DUNE Computing Usage in 2021

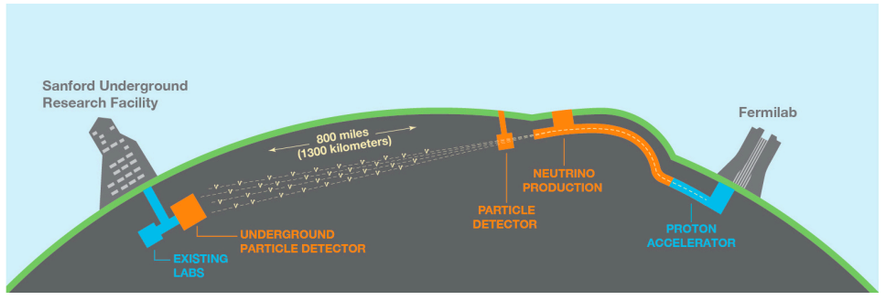
Michael Kirby, Fermilab/Scientific Computing Division

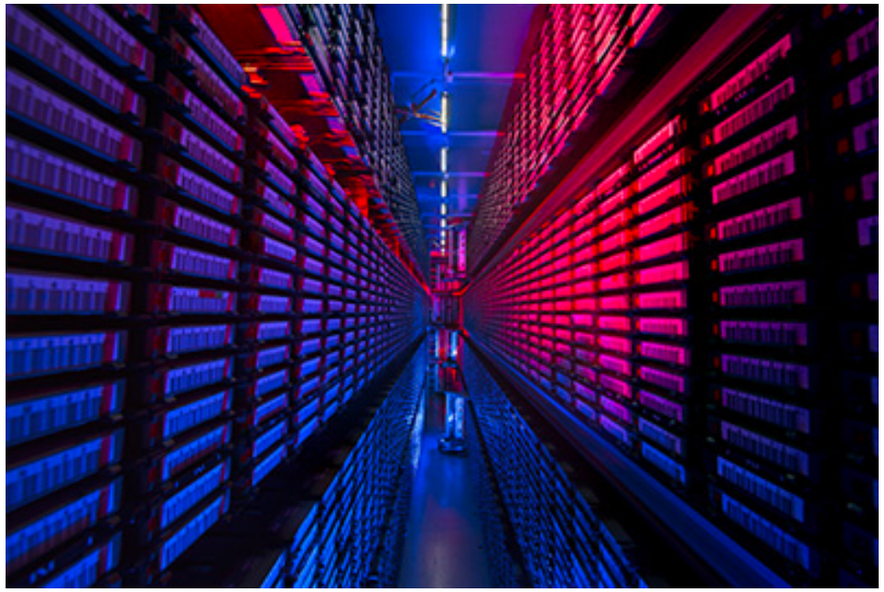
DUNE Computing Contributions Board

Jan 26, 2022

Outline

- general overview of computing plan (2020-2022) from November 2020 CCB meeting
- estimated resource requests from Nov 2020
- Pledged resources for 2021
- Deployed resources in 2021
- Used resources in 2021
- discussion





Fermilab

Fermilab Computing Resource Scrutiny Group

- Fermilab CRSG review occurred May 4-5, 2020
- CRSG review of resource requests from FNAL experiments (DUNE, NOvA, Mu2e, etc)
- reviewed DUNE computing model and resource projections for 2020-2022
- commended FNAL and DUNE on choice of Rucio for Data Management
- emphasized importance of data lifetime policy
- focus on dCache (disk) usage requests
- stated that experiment resource requests were reasonable
- may want to delve deeper into the larger experiment resource justifications

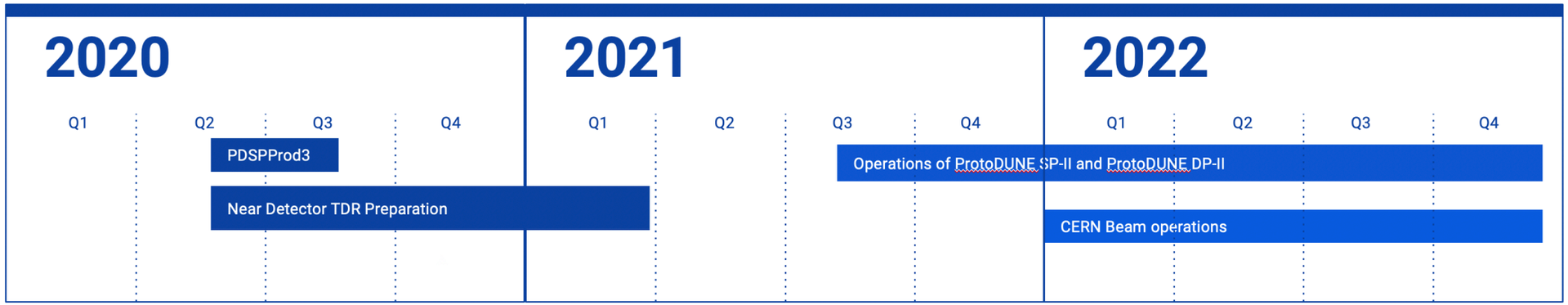
Assumptions from Nov 2020

These are NOT current projections

General Plan for DUNE Computing 2020 - 2022

- Support the operation and analysis of protoDUNE SP and DP (beam and non-beam data)
  - archive of raw, derived, and simulated data - successful keep-up of data taking
  - production processing of SP and DP data twice per year - Three PD-SP Prod Campaigns
  - production processing of simulation for SP and DP - all sim requests for PD complete
  - Successfully supported numerous PD-SP results
- Support the DUNE Far and Near Detector for design and sensitivity improvements
  - simulation of FD and ND samples - DUNE TDR published - ND CDR posted on arXiv
  - sensitivity studies for detector improvements and reviews
- Utilize the OSG, WLCG, and additional compute resources - many new sites available
- Improve access to HPC resources for specific analysis workflows and production workflow development
- With more robust implementation of Rucio, start to replicate data to global storage elements

Beam and Cosmic Operations assumptions



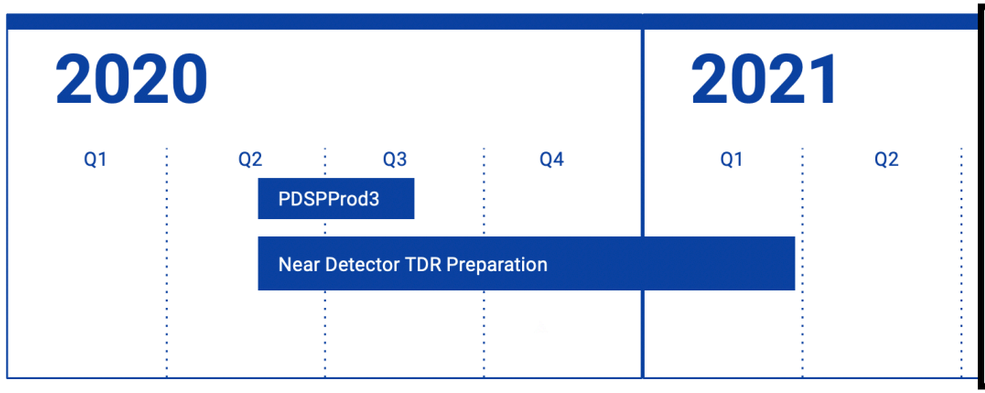
- ProtoDUNE II
  - 150 days of cosmic running in 2020/2021/2023 - both SP (6.5 M evts) and DP (21.6 M evts)
  - 50 days of beam in 2021 and 2022 - both SP (17 M evts) and DP (40 M evts)
  - cosmic running in 2022 - SP (300 days - 13 M evts) and DP (200 days - 24.4 M evts)
- Commissioning of one SP-FD module begins with 150 APAs in 2026
- Commissioning of one DP-FD module begins in 2028
- “ProtoND” operations in 2022 and 2023
- Commissioning of ND starts in 2028

