



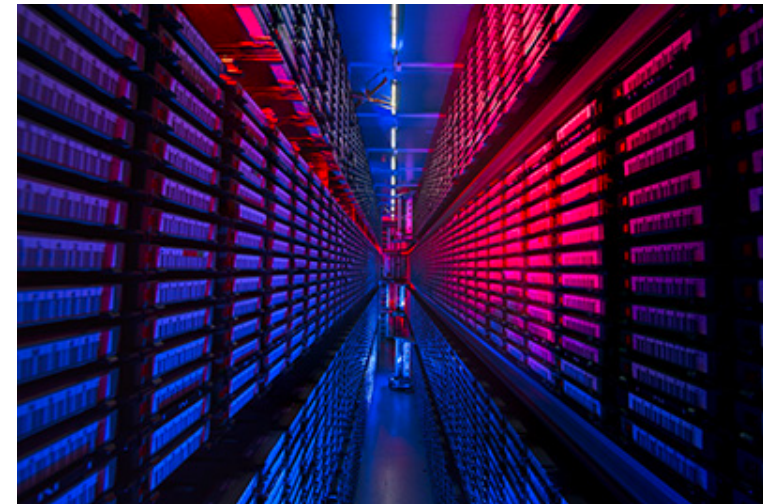
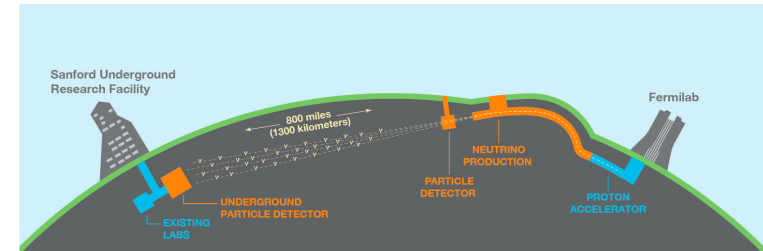
DUNE Computing Capacity Requirements

Michael Kirby, Fermilab/Scientific Computing Division

Jan 24, 2020

Outline

- development of the Resource Requirement Document
- general overview of computing plan (2020-2022)
- protoDUNE operation assumptions
- overview of Computing Model
- estimated resource requests
- discussion



Development of **Draft** Resource Requirements Document

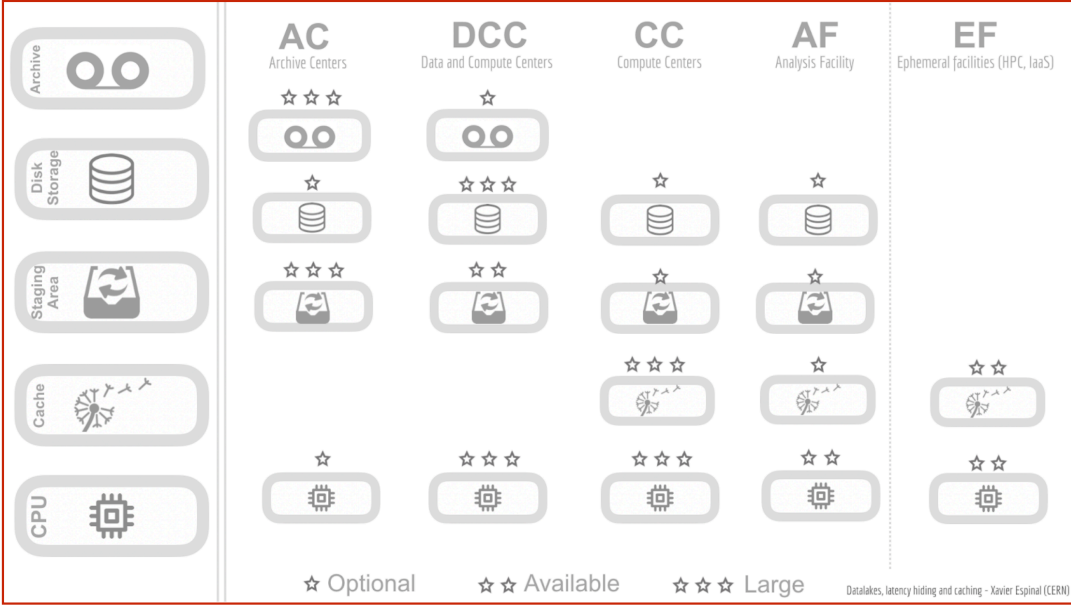
- Fermilab Computing Resource Scrutiny Group (CRSG)
 - CRSG review of resource requests from FNAL experiments (DUNE, NOvA, Mu2e, etc)
 - formation is in progress
 - anticipate that membership will be similar to previous SC-PMT
 - 2 reviewers from FNAL PPD and 4 external reviewers
 - membership composition at the discretion of CIO and FNAL Directorate
- As the review has not yet occurred, numbers presented here are preliminary
 - anticipated to happen in late March/early April
 - anticipate will prepare and submit plan for 3 years of protoDUNE and DUNE computing requirements
 - after review and revision DUNE will present the RRD to the CCB

