DUNE Offline Computing Model Calculations

H. Schellman for the Computing Consortium

November 27, 2022

Draft

1 Introduction

This is an annual projection for DUNE CPU and storage needs intended for use at the Computing Contributions Board meeting in December 2022. It projects needs for 2023 onwards.

The overall computing model and 2022 projections for DUNE are described in chapters 6-13 of the recent (Oct. 2022) DUNE Conceptual Design Report [1]. This document provides updates on resource needs for 2023.

The 2023 projection is done using codes at: https://github.com/DUNE/CCB-data/tree/master/Numbers-2023 from parameters stored in a json file. We use CPU and storage sizes derived from protoDUNE and simulation experience and apply them to projected numbers of events from the various DUNE detectors.

Details are provided in the appendices while the main body of this note summarizes pledges, usage and projected need for the CCB.

Changes since the last report include:

- A later start for ProtoDUNE-2 running at CERN. This leads to a spike in needs for storage/CPU in late 2023.
- Use of memory-weighted-core time instead of CPU time as our codes often require more memory than is available for a single core. This requires the introduction of a memory-weighted wall-time unit for contributions as different sites will need to provide different amounts of wall-time to perform the same processing.
- Revisions to near-term requests based on the 2022 experience including a hold on tape requests from the collaboration during protoDUNE activities.
- Simulation disk copies has been reduced from 2 to 1.5 to fit within a reasonable profile. If more disk becomes available we can restore more simulation copies.
- A change in future tape requests to reflect the greater accessibility of tape archives at CERN and FNAL relative to other sites.

Proposed pledges for 2023 are detailed in sections: Disk(2.1.1), Tape(2.2.1) and CPU(3.1)below. A summary of requests for 2023 is shown here in Table 1.

	Disk (PB)	Modified Disk (PB)	Tape(PB)	CPU (MWC-years)
Model	25.80	25.80	45.60	12765
Request				
FNAL	7.80	8.86	36.30	3191
CERN	2.60	4.00	9.20	3191
National	15.40	12.94	0.10	6383
Total	25.80	25.80	45.60	12765

Table 1: Proposed pledges for 2023. Disk pledges are based on existing CERN and FNAL contributions with National contributions making up the rest of the model request. Tape pledges reflect the dominant use of CERN and FNAL for archival storage of data. CPU pledges are in units of memory-weighted-core-years and assume Fermilab and CERN each pledge 25%.

Draft

2 Disk and Tape

Generally, raw data are stored on tape at both CERN and FNAL. Simulation and reconstructed data have one tape copy at Fermilab and recent reconstructed and simulated samples have one (or two) disk copies with one at Fermilab and one in Europe. Appendix A gives details on the size and types of data from the SAM data catalog.

CERN and FNAL have special responsibilities for archival data storage and for disk space for raw data while contributions from other collaborating institutions are aggregated under the heading National, which includes US sites outside of FNAL. The traditional split between FNAL, CERN and National contributions until 2027 is shown in Table 2. The pledges proposed here deviate slightly from those numbers with larger contributions from the host laboratories in 2023 due to resources already in place.

		FNAL	CERN	National
Tape	Raw:	0.5	0.5	0.0
	Sim:	1.0	0.0	0.0
	Reco-Data:	1.0	0.0	0.0
	Test:	0.5	0.5	0.0

		FNAL: 0.5	CERN	National
Disk	Raw	0.50	0.50	0.00
	Sim:	0.25	0.0	0.75
	Reco-Data:	0.25	0.00	0.75
	Test:	0.50	0.50	0.00

Table 2: Proposed division between FNAL/CERN/National for storage until 2027.

2.1 Disk

Table 3 summarizes the disk utilization known to rucio, augmented with information gathered from individual sites at a meeting on 2022-11-21. Some sites, notably TIFR, are not yet fully integrated so do not show up in the rucio reports, while others (PIC) are still being filled. The contributions listed in Table 4 sum the rucio and non-rucio disk known to be allocated to DUNE.

Figure 1 summarize the cumulative disk needs and requests projected by our model. These numbers are used to generate the request for 2023. They are divided into the two host laboratories (CERN and FNAL) and National, which includes contributions from the rest of the collaboration, including OSG, BNL and NERSC in the US.

Table 4 summarizes the pledges from previous years compared to the actual amounts allocated and used from Table 3 . The 2023 request has been re-evaluated in light of underuse in 2021 and 2022 and should better match the likely capacity of the collaborating sites. It is however, still higher than 2022 due to ProtoDUNE running and increased simulation for the far and near detectors.

Draft

RSE	Rucio Allocation (PB)	Rucio Used (PB)	percent full	Non Rucio (PB)	Total Alloc
CA				0.050	0.050
CERN_EOS	4.000	2.500	63%		4.000
CZ_PRAGUE	1.130	0.514	45%		1.130
ES_DUNE_ES_PIC	0.719	0.005	1%		0.719
FR_IN2P3	0.500	0.128	26%		0.500
IN_TIFR	0.100	0.000	0%	0.000	0.100
NL_SURFSARA	0.900	0.001	0%		0.900
NL_NIKHEF	1.010	0.417	41%		1.010
RU_JINR				0.500	0.500
UK Total	3.828	3.121	82%		3.828
UK_LANCASTER	0.549	0.503	92%		0.549
UK_MANCHESTER	1.080	1.000	93%		1.080
UK_QMUL	1.100	0.679	62%		1.100
UK_RAL-PP	0.099	0.099	100%		0.099
UK_RAL_ECHO	1.000	0.840	84%		1.000
US_BNL	1.000	0.501	50%		1.000
US_FNAL	7.029	4.654	66%	1.829	8.858
National	9.237	4.687	2.448	0.500	9.687
CERN_EOS	4.000	2.500	0.625	0.000	4.000
FNAL	7.029	4.654	0.662	1.829	8.858
Total	24.044	14.962	8.035	2.379	22.595

Table 3: Disk allocations and usage across sites. These numbers are derived from rucio reports and from cross-checks with individual sites on 2022-11-21.

