Fermilab U.S. DEPARTMENT OF ENERGY Office of Science

Analyzing how we do Analysis and consume data Results from the SciDAC-Data Project

L. Aliaga, P. Ding, A. Tsaris, M. Mubarak, A. Lyon, A. Norman, R. Ross

CHEP 2016

10 October 2016

In partnership with:

Argonne NATIONAL LABORATORY

SciDAC-Data Project

Scientific Discovery through Advanced Computing

SciDAC-Data is a specially designed program within the Office of Science of the U.S. Department of Energy:

- *to develop the Scientific Computing Software and Hardware Infrastructure needed to use HPC computers.*

- *to advance DOE research programs in basic energy sciences, biological and environmental research, fusion energy sciences, and high-energy and nuclear physics.*

http://www.scidac.gov

SciDAC-Data Project at Fermilab

- Analyze more than **8.7 M datasets** of high energy physics collected, generated and defined in experiments at Fermilab.

- The analytics and meta information for these datasets and analysis projects are being combined with knowledge of their part of the HEP analysis chains for major experiments to understand how modern computing and data delivery is being used.

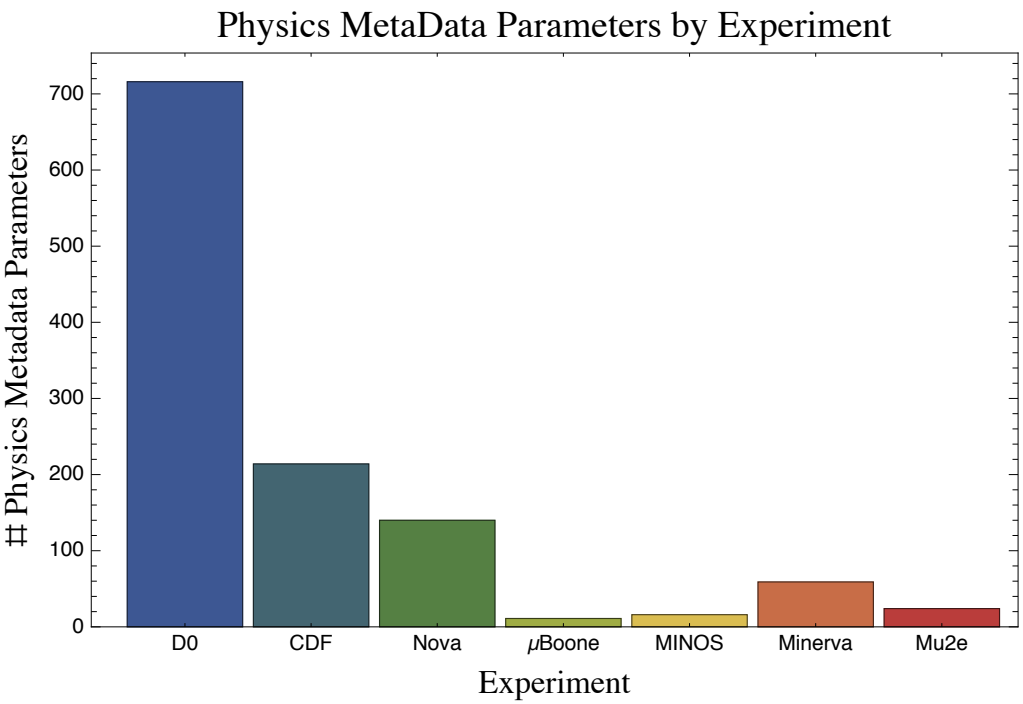
- In this talk, we present the first results of this project on how the **CDF, D0 and NOvA, MicroBooNE and MINERvA** experiments have organized, classified and consumed their datasets to produce their physics results.

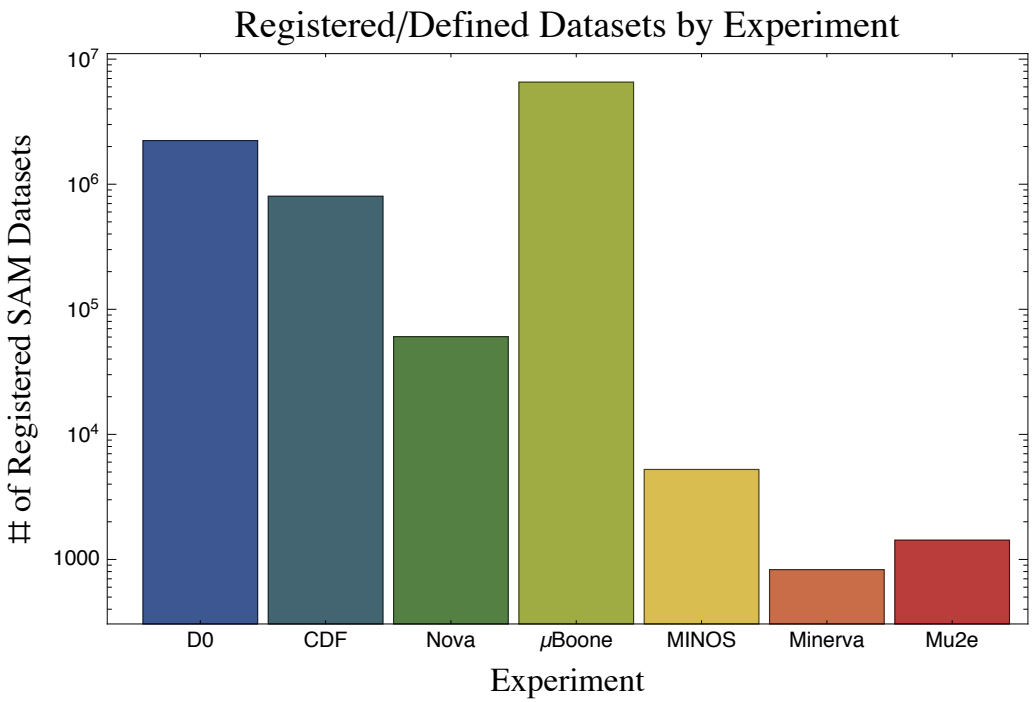
SAM

- SAM (**Sequential Access via Metadata**) is a unified data catalog / replica catalog
  - *Specially, optimize the restoration of data from slow archival media (tape) and manage or interact with disk caching layers which front the slow archival systems.*
  - *Provides organizational associates of data files into "**datasets**" based on metadata selection criteria using the defined **parameters**.*
  - *This collection is referred to as a "snapshot" and represents the deterministic evaluation of the dataset on at a specific point in time.*

Metadata Parameters and Datasets by Experiments

- These correspond to data taken at Fermilab from 2004-2016 for different experiments:





- *These parameters are used to create the definitions of datasets used in the SAM data handling system.*
- *The total number of definitions that have been defined by each experiment.*

