Study of xrootd transaction

Introduction. The SAM data access system was designed to provide granular information about file access as well as static metadata about file characteristics. The DUNE data management group is designing a MetaCat catalog and Data Dispatcher which will decouple these two role.

This note describes a system for studying file access using the xrootd streaming file access for DUNE compute jobs run on the grid and at Fermilab between Jan 1, 2021 and June 28, 2021.

Individual transfer action communicated to the sam system are logged in ElasticSearch and accessible through https queries such as this example.

https://fifemon-es.fnal.gov/sam-events/2021.04/\_search?q=experiment:dune%%20and%%20project\_id:%s&size=10000"%( projectID)

Elastic search can only return 10000 entries per query so queries are made by first selecting individual sam projects and then filtering on projectID. Sam “events” that are directly relevant for file transfers are then recorded in a json file for each project. Those events are then grouped by the name of the file transferred and sorted in time to determine the start and end of the transfer.

For this sample a total of 1305010 file transactions were studied. Of these 1031606 had successful final state “consumed”, 52106 had final state “skipped” which indicate failure after file processing has started. The remaining transfers were of extremely short duration and had state “transferred” or “delivered” which, in streaming mode, means that sam successfully communicated the file location to the process.

The following two examples show the information recorded about file prccessing based on the ElasticSearch information (durations, locations) and general information (application) about the project.

This is an example file transfer. User “drivera” ran a job at site “pp.rl.ac.uk” in the UK which took 49,013 seconds to process a 4.7 GB file from the PDSPProd4 production through application neutronana. The data were streamed from disk “fndca1.fnal.gov". The final state of the file was “consumed” which means that the process reported success in processing the file. The average transfer rate was 0.097 MB/sec which indicates that the job was completely CPU bound.

{"disk": "fndca1.fnal.gov",

"user": "drivera",

"date": "2021-06-27",

"process\_id": 16059492,

"timestamp": "2021-06-27T23:57:58.047Z",

"duration": 49013.56700015068,

"file\_size": 4767154720,

"username": "drivera",

"application": "neutronana",

"version": "v09\_16\_01",

"final\_state": "consumed",

"site": "uk\_pp.rl.ac.uk",

"rate": 0.09726194218807507,

"project\_name": "drivera\_protodune-sp\_runset\_5842\_reco\_v09\_09\_01\_v0\_20210627181334",

"file\_name": "np04\_raw\_run005842\_0013\_dl3\_reco1\_14895769\_0\_20201208T225647Z.root",

"data\_tier": "full-reconstructed",

"node": "heplnc123.pp.rl.ac.uk",

"country": "uk",

"campaign": "PDSPProd4"}

This is an example file transfer. User “spurgeon” ran a job at site “usatlas.bnl.gov” in the US which took 226 seconds to process a 4.1 GB raw data file through application twocrtmatching. The data were streamed from disk “fndca1.fnal.gov". The final state of the file was “consumed” which means that the process reported success in processing the file. The average transfer rate was 18 MB/sec which indicates that the job was completely IO bound.

{

"disk": "fndca1.fnal.gov",

"user": "spurgeon",

"date": "2021-01-01",

"process\_id": 14503418,

"timestamp": "2021-01-01T00:46:54.253Z",

"duration": 225.9170000553131,

"file\_size": 4101329211,

"username": "spurgeon",

"application": "twocrtmatching",

"version": "v08\_40\_00",

"final\_state": "consumed",

"site": "us\_usatlas.bnl.gov",

"rate": 18.15414160951074,

"project\_name": "spurgeon\_stable\_xe\_small\_20201231183416",

"file\_name": "np04\_raw\_run011184\_0006\_dl6.root",

"data\_tier": "raw",

"node": "acas1091.usatlas.bnl.gov",

"country": "us",

"campaign": null

}

The following table summarizes the average transfer rate for different applications measured in this data sample. Transfers internal to FNAL are summarized on the left while transfers to sites outside the US are summarized on the right. The ratio of external to internal is in the last column. Applications with no job runs outside the US are not shown.

