LCG Applications Area Overview

Applications Area Internal Review

30 March – 1 April 2005

Pere Mato/CERN

Outline

- Charge of the Review
- Applications area organization and overview
- Phase 2 programme of work preparation Proposed changes in AA
- Personnel resources, participation
- Concluding remarks

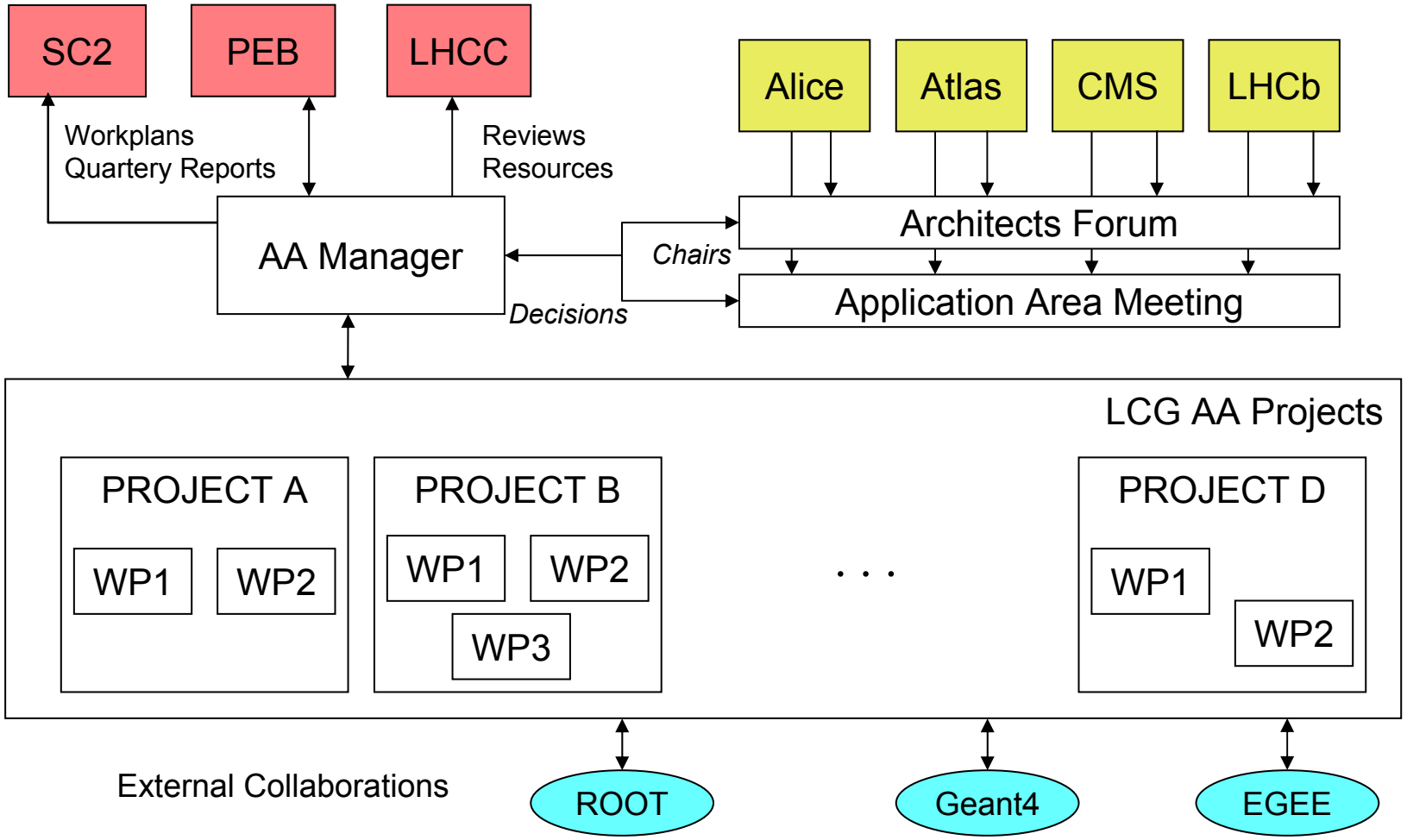
Charge of the AA internal Review

- Technical review of the software and infrastructure of the Applications Area
- Follow-up of the review that took place in October 2003
  - Progress that has been made since the last review
  - Adoption of the recommendations
- Feedback on the program of work for the second phase of the LCG project
  - About the new project organization
  - The objectives to serve best the LHC experiments
  - The technical choices