Analysis Using SAM

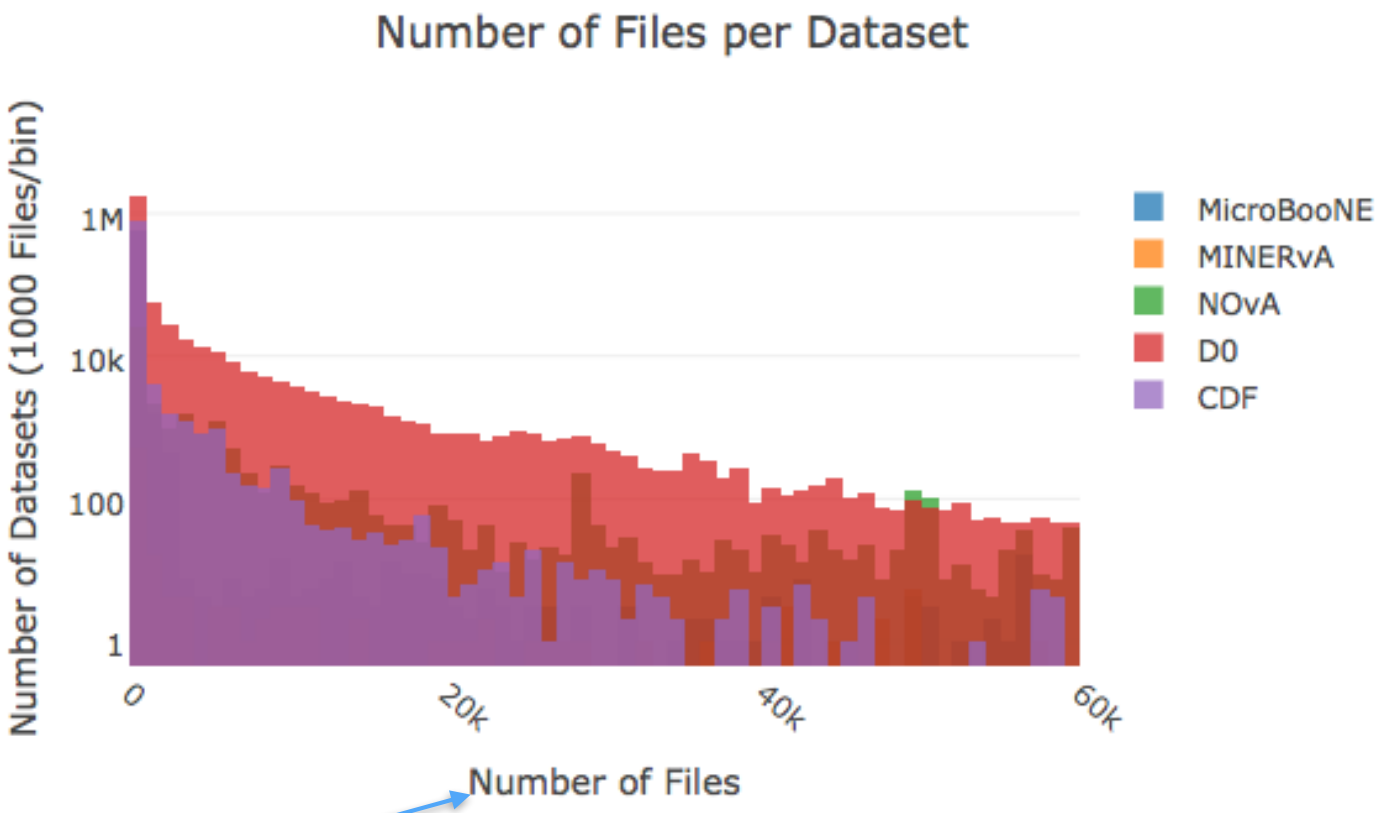
- SAM workflow management and bookkeeping perform an evaluation and start a corresponding "**project**".

- Each analysis client connects to the project and requests "the next file" in the dataset that is available for analysis.

- Files are delivered in an optimal order for any high storage systems. Cache layers can be leveraged and predictive pre-staging are done to keep pipelines near full capacity.

Data Organization

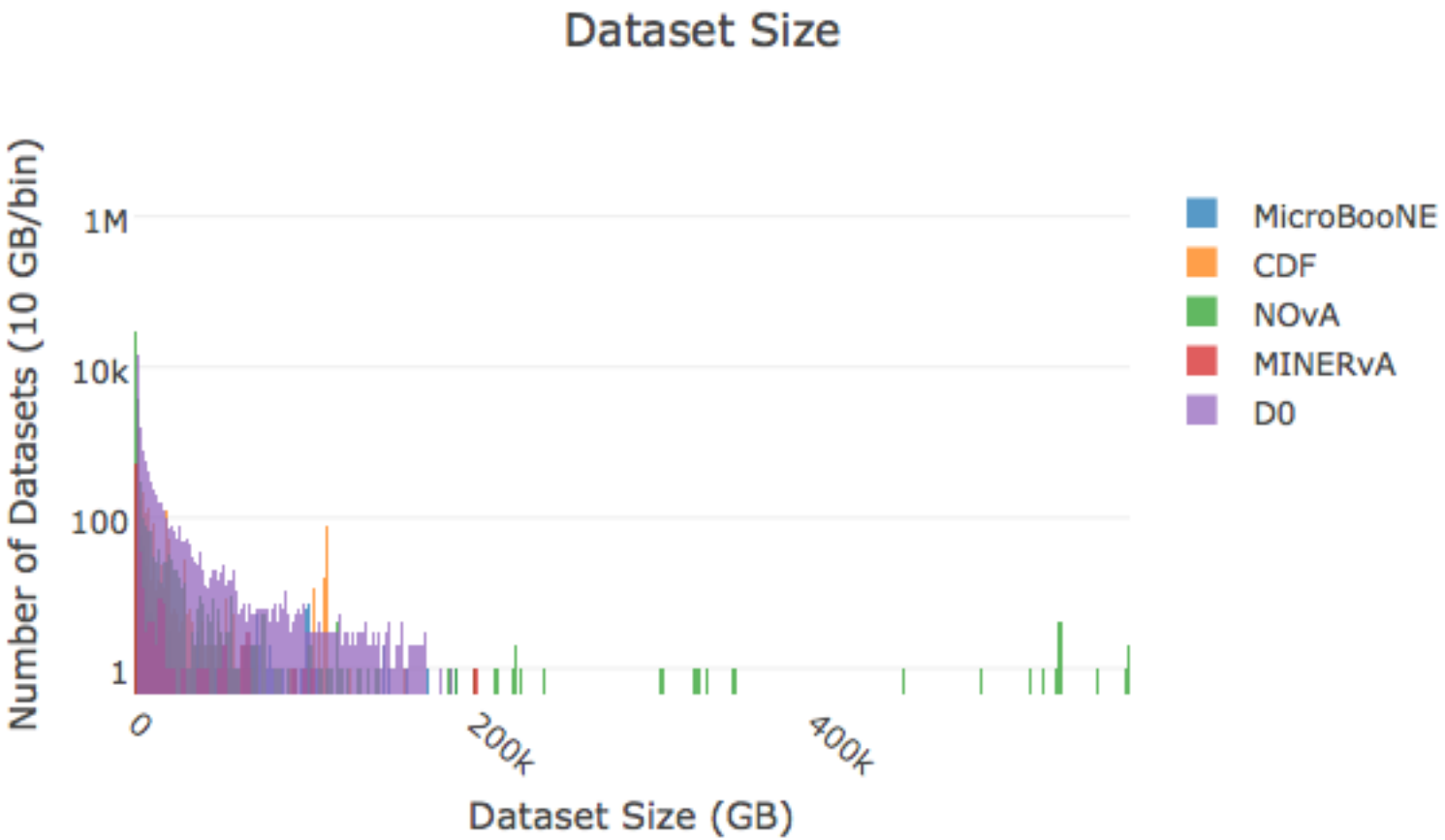
- Distribution representing the data organization within the CDF, D0, NOvA, MicroBooNE and MINERvA experiments.



Number of unique files that belong to each dataset that was defined by the experiments

