Status of ICARUS Production

Maya Wospakrik

ICARUS Collaboration Meeting

12 September 2019

Status of Monte Carlo Production

- The current MCC1 campaign is a year-long production that was launched to simulate and fully reconstruct events, including TPC, PMT and CRT information and are used to study the simulation, calibration, and reconstruction.

- Starting on **November 2018**, ICARUS production have fully migrated to the **Production Operations Management Service (POMS)** for the fully simulated production campaign

  - POMS enables automated jobs submission on distributed resources by setting up dependency between different stages of production.

    - **Jobsub** - provides support for the job lifecycle enabling the management of jobs on the Grid.

    - Sequential Access via Metadata (**SAM**) - the data handling system, to keep track of files, their meta-data and processing.

    - **FIFEmon** for monitoring.

🔷 **Fermilab**

Status of Monte Carlo Production

- POMS MC Campaign:
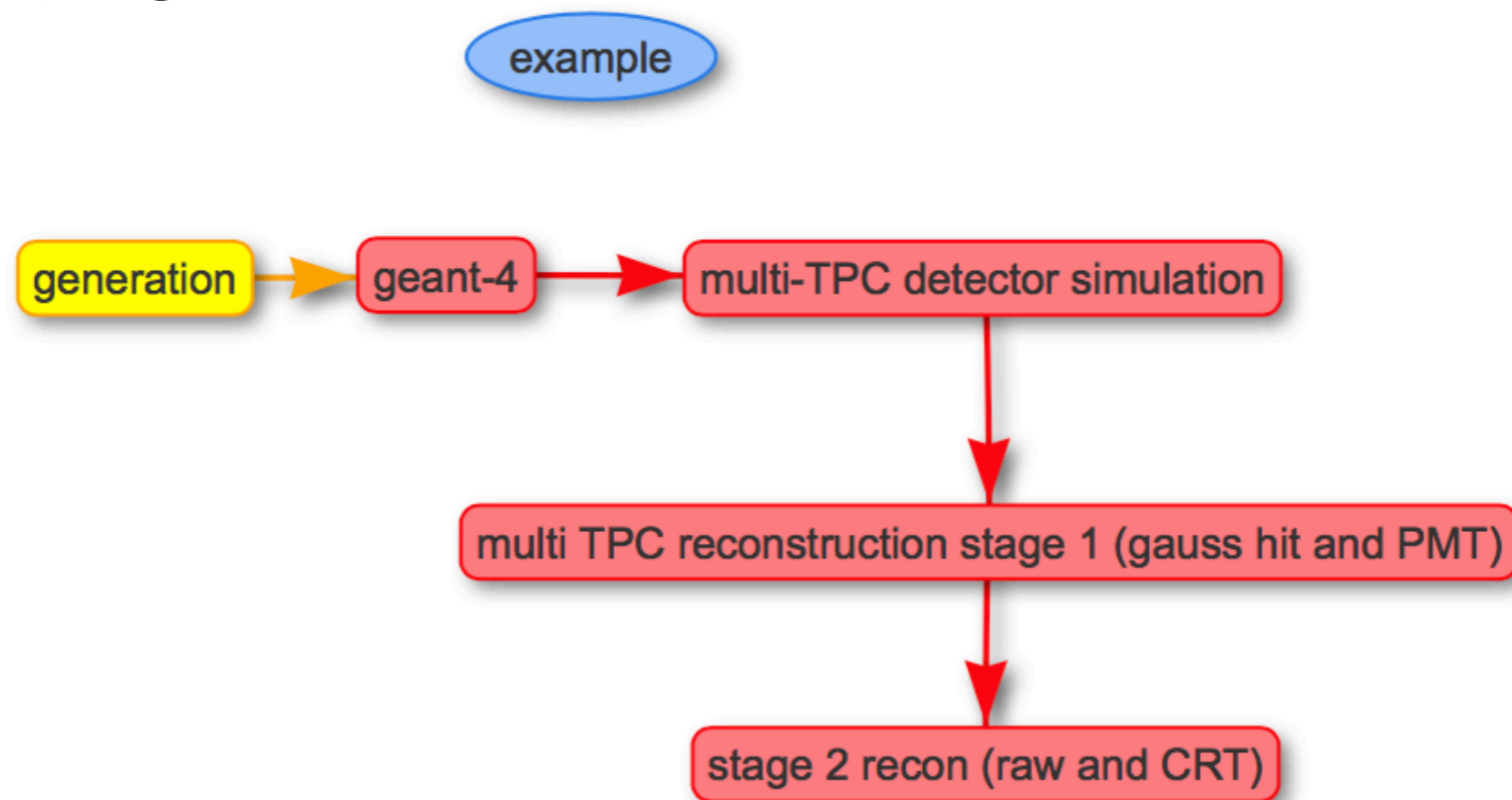
The production instance of POMS is at https://poms.fnal.gov