Assumptions from Nov 2020

These are NOT current projections

Fermilab

Beam and Cosmic Operations assumptions



Important change to detector design:
ProtoDUNE SP -> ProtoDUNE HD
ProtoDUNE DP -> ProtoDUNE VD
DUNE FD-SP -> FD1-HD
DUNE FD-DP -> FD2-VD

- ProtoDUNE II
  - 150 days of cosmic running in 2020/2021/2023 - both SP (6.5 M evts) and DP (21.6 M evts)
  - 50 days of beam in 2021 and 2022 - both SP (17 M evts) and DP (40 M evts)
  - cosmic running in 2022 - SP (300 days - 13 M evts) and DP (200 days - 24.4 M evts)
- Commissioning of one SP-FD module begins with 150 APAs in 2026
- Commissioning of one DP-FD module begins in 2028
- “ProtoND” operations in 2022 and 2023
- Commissioning of ND starts in 2028

Assumptions from Nov 2020

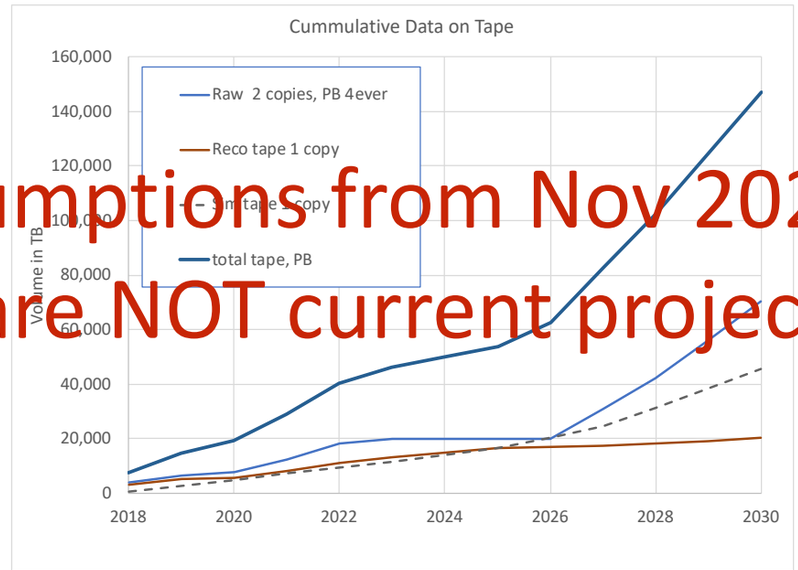
These are NOT current projections

Fermilab

Computing Model and Data Model Policies for 2020 and 2021

- ProtoDUNE
  - reconstruction processing twice/year
  - four active data datasets on disk
  - simulation processing once/year
  - two active MC datasets on disk
  - production processing continues though 2025 when datasets are frozen
- DUNE Far Detector
  - reconstruction processing twice/year
  - process full dataset - not seen til 2030
  - simulation processing once/year
  - if data lifetime is shorter than 10 years, effect not seen in model

- Can estimate tape, disk, and CPU usage for each year
- analysis CPU estimated as equal to production processing
- analysis storage is considered < 10% perturbation



Assumptions from Nov 2020

These are NOT current projections

Fermilab

Total DUNE Resource Needs for 2020 - 2022 presented in Nov 2020

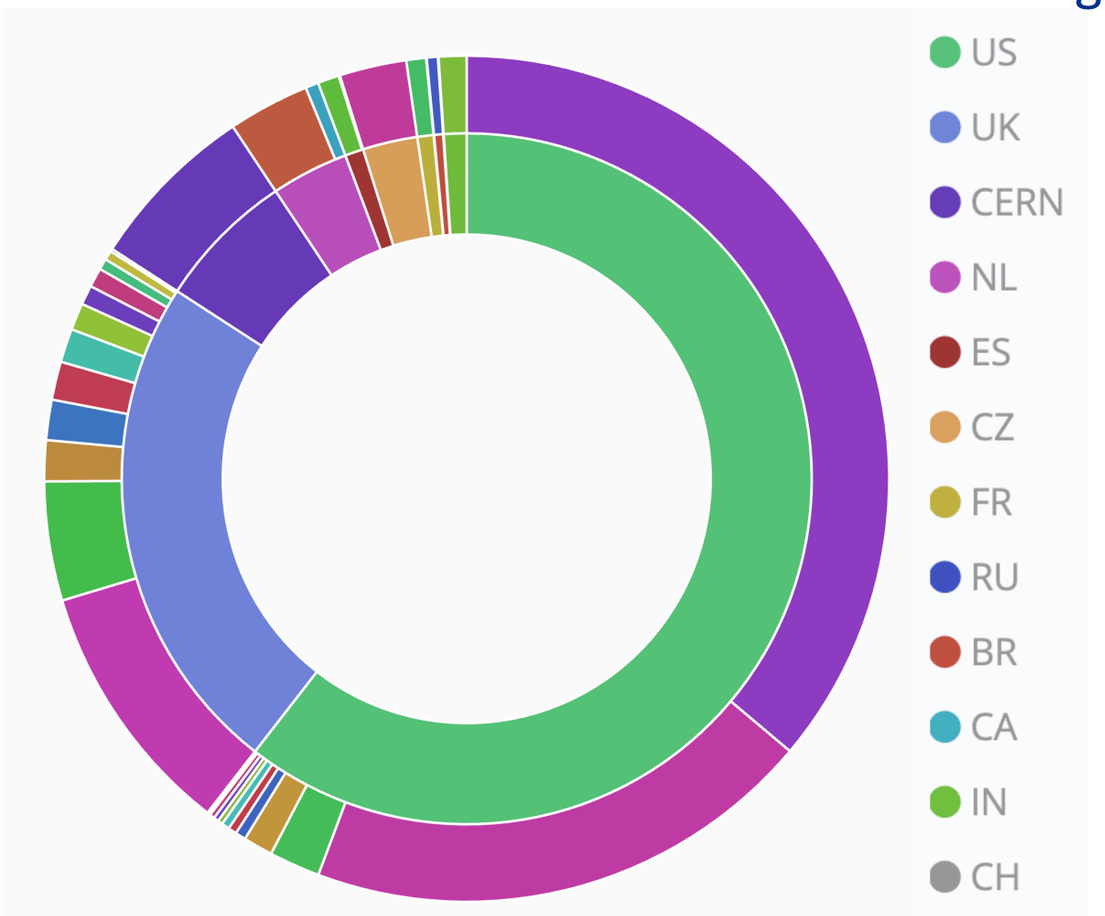
Resource	2020	2021	2022
Disk (PB)	16	20	26
Tape (PB)	19	29	40
CPU (kHS06-years)	33.1	50.2	52.5
CPU Cores	2200	3460	3540