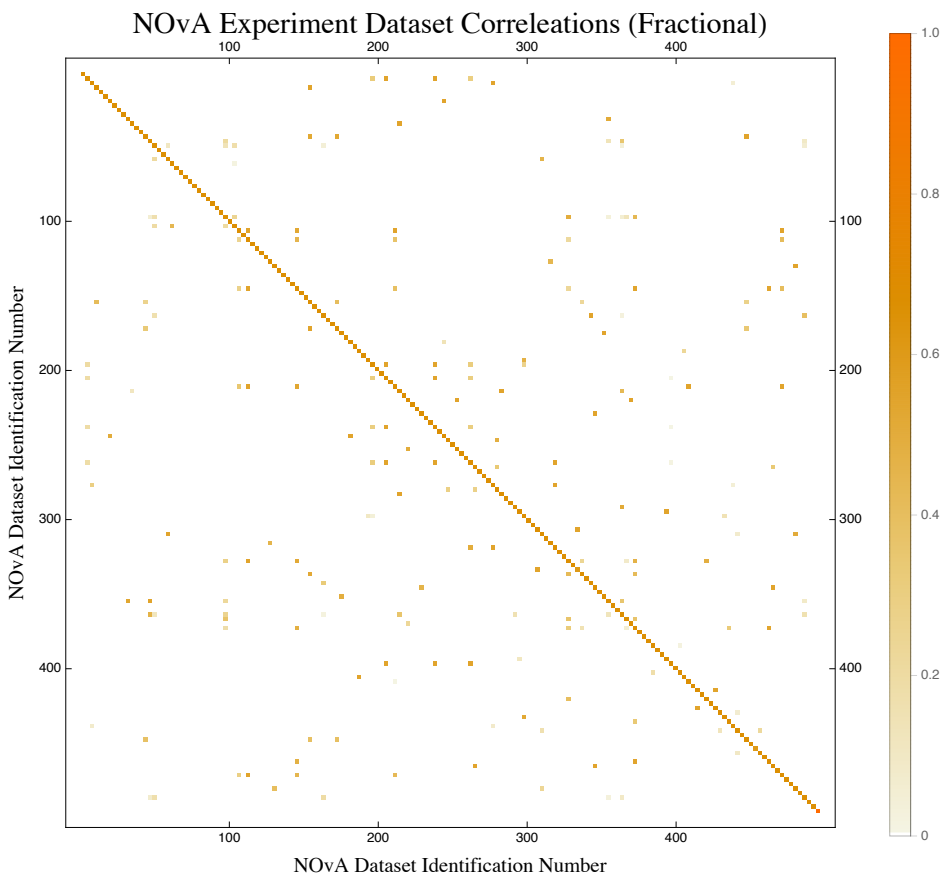
Dataset Size

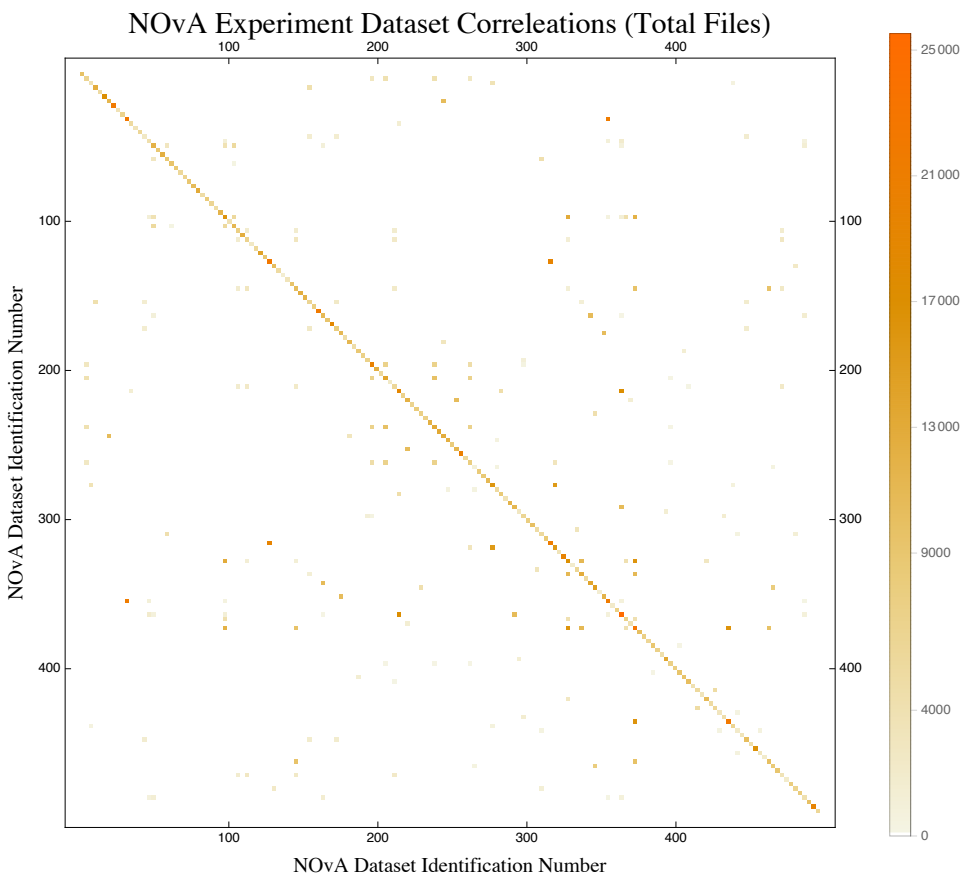
- Distribution of the total size of datasets defined and used by CDF, D0, NOvA, MicroBooNE and MINERvA experiments.



Dataset Correlation for NOvA

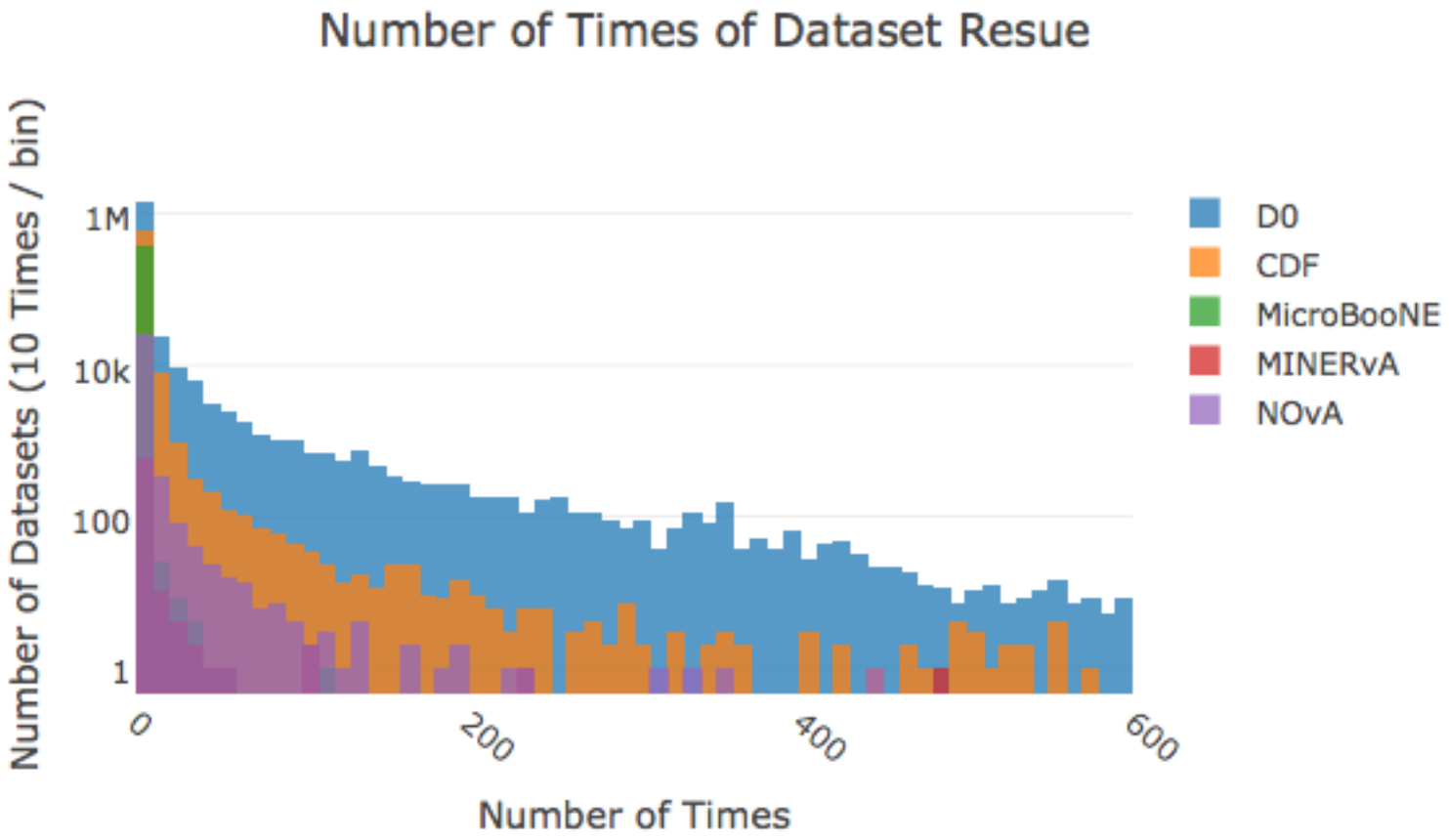
- Correlation matrix representing the fractional and absolute overlap (in number of files) between all datasets for the NOvA experiment.





