Operations

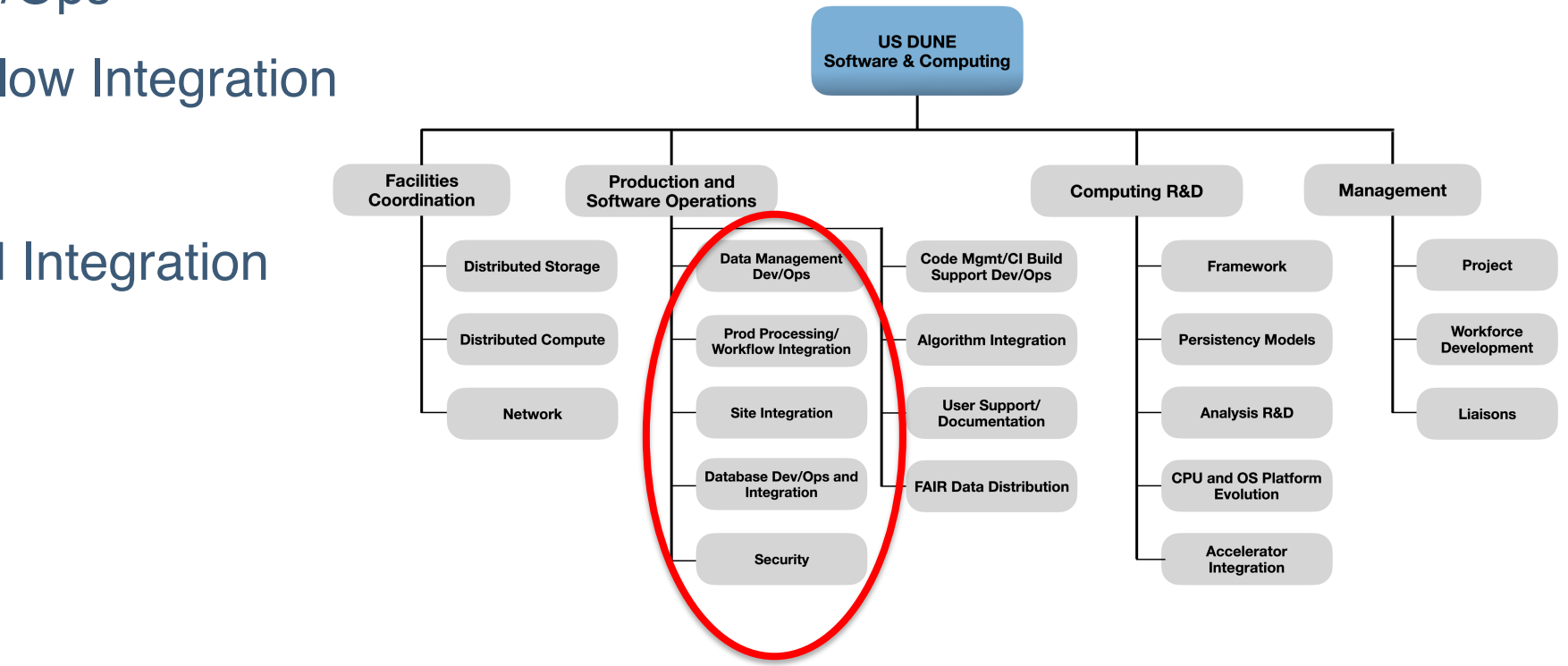
Doug Benjamin (BNL) on behalf of the DUNE computing consortium

US DUNE Preliminary Design Review: Software and Computing

July 13, 2023

Introduction

- Data Management Dev/Ops
- Prod Processing/Workflow Integration
- Site Integration
- Database Dev/Ops and Integration
- Security
- Data Challenges
- Effort
- Conclusion



Acknowledgements

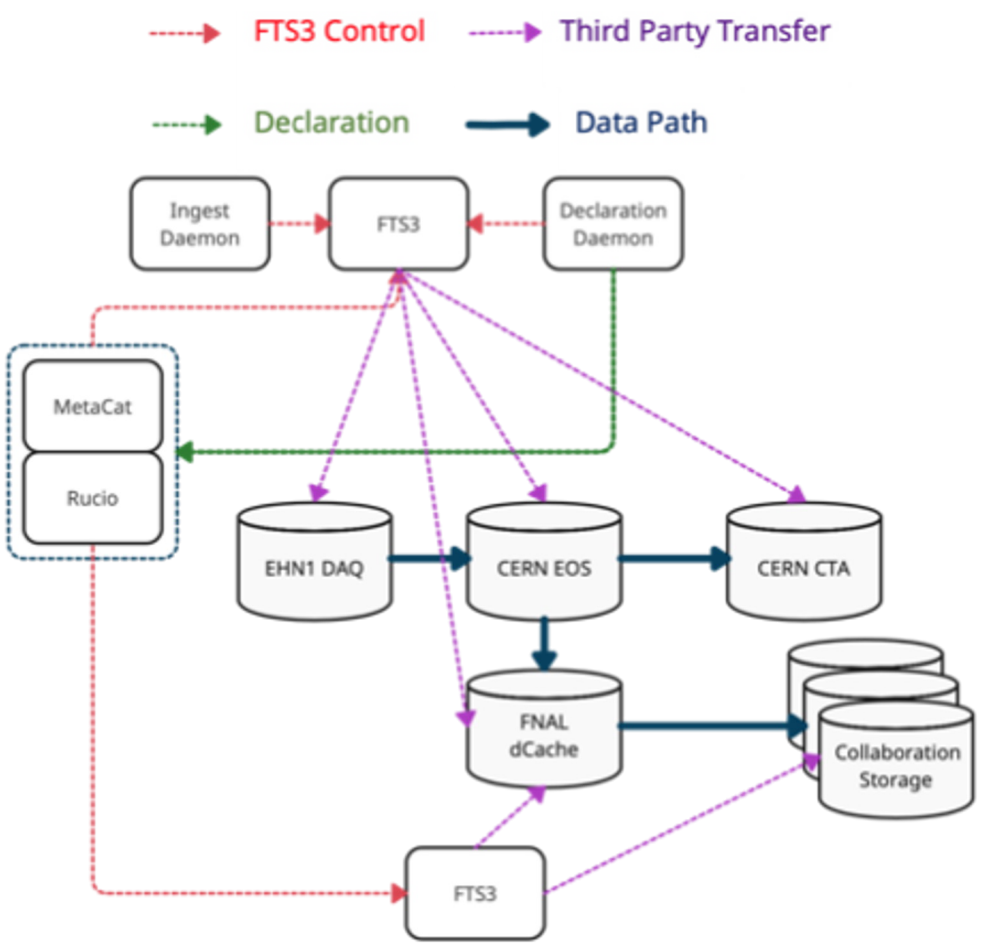
- Many Thanks to the following slide authors
- Database –Ana Paula Vizcaya Hernandez, Norm Buchanan, Lino Gerlach, Paul Laycock, Igor Mandrichenko
- Production – Elisabetta Pennacchio, Heidi Schellman
- Data Management – Steve Timm, Heidi Schellman
- Site Integration – Steve Timm
- Workflow – Andrew McNab, Mike Kirby

