Polish EU Grid Projects

Marian Bubak
Institute of Computer Science AGH
and
ACC CYFRONET AGH

bubak@agh.edu.pl

Kraków, ICFA Workshop, October 09, 2006

Overview

1. CrossGrid

2. Overview of FP6 Projects
   - K-WfGrid
   - EGEE, EGEE II
   - CoreGrid
   - Ambient Network
   - ViroLab
   - (int.eu.grid, GREDIA)

3. Projects at ICM and PSNC

4. Summary

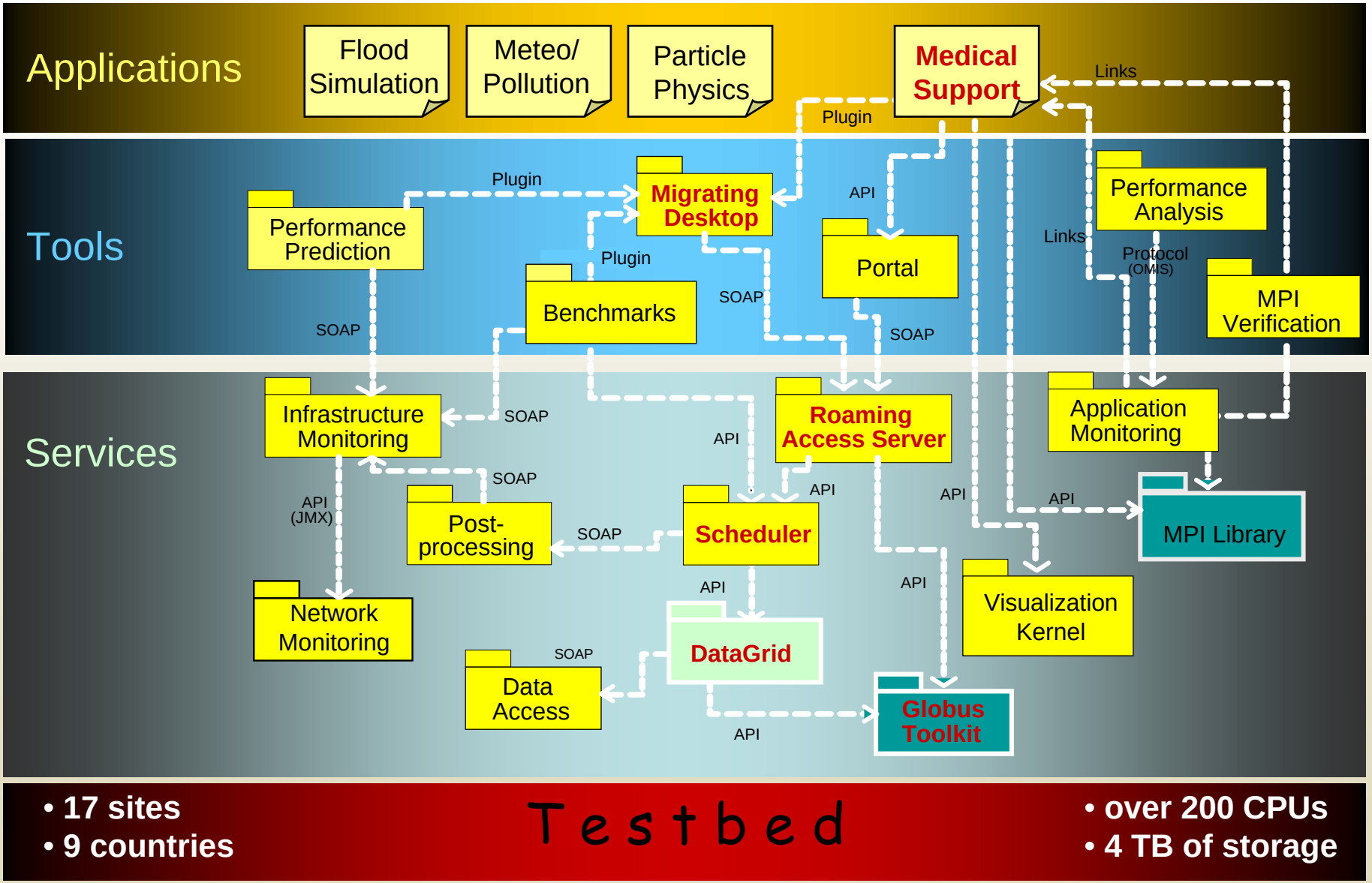
CrossGrid

www.crossgrid.org

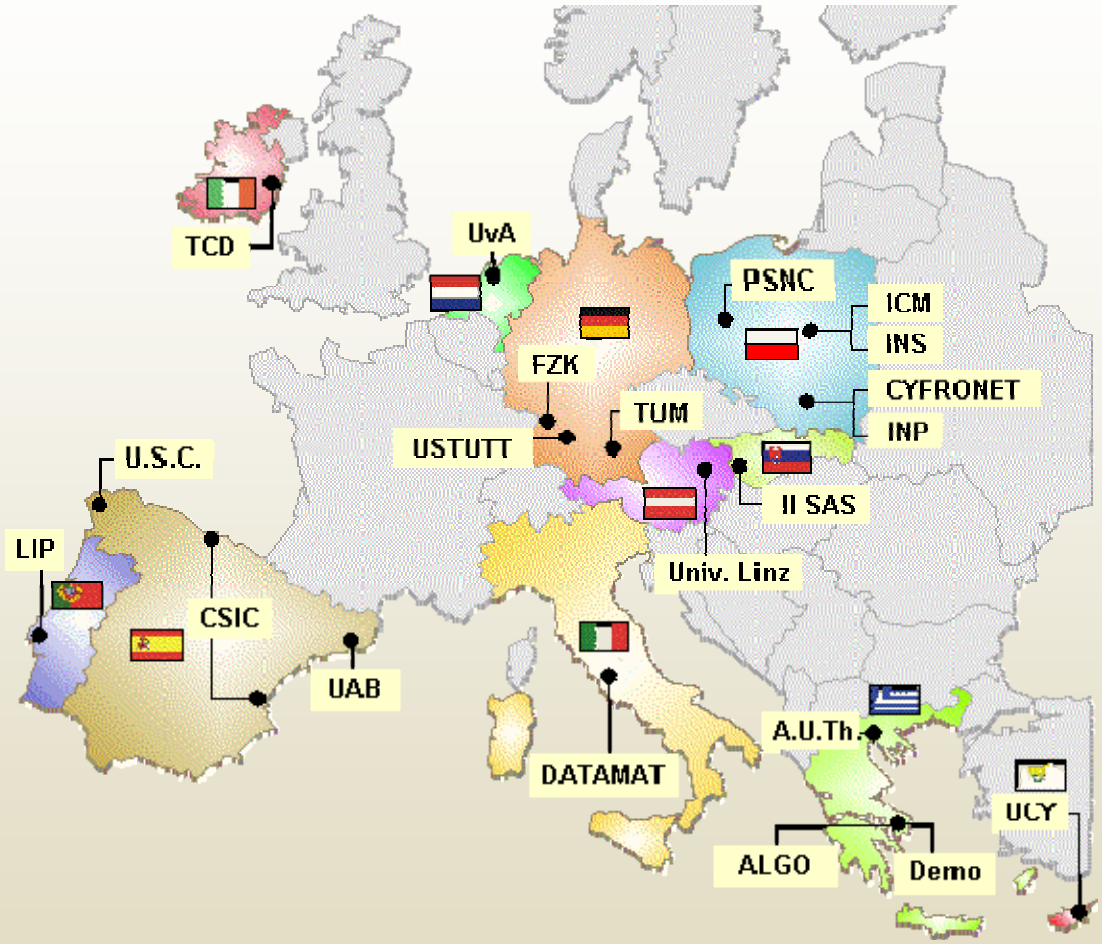
Development of Grid Environment for Interactive Applications