General Plan for DUNE Computing 2020 - 2022

- Support the operation and analysis of protoDUNE SP and DP in 2020 (non-beam data) through disassembly for protoDUNE II
 - archive of raw, derived, and simulated data
 - production processing of SP and DP data twice per year
 - production processing of simulation for SP and DP (equal stats to raw data) twice per year
 - provide infrastructure and access to analysis computing resources for physics publications
- Support the DUNE Far and Near Detector for design and sensitivity improvements
 - simulation of FD and ND samples
 - sensitivity studies for detector improvements and reviews
- Utilize the OSG, WLCG, and additional compute resources when available
- Improve access to HPC resources for specific analysis workflows and production workflow development

DUNE Computing Model for Institutional Sites

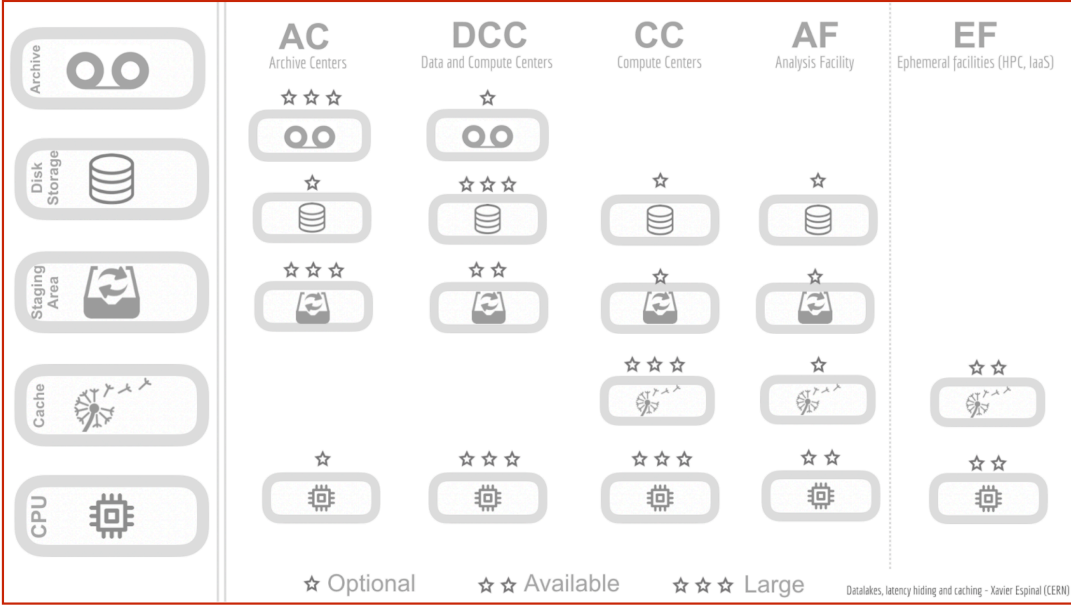
- Based the HSF DOMA model for sites
 - Archive Center (AC) - tape/staging
 - Data and Compute Center - disk + CPU
 - Computer Center (CC) - CPU + cache
 - Analysis Facility (AF) - cpu + cache
 - Ephemeral Facilities (EF) - (HPC, IaaS)
- Goal is to have resource split between FNAL and other institutions - 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