Assumptions from Nov 2020

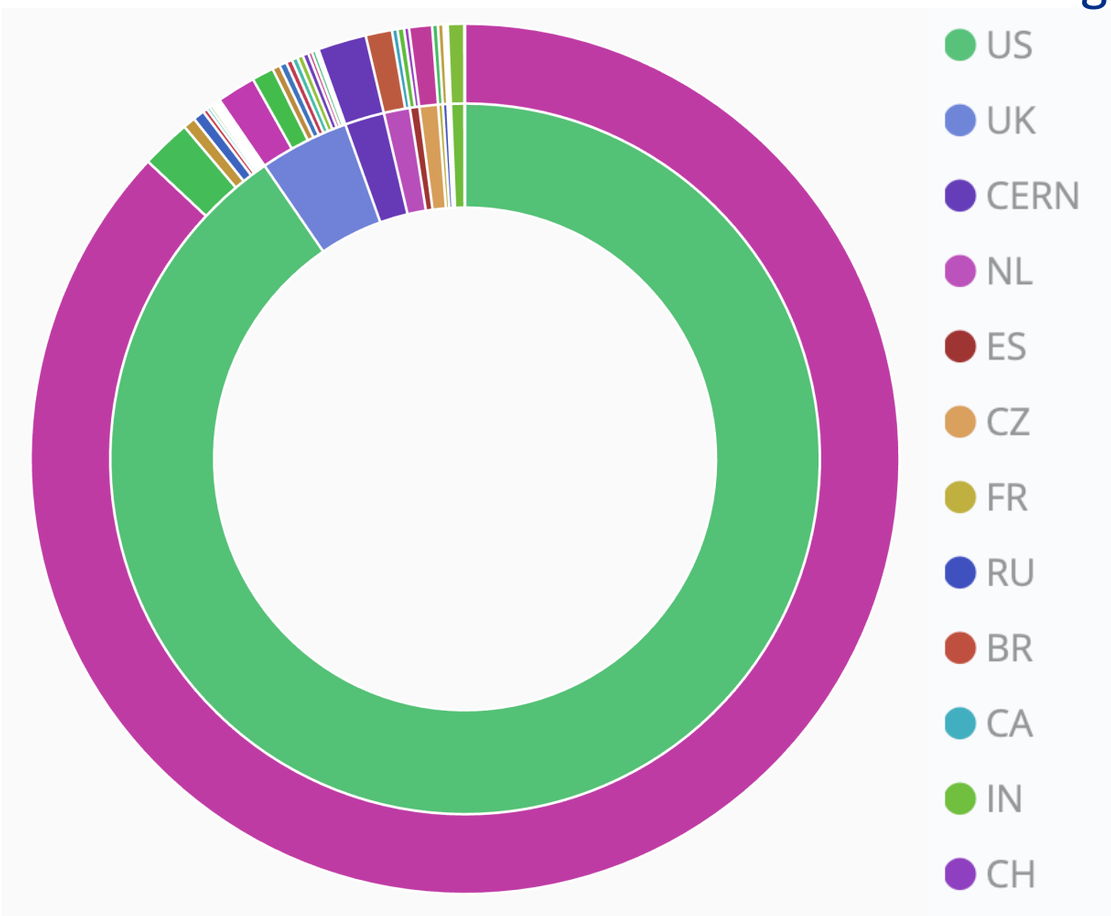
These are NOT current projections

CPU Usage (Wall hours) by Production and Analysis - 2021

Production

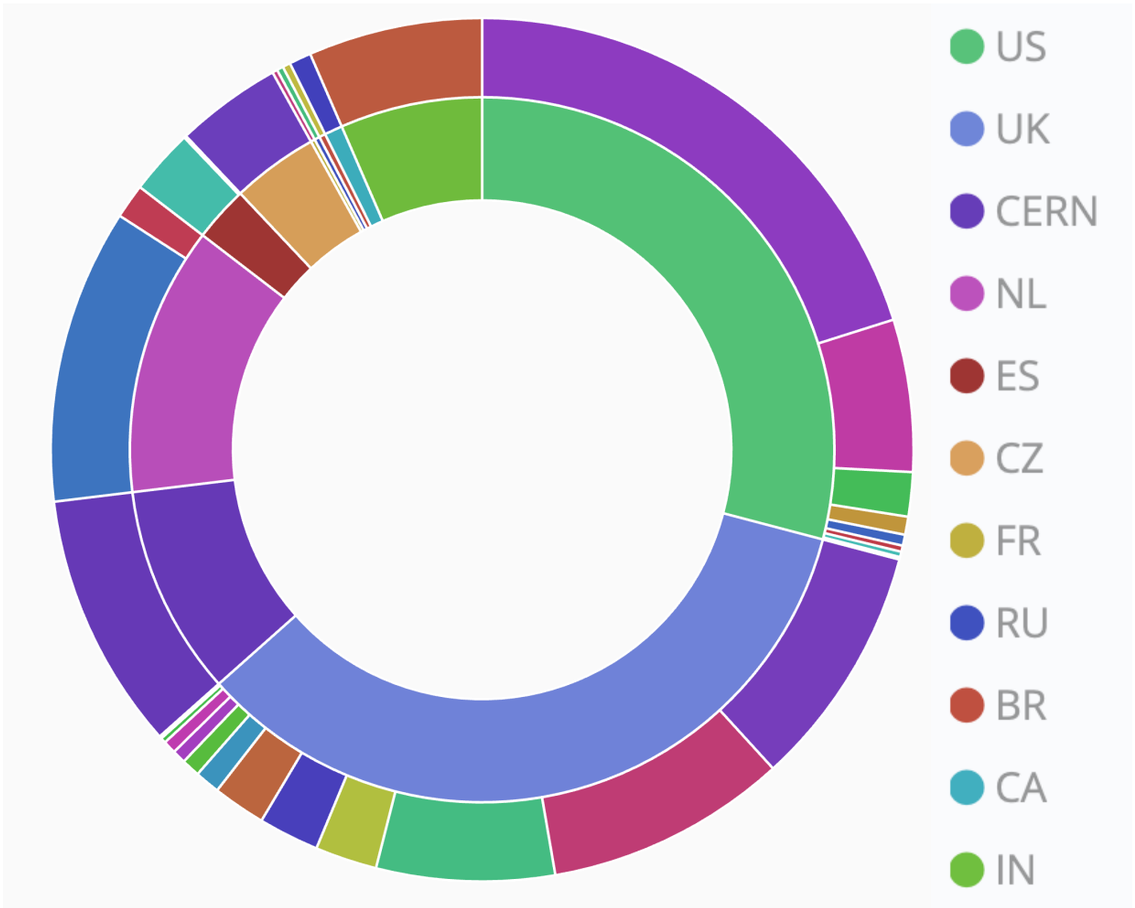


Analysis

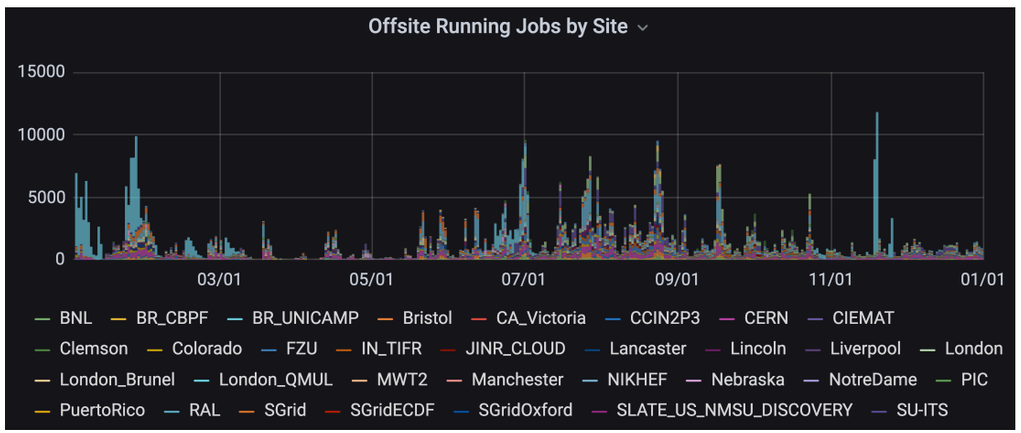


66.7M Slot weighted hours / 41.2M Wall Hours

Recent CPU Usage (Wall Hours) by site - Production



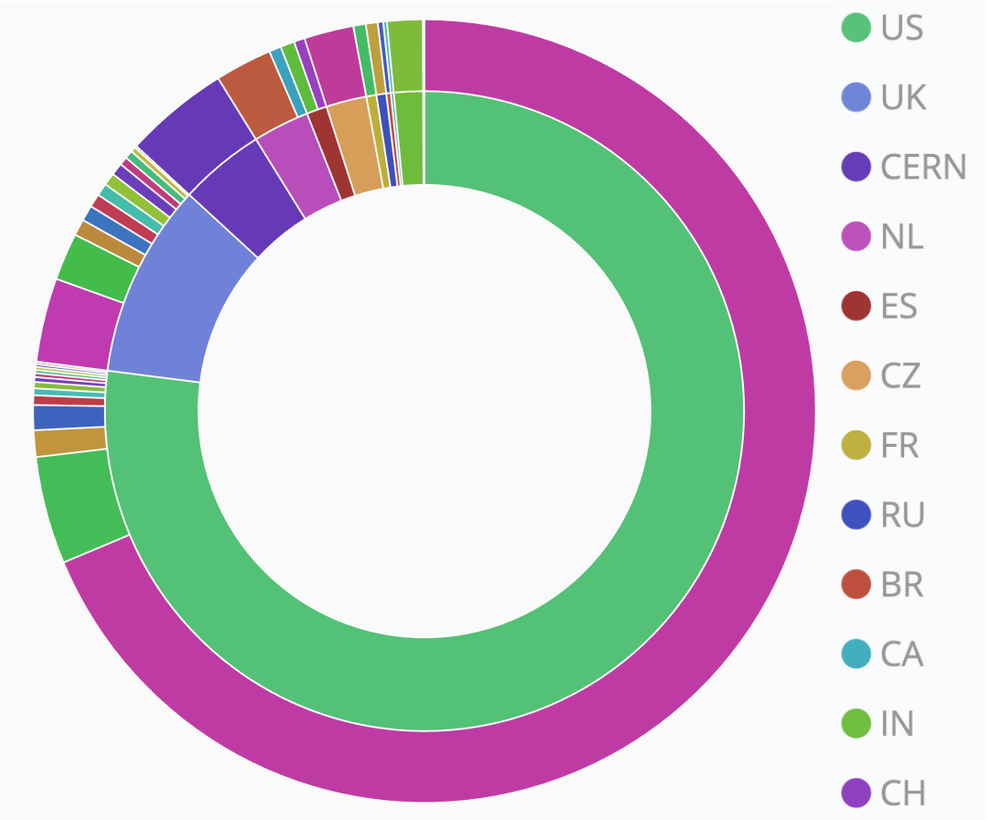
- production CPU utilization for Dec 1, 2021 - Jan 22, 2022
- balance of CPU is close to target of 25% - 75%
- removed NERSC HPC processing
- analysis slightly different



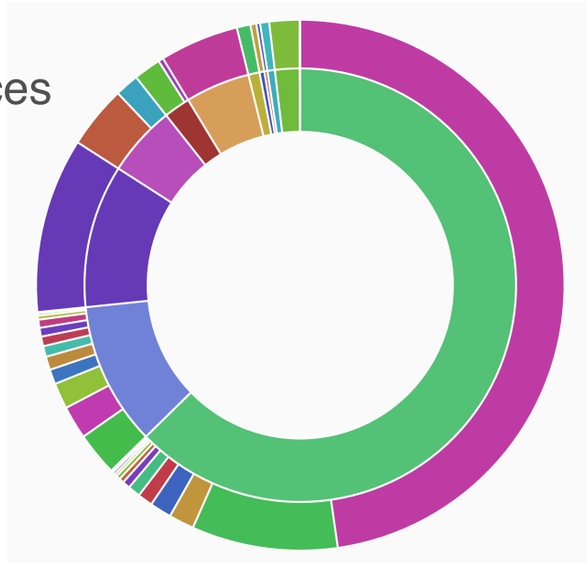
Fermilab

Analysis CPU (Wall hours) excluding MARS jobs

Analysis



- MARS software is under export control and so has a limited number of site where it can run
- almost all MARS jobs end up restricted to running on FermiGrid
- Exclude those jobs from Analysis submissions, and distribution of jobs over year, much broader