- New category of Grid-enabled applications:
  - compute- and data-intensive,
  - distributed,
  - near-real-time response (person in a loop)
- New programming tools
- Grid more user friendly and efficient
- Interoperability with other Grids
- Implementation of standards

CrossGrid Architecture



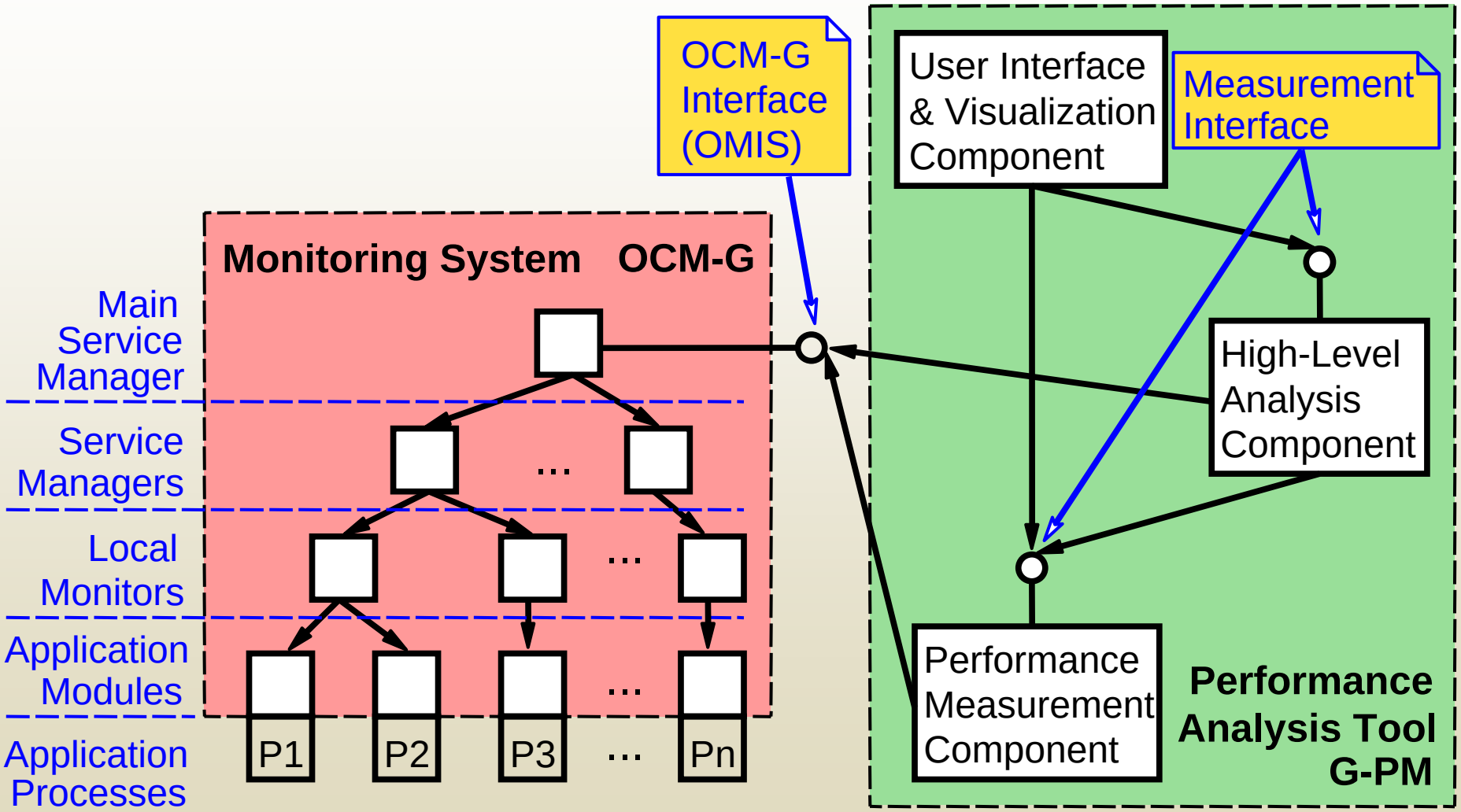
CrossGrid



- 21 partners
- 2002-2005
- Coordination
  - CYFRONET
  - Michał Turała
- Research areas
  - CrossGrid Applications
  - Grid Tool Environment
  - New Grid Services
  - International Testbed
  - Architecture

