Fermilab

Managed by Fermi Research Alliance, LLC for the U.S. Department of Energy Office of Science

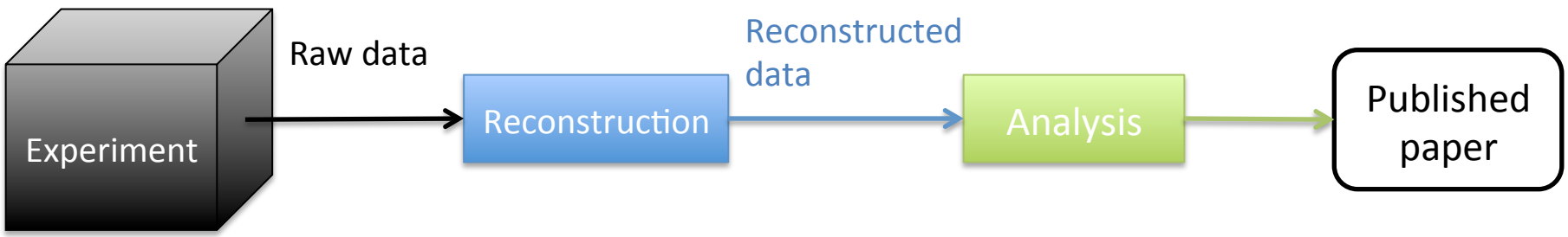
Data management and SAM for Simons Foundation

Robert Illingworth

7 October 2015

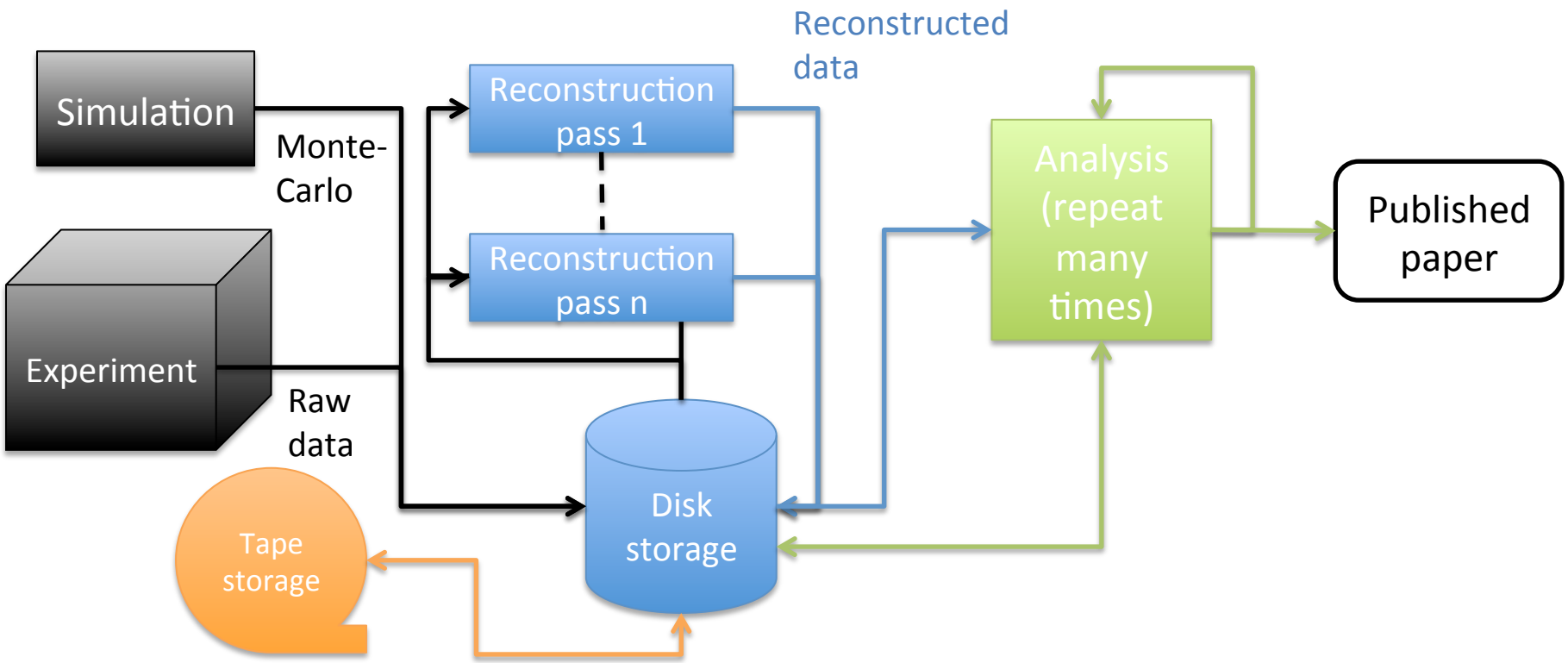
HEP Analysis dataflow (user view)