  - 68.5% at FermiGrid
  - 31.5% at other resources
- restricting only to jobs that have been run after Dec 2021 then FermiGrid is < 50%



Analysis Dec ‘21 - Jan ‘22

Fermilab

Pledged, deployed, and utilized resources (CPU cores)

CPU Cores	Site/Cluster	Pledged	Deployed	Utilized
FNAL	FermiGrid	4000	4000	3541
CERN	Tier-0	1650	1650	133
BNL	BNL	100	100	15
USA	OSG - opportunistic	1150	***	182
UK	GridPP	1000	1000	394
FR	CC-IN2P3	310	310	16
ES	PIC Tier-1	500	500	25
NL	NL/LHC Tier-1	696	696	81
CZ	CZ-Prague-T2	1560	1560	58
CH	Bern	200	200	—
BR	CBPF	100	100	8
IN	Tata	450	450	33
RU	JINR	N/A	N/A	7
Total		11716	10566	4685

"core" based on wall hours, year averaged, e.g. 8760 CPU Wall hour = 1 core

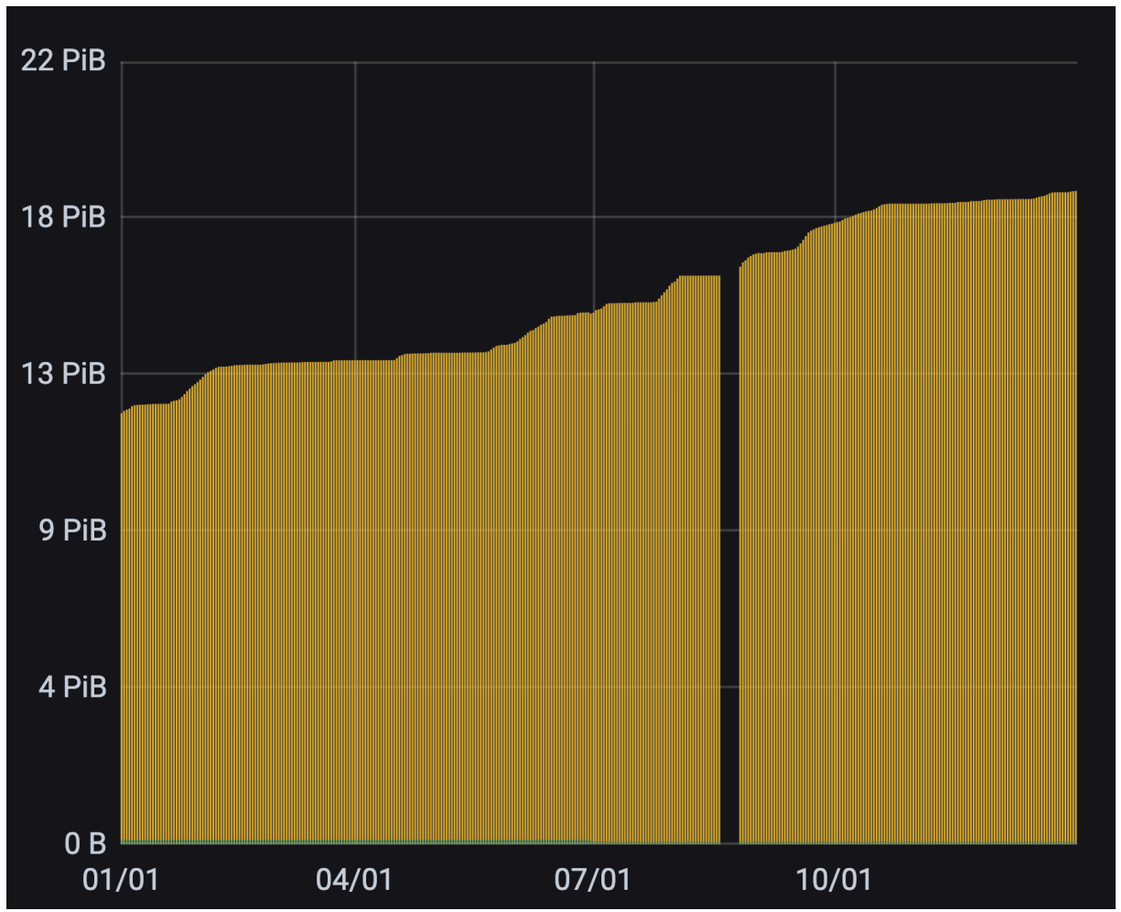
Pledges taken from here: https://docs.google.com/spreadsheets/d/180JGdW4ezFfx9BEzMRlzTDPbYjjlk27U/edit#gid=1610319714

