Resource Selection in OSG & SAM-On-The-Fly

Parag Mhashilkar

(Computing Division, Fermilab)

Condor Week 2006

April 25, 2006

Resource Selection in OSG

- Overview
  - Why Resource Selection?
  - Resource Selection Service in OSG
  - Collaborators Involved
  - Resource Selection Service Architecture
  - Current Status
  - Future Work

Why Resource Selection?

- Several Grid resources available to run the job. User needs to know and keep track of availability of each resource.
- Each resource can provide specific services that can be advertised as a part of resource advertisement.
- Jobs can have specific requirements that can be advertised as a part of job advertisement.
- There is a need for a service to match your jobs to one of the available resources based on the information available in job and resource advertisements.

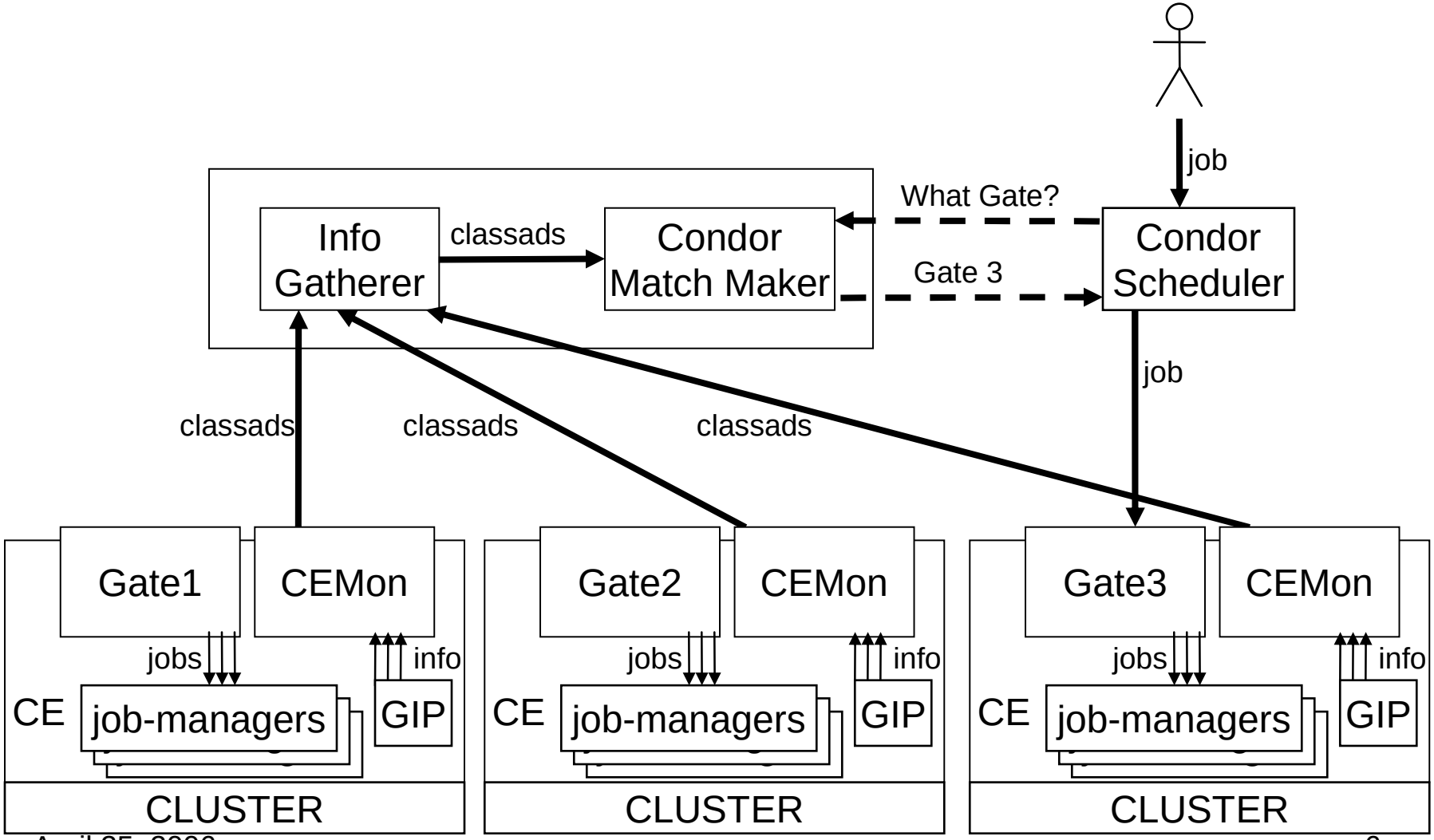
Resource Selection Service (ReSS) in OSG