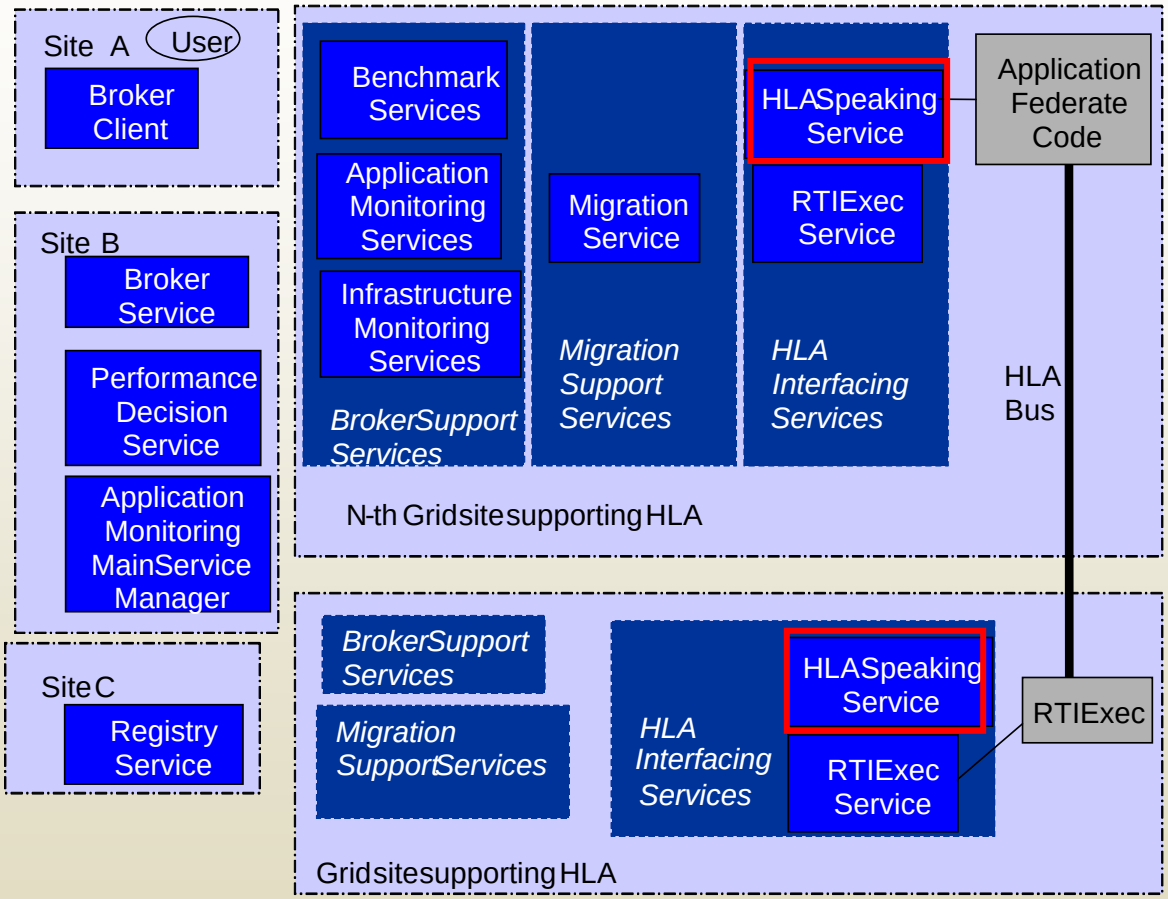
http://www.crossgrid.org

Performance Analysis Tool



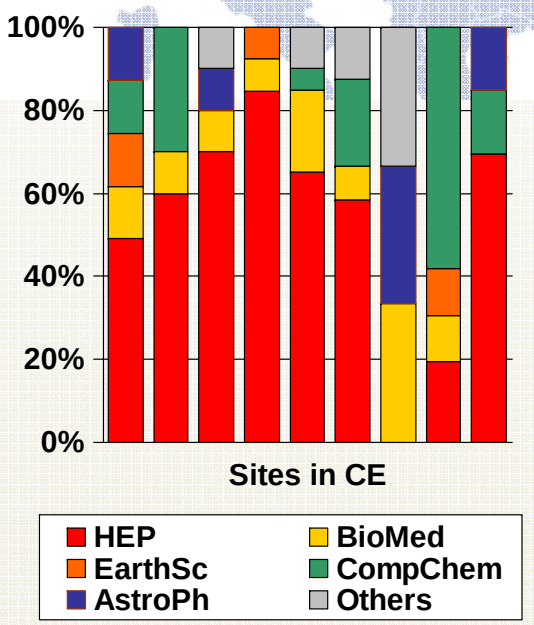
Grid HLA Management System

- **HLA interfacing services**
  - HLA-Speaking Service for managing federates
  - RTIExec Service for managing RTIExec (coordination process in RTI)
  - Broker for setting up a federation and deciding about migration
- **Broker decision services**
  - Registry for storing location of HLA-speaking services
  - Infrastructure Monitoring/Benchmarks for checking environment of HLA service
  - Application Monitoring for monitoring performance
- **Migration support services**
  - Migration Service for performing migration



Enabling Grids for e-Science in Europe

- CE Region Consortium consists of 12 institutes from 7 countries: CESNET, CYFRONET, IISAS, GUP ICM, JSI, KFKI-RMKI, PSNC, MTA-SZTAKI, NIIF, SRCE, UNIINNSBRUCK
- Operate, maintain and support EGEE Grid Infrastructure in CE region
- Provide researchers from a variety of scientific disciplines with computing resources of more than 1000 CPUs and 20 TB disk space
- Virtual Organizations supported in Central European region: HEP, computational chemistry, biomedicine, pharmacology, astrophysics, earth science and regional users (VOCE VO)

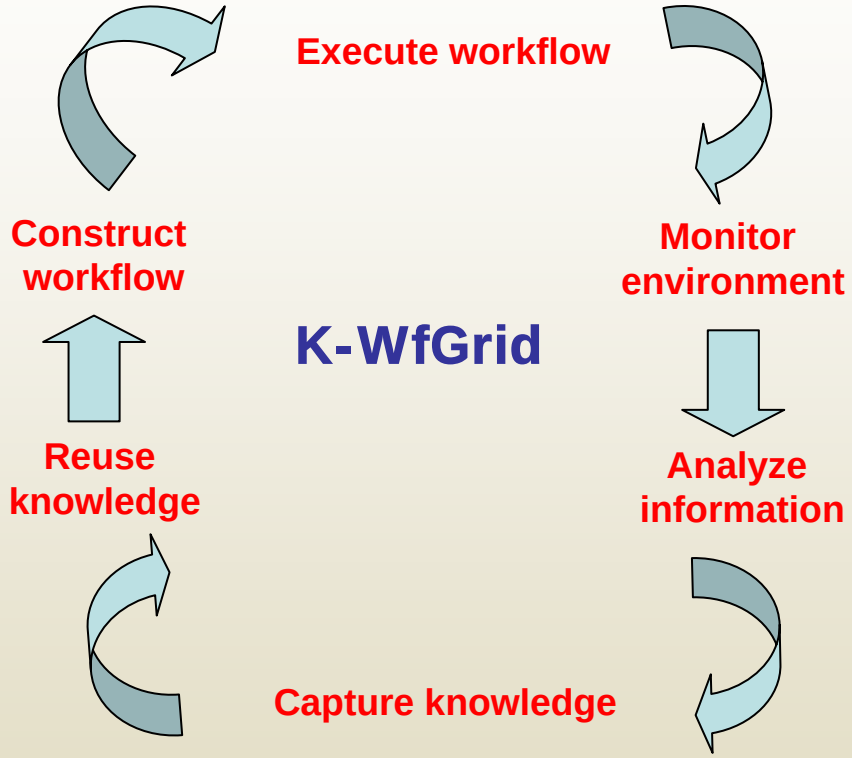


EGEE activities in CE ROC

- Coordination of Grid Operations within the region
  - CYFRONET
- Middleware deployment
  - all sites coordinated by CYFRONET
- Regional certification of middleware releases and integration of local additions (OCM-G, glogin)
  - CYFRONET, CESNET, GUP-JKU
- Monitoring and management of operational problems
  - PSNC
- User Support
  - ICM
- Pre-production service (testing gLite version of middleware)
  - CYFRONET, CESNET
- Running essential grid services
  - CYFRONET, CESNET (VOCE VO), RMKI-KFKI

