Software & Grid Middleware for Tier 2 Centers

Rob Gardner
Indiana University

DOE/NSF Review of U.S. ATLAS and CMS Computing Projects
Brookhaven National Laboratory
NOVEMBER 14-17, 2000

Motivation

- The distributed LHC computing model developed by MONARC, and adopted by all 4 LHC experiments and the Hoffman review panel consists of distributed computational and data handling stores organized hierarchically (but flexibly) into "tiers"
  - Tier 0 CERN
  - Tier 1 A national facility
  - Tier 2 A regional facility
  - Tier 3 An individual institute
  - Tier 4 An individual

- The software which enables the experiments' software frameworks to work effectively, and to provide efficient access to the and computation in a distributed environment is the called

  Grid Middleware

- Middleware is broadly:
  - Distributed Computing Management and Intersite Load **Balancing**
  - Distributed Data Management
  - Task and System Tracking, Workflow Management

Grid Middleware

- Distributed computing
  - Workflow scripts, schedulers: intersite load balancing
  - Resource estimation, allocation
  - Security, authentication, authorization, prioritization
- Distributed data management
  - replication
  - transport
  - mass store APIs for integration
- Packages :
  - CONDOR – distributed computing package
  - Globus – wide security and authentication, file replication, information services
- Projects:
  - PPDG – adapt existing tools; first HEP Data Grid components
  - GriPhyN – petascale virtual data; the Grid
  - DataGrid – large scale EU initiative to develop and deploy the "Grid"

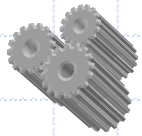
Grid Middleware functionality

- Physicist submits a job
- Middleware:
  - Estimates resource requirements and performance
  - Finds convenient places for it run
  - Organizes efficient access to data
    - Caching, migration, replication
  - Deals with authentication to the different sites
  - Interface to local site resource allocation mechanisms, policies
  - Run jobs
  - Monitor progress
  - Recover from problems
  - Collects, manages output
  - Archival bookkeeping (catalogs)

Software for Tier 2 Centers

- Local site management
  - Automated installation, configuration management, system maintenance
  - Automated monitoring and error recovery
  - Performance monitoring
  - Expressing and managing local Grid resources
- Mass storage management
  - data storage
  - uniform mass storage interface
  - exchange of data and meta-data between mass storage systems

