



# Early Thinking on ARDA in the Applications Area

Torre Wenaus, BNL/CERN  
LCG Applications Area Manager

PEB

Dec 9, 2003



# So far...

- ARDA AA Meeting Nov 27
  - Small first meeting with the AA LCG people we expect to be involved with a substantial amount of their time
    - Derek Feichtinger, Juha Herrala, Kuba Moscicki, Frederick Orellana
    - Plus Frederic Hemmer, Predrag Buncic, Dirk Duellmann, Alberto Aimar and myself
  - Had an overview of ARDA from Predrag, then discussed possible areas of AA activity
- Frederic's initial ARDA middleware meeting
  - TW attended (some of it) as AA rep
- The following is based mainly on Nov 27 meeting notes + feedback



# General ARDA AA Objectives

- **Common software above the middleware layer**
  - Adapting, extending, interfacing AA software for ARDA
  - Participating in ARDA interface definition; ensuring AA requirements met
    - **ARDA interfaces insulate users from underlying technology while allowing to immediately leverage existing implementations**
  - Applying lower level middleware services to provide specialized higher level services directed at HEP and analysis
- **Integration and validation**
  - Integrating ARDA middleware services and analysis application level services into end-to-end distributed analysis prototype
  - Assisting integration of distributed analysis prototype or components thereof into experiment environments
  - Validation of the prototype and feedback to middleware providers



# General work areas

- 1) Event data management and access
  - 2) Framework integration services
  - 3) Provenance and session state information management
  - 4) Interactive analysis tools
  - 5) Analysis environment integration and validation
- ... and including first thoughts on work package organization

We expect ARDA will use SPI services and policies



# Event data management and access

- Event collections, physics-level datasets, physics queries
- Efficient sparse data access
- Data access below file level (event objects)
- Splitting at physics dataset level
- A mix of interface development, POOL work, ROOT work
- Collections work currently going on in a POOL WP, but this work needs an 'analysis' perspective and not just a 'persistency' perspective - ARDA can provide that
  - Make this ARDA work package a joint work package with POOL Collection WP



# Framework integration services

- Interfacing/integrating framework-level distributed services
  - Distributed messaging, error handling, logging, ...
- Interactive interface; Python, ROOT bindings
- Framework access to more sophisticated middleware services?
  - Workflow management, replication, ...
- Probably mostly a very 'thin' activity
  - not developing services, or even probably the interfaces
    - the middleware people will do this, though this WP will probably contribute to interface definition
  - just packaging/integrating them for the AA architecture
  - Maybe some specialization of generic services (such as next area...)
- The long-empty 'grid based services' box in SEAL
  - Joint ARDA/SEAL WP



# Provenance and session state info management

- Higher level provenance info services as an application of the generic provenance service provided by middleware
  - HEP specificity
  - Presentation of provenance info to the user
  - Unless everything above the generic service level is regarded as experiment specific
- Persistent analysis session support
  - Again adding ARDA analysis environment specificity above generic services, if all is not experiment specific
  - Customization of analysis environment
    - Support for non-standard algorithms, configurations
- Fold into 'framework integration services' WP



# Interactive Analysis Tools

- Interfacing to tools supporting interactive (low-latency, rapid-response) analysis
- ROOT, PROOF integration
- Interfacing to tools supporting 'chaotic' workload management
  - User level management/monitoring
  - User level reservations ('what' and 'when')
- Interfacing to tools supporting dynamic job interaction/control
- AIDA integration
- Needs will vary from experiment to experiment; maybe mostly experiment-specific integration
- Fold into the next WP...





# Analysis Environment Integration & Validation

- ARDA integration as an analysis system in experiment environments
  - Integrating experiment specific front end with ARDA back end
- Early priority: users in experiments testing detailed use cases using experiment-integrated ARDA
  - Get ARDA in the hands of (select) physicists doing analysis as soon as possible (as soon as there is a tool of interest to attract them - experiment ARDA teams need to sell the product)
- The key work package
  - Support four distinct but collaborative ARDA integration efforts in the experiments
  - Coordinate gathering of feedback from experiment ARDA teams/users
  - Provide overall coordination/coherence for AA ARDA
- No 'joint WP' arrangement with existing AA project



# Summarizing My Current Thoughts on WPs

## 1) Integration and Validation

- Main driver for ARDA in AA
- Primarily providing coordination, communication, coherence for integration efforts residing in the experiments
  - And ensuring close communication/feedback to middleware part of the project
  - Some similarity to Physics Validation in the simu project

## 2) Event data management

- Physics-driven event collections
- Incorporating POOL Collections WP

## 3) Framework integration

- 'Thin' adaptation of middleware services to whatever is required for integration in experiment analysis frameworks



# Next Steps

- Take account of PEB and SC2 feedback, circulate to apps area, discuss in AF
- Take account of general feedback, flesh it out, and use as a starting point for workshop discussion



# LCG Service Time-line

computing service

physics

open LCG-1 (achieved) – 15 Sept

LCG-2 - upgraded middleware,  
mgt. and ops tools

Second generation  
middleware prototyping,  
development

LCG-3 – second generation  
middleware

TDR for the Phase 2 grid

Phase 2 service acquisition,  
installation, commissioning

Phase 2 service in production

2003

2004

2005

2006

2007

Testing, with simulated  
event productions

principal service for  
LHC data challenges

Computing models

validation of  
computing models

experiment setup &  
preparation

first data



# High-Level Strategy for Middleware

- LCG-2 middleware package strongly supported and evolved  
demonstrating a base solution for LHC start-up  
supported until overtaken by ARDA
- ARDA -
  - Re-engineered generic middleware package
  - Incorporating experience from AliEn, EDG, ....., VDT
  - Architected for scale and performance requirements of LCG
  - "batch" and "analysis"

Fast prototyping approach - with clear end-to-end goals



# Workshop 21-22 January 2004

## Goals

- Explain and agree on middleware scope, approach, aims, target timescales
- Agree on what *should* and what *could* be done in common in the LCG AA (e.g. POOL, collections, meta-data, SEAL, ..)
- Agree on what is the responsibility of the experiments
- Explore a framework for an ARDA implementation project, coordinating -

Middleware  $\leftrightarrow$  LCG AA  $\leftrightarrow$  experiment analysis s/w  $\leftrightarrow$  end-users