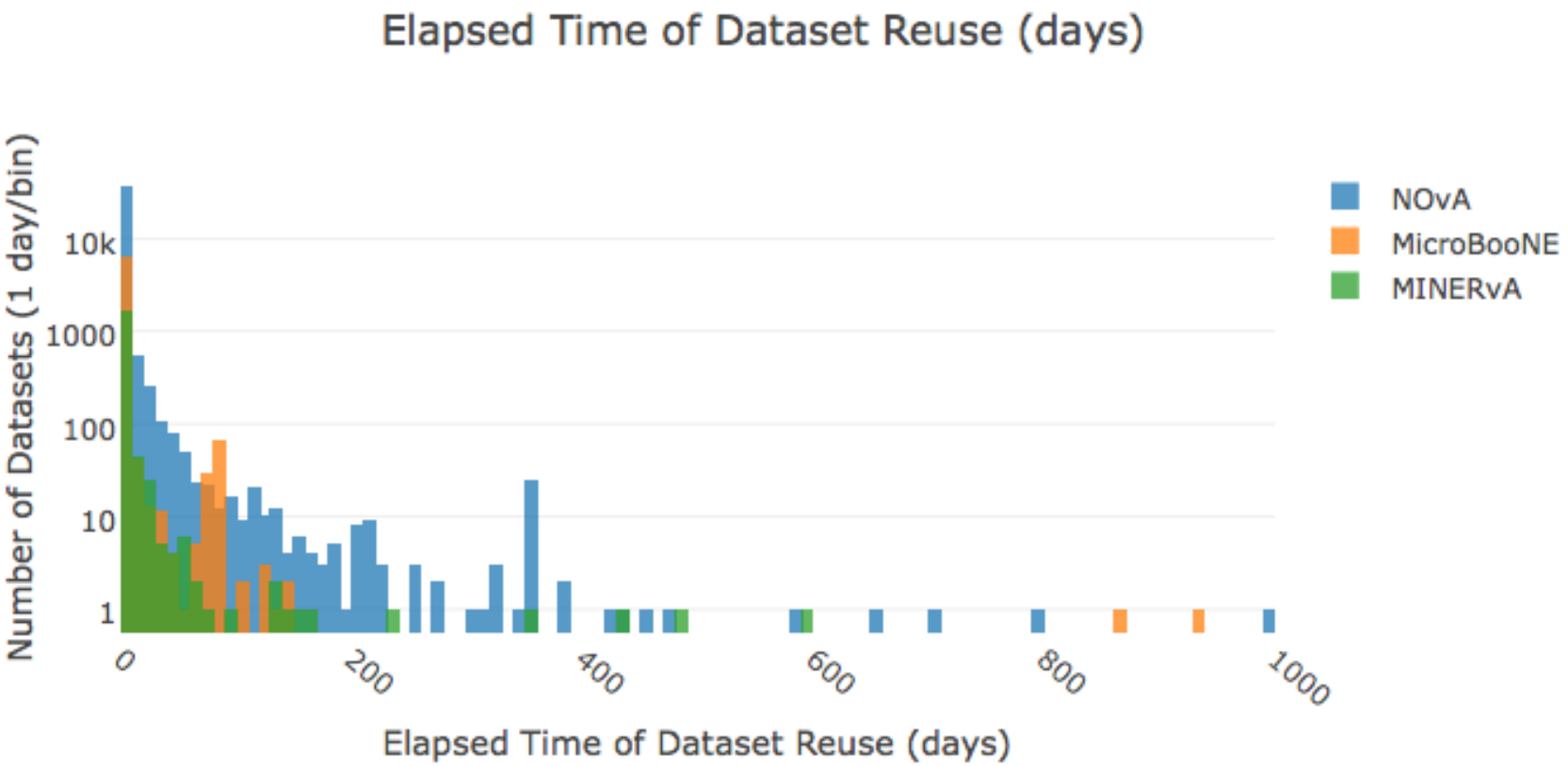
Dataset Usage

- Distribution of the number of times that a given dataset was processed by an analysis application for the NOvA, MINERvA, MicroBooNE, CDF and D0 experiments.



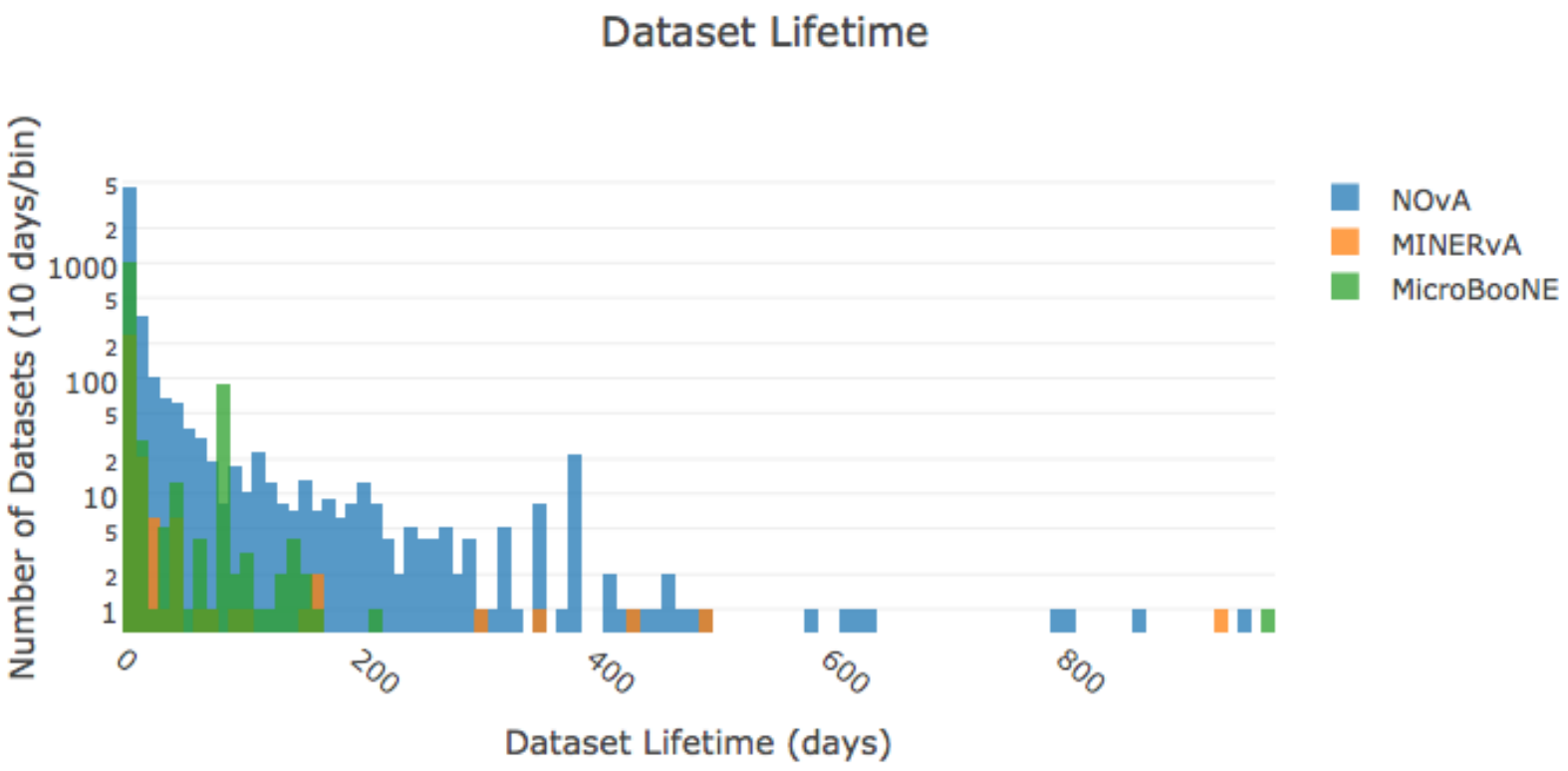
Elapsed Time of a Dataset Usage

- Distribution of the elapsed time between successive analysis projects being started on a dataset for the NOvA, MicroBooNE and MINERvA experiments.



