

# Status and Plans for Data Management

Qizhong Li

Fermilab

Nov. 14, 2013

LBNE S&C Meeting

# Outlines

- Metadata for lbne
- Enstore
- SAM
- Put existing MC files into Enstore through SAM
- Plans for next steps

## Metadata for LBNE

- Metadata is a set of parameters that describes data files.
- Metadata serves two purposes:
  - To be used to search data files with required information or parameters
  - To be used by the data catalog system to manage data files
- Designed lbne metadata (Jonathan Insler, Tom Junk and Qizhong Li), so that the existing MC files can be stored onto tapes in ENSTORE, through a data catalog system SAM.
- A document for lbne metadata is in DocDB, **LBNE-doc-8093**
  - Listed lbne metadata parameters
  - Defined parameter values
  - Showed example metadata in JSON format

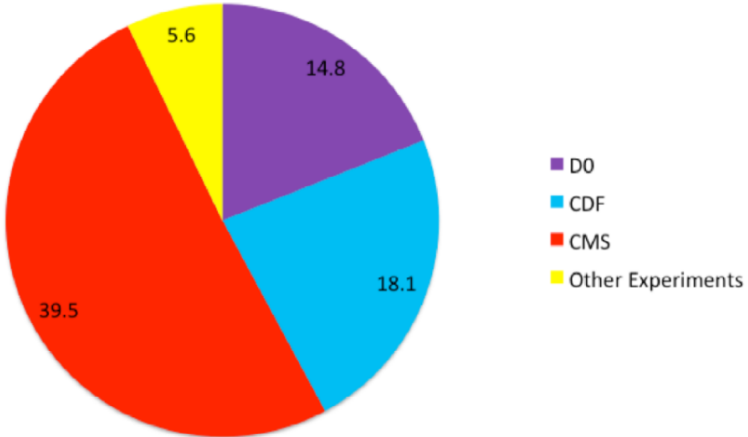
## Enstore Tape Storage

- Enstore is the mass storage system developed and implemented at Fermilab as the primary data store for scientific data sets.
- It provides access to data on tape to/from a user's machine on-site over the local area network, or over the wide area network through the dCache disk caching system.
- It is a very reliable tape storage system with a strong support team
  - on-call 24x7 and
  - the tape libraries and tape drives have 24x7 support from the vendor.