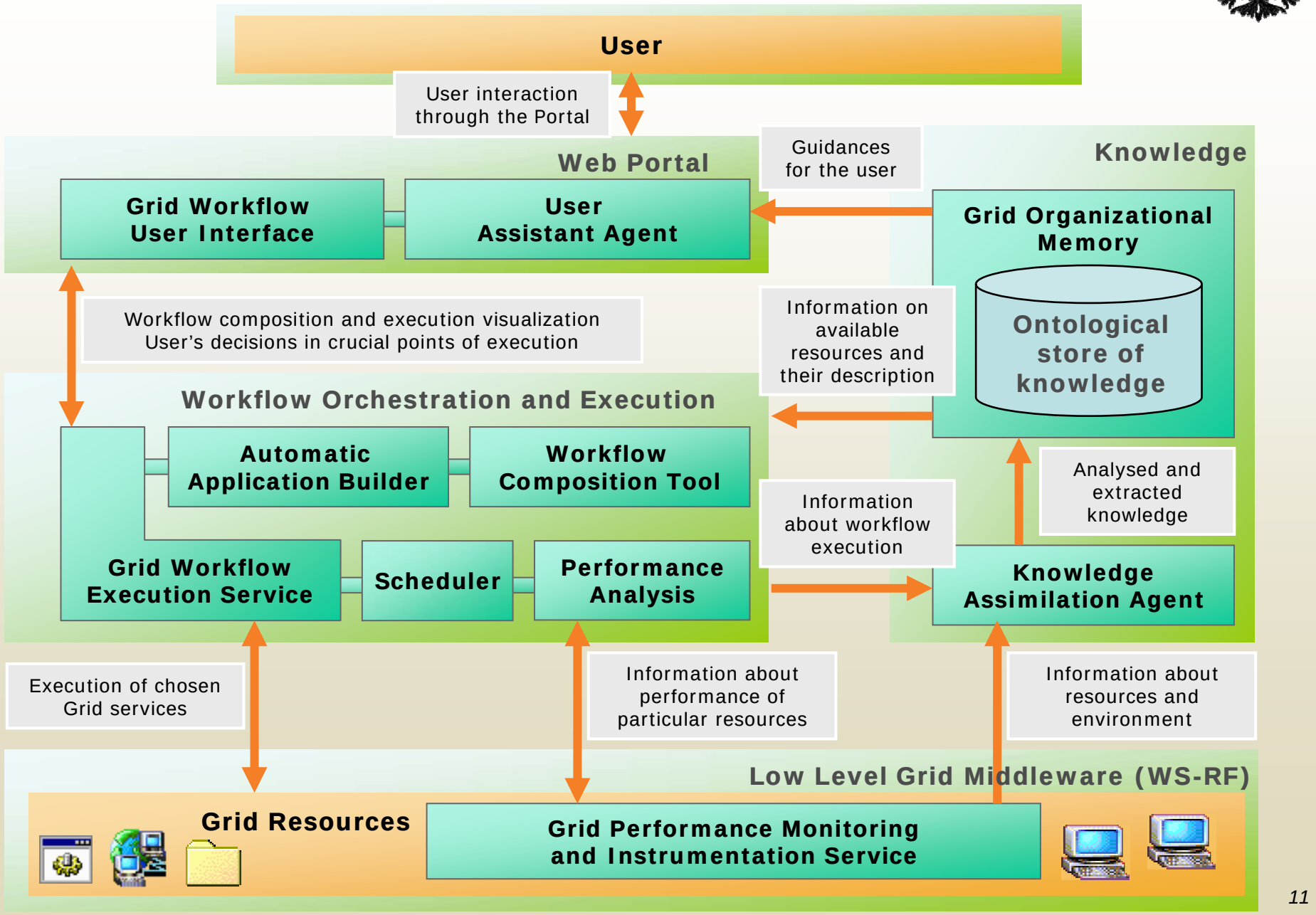
K-WfGrid–Workflow Applications and Knowledge

- Integrating services into coherent application scenarios
- Enabling automatic construction and reuse of workflows with knowledge gathered during operation
- Involving monitoring and knowledge acquisition services in order to provide added value for end users



Technologies: service-oriented Grid architecture, software agents, ontologies, dynamic instrumentation.

K-WfGrid - Architecture and Flow of Actions



K-WfGrid - Partners

- Fraunhofer FIRST
  Berlin, Germany

- Institute of Computer Science,
  University of Innsbruck
  Innsbruck, Austria

- Institute of Informatics of the Slovak
  Academy of Sciences
  Bratislava, Slovakia

- ACC CYFRONET AGH
  Kraków, Poland

- LogicDIS S.A.
  Athens, Greece

- Softeco Sismat SpA
  Genova, Italy

www.kwfgrid.net

In CoreGRID – Support for Grid PSE

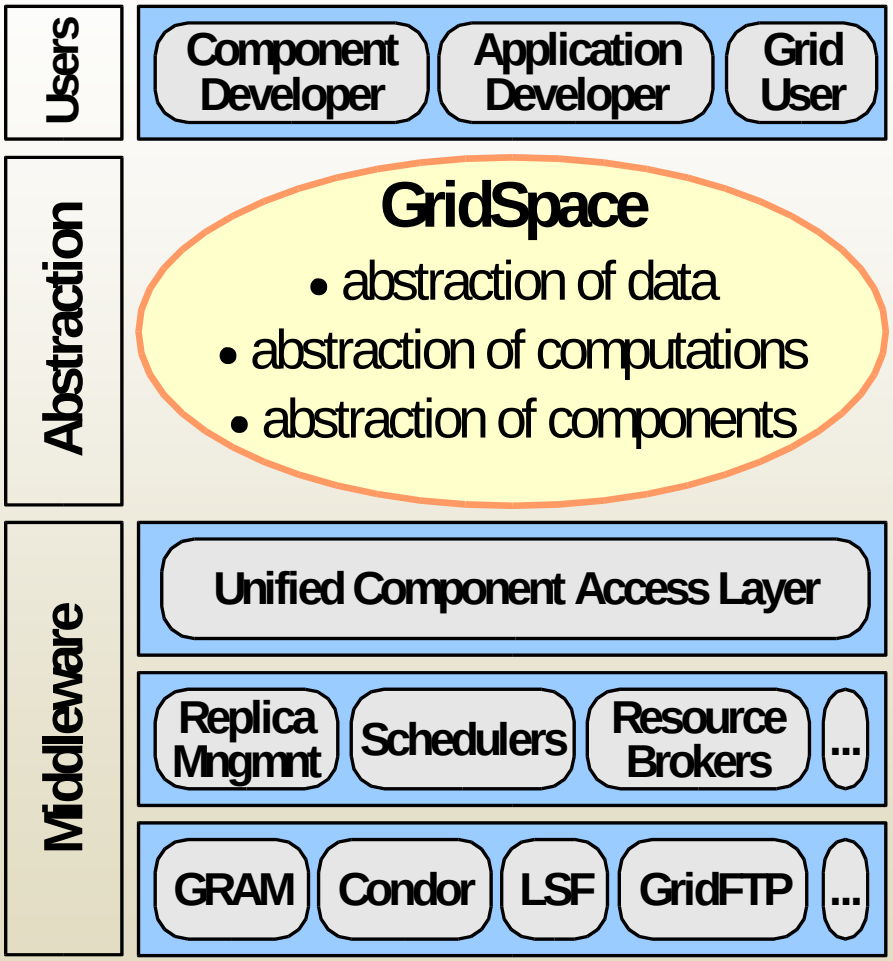
- **High Level Programming Language**
  - Interpreted on Grid as a runtime system
  - Routine call equals Grid component execution
  - Dynamic binding with implicit syntax and semantics matching
  - Allows programming of various types of computation directly on Grid

