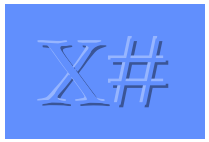
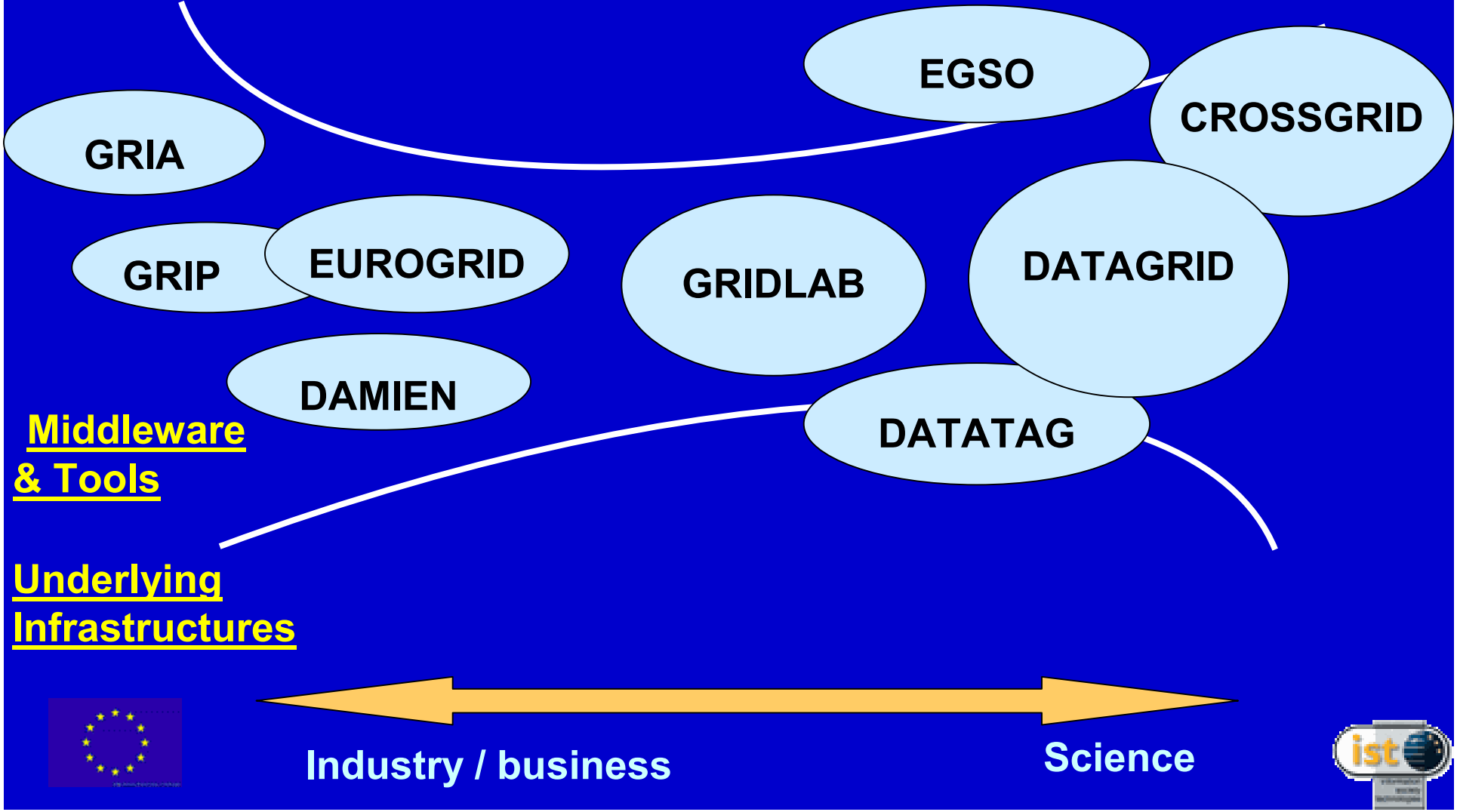


Towards the CrossGrid Architecture

Marian Bubak, Marek Garbacz, Maciej Malawski, and Katarzyna Zajac
X# TAT
Institute of Computer Science & ACC CYFRONET
AGH, Kraków, Poland

A new IST Grid project space (Kyriakos Baxevanidis)

- *Links with European National efforts*
- *Links with US projects (GriPhyN, PPDG, iVDGL,…)*



CrossGrid Collaboration

Main Objectives

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

Key Features of X# Applications

- **Data**
  - Data generators and data bases geographically distributed
  - Selected on demand
- **Processing**
  - Needs large processing capacity; both HPC & HTC
  - Interactive
- **Presentation**
  - Complex data require versatile 3D visualisation
  - Support interaction and feedback to other components

WP1 – CrossGrid Application Development

Tasks

1.0 Co-ordination and management (Peter M.A. Sloot, UvA)

1.1 Interactive simulation and visualisation of a biomedical system (G. Dick van Albada, Uva)

1.2 Flooding crisis team support (Ladislav Hluchy, II SAS)

1.3 Distributed data analysis in HEP (C. Martinez-Rivero, CSIC)

1.4 Weather forecast and air pollution modelling (Bogumil Jakubiak, ICM)

WP2 - Grid Application Programming Environments

Tasks

2.0 Co-ordination and management (Holger Marten, FZK)

2.1 Tools requirement definition (Roland Wismueller, TUM)

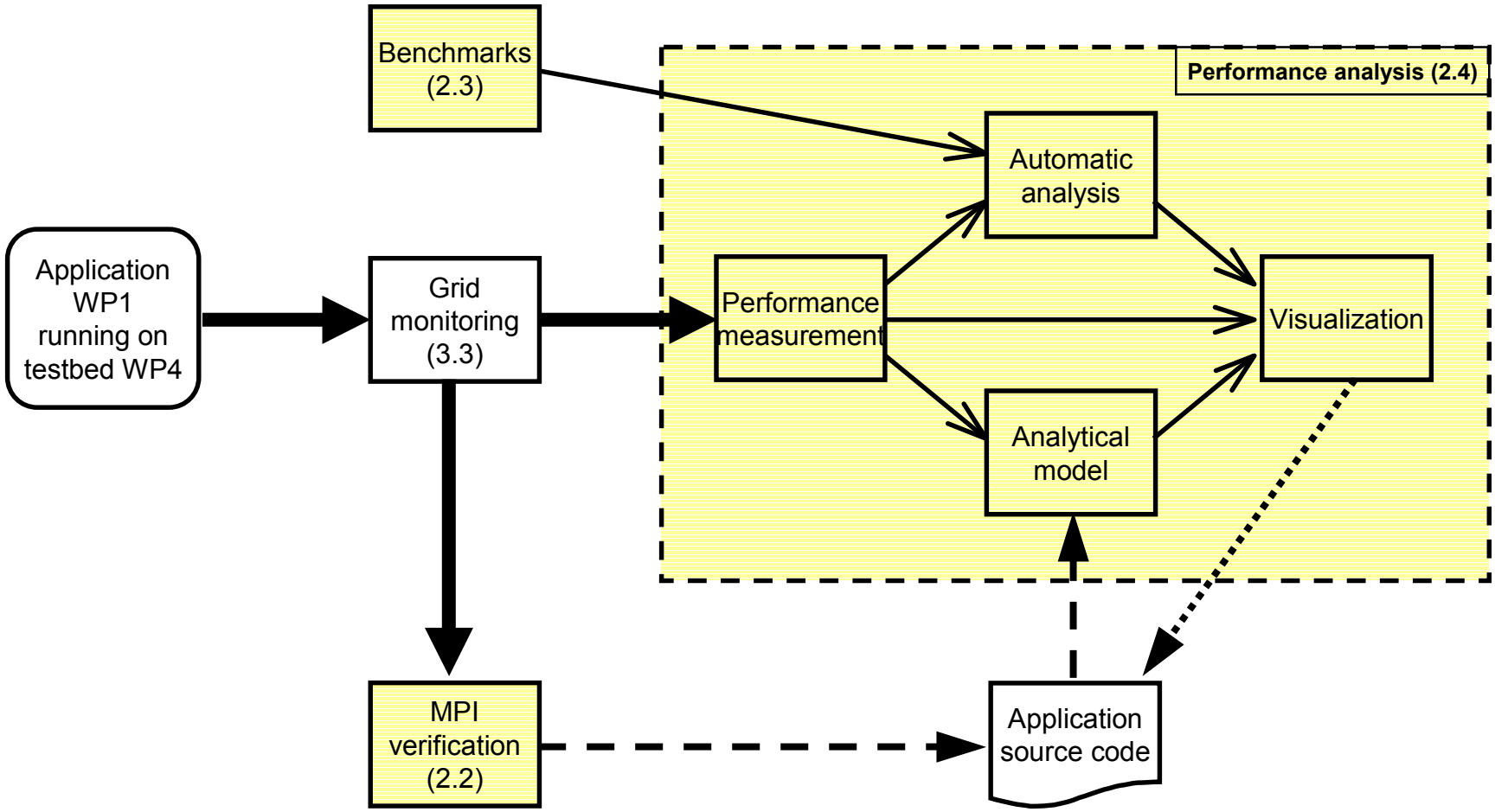
2.2 MPI code debugging and verification (Matthias Mueller, USTUTT)

2.3 Metrics and benchmarks (Marios Dikaiakos, UCY)

2.4 Interactive and semiautomatic performance evaluation tools (Wlodek Funika, Cyfronet)

2.5 Integration, testing and refinement (Roland Wismueller, TUM)

WP2 - Components and relations to other WPs



WP3 – New Grid Services and Tools

Tasks

3.0 Co-ordination and management (Norbert Meyer, PSNC)

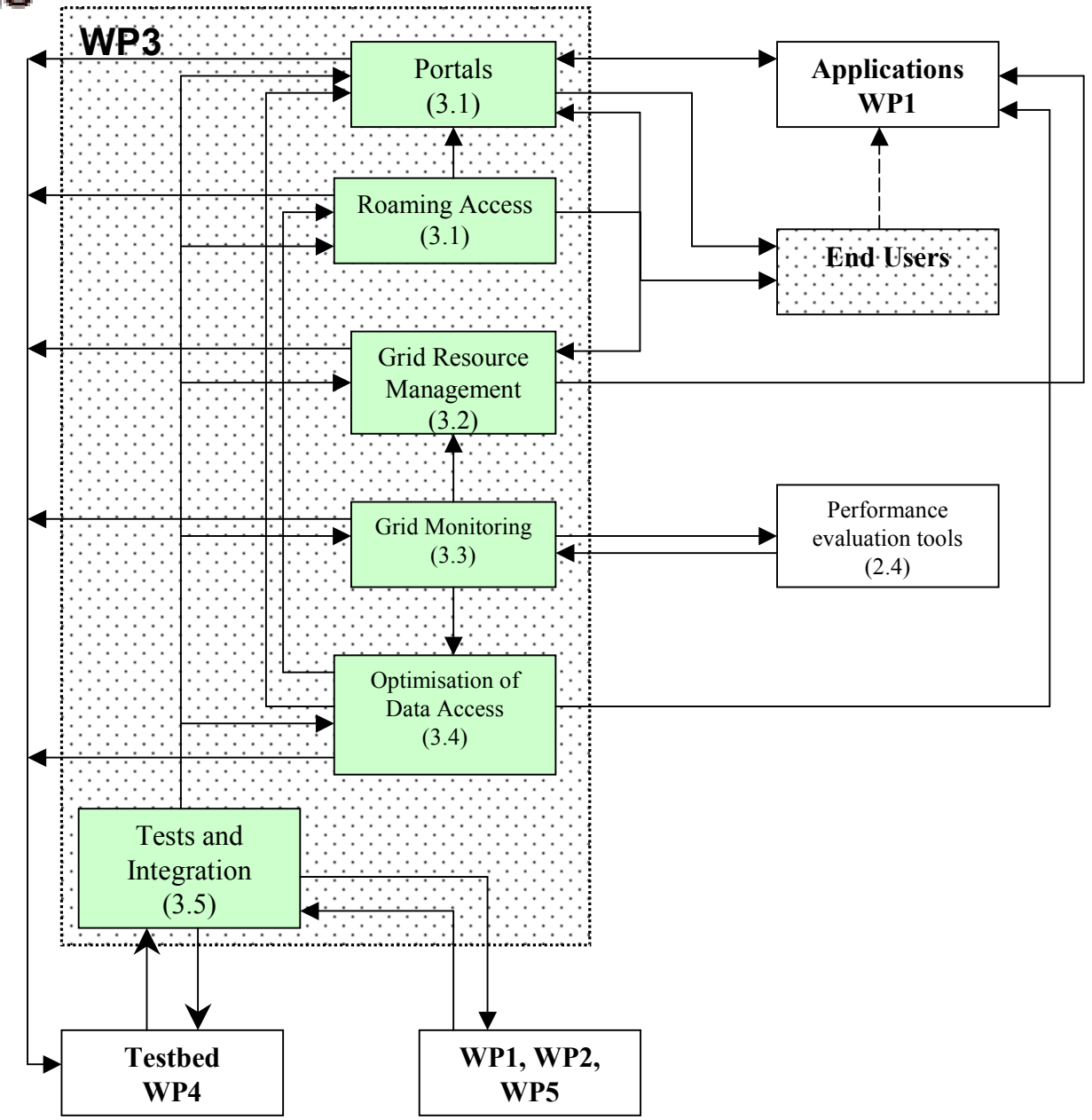
3.1 Portals and roaming access (Miroslaw Kupczyk, PSNC)