Application Area Focus

- Deliver the common physics applications software
- Organized to ensure focus on real experiment needs
  - Experiment-driven requirements and monitoring
  - Architects in management and execution
  - Open information flow and decision making
  - Participation of experiment developers
  - Frequent releases enabling iterative feedback
- Success defined by experiment validation
  - Integration, evaluation, successful deployment

Applications Area Organization



Current AA Projects

- SPI – Software process infrastructure (A. Aimar)
  - Software and development services: external libraries, savannah, software distribution, support for build, test, QA, etc.
- SEAL – Core Libraries and Services (P. Mato)
  - Foundation class libraries, math libraries, framework services, object dictionaries, python services, etc.
- POOL – Persistency Framework (D. Duellmann)
  - Storage manager, file catalogs, event collections, relational access layer, conditions database, etc.
- PI - Physicist Interface (V. Innocente)
  - Analysis services (AIDA), Analysis environment (CINT, Python)
- SIMU - Simulation project (G. Cosmo)
  - Simulation framework, physics validation studies, MC event generators, participation in Geant4, Fluka.

AA Organization

- Application Area Meetings (AAM)
  - Informal forum of exchange of information between the AA projects and experiments, etc.
    - » Project status, release news, results, new ideas, evaluations, new requirements, general discussions, experiment feedback, etc.
  - Encourage presentations from the projects and experiments
  - Every two weeks on Wednesdays @ 16:30
  - Each meeting should have a "theme" defined well in advance

AA Organization (2)

- Architects Forum Meetings
  - Formal decision and action taking meeting
  - Consists of the experiment architects, AA projects leaders, computing coordinators with an standing invitation and other invited participants
  - Experiments participate directly in the planning, management, and architectural and technical direction of AA activities
  - Public minutes after internal circulation
  - Good atmosphere, effective, agreement generally comes easily. No problems so far.

AA Organization (3)

- Work Plans
  - AA Projects must prepare yearly work plans
  - Opportunity to re-think strategic decisions, change of direction, introduce new work packages
  - Approved by PEB
- Quarterly reports
  - To monitor progress of the projects
  - Scrutinized by SC2 committee and generation of feedback
- Reviews
  - Internal AA reviews and LHCC reviews

ROOT Collaboration

- User-Provider relationship with LCG AA
  - Defined in the Blueprint RTAG
- Close and effective collaborative effort between POOL and ROOT teams
- AA internal review recommended an evolution towards a closer cooperation
- Planned the convergence in a number of areas
  - Mathematical libraries, dictionaries, POOL reference support, etc.
  - Currently in these areas most work is being done on the LCG side; objectives have not been reflected in ROOT plans and work commitments

Current Status

- The status of SPI, SEAL, POOL and Simulation projects will be covered in detail in separated presentations
- PI status will be coved briefly with the next slides

Validation Highlights

- POOL successfully used in large scale production in ATLAS, CMS, LHCb data challenges in 2004
  - ~400TB of POOL data produced
  - Objective of a quickly-developed persistency hybrid leveraging ROOT I/O and RDBMSes has been fulfilled
- Geant4 firmly established as baseline simulation in successful ATLAS, CMS, LHCb production
  - EM & hadronic physics validated
  - Highly stable: 1 G4-related crash per O(1M) events
- SEAL components underpin POOL's success, in particular the dictionary system
  - Now entering a second generation with Reflex
- SPI's Savannah project portal and external software service are accepted standards inside and outside the project

Status of main issues identified by 2003 review

- Build/config tool
  - XML-based config description now in development will support SCRAM, CMT and config/make
- Schema evolution
  - Addressed by ROOT4 migration
- POOL collections
  - Not resolved
- SEAL libraries management
  - Dependency management and reduction ongoing. SEAL partitioning work.
- Dictionary convergence
  - Ongoing as a priority
- Generic SIMU framework
  - Dropped due to ATLAS, CMS, LHCb disinterest
- Redefine PI in light of ARDA
  - PI in maintenance mode

Current Status of PI

- Developments of the Analysis Services component is completed
  - Set of component libraries implementing AIDA interfaces
  - Flexible: choose implementation at runtime using plug-in manager system from SEAL
    - » ROOT and Native implementations for histograms
    - » Storage (I/O) for Histograms and Tuples in ROOT, HBook and XML compressed format
- Provide easy conversion between all formats
  - Well tested, large number of unit tests
    - » Failures due to differences in implementations
  - Provided Python bindings to AIDA interfaces using LCG Dictionary and PyLCGDict from SEAL
- Easy interoperability with external tools
  - Can display histogram in Python using ROOT, JAS and HippoDraw