- **Common Grid Component Model**
  - Syntax and meaning description of functionality
  - Explicit model of parameter-to-output relationship
  - Functional and non-functional properties described with ontologies
  - Data type system strictly based on XML Schema web standard

- **CCA / H2O Execution Framework**
  - Based on well-known CCA component standard
  - Support for multi-paradigm computations
  - Supports interactive components for user control
  - Provides dynamic self-reconfiguration of components during runtime

Programming Grid Applications

- Separates the developer from ever-changing Grid resource layer
- Seamlessly introduces dynamism into newly created applications
- Provides unified access to resources by means of semantically described abstractions
- Supports evolving and well organized library of applications used up-to-date
- Allows easy reuse of already built applications



MOCCA Component Framework

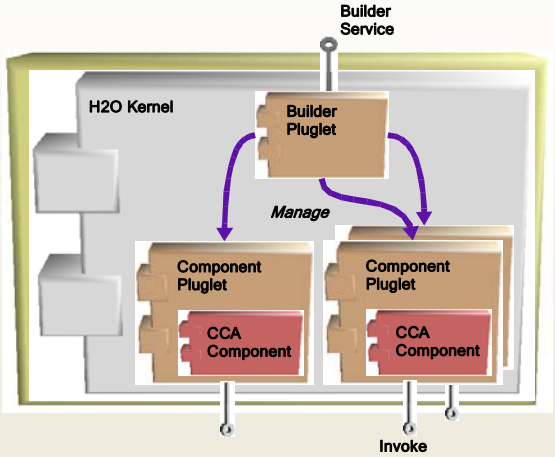
Goals

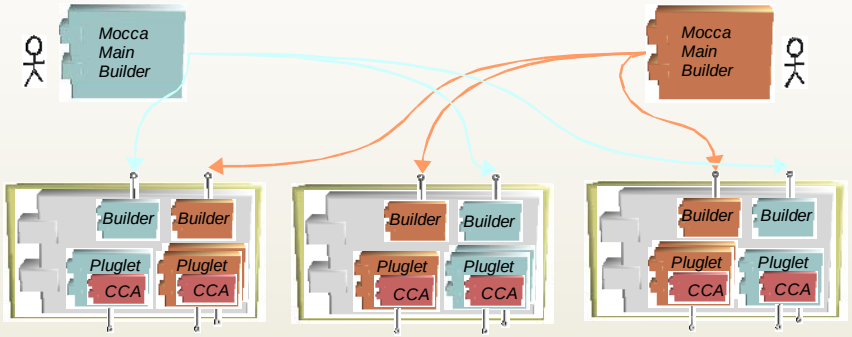
- Provide easy mechanisms for creation of components on distributed shared resources;
- Provide efficient communication mechanism both for distributed and local components;
- Allow flexible configuration of components and various application scenarios;
- Support native components, i.e. components written in non-Java programming languages and compiled for specific architecture.

Solutions

- Distributed CCA-compliant framework
- Based on H2O metacomputing platform
- Uses RMIX for efficient communication
- Java and Jython scripting interface for assembling applications
- Work in progress to support native components using Babel

MOCCA implementation





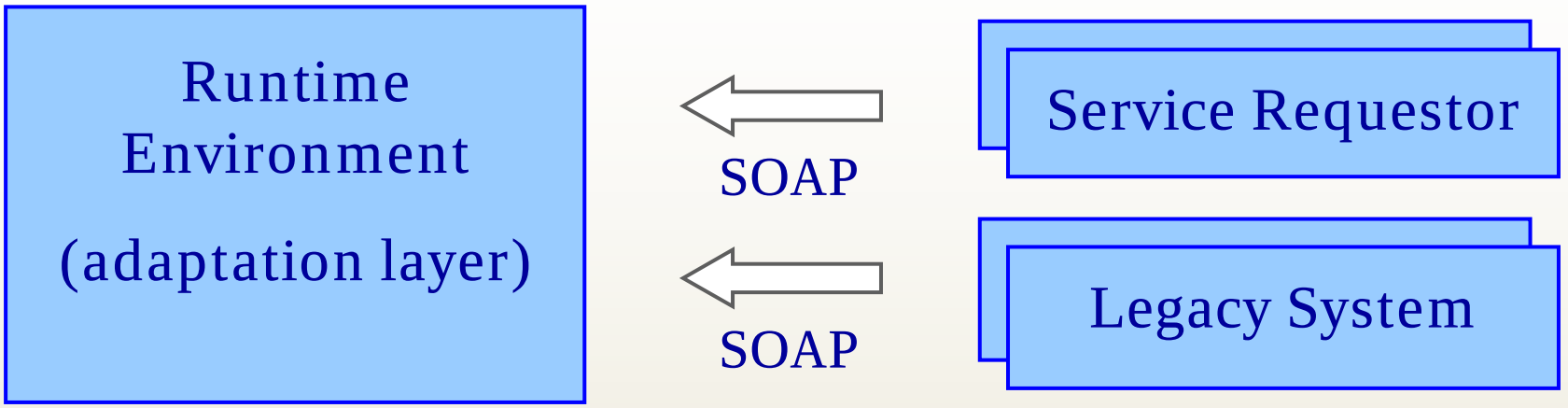
- CCA components instantiated as H2O pluglets
- Each user can create own arena where components are deployed
- Thanks to H2O kernel security mechanisms, multiple components may run without interfering

Legacy Software to Grid Services

- Legacy software
  - Validated and optimized code / binaries
  - Follows traditional process based model of computation (language & system dependent)
  - Scientific libraries (e.g. BLAS, LINPACK)

- Service oriented architecture (SOA)
  - Enhanced interoperability
  - Language independent interface (WSDL)
  - Execution within system neutral runtime environment (virtual machine)

Legacy Software to Grid Services



- Universal architecture enabling to integrate legacy software into service-oriented architecture
- Novel design enabling efficiency, security, scalability, fault-tolerance, versatility
- Current implementation: LGF framework which automates the process of migration of C/C++ codes to GT 3.2
- Further work: WSRF, message level security, optimizations, early process migration

Ambient Networks

- IP, 45 partners from all over Europe,
- Started in 2004, it has 3 2-years phases
- The Project aims at an innovative, industrially exploitable new network vision based on the dynamic composition of networks
- It uses the **Overlay Networks** to provide access to any network, including mobile personal networks, through instant establishment of inter-network agreements
- Key element of A.N. architecture is the **Ambient Control Space** (**ACS**) as an environment within which a set of modular control functions can co-exist and cooperate. These control functions include **SATO** modules (**Service Aware Transport Overlays**), **Network Context** management, and others.
- **Network Composition** is the core approach to achieve the dynamic integration of control functionality – the ACS – across a heterogeneous set of networks.
- Integrates 3G networks and uses **IMS** (**IP Multimedia Subsystem**)