3.2 Grid resource management (Miquel A. Senar, UAB)

3.3 Grid monitoring (Brian Coghlan, TCD)

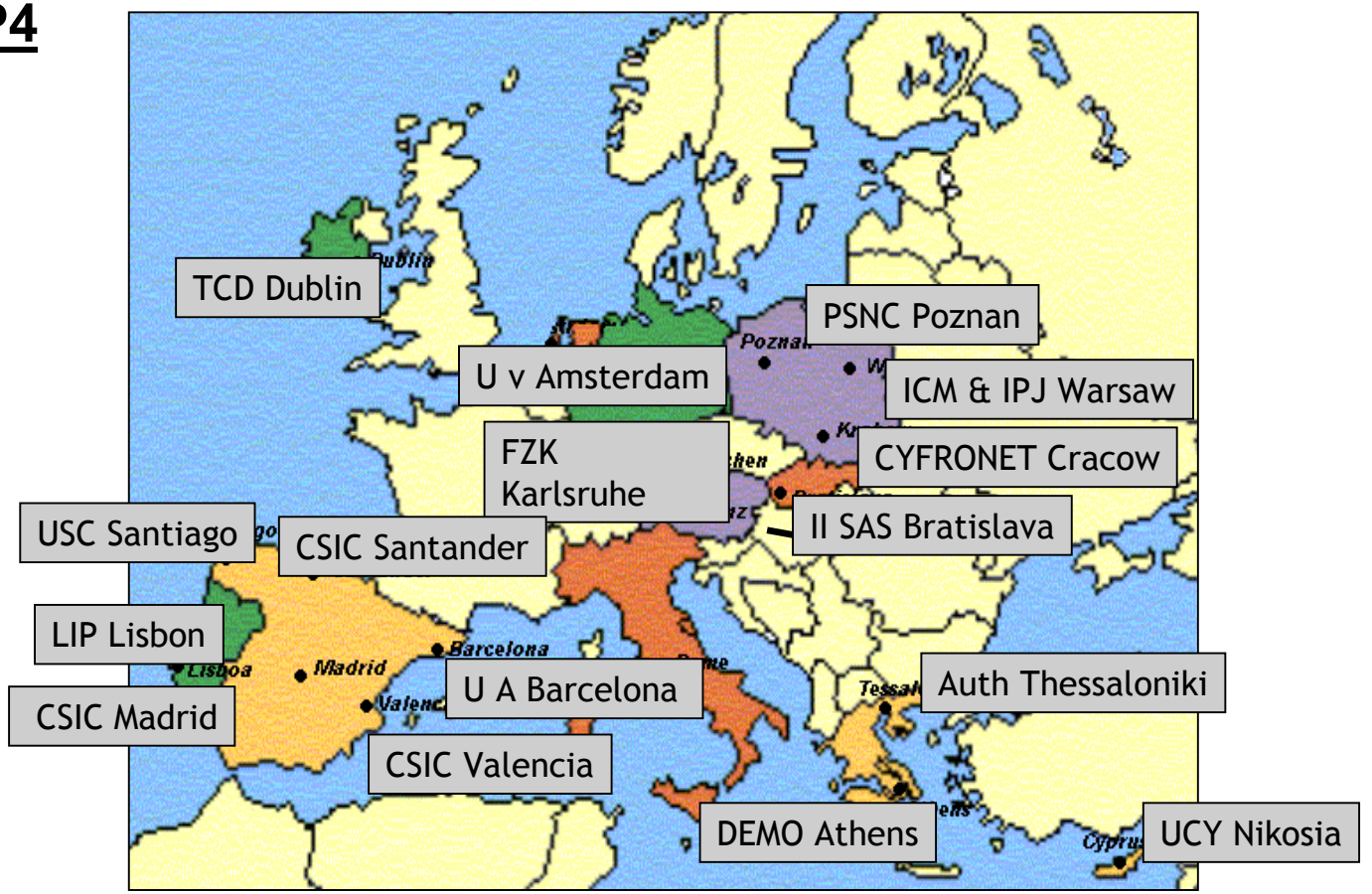
3.4 Optimisation of data access (Jacek Kitowski, Cyfronet)

3.5 Tests and integration (Santiago Gonzalez, CSIC)


WP3
Portals
(3.1)
Roaming Access
(3.1)
Grid Resource
Management
(3.2)
Grid Monitoring
(3.3)
Optimisation of
Data Access
(3.4)
Tests and
Integration
(3.5)
Applications
WP1
End Users
Performance
evaluation tools
(2.4)
Testbed
WP4
WP1, WP2,
WP5

WP4 - International Testbed Organization

Partners in WP4



WP4 lead by

CSIC (Spain)

WP4 - International Testbed Organization

Tasks

4.0 Coordination and management

(Jesus Marco, CSIC, Santander)

–Coordination with WP1,2,3
–Collaborative tools
–Integration Team

4.1 Testbed setup & incremental evolution

(Rafael Marco, CSIC, Santander)

–Define installation
–Deploy testbed releases
–Trace security issues

Testbed site responsibles:

- CYFRONET (Krakow) A.Ozieblo
- ICM(Warsaw) W.Wislicki
- IPJ (Warsaw) K.Nawrocki
- UvA (Amsterdam) D.van Albada
- FZK (Karlsruhe) M.Kunze
- IISAS (Bratislava) J.Astalos
- PSNC(Poznan) P.Wolniewicz
- UCY (Cyprus) M.Dikaiakos
- TCD (Dublin) B.Coghlan
- CSIC (Santander/Valencia) S.Gonzalez
- UAB (Barcelona) G.Merino
- USC (Santiago) A.Gomez
- UAM (Madrid) J.del Peso
- Demo (Athenas) C.Markou
- AuTh (Thessaloniki) D.Sampsonidis
- LIP (Lisbon) J.Martins

WP4 - International Testbed Organization

Tasks

4.2 Integration with DataGrid (Marcel Kunze, FZK)

- –Coordination of testbed setup
- –Exchange knowledge
- –Participate in WP meetings

4.3 Infrastructure support (Josep Salt, CSIC, Valencia)

- –Fabric management
- –HelpDesk
- –Provide Installation Kit
- –Network support

4.4 Verification & quality control (Jorge Gomes, LIP)

- –Feedback
- –Improve stability of the testbed

WP5 – Project Management

Tasks

5.1 Project coordination and administration (Michal Turala, INP)

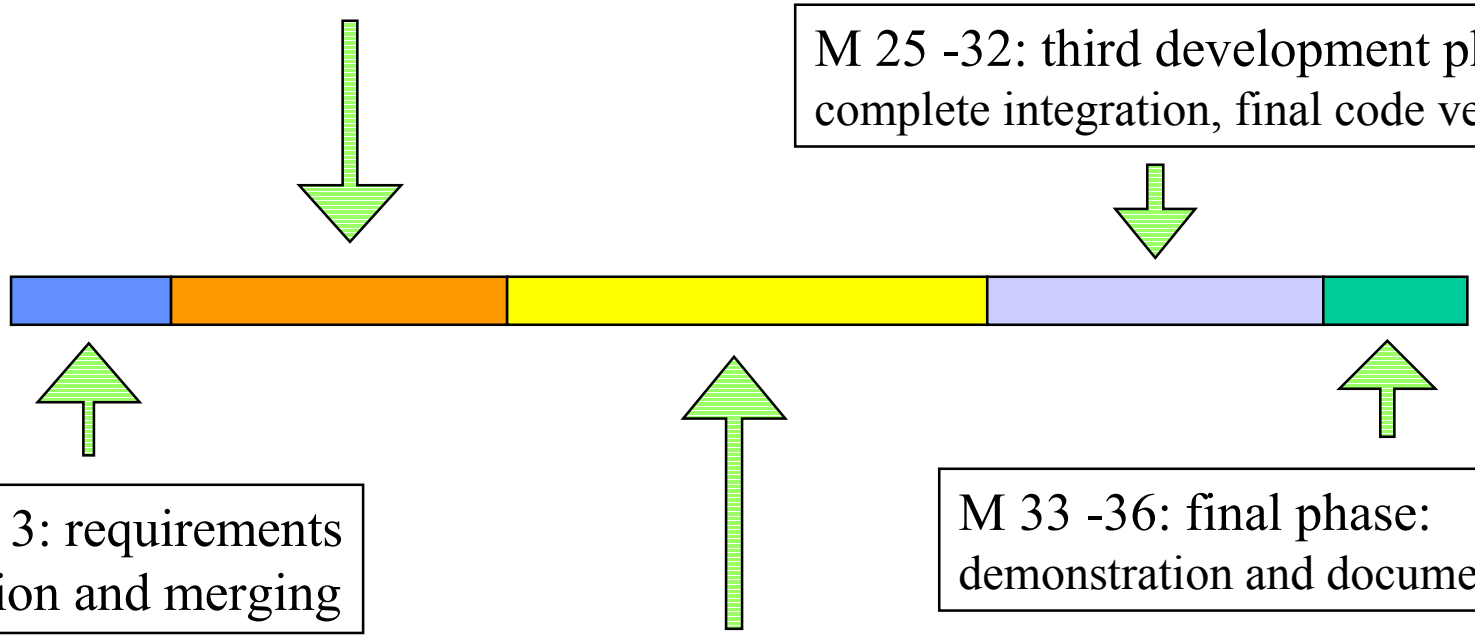
5.2 CrossGrid Architecture Team (Marian Bubak, Cyfronet)

5.3 Central dissemination (Yannis Perros, ALGO)

Project Phases

M 4 - 12: first development phase: design, 1st prototypes, refinement of requirements