- Why?
  - Emphasis on supporting several Virtual Organizations (VO) based on policies.
  - VOs can tag resources which are certified to run their jobs making resource selection more manageable.
  - VOs can use resources that provide specific features.
  - Resources are advertised using GLUE Schema which cannot be used by existing match making services directly.
- ReSS
  - The Resource Selector is a component of the OSG Job Management Infrastructure.
  - The project started in Sep 2005.
  - Sponsored by PPDG as a DZero contribution to the Common Project
  - Develop and deploy a Resource Selection Service that VOs with requirements on job management similar to DZero can use.

Collaborators Involved

- VOs
  - Dzero
  - Atlas
  - LIGO
  - FermiGrid
- Fermilab
- OSG TG-MIG group
- CEMon group from INFN
- Condor group from UW Madison

Resource Selection Service Architecture



Architecture …

- Generic Information Provider (GIP) describes resources in LDIF format using GLUE Schema.
- CEMon provides flexible plug-in mechanism to translate classads. Plug-in developed to convert the GLUE schema to condor classads.
- Information Gatherer subscribes to several CEMons to gather the information about the CEs and advertises it to several condor pools.
- Work is being done to expose the web service interfaces for the condor_collector and condor_negotiator.

Current Status

- First release of the ReSS is scheduled to be included in OSG ITB-0.5.0
  - Focus on testing functionality, scalability and stress test of Information Gatherer.
  - Validate Classads from different sites so they can be used for common resource selection criteria.
  - Study the scalability and investigate how IG handles O(10) CEMon registrations and O(100) classad processing and transferring to the condor_collector.
  - Stress test study of the IG. Simulate the load of the production environment by increasing 10 times the frequency of classad publication by the O(10) CEMon's.
  - Stress test the match making infrastructure submitting O(1) job/sec for 1 hour. In particular, evaluate the efficiency of the condor_negotiator call-out code, to match elements of an attribute list.

Future Work

- Improve installation procedures and integration with VDT, possibly by OSG-0.6.0
- Work with other VOs with similar requirements and extend the support of ReSS service.
- Use the web service interfaces of condor_collector and condor_negotiator to publish classads to condor pools