General Theme of this talk

- DUNE computing operations strives to avoid bespoke solutions to our computing operations.
- Use when ever possible common tools from across FNAL and HEP in general
- Add DUNE specific components ie “the special sauce”.

Data Management Dev/OPS

ProtoDUNE Run 2 and DUNE Data Pipeline Diagram

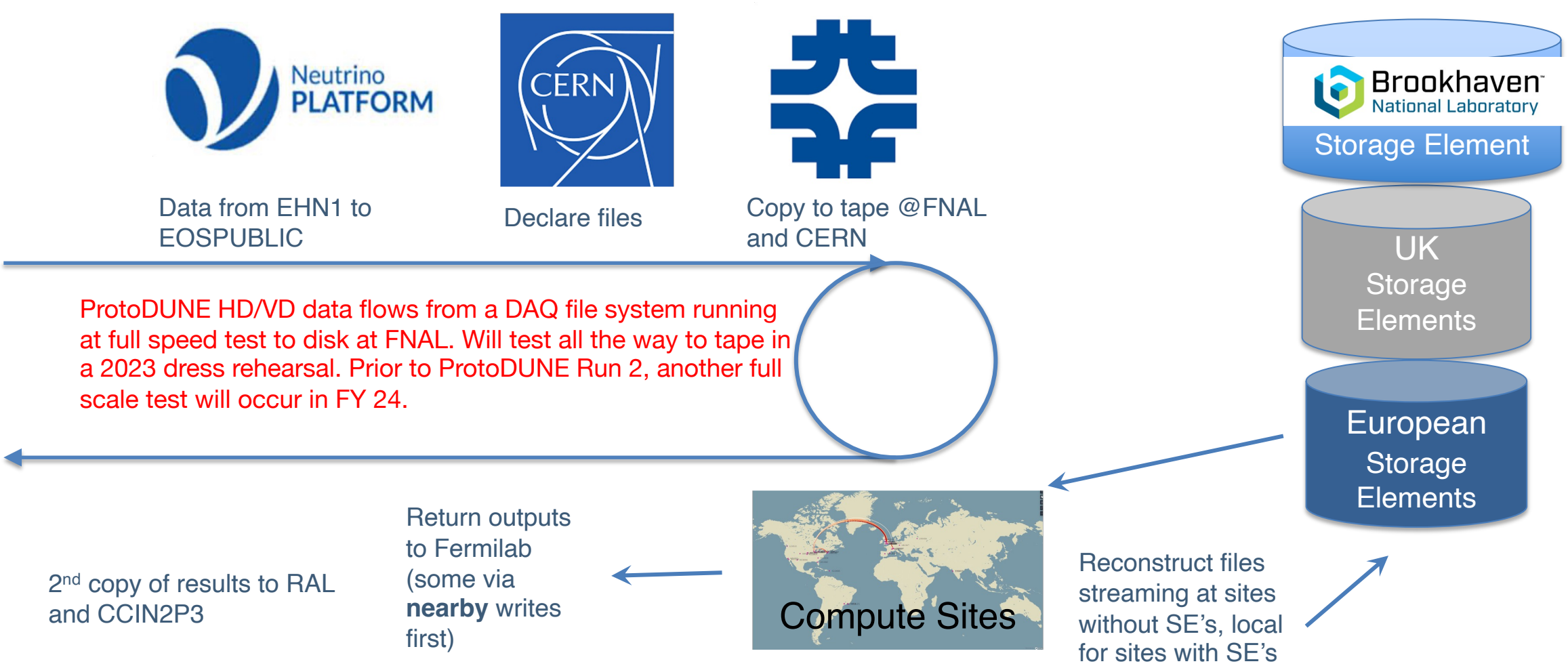


- Ingest Daemon and Declaration Daemon
  - Ingest daemon brings files from experimental systems to dropbox
    - Can operate without connection to Fermilab.
    - All transfers done via FTS3
  - Declaration daemon declares them to MetaCat and Rucio and makes rules to get them to the final destinations.
  - being adapted for 2x2 test beam at FNAL
- 2 copies of raw data on tape
- 1 copy of sim/reco on tape
- 2 copies of sim/reco on disk distributed across global storage elements.