Draft

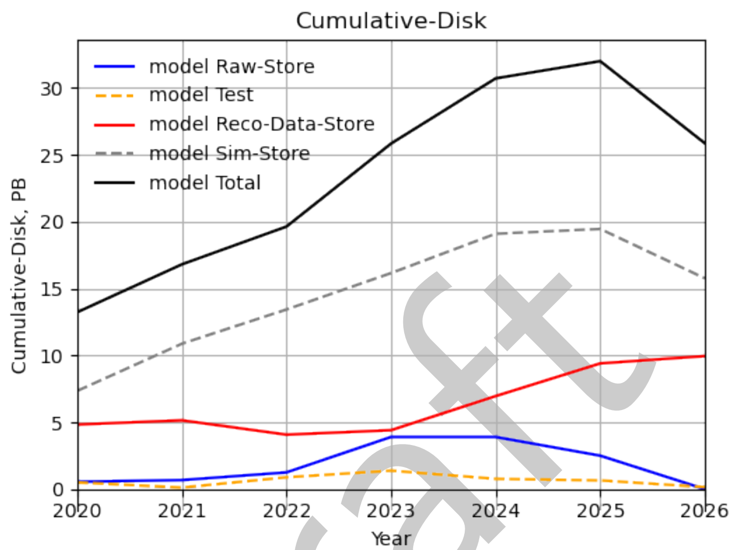
2.1.1 Conclusion

The overall request for 2023 is 25.8 PB vs. the 22.6 PB already on the floor so we need to find 3.2 PB or descope the number of copies on disk. Currently CERN and FNAL contribute more (12.9 PB on the floor vs. 9.1 PB request) than they would under the current divisions in Table 2 while national contributions are currently 9.7 PB vs 15.4 in the request based on the allocations in Table 2. A suggestion is to ask the collaborating institutions for the extra 3.2 PB of additional disk instead of the 5.7 PB that the current division of responsibility would suggest. This is shown in Table 4 as a modified request.

	2021	2021	2022	2022	2022	2023	2023
	Pledge (PB)	Disk Actual	Pledge (PB)	Disk Alloc (PB)	Disk Used	Standard Request	Modified Request
BR	0.00						
CA				0.05	0.05		
CH	0.20		0.20				
CZ	0.30		1.00	1.13	0.51		
ES	0.50		0.72	0.72	0.01		
FR	0.50		0.50	0.50	0.13		
IN	0.75		0.75	0.10	0.00		
IT							
NL	1.90		1.90	1.90	0.42		
RU			0.50	0.50	0.50		
UK	4.00		4.00	3.83	3.12		
US BNL	0.50		0.50	1.00	0.50		
US - other							
National	8.65	0.00	10.07	9.73	5.24	15.40	12.94
CERN	2.20		3.00	4.00	2.50	2.60	4.00
FNAL	2.20		7.60	8.86	8.85	7.80	8.86
Total	13.05	0.00	20.67	22.59	16.59	25.80	25.80

Table 4: Summary of disk pledges, allocations and usage for 2021-2022 with model request for 2023. This is based on the 2022 CCB tables which are available in indico [2, 3]. These numbers are derived from the rucio reports in Table 3 and may not be complete.

Draft



Cumulative-Disk	2020	2021	2022	2023	2024	2025	2026
Raw-Store(PB)	0.6	0.7	1.2	3.9	3.9	2.5	0.0
Test(PB)	0.5	0.1	0.9	1.4	0.8	0.7	0.1
Reco-Data-Store(PB)	4.8	5.1	4.1	4.4	7.0	9.4	9.9
Sim-Store(PB)	7.4	10.9	13.4	16.2	19.1	19.5	15.8
Total(PB)	13.2	16.8	19.6	25.8	30.7	32.0	25.

Cumulative-Disk	2020	2021	2022	2023	2024	2025	2026
National(PB)	9.1	12.0	13.1	15.4	19.5	21.6	19.3
FNAL(PB)	3.6	4.4	5.4	7.8	8.8	8.8	6.5
CERN(PB)	0.5	0.4	1.1	2.6	2.3	1.6	0.1
Total(PB)	13.2	16.8	19.6	25.8	30.7	32.0	25.

Figure 1: Cumulative Disk needs in PB. Includes data lifetimes. The top table shows the source of the data while the bottom table shows the proposed split using the fractions from Table 2.

Draft

2.2 Tape

Figure and Table 2 summarize the cumulative tape need projected by our model. These numbers are used to generate the requests for 2023. They are divided into the two host laboratories (CERN and FNAL) and National, which includes contributions from the rest of the collaboration, including OSG, BNL and NERSC in the US.

