ATLAS and ARDA

ARDA Workshop

David Adams
BNL
June 21, 2004

David Adams

ATLAS

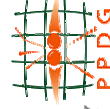BROOKHAVEN
NATIONAL LABORATORY

PPDG
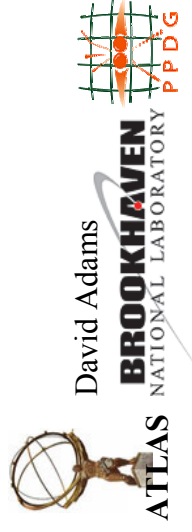
Contents

Introduction

ATLAS ARDA plan

Distributed analysis model

Analysis service scenario

Current status of ARDA

Plans for ARDA

Catalog services

ARDA analysis service

Using other gLite services

EGEE MW architecture

Conclusions

David Adams

BROOKHAVEN
NATIONAL LABORATORY

ATLAS

PPDG

ATLAS and ARDA   ARDA Workshop   June 21, 2004   2

Draft 2

Introduction

ATLAS looks to ARDA to provide

- Bridge to the EGEE middleware (gLite)
- Prototype distributed analysis system based on gLite

ATLAS looks to gLite to provide

- Middleware services for EGEE sites
  – Specifics to follow

- Support for implementing our own services on EGEE
  – Alternate implementations of EGEE services
  – Services not included in EGEE model

- Support for all ATLAS grids
  – Including U.S. (grid3+, OSG) and Nordugrid
  – Or we must support their middleware as well
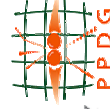
David Adams   ATLAS and ARDA   ARDA Workshop   June 21, 2004   3

Draft 2

ATLAS ARDA plan

ATLAS-ARDA activities include

- Delivery of an end-to-end distributed analysis system
- Study of metadata services
  - Later talk by Solveig

ARDA team is responsible for

- An analysis service based on gLite
- Deploying this service with significant compute resources

ATLAS is responsible for

- Distributed analysis model
- Software to support this model (AJDL/DIAL/GANGA)
- User interface including client(s) for analysis service
- Data, metadata and transformations
- Serious evaluation of end-to-end system using gLite

Draft 2

David Adams

BROOKHAVEN NATIONAL LABORATORY

PPDG
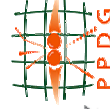
ATLAS and ARDA   ARDA Workshop   June 21, 2004   4

Distributed analysis model

User interactions

- **Select a dataset**
  - AJDL, not HEPCAL

- **Examine dataset**
  - Display content (event ID's, type-keys)
  - Location of data (logical files, physical files, DB, …)
    - > Including means to localize data

- **Define transformation**
  - Application that processes the data
  - Task carrying data to configure application
  - Select from catalog and modify

- **Create job to apply transformation to dataset**
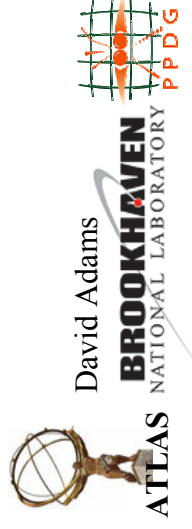 - Result is a new dataset

Distributed analysis model (cont)

User interactions (cont)

- Monitor job
  - Status (running, done, failed, …)
  - Status of sub-jobs
  - Examine partial result
- Examine result (output dataset)
- Repeat