Data Management tools

- Ingest Daemon and Declaration Daemon developed by FNAL Computing Division
  - Designed according to DUNE specifications
- MetaCat – meta data catalogue developed by FNAL for Intensity Frontier experiments
  - MetaCat documentation
- CERN File Transfer Service – (FTS3) – Software used to manage data transfers between Data Centers worldwide - FTS3 Documentation
- Rucio – Scientific Data management system – developed for LHC experiments now used by many scientific communities (HEP, NP, Astronomy) Rucio Documentation
- FNAL manages a FTS3 Instance and the DUNE Rucio service.
  - DUNE DDM operations
- 2nd FTS instance at CERN is also used by DUNE

ProtoDUNE HD/VD data flows



DUNE Phase I data flows from SURF

SURF

Declare files

Copy to tape @FNAL and CERN

Brookhaven National Laboratory Storage Element

UK Storage Elements

European Storage Elements

This data pipeline will be well tested before beam comes to both the DUNE Far Detector and Near Detector. Two copies of raw data will be on tape. Transient copies of data files for efficient processing will be on both in the US and non-US sites.

Reconstruct files streaming at sites without SE's, local for sites with SE's

Return outputs to Fermilab (some via **nearby** writes first)

2nd copy of results to other sites as dictated by DUNE policy.

Data Collections Manager

- Data comes in collections defined by:
  - Experiment (PDHD/PDVD)
  - Run conditions
  - MC or Data
  - Processing stage (raw->full reco)
  - Version
- Defining and documenting these samples depends on physics needs of the experiment
- Finite disk space means samples need to be prioritized

- This role interfaces with the physics coordination, physics groups, production team and data management to set and implement the policies/strategies.
- Experience with data analysis
- Physicist – member of DUNE collaboration
- User-centered approach to identify and define policies and procedures

University effort

Milestones: Data management (Prototype testing at CERN and FNAL)

- Summer 2022 - Commission data pipeline for ProtoDUNE-HD/VD. First tests **completed** Summer 2022 (FNAL/BNL/UK/FR/CERN)
- Summer 2023 - Data pipeline from FNAL to NERSC storage for 2x2 Demonstrator (FNAL/BNL/LBNL)
- Fall 2023 - Transition production data management from SAM to MetaCat/Rucio (FNAL/BNL/UK)
- Fall 2023 - Data challenges before ProtoDUNE-HD/VD beam operation (FNAL/BNL/UK/FR/CERN)

Milestones: Data management (DUNE Far Detector data taking)

- 2024-25 - Complete move to MetaCat/Rucio for all activities including user analysis
- 2027 - Full data management system in place for detector commissioning and simulation campaigns
- 2028 - Full data management system operations for physics and supernova data taking

Milestones: Data management (Near and Far Detector Physics with beam)

- 2028 - data management schema for ND data - complete

Production Processing/Workflow Integration

DUNE Production team

- The production team validates and runs large scale data and MC processing
- Vital for any physics study we do
- Can do this remotely but requires checking on jobs several times/day when running a “campaign”
- Need 6-10 people from the collaboration members to allow rotations and spread the load

- Team members learn how to run large jobs at scale
- Need knowledge of Unix/python commands
- Learn how big compute systems work
- Some operations funded labor will be used for continuity

Predominantly University effort

Production

Monte Carlo since January 2022

- VD 48deg
- VD 30deg+48 deg
- VD low energy
- HD Supernova neutrino production
- nucleon decay (ndk)+ 2nd reconstruction pass
- ProtoDUNE Prod4a
- ProtoDUNE-SP 1 GeV

111 Million events

Real Data since January 2022

- ProtoDUNE-SP CRT data taken at different E-fields for the diffusion analysis
- ProtoDUNE-SP cosmic data (same version of Prod4a)
- ProtoDUNE-SP Reconstruction of 1 GeV/c data with alternate space charge correction
- ProtoDUNE-DP 6m tracks
- VD cold box data reconstruction

28 Million events

In the following months

- MC production in both FD1-HD and FD2-VD
  - 24 M neutrinos split between both detector types - 6 months starting now.
  - The production is separated into LBL and Low energy samples
- MC production for ProtoDUNE-VD - expect a plan – Summer 2023
- MC production for ND

- Keep-up reconstruction campaign for ProtoDUNE-HD and ProtoDUNE-VD

The computing model assumptions:

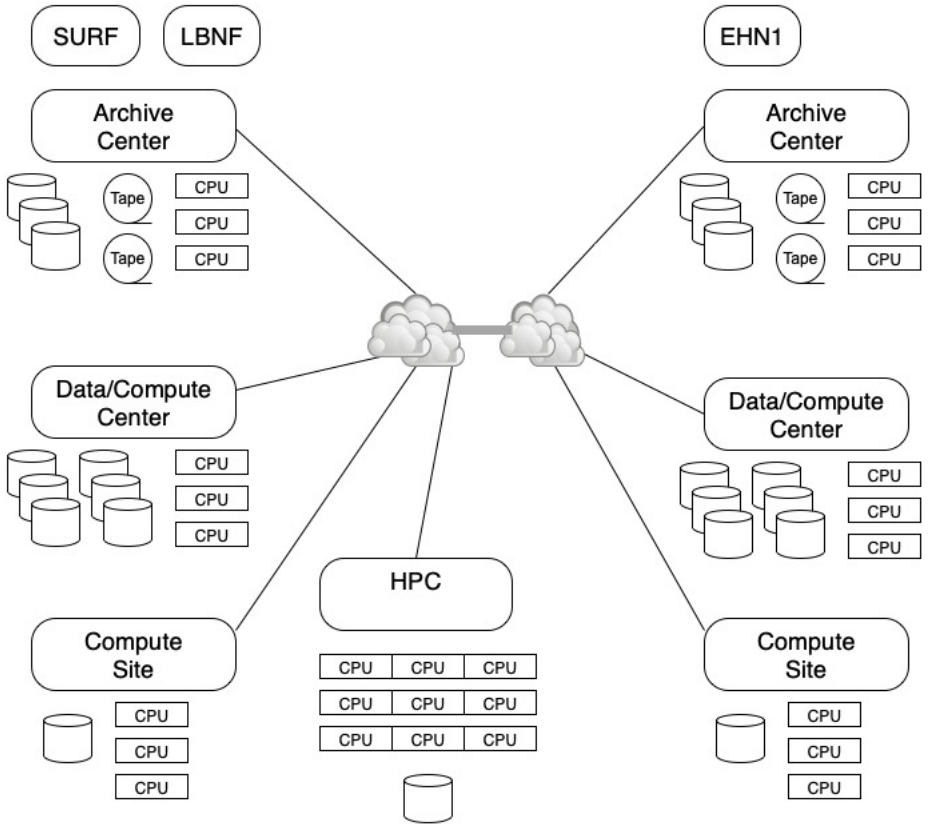
- one reconstruction pass per year over the full data sample
- one simulation pass per year