- Take raw data
- Reconstruct it
- Analyze reconstructed data
- Publish



HEP Analysis dataflow (more realistic view)

Not usually a nice linear progression
Even a small HEP experiment has millions of data files to keep track of and feed the correct ones into the next stage of processing



Why is data movement difficult?

Your files don't reside where your jobs run

Your files reside here

Tape Robot

Disk Storage

Not here, where you need them

Compute Farm of worker nodes

Files must be moved to where a worker node can access them with efficiency and scalability

Output files must be returned from worker node

NOT a simple process, especially the efficiency and scalability part!

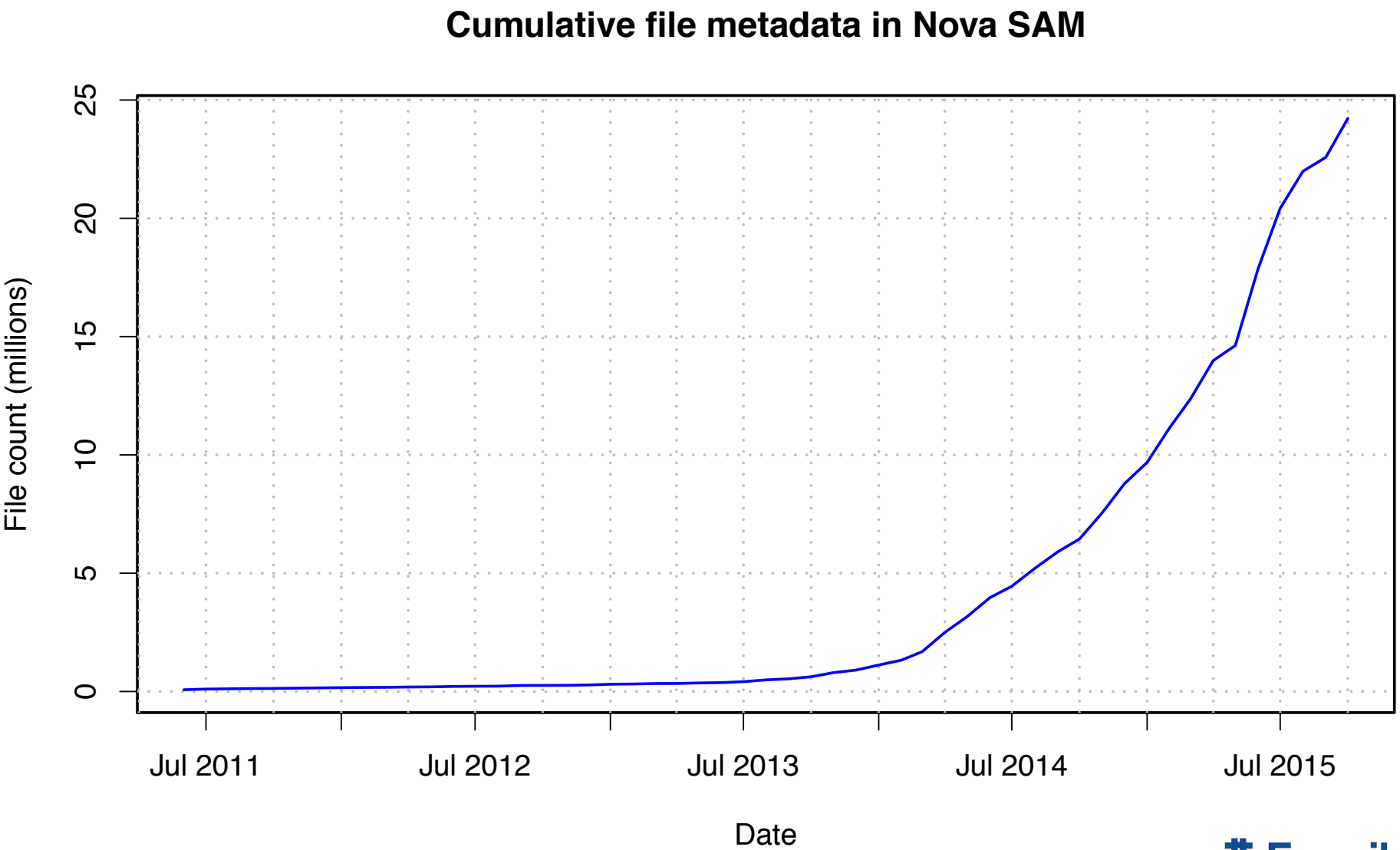
Background

- High energy physics experiments produce a lot of data
- Traditionally experiments with enough data to require it ended up writing their own management system for keeping track of it
- However, now even a “small” HEP experiment can produce petabyte scale datasets
  - Writing their own is not a option here
- As Fermilab has multiple experiments requiring data management services we needed a common solution suitable for as many as possible
  - Some of the data is unusual for modern particle physics; for example the COUPP dark matter search produces what are essentially movies

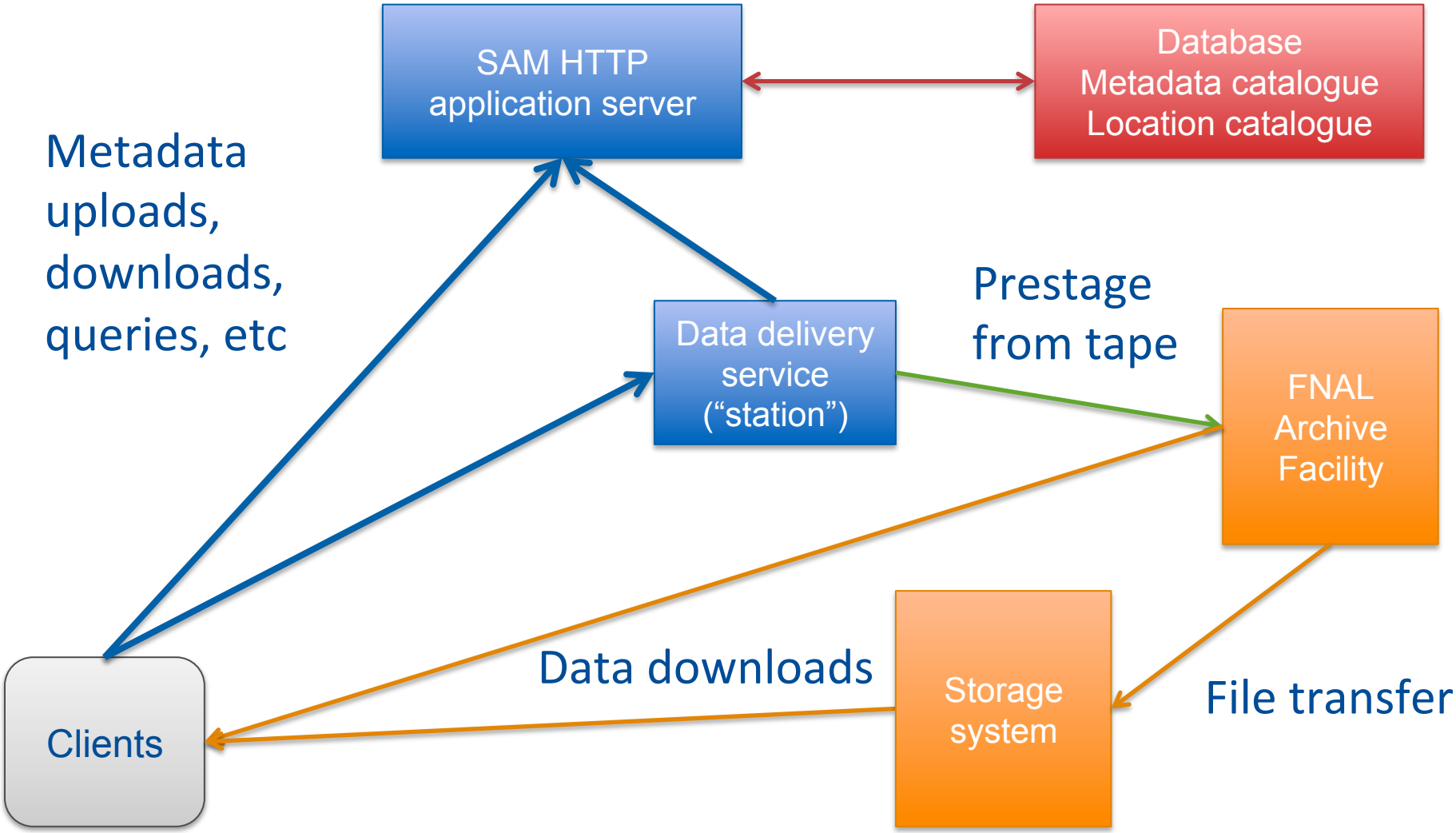
SAM data management as a common solution