Following block diagram shows major components

- Client
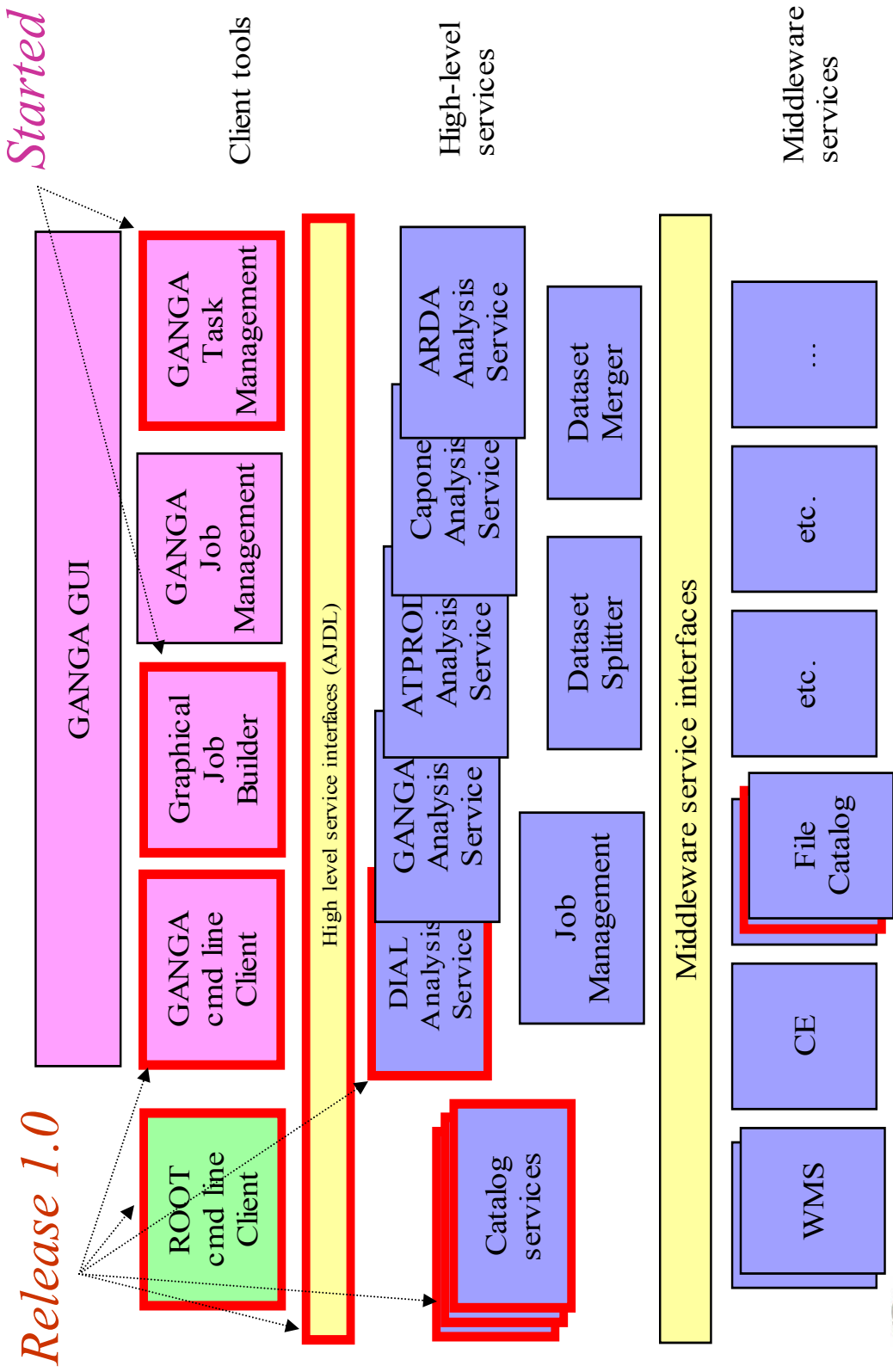- High-level services
- Middleware services

David Adams

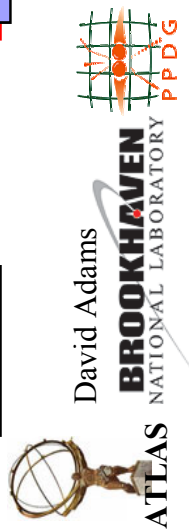ATLAS and ARDA   ARDA Workshop   June 21, 2004   6

Distributed analysis model (cont)

Release 1.0

Started

GANGA GUI

| GANGA cmd line Client | Graphical Job Builder | GANGA Job Management | GANGA Task Management |

Client tools

ROOT cmd line Client

High level service interfaces (AJDL)

| DIAL Analysis Service | GANGA Analysis Service | ATPROD Analysis Service | Capone Analysis Service | ARDA Analysis Service |

High-level services

Catalog services

| Job Management | Dataset Splitter | Dataset Merger |

Middleware service interfaces

| WMS | CE | File Catalog | etc. | etc. | ... |

Middleware services

Draft 2

David Adams

ATLAS and ARDA

ARDA Workshop

June 21, 2004   7

Analysis service scenario

Typical job processing scenario in analysis service

- (There are other possibilities)

Job request received

- Application, task, dataset and job preferences

Service locates application

- Application includes
  - Build and run scripts (preinstalled at present)
  - List of required software packages (not yet)
- Use existing location if installed
- Otherwise install from application description
  - Copy build and run scripts
  - Install missing software packages
    > Using package management service

David Adams

BROOKHAVEN
NATIONAL LABORATORY

PPDG
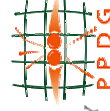
ATLAS and ARDA

ARDA Workshop

June 21, 2004   8

Draft 2

Analysis service scenario (cont)