Current Status of PI (2)

- Ongoing maintenance
  - Low effort, mainly bug fixes
  - Release new versions following SEAL, AIDA, etc. releases
- Users of PI :
  - Histogram libraries are used by Gaudi (LHCb + ATLAS) and CMS
    - » LHCb is evaluating to use also AIDA Tuple libraries
  - Used by Geant4 in advanced examples for storing histograms and tuples
  - Python layer is used by some physicists for analysis (fitting, etc…)
- Need to provide customized download and installation for external users

Milestones 2004 (level2)

SPI		Expected Date	Done
	savannah.cern.ch migrated to GNU savannah	20-02-2004	04-02-2004
	RH 7.3 gcc 3.2.3 supported	15-05-2004	15-05-2004
	SPI/EGEE collaborative workplan complete	31-05-2004	28-10-2004
	Certification of external software for the new Linux platform	01-07-2004	01-06-2004
	LCG AA build system selection	15-07-2004	
	External software guideline document	31-12-2004	
SEAL			
	Workbook for SEAL	15-06-2004	30-06-2004
	New Dictionary API and reference implementation	30-06-2004	03-09-2004
	mathlib project web	15-07-2004	15-07-2004
	First version of the C++ mathlib package	01-10-2004	29-01-2005
POOL	**Persistency Framework**		
	POOL hierarchical cataloging production release	15-03-2004	09-03-2004
	POOL RDBMS abstraction layer completed	31-05-2004	31-05-2004
	RDBMS independency achieved for POOL relational components	30-06-2004	30-06-2004
	First release of the POOL Relational Storage Manager	31-08-2004	17-11-2004
	POOL meets scalability requirements	31-10-2004	31-10-2004
	POOL integrates ROOT4	31-10-2004	28-02-2005
	Conditions Database		
	Common interface for Conditions DB defined	30-06-2004	12-12-2004

Milestones 2004 (level 2) (2)

