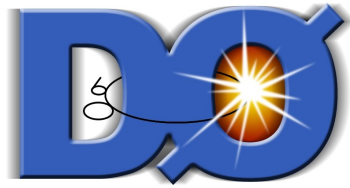# DZero Monte Carlo Data Handling On OSG
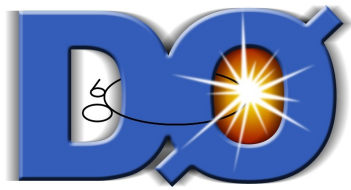

OSG All Hands Meeting
March 3, 2009

Joel Snow
Langston University

Outline

- SAM

- SAMGrid

- MC Applications

- Grid Job Flow

- Storage Element Use

- Summary

# Sequential data Access via Metadata

- Fermilab system first used by DZero

- SAM data handling system predates grid

- Set of servers working together to store and retrieve files and metadata

- Permanent storage and local disk caches

- Database tracks location, metadata of files, job processing history

- Delivers files to jobs (using GridFTP over WAN), provides job submission capabilities
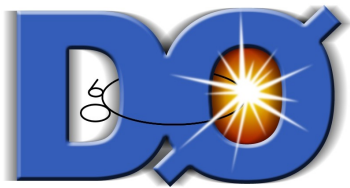
SAMGrid

- Fermilab developed grid predates OSG, first used by DZero for globally distributed computing

- SAMGrid = SAM + Job and Information Management (JIM) components

- Provides the user with transparent remote job submission, data processing, and status monitoring.

- VDT based (Globus + Condor)

- As Open Science Grid (OSG) and LHC Computing Grid (LCG) became operational it was desirable to leverage these resources for DZero ⟶ interoperability

- Forwarding node packages SAMGrid jobs for OSG/LCG job submission
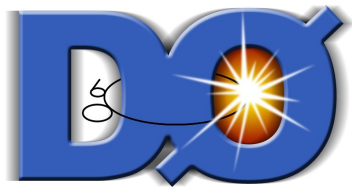
MC Applications

- Typical request has 4 phases – 1 appl. each

  - Generator – physics of interest is created
  - Simulator – propagation of particles of interest and decay products through the detector
  - Digitizer – Put simulated data in the form of raw data and overlay with generic background
  - Reconstruction – Reconstruct with first pass data analysis code

- Metadata of all phases saved in SAM DB

  - Typically only the output file of the last phase is eventually saved on tape at FNAL

