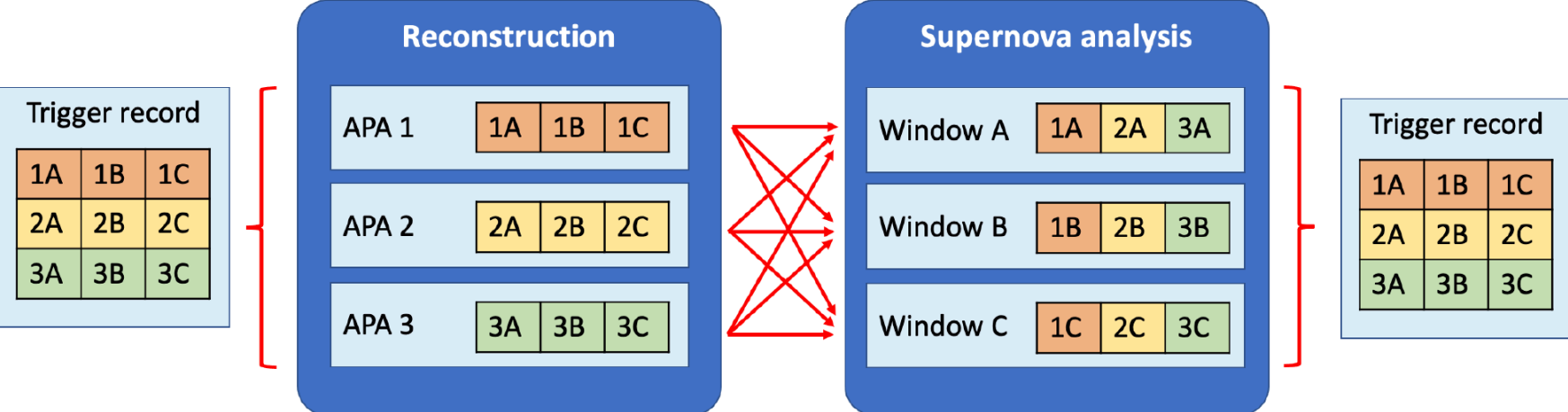
- vLAN service provided by REED/GPN (shared)
- Dedicated circuit Ross Dry Bldg. to Chicago
- Dedicated circuit Yates Complex to Denver (10GE or 100GE)

Offline Computing Frameworks

- thanks to **Frameworks Requirements Working Group**
- document available on docdb at [DocDB 24423](#)
- review by the HEP Software Foundation (HSF)
 - recommend Production and Analysis frameworks may be different
 - develop unified framework for all offline prod
 - plan interface with Workflow Management

- handle very large events on complex architectures
 - partial region and subsetting of trigger records
 - temporal stitching of trigger records
- threadsafe + i/o management
- work plans for development of dynamic time-window processing
 - FNAL LDRD funded for a prototype framework
 - DUNE coordinating closely with FNAL effort

Diagram from Kyle Knoepful FNAL LDRD



Laying out paths to solutions

WEB Form

Databases

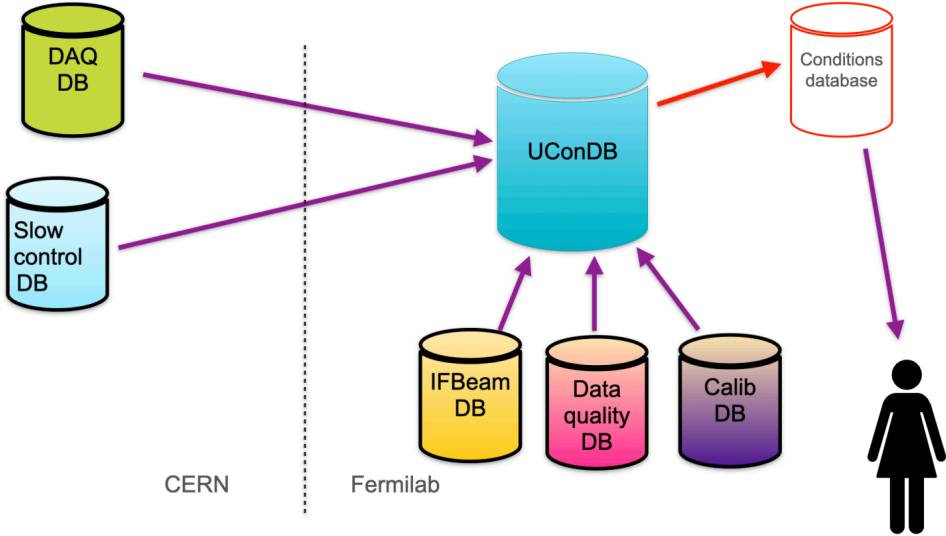
- layout of current and planned databases
- define the interaction with online/offline/analyzers
- important to complete interface w/ consortia & physics groups
- plan in place for development, service, & maintenance effort
- one coherent interface/point of access for all conditions needed for data processing and analysis
- new postdocs making significant progress
 - **Alex Wagner** - Hardware DB interface
 - Swift/python entry framework
 - **Ana Paula Vizcaya Hernandez** - UConDB and Conditions DB
 - input from RunConfigDB and IFBeamDB
 - **Lino Gerlach**- Conditions DB
 - Slow Controls (SCADA)
 - Calibration inputs into Conditions

- WEB forms are provided by HWDBI.