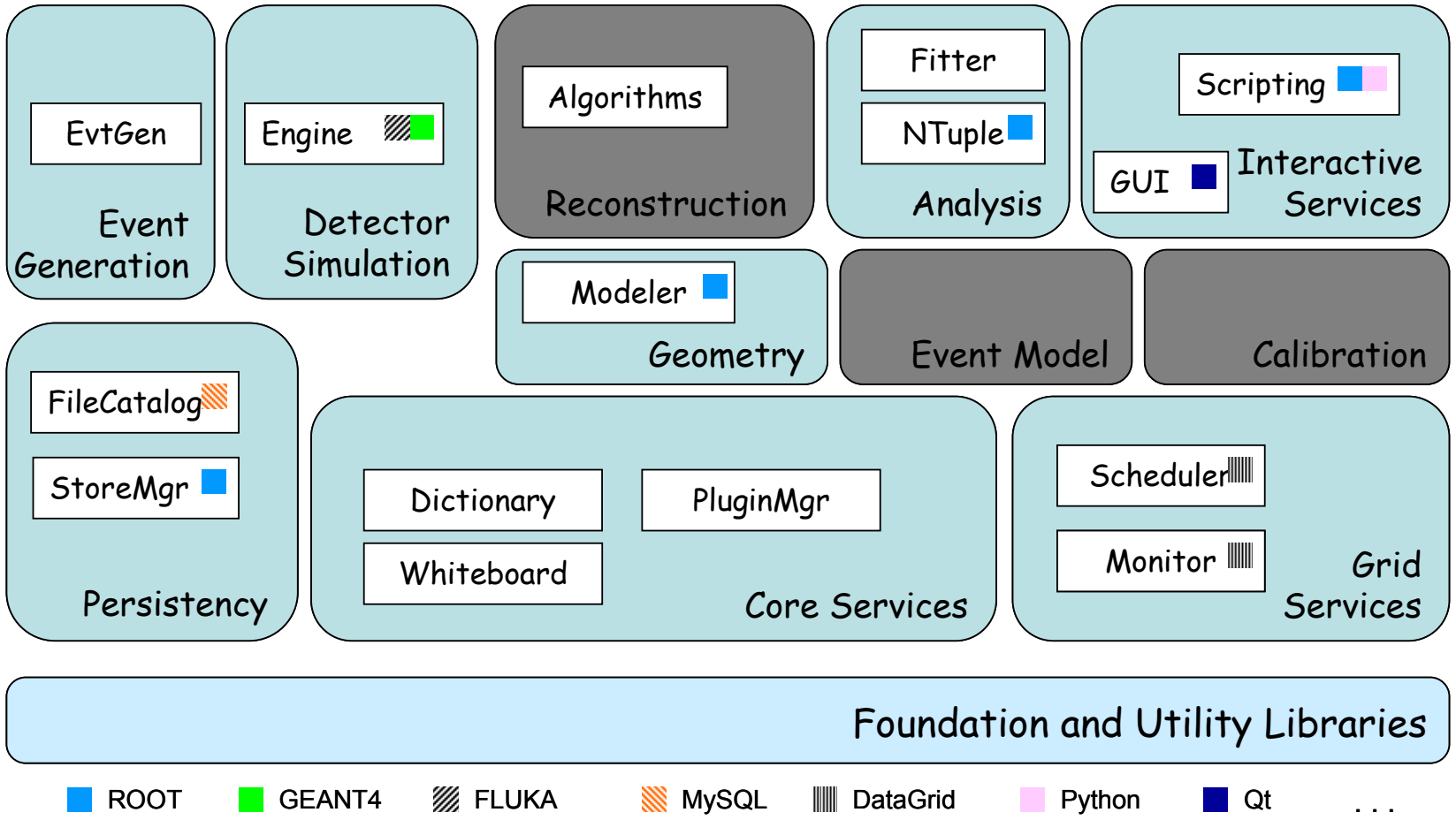
SIMU		Expected Date	Done
	Geant4 development and integration		
	SPI-G4 collaborative infrastructure pilot	16-02-2004	16-02-2004
	Geant4 6.2 release - resource usage refinements	25-06-2004	25-06-2004
	First consolidated G4 acceptance suite for LHC applications	15-10-2004	14-12-2004
	Geant4 7.0 release - physics models and geometry	17-12-2004	17-12-2004
	Geant4 validation in LHC production	31-12-2004	32-03-2005
	Physics validation		
	First cycle of hadronic physics validation complete	01-02-2004	30-06-2004
	Review/prioritization of simple benchmarks for simu physics validation	31-05-2004	
	Comparison of LHC calorimeters for EM shower development	15-09-2004	
	Second iteration of hadronic physics validation complete	31-12-2004	
	Simulation test and benchmark suite available	31-12-2004	31-12-2004
	Generator services		
	Agreement on formats for event generator common samples	01-03-2004	01-03-2004
	Beta version of MCDB in production in the LCG environment	01-07-2004	28-06-2004
	Agreement on parton-level event generator file format	01-09-2004	01-09-2004
	Generator production framework beta	01-12-2004	