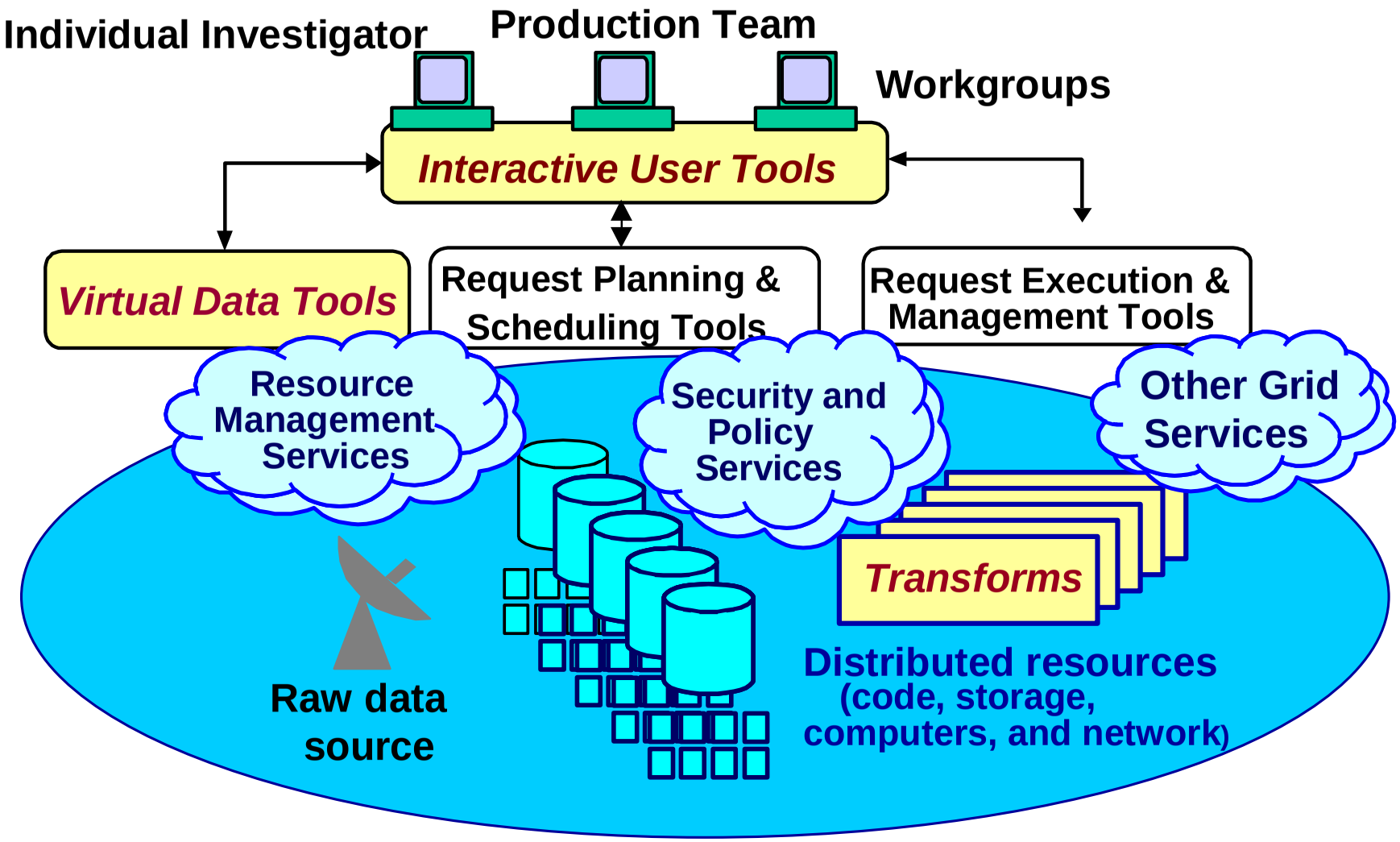
Frameworks, Interfaces and Adapters



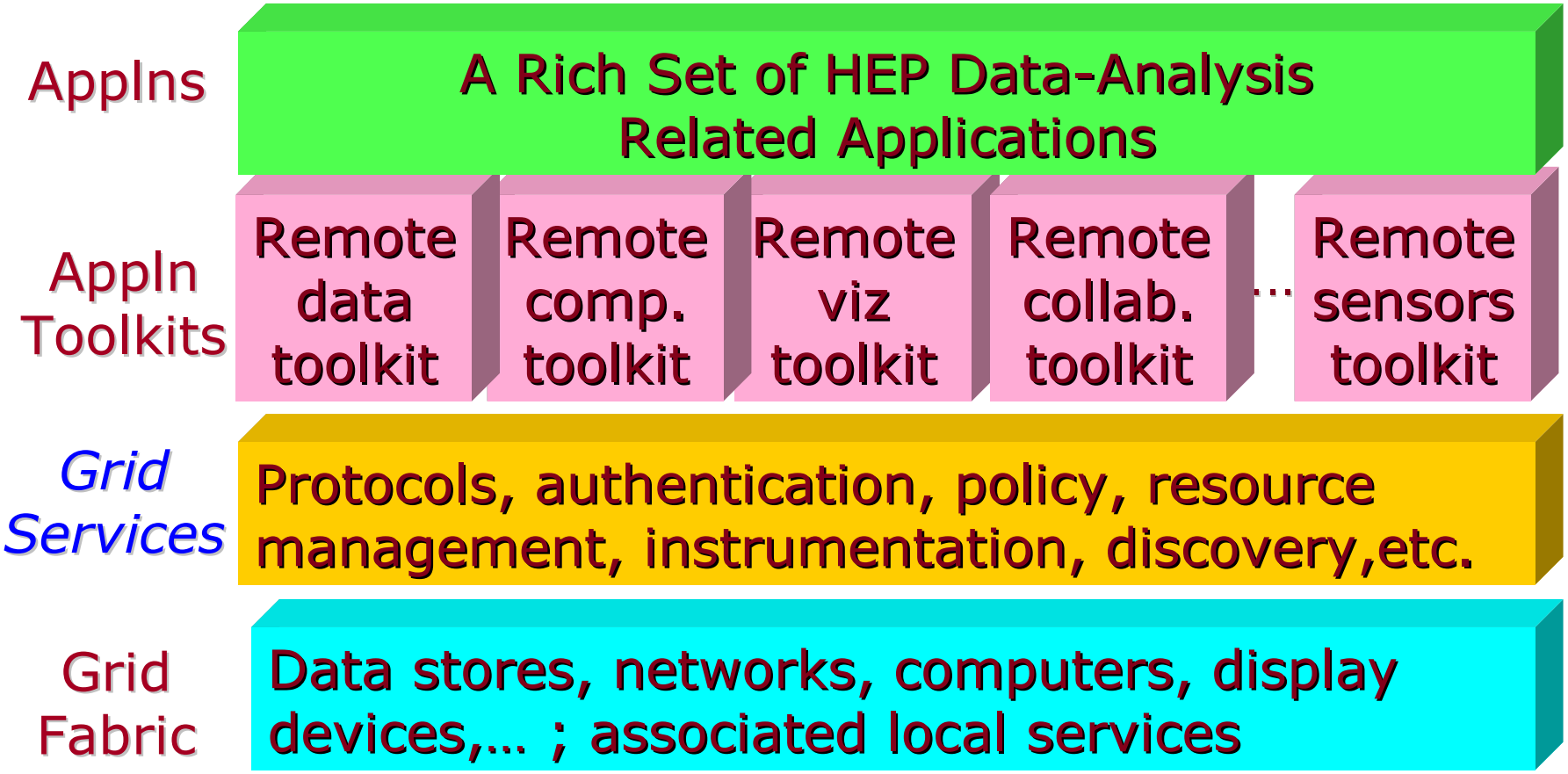
- Middleware is co-developed and written by "off project" computer scientists and some software engineers
- Interfaces to Software Framework are specified by core software engineers and physicists in consultation with grid software integration engineers and physicists– all "on project"
- Adapters, HEP or experiment specific, are written by grid software integration engineers and physicists
- Requirement: Core software can function completely independent of middleware; likewise, an app can be built grid-enabled without without specialized knowledge.

- WBS organization:
  - In CMS, grid software integration is initiated done by CAS engineers and by off project researchers (such as GriPhyN staff and postdocs), deployment and operation of final production codes in UF
  - In ATLAS, grid software integration engineers are accounted for under Facilities

GriPhyN: PetaScale Virtual Data Grids



Grid Services Architecture [*]