M 25 -32: third development phase: complete integration, final code versions



M 1 - 3: requirements definition and merging

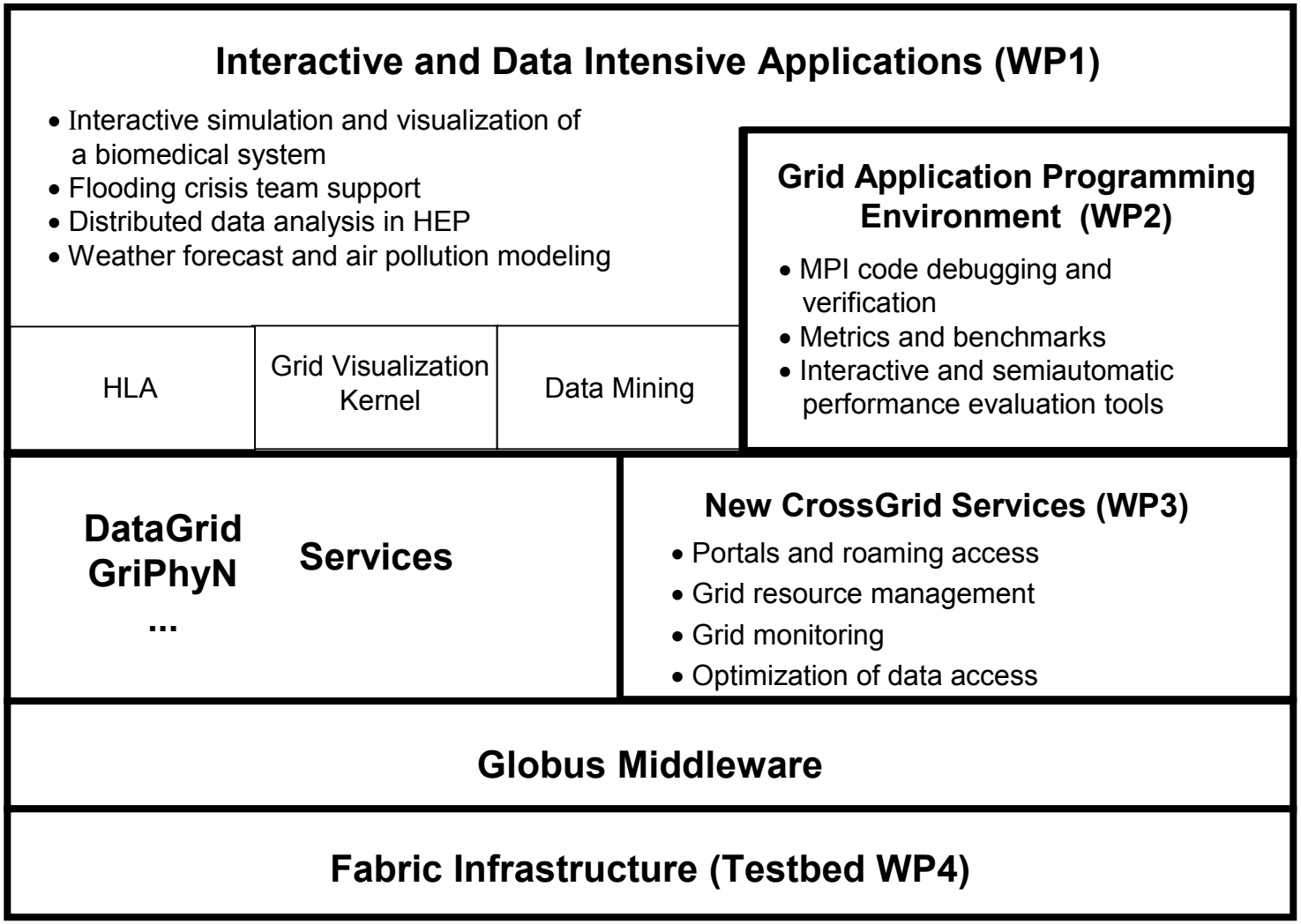
M 33 -36: final phase: demonstration and documentation

M 13 -24: second development phase: integration of components, 2nd prototypes

Person-months

WP	WP Title	PM Funded	PM Total
WP1	CrossGrid Applications Development	365	537
WP2	Grid Application Programming Environment	156	233
WP3	New Grid Services and Tools	258	421
WP4	International Testbed Organization	435	567
WP5	Project Management	102	168
	Total	1316	1926

Layered Structure of X#



Two important questions

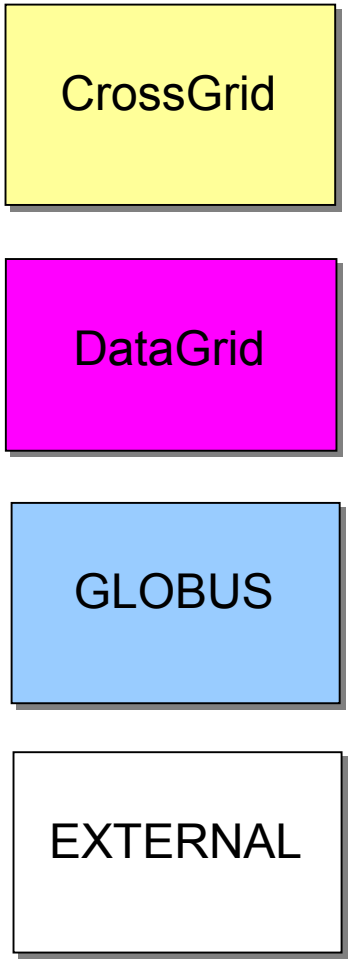
- How to build interactive Grid environment ? (Globus is more batch-oriented than interactive-oriented; performance issue)
- How to use with Globus and DataGrid SW, how to define interfaces ?

Layer Approach

According to the Global Grid Forum Grid Protocol Architecture Working Group – „computing/simulation grid”

- Applications and Supporting Tools
- Applications Development Support
- Grid Common Services
- Local Resources

Building Blocks



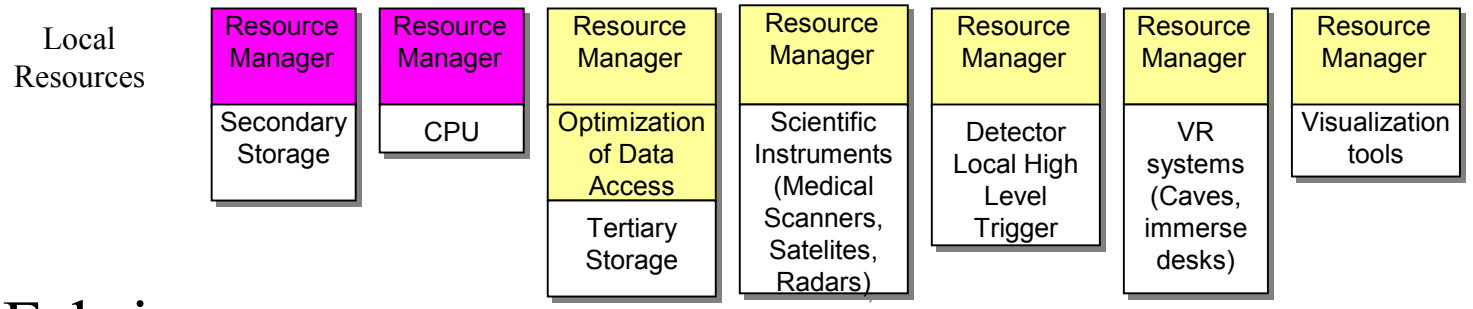
To be developed in X#

From DataGrid

Globus Toolkit

Other

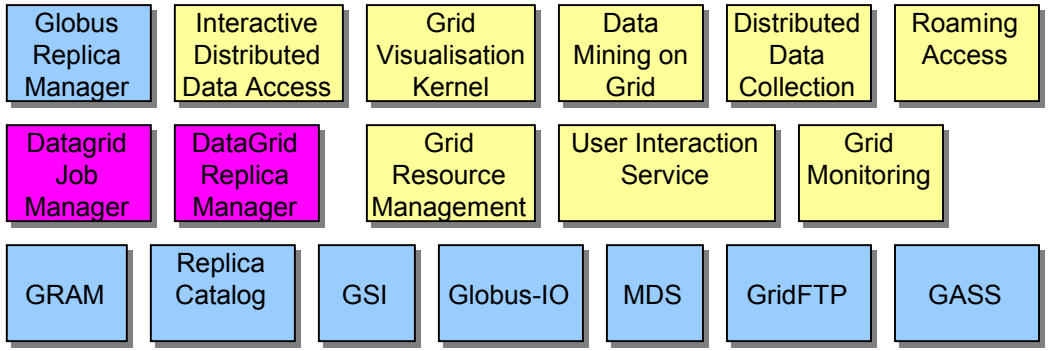
Local Resources



- Fabric
  - CPU,
  - Storage,
  - Instruments,
  - VR Systems
- Local Resource Managers – make them „gridable”