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

Computing Model Policies

- Tape Storage
 - two copies of all raw data for security
 - FNAL provides storage for an archival copy of all raw data for DUNE (ND, FD, protoDUNE)
 - Rucio Storage Elements (RSEs) around world provide storage for 2nd copy
 - FNAL provides storage of derived datasets with lifetime of 2 years
 - FNAL provides storage for single copy of simulated data
 - RSEs around the world provide storage for second copy of simulated data
- Disk Storage
 - two or three copies of every active derived dataset on disk at any time
 - two derived datasets will be active at any one time
 - latest two active derived dataset staged to disk at FNAL
 - two or three copies of every active simulated dataset on disk at any time
 - two simulated datasets will be active at any one time (matching active derived dataset)
- From these policies

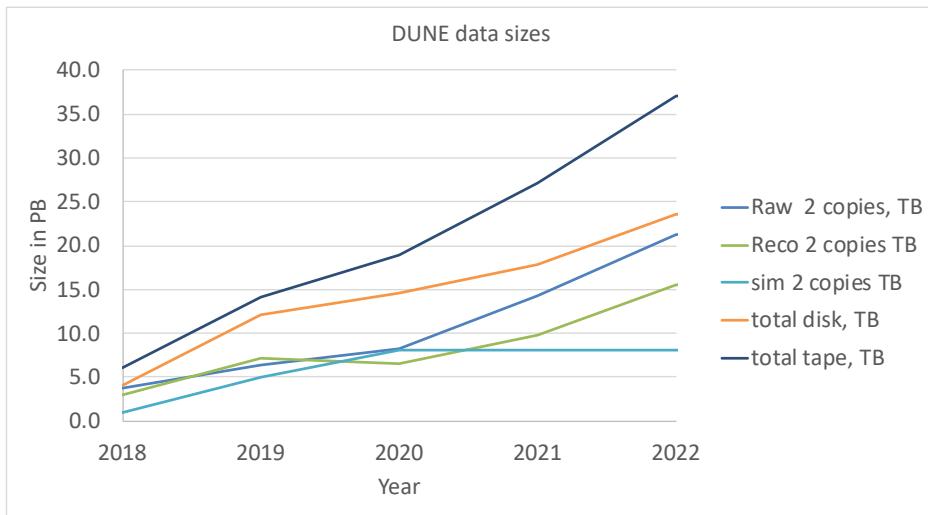
Assumptions about protoDUNE SP and DP operations

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
total data (TB)	450	1625	2070

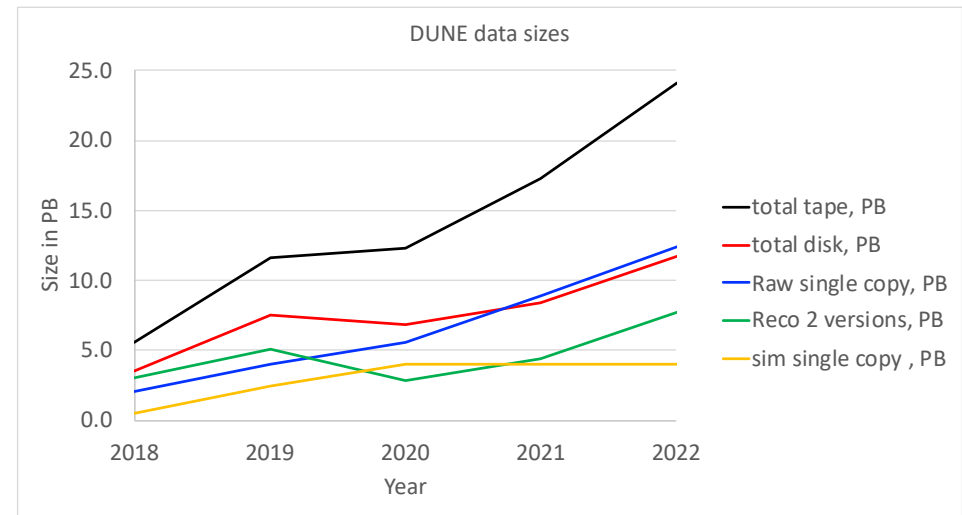
protoDUNE Dual Phase	2020	2021	2022
cosmic rate (Hz)	5	1	1
beam rate (Hz)	0	18.3	18.3
uptime (days)	100	150+50	150+50
events	21 M	62 M	62 M
event size (MB)	220	220	220
compression	10	10	10
total data (TB)	240	1360	1360