- Physics working group conveners can make request MC production or detector data reconstruction

Workflow Integration

- Workflow system - justIN (UK deliverable):
  - justIN designed according to DUNE requirements to manage processing campaigns for users, working groups, and the production team
  - Matches workflows to sites “near” unprocessed files
  - Uses a “just in time” philosophy (late binding)

- **Workflow system Integration – (US deliverable)**
  - Integrate justIN with the DUNE data management system (MetaCat/Rucio/FTS)
  - Use the DUNE global HTCondor pool via justIN
  - Integrate justIN w/ DUNE central monitoring



DUNE Computing (flat) model
Prioritizes local file access when possible
Will use “nearby” sites to access data as needed

Milestones: Prod Processing / Workflow Int. (Prototype testing at CERN and FNAL)

- Summer 2023 - start simulation campaigns based on delivery of software from Reco/Sim for September LBNC review (DUNE Production group (US ,FR,CA))
- Fall 2023 - Down select production integrated workflow/data delivery system based on testing and experience from ProtoDUNE-HD/VD (owner: UK + Production group (US,FR,CA))
- Winter 2023 - Move production processing to selected workflow/data delivery system (UK + Production group (US,FR,CA))
- Spring 2024 - run processing campaigns for either ProtoDUNE-HD or ProtoDUNE-VD (based on operating order) (DUNE Production group (US ,FR,CA))
- Fall 2024 - run processing campaigns for second beam run of ProtoDUNE-HD or ProtoDUNE-VD (DUNE Production group (US ,FR,CA))
- 2024-2028 Continued reprocessing campaigns

Milestones: Prod Processing / Workflow Integration (DUNE Far Detector data taking)

- 2024-27 - Annual initial simulation campaigns
- 2027 - Workflow system ready for simulation, reconstruction and analysis prior to FD data
- 2028-onwards - Annual full-scale simulation in anticipation of FD data
- 2028 - Tests of calibration in production and simulation
- Late 2028 - Start of data reconstruction and calibration

Milestones: Prod Processing / Workflow Integration (Near Det and Far Det Physics with beam)

- 2024-29 - Iterative simulation campaigns of ND design
- 2029 - Tests of calibration and detector conditions in production
- 2029 - Full workflow for simulation/reconstruction of ND data
- 2030-onwards - Full scale simulation in anticipation of ND data
- 2031-onwards - Reconstruction of data

Site Integration

US DUNE Site Integration

- DUNE has had a unified weekly site management / operations meeting since fall of 2018 (US and Foreign sites together)
- Initially focused on onboarding
- Currently covers both facilities and operations
- Short meeting focused on major changes that will affect sites, and operations issues.
- Trying to figure out a unified way to track site contacts and ticketing systems.

Milestones:

- Technical Evolution Activities/Milestones
  - Transition SL7-> Alma 9 2024
  - Unified documentation for DUNE-specific configurations for compute and storage. 2024
  - Improve ticketing system 2025-2026
  - Unified interface to downtime system 2026-2027

Site Integration

Milestones: (Prototype testing at CERN and FNAL)

- 2018-2023 Integrate global sites (WLCG, OSG, etc.) for ProtoDUNE-HD/VD simulation and production - largely complete
- Summer 2023 - Integration of production operations at NERSC for 2x2 Demonstrator workflows - in progress (FNAL/LBNL)

Milestones: (DUNE Far Detector data taking)

- 2027- 2028 – integrate SURF systems into offline computing (FNAL/SURF)

Milestones: (Near and Far Detector Physics w/ beam)