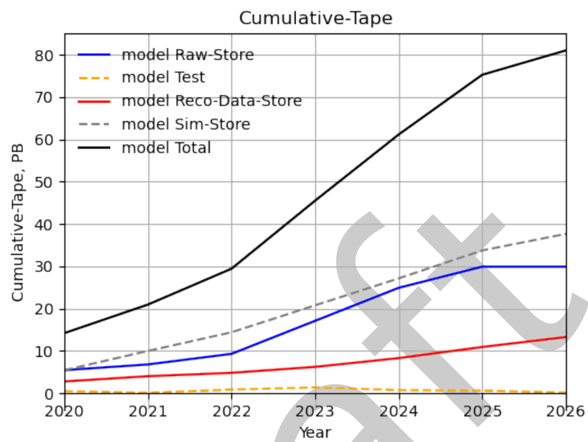
Table 5 shows pledges and utilization for 2021-2022 and the request for 2023.

2.2.1 Conclusion

DUNE currently has ~23 PB of data on tape at Fermilab and 5 PB of protoDUNE data as a second copy at CERN. We anticipate needing up to 45 PB of tape (an increase of 16 PB from 2022) to accommodate the ProtoDUNE run 2 data and increased simulation.

The UK and the IN2P3 have made ~ 3 PB of tape archive available but it has not yet been smoothly integrated into our data flow. We therefore substantially reduce our request for these resources to ~ 100 TB per site, to allow testing of integration, with an increased request anticipated future years.

Draft



Cumulative-Tape	2020	2021	2022	2023	2024	2025	2026
Raw-Store(PB)	5.5	6.8	9.3	17.1	24.9	29.9	29.9
Test(PB)	0.5	0.1	0.9	1.4	0.8	0.7	0.1
Reco-Data-Store(PB)	2.8	4.0	4.8	6.2	8.3	10.9	13.3
Sim-Store(PB)	5.5	10.0	14.4	20.8	27.1	33.7	37.6
Total(PB)	14.2	21.0	29.4	45.5	61.1	75.2	81.

Cumulative-Tape	2020	2021	2022	2023	2024	2025	2026
National(PB)	0.0	0.0	0.0	0.0	0.0	0.0	0.0
FNAL(PB)	11.2	17.5	24.3	36.3	48.3	60.0	66.0
CERN(PB)	3.0	3.5	5.1	9.2	12.8	15.3	15.0
Total(PB)	14.2	21.0	29.4	45.5	61.1	75.2	81.

Figure 2: Cumulative Tape needs in PB. Includes data lifetimes. The top table shows the source of the data while the bottom table shows the proposed split.

Draft

3 CPU Needs

DUNE differs from other HEP experiments in frequently requiring more memory/core than is available at particular sites. For example an 8-slot pilot with 16 G of available memory could only accommodate four reconstruction processes. As a result, we make our requests in terms of memory-weighted-core wall time (MWC) with the base memory being 2000 MB. Sites that offer more than 2000 MB/core can scale their contributions up by the local memory/core.

Wall-time estimates are created by estimating the number of simulated and raw events taken and then scaling by the measured CPU time on a gpvm corrected to wall-time by the estimated efficiency (default 70%) and for the memory utilization factor that takes into account the differing memory needs for different applications and the number of cores needed to satisfy those needs. Here we assume that analysis takes 3000 MB, reconstruction takes 4000MB, and simulation takes 6000MB. Contributions are then requested in units of MWC-time which is wall-time$\times$2000 MB units.

Figure/Table 3 shows the projected memory-weighted wall-time (MWChr) need projections through 2026. This is different than in previous years where memory weighting was not applied. They are divided into the two host laboratories (CERN and FNAL) and National, which includes contributions from the rest of the collaboration, including OSG, BNL and NERSC in the US.

Table 5 summarizes the pledges[2] and measured usage using FNAL's HTCondor memory-weighted wall-time statistics[4]. The usage numbers for 2022 are Nov 2021 to Oct 2022.

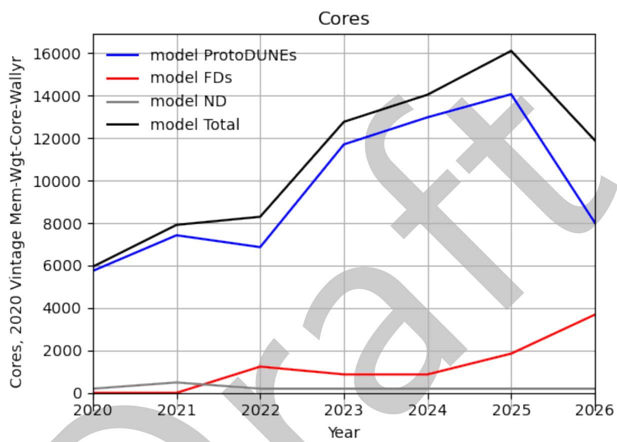
Table 6 summarizes the statistics for European sites from Nov 2021 to Oct 2022 derived from the EGI accounting[5] which uses the number of cores allocated to a pilot. If four 4000 MB reconstruction jobs were sent to an 8-core pilot on a system with 2000MB/core, this would be equivalent to using 8 MWC (memory-weighted wall-time units).

