DUNE Computing Projects

Michael Kirby, Fermilab/SCD

July 2, 2020

DUNE Projects in Service Now

- Not all of these fall under the guise of the Computing Consortium (DAQ, ARAPUCA, Speakers Database)
- Some projects not part of Service Now yet - Conditions and Slow Controls DB
- As well, there are several projects happening outside of SCD/FNAL (DB interface, Documentation, Tutorials, etc)

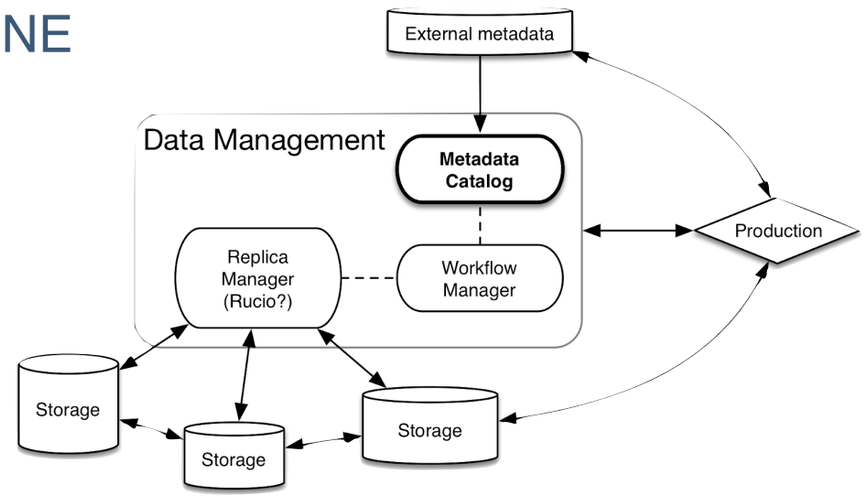
Four main thrusts of projects

- Data management
- Databases
- DAQ
- Production/Ops
- Frameworks (counting is not my strong suit)

Slides from Steve Timm with help from Robert Illingworth

DUNE Data Management Project work @ FNAL

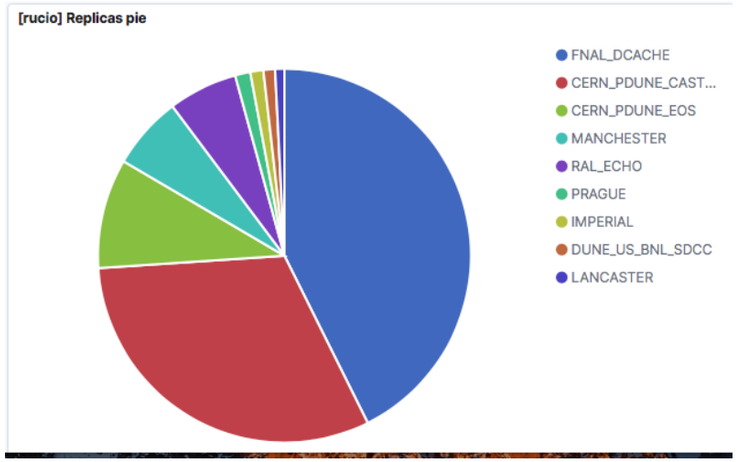
- Overall goal: replace all functions of SAM in time for ProtoDUNE II run, currently scheduled Jan 2022.
- File Replica Catalog and Data Movement-> Rucio
- Metadata Catalog -> New metadata server “MetaDB”
- Project Management (Workflow)-> Currently being specified
- Data Ingest (FTS replacement) -> Currently being specified
- Already other customers identified for both Rucio and MetaDB.
- Biggest driver is proposed DUNE data structure in which one trigger record may be split across many files and/or parts of many trigger records may be contained in one file.



Slides from Steve Timm with help from Robert Illingworth

Rucio Development and Deployment

- Features needed to transfer into tape-backed “non-deterministic” path hosts
- configuration flexibility (J. Perry)
- Quality of Service
- Deployment of OKD-based containerized Rucio [when] (Brandon)
- Adding significant monitoring
- In current deployment Rucio used to distribute files from FNAL and CERN to 16 other Rucio Storage Elements around world.
- Automation of more types of data ingest into Rucio.