🔱 Fermilab

Status of Monte Carlo Production

- The production campaign started in November is used to:
  - Produce calibration sample and numi off-axis sample (**v08_01_00**)
  - Fully iron out any remaining bugs related to POMS.
  - Exercise the use of necessary tools needed for production:
    - SAM data cataloging
    - Metadata format
    - Fermilab FTS (File Transfer Service) - daemon process that automates the transfer of files from one storage system to another
  - Common developments with SBND

- Fully simulated MC campaign was started at the end of March to the beginning of April 2019 with aim to provide various sample for the SBN workshop held at Oxford University on April 2019 (**v08_13_00**)
    - Issues identified related to the geometry and memory leakage when using photon library at the g4 stage and there's no cosmic muon sample

🔷 Fermilab

Status of Monte Carlo Production

- Another MC campaign was started on June 2019 to provide several sample for reconstruction and calibration studies (v08_22_00) which includes the following improvements:

    - Splitting the reco step into two sections, the first primarily does the gauss hit finding path (plus optical), the second does the "raw" path (plus CRT)

    - Utilize the old photon library covering a single cryostat, mapping it also to the second one.

    - Memory reduced from **12GB** in average to **4.5GB** when using photon library for light simulation.

- Current MC campaign started in **September 2019** to provide several sample for the upcoming SBN workshop in Fermilab (see following slide).

🔀 **Fermilab**

Available Production Sample

| Sample	Software version	Full Simulation	Number of Events	Size of data/event average after track/shower reco
Isotropic muon	v08_22_00	Yes	52k	~18MB/event
Purity sample 1ms, 2ms, 4ms, 6ms, 8ms, 15ms	v08_22_00	Yes	~10k each	~600MB/event
Numi off axis (mostly numu)	v08_22_00	Yes	~60k	~20MB/event
BNB neutrino sample	v08_13_00	Yes	~11k	~18MB/event
BNB oscillated electron-neutrino	v08_13_00	Yes	~60k	~35MB/event
BNB intrinsic electron-neutrino	v08_13_00	Yes	~75k	~45MB/event
electron and pi + vertex sample	v08_13_00	Yes	~26k	~45MB/event
single electron particle gun (bnb-like)	v08_13_00	Yes	~25k	~25MB/event
single pi0 particle gun (bnb-like)	v08_13_00	Yes	~26k	~45MB/event
single muon isotropic particle gun (bnb-like)	v08_13_00	Yes	~26k	~18MB/even

చ్ఛ Fermilab

Sample scheduled for Production

Sample	Software version	Tested	Status	Number of Events requested
BNB muon neutrino	v08_30_01	Yes	In production	25k
BNB electron-neutrino	v08_30_01	Yes	In production	25k
cosmic muon	v08_30_00	Yes	In production	50k
electron neutrino + cosmic	v08_30_01		Test stage	25k
muon neutrino + cosmic	v08_30_01		Test stage	25k
single photon with and without new noise model	v08_30_01		Test stage	100k
Michel electron with and without new noise model	v08_30_01		Test stage	100k

Fermilab

Status of Monte Carlo Production

We have about **91% job success rate** for the current production.