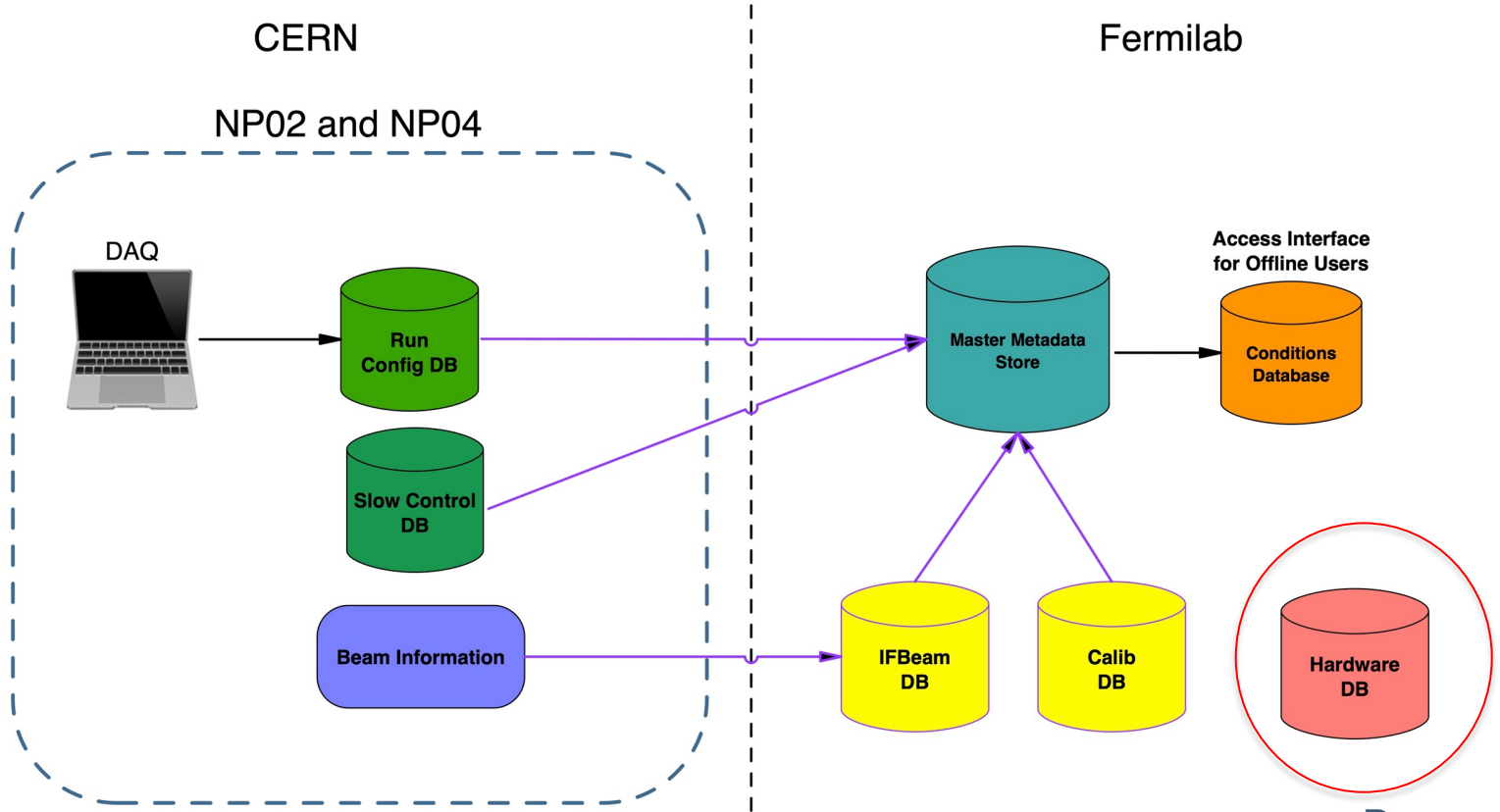
- 2030 Provision specialized resources for analysis of near detector data

Database Dev/Ops and Integration

ProtoDUNE database system

Strong University involvement

- DUNE will produce vast amounts of **metadata**, which describe the data coming from the read-out of the primary DUNE detectors.
- An unstructured database (uconDB) holds the **master store** of metadata which collects info from all databases, allowing for maximal flexibility.



Same design for DUNE database system

Delivered on schedule ahead of detector construction

Dune specific component Developed at Colo. State U

Brookhaven National Laboratory

(Offline) Conditions database

The subset of all metadata that is accessed during offline data reconstruction and analysis is referred to as **conditions data** and it is stored in a dedicated database

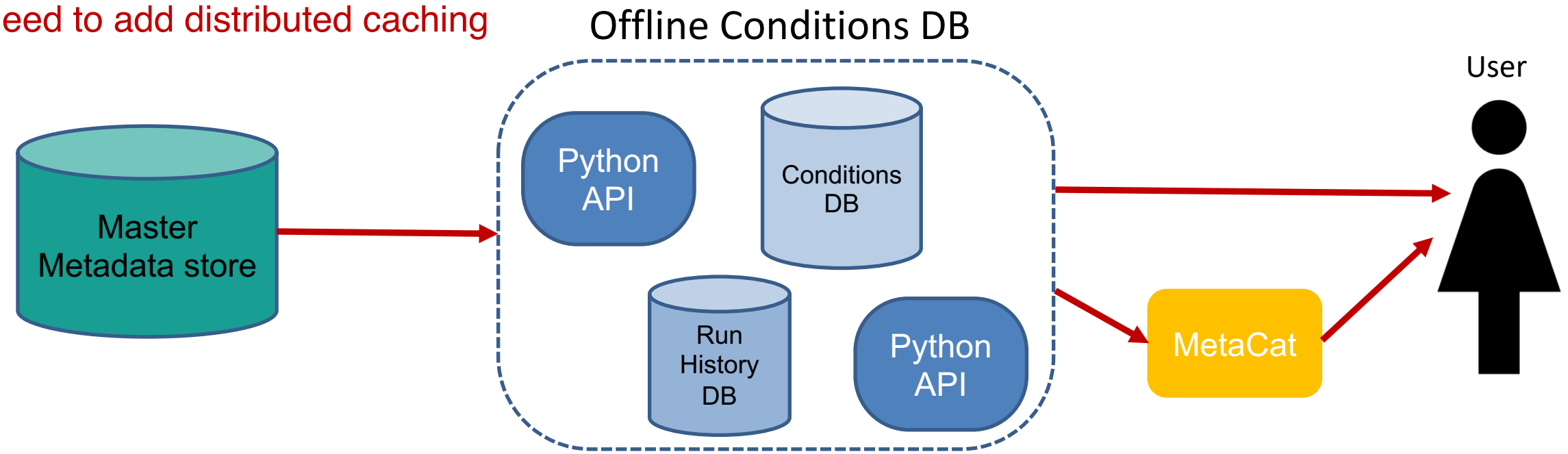
Purposes:

- Provide all of the non-primary data stream needed for offline data processing
- Identify runs which fulfill specific configurations, like a chosen high voltage setting, with the aim of file discovery.

Most of the data flow has been tested. Similar approach was used during ProtoDUNE run I

Need to add distributed caching

.