[*] Adapted from Ian Foster

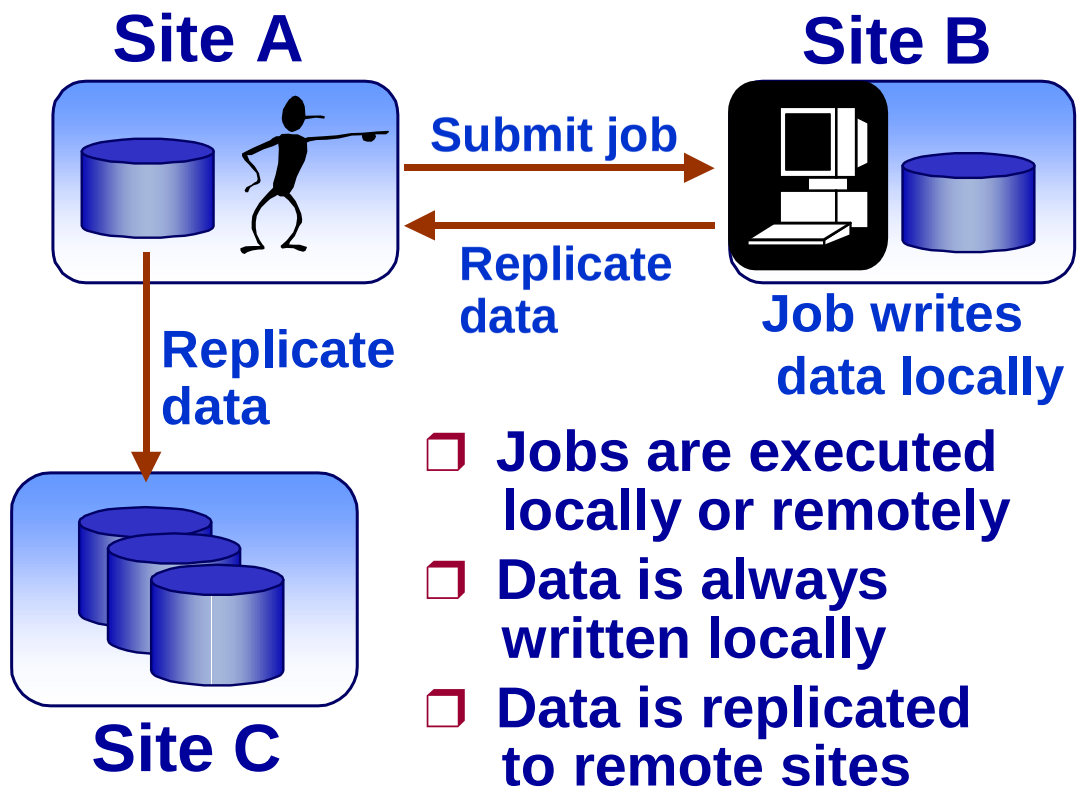
EU DataGrid

- Work areas:
  - Workload management
  - Grid data management
  - Grid monitoring services
  - Grid fabric management
  - Grid mass storage management
  - Grid integration testbed
  - Two application areas (HEP & Bioinformatics)
- Scale of effort:
  - 3 year effort
  - Many man-years
- National initiatives to build Tier 1 & 2 facilities across the EU

Grid Data Management Prototype (GDMP)

Distributed Job Execution and Data Handling:

- Transparency
- Performance
- Security
- Fault Tolerance
- Automation



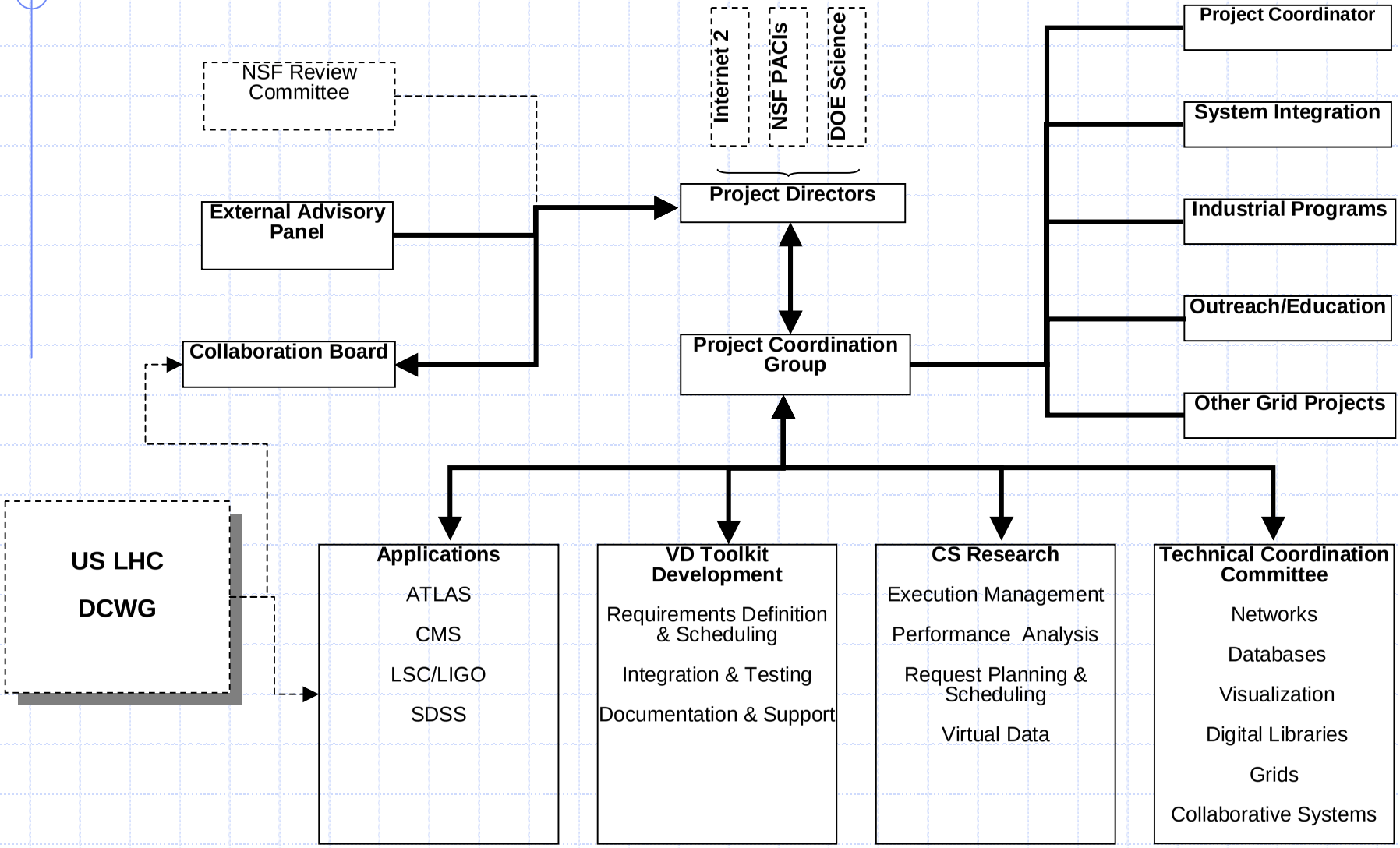
- Jobs are executed locally or remotely
- Data is always written locally
- Data is replicated to remote sites

GDMP V1.1: Caltech + EU DataGrid WP2
Tests by CALTECH, CERN, FNAL, Pisa for CMS "HLT" Production 10/2000;
Integration with ENSTORE, HPSS, Castor