Most of the failures are due to test samples using custom production workflows.
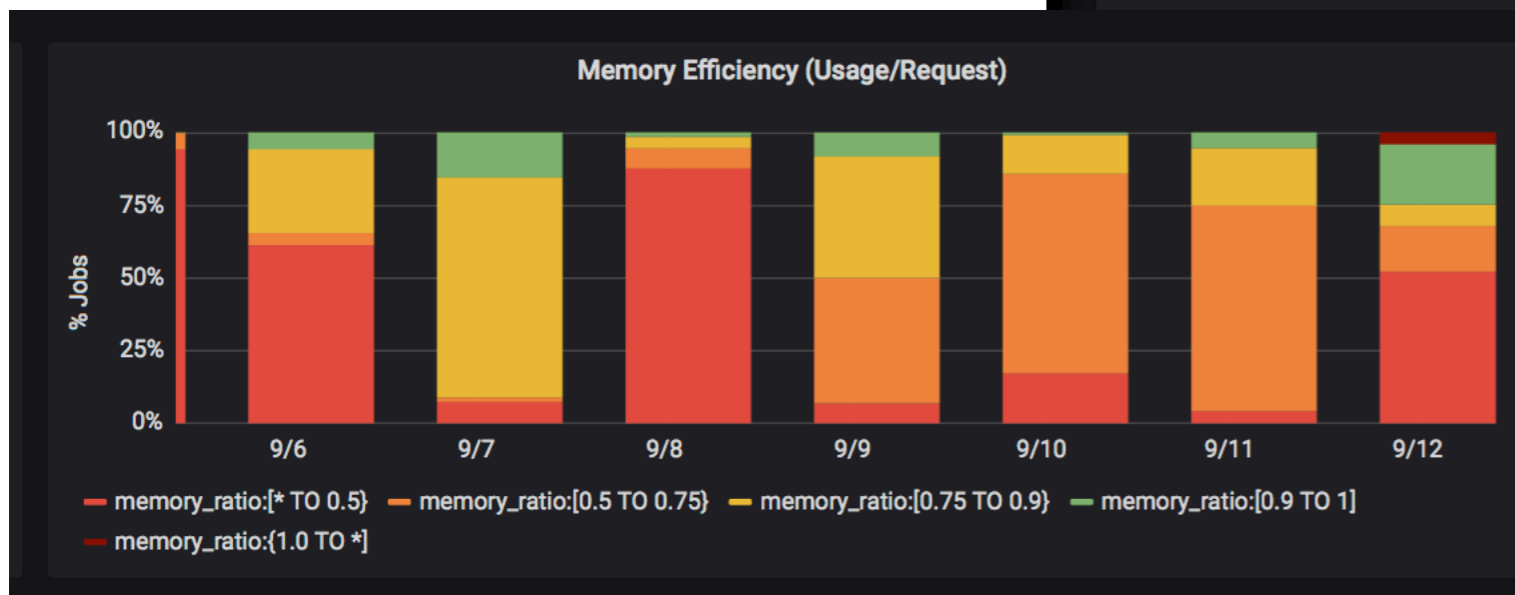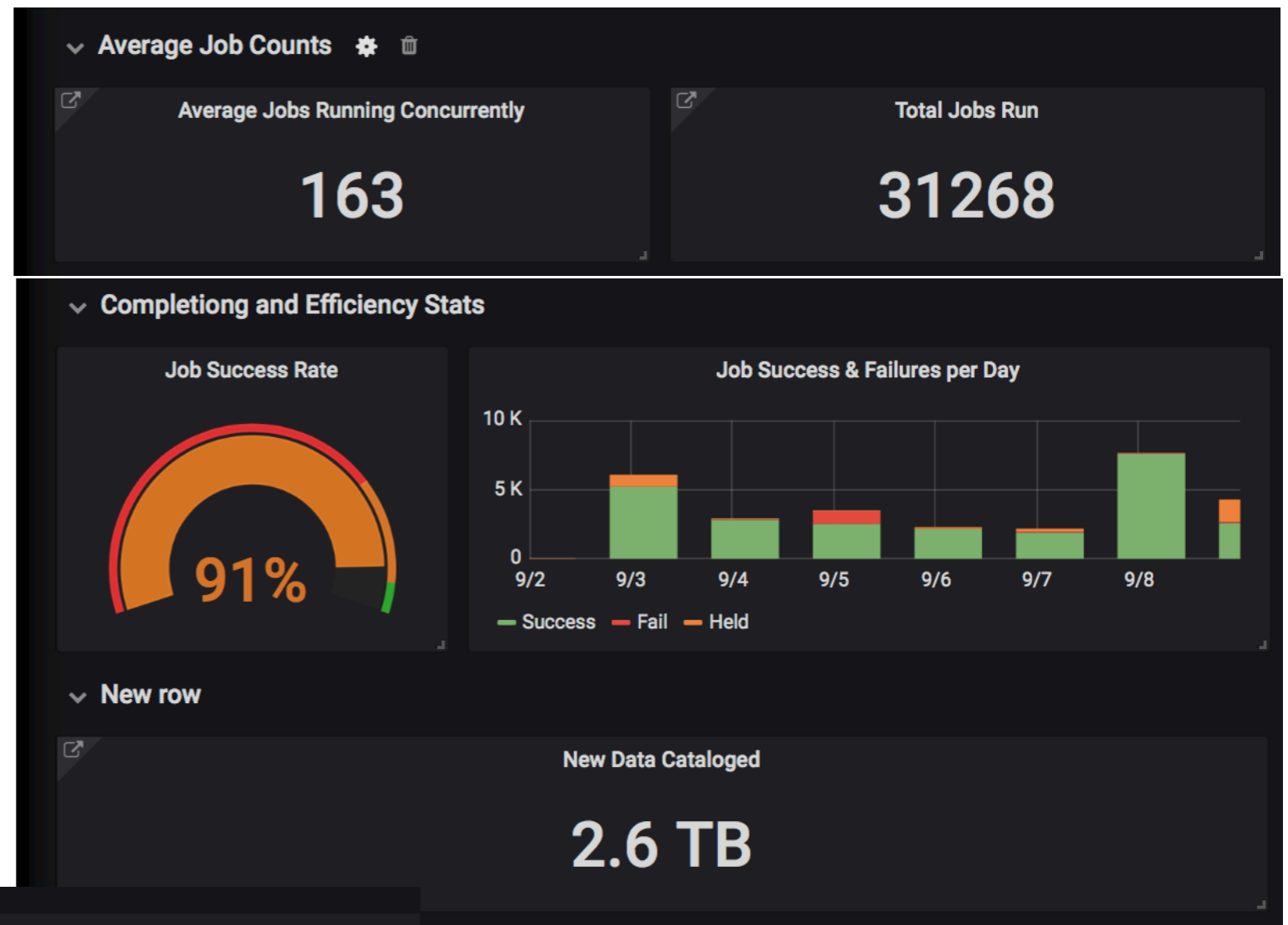
Test samples are essential for **memory profiling** and **unraveling any bugs/ failures** before launching the large scale production