Install task and build transformation

- Task includes
  - Named text files
  - Named parameters (not yet)
  - Named logical files (not yet)

- Copy these if not already installed
- Transformation is build of task for given application
- Locate previous build of transformation
- If absent, use script in application to build from task

Compare dataset content with that required for xform

- Both dataset and xform (not yet) advertise content
- If not same or superset, job fails
- If superset, create dataset with only the required subset

David Adams

BROOKHAVEN
NATIONAL LABORATORY

PPDG

ATLAS and ARDA   ARDA Workshop   June 21, 2004   9

Draft 2

Analysis service scenario (cont)

Locate concrete representation of dataset

- If given dataset is already concrete, use it
- Otherwise consult DRC (dataset replica catalog)
- Choose between available replicas depending on
  - > Data placement
  - > Available CPU resources
  - > Both (matchmaking)
    - – Use information services in decision making
- May be necessary (or easier) to submit job to build a representation of the input dataset
- Location for the concrete dataset is typically a collection of logical files

David Adams   ATLAS and ARDA   ARDA Workshop   June 21, 2004   10

Draft 2

Analysis service scenario (cont)

Split dataset

- Typically along dataset or file boundaries
  - But not necessarily
  - Might have finer splitting for interactive response
  - Or coarser to generate bigger output files
- Result is a collection of sub-datasets
- Might match sub-datasets to sites
- Might be a service to do splitting

Create compound job

- From list of sub-datasets
- CJ creates one sub-job for each sub-dataset
- CJ manages sub-jobs

David Adams

ATLAS and ARDA   ARDA Workshop   June 21, 2004 11   Draft 2

Analysis service scenario (cont)

Return compound job ID to user

Manage compound job

- Start job
- Update regularly until finished

Compound job manages sub-jobs

- Start sub-jobs when slots are available
- Merge results from sub-jobs when they complete
  – Result is output dataset
  – Might use merging service
  – Partial result is available

Return job status to user on request

- Includes state and partial result
- Either compound job or any of the sub-jobs

David Adams

BROOKHAVEN NATIONAL LABORATORY

ATLAS

P P D G

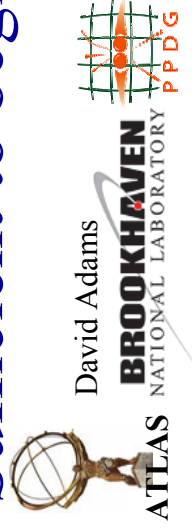ATLAS and ARDA   ARDA Workshop   June 21, 2004 12
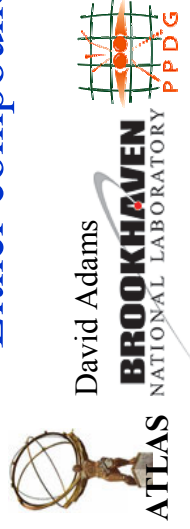
Draft 2

Current status of ADA

The current ADA (ATLAS Distributed Analysis):

- Provides service using fork, LSF or condor

- Command line client using root

- Dataset catalogs in place
  - Mostly DC1 data

- Applications
  - DC1 CBNT → histograms in place
  - DC2 reconstruction for release 8.0.1
  - Working on DC2 event data → histograms

- Services
  - Running interactive service at BNL

Sufficient to begin ARDA development

David Adams

ATLAS and ARDA   ARDA Workshop   June 21, 2004 13
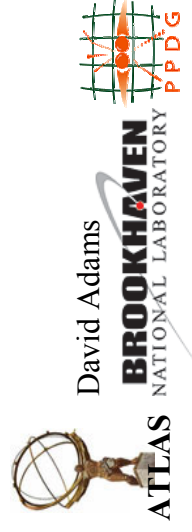
Plans for ADA

Over the summer we plan to do the following

- Add remaining catalog services
- Move catalogs to AMI
- Add file catalog support for DC2
  - Most likely using Don Quijote
- Populate dataset catalogs with DC2 data
- Develop and deploy ARDA analysis service

Summer or fall

- Add user-definable applications
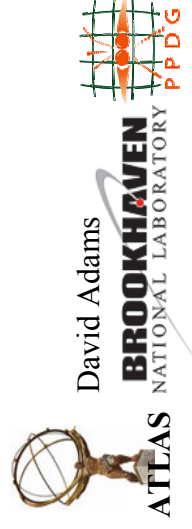- Add package management service

Draft 2

Catalog services

Repositories

- Store XML representation indexed by ID
- Needed for the following components
  - Datasets
  - Tasks
  - Applications
  - Jobs