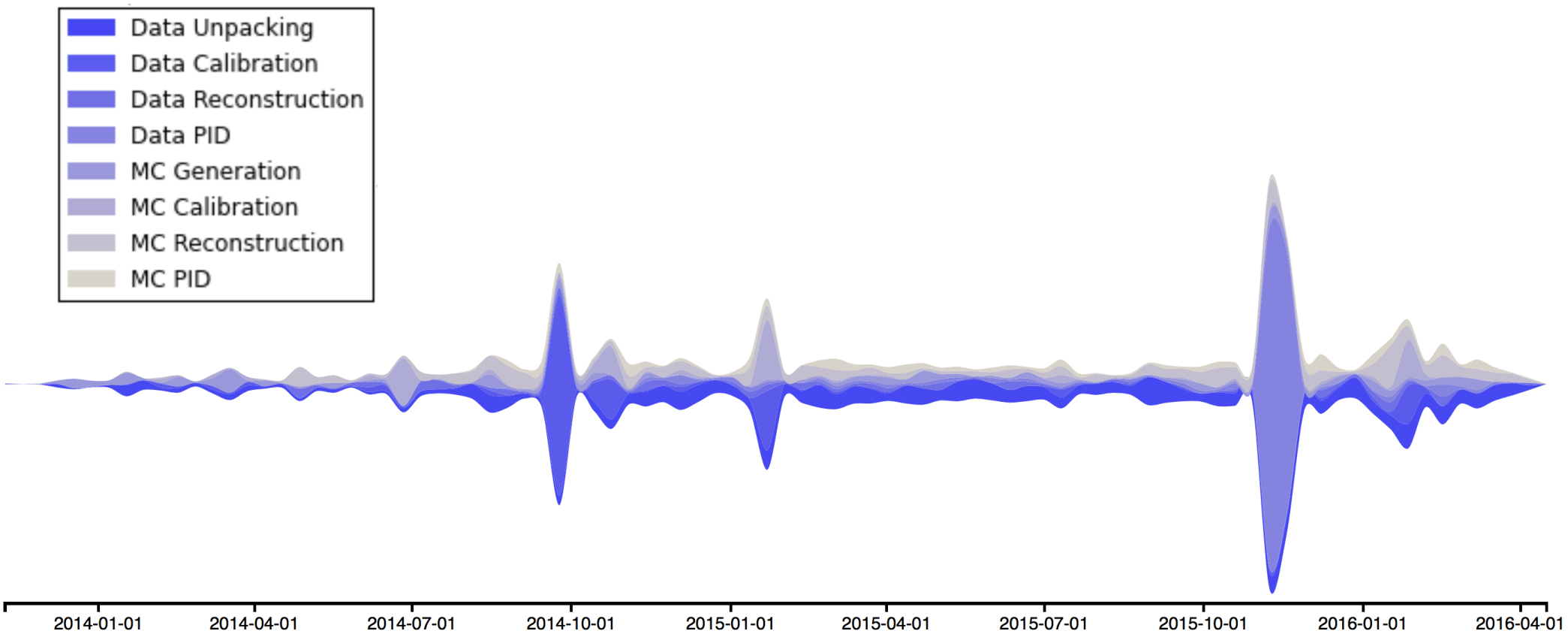
Dataset Lifetime

- Effective operational lifetime of the datasets: time until the dataset is not longer actively analyzed for NOvA, MINERvA and MicroBooNE.



User Request Time for NOvA

- 9536 different dataset requests over 5 years that represents 53 K analysis campaigns for NOvA experiment.



Conclusions

- We have presented the **first results of the SCIDAC-Data project**, examining in detail how experiments at Fermilab are organized, classified and consume petascale datasets.

- **We will release this data soon** to allow the community further analysis and simulations.

- Our first results of a simulation using this data will be also presented in CHEP 2016:

  - Title: "***SciDAC-Data, A Project to Enabling Data Driven Modeling of Exascale Computing***".
  - Date: October 11, 2016 at 11:45 AM.
  - Place: GG C3 (Track4).

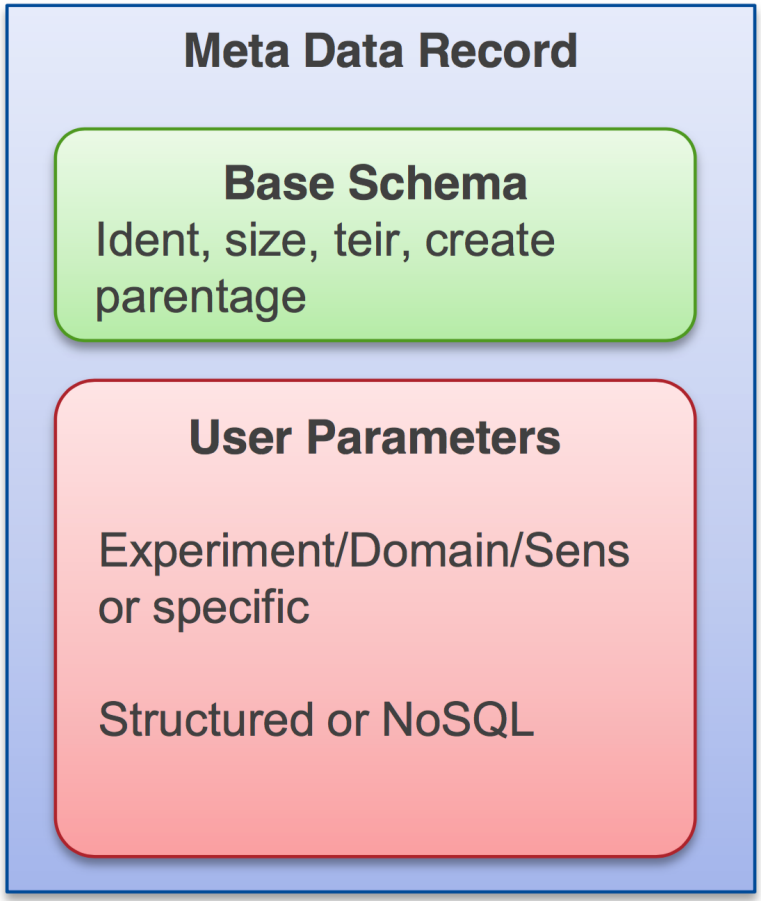
Backup

Dataset Organization and Metadata

- SAM contains a highly extensible schema to allow for classification and organization of the experimental data.
- The metadata falls into two categories:
  A. ***Core/Physical characteristics:*** creation, start, end times of the files, file sizes and format, data integrity checksums, data tier, number of events, etc.
  B. ***Physics and user defined characteristics:*** values ranging from the state/configuration of the accelerator beam to MC generators used in the simulation.
- For instance, some numbers of Metadata physics parameters are:
  - NOvA: 140 parameters.
  - D0: 700 parameters.
  - CDF: 214 parameters.

Data Registering and Catalog

- "Object based" data, replica and project catalog.
- Each data object is registered in the catalog along with metadata describing it.
  - Two components to the metadata.
    - Base schema - General Object Information
      - identifier, size, data tier, begin/end times, parentage/provenance.
    - User parameters - Data content specific fields.
      - detector type, location, trigger stream, etc.
  - Only base schema is required
    - Simplifies registration of foreign/legacy data with catalog systems.
- "Datasets" are then defined via queries against the meta data.
  - Evaluate to the set of objects to retrieve/analyze.



Data Search Classification/Association (SAM)

- Leverage SAM data handling service.
  - Provides full metadata based service for data handling.
- Facilities for
  - Defining arbitrary string-value pairs which can be associated with each file in the system.
  - Storing location information for each file.
  - Searching the database for files which match logical constraints on the metadata.
  - Storing the results of such searches as dataset definitions.
  - Recording the processing history of files accessed via the stored dataset definitions

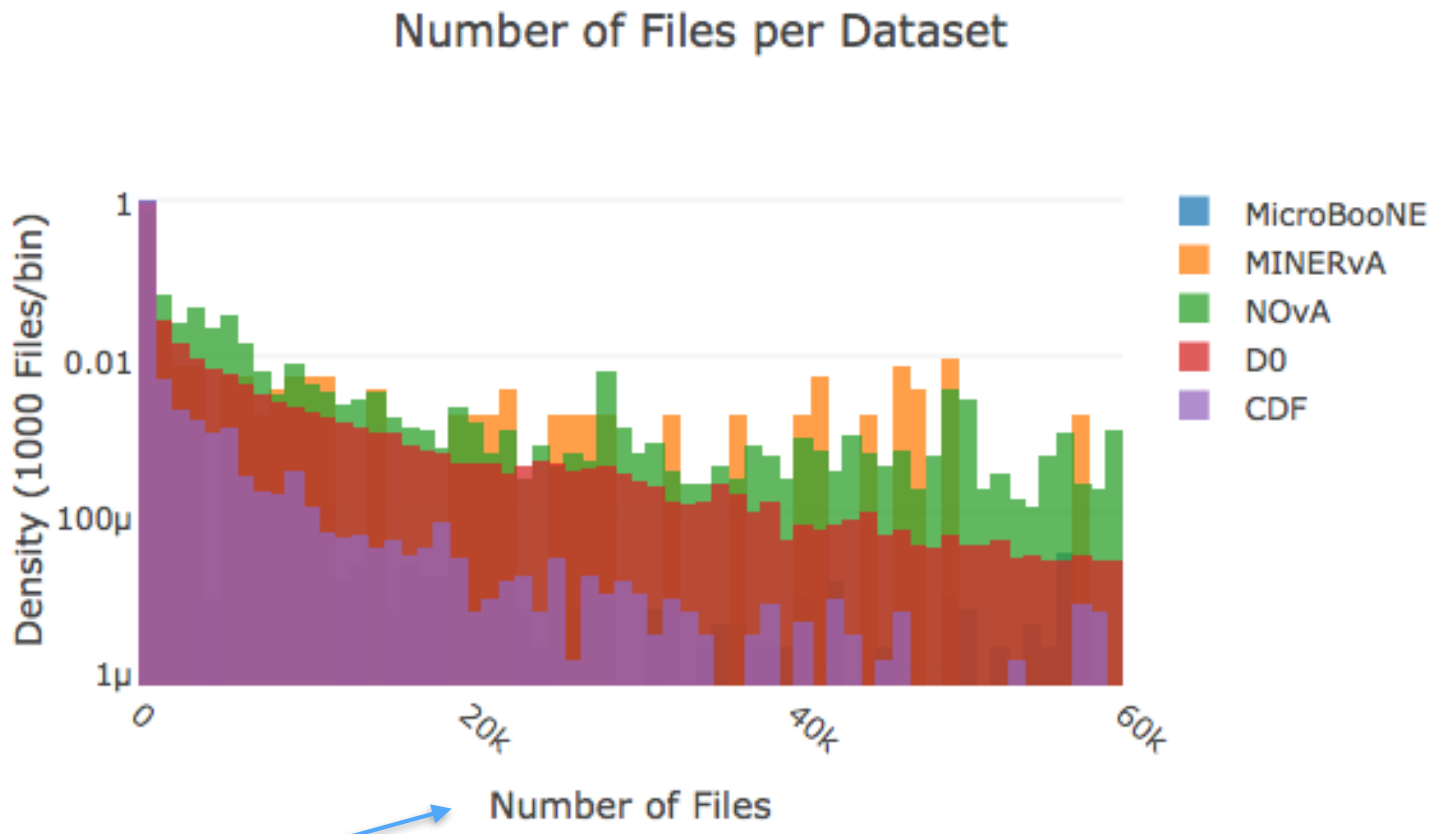
All the associations and structure between objects can be captured and searched w/SAM catalog.