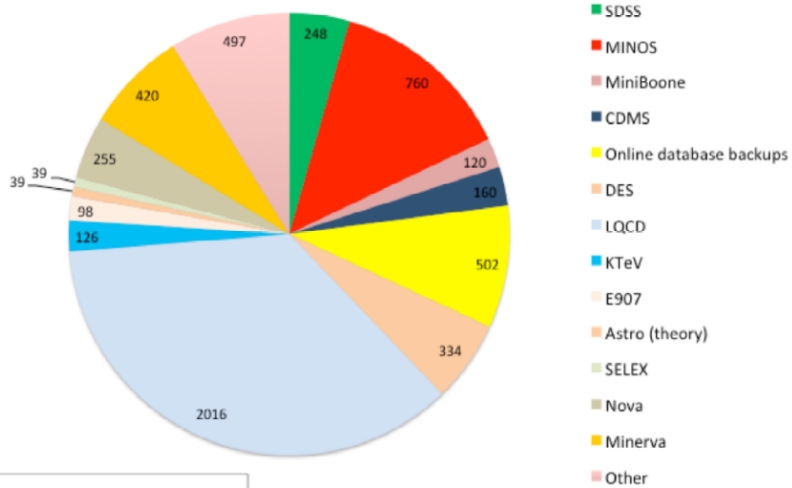


# Enstore Tape Usage

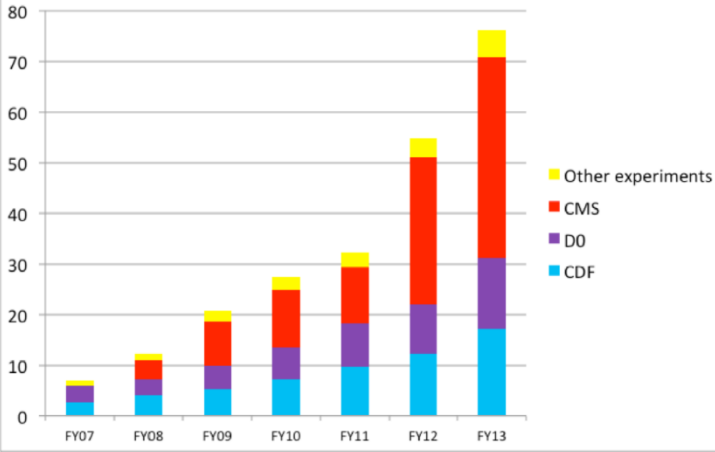
**77.9 Petabytes on Tape 11/1/2013**



**5615 Terabytes on Tape, Other Experiments 11/1/2013**



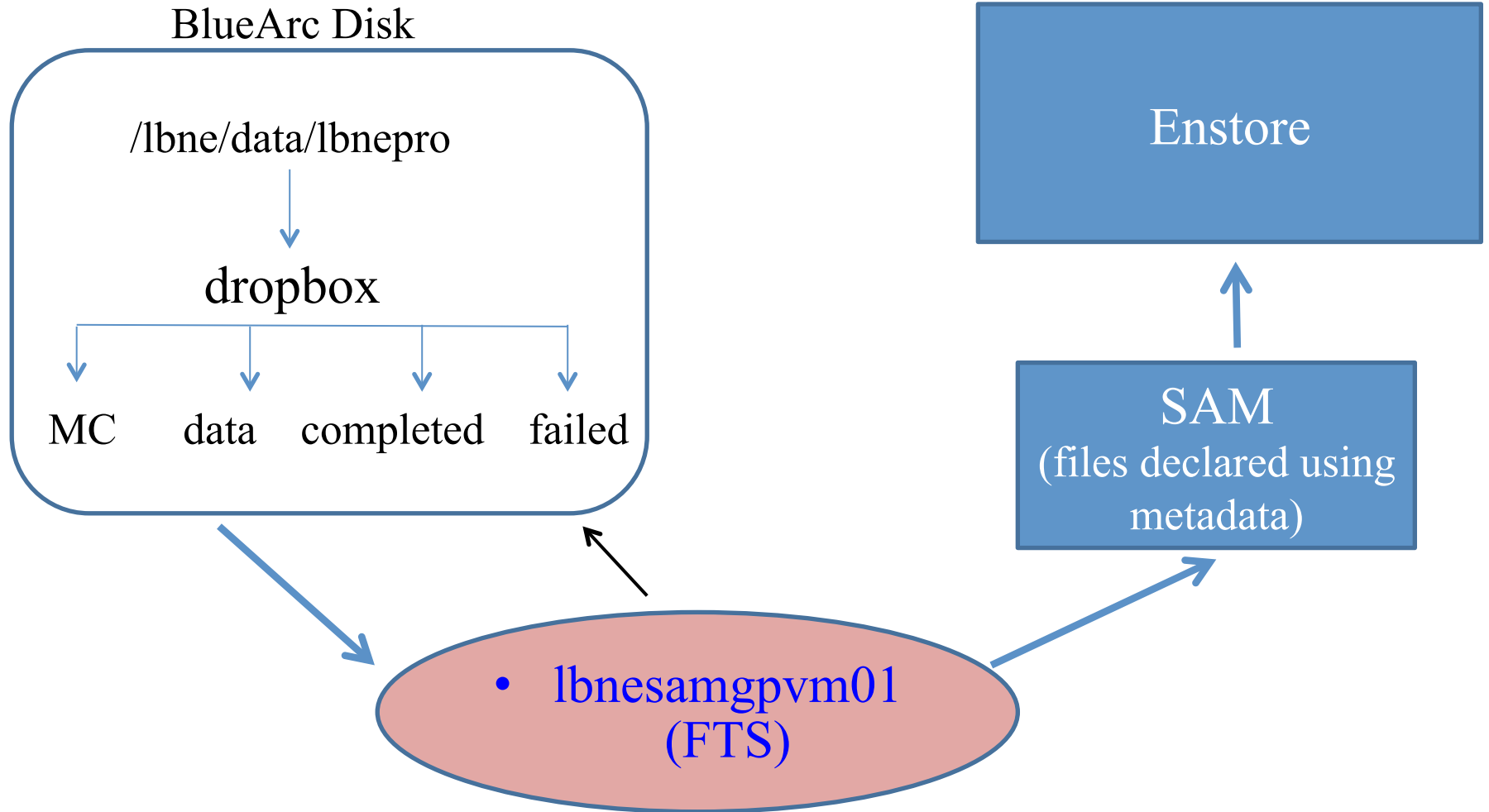
**Petabytes on Tape per Fiscal Year**



# SAM – A Data Handling System

- Files stored into Enstore through SAM.
- We use lbne SAM as our data catalog system to store and retrieve files.
- Users interact with SAM db through SAMWEB server using SAMWEB commands.
- SAM database has 24x7 support.
- To setup SAM for lbne:
  - Created an LBNE library in Enstore
  - Created an lbne SAM database.
  - Defined lbne metadata into SAM.