YOU define the form

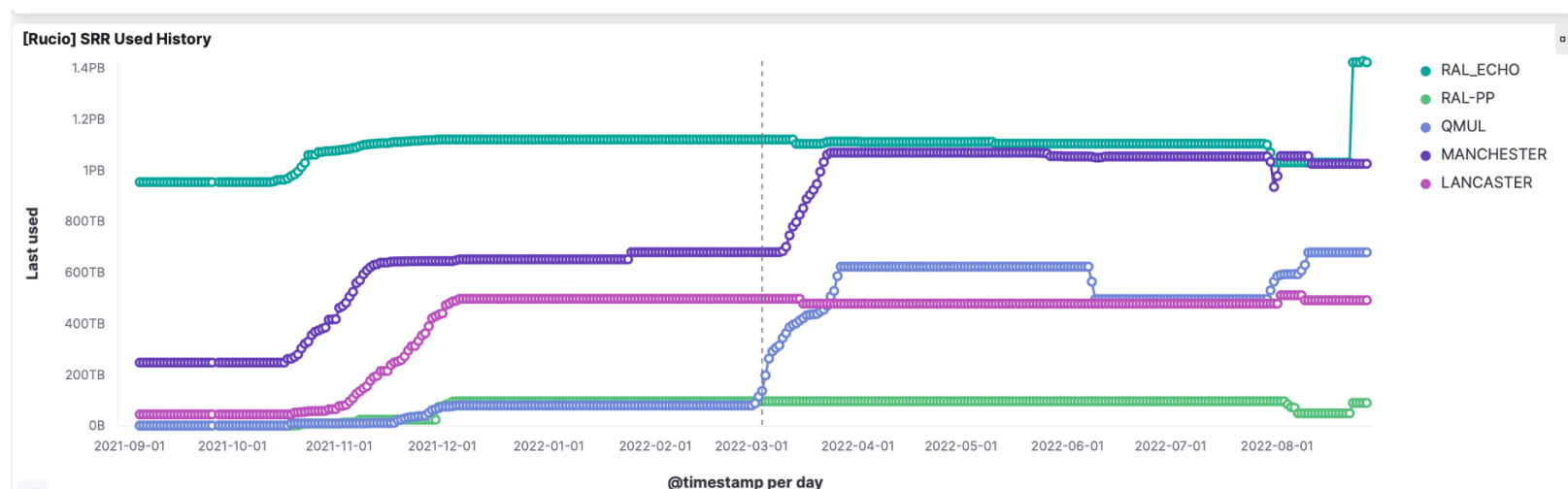
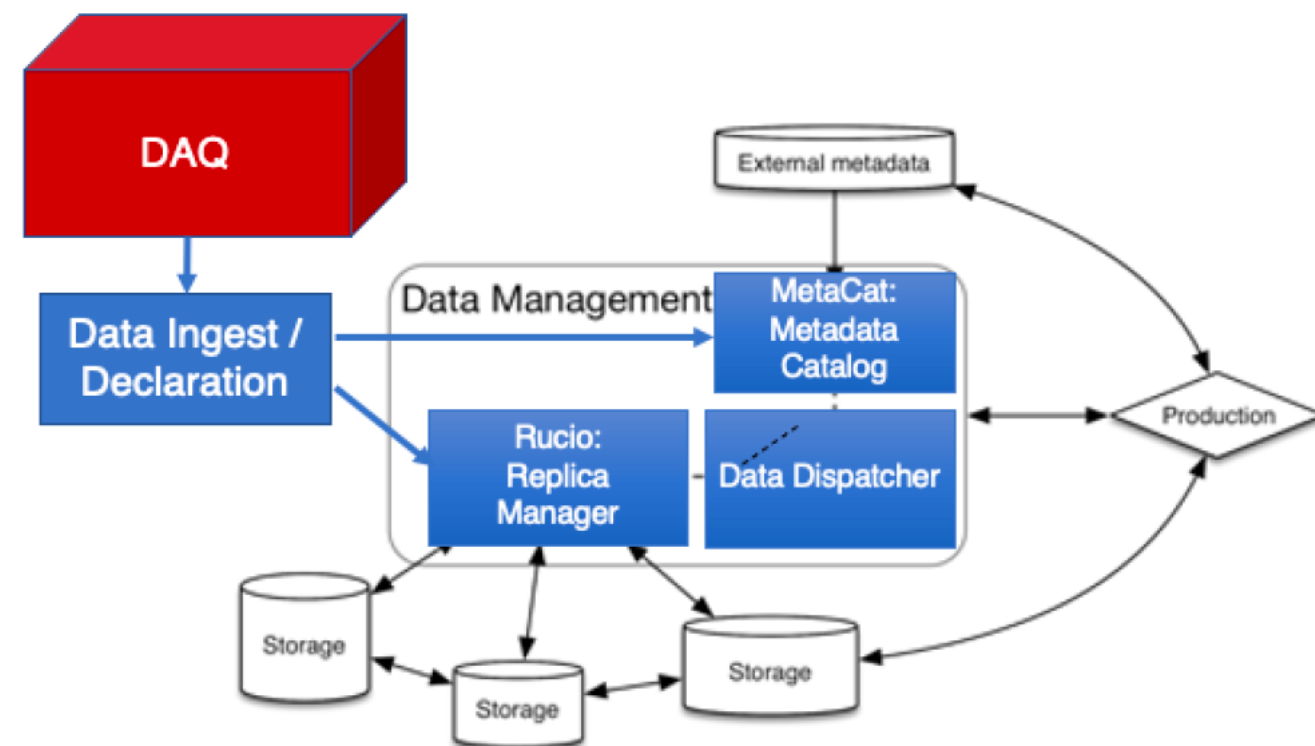
User(s) can then fill the form

Hajime



Data Management

- DUNE's large data volumes and global computing resources require replica manager and updated data management
- recognized early in process that current DM tools would not scale
 - adopted Rucio for replica mgmt
- defined requirements for future system and developed plan for development
- all components are at least at the testing stage with some very close to being declared production



Effort from FNAL, BNL, Edinburgh, and Oregon State involved in accomplishing transfers