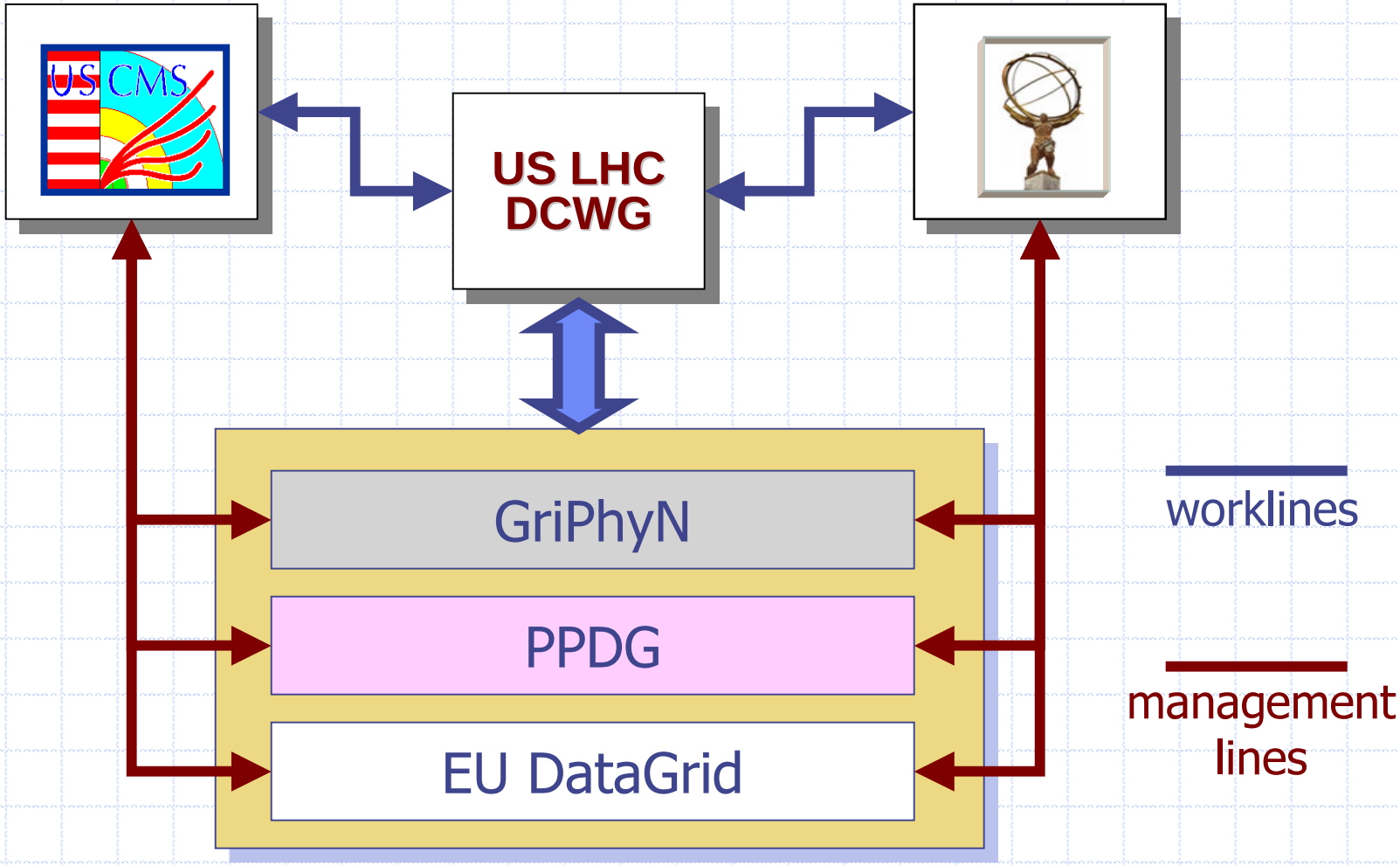
GriPhyN and LHC Computing

- Request planning and execution in a large scale production (ordered) and chaotic user analysis
  - Large numbers of LHC physicists
  - Wide area infrastructure
- Execution management & fault tolerance
- Performance analysis
- Strategy:
  - Physicists interact closely with CS and middleware developers
  - Develop and deploy prototype
    - tier 2 centers and testbeds to provide platform for testing and performance assessment, comparison with MONARC simulations
  - Integrate toolkits into Core software from the beginning
    - Grow the infrastructure adiabatically

GriPhyN Management



US LHC Distributed Computing Working Group



Summary

- The hierarchical LHC computing model is essential for physics and requires software which works in a distributed environment
- Close interaction and collaboration are required
  - Between physicists and computer scientists
  - Between the two LHC experiments
    - Many common problems
- GriPhyN collaboration is off to a good start and is an excellent opportunity for US ATLAS and US CMS to collaborate on common projects