Ambient Network Abstractions

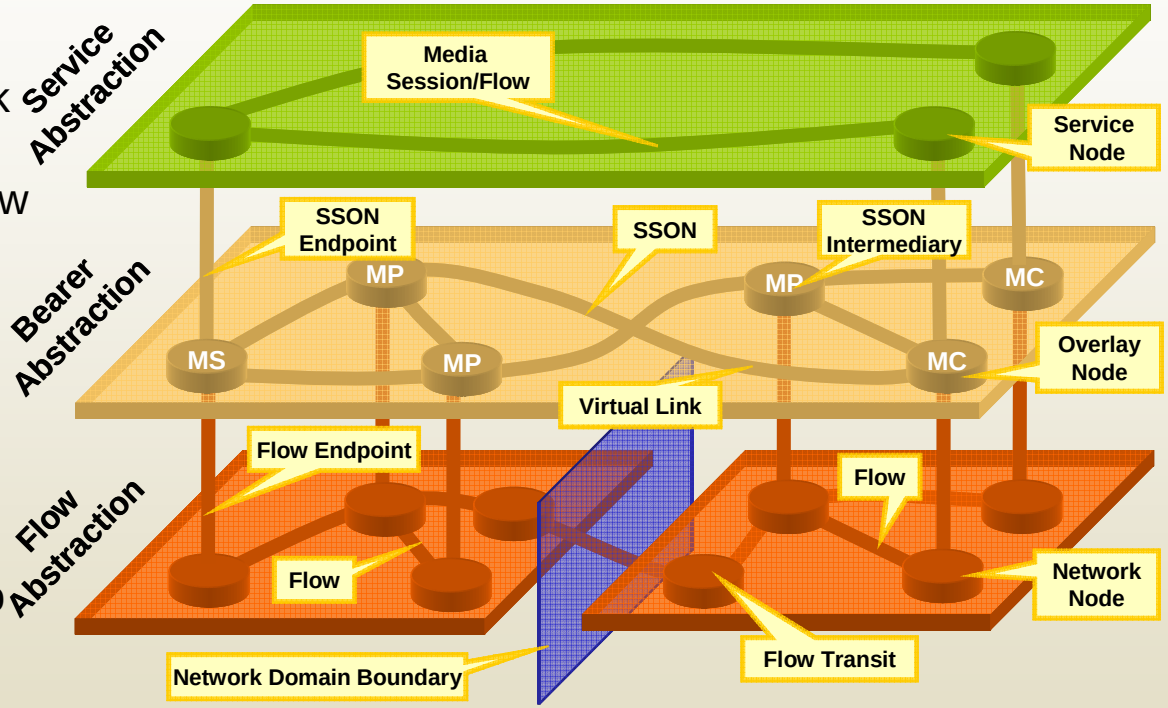
Overlay Networks introduced to improve network reliability and to derive a network structure enabling new routing paradigms.

SSON – Service Specific Overlay Network

Bearer – abstract flow in overlay network

Service Abstraction – end to end connection supporting particular service

MS/MP – Media Server, Media Port (content and processing)



http://www.ambient-networks.org/

ViroLab - Project Objectives

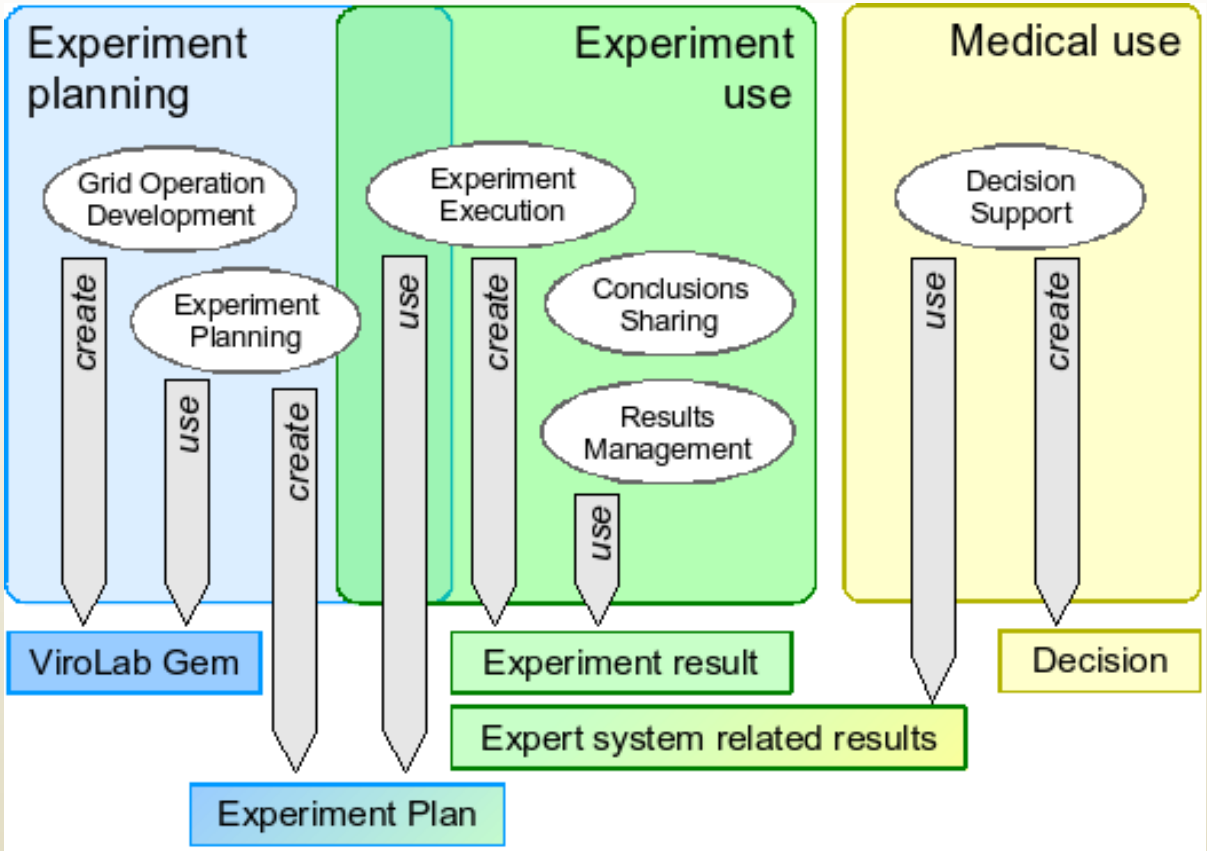
The long term mission of ***ViroLab*** is to provide researchers and medical doctors in Europe with a virtual laboratory for infectious diseases.