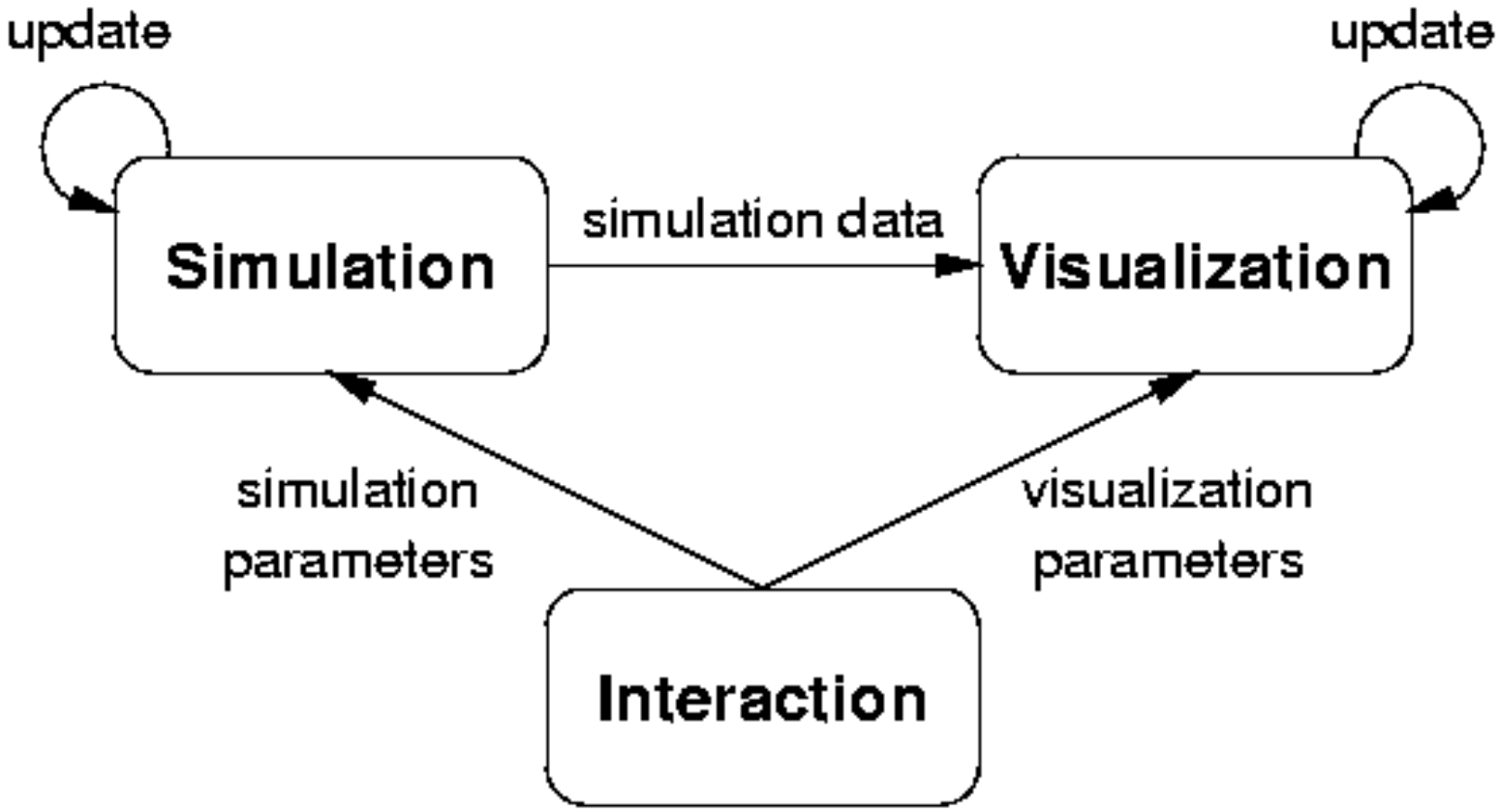
Grid Services

Grid Common Services

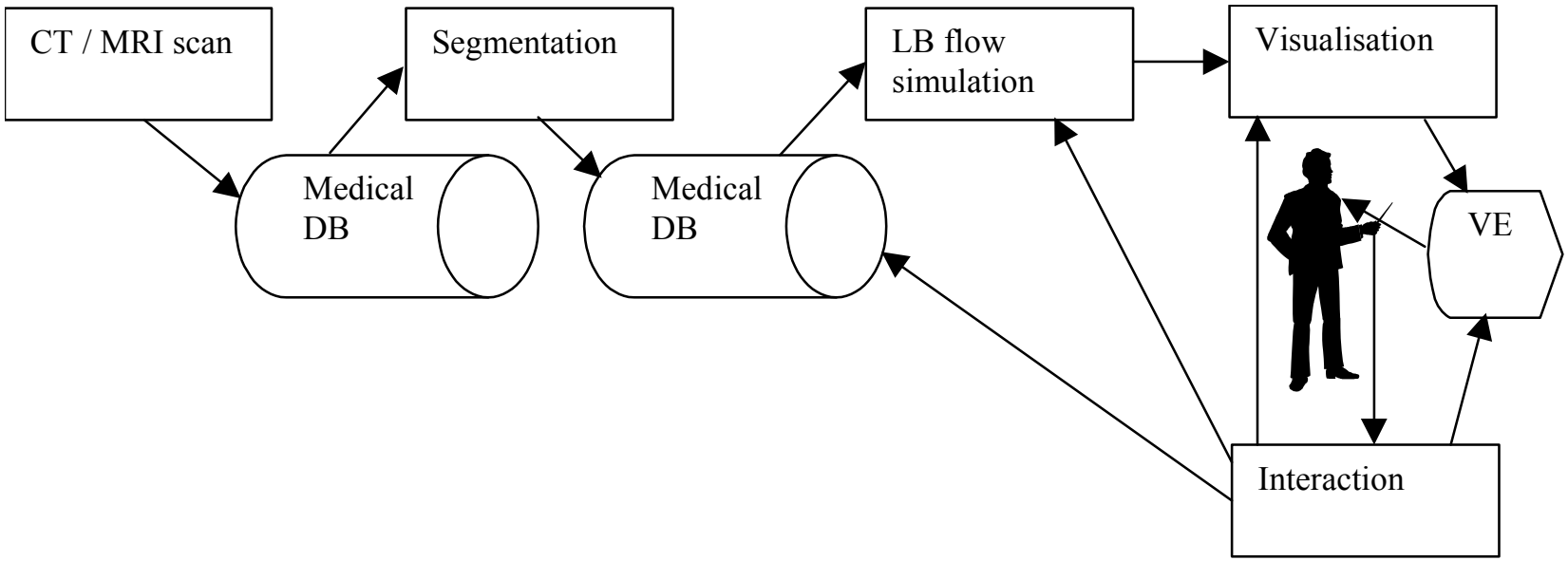


- Service = protocol + behavior (Foster: *Anatomy*...)
- Protocol – rules for exchanging information (interoperability)
- Behavior expected in response to protocol messages
- Service definition permits variety of implementations

Interaction in Biomedical Application



Biomedical Application Use Case



Step 1

- **Action**: An MRI scan ("Angiogram") is obtained for the patient.
- **Data:** 3D image 512 * 512 * 128 pixels.
- **Resource:** A 3D-visualisation system is reserved for use in step 3.

Step 2

- **Action:** The image is segmented so that a clear picture of the important blood vessels and the location of aneurisms and blockages is obtained.
- **Data:** 10% of the original image.

Step 3

- **Action:** Using the segmented image, a computational grid for a LB simulation is generated
- **Action:** A simulation of the normal pulsatile blood flow in the vessels is started.
- **Input from the physician:** parameters like pressure drop (possibly time dependent)
- **Run time:** several hours on a fast 16 node Beowulf cluster.

Step 4

- **Resource:** Interactive 3D-visualisation system
- **Interction:** the physician studies the vascular structure and proposes several (3 to 5) bypass designs. They are used to generate alternative computational grids.
- **Estimated duration:** order of 1 hour.

Step 5

- **Action:** The blood flow simulations for the bypasses are initialised using the new grids and the (partially) converged results from step 3.
- **Time:** several hours on a fast 16 node Beowulf cluster for each simulation

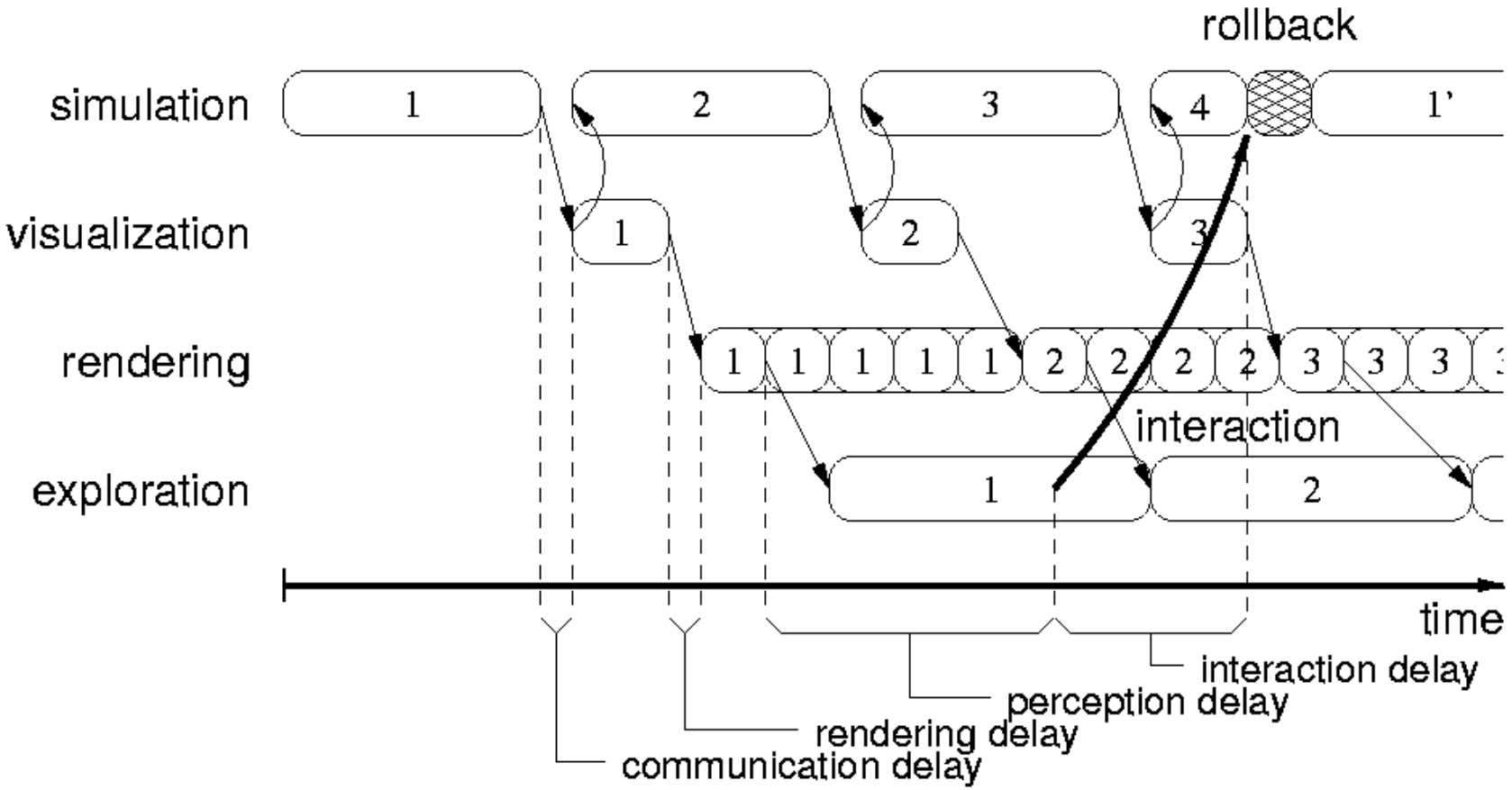
Step 6

- **Action:** The physician can monitor the progress of the simulations through his portal. He will be informed automatically, e.g. through an SMS message of their completion.
- **Resource:** The physician can use the advance information to reserve a 3D-visualisation environment for step 7.

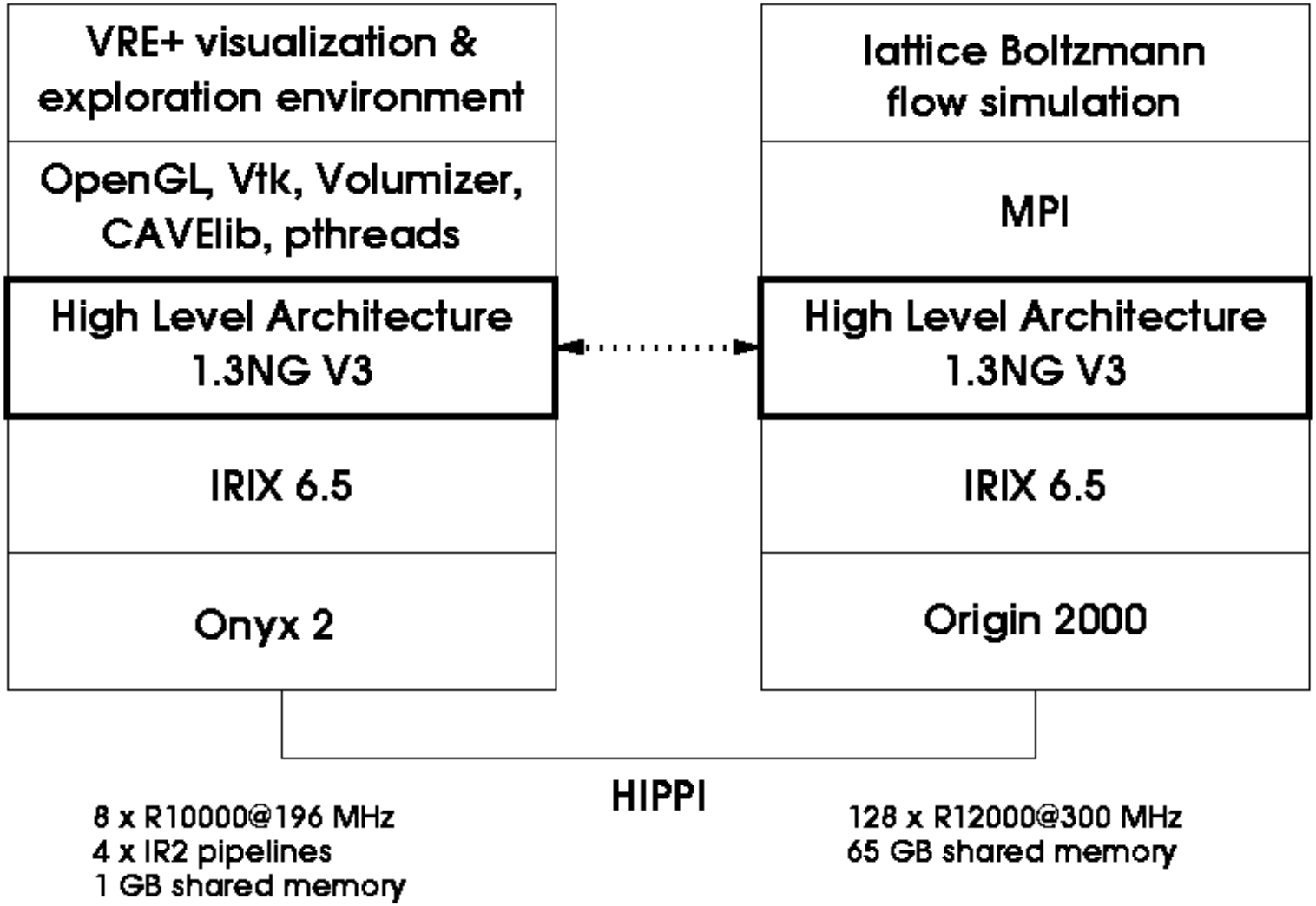
Step 7 – human in the loop

- **Action:** The results of the simulation are presented in the 3D-visualisation system
- **Input:** stored history or the running simulation
- **Interaction:** The physician can apply small modifications to the proposed bypass structure that should still allow a fast convergence of the blood-flow simulation
- **Time:** Simulations of the resulting changes in the blood flow should be initiated immediately, so that the results will be available within minutes.