Pledged, deployed, and utilized resources (Disk in TB)

Disk	Pledged	Deployed in Rucio	Volume in Rucio
FNAL	2200	4800	4800
CERN	2200	0	975
BNL	500	0	0
USA	0	***	0
UK	4000	3136	2178
FR	500	0	0
ES	500	0	0
NL	1900	207	0
CZ	300	0	300
CH	0	0	0
BR	0	0	0
IN	750	0	0
RU	0	0	0
Total	12850	10343	8253

Pledges taken from here: https://docs.google.com/spreadsheets/d/180JGdW4ezFfx9BEzMRlzTDPbYjjlk27U/edit#gid=1610319714

Tape usage at FNAL



Fermilab

Pledged, deployed, and utilized resources (Tape in TB)

Tape	Pledged	Deployed in Rucio	New Volume in Rucio
FNAL	9000	19260	6300
CERN	9000	5100**	71
BNL	0	0	0
USA	0	0	0
UK	3000	0	0
FR	2000	0	0
ES	0	0	0
NL	0	0	0
CZ	0	0	0
CH	0	0	0
BR	0	0	0
IN	0	0	0
RU	0	0	0
Total	23000	24360	11300

Pledges taken from here: https://docs.google.com/spreadsheets/d/180JGdW4ezFfx9BEzMRlzTDPbYjjlk27U/edit#gid=1610319714

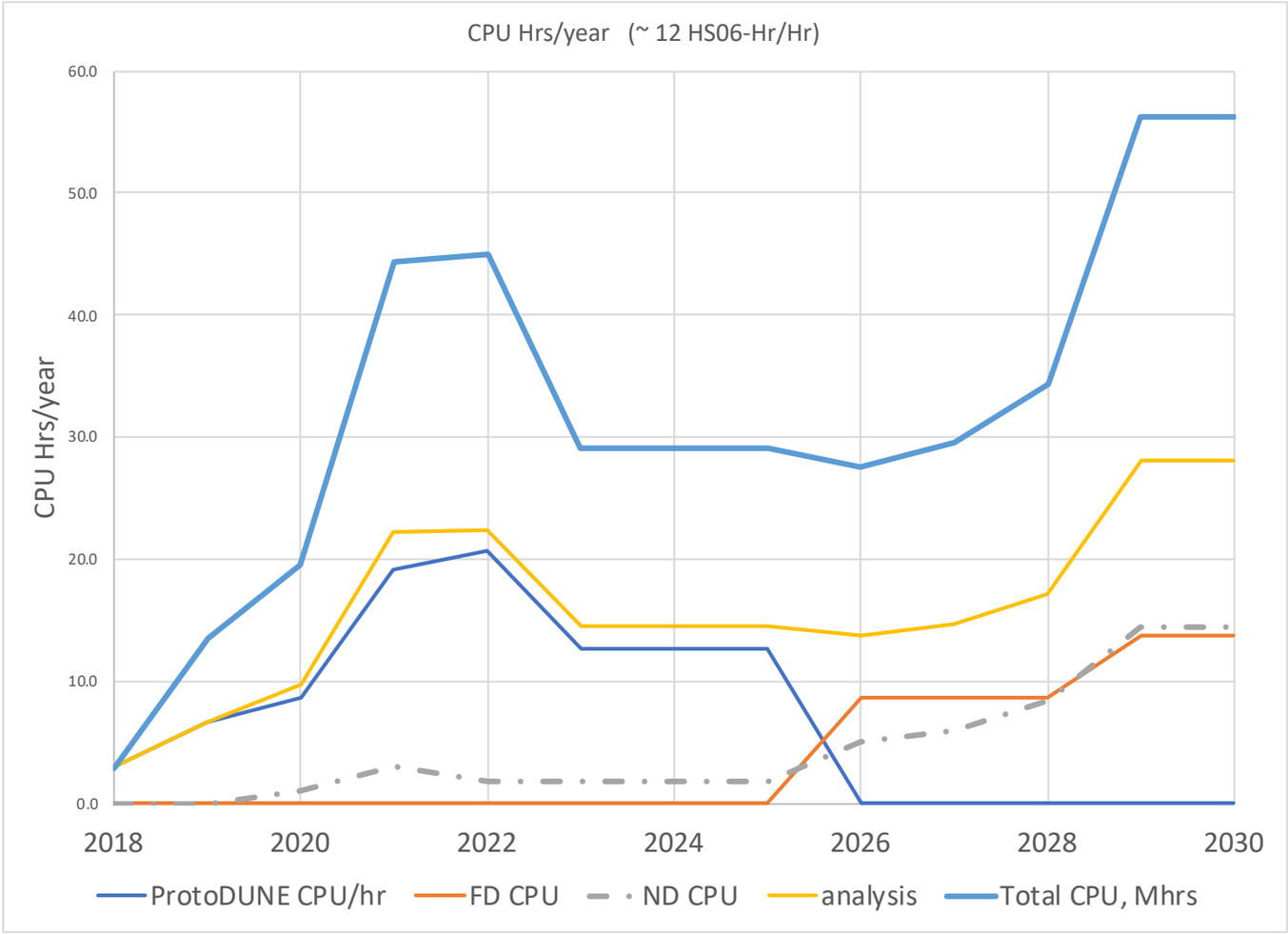
Summary

- Integration of CPU moves much more quickly
  - not a surprise as processing is much more flexible
  - integrating a new site can take as little as one day
- Integration of storage elements is much more challenging
  - stability of Rucio service at FNAL was an issue - now resolved
  - implementation of data replication plan just recently started
    - limited by throughput and non-optimization of tape reads
  - adapting data delivery to take advantage of distributed disk volumes
  - data challenges proposed
  - new Data Management convener (Doug Benjamin, BNL)
- Tape need was close to projection, but failed to utilize non-FNAL/non-CERN RSEs

Discussion

CPU Resources 2018 - 2030

- ProtoDUNE CPU needs dominate for the foreseeable future
- May actually be larger than FD requirements
- Near Detector CPU needs for reconstruction and simulation become significant in 2026 and will content with FD CPU needs