Fermilab

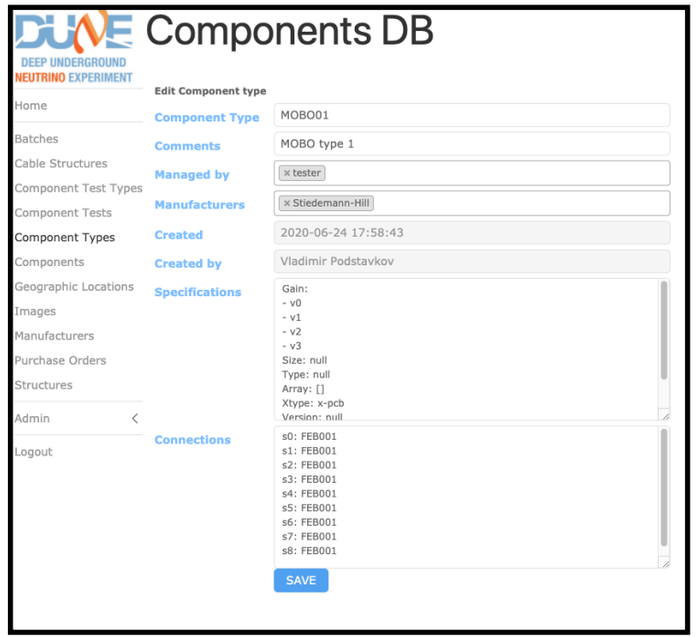
Slides from Steve Timm with help from Robert Illingworth

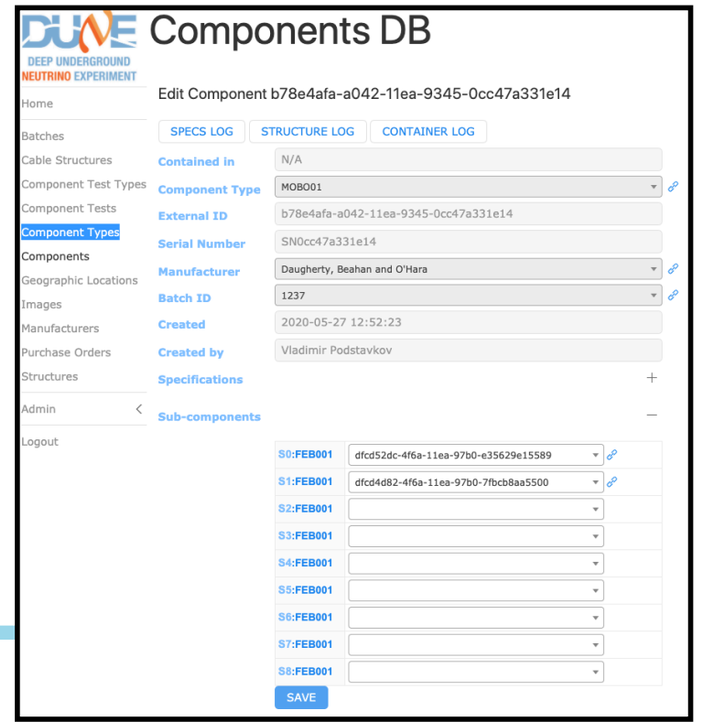
Metadata server MetaDB (Igor M.)

- Requirements document complete and being reviewed by stakeholders
- Proof of principle demonstration has been constructed and successfully ingested all existing DUNE Metadata
- Joint taskforce between DAQ, Data Mgmt, DB, and Offline is being formed to specify the contents of the metadata
- Remaining SAM functionality (sam project, sam get_next_file, etc) currently being specified.
  - This is not meant to be a full workflow management solution for DUNE, rather just a way to replace current SAM project and file location delivery solutions.