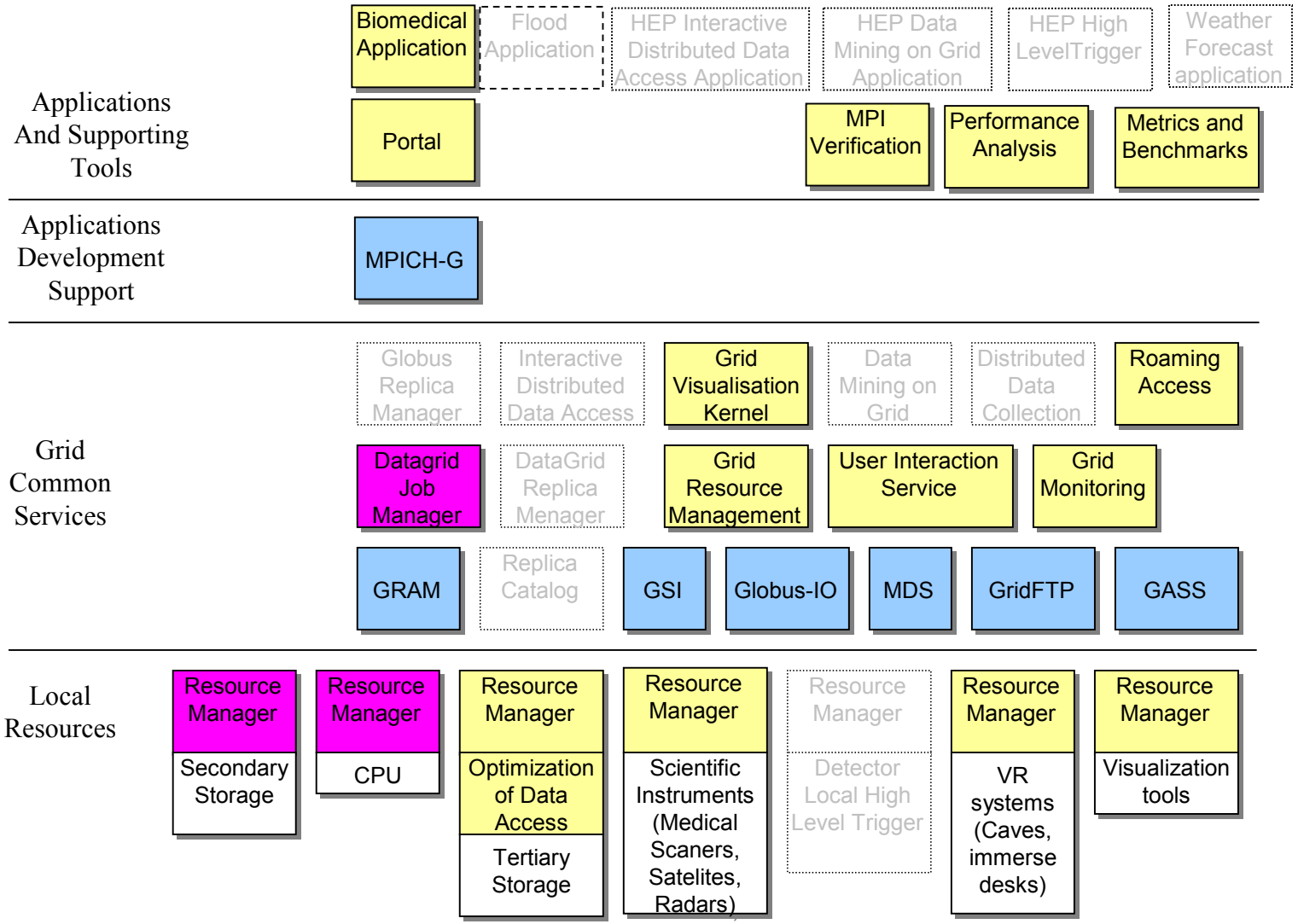
Asynchronous Execution of Biomedical Application



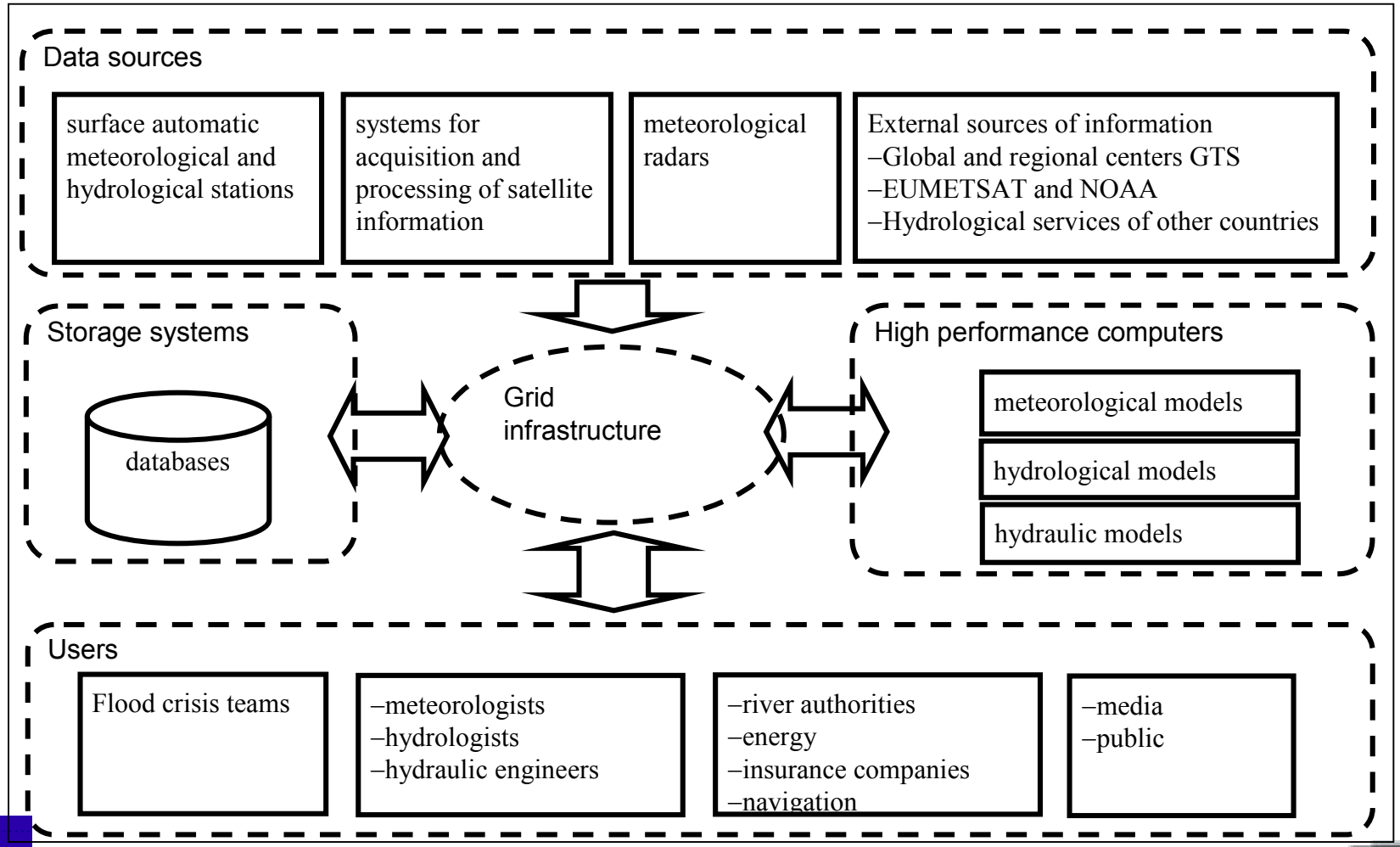
Current architecture of biomedical application



