Data requirements for MonteCarlo production by D0

A. Baranovski

Goal of this presentation

- Discuss D0 data requirements
- Advocate for support of D0 use of local srm enabled storage elements
- Share operational experience

Processing model

- Process starts and requests runtime execution environment.
  - Fairly large set of files totaling 800Mb
- Request is processed by D0 data handing to srmcp that data to “local” storage
  - “local” defined by workflow conf. (not ness. local to site)
- When files arrive - srmget and globus-url-copy are issued to stage data to local scratch

Processing model

- Stage “Zero bias” data- integral part of the detector simulation requirements
  - 2 ZeroB files are selected per job
  - Each file is 200Mb
  - Selection probability is 1/2800
    - ZeroB dataset is 560Gb

Input data requirements

- 800Mb + 400Mb (ZeroB) = 1200Mb of input data is required to start generation phase
- Processing time is 5 hours
- Timeout is 3600 secs < proc time

- Job must have access to at least 0.33 Mbytes/sec of throughput capacity
- Each job generates 10 requests to storage system (5 srmgets/ 5 gftp copies)

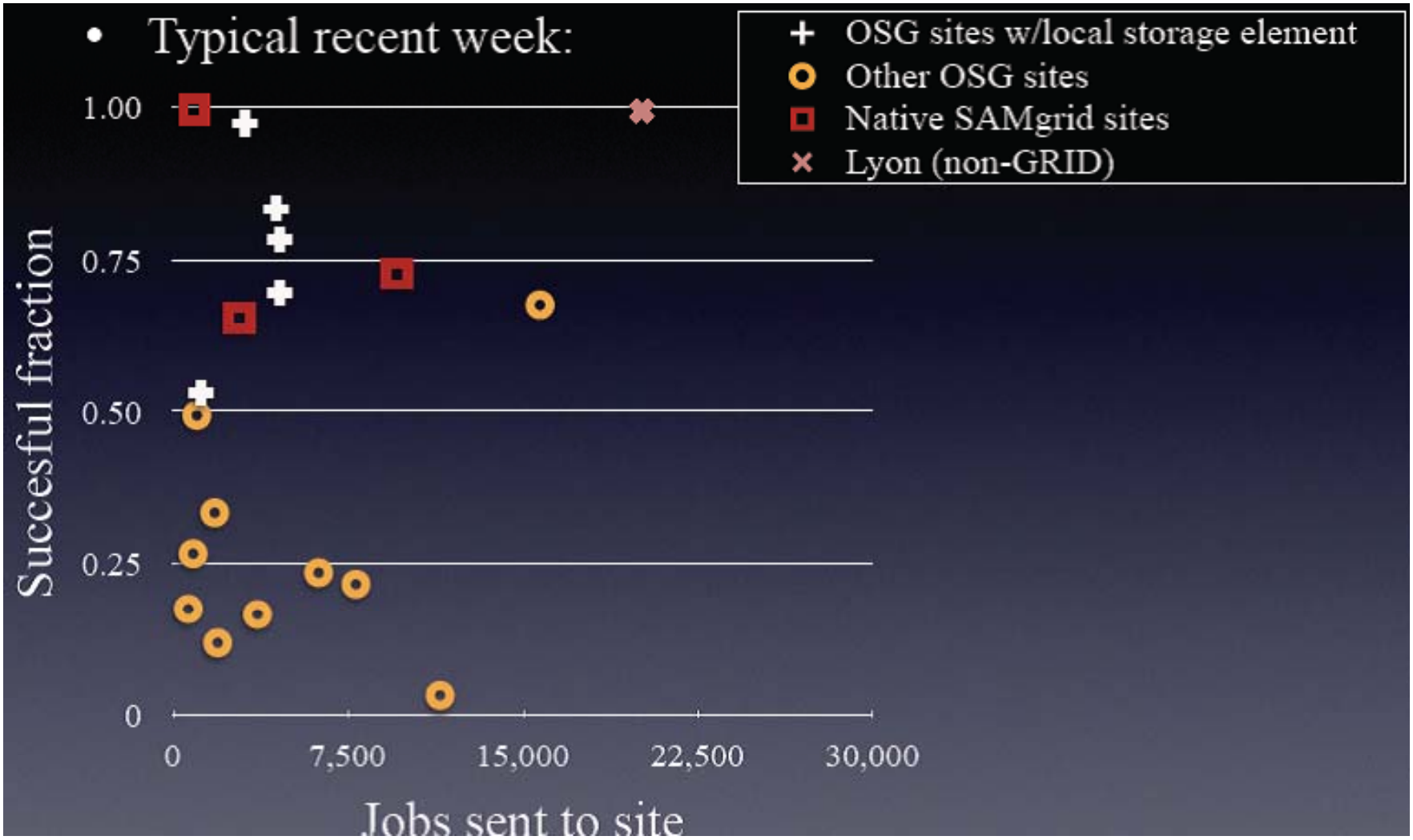
Summary

- For 100 batch slots D0 should be able to access storage at
  - 33Mbytes/sec
    - Is an issue for WAN storages
  - Minimum storage response time 3.6 secs per request.
    - 0.27 requests per sec
- Scales with number of slots

Current deployment

- Nodes that serve both onsite and offsite processing needs
- dCache/SRM at UNL, IU , Wisc
  - "60% of the time, it works all the time"
  - Sensitive to outage of central services
    - GUMS as a recent example
  - Pool downtime cannot be detected by storage users
    - srmcp requests for files at unavailable pools do not transition to terminal states
      - Have to use ambiguous handling policies i.e timeout and remove
    - Srmget requests continue returning urls for inaccessible data.
  - Transient auth. errors during peak usage
  - Not sufficient diagnostic output of the srm v2 interface calls (compared to v1)

Current efficiency
(plot credit goes to Joel Snow)



Conclusion

- Efficiency increases by about a factor of two when there is a local Storage Element (SRM interfaced) - on the site LAN - where D0 data can be moved and then accessed by the application on the local worker nodes through the use of GridFTP.
- The space needed is ~500 Gigabytes per site.
  - D0 then manages this as part of the job submissions.
- Have tested with dCache Ses(srmv1 and v2), should work with Bestman and xrootd and D0 are happy to test with these if storage is available.