- The common solution is based on the SAM data management service
- Originally developed in the late 90's for Fermilab collider experiments; used to track ~250 million files over 10 year running period
- Recently updated to be more flexible and to make use of more modern interfaces (http)
- Added another ~50 million files over a dozen instances with the updated system (and increasing)
- Aim to minimize human effort required to operate the system

Metadata in NOvA SAM instance



Basic architecture



SAM components

- Metadata catalogue
  - What is in the files?
- Location catalogue
  - Where are the files?
- Data delivery service
  - Get the files to the processing tasks, and keep track of which ones were used by whom to do what

Metadata catalogue

- Stores data about the data
- Tracks both physical information about the files (name, size, checksums, …), and scientific data
- Some parameters are predefined
- Each instance can independently define attributes for key/value pairs

Example metadata

```
$ samweb -e nova get-metadata fardet_r00020577_s51_t02_S15-03-11_v1_data.artdaq.root
                File Name: fardet_r00020577_s51_t02_S15-03-11_v1_data.artdaq.root
                  File Id: 166148526
              Create Date: 2015-10-05T20:43:38+00:00
                     User: novapro
                File Type: importedDetector
              File Format: artroot
                File Size: 1605524033
                 Checksum: enstore:3193513656
           Content Status: good
                    Group: nova
                Data Tier: artdaq
             Application: nova rawconvert s15-03-11
             Event Count: 2115
                      … : …
            DQ.IsGoodRun: true
          NOVA.DetectorID: fd
          Online.Detector: fardet
                     Runs: 20577.0051 (physics)
                  Parents: fardet_r00020577_s51_t02.raw
```

Can store one or more checksums per file

Experiment defined parameters

Query metadata catalogue

- Provides a flexible mini language for querying metadata information

```
$ samweb -e nova list-files "run_number 20577.0051 and
data_tier artdaq and nova.detectorid = fd with limit 3"
fardet_r00020577_s51_DDContained_S15-03-11_v1_data.artdaq.root
fardet_r00020577_s51_DDnumu_S15-03-11_v1_data.artdaq.root
fardet_r00020577_s51_DDenergy_S15-03-11_v1_data.artdaq.root
```