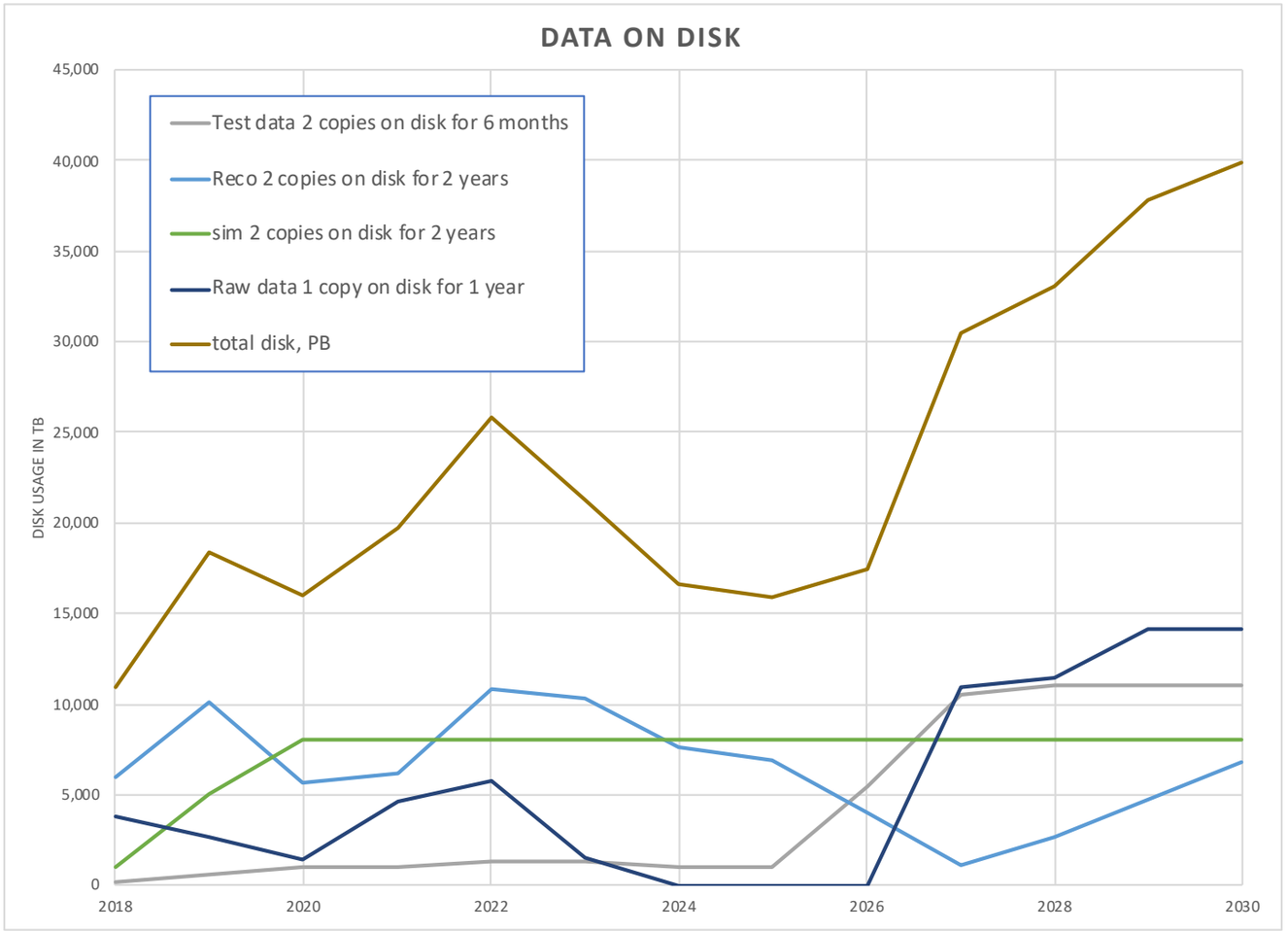
Assumptions from Nov 2020

These are NOT current projections

Fermilab

Summary

- Proposed resources needs for 2020 - 2022 were reviewed by Fermilab CRSG
- estimates based upon experience from ProtoDUNE, FD TDR, and ND CDR development
- tape and disk estimates may be the largest driver of dedicated resources
- CPU needs on the WLCG/ OSG are being met currently



Assumptions from Nov 2020

These are NOT current projections

Fermilab

How do we estimate our raw data volume?

ProtoDUNE II Data Volume

- uncompressed SP data Run I - 178 MB/evt
  - expected from 3 ms of 6 APAs
  - 20% overhead - other PDS and headers
  - compressed SP data - 71 MB/evt
  - compression factor of 2.5
- uncompressed DP data Run I - 110 MB/evt
  - 2 CRPs readout - 4 CRPs full detector
  - compression factor of 10 expected in 2021-2022

Far Detector Data Volume

- Single Phase FD Module scaled from the ProtoDUNE SP
  - scale readout 6 —> 150 APA
  - scale readout 3 ms —> 5.4 ms
  - compression factor of 2.5 for standard operations
  - assume start commissioning of one SP-FD module in 2026
  - assume start commissioning of one DP-FD module in 2028

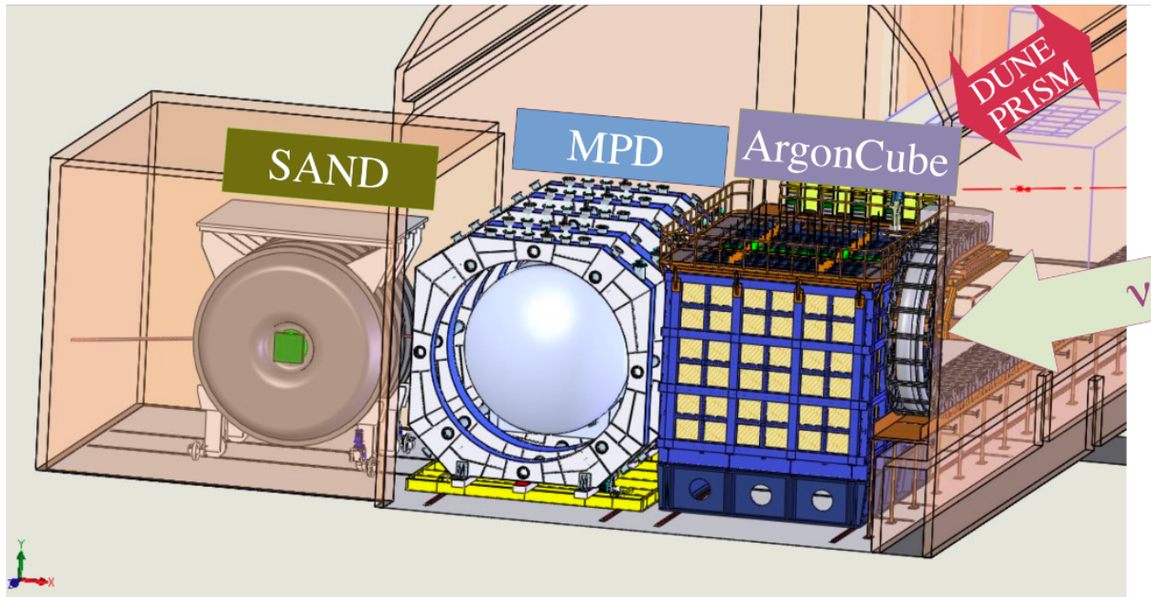
Assumptions from Nov 2020

These are NOT current projections

How do we estimate our raw data volume?