Data Retrieval

- Retrieval is through the IFDH tools set.
  - Interacts with data and replica catalog to find data elements
  - Handles “last miles” of data movement between storage facilities -> “local storage”
    - Can move data between arbitrary elements, e.g. local disks, disk caches such as dCache, or tape libraries.
    - Acts as protocol abstraction layer.
      - Will select a transport protocol suitable for the storage elements.
      - Modular support for protocols: gridftp, srm, dccp, aws S3, cp, dd, etc.
    - Can instigate transfers as local copies, copies to or from remote nodes, or as third party transfers between remote nodes.
    - Understands data storage locations as provided by queries to the SAM metadata system.
    - Incorporates load leveling mechanisms in multi-file transfer to prevent overloading os storage resources.

Data Organization

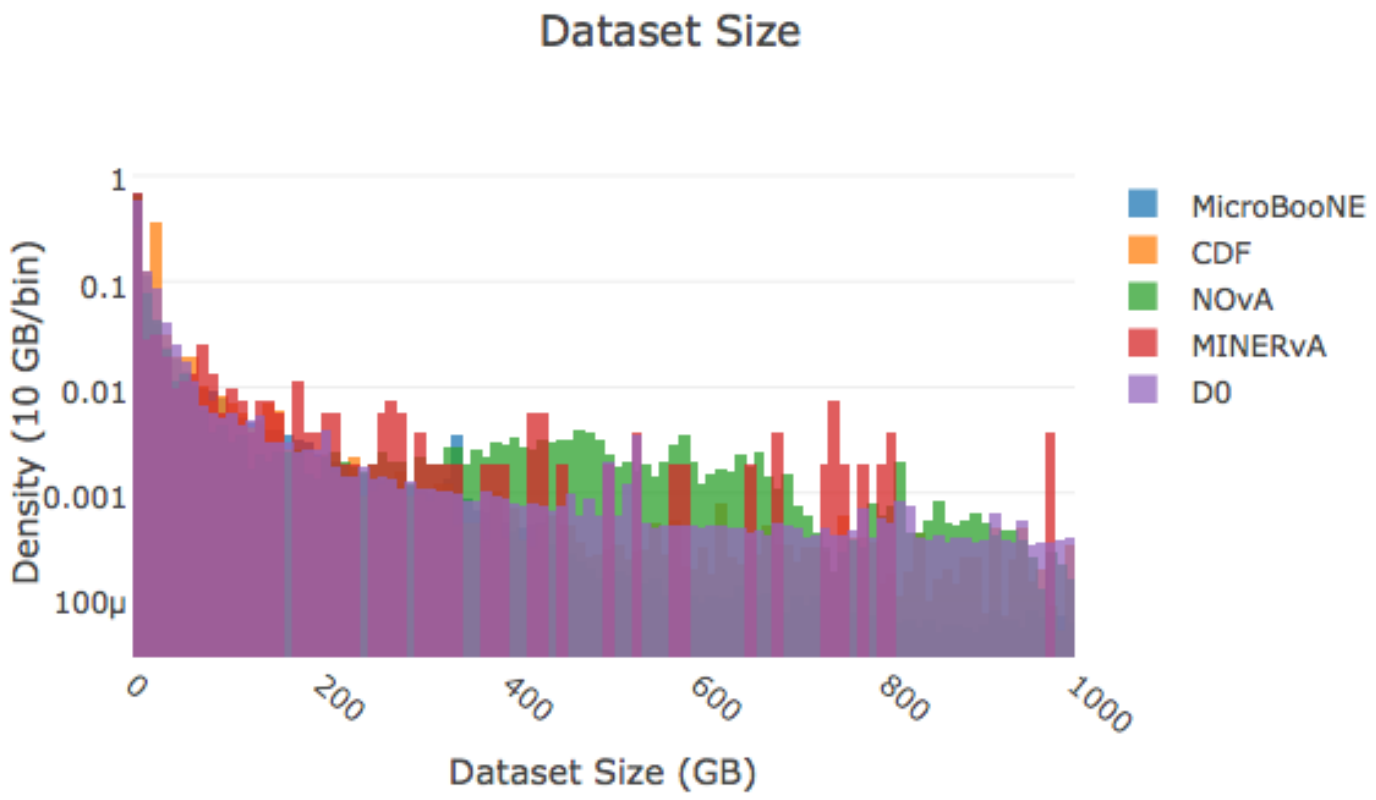
- Distribution representing the data organization within the CDF, D0, NOvA MicroBooNE and MINERvA experiments.



Number of unique files that belong to each dataset that was defined by the experiments

