

Data management summary

Steven Timm, Wenlong Yuan, Doug Benjamin

DUNE CRAB

17 Jan 2025

Overall RSE usage (TB) Jan 16

RSE	RuleUsage	RucioUsage	StorageUsage	DuneproLimit	StorageLimit	Change (R.U)	Free
BNL	0.04	10	3	830	831	0%	828
FNAL_DISK	1453	1575	1575	2100	2674	-1.4%	1147
CA_SFU	4	4	4	200	200	0%	196
PRAGUE	543	601	623	1126	1161	-0.6%	538
NIKHEF	1005	1098	1098	1143	1143	0%	45
SURFSARA	405	563	561	781	756	-7%	195
GLASGOW	495	495	495	500	500	0%	5
LANCASTER	452	462	462	700	700	0%	238
MANCHESTER	603	603	603	1000	1077	0%	474
QMUL	494	641	614	1100	1100	0%	486
RAL-PP	361	499	516	1000	1000	-5.9%	484
RAL_ECHO	699	849	845	1050	1000	-2%	155
CNAF	8	16	8	300	550	0%	542
CCIN2P3	1000	1059	1057	1100	1181	-0.3%	124
PIC	1181	1205	1271	1300	1439	-6%	168
CERN(new)	4215	4241	4241	4500	4500	-0.1%	259



Summary by type of data

Data type	Data Tier	Retention	Size(TB)
ProtoDUNE Raw	raw	physics	5772
ProtoDUNE trigprim	trigprim	test	1787
Coldbox raw	raw	study	609
Monte Carlo	Full-reconstructed	physics	2866
Transient files	miscellaneous	n/a	815
Est. dark data			159
Other (logs, per-rse)			2252*
Free space		n/a	5552

Decisions pending from CRAB or Conveners

- OK to remove hitreco level fardet MC off of disk? (300TB) Fardet reco/sim
- Are Histogram files from fardet reco-sim pass summer 2023 still needed? Fardet reco/sim—small but a lot of files.
- OK to get legacy protodune-dp raw data off of disk? (300TB) (Protodune reco/sim)
- OK to get most hd-protodune trigger primitives off of disk (1.7PB) (Protodune reco/sim, DAQ)

How Much Disk Space Do We Have:

- 5552 TB free
- 19812 TB total.
- Long-term datasets shown in previous slides account for 11869TB
 - In general we only have one copy of reco (MC or data) on disk where there should be two.
- In FY2025 expect O(5PB) from ProtoDUNE Vertical Drift
- Plus 1.3PB for Low Energy fardet-hd and vd production now in process.
- Above figure of total and free disk space does not count the 7.5 PB of disk in front of Fermilab tape-backed dCache which currently is the bulk of the US disk.
 - Have begun shifting 5PB of the 7.5PB of disk in front of tape-backed Fermilab dCache to persistent to follow CMS model that disk is disk, both in pledges and reality.