Near Detector Data Volume

- incorporated from the ND Conceptual Design Report
- ND LAr Detector
  - TPC - 3 MB/spill
  - Photon Detector System - 5 MB/spill
  - assume compression of 3 - 2.6 MB/spill
- ND GAr Detector
  - Gas TPC - 2 MB/spill
  - ECAL - 1 MB/spill
  - both uncompressed



- SAND 3D Scintillator Tracker
  - 0.3 MB/spill
- 10 MB/spill raw data total
- assumptions 5% downtime of each subdetector

Assumptions from Nov 2020

These are NOT current projections

Fermilab

Reconstructed data volume and processing

- Reconstruction algorithms for ProtoDUNE are well defined (three successful campaigns)
- dropping waveforms/raw data after reconstruction leads 10x reduction from raw volume - FD 100x reduction
- run time is 600 s/evt for both SP and DP
  - 180 s/evt his finding
  - 420 pattern recognition
  - matching performance found on WLCG/ OSG worker nodes
- ND reco data estimate equal to raw data
- ND reconstruction estimate 172 s/evt

- ProtoDUNE Simulation - 10 M evts
  - 2700 s/evt current experience
  - 200 MB/evt current experience
- FD simulation - 10 M evts
  - 2700 s/evt from ProtoDUNE
  - 200 MB/evt
- ND simulation
  - 25 M evts 2021
  - 10 M evts - 2022 - 2025
  - 50 M evts - 2026 —>2030
  - 300 s/evt (2.5 x CPU for data)
  - 20 MB (2x data reco/evt)

Assumptions from Nov 2020

These are NOT current projections

ProtoDUNE SP and DP Operations and Raw Data

protoDUNE Single Phase	2020	2021	2022
cosmic rate (Hz)	1	1	1
beam rate (Hz)	0	5.6	5.6
uptime (days)	150	150+50	300+50
events	6.5 M	24 M	30 M
event size (MB)	173	173	173
compression	2.5	2.5	2.5
yearly data (TB)	450	1625	2070

protoDUNE Dual Phase	2020	2021	2022
cosmic rate (Hz)	5	1	1
beam rate (Hz)	0	18.3	18.3
uptime (days)	100	150+50	300+50
events	21 M	62 M	65 M
event size (MB)	110	110	110
compression	10	10	10
yearly data (TB)	240	681	710

Assumptions from Nov 2020

These are NOT current projections

DUNE ND Simulation Data Volume

Near Detector Simulation	2020	2021	2022
events	10 M	25 M	10 M
event size (MB)	20	20	20
yearly data (TB)	200	500	200
CPU (MHrs)	1	3	1.2

Near Detector Prototypes	2022	2023	2024
events	10 M	10 M	0
event size (MB)	10	10	10
yearly data (MB)	100	100	0
test data (MB)	300	300	0
derived data (MB)	200	200	200
CPU (MHrs)	0.3	0.3	0.3

Assumptions from Nov 2020

These are NOT current projections

DUNE CPU requirements

Assumptions from Nov 2020

These are NOT current projections

protoDUNE Single Phase	2020	2021	2022
data events	6.5 M	24 M	30 M
CPU (hr)	1 M	4 M	5 M
MC events	5 M	5 M	5 M
CPU (hr)	3.8 M	3.8 M	3.8 M
analysis (hr)	5 M	8 M	9 M
Total CPU (hr)	13.8 M	16 M	17 M
Total CPU (HS06 years)	15750	18260	19400

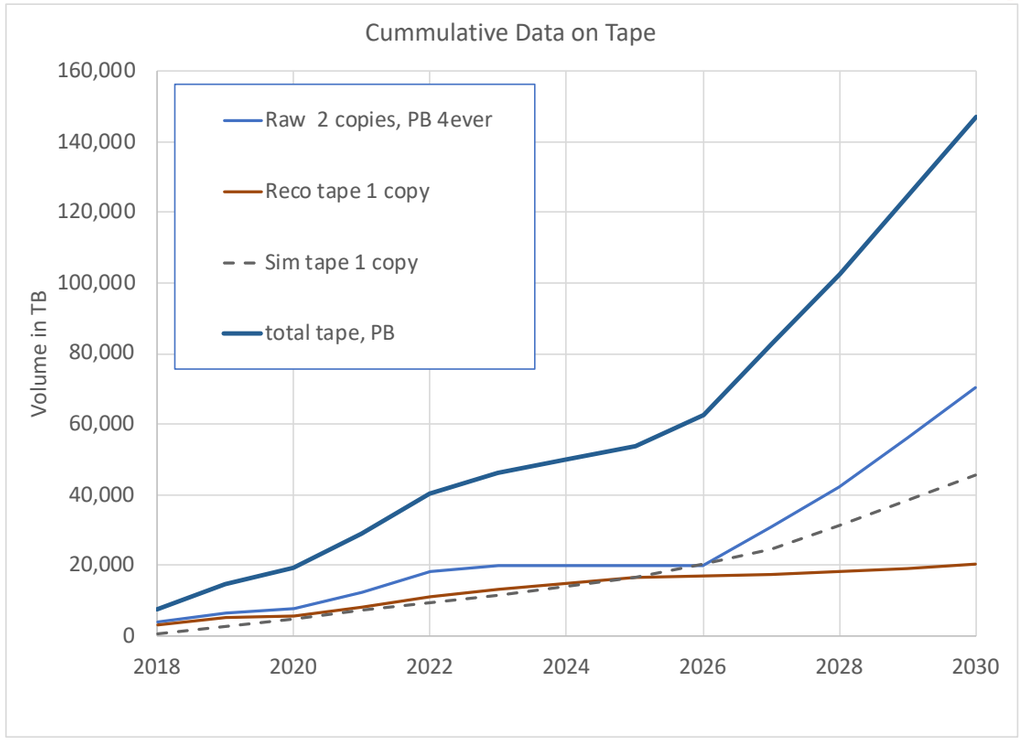
protoDUNE Dual Phase	2020	2021	2022
data events	21 M	62 M	64 M
CPU (hr)	3.5 M	10.5 M	10.8 M
MC events	5 M	5 M	5 M
CPU (hr)	3.8 M	3.8 M	3.8 M
analysis (hr)	8 M	14.3 M	14.6 M
Total CPU (hr)	15.3 M	28 M	29 M
Total CPU (HS06 years)	17470	32000	33100