Databases

- HardwareDB (Components DB in SCD terms) - Steve White, Vladimir Podstavkov
  - We have a development (prototype) server up and running and soon to be available to DUNE DB members -- https://rexdbsrv4.fnal.gov:9443/cdb/login/
  - Working on toughest part of the web forms - This is very near done – 90%.
    - Adding Component types
    - Adding Components, includes adding child components and traversing the entire component structure.
  - We will be duplicating the above work for tracking testing of components.
  - Dynamic form generation, based on the json fields is nearly complete.
  - Security
    - who can login is currently ldap – because its easy but will be Shibboleth
    - who can edit what component and role is working for component types and compnents.
  - Just starting on design of an API for adding/updating components via the web and possibly a CSV file.
- Conditions and Slow Controls Databases - Igor Mandrichenko





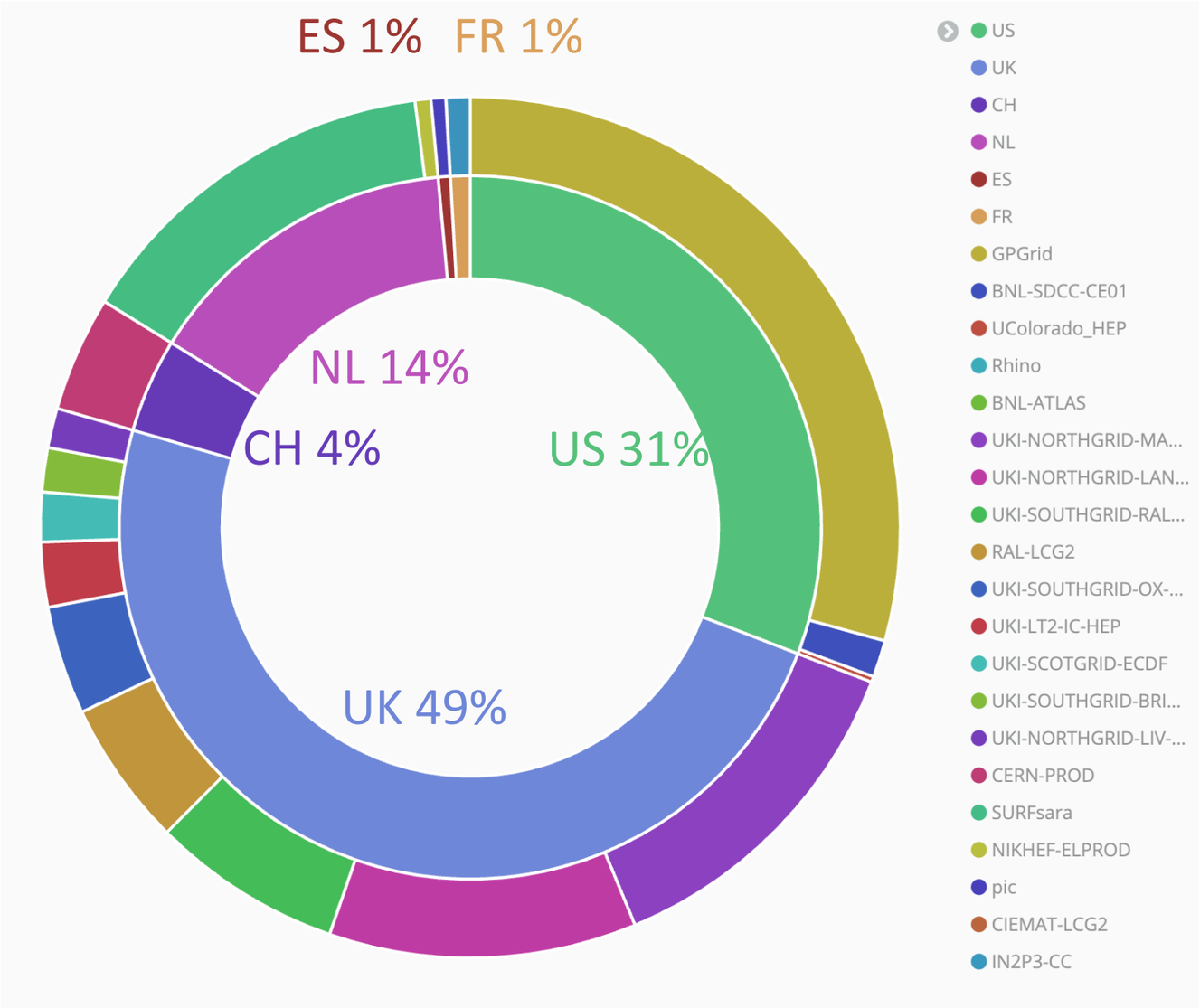
DAQ Work - Kurt Biery

- protoDUNE Single Phase will be drained of LAr around 20-July-2020
- protoDUNE-II will run in 2022(?), DAQ goal is to be ready by Sep 2021