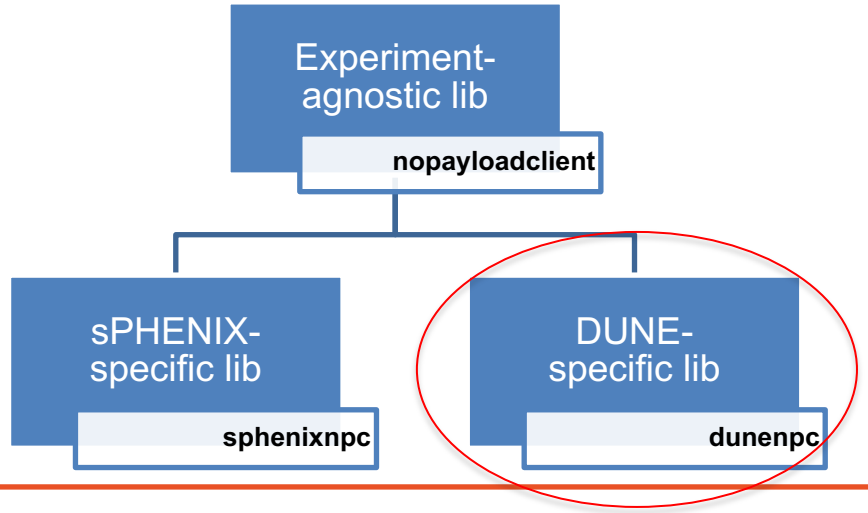


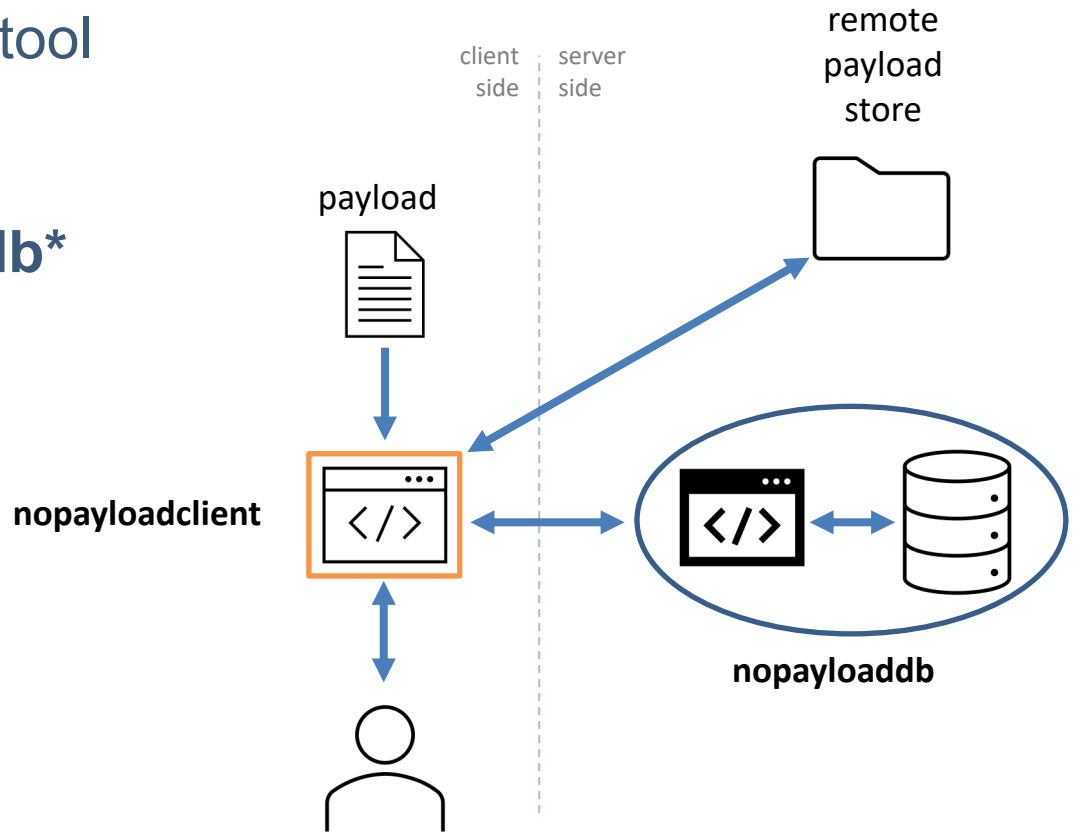
Conditions Database modernization

HSF Recommended design

nopayloadclient: client-side stand-alone C++ tool

- Experiment agnostic
- Communicates with server-side **nopayloaddb***
- Local caching
- Handling of payloads





Example of Dune specific component

Brookhaven National Laboratory

Milestones: Databases (Prototype testing at CERN and FNAL)

- Spring 2023 – Hardware database in production (DUNE Database group) – **COMPLETED** – Spring 2023
- Early 2024 – Runs database in production (DUNE Database group, US Univ, FNAL, BNL)
- Early 2024 – Conditions database in production (DUNE Database group, US Univ, FNAL, BNL)
- Fall 2024 - Explore scalability of Conditions DB solutions with ProtoDUNE-HD/VD data and simulation (DUNE Database group, US Univ, FNAL, BNL)

Milestones: Databases (DUNE Far Detector data taking)

- 2027-28 – Final Integration of FD slow controls/run configuration prior to FD commissioning
- 2028 – Full integration of calibration and detector conditions systems

Milestones: Databases (Near and Far Detector Physics with beam)

- 2029-30 – Final Integration of ND slow controls/run configuration prior to ND commissioning
- 2030 – Full integration of calibration and detector conditions systems

Security

Computer Security

- DUNE security policy inherits from FNAL security policy
- Security alerts come from FNAL and from partner compute grids – OSG and WLCG
- DUNE security contact (currently S. Timm FNAL) follows up as necessary.