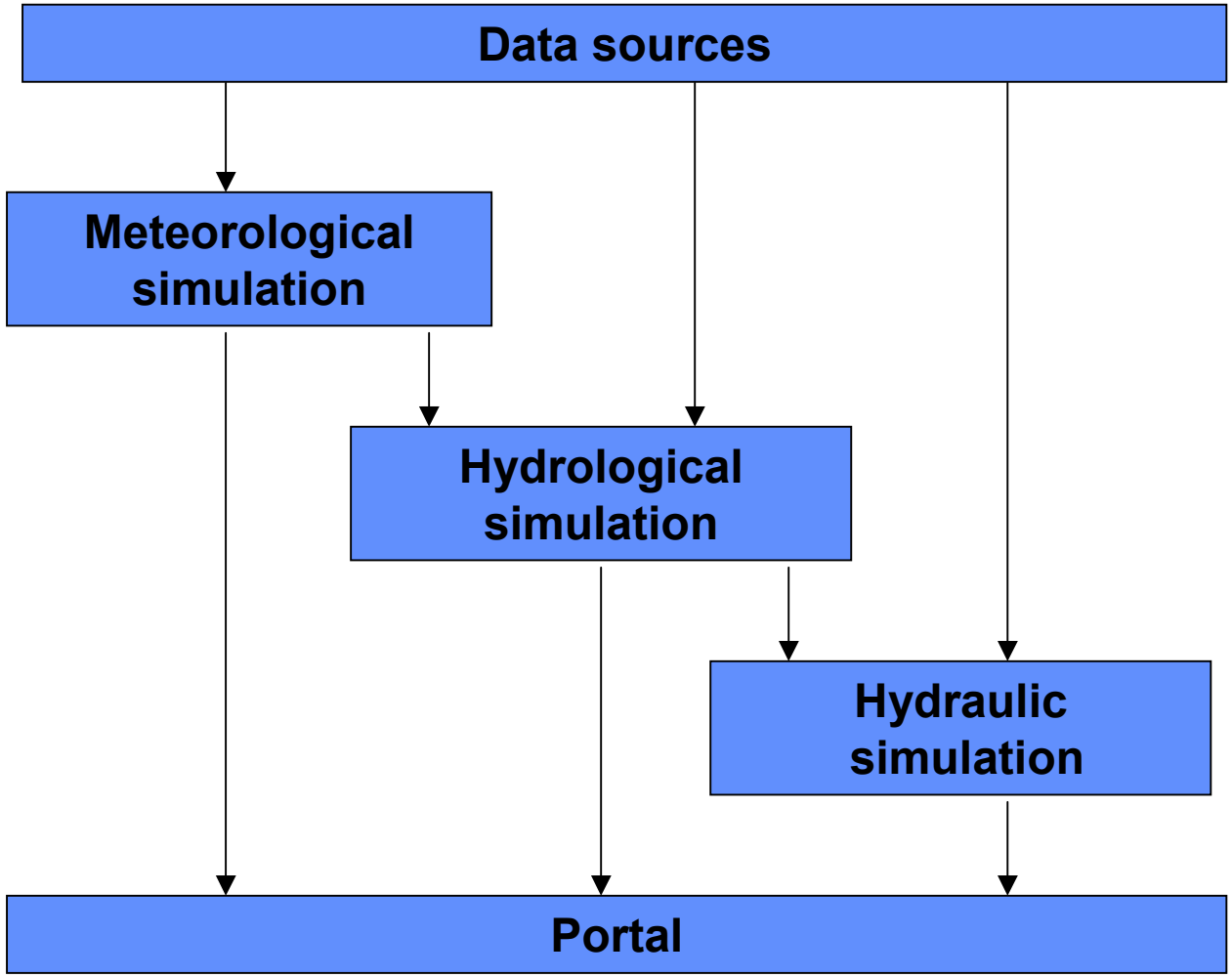
Biomedical Application



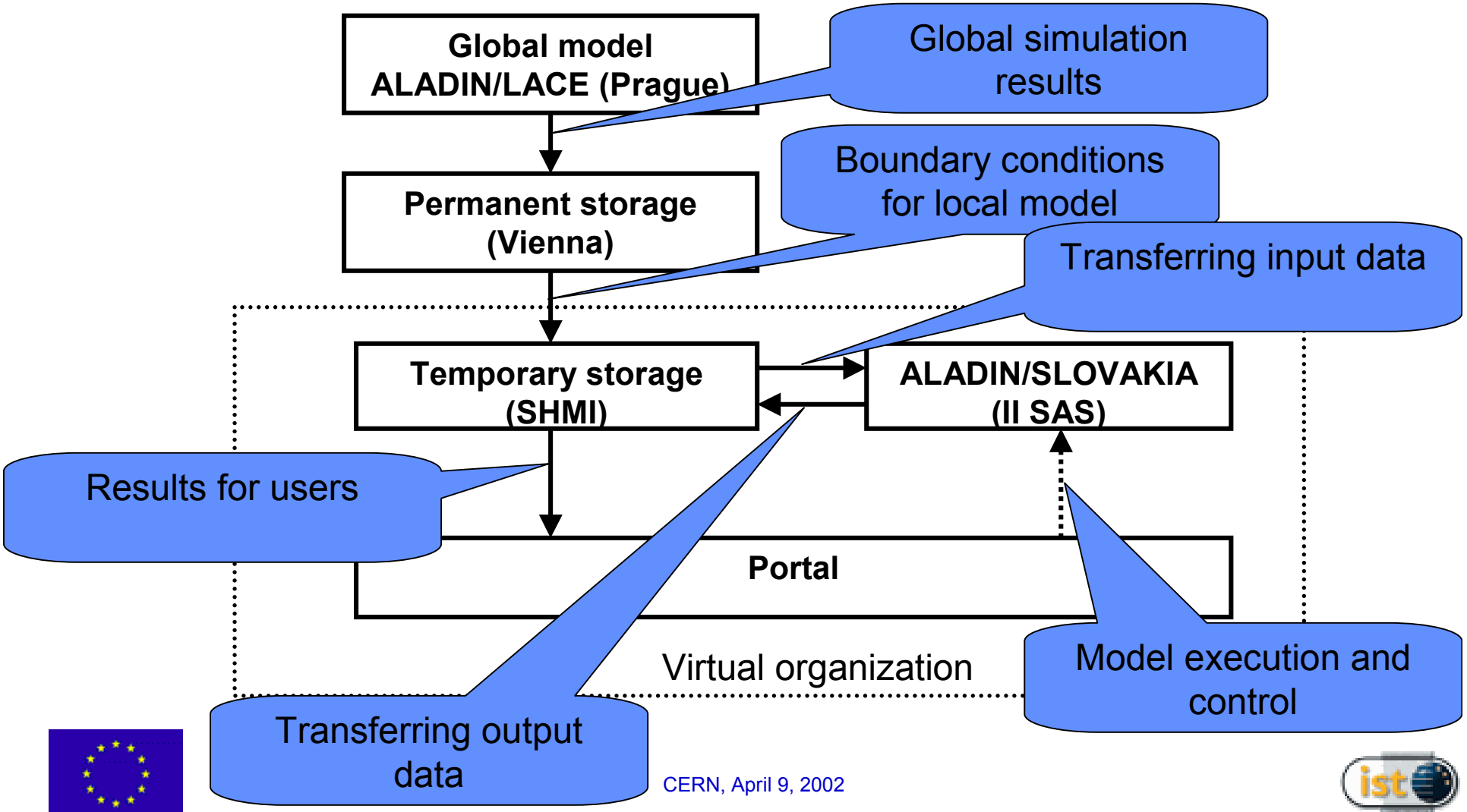
Flooding Crisis Team Support



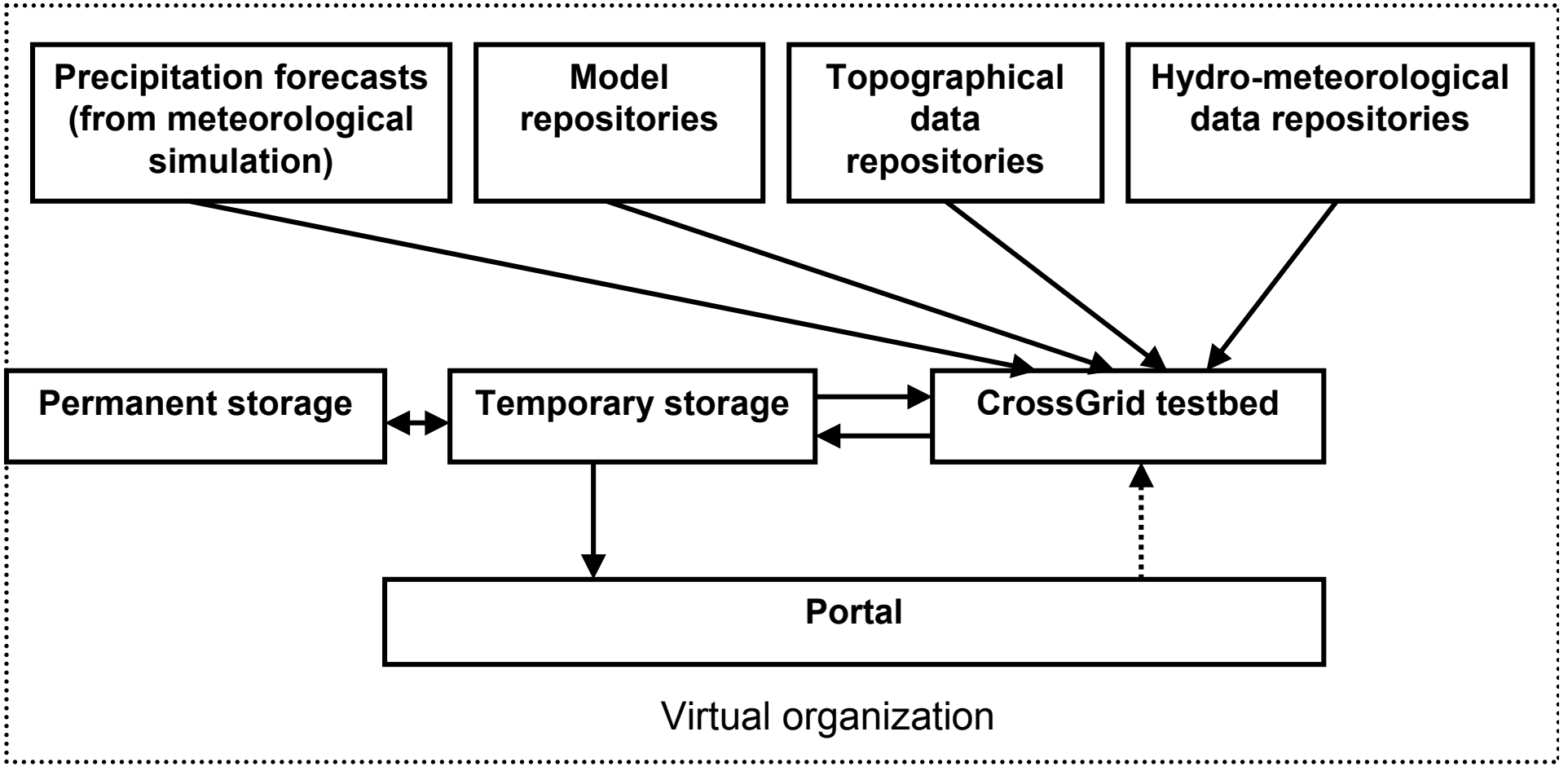
Simulation Flood Cascade



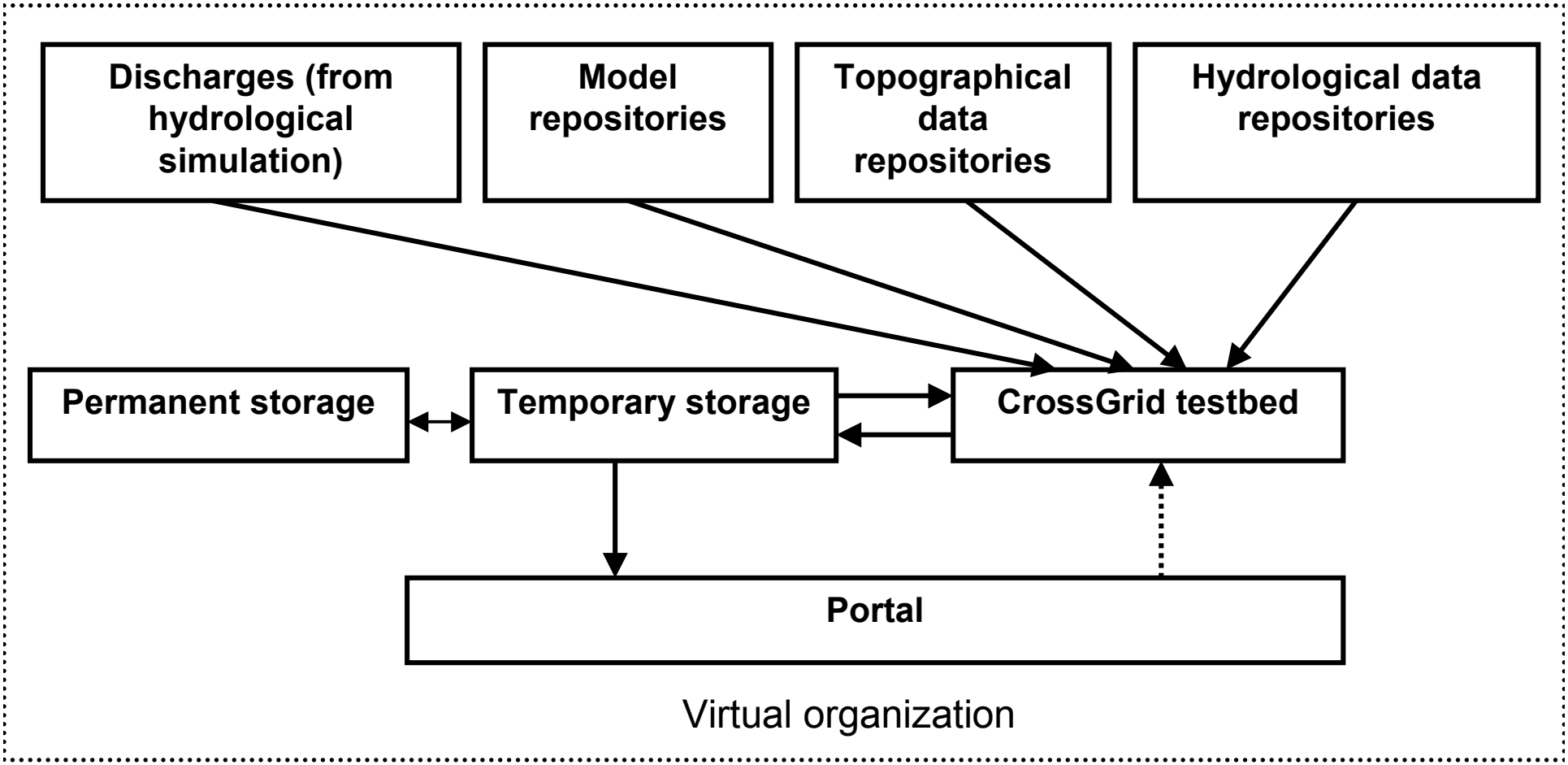
Meteorological Simulation (ALADIN Model)