🔷 Fermilab

ICARUS Data Storage

Volume	Quota	Used	Use %	Usage	Grid accessible
Tape backed dCache	~2 PB	37 TB	2%	Long-term archive	Yes
Persistent dCache	72 TB	53 TB	72%	immutable files with long lifetime	Yes
BlueArc Data	20 TB	1.7 TB	9%	Storing final analysis samples	No
BlueArc App	2 TB	1.1 TB	55%	Storing and compiling software	No
Persistent scratch	No limit	Shared across all experiments		immutable files w/ short lifetime	Yes
Persistent resilient	No limit			input tarballs with custom code for grid jobs (do NOT use for grid job outputs)	Yes

🔷 **Fermilab**

Future Data Storage Production
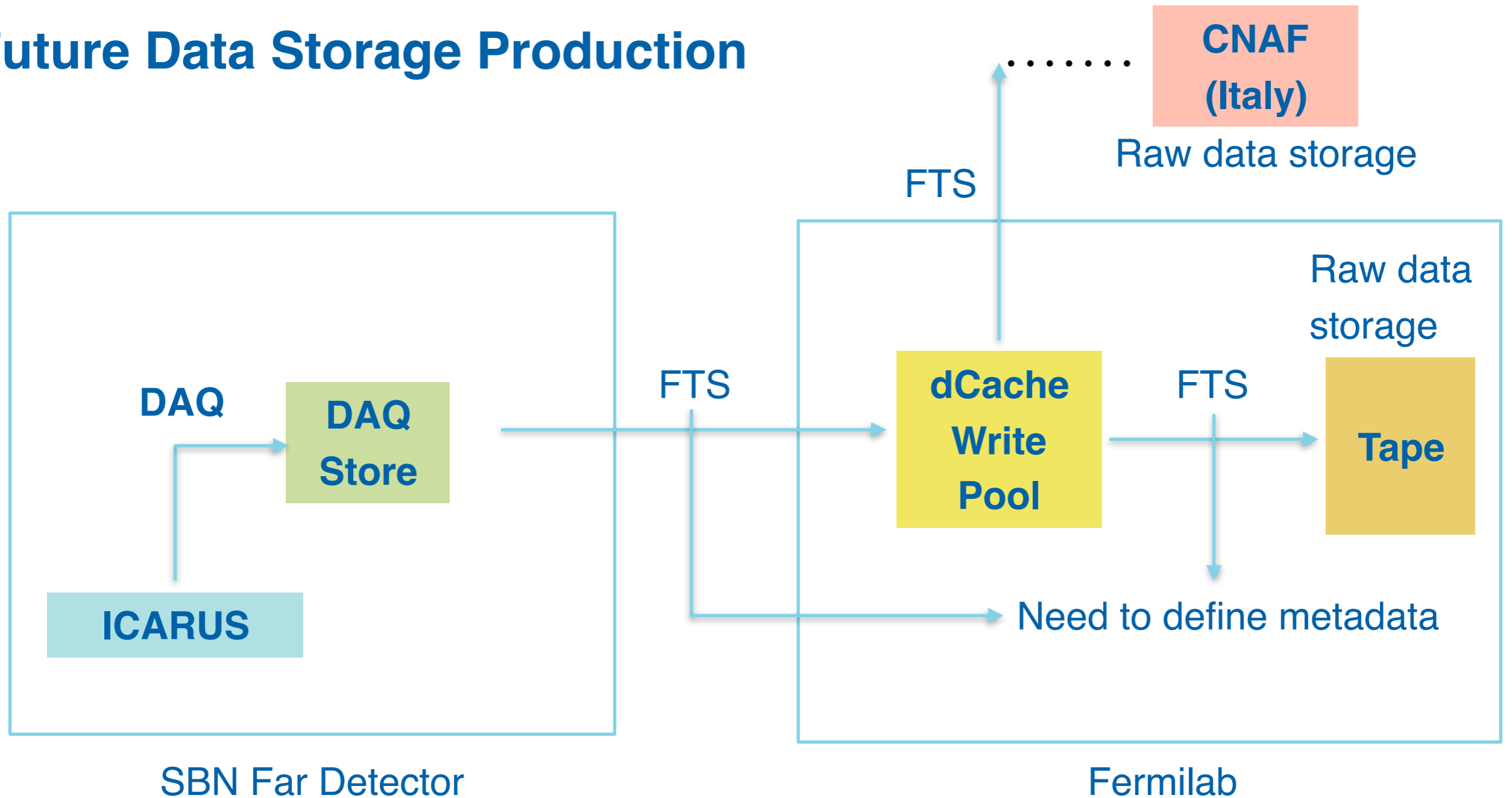
Data Files that need to be stored:

1. Raw Data
2. Wires Hits
3. Tracks and Showers

🟦 **Fermilab**

# Production plans beyond commissioning/first data

- We are beginning to work with **experts from CNAF** to integrate additional storage and processing power

- Based on previous production campaigns and expected data-taking rates, we have estimates for our total computing footprint

  - In normal operations, we expect to take **~2.2 PB raw data per year**

  - We assume that we will move quickly to using our own data taken outside of the beam spill to model and measure cosmic background interactions

    - Simulating cosmics at large scale will likely be too burdensome on our available computing

  - We are planning for **major simulation and reconstruction campaigns** to be achievable for the full collected datasets at a frequency of about **once per year**

- What seem to be small changes can affect the final numbers considerably, and so we will need to be careful to keep all estimates up-to-date

🔷 **Fermilab**

From: Wes Ketchum

Summary

- Production are currently run via POMS interface.

- Various samples have been produced to study the simulation, calibration, and reconstruction and new samples with current best framework are in production or have been scheduled.

  - Detailed information about the samples, tutorials, requesting samples: https://cdcvs.fnal.gov/redmine/projects/icarus-production/wiki

- We have begun working with experts from CNAF Italy to integrate additional data storage and processing power.
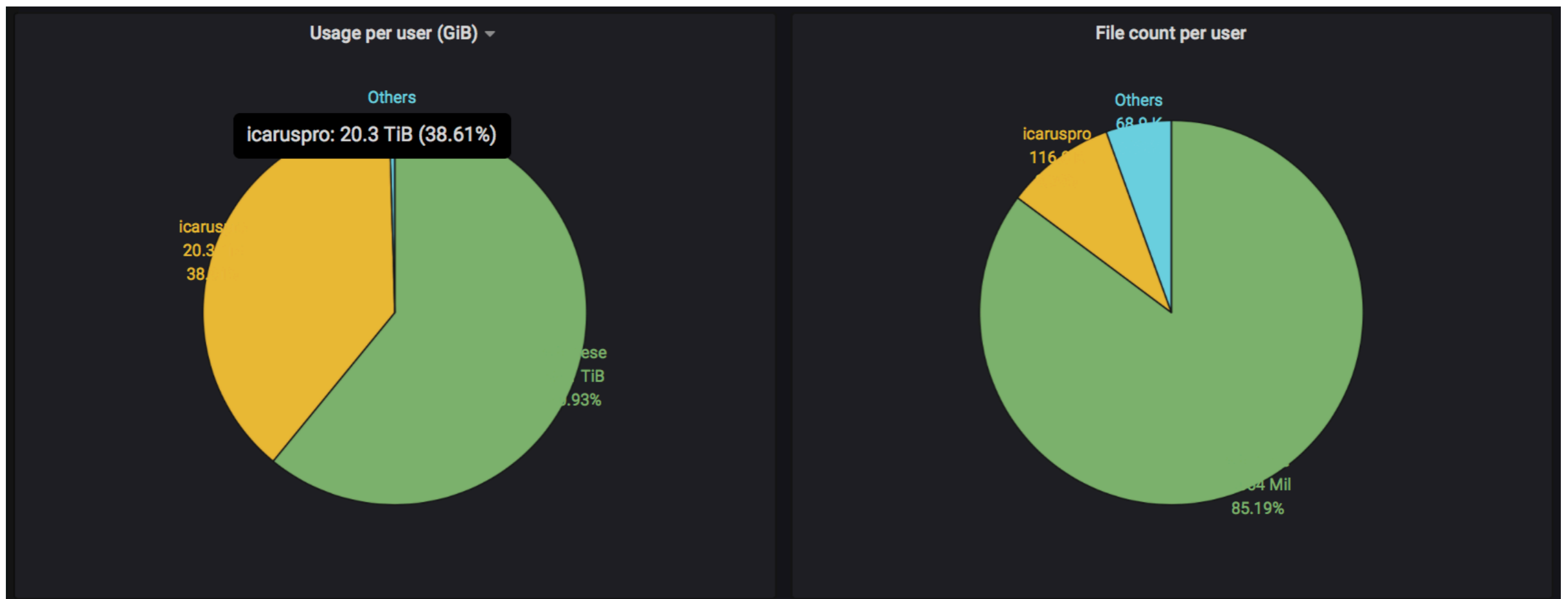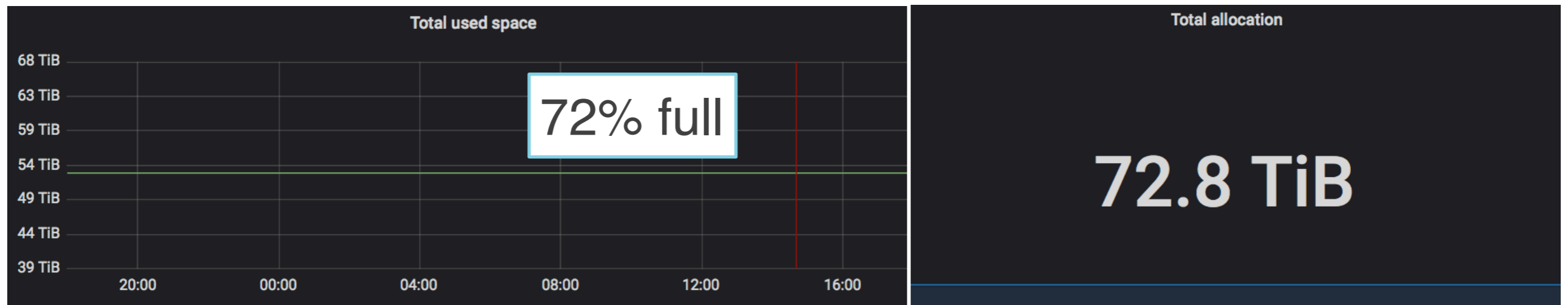
🔷 **Fermilab**