Bringing it all together via Data Challenges

- Data Challenges (Dress rehearsals) done on a periodic cadence are perfect tools for exercising multiple parts of the system.
  - Fall 2023 - Data challenge of one-week of operation at the scale of ProtoDUNE-HD/VD beam operations including keep-up processing
  - 2024 - Combined data challenge with WLCG at 25% of expected FD data rate.
- Increasing complexity including supernova tests as we get close to FD data taking
- Near Detector simulation/reconstruction data challenge with simulated data in 2028-2029
- University effort required to make the data challenges a success

Milestones: Data Challenges (Prototype testing at CERN and FNAL)

- Fall 2023 - Data challenge of one-week of operation at the scale of ProtoDUNE-HD/VD beam operations (FNAL/BNL/UK/FR/CERN)

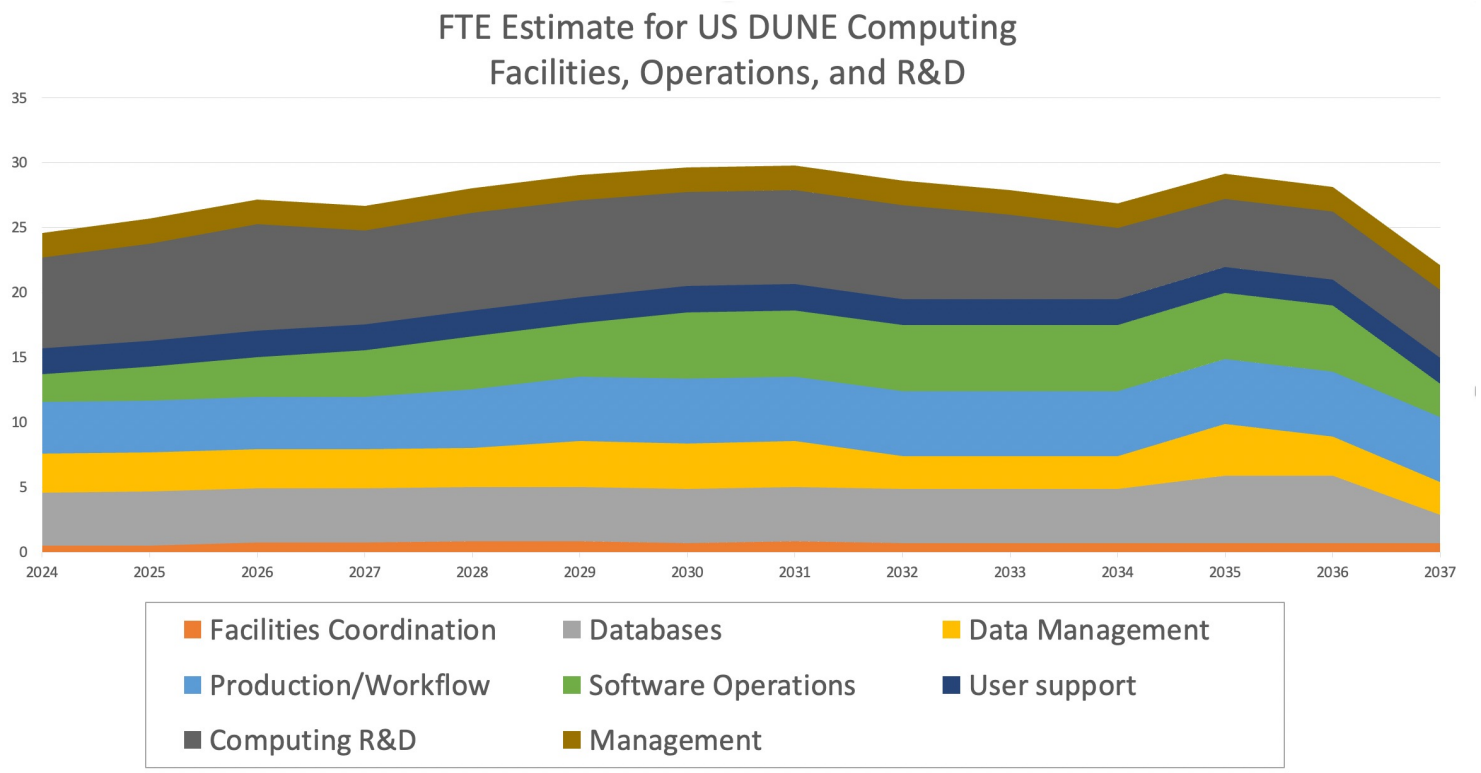
Milestones: Data Challenges (DUNE Far Detector data taking)

- 2024 - Combined data challenge with WLCG at 25% of expected FD data rate (ES-NET /DUNE/CERN)
- 2025-28 - continued data challenges leading up to full supernova tests.
- 2028 - Data challenge with 100Gbs data from SURF→FNAL→ sites
- 2027-29 - Data challenge of analysis of distributed data using workflow systems

Milestones: Data Challenges (Near and Far Detector Physics with beam)

- 2028-29 - ND simulation/reconstruction data challenge with simulated data
- 2029-31 - Data challenge of analysis of distributed data using workflow systems
- 2030 - Full scale parameter estimate challenge on simulated data

Operations Effort



(Effort in FTE)	Next Year 2024	Start of FD Ops 2029
Data management Ops	1.5	2.0
Data management Dev/Int	1.5	1.5
Production Processing Ops	1.5	2.0
Workflow & Monitoring Ops	1.5	2.0
Workflow Dev/Int	0.5	0.5
Monitoring Dev/Int	0.5	0.5
Site Integration	0.1	0.1
Database Administration	0.1	0.2
Database Ops	1.0	3.0
Database Dev/Int	3.0	1.0
Security	0.1	0.1
Total	11.3	12.9

Conclusions

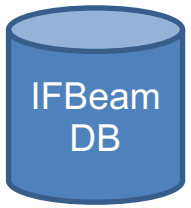
- Most of the components of the DUNE computing system have been tested ProtoDUNE Run 1 data and subsequent processing.
- DUNE uses common tools from across FNAL and HEP in general adding DUNE specific components to provide solutions.
- ProtoDUNE Run 2 will be used to further test additional features.
- Additional effort is needed to get ready for data from SURF