Hydrological Simulation



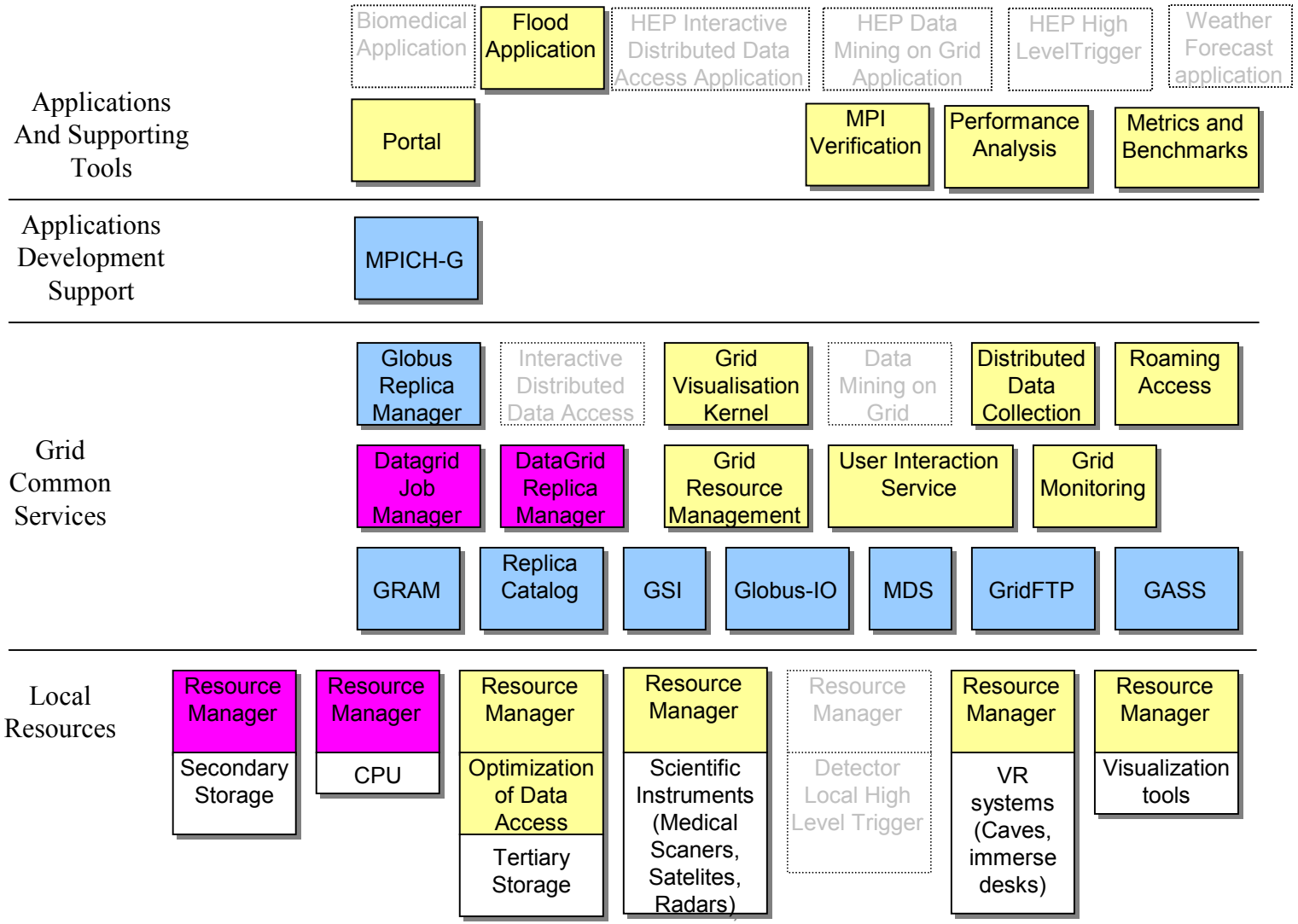
Hydraulic Simulation



Basic Characteristics of Flood Simulation

- **Meteorological**
  - intensive simulation (1.5 h/simulation) – maybe HPC
  - large input/output data sets (50MB~150MB /event)
  - high availability of resources (24/365)
- **Hydrological**
  - Parametric simulations - HTC
  - Each sub-catchment may require different models (heterogeneous simulation)
- **Hydraulic**
  - Many 1-D simulations - HTC
  - 2-D hydraulic simulations need HPC

Flooding Crisis Team Support



Distributed Data Analysis in HEP

Complementarity with DataGrid HEP application package:

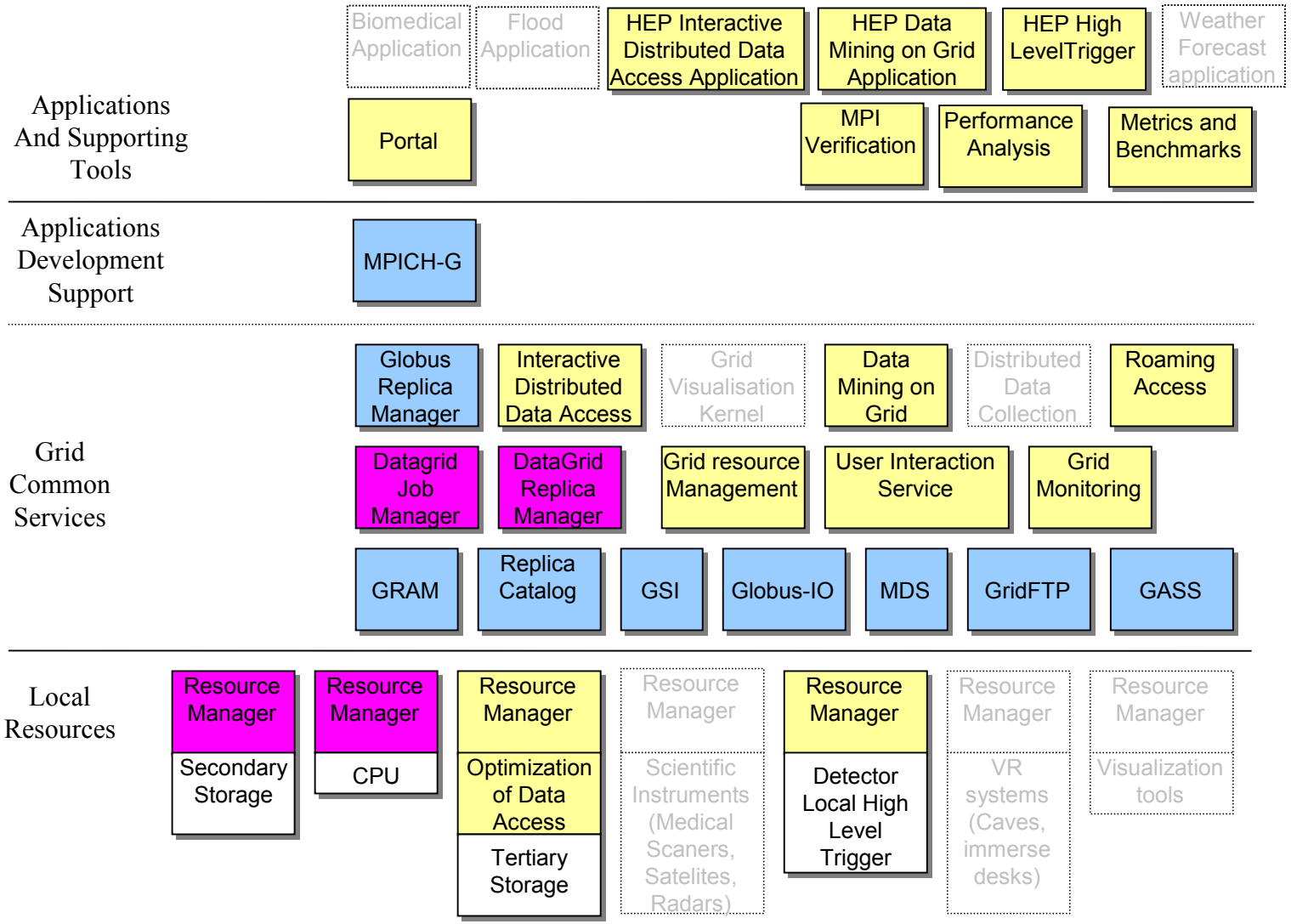
- Crossgrid will develop **interactive** final user application for **physics analysis**, will make use of the products of non-interactive simulation & data-processing preceeding stages of Datagrid
- Apart from the file-level service that will be offered by Datagrid, CrossGrid will offer an object-level service to **optimise the use of distributed databases**:
  - Two possible implementations (will be tested in running experiments):
    - Three-tier model accesing OODBMS or **O/R DBMS**
    - More specific HEP solution like ROOT.
- User friendly due to specific **portal** tools

Distributed Data Analysis in HEP

- **Several challenging points:**
  - –Access to large **distributed databases** in the Grid.
  - –Development of distributed **data-mining** techniques.
  - –Definition of a **layered application** structure.
  - –Integration of **user-friendly** interactive access.

- **Focus on** LHC **experiments (ALICE, ATLAS, CMS and LHCb)**

Distributed Data Analysis in HEP



Weather Forecast and Air Pollution Modeling

- Porting distributed/parallel codes on Grid
  - Coupled Ocean/Atmosphere Mesoscale Prediction System
  - STEM-II Air Pollution Code
- Integration of distributed databases
- Migration of data mining algorithms to Grid
- Integration, testing and running on the X# testbed

COAMPS

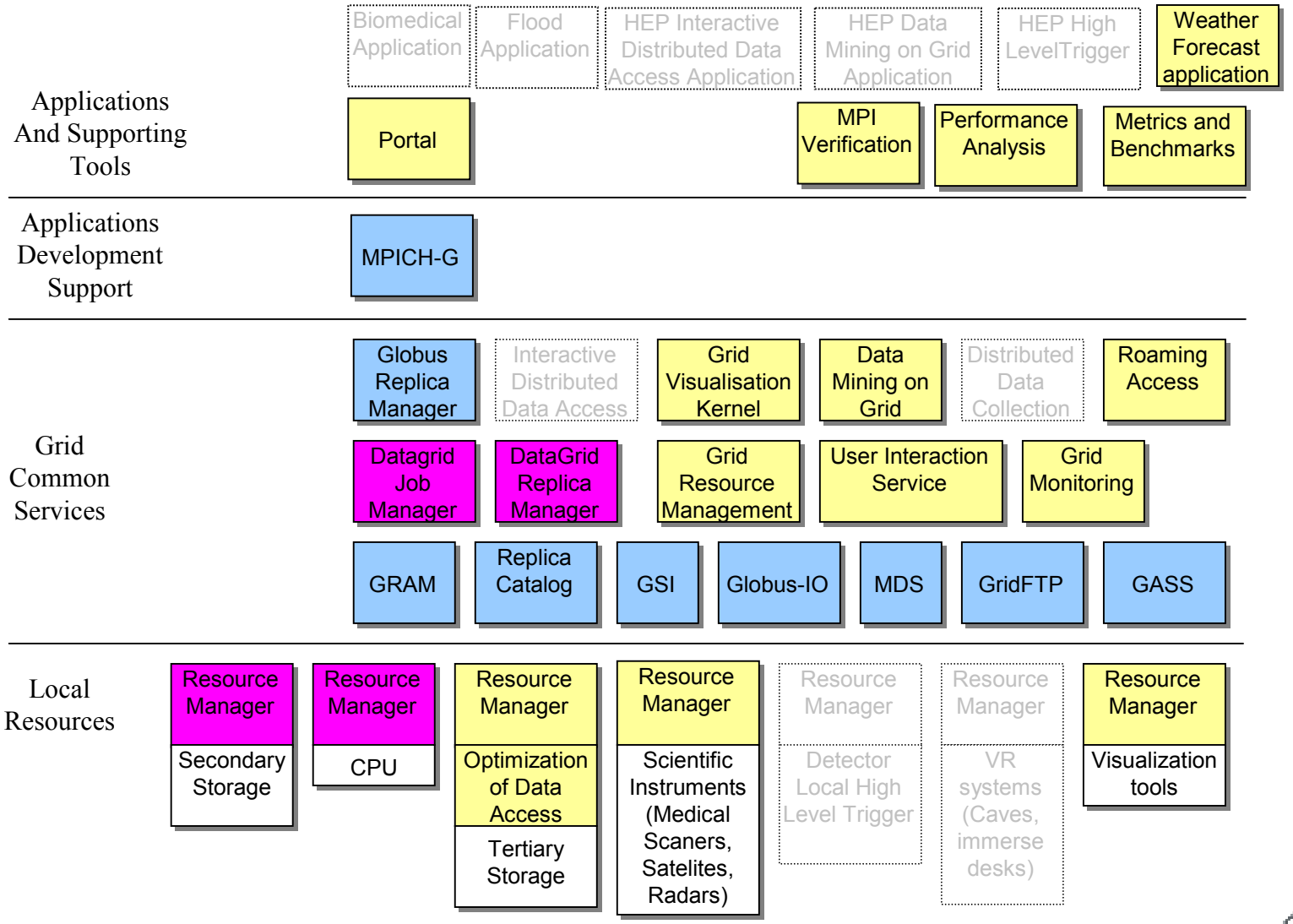
Coupled Ocean/Atmosphere Mesoscale Prediction System: **Atmospheric Components**