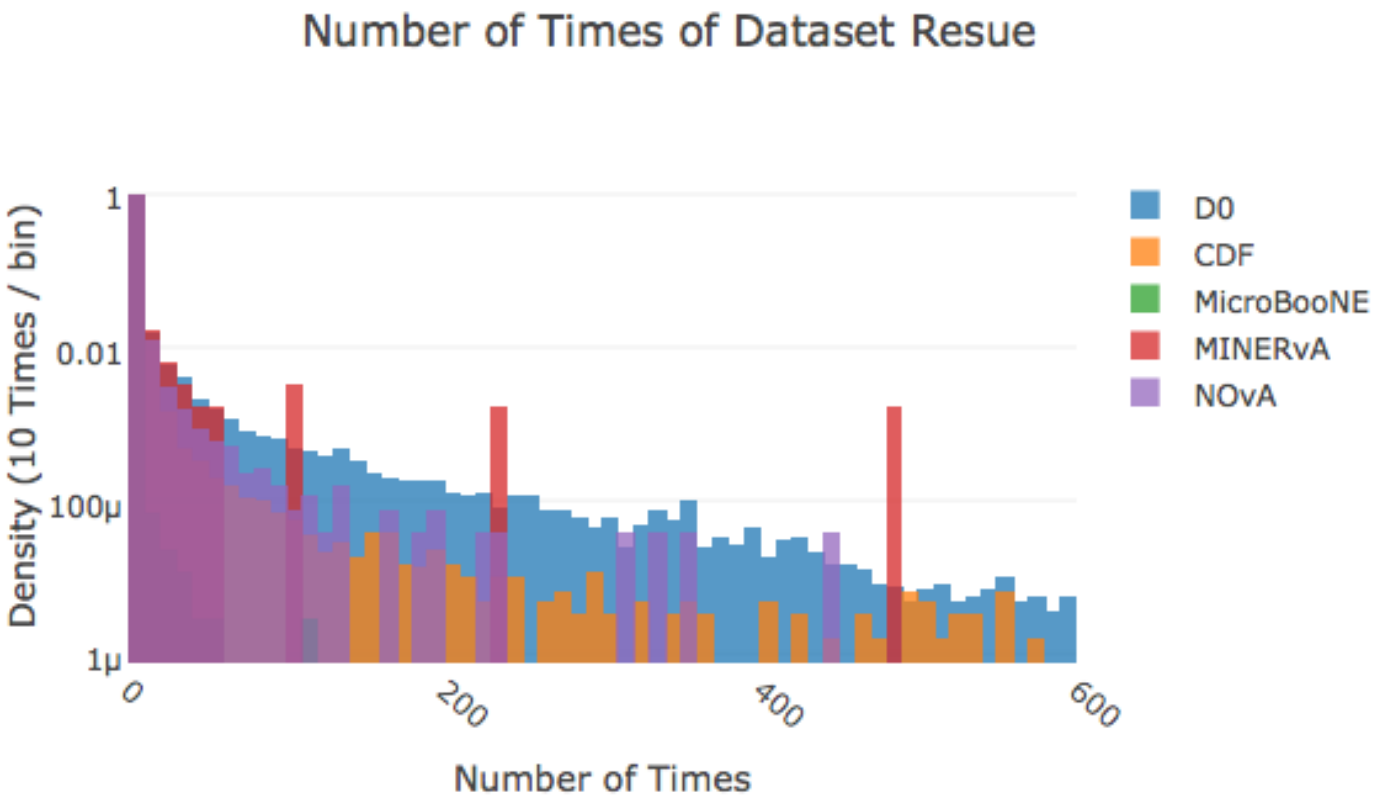
Dataset Size

- Distribution of the total size of datasets defined and used by CDF, D0, NOvA MicroBooNE and MINERvA experiments.



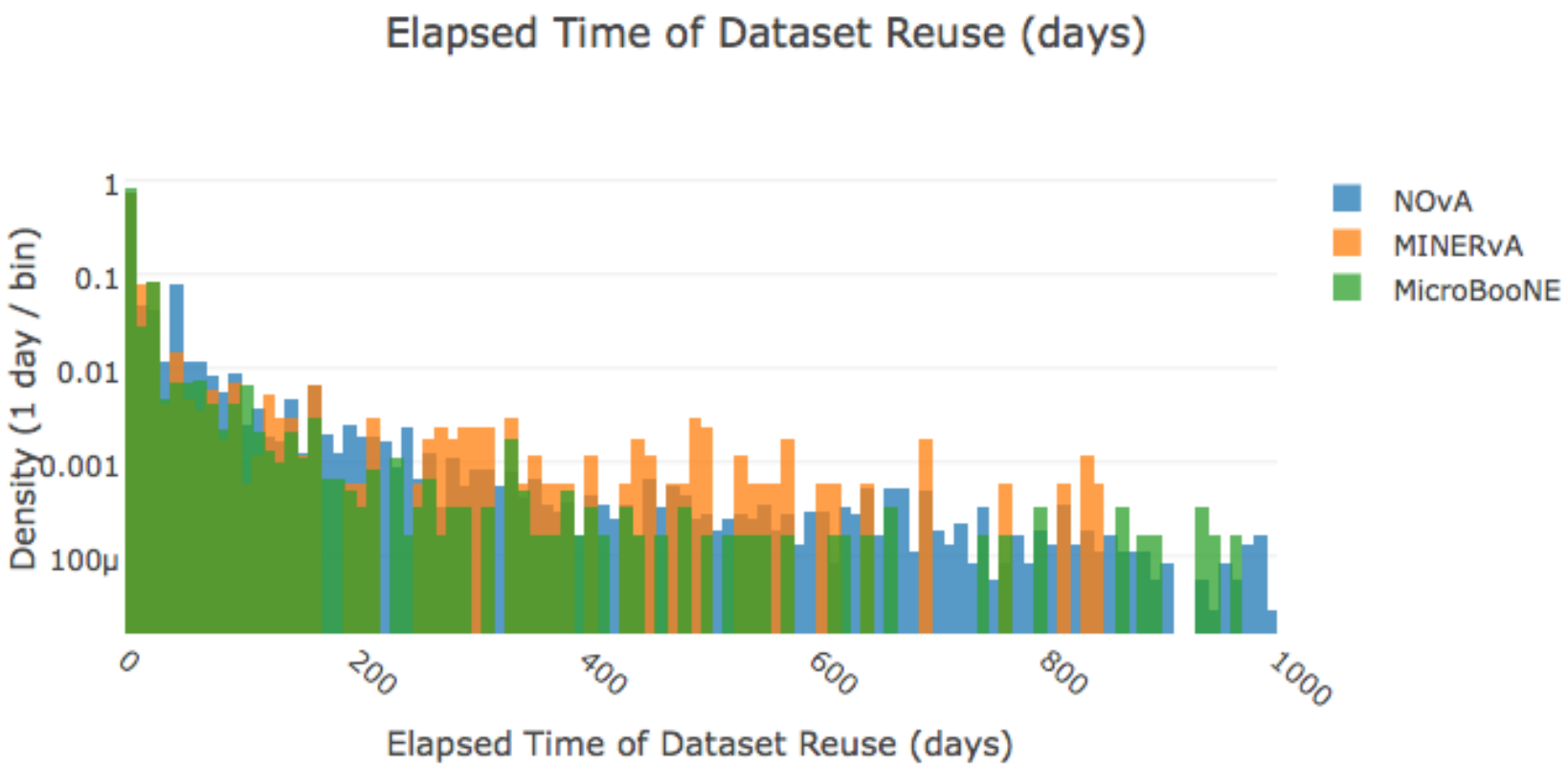
Dataset Usage

- Distribution of the number of times that a given dataset was processed by an analysis application for the NOvA, MicroBooNE, CDF and D0 experiments.



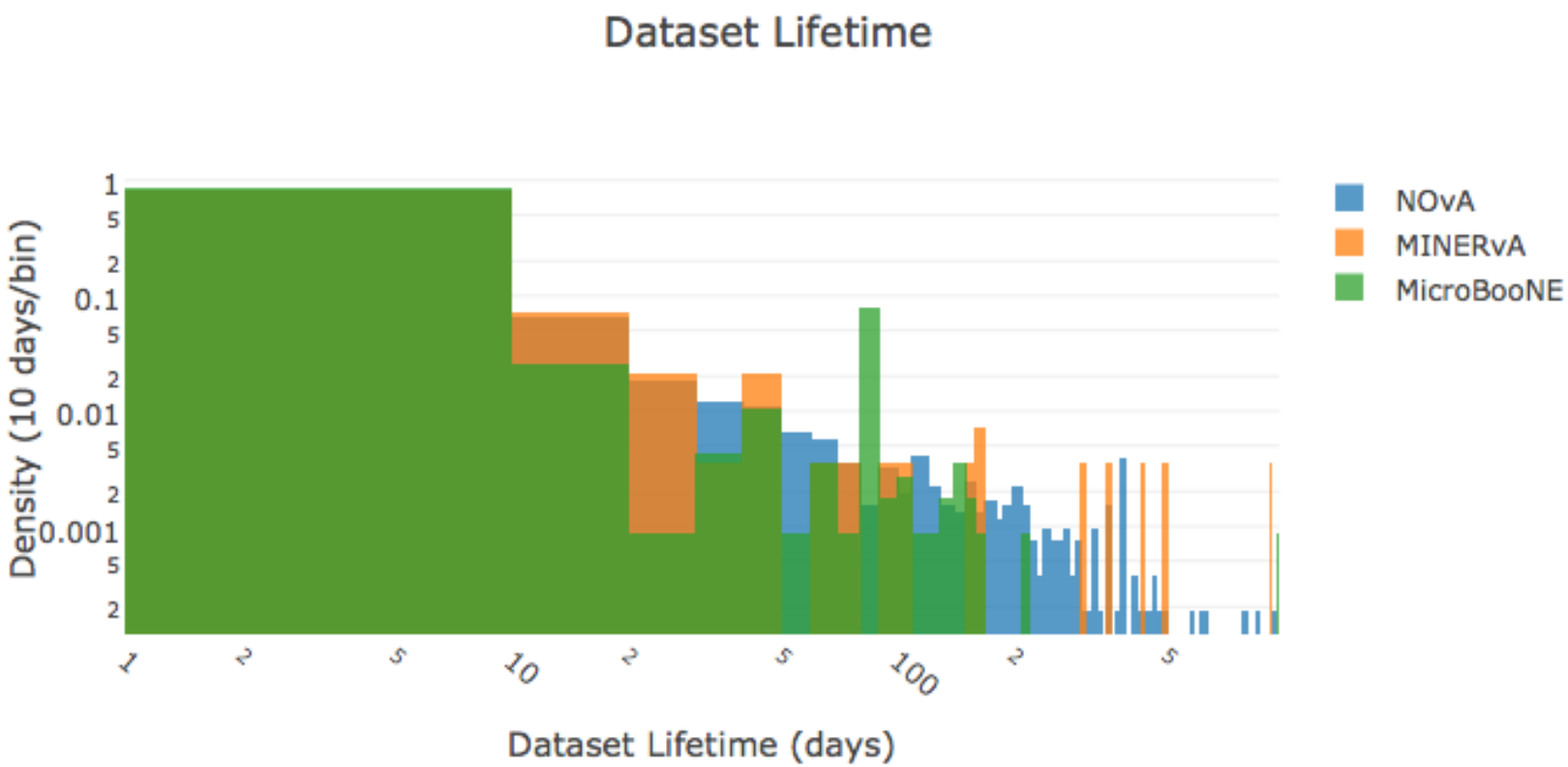
Elapsed Time of a Dataset Usage

- Distribution of the elapsed time between successive analysis projects being started on a dataset for the NOvA, MicroBooNE and MINERvA experiments.