Backup material

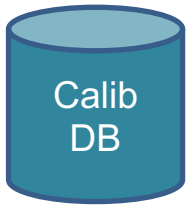
Overview of ProtoDUNE databases

- Run Configuration DB
  - DAQ and wibs front end electronics configurations
  - Metadata collected, sent to the UconDB, and sent to the conditions DB

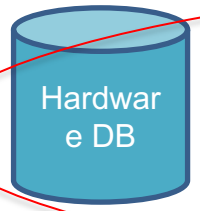
- Slow Control DB
  - Device values stored at high rate giving O(10 GB)/day
  - Subset will be moved to UconDB, we need better understating of sensor list



- IFBeam DB
  - Beam related device metadata – subset will be sent to UconDB



- Offline calibration
  - Calibration constants derived from data
  - ProtoDUNE run I calibrations stored at the conditions database



- Hardware DB
  - Completed and available for use (documentation and training have been provide)
  - Utilizes FNAL system created in 2008, with updates for DUNE requirements
  - iOS interface and Python APIs for user insertion/extraction

Dune specific component

Conditions Database - Introduction

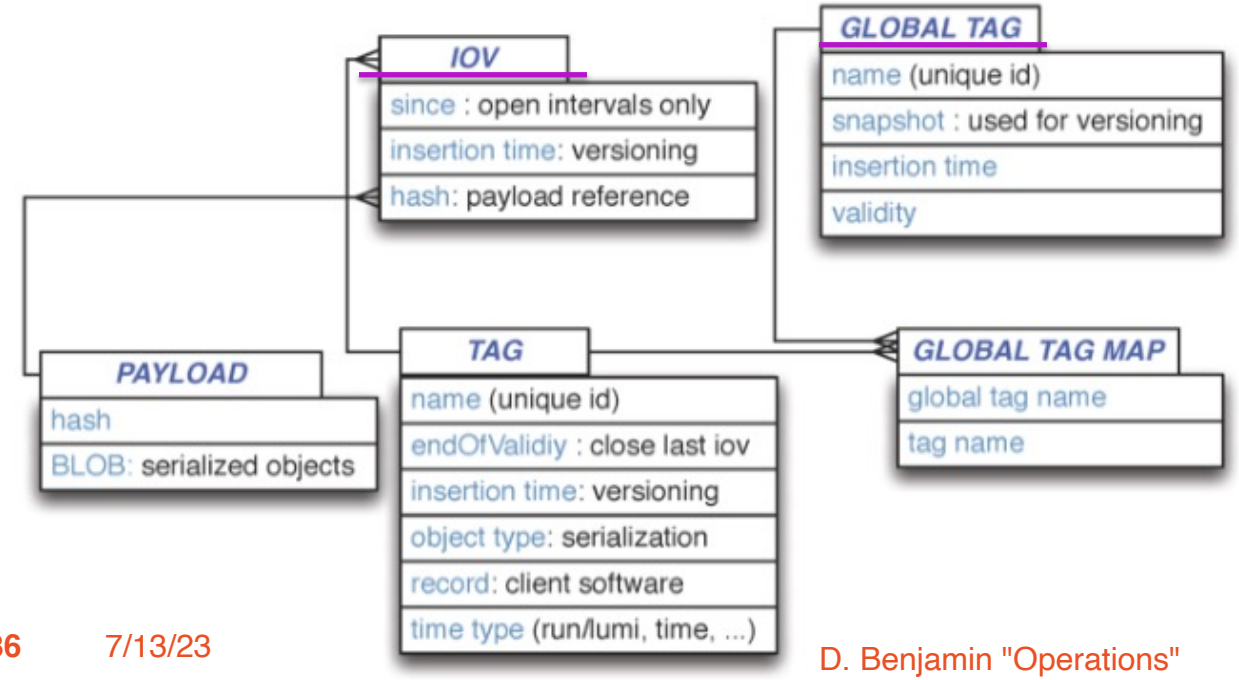
Purpose: provide all of the non-event data stream needed for offline data processing

Changes over time	High access rates	Heterogenous data
• Repeat detector calibration with larger cosmic dataset • Improve calibration algorithms	• Distributed computing jobs access same conditions data simultaneously • Access rates up to ~kHz	• Granularity varies (time indexed, run-indexed, constant) • Structure of payload varies (3D map, single number, …)
Versioning & configuration	Fast DB queries & effective caching	Payload agnostic by design

Similar challenges for various HEP experiments

Conditions Database – HSF Recommendations

- Dedicated HEP Software Foundation (HSF) conditions data activity: https://hepsoftwarefoundation.org/activities/conditionsdb.html
- Key recommendations for conditions data handling
  - Separation of payload queries from metadata queries
  - Schema below to enable appropriate configuration



HSF
HEP Software Foundation

HEP Software Foundation
Community White Paper Working Group – Conditions Data

HSF Conditions DB Integration

- Deploy **nopayloaddb** on Kubernetes container orchestation system (OKD) @ FNAL
- Publish **nopayloadclient** & **dunenpc** on /cvmfs/
- Organize existing conditions data in payloads (files)
  - Publish them on /cvmfs/
- Allow seamless change between classic and HSF design
  - Unify data access API across the two
  - Develop additional API layer in LArSoft / art
- Use ProtoDUNE-HD/VD DB data to compare the performance between the designs

justIN workflow operations

- Automated Workflow Tests
  - justIN sends out AWT test jobs to all sites every day
  - Each job tries to download and upload with each storage
  - Dashboard displays results, with links to drill down to site or storage history, and individual job logs
- Future plans
  - Use justIN's global overview to impose real-time bandwidth quotas on use of sites and storages, by selecting
  - Impose limits on the scale of campaigns that individuals and working groups can start without some form of review
  - Need to ensure proper testing of large campaigns, planning of resource usage, realistic tape recall campaigns etc …
  - … while still supporting smaller scale "chaotic" user analysis