MC Application Data Flow
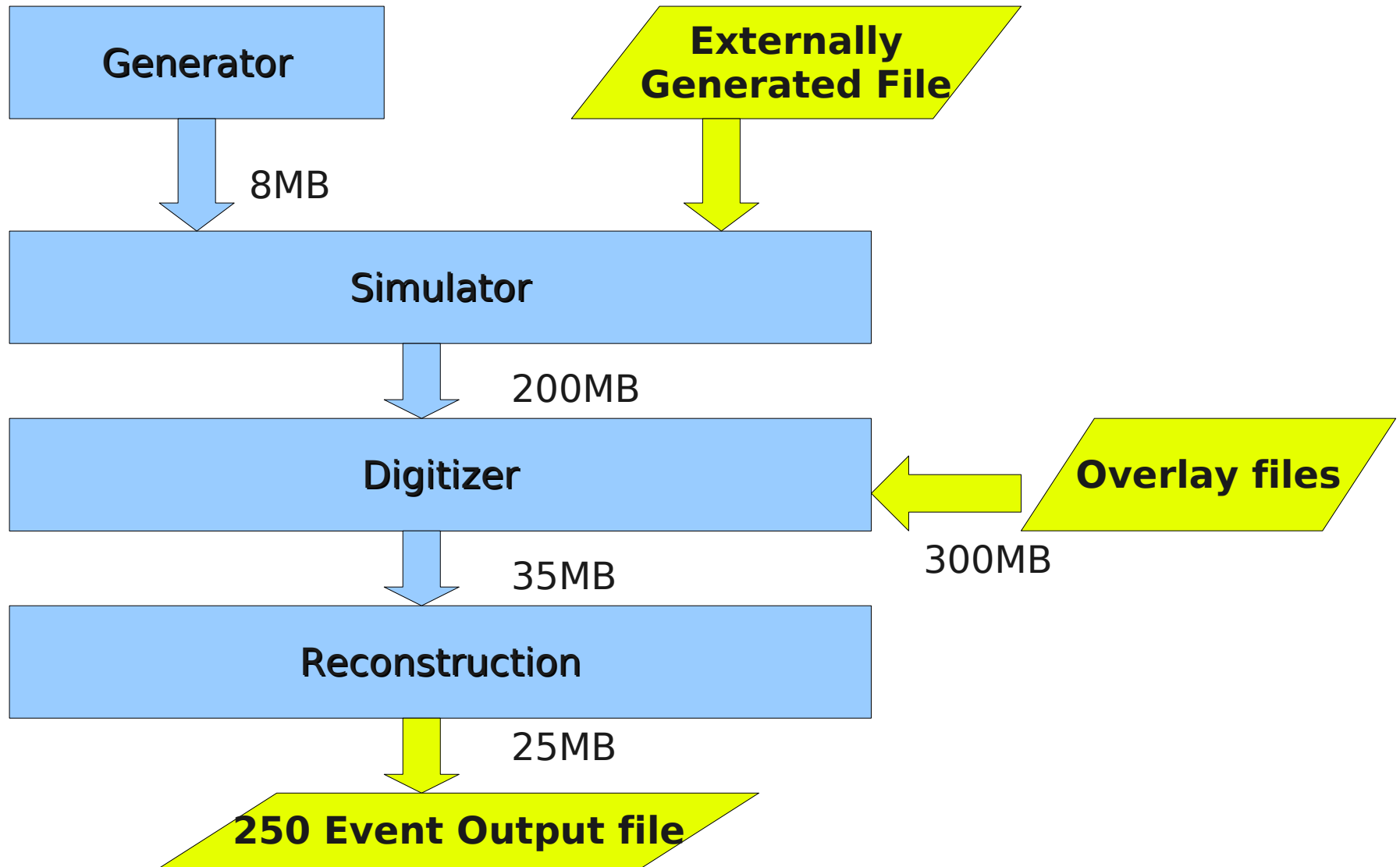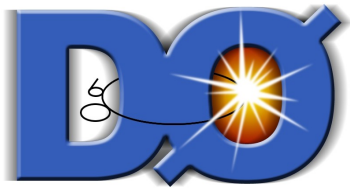
MC Grid Job Flow

- Bootstrap executable from Exec/Fwd node 4MB

- Initial environment/utility files from Exec/Fwd node 20MB

- Applications and execution environment from SAM cache (local or remote) 800MB

- Optional input data file, overlay files from SAM cache 0.2-1GB + 300 MB

- For OSG or LCG jobs no VO specific pre-installed software required

- Output data file stored in "durable location" until merged
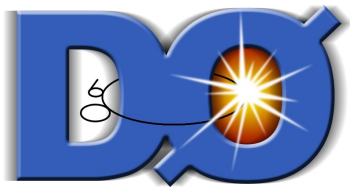
MC Grid Job Flow

- SAMGrid jobs are broken into 250 event chunks at the Exec/Fwd node for submission to the batch/Condor_g system

  – Execution time trade-off

- Output file size too small for efficient tape storage (20-30MB)

  – Merged in separate grid job

- The 10k event merged files (1GB) are stored on tape via SAM and unmerged files are deleted.

Storage Element Use

- WAN transfer failures to worker can lead to low efficiency

- LAN transfers to worker usually have high efficiency

- Use SRM at SE's to serve as local SAM cache for applications and environment files

- SE's in use at MSU, MWT2-IU, Purdue, UCSD, UNL

- Deployment resulted in significantly higher job efficiency

Summary

- DZero has adapted existing FNAL pre-OSG data handling (SAM) and grid (SAMGrid) to operate on OSG resources

- MC jobs require no pre-installed software at the OSG site

  - Everything needed for the job is retrieved from SAM caches or the forwarding node

  - Allows running opportunistically at many sites

- Use of SE's has increased job efficiency

- Has been a successful strategy for DZero

SAMGrid OSG/LCG

MC Job Flow On Worker

- Bootstrap executable unpacks to provide initial environment to obtain applications
  - gridftp client fetches job's sandbox files from SAMGrid-OSG forwarding node
- Sandbox SAM client (gridftp transport) retrieves application environment, executables, input files from SAM caches
- MC application chain runs writing metadata of files produced to SAM DB and persistent output files to SAM storage
- No pre-installed software at site required