# File Transfer



## Storing Files

- Created a File Transfer System (FTS) on lbnegpvm01 machine.
- Created a drop box area on /lbne/data.
- The FTS scans the dropbox periodically, checking if any files need to be transferred to tape.
- Currently, for storing the existing MC files:
  - Need to create the metadata in JSON format
  - Declare files to SAM using the metadata
  - Put the files into the drop box
  - Files will automatically be put onto tape.
- All Tom's existing MC file (about 3,000 files) have been stored into Enstore through SAM.
- These files can be retrieved from Enstore.

**Thanks to Enstore team and SAM team for the help!**



# FTS Monitoring

## FTS status for lbnesamgpvm01.fnal.gov

Generated at 2013-11-12 14:43:20 CST ([refresh](#))

### Summary

FTS: OK	FSS: OK	Stager: OK
---------	---------	------------

Completed files:	2736
Failed transfers:	0
All error files:	0
Pending files:	0
New files:	0

### - Recent completed transfers

Time	File name	Destination
<b>mc</b>		
2013-11-12 09:49:20 CST	<a href="#">genieNtp_20130423_lbne_lbneCD1_anuflux_nuebarflux_nuebar_g280_Ar40_10kt36_5000_0_002.root</a>	enstore:/pnfs/lbne/mc/lbne/simulated/001
2013-11-12 01:31:29 CST	<a href="#">genieNtp_20130423_lbne_lbneCD1_anuflux_numubarflux_nuebar_g280_Ar40_10kt36_5000_9_100.root</a>	enstore:/pnfs/lbne/mc/lbne/simulated/001
2013-11-12 01:30:56 CST	<a href="#">genieNtp_20130423_lbne_lbneCD1_anuflux_numubarflux_nuebar_g280_Ar40_10kt36_5000_9_099.root</a>	enstore:/pnfs/lbne/mc/lbne/simulated/001
2013-11-12 01:30:53 CST	<a href="#">genieNtp_20130423_lbne_lbneCD1_anuflux_numubarflux_nuebar_g280_Ar40_10kt36_5000_9_098.root</a>	enstore:/pnfs/lbne/mc/lbne/simulated/001
2013-11-12 01:30:43 CST	<a href="#">genieNtp_20130423_lbne_lbneCD1_anuflux_numubarflux_nuebar_g280_Ar40_10kt36_5000_9_097.root</a>	enstore:/pnfs/lbne/mc/lbne/simulated/001
2013-11-12 01:30:39 CST	<a href="#">genieNtp_20130423_lbne_lbneCD1_anuflux_numubarflux_nuebar_g280_Ar40_10kt36_5000_9_096.root</a>	enstore:/pnfs/lbne/mc/lbne/simulated/001
2013-11-12 01:30:35 CST	<a href="#">genieNtp_20130423_lbne_lbneCD1_anuflux_numubarflux_nuebar_g280_Ar40_10kt36_5000_9_095.root</a>	enstore:/pnfs/lbne/mc/lbne/simulated/001
2013-11-12 01:30:31 CST	<a href="#">genieNtp_20130423_lbne_lbneCD1_anuflux_numubarflux_nuebar_g280_Ar40_10kt36_5000_9_094.root</a>	enstore:/pnfs/lbne/mc/lbne/simulated/001
2013-11-12 01:30:27 CST	<a href="#">genieNtp_20130423_lbne_lbneCD1_anuflux_numubarflux_nuebar_g280_Ar40_10kt36_5000_9_093.root</a>	enstore:/pnfs/lbne/mc/lbne/simulated/001
2013-11-12 01:30:25 CST	<a href="#">genieNtp_20130423_lbne_lbneCD1_anuflux_numubarflux_nuebar_g280_Ar40_10kt36_5000_9_092.root</a>	enstore:/pnfs/lbne/mc/lbne/simulated/001

+ Failed transfers (0 hidden)

+ All errors (0 hidden)

+ Pending (0 hidden)

+ New (0 hidden)

+ Configuration

## Plans

- To generate metadata automatically when a new MC file is produced
  - Art and larsoft have ability to create metadata for SAM
  - Need to make this work for lbne.
- Define datasets using metadata queries.
- Then, to be able to interface art and larsoft programs to read input files from SAM.
- Maxim plans to also store files at BNL.
- Craig plans to also store files at LBL.