BACKUP SLIDES

ICARUS Tape Data Storage
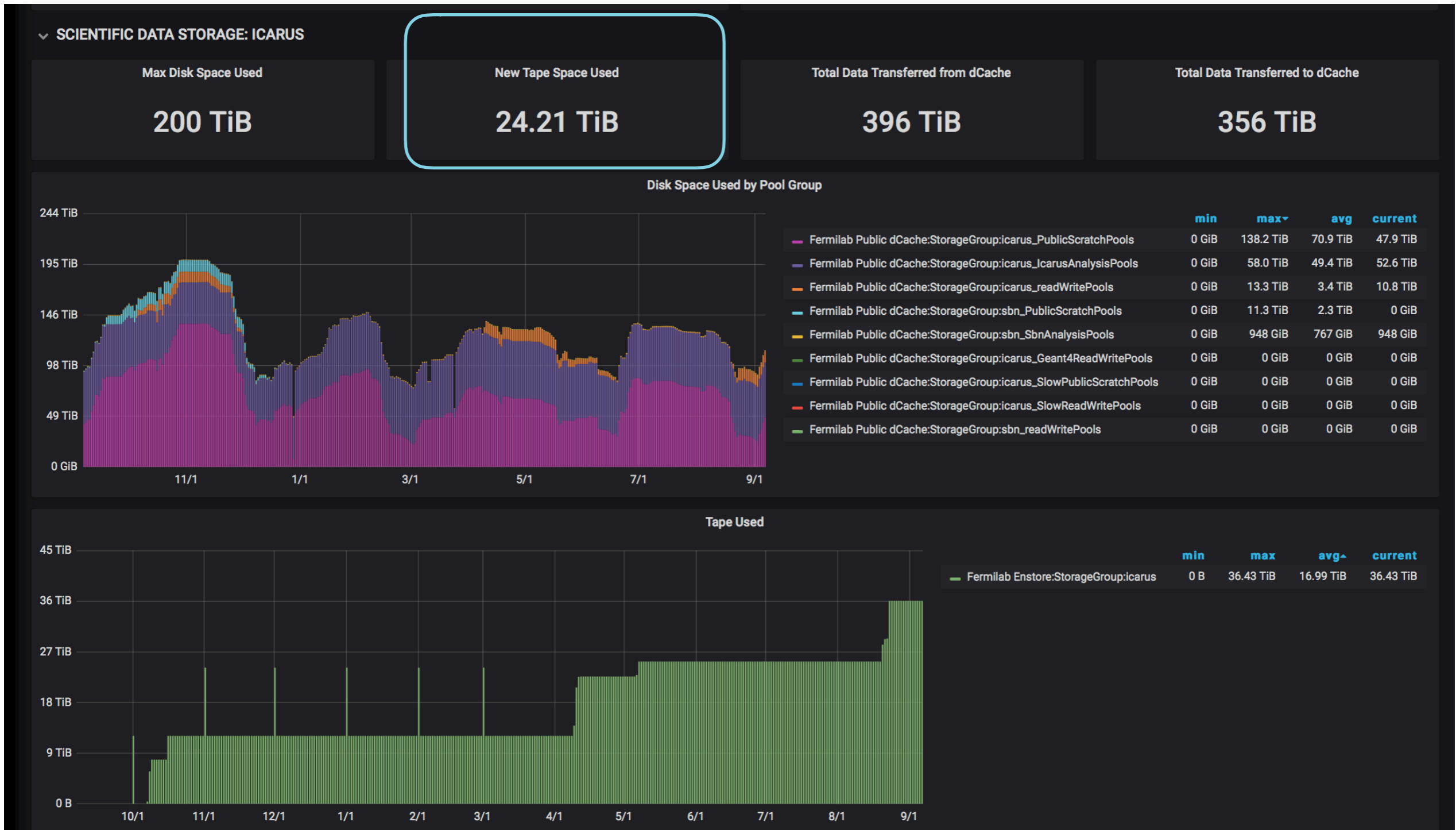
Storage volumes used for long-term archive (2PB storage currently allocated)

ICARUS DCache Persistent Data Storage (accessible by grid)

Total used space

72% full

Total allocation

72.8 TiB

Usage per user (GiB)

Others

icaruspro: 20.3 TiB (38.61%)

icarus...
20.3
38...