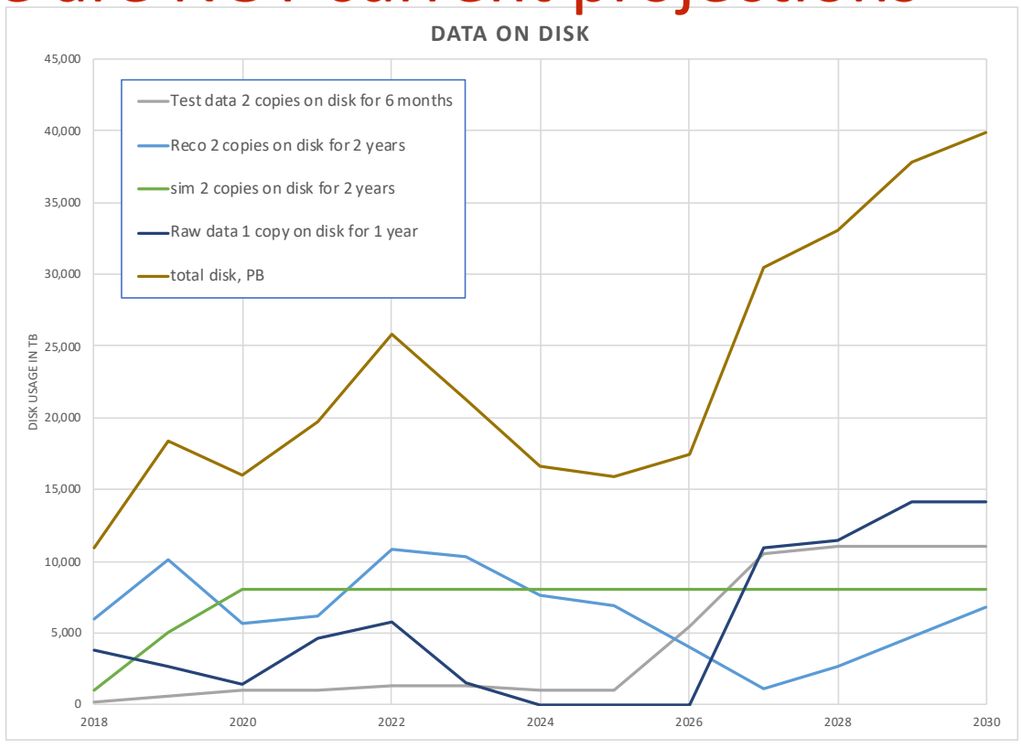
Fermilab

- SP Data 0.16 hr/evt - SP MC 0.75 hr/evt - DP 0.16 hr/evt - DP 0.75 hr/evt

Tape and disk storage 2018-2030

Assumptions from Nov 2020
These are NOT current projections



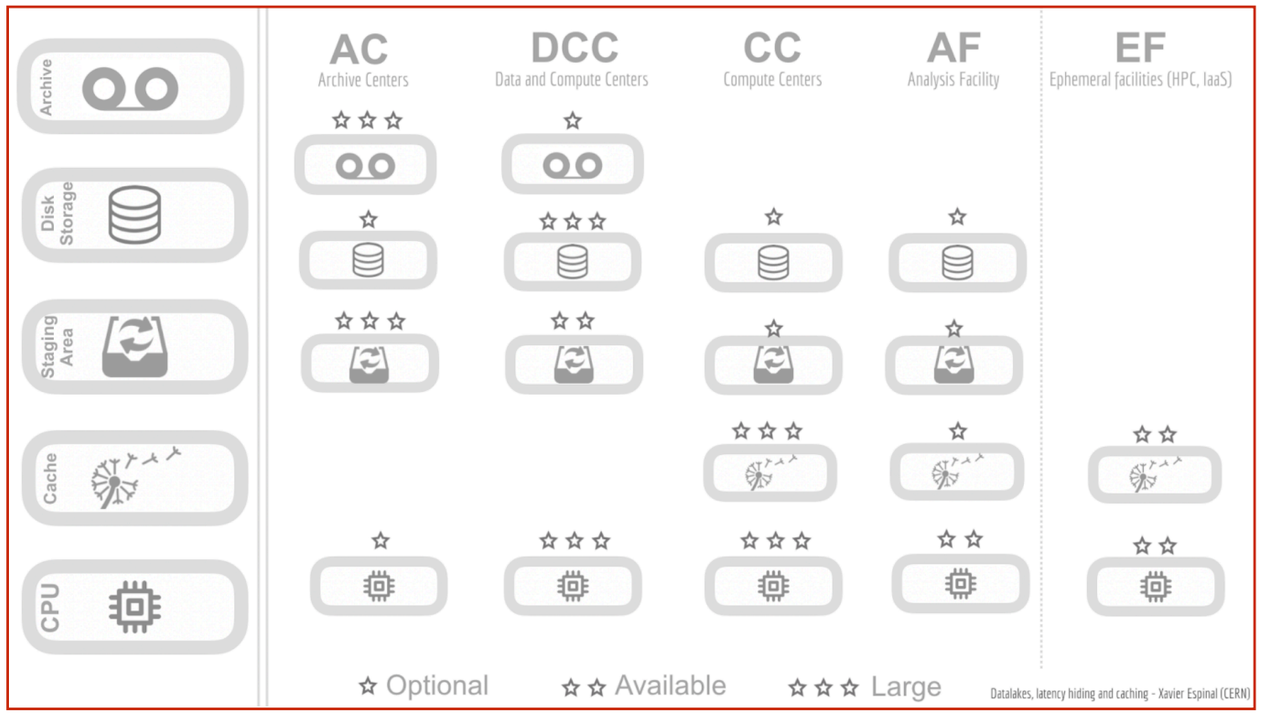


- Computing Model for DUNE Storage
  - 2 archival copies of raw, 1 for derived & simulated data - 1 copy at FNAL, second copy distributed institutions
  - production processing of SP and DP data twice per year and once matching simulation
  - 4 copies of active derived and simulated datasets on disk - dataset stays active for 2 year

Fermilab

DUNE Computing Model for Institutional Sites

- Based the HSF DOMA model for sites
  - Archive Center (AC) - tape/staging
  - Data and Compute Center - disk + CPU
  - Computer Center (CC) - CPU + cache
  - Analysis Facility (AF) - cpu + cache
  - Ephemeral Facilities (EF) - (HPC, IaaS)
- Goal is to have resource split between FNAL and other institutions - 25%/75%
- FNAL has some custodial responsibilities from the Dept of Energy that make this not possible for tape



Data Access in DOMA, HSF/OSG/WLCG Joint Workshop J-LAB Newport News, VA 19-23 March 2019

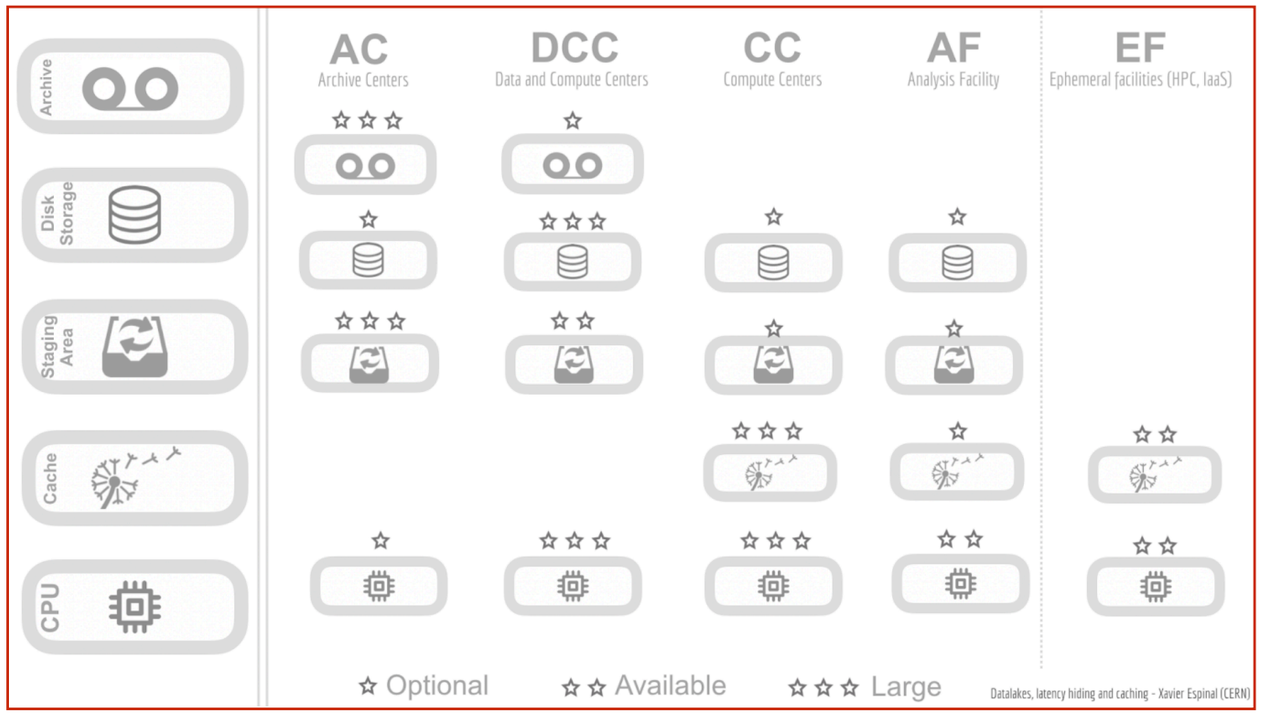
Assumptions from Nov 2020

These are NOT current projections

Fermilab

DUNE Computing Model for Institutional Sites

- Simplified terms for current DUNE sites
  - Tape Site - tape/staging
  - Disk Site - disk + CPU
  - Compute Site - CPU + cache
  - Analysis Site - cpu + cache
  - HPC - (HPC, IaaS)
- Goal is to have resource split between FNAL and other institutions - 25%/75%
- FNAL has some custodial responsibilities from the Dept of Energy that make this not possible for tape



Data Access in DOMA, HSF/OSG/WLCG Joint Workshop J-LAB Newport News, VA 19-23 March 2019

Assumptions from Nov 2020

These are NOT current projections