- **Complex Data Quality Control**
- **Analysis:**
  - Multivariate Optimum Interpolation Analysis (MVOI) of Winds and Heights
  - Univariate Analyses of Temperature and Moisture
  - OI Analysis of Sea Surface Temperature
- **Initialization:**
  - Variational Hydrostatic Constraint on Analysis Increments
  - Digital Filter
- **Atmospheric Model:**
  - Numerics: Nonhydrostatic, Scheme C, Nested Grids, Sigma-z, Flexible Lateral BCs
  - Physics: PBL, Convection, Explicit Moist Physics, Radiation, Surface Layer
- **Features:**
  - Globally Relocatable (5 Map Projections)
  - User-Defined Grid Resolutions, Dimensions, and Number of Nested Grids
  - 6 or 12 Hour Incremental Data Assimilation Cycle
  - Can be Used for Idealized or Real-Time Applications
  - Single Configuration Managed System for All Applications
  - Operational at FNMOC:
    - 7 Areas, Twice Daily, using 81/27/9 km or 81/27 km grids
    - Forecasts to 72 hours
  - Operational at all Navy Regional Centers (w/GUI Interface)

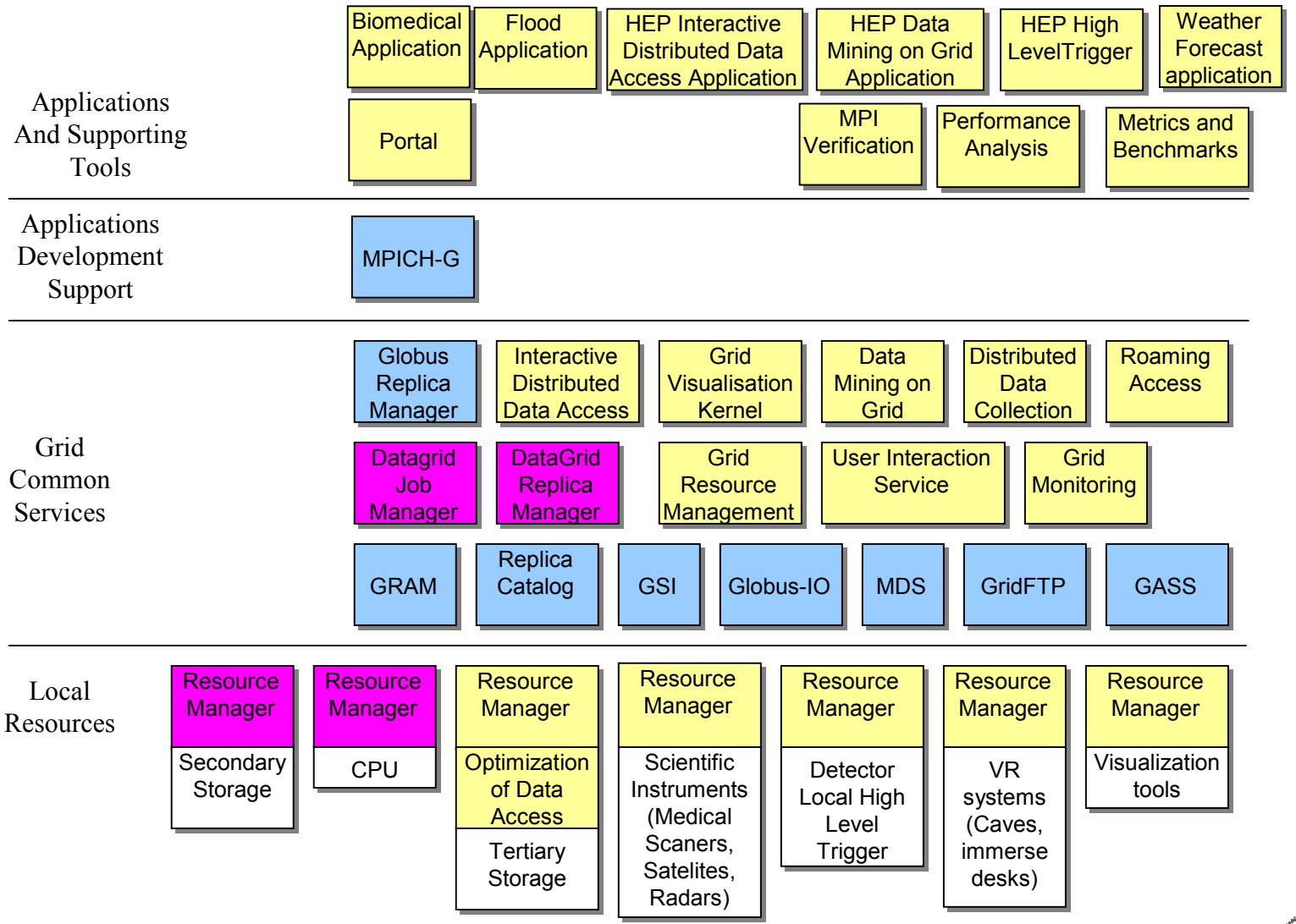
Air Pollution Model – STEM-II

- Species: 56 chemical, 16 long-lived, 40 short-lived, 28 radicals (OH, HO_2)
- Chemical mechanisms:
  - 176 gas-phase reactions
  - 31 aqueous-phase reactions.
  - 12 aqueous-phase solution equilibria.
- Equations are integrated with locally 1-D finite element method (LOD-FEM)
- Transport equations are solved with Petrov-Crank-Nicolson-Galerkin (FEM)
- Chemistry & mass transfer terms are integrated with semi-implicit Euler and pseudo-analytic methods

Weather Forecast and Air Pollution Modeling



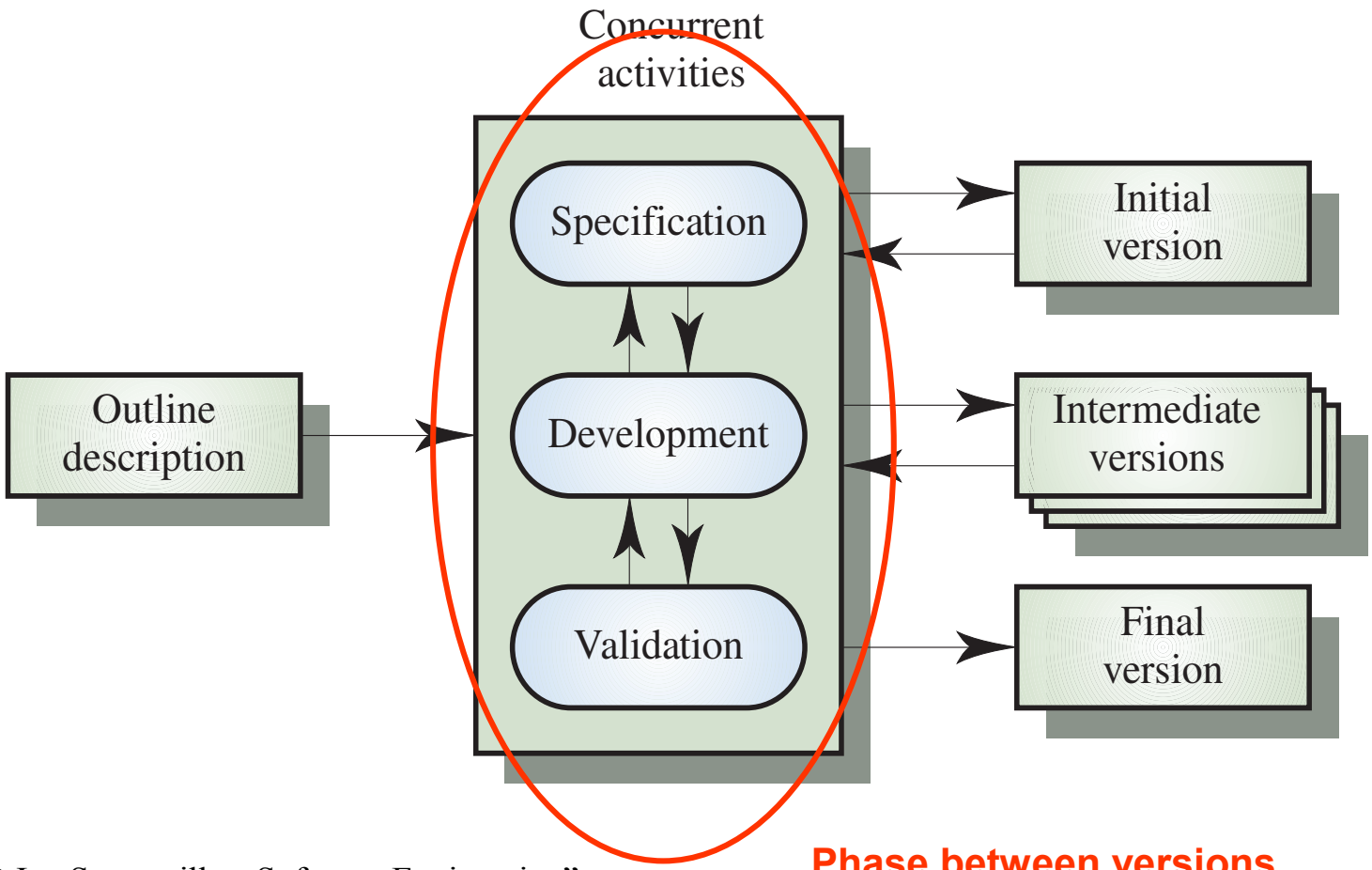
CrossGrid Architecture



Rules for X# SW Development

- Iterative improvement:
  - development, testing on testbed, evaluation, improvement
- Modularity
- Open source approach
- SW well documented
- Collaboration with other # projects

Evolutionary life-cycle model



Phase between versions

© Ian Somerville, „Software Engineering”

Architecture Team - Activity

- Merging of requirements from WP1, WP2, WP3
- Specification of the X# architecture (i.e. new protocols, services, SDKs, APIs)
- Establishing of standard operational procedures
- Specification of the structure of deliverables
- Improvement of X# architecture according to experience from SW development and testbed operation

Architecture Team - Organization

- Technical Architecture Team (at Cyfronet) – elaboration of proposals
  - Marian Bubak
  - Marek Garbacz
  - Maciej Malawski
  - Katarzyna Zając
- Representatives of WPs (persons responsible for integration in WPs) – evaluation of TAT proposals
  - WP1 – Dick van Albada
  - WP2 – Roland Wismueller
  - WP3 – Santiago Gonzalez
  - WP4 – Rafael Marco

Schedule

- Pre-final versions of SRS – April 9
- X#TAT+ meeting with DG ATF and PTB – April 11-16, CERN
- X#AT comments to SRS – April 17
- ICCS'2000 Amsterdam
- Final versions of SRS (deliverables!) – April 25
- 1st def of X# Architecture – May 17-18, Cracow