

Production report

Aaron Higuera Pichardo, Elisabetta Pennacchio, Matt Kramer
CRAB Meeting, September 27th, 2024

Current productions

Last [CRAB meeting](#) (August 23th) :

Tr Campaign	Priority	Tr Shifter(s)	Status	Start date	End date	Tr Deliverable	Tr Notes
CAF Production	P0	Elisabetta/Aaron	testing	8/2/2024	m/d/yyyy	Deliverable	Notes
ProtoDUNE keep-up	P1	Jake	Running	8/1/2024	m/d/yyyy	Deliverable	Notes
<i>Campaign</i>		<i>Shifter(s)</i>		<i>m/d/yyyy</i>	<i>m/d/yyyy</i>	<i>Deliverable</i>	<i>Notes</i>

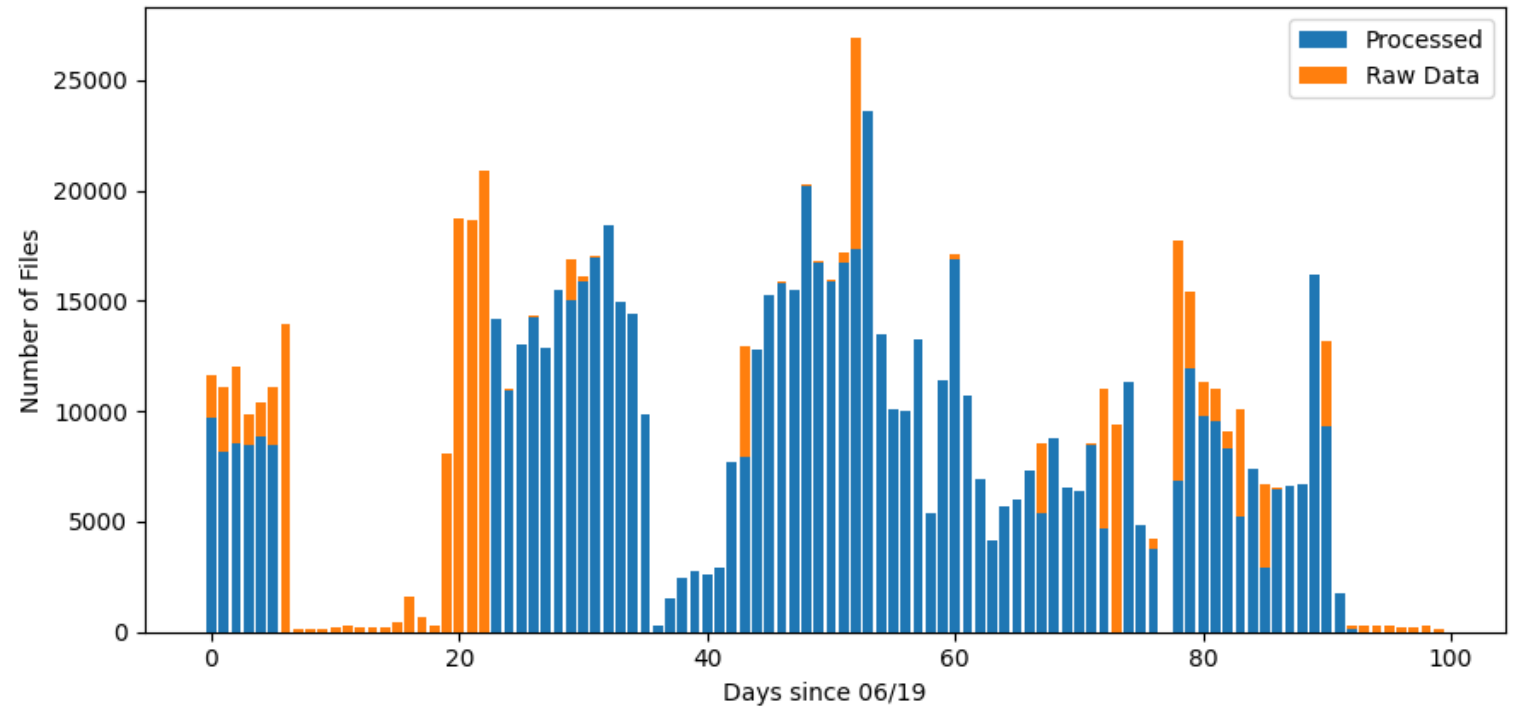
Now :

Production	Tr Campaign	Priority	Tr Shifter(s)	Status	Start date	End date	Tr Deliverable	Tr Notes
	CAF Production	P0	Heidi/Elisabetta/Aa	Running	8/2/2024	m/d/yyyy	CAF atmo stage 1	Notes
	ProtoDUNE keep-up	P1	Jake	Running	8/1/2024	m/d/yyyy	Deliverable	Notes
	LE Request	P0	Elisabetta	testing	9/23/2024	m/d/yyyy	Deliverable	Notes
	<i>Campaign</i>		<i>Shifter(s)</i>		<i>m/d/yyyy</i>	<i>m/d/yyyy</i>	<i>Deliverable</i>	<i>Notes</i>

CAF files production for LBL and atmospheric productions

- CAF production: 13 samples (1 Fardet-HD atmos, 6 Fardet-VD, 6 Fardet-HD)
- Processing is relatively quick and output file size is small (stage 1), therefore we need to merge CAF files (stage 2), merging is a challenging operation
 - HD atmospheric sample has been processed and CAF files produced (stage 1). Final merging is completed (stage 2, Heidi), files to be declared to RUCIO and MetaCat
 - One VD sample has been processed and merged (stage 1 and 2), CAF files to be declared to RUCIO and MetaCat
 - Processing of remaining VD and HD samples will continue next weeks

- Automated submission (twice a day) using raw data with `core.data_stream` in (physics, cosmics)
- +640 TB of reco files have been produced (reco2 stage)
- Finishing-up in progress
 - Early time period before keep up started
 - Resubmitting failures due to site issues



LE production

- Complex production: 14 different samples to be generated, corresponding to FD1-HD, FD2-VD, with and without radiological background (volume ~1.4 PB)
- It is a 4 stage production: generation, g4stage1, g4stage2, detsim, reco1
- Further processing is foreseen once Pandora training is complete
- To allow early training, the requester asked to produce 50k events for each sample before the full sample set is produced (volume ~60TB).
 - Jobscripts prepared, results being investigated now, to validate workflows implementation and metadata
 - Event size is very different among samples; samples without background (both FD1-HD and FD2-VD) require merging → in progress
 - If no unexpected issues show up, we can start production within one/two weeks . Possibility to produce some samples at NERSC under study

- 2x2: Aiming to collect $1.5E20$ POT RHC of NuMI in 2025
Need $1.5E21$ POT of MC for analysis:
 - 15,000 CPU + 15,000 GPU node hours at Perlmutter
 - Dominated by rock GENIE+Geant4 (CPU), larnd-sim (GPU)
 - ~0.7 PB of storage under current practices
 - Without keeping intermediate files, ~0.4 PB
- ND: 50x more LAr
Need $1E19$ POT RHC + $1E19$ POT FHC for summer IPR
Scaling vs. 2x2 per POT: 20x CPU (rockbox helps), 50x GPU
Assuming FHC costs 3x RHC
 - 8,000 CPU + 20,000 GPU node hours at Perlmutter
 - 150 TB
- Production moving to dunepro NERSC allocation; ERCAP request in preparation