Data movement in ATLAS

Some notes from Hironori Ito

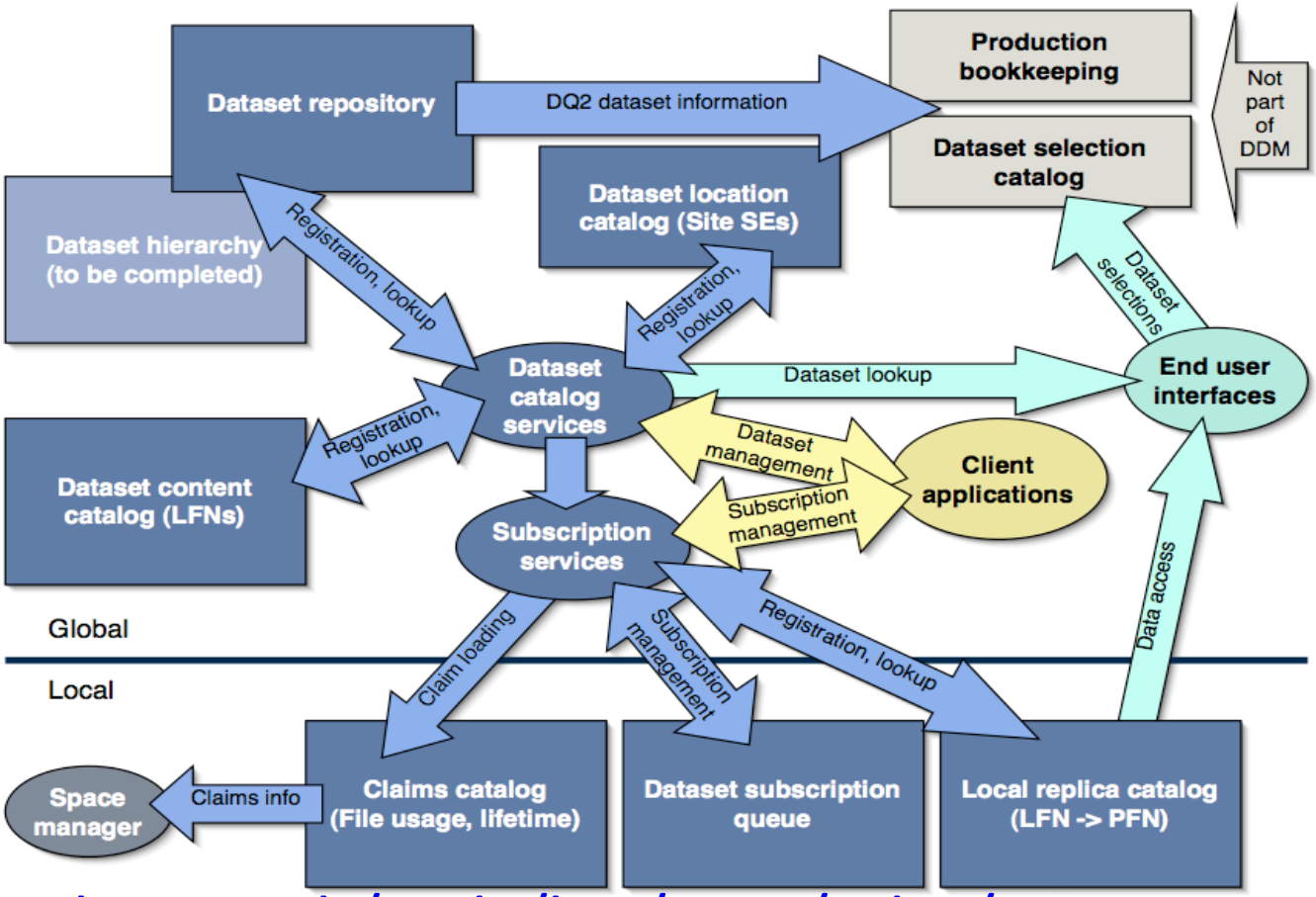
Arranged by Marco Mambelli

Data movement

- It is moved by DQ2.
- The actual transfers between two SRM are controlled by FTS.
- Besides DQ2, some data (very small fraction) are moved by panda mover (for any missing inputs for PANDA jobs)
- Also, users use dq2-get with/without FTS to move files to the local storage

DDM

- DQ2
- FTS
- LFC



- https://twiki.cern.ch/twiki/bin/view/Atlas/DistributedDataManagement

Transfer rate

- Some day BNL, ATLAS Tier1, can push at 20Gb/s to all T2s. While other times, it might be doing 100MB/s.
- Looking at last 30 days, the average rate is of about 500MB/s from BNL to US T2 and 200MB/s from US T2s to BNL.

Strengths and weaknesses

- Achieving high throughput rate is very easy
- The priority and fair share is not quite up-to what we want
- Management of files at many sites is not as efficient: deletion, consistency, etc...

Adoption by other users in OSG

- DQ2 is ATLAS specific. It requires the concept of dataset or at least some grouping of files. And, it requires some central services.
- FTS can be used in any VOs. In fact, it is used by other VOs in non-OSG grid.
- The client command, dq2-get with FTS, is the light-weight version of DQ2 without the over head of DQ2 site service. Similar thing can be done for general purpose without the concept of dataset.
- FTS works with SRM and GridFTP end points. US ATLAS T3s are expected to use this method of getting files/ datasets from T1/T2s.