David Adams

ATLAS and ARDA   ARDA Workshop   June 21, 2004 15

Catalog services (cont)

Selection catalogs

- Enable users to select a component ID based on assigned metadata

- Need the following
  - DSC – dataset selection catalog
  - TSC – task selection catalog
  - Maybe ASC and JSC

- DSC includes VDC (virtual data catalog)
  - Output dataset ID indexed by application, task and input dataset ID's
  - Datasets are virtual

Draft 2

Catalog services (cont)

Other catalogs

- **DRC – dataset replica catalog**
  - Associate virtual dataset ID with concrete replica ID's
- **SFDC – single file dataset catalog**
  - Record associations between single-file datasets and their files
  - To avoid duplicating datasets
- **Dataset annotation**
  - Allow users to attach annotations to aid in dataset selection
  - For user examination rather than query
- **Standard file replica catalog**
  - From middleware provider(s)

Draft 2

ARDA analysis service

Following outlines ideas for implementing the ARDA analysis service

- In order of increasing complexity
- And making use of more and more EGEE functionality
- All build from DIAL

1. EGEE fork

- Instead of forking jobs, fork script to send to prepare jobs and send to WMS
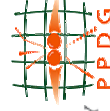- Not scalable—one sub process for each unfinished job

2. EGEE job

- Provide EGEE job subclass
- Again prepares job and sends to WMS
- Scalable—poll or callback to know when jobs finish

David Adams

ATLAS and ARDA   ARDA Workshop   June 21, 2004 18

Draft 2

ARDA analysis service (cont)

3. EGEE compound job
   - Make EGEE-specific compound job
   - Merge jobs as they finish instead of polling
   - Take advantage of EGEE merging strategy

4. EGEE scheduler
   - Scheduler is the basis for the DIAL analysis service
   - Could make use of EGEE staging or splitting
   - Want to encapsulate and re-use DIAL transformations

5. EGEE web service
   - Base on WSDL instead of DIAL web service infrastructure
   - Otherwise similar to the previous
   - Might be a better match to other EGEE services

David Adams

ATLAS and ARDA   ARDA Workshop   June 21, 2004 19

Using other gLite services

We might also make other direct uses of other EGEE middleware services

- In analysis and/or production
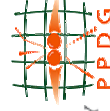- May or may not be part of ARDA prototype

Examples:

- WMS for production
  - Analysis is covered by ARDA prototype
- Data movement and cataloging
  - File catalog interface in DIAL
- Metadata cataloging
- Package management

Draft 2

EGEE MW architecture

Following are comments on EGEE MW architecture

- Mostly my comments—not sanctioned by ATLAS
- Indicate some of what is not understood or needed
- Also things that would be nice to have
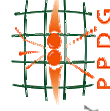- Intended to be constructive...

General comments

- Concern about our ability to influence GLite
  - Single well-advertised mailing list for LHC comments
    > With public responses from the gLite team
    > (Thanks for those so far)
- What are plans for deploying gLite outside EGEE?
- Concerns abut replicating rather than supporting AliEn

David Adams   ATLAS and ARDA   ARDA Workshop   June 21, 2004 21

Draft 2

EGEE MW architecture (cont)

GAS

- Is authorization this more than forwarding GSI proxies?
  - If so, why do it?
- For service discovery and creation, should we wait for evolution to WSRF?
- Concern about integration with non-EGEE services

Global file system

- Arguably a useful API for some activities but services should not depend on this
- Concern about integration with non-EGEE systems

CE and WMS

- Are these really different?
- If not, could I a build a hierarchy more than two deep?

Draft 2

EGEE MW architecture (cont)

Data management

- Very large collection of services
- Not clear which are intended for external use

Metadata

- Would like to have a generic metadata service
- Not one restricted to files

Package management service

- A service which actually did management including location, installation and lifetime management would be most useful
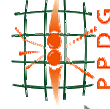
Use cases

- Helpful: like to have many more

David Adams   ATLAS and ARDA   ARDA Workshop   June 21, 2004 23

Draft 2

Conclusions

ATLAS is well on its way to delivering distributed production and analysis systems

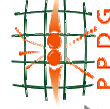- DC2 production starting now
- First release of ADA two weeks ago

Look forward to working with ARDA to deliver a distributed analysis system based on gLite

- ATLAS providing clients and infrastructure
- ARDA providing EGEE analysis service
- ATLAS will use prototype to provide feedback to gLite
- ATLAS users get a gateway to EGEE resources
- Mutual exchange of ideas

David Adams   ATLAS and ARDA   ARDA Workshop   June 21, 2004 24

Draft 2