...ese
...TiB
...93%

File count per user

Others
68.9 K

icaruspro
116...

...4 Mil
85.19%

Storage volumes used for storing immutable files with long lifetime
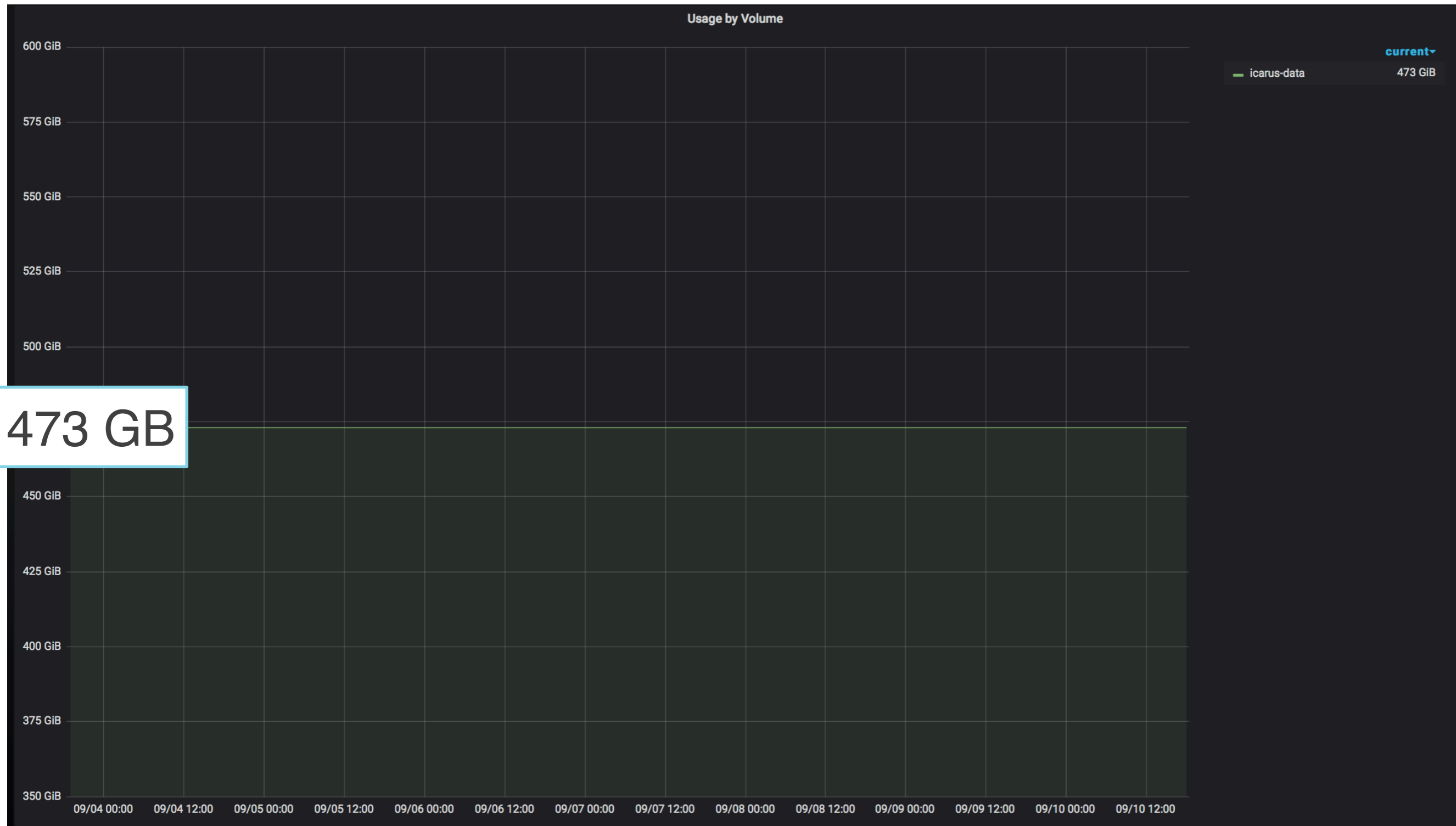
🔷 Fermilab