- v1 of the DAQ Application Framework released last week. Good start, but not full functionality. Inter-process communication interfaces and wrappers to existing technologies next. Logging and configuration also next. [Eric Flumerfelt, John Freeman, Ron Rechenmacher]
- software management topics like builds and deployments under active discussion & development [JohnF, Pengfei]
- HDF5 data format demonstrated at PDSP; needs further joint offline/online validation; goal is to pick a PD-II raw data format in Nov 2020 [Kurt, Kirby, others]
- online computing coordination (e.g. sysadmin of PDSP DAQ servers at CERN) going well [BonnieK, Pengfei]
- good interactions with facilities and networking about needs and plans (space and power still quite tight) [Bonnie]
- requirements being gathering for DAQ storage and compared to existing technologies

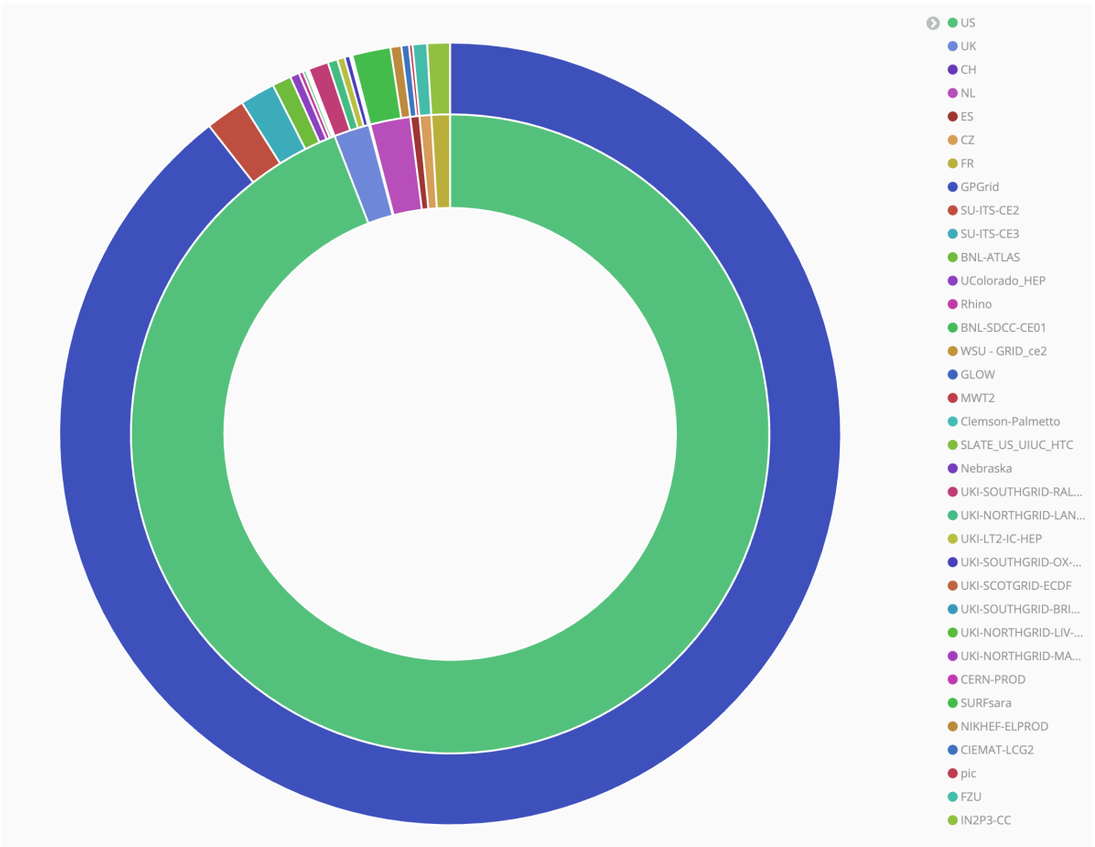
Software Framework Requirement Task Force - Andrew Norman

- Finalizing transformation from a use case document into a requirement document
- most requirements are inline with current frameworks
- some interesting requests coming from Dave Adams
- memory management/clean up requirements are interesting - large event size makes for discussion of transient memory objects and intermittent clean up during processing
- anticipate a first version at the end of this week

Production - Ken Herner



Analysis



Fermilab