1. To develop a VO that provides the "glue" for binding the various components of our virtual laboratory
2. To develop a virtual laboratory infrastructure for transparent workflow, data access, experimental execution and collaboration
3. To virtualize and enhance state-of-the-art in genotypic resistance interpretation tools and integrate them into the virtual laboratory
4. To establish epidemiological validation to correctly and quantitatively predict virological and immunological outcomes and disseminate the results to European stakeholders

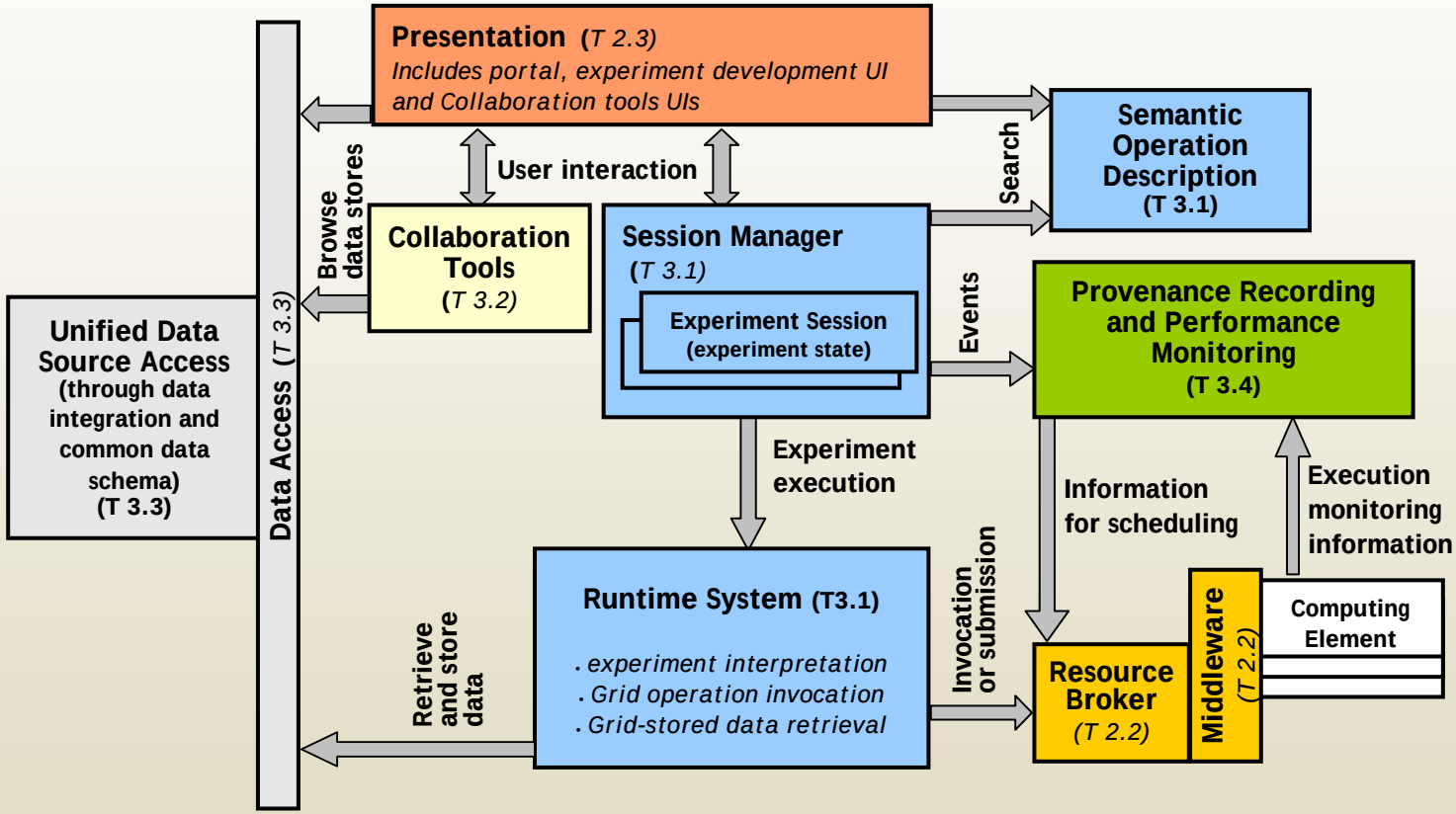
ViroLab Users

- Experiment developer
  - Plans a ViroLab experiment using the VLvl development tools
  - Knows both how to script an experiment and the modelled domain
  - May prepare dedicated UI for the experiment user
- Scientist - experiment user
  - Uses prepared experiment to gather scientific results
  - The results may be shared through collaboration tools with others
  - The results provenance could be tracked and recorded
  - The results may be stored in ViroLab data store
- User of the ViroLab decision support system
  - Uses dedicated web GUI (only web browser required)
  - The system seamlessly uses some ViroLab applications (BAC, Rule Miner) to provide better support over time