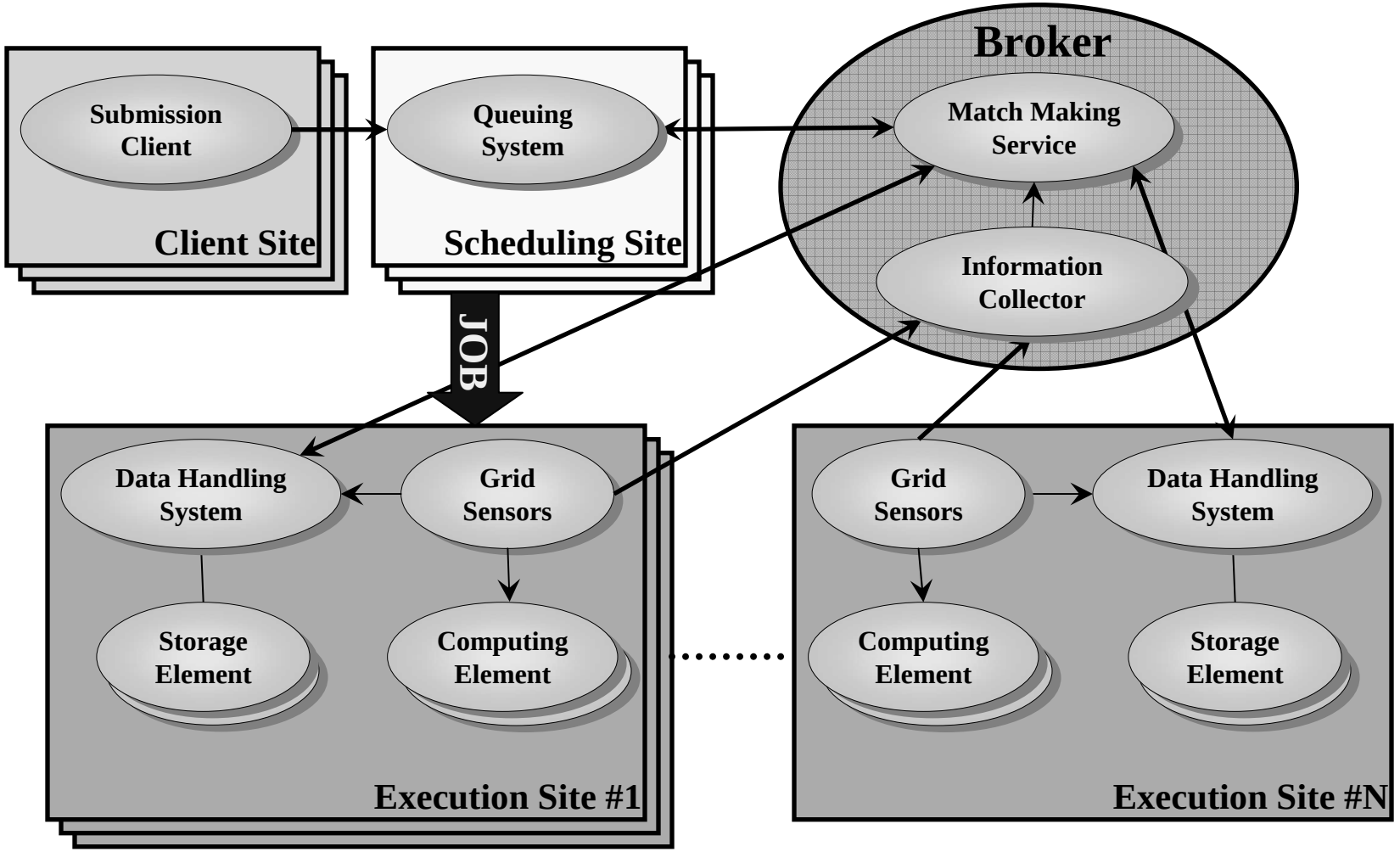
Sam-on-the-fly

- Overview
  - What is SAM?
  - SamGrid Architecture
  - SAM as a Distributed System
  - Why sam-on-the-fly?
  - Challenges
  - Current Status

What is SAM?