An example of a pledge in MWC might be 1000 cores with 2GB available/core, 500 cores with 4 GB available/core and 100 cores with 8 GB available/core. The MWC pledge would then be 2000+2000+800 or 4800 MWC units.

Draft

3.1 Conclusion

The advent of ProtoDUNE-2 running in 2023 and ramp-up of simulation for the FD and ND will lead to somewhat increased needs for CPU resources. For 2023, we are requesting $\sim$12,800 memory-weighted core-years (MWC-years) where 2000 MB/core is assumed to be 1 core.



Cores	2020	2021	2022	2023	2024	2025	2026
ProtoDUNEs (2020 Vintage Mem-Wgt-Core-Wallyr)	5746	7423	6863	11701	12979	14066	8018
FDs (2020 Vintage Mem-Wgt-Core-Wallyr)	0	0	1235	868	868	1846	3681
ND (2020 Vintage Mem-Wgt-Core-Wallyr)	195	489	195	195	195	195	195
Total (2020 Vintage Mem-Wgt-Core-Wallyr)	5941	7912	8294	12765	14043	16109	11896

Figure 3: Proposed memory-weighted wall-time needs in number of 2000 MB cores (MWC-years). Memory-weighted wall-time takes into account memory and efficiency.

Draft

CPU	2021 pledge	2021 use	2022 pledge	2022 MWC used	2023 request (MWC)
BR	100	19	200	16	
CA		6		40	
CH	200	1		0	
CZ	1560	128	2400	78	
ES	500	47	512	122	
FR	310	34	250	48	
IN	450	79	450	126	
IT	-	-		-	
NL	696	210	788	340	
RU		8	1000	39	
UK	1000	947	1000	912	
US OSG	1250	593	1000	354	
National	6066	2072	7600	2075	6383
CERN	3310	306	950	298	3191
US Fermilab	3310	3059	1945	2966	3191
Total	12686	5437	10495	5339	12765
Model		7912		8294	12765

Table 5: Summary of DUNE wall-time pledges and contributions for 2021 and 2022. The 2022 actual numbers are memory-weighted. Individual nations are listed and then merged (with US OSG) into a National section.

Draft

Nation	Slot-hrs	Slot-years	Percent
Brazil	114904	13.1	0.7
Canada	3163	0.4	0
Czech Republic	1053867	120.3	6.1
France	871042	99.4	5
Italy	14295	1.6	0.1
Netherlands	1668058	190.4	9.6
Poland	58904	6.7	0.3
Portugal	47529	5.4	0.3
Russia	43757	5	0.3
Spain	1769349	202	10.2
Switzerland	21776	2.5	0.1
Turkey	6544	0.7	0
United Kingdom	11712533	1337	67.4
Total	17385721	1984.7	100

Draft

Table 6: Summary of DUNE memory-weighted-core hours from European collaborators, Nov. 21 to Oct. 22, using the EGI accounting[5]. These numbers differ, and are generally higher, than the FNAL numbers in the previous table.

References

[1] **DUNE** Collaboration, A. Abed Abud *et al.*, "DUNE Offline Computing Conceptual Design Report," `arXiv:2210.15665 [physics.data-an]`.

[2] P. Clarke, "DUNE Computing Contributions Board Meeting for 2022." https://indico.fnal.gov/event/52519/.

[3] P. Clarke, "DUNE Computing Contributions Board Meeting for 2023." https://indico.fnal.gov/event/52519/.

[4] K. Herner, "Kibana sum of slot hours for DUNE with MARS removed, Nov 1, 2021 to Oct 31, 2022." https://fifemon.fnal.gov/kibana/goto/5ac70f9ccc22e447632c7276cea83ef5.

[5] "EGI accounting for DUNE slot time, November 2021 to October 2022. slot time refers to the core allocated to pilots. it is partially equivalent to the memory-weighted wall-time from htcondor if the pilot fills the requested memory.." https://accounting.egi.eu/wlcg/countries/elap_processors/COUNTRY/DATE/2021/11/2022/11/custom-biomed,dune/localinfrajobs/.

Draft

A Information about storage from SAM

This section provides information on the sizes of data samples known to the SAM data catalog as of Nov. 1, 2022. If a file has multiple copies, that is not shown here. Tables 7 and 8 show the total across all streams and data tiers while table 9 shows the distribution of the largest samples.

expt	size	TB	size/event	MB
fardet	14.252	TB	0.358	MB
neardet	76.176	TB	0.178	MB
physics	1540.003	TB	35.161	MB
fardet-hd	45.097	TB	1.511	MB
fardet-sp	171.612	TB	2.4	MB
fardet-vd	1837.767	TB	31.109	MB
protodune-sp	6693.574	TB	90.061	MB
Total	10378.48	TB	13.903	MB

Table 7: Summary of total simulation in SAM by detector type as of Nov 1, 2022.

expt	size	TB	size/event	MB
iceberg	3.339	TB	3.461	MB
hd-coldbox	79.891	TB	40.247	MB
protodune-dp	413.914	TB	135.997	MB
protodune-sp	8798.856	TB	29.891	MB
vd-coldbox-top	319.305	TB	34.291	MB
vd-coldbox-bottom	26.843	TB	23.858	MB
dc4-vd-coldbox-top	102.972	TB	-1	MB
dc4-vd-coldbox-bottom	211.359	TB	69.149	MB
Total	9956.601	TB	31.721	MB

Table 8: Summary of total detector data in SAM by detector type as of Nov 1, 2022.