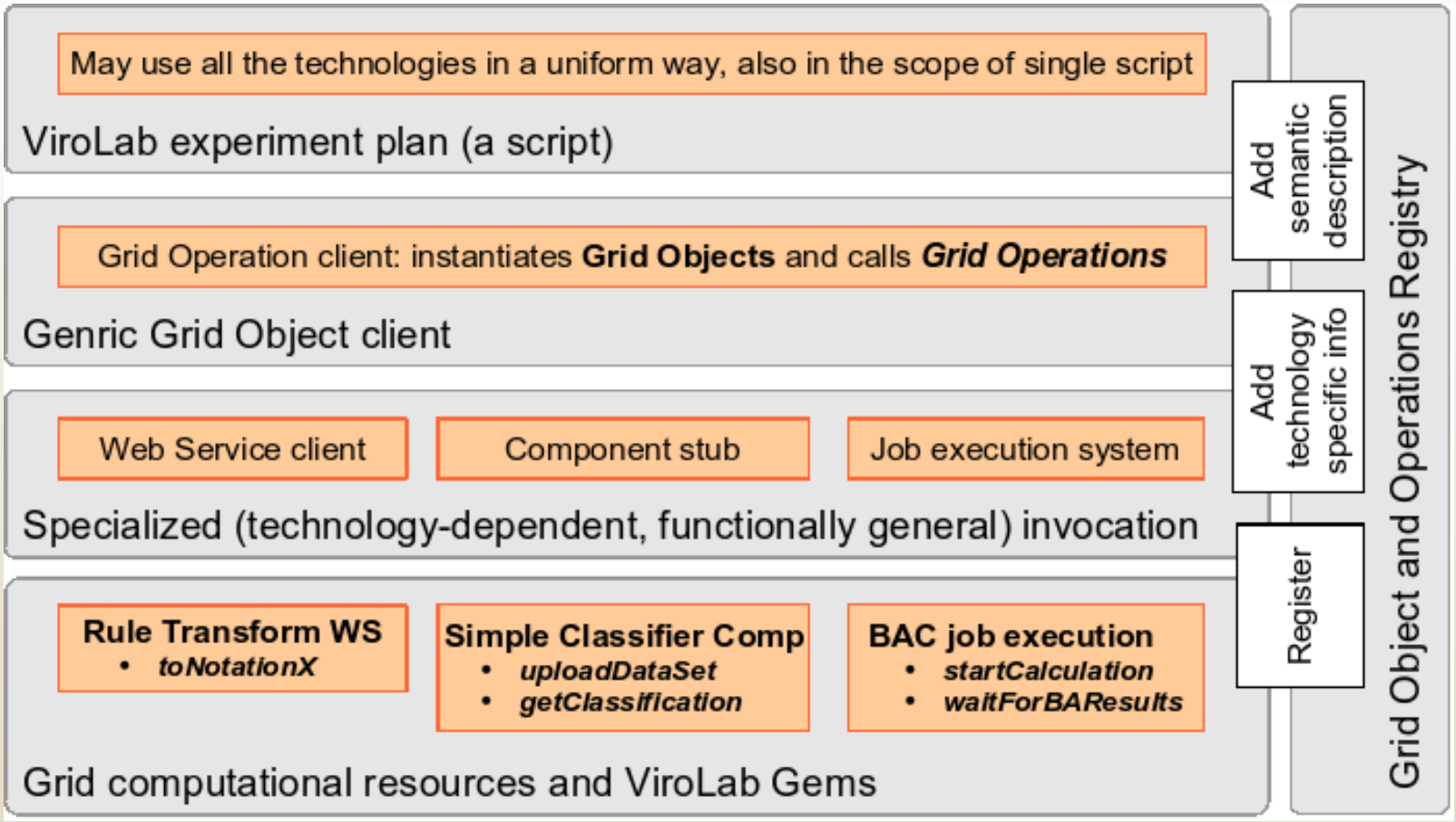
ViroLab Use Cases



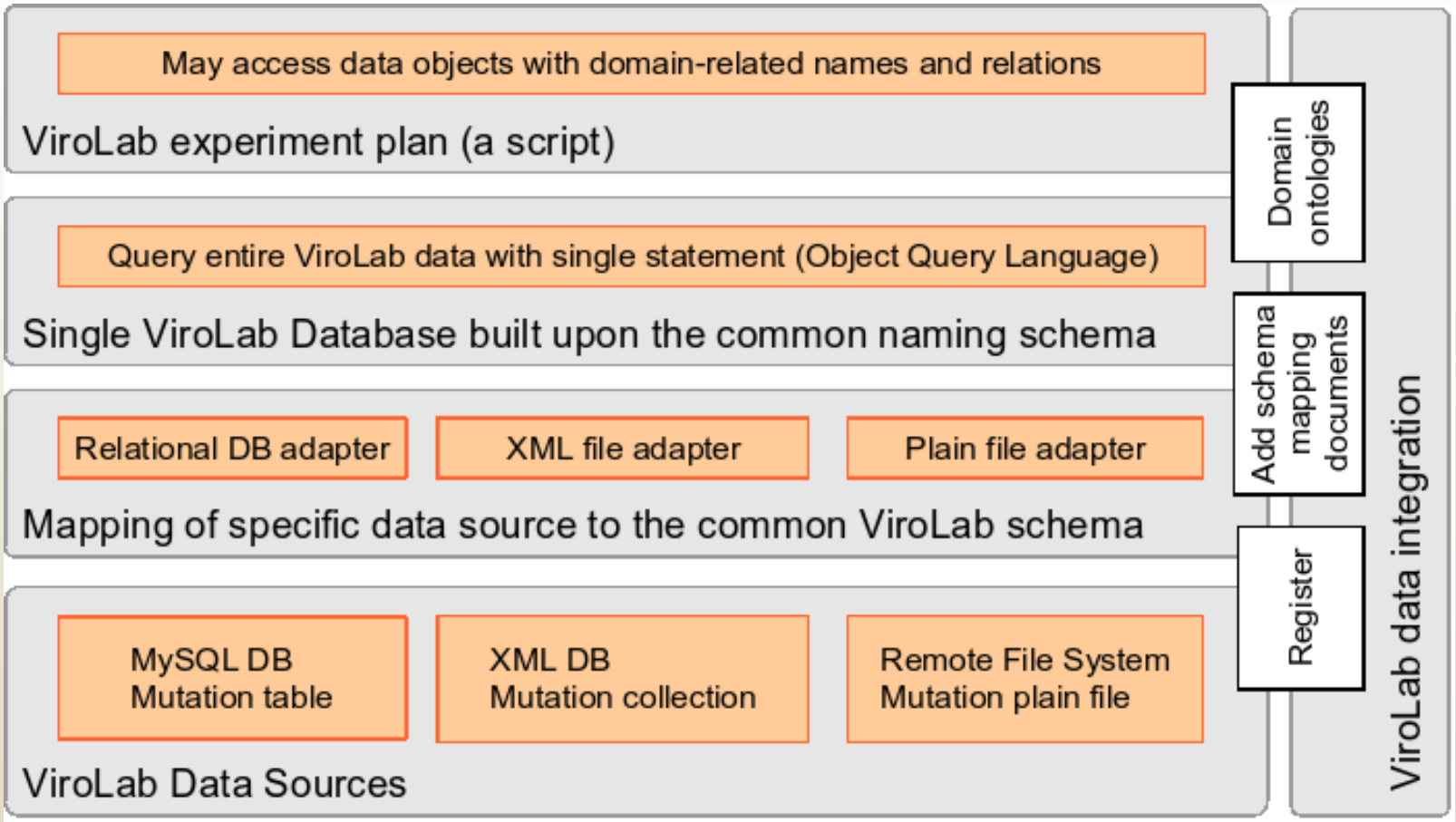
Architecture of VL



Access to Computation (Middleware)



Access to Data (Data Virtualization)



ViroLab - Partners

- Universiteit van Amsterdam
- University Medical Centre Utrecht
- High Performance Computing Center Stuttgart
- CYFRONET AGH
- Gridwise Technologies
- Institute de Recerca de la SIDA
- Catholic University Leuven
- University College London
- Catholic University Rome
- Eötvös Loránd University
- Institute of Infectious and Tropical Diseases, University of Brescia
- Virology Education B.V.

GRID projects at ICM (current)

Current European projects (EU IST):

- Centre of Excelence for Multiscale Biomolecular Modelling, Bioinformatics and Applications (2002 – 2005)
  - grid workpackage – deployment of grid infrastructure for life sciences
- UNIGRIDS (2004 – 2006)
  - ICM develops high level services (visualization, database access, access to remote instruments)
  - ICM deploys applications for UNICORE/GS
- EGEE (2004-2006)
  - ICM operates HPC resources
- ATVN (2004-2006)
  - ICM operates HPC resources

New project coordinated by ICM (EU IST call 5):

- Chemomentum

R&D CENTER IN GRIDS

- GRIDLAB

Development of the Grid Application Toolkit and middleware secure and efficient tools for the grid applications tested in the transatlantic testbed

- GRIDSTART
- CROSSGRID
- ENACTS
- CoreGrid
- ACGT
- GridCoord
- EGEE

Enabling Grids for E-science in Europe

- HPC-Europa

European grid computing infrastructure, data management, monitoring and efficiency analysis, portal access to the grid resources

- InteliGrid

Improvement of the project wotk for the aircraft industry, automotive indystry and construction engineering computing

- CLUSTERIX
- SGIgrid
- VLab
- PROGRESS

Summary

- Large scale numerical simulations
- Computationally demanding data analysis
- Distributed computing and storage
- Remote access to experimental equipments
- A need for integration heterogeneous environments into one application
- Collaborative problem solving
- Virtual organisations

www.cyfronet.krakow.pl