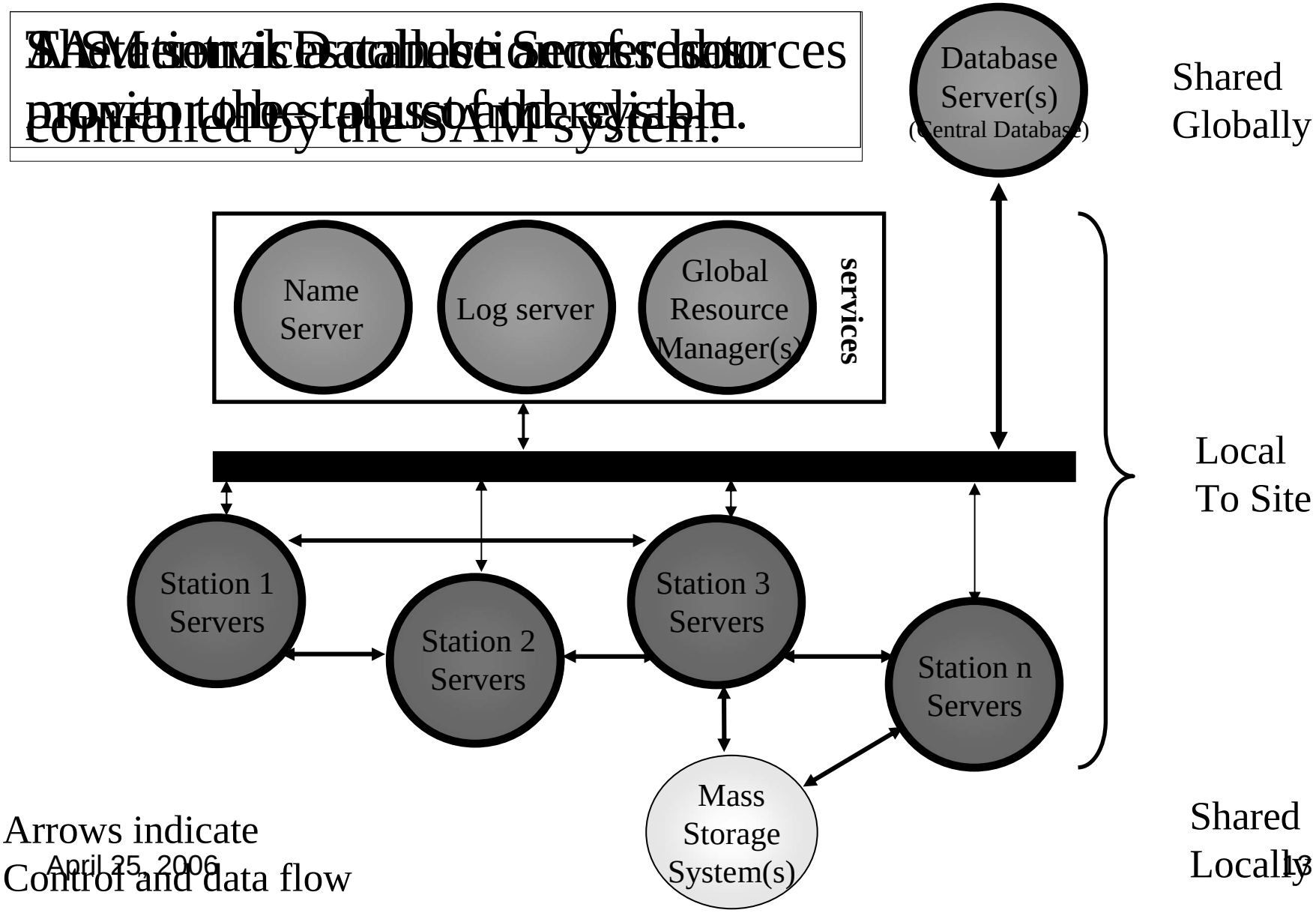
- Stands for ‘Synchronous Access via Metadata’ (SAM). The project was started in 1997 by DZero
- In SAMGrid
  - Condor forms the back-bone for providing scheduling, match making services at grid level.
  - SAM is the data handling component.
- Experiments: Dzero, CDF, MINOS

SAMGrid Architecture



SAM as a Distributed System

SAM is a distributed system that provides access to data and computing resources controlled by the SAM system.



Why Sam-on-the-fly?

- Grid Interoperability
  - SAMGrid-LCG Interoperability was a success
  - SAMGrid-OSG Interoperability moving towards production
  - Exposed lot of resources that SAMGrid users can use.
- Sites have resources that are available for a VO for longer duration.
- Need to dynamically
  - Deploy and configure services
  - Register services
  - Start and Stop services.
  - Do the cleanup after everything is done.

Challenges

- Current deployment of SAM services
  - Complex and based on the resource.
  - Manual intervention required for
    - Registering new services
    - Configuring new services
- Firewall configurations

Current Status

- Automated the deployment steps.
- Automated the registration steps.
- We hope to have working prototype soon.
- This project is a work in progress.
- People: Fermi National Accelerator Laboratory, University of Wisconsin Madison.

References

- Resource Selection Service for OSG
  - http://www.opensciencegrid.org
  - http://osg.ivdgl.org/twiki/bin/view/ResourceSelection/WebHome
- SAM
  - http://d0db.fnal.gov/sam