Draft

year	det/mc	expt	stream	tier	size (TB)	size/event (MB)
ALL	mc	protodune-sp	out1	detector-simulated	3509	238.8
ALL	mc	protodune-sp	out1	full-reconstructed	3093	177.2
ALL	detector	protodune-sp	cosmics	raw	2310	51.4
ALL	detector	protodune-sp	physics	full-reconstructed	2003	66.4
ALL	mc	fardet-vd	out1	full-reconstructed	1826	62.4
ALL	detector	protodune-sp	cosmics	full-reconstructed	1449	45.9
ALL	detector	protodune-sp	physics	raw	864	83.6
ALL	detector	protodune-sp	test	raw	780	35.6
ALL	mc	physics	out1	full-reconstructed	735	29.5
ALL	mc	physics	out1	detector-simulated	537	39.8
ALL	detector	protodune-sp	physics	reco-recalibrated	360	37.9
ALL	mc	physics	out1	simulated	268	52.9
ALL	detector	protodune-dp	test	raw	259	124.2
ALL	detector	vd-coldbox-top	study	raw	225	49.4
ALL	detector	dc4-vd-coldbox-bottom	test	raw	211	69.1
ALL	mc	fardet-sp	out1	full-reconstructed	169	3.7
ALL	detector	protodune-sp	test	full-reconstructed	151	37.4
ALL	detector	protodune-dp	cosmics	raw	140	326.4
ALL	detector	protodune-sp	calibration	raw	137	4.0
ALL	detector	protodune-sp	physics	hit-reconstructed	127	50.0
ALL	detector	protodune-sp	physics	decoded-raw	125	49.2
ALL	detector	protodune-sp	cosmics	reco-recalibrated	107	22.5
ALL	detector	dc4-vd-coldbox-top	test	raw	103	-1.0
ALL	detector	hd-coldbox	test	raw	80	40.2
ALL	mc	neardet	physics	simulated	76	0.2
ALL	detector	protodune-sp	physics	pandora_info	75	6.9
ALL	detector	vd-coldbox-top	test	raw	60	28.8
ALL	mc	protodune-sp	out1	pandora_info	57	6.0
ALL	detector	protodune-sp	cosmics	hit-reconstructed	42	25.9
ALL	detector	protodune-sp	noise	decoded-raw	40	27.0
ALL	detector	protodune-sp	cosmics	decoded-raw	40	24.6
ALL	detector	protodune-sp	cosmics	pandora_info	39	6.0
ALL	mc	fardet-hd	out1	full-reconstructed	36	5.1
ALL	detector	protodune-sp	noise	raw	30	30.6
ALL	detector	vd-coldbox-bottom	test	raw	27	23.9
ALL	detector	vd-coldbox-top	test	full-reconstructed	26	12.8

Draft

Table 9: Classification of the largest data samples in SAM. They are classified as detector(data) or mc, by the detector producing the data, by the stream (readout time) and by the data tier. Some types, test and noise for example are archival only.

B Model Details

This appendix shows the parameters used in the model and plots of all the input and derived quantities as a function of time.

Resource needs for reconstructed data for a given year are based on the number of events produced over the previous "Reprocess" years. For ProtoDUNEs that is 2-4 years.

Simulation resource needs are instead calculated based on a number of simulation events each year. The assumption is that new software versions imply resimulation.

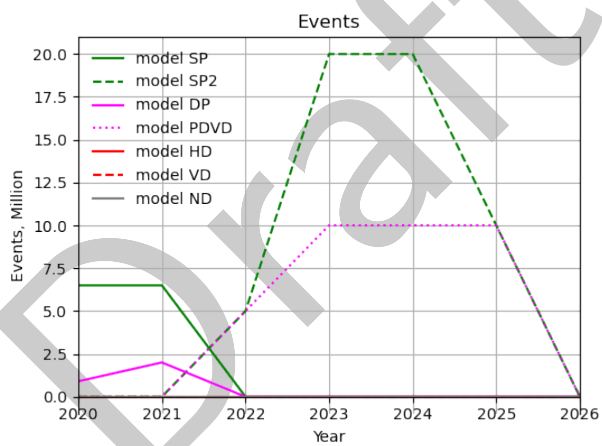
Disk and tape lifetimes for different data types are specified as well as the desirable number of copies.

The splits parameters make CERN responsible for raw data until 2027 with the collaboration taking over after that point.