ICARUS BluArc Usage (/icarus/app) - not accessible by grid
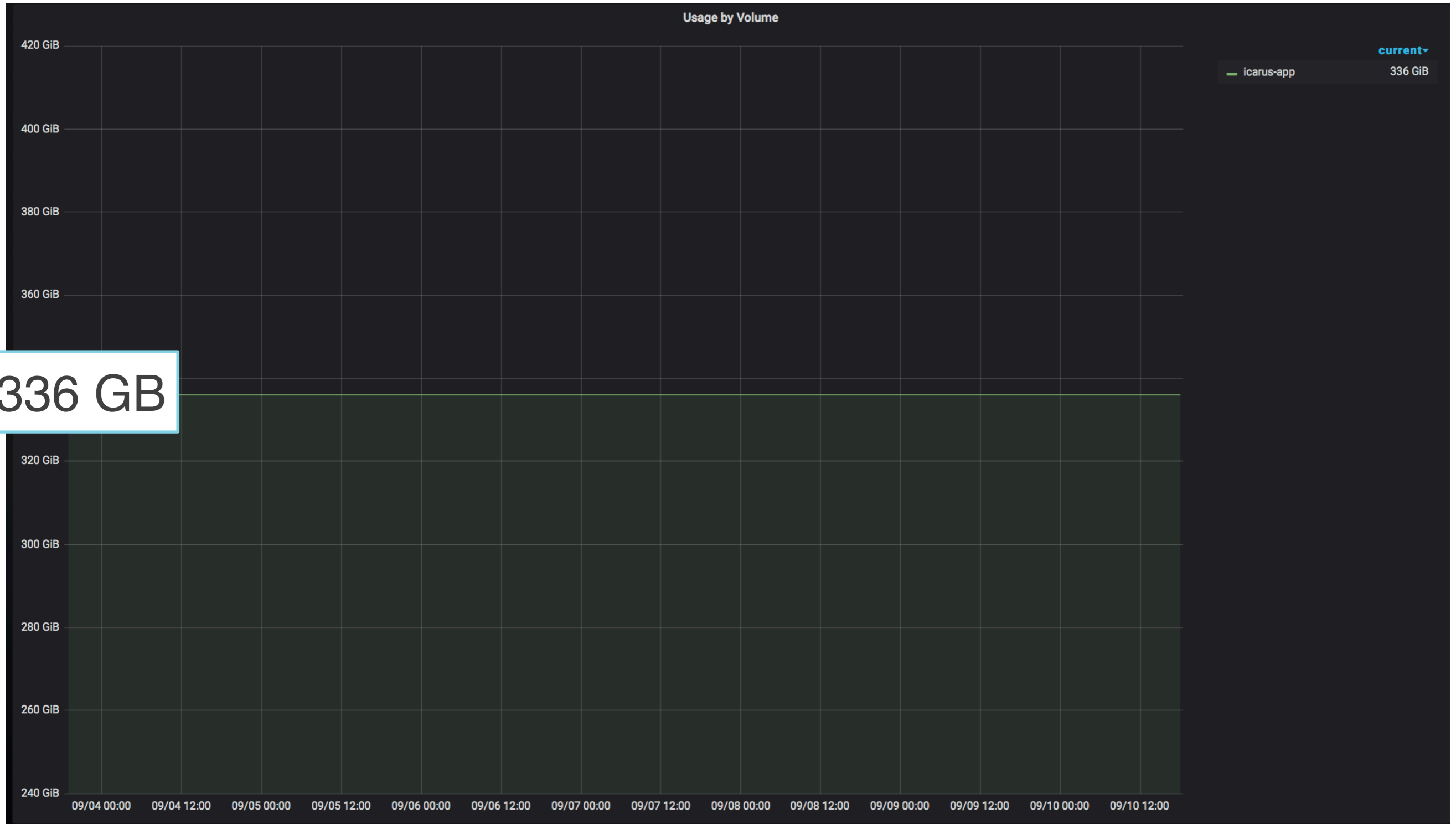
Storage volumes used for storing and compiling software

🔷 Fermilab

ICARUS BluArc Usage (/icarus/data) - not accessible by grid

Storage volumes used for storing final analysis samples

Looking towards data produced during commissioning

Inputs data	
Raw Data Size (MB/event)	70
Reco data size (MB/event)	18
Sim data size (MB/event)	80
BNB Neutrinos per year	200,000
BNB rep rate (Hz)	5
NuMI neutrinos per year	150,000
NuMI rep rate (Hz)	0.5
Off-beam final trigger rate (Hz)	1
On-beam final trigger rate (Hz)	1
Shutdown fraction	0.25
Reco processing time per event (s)	240
Reco processing time per MB data (s/MB)	3.428571429
Simulation processing time per event (s)	240

Commisioning	
Commissioning data events	25920000 (based on 5Hz on first month, 3Hz second month, and 2Hz last month)
Commissioning data volume (PB)	1.8144
TOTAL DATA PER YEAR (PB)	1.8144
Cumulative "raw" data events	25920000
CUMULATIVE "RAW" DATA (PB)	1.8144
Keep-up events	25920000
KEEP-UP PROCESSING (CPU Mhr)	1.728
KEEP-UP DATA PRODUCED (PB)	1.8144
KEEP-UP DATA CUMULATIVE (PB)	1.8144

🔷 Fermilab