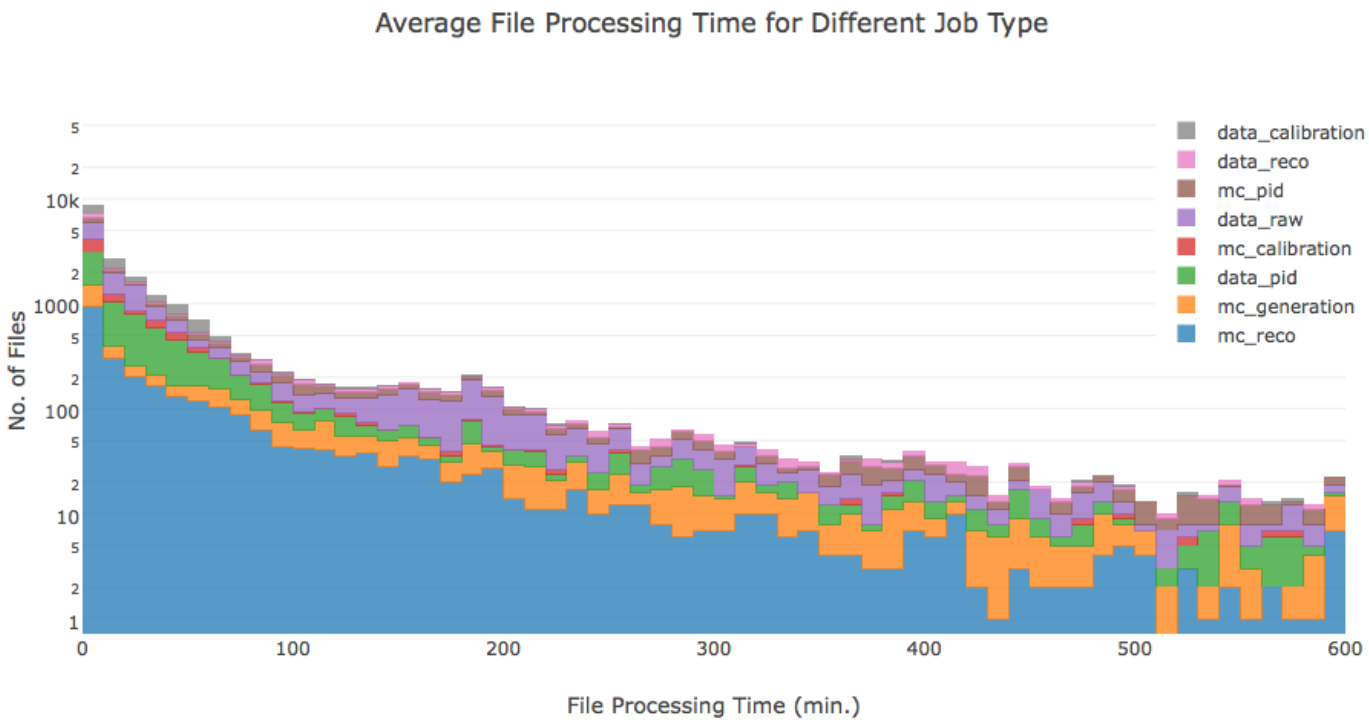


Dataset Lifetime

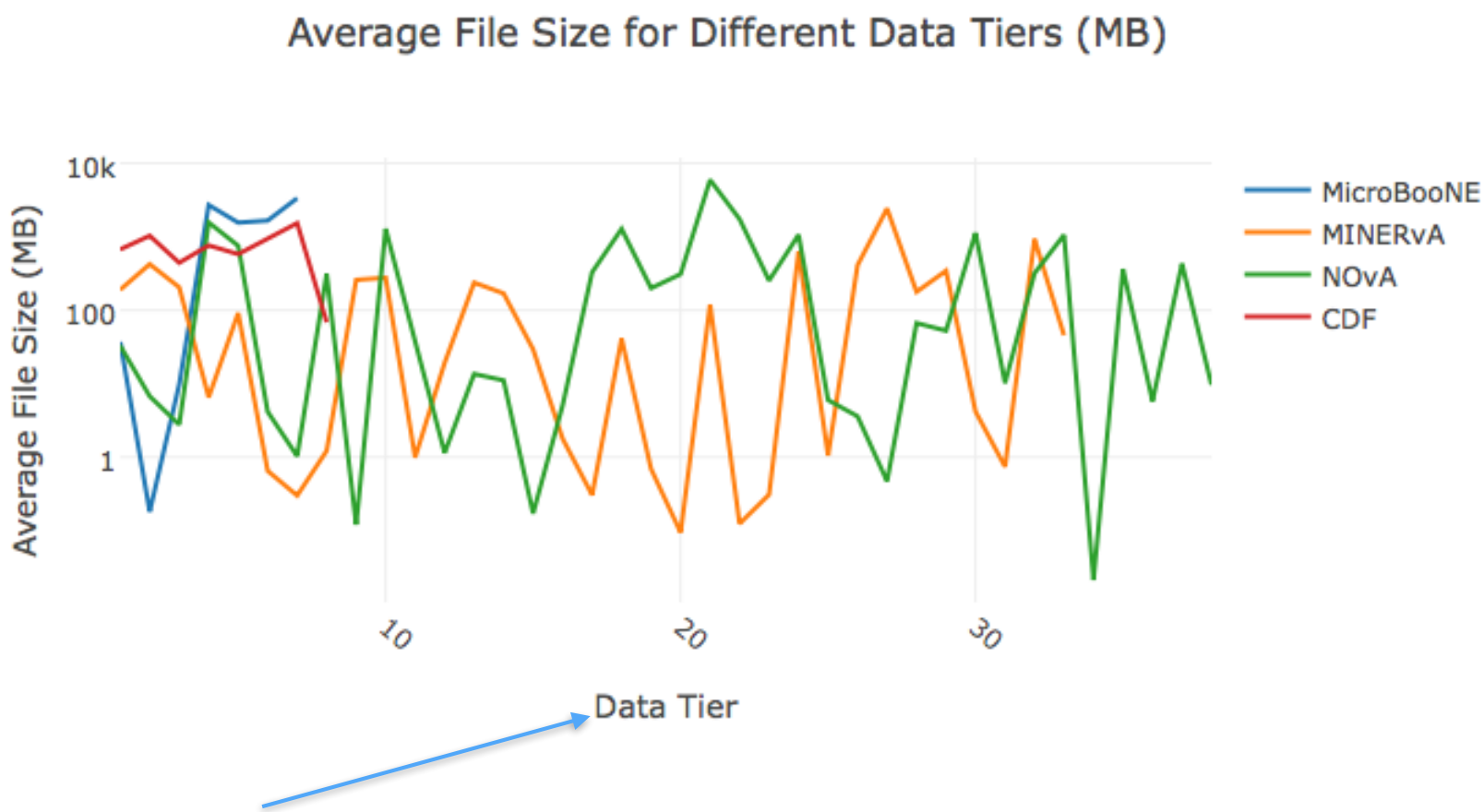
- Effective operational lifetime of the datasets: time until the dataset is not longer actively analyzed.



Average Processing File Time for NOvA



Average File Size Per Data Type



*"**Data tier"** correspond the the data type: raw detector data, Monte Carlo simulation, reconstructed event data) and use for the data within the experiment*