`Detectors`: Detectors included in the calculation = ['SP', 'SP2', 'DP', 'PDVD', 'HD', 'VD', 'ND']
`Cap`: Cap on Raw data/year in PB = 30
`Base-Memory`: MB of memory per slot assumed as the average = 2000
`MaxYear`: Plot until year = 2026
`MinYear`: Plot starting with year = 2020
`Reprocess`: Number of years of data reprocessed when doing a new pass = 'SP': 3, 'DP': 2, 'SP2': 4, 'PDVD': 4, 'ProtoDUNEs': 3, 'VD': 100, 'HD': 100, 'FDs': 100, 'ND': 100, 'MARS': 1
`PatternFraction`: Fraction of time taken in pattern recognition = 'SP': 0.7, 'SP2': 0.7, 'DP': 0.7, 'PDVD': 0.7, 'ProtoDUNEs': 0.7, 'HD': 0.1, 'VD': 0.1, 'FDs': 0.1, 'ND': 0.9, 'MARS': 0
`TapeLifetimes`: Number of years kept on tape = 'Raw-Store': 100, 'Test': 0.5, 'Reco-Data-Store': 15, 'Sim-Store': 15
`DiskLifetimes`: Number of years kept on disk = 'Raw-Store': 1, 'Test': 0.5, 'Reco-Data-Store': 2, 'Sim-Store': 2
`TapeCopies`: Number of copies kept on tape = 'Raw-Store': 2, 'Test': 1, 'Reco-Data-Store': 1, 'Sim-Store': 1
`DiskCopies`: Number of copies kept on disk = 'Raw-Store': 1, 'Test': 1, 'Reco-Data-Store': 2, 'Sim-Store': 1.5
`PerYear`: Number of reprocessing done per year = 'Raw-Store': 1, 'Test': 1, 'Reco-Data-Store': 1, 'Sim-Store': 1, 'Events': 1, 'Sim-Events': 1, 'Reco-Data-CPU': 1, 'Sim-CPU': 1, 'Analysis': 1, 'Analysis-CPU': 1
`Cores`: Description of cores, efficiency and speed relative to 2020 vintage = 'Efficiency': 0.7, '2020Units': 1
`kHEPSPEC06PerCPU`: kHEPSPEC06 per core assumed = 0.011
`SplitsYear`: Year CERN no longer responsible for disk or tape = 2027
`SplitsEarly`: Division between FNAL/CERN/National for storage until SplitsYear = 'Tape': 'Raw-Store': 'FNAL': 0.5, 'CERN': 0.5, 'National': 0.0, 'Sim-Store': 'FNAL': 1.0, 'CERN': 0.0,

Draft

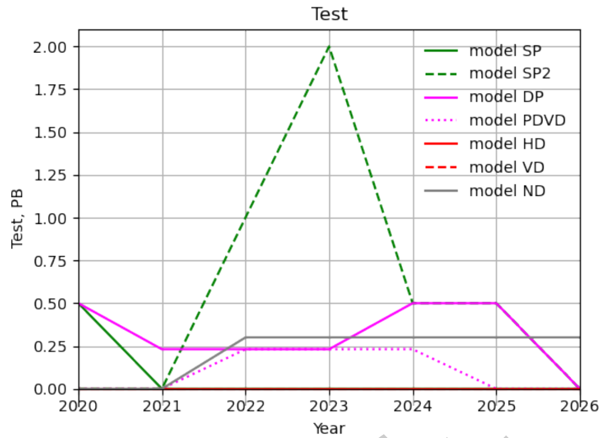
’National’: 0.0, ’Reco-Data-Store’: ’FNAL’: 1.0, ’CERN’: 0.0, ’National’: 0.0, ’Test’: ’FNAL’: 0.5, ’CERN’: 0.5, ’National’: 0.0, ’Disk’: ’Raw-Store’: ’FNAL’: 0.5, ’CERN’: 0.5, ’National’: 0.0, ’Sim-Store’: ’FNAL’: 0.25, ’CERN’: 0.0, ’National’: 0.75, ’Reco-Data-Store’: ’FNAL’: 0.25, ’CERN’: 0.0, ’National’: 0.75, ’Test’: ’FNAL’: 0.5, ’CERN’: 0.5, ’National’: 0.0
SplitsLater: Division between FNAL/CERN/National for storage after SplitsYear = ’Tape’: ’Raw-Store’: ’FNAL’: 0.5, ’CERN’: 0.0, ’National’: 0.5, ’Sim-Store’: ’FNAL’: 0.5, ’CERN’: 0.0, ’National’: 0.5, ’Reco-Data-Store’: ’FNAL’: 0.5, ’CERN’: 0.0, ’National’: 0.5, ’Test’: ’FNAL’: 0.5, ’CERN’: 0.0, ’National’: 0.5, ’Disk’: ’Raw-Store’: ’FNAL’: 1.0, ’CERN’: 0.0, ’National’: 0.0, ’Sim-Store’: ’FNAL’: 0.25, ’CERN’: 0.0, ’National’: 0.75, ’Reco-Data-Store’: ’FNAL’: 0.25, ’CERN’: 0.0, ’National’: 0.75, ’Test’: ’FNAL’: 0.5, ’CERN’: 0.0, ’National’: 0.5
filename: Input configuration file = MoreSim_2022-11-21-2040.json



Draft

Events	2020	2021	2022	2023	2024	2025	2026

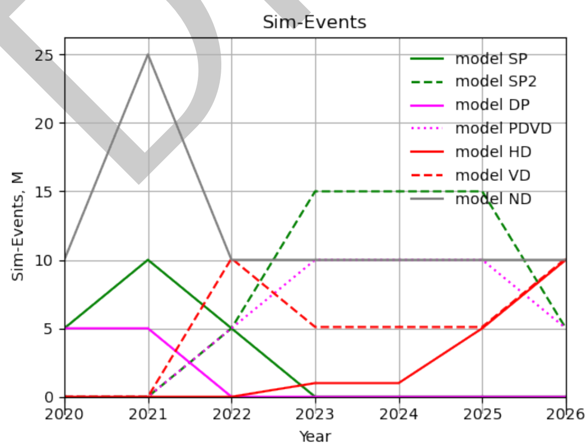
Figure 4: Projected million of detector events per year. Reconstructed data resources are based on this number.