- End users do not need to understand the database structure or anything about SQL to query the data

Dataset creation

- Dataset definitions are essentially stored metadata queries

```
$ samweb -e nova create-definition illingwo_test_20151006_1
"run_number 20577.0051 and data_tier artdaq and nova.detectorid = fd
with limit 3"
Dataset definition 'illingwo_test_20151006_1' has been created with
id 242531
```

- The definition can be used later to find and process the associated files

```
$ samweb -e nova count-definition-files illingwo_test_20151006_1
3
```

Replica catalogue

- Keeps track of where the files are

```
$ samweb -e nova locate-file fardet_r00020577_s51_t02_S15-03-11_v1_data.artdaq.root
enstore:/pnfs/nova/production/raw2root/S15-03-11/fardet/000205/20577/02(4751@vpo454)
```

- Can have multiple locations per file

- Also knows how to turn a location into an accessible URL

```
$ samweb -e nova get-file-access-url --schema=gsiftp
fardet_r00020577_s51_t02_S15-03-11_v1_data.artdaq.root
gsiftp://fndca1.fnal.gov:2811/pnfs/fnal.gov/usr/nova/production/raw2root/S15-03-11/
fardet/000205/20577/02/fardet_r00020577_s51_t02_S15-03-11_v1_data.artdaq.root
```

- Can support multiple access methods for the data this way

Data delivery service

- Ask for all the files in a dataset to be given to you
- Uses a combination of where the files are located and the arguments you pass in to pick the best source locations
- Also able to initiate prestaging from tape and (potentially) copy files around before giving them out
- The end user doesn't have to worry about where the data is to get at it
- All the activity is stored in the database so you can go back later and discovered which files were processed

Example

- Example that simply prints the access urls

```
$ samweb -e nova  run-project --defname=illingwo_test_20151006_1
Started project illingwo_test_20151006_1_20151006192649
Started consumer processs ID 14123779
gsiftp://fndca1.fnal.gov:2812/pnfs/fnal.gov/usr/nova/production/raw2root/S15-03-11/fardet/000205/20577/
all/fardet_r00020577_s51_DDContained_S15-03-11_v1_data.artdaq.root
gsiftp://fndca1.fnal.gov:2811/pnfs/fnal.gov/usr/nova/production/raw2root/S15-03-11/fardet/000205/20577/
all/fardet_r00020577_s51_DDnumu_S15-03-11_v1_data.artdaq.root
gsiftp://fndca1.fnal.gov:2811/pnfs/fnal.gov/usr/nova/production/raw2root/S15-03-11/fardet/000205/20577/
all/fardet_r00020577_s51_DDenergy_S15-03-11_v1_data.artdaq.root
Stopped project illingwo_test_20151006_1_20151006192649
```

- Monitor progress here: http://samweb.fnal.gov:8480/station_monitor/nova/stations/nova/projects/illingwo_test_20151006_1_20151006192649
- Advanced usage: can have multiple consumers pulling files from the same dataset simultaneously
  - Hand out files according to the rate they are being processed rather than splitting it up in advance

Modifying SAM to meet your needs

- Some aspects seem a good fit to your requirements
  - Metadata catalogue with datasets defined in terms of metadata
- Some aspects may need improvements
- One limitation is that currently file names must be unique (there can only be one active file with a given name, although there may be older, inactive versions)
  - If necessary this restriction could be relaxed, although there must be some sort of unique identifier, such as the trailing path segment
- Definable metadata is limited to key/value pairs
  - If you want more structured metadata it involves more work, although the structure is flexible enough to allow this

Modifying SAM to meet your needs (cont)

- Authentication is based on certificates
- Authorization is relatively limited in scope
  - Usually any member of an experiment can read everything from that experiment
- Client support
  - We currently only support Linux (and unofficially the client works on OS X)
  - Is this sufficient?
- The web browser interfaces are not used much
  - Would they need to be improved?
- Other requirements?