Tape and disk storage 2018-2022

Total DUNE Storage



FNAL DUNE Storage



- Computing Model for DUNE Storage
 - 2 archival copies of raw, derived, and simulated data - 1 copy at FNAL, second copy distributed institutions
 - production processing of SP and DP data and matching simulation twice per year
 - 2 or 3 copies of active derived and simulated datasets on disk - dataset stays active for 1 year

DUNE CPU requirements

protoDUNE Single Phase	2020	2021	2022
data events	6.5 M	24 M	30 M
CPU (hr)	2 M	4 M	5 M
MC events	5 M	10 M	10 M
CPU (hr)	3.8 M	7.5 M	7.5 M
analysis (hr)	8 M	8 M	8 M
Total CPU (hr)	13.8 M	19.5 M	20.5 M
Total CPU (HS06 years)	15750	22260	23400

protoDUNE Dual Phase	2020	2021	2022
data events	21 M	62 M	62 M
CPU (hr)	3.5 M	10.5 M	10.5 M
MC events	5 M	10 M	10 M
CPU (hr)	3.8 M	7.5 M	7.5 M
analysis (hr)	8 M	8 M	8 M
Total CPU (hr)	15.3 M	26 M	26 M
Total CPU (HS06 years)	17470	29680	29680



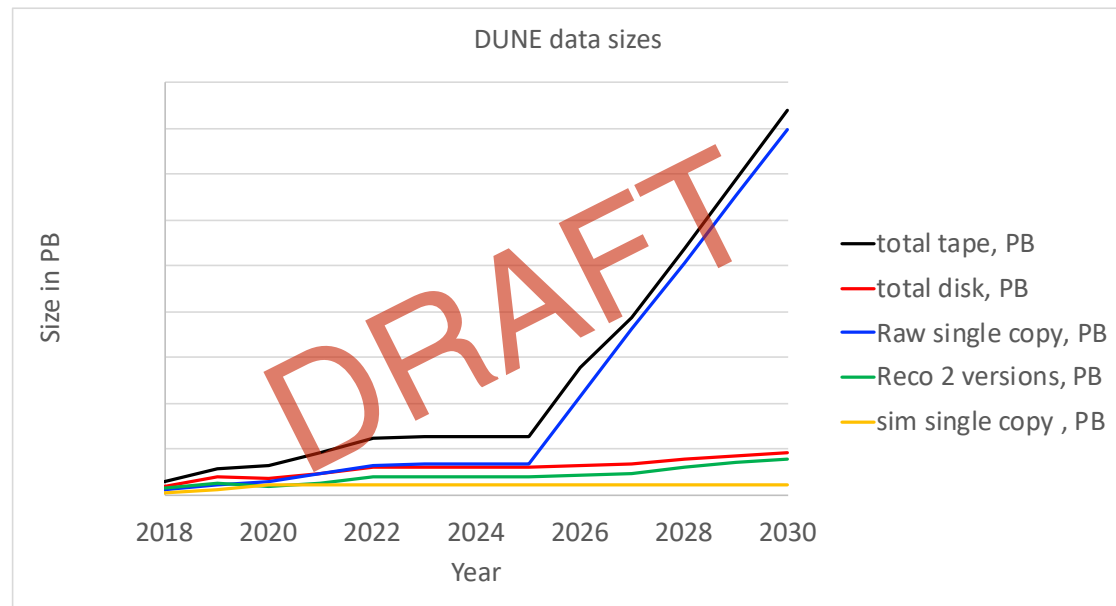
- SP Data 0.16 hr/evt - SP MC 0.75 hr/evt - DP 0.16 hr/evt - DP 0.75 hr/evt

Total DUNE Resource Needs

Resource	2020	2021	2022
Disk (PB)	15	18	24
Tape (PB)	19	28	37
CPU (kHS06-years)	33.1	51.9	53.1
CPU Cores	2200	3460	3540

Longer term estimates

- We are actively working on longer term projections for resource requirements
- Fermilab has a specific request from Dept of Energy for a pre-Operations plan with a focus on the resource needs for Software & Computing
- This is a draft and only intended to inform that the consortium is working on estimates for resources needs during operations



Discussion