Test	2020	2021	2022	2023	2024	2025	2026

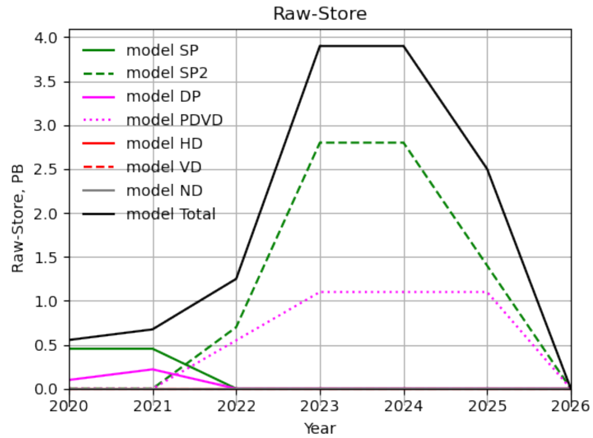
Figure 5: Projected PB of Test data per year.

Draft



Sim-Events	2020	2021	2022	2023	2024	2025	2026

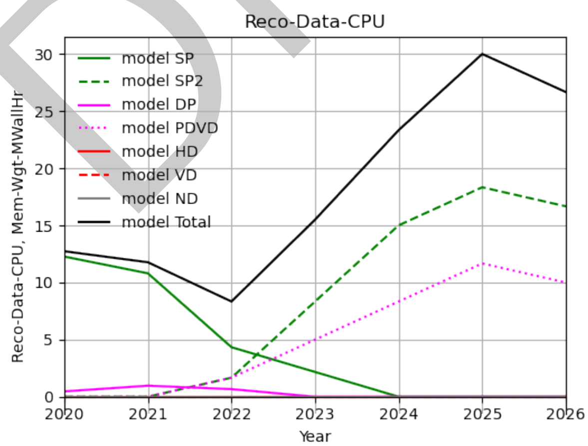
Figure 6: Projected millions of simulated events per year. Simulated data resources are based on this number.



Raw-Store	2020	2021	2022	2023	2024	2025	2026

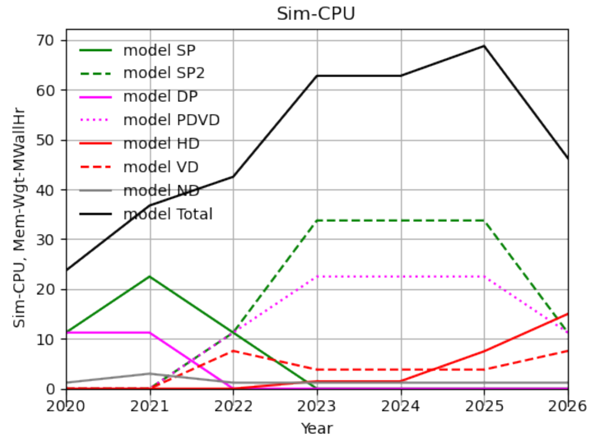
Figure 7: Projected raw data written per year in PB, derived from the number of events.

Draft



Reco-Data-CPU	2020	2021	2022	2023	2024	2025	2026

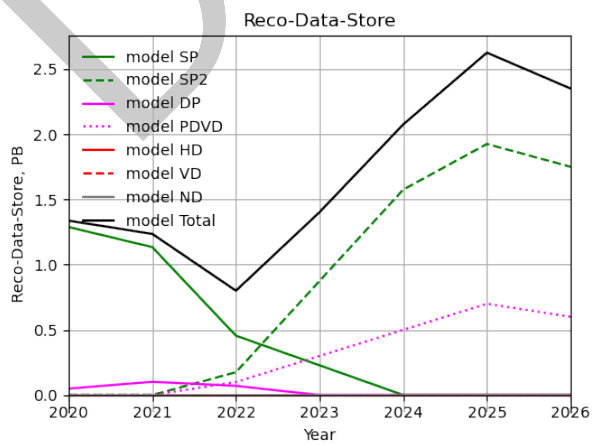
Figure 8: Projected CPU needs in core-years for data reconstruction. Slot weighted wall time takes into account memory use and an efficiency correction. Assumes rereconstruction of several years of older data.



Sim-CPU	2020	2021	2022	2023	2024	2025	2026

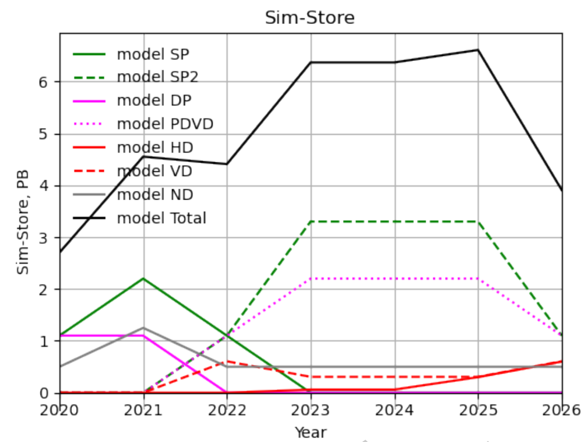
Figure 9: Projected CPU needs in core-years for simulation and reconstruction. Slot weighted wall time takes into account memory use and an efficiency correction. Based directly on the number of simulated Events.

Draft



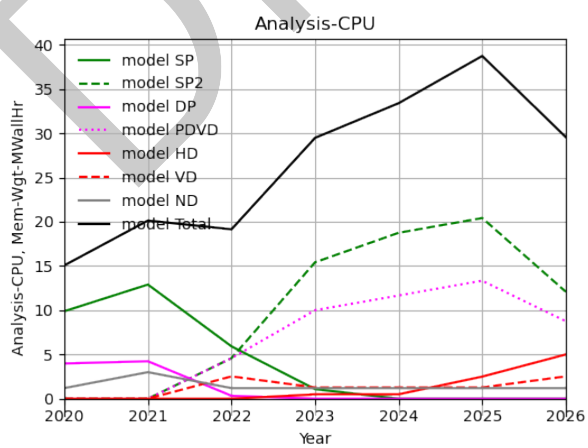
Reco-Data-Store	2020	2021	2022	2023	2024	2025	2026

Figure 10: Projected PB of reconstructed data per year. Includes reprocessing.



Sim-Store	2020	2021	2022	2023	2024	2025	2026

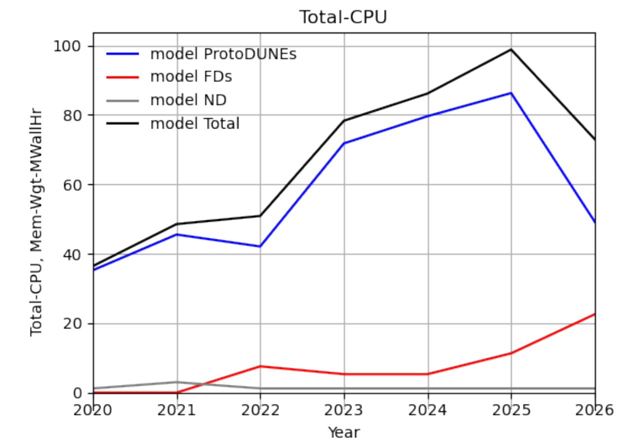
Figure 11: Projected PB of simulated data/year



Analysis-CPU	2020	2021	2022	2023	2024	2025	2026

Figure 12: Slot weighted analysis CPU needs in core-years. Assumed to be a weighted fraction of reco+sim needs.

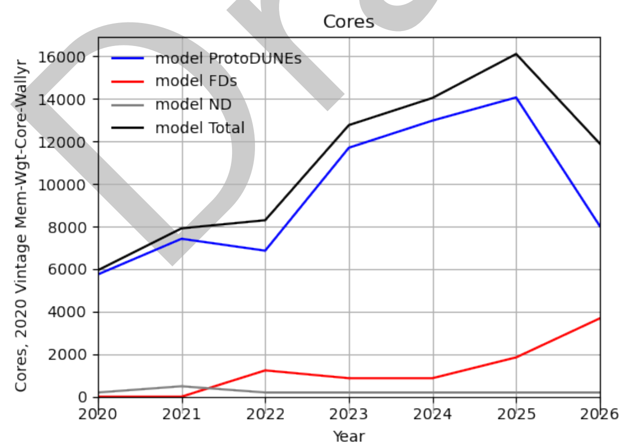
Draft



Total-CPU	2020	2021	2022	2023	2024	2025	2026

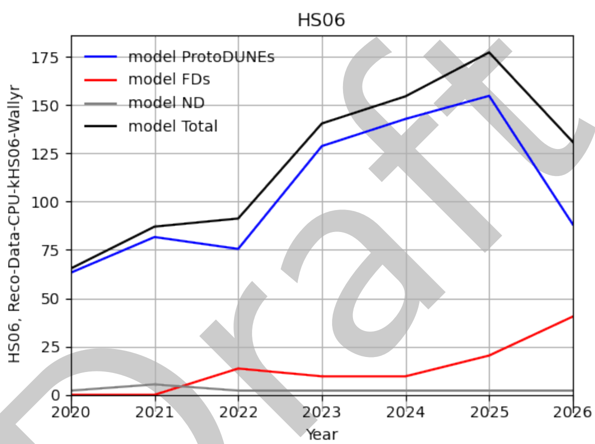
Figure 13: Slot weighted CPU needs in core-years. Slot weighted wall time takes into account memory and efficiency.

Draft



Cores	2020	2021	2022	2023	2024	2025	2026
ProtoDUNEs (2020 Vintage Mem-Wgt-Core-Wallyr)	5746	7423	6863	11701	12979	14066	8018
FDs (2020 Vintage Mem-Wgt-Core-Wallyr)	0	0	1235	868	868	1846	3681
ND (2020 Vintage Mem-Wgt-Core-Wallyr)	195	489	195	195	195	195	195
Total (2020 Vintage Mem-Wgt-Core-Wallyr)	5941	7912	8294	12765	14043	16109	11896

Figure 14: Slot weighted CPU needs in number of cores. Slot weighted wall time takes into account memory and efficiency.



Draft

HS06	2020	2021	2022	2023	2024	2025	2026

Figure 15: Slot weighted CPU needs in kHS06 hrs. Slot weighted wall time takes into account memory and efficiency.