Planning AA in LCG Phase 2

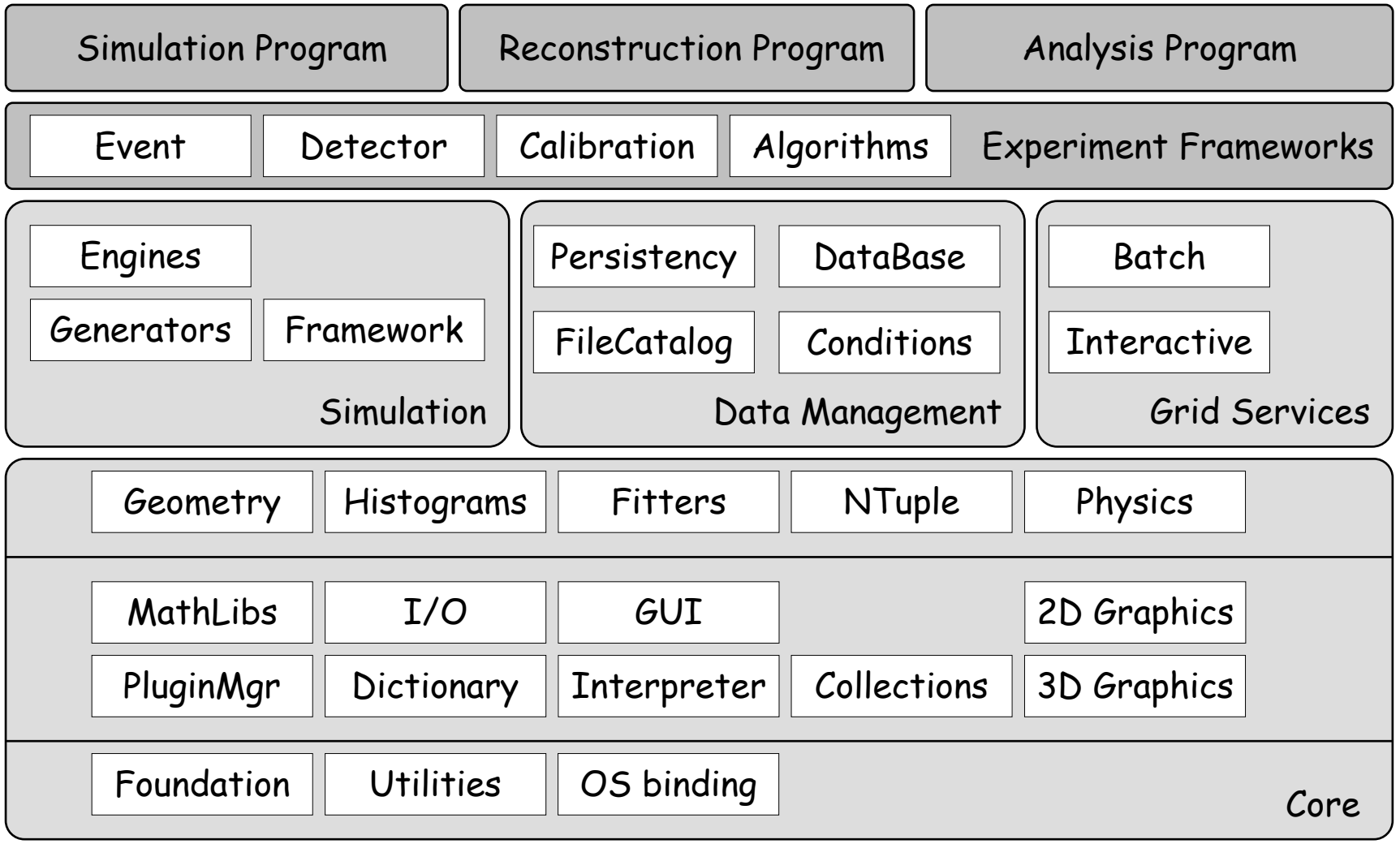
Phase 2 preparation

- Phase 2 covers from mid-2005 to mid-2008
- Need to establish the level of long-term support that is required for the products that are essential for the experiments
  - Re-use software and infrastructure across projects. Minimize duplication
  - Ease maintenance of AA software at the end of the LCG
- More emphasis in development of **Physics Analysis**
- AA Internal Review
  - Presentations with the plans for each project
  - Opportunity to receive advise and recommendations
- Written plans
  - Planning document will be produced
- Approval by PEB in April

Old Domain Decomposition (Blueprint)



New Domain Decomposition



Proposed Changes in AA Projects

- Round of discussions with PH management, project leaders, experiment architects, experiment representatives, presentations to experiments, etc.
- The proposed changes the Applications Area are:
  1. SEAL and ROOT projects merge
  2. Some redefinition of SPI role
  3. Some adaptations of POOL required
  4. PI libraries absorbed by client projects
  5. SIMULATION project basically unchanged

SEAL + ROOT project merge

- Both SEAL and ROOT projects have a big overlap
  - The objectives are very similar
  - Avoid duplication by construction
- Single AA project to provide all the **core and framework software**
  - Put all the people involved in a single TEAM
  - Select or evolved each provided functionality to the best technical solution
  - Make sure that all clients receive a good service
  - Encourage the usage of the core software by the other domains

Why the SEAL+ROOT Merge?

- Optimization of resources
  - Avoid duplicate developments
- Better "coherency" vis-à-vis our clients, the LHC experiments
- ROOT activity fully integrated in the LCG organization
  - Planning, milestones, reviews, resources, etc.
- Ease long-term maintenance and evolution of a single set of software products
  - Thinking on the post-LCG era

What it means in practice?

- Single team lead by Rene Brun
- Combined program of work, single deliverable
  - Initially the union of what is available in SEAL+ROOT
  - Continuation of the convergence work in Dictionary, MathLibs, etc.
  - Towards a single set of functionalities in a time scale of 1-2 years
- User-level compatibility MUST be maintained
  - Adiabatic transition for the LHC collaborations
- Software **evolution** decided by the new project team
  - The "WHAT" is agreed with the experiments
  - The "HOW" is mostly left to the developers with input from experiment core people
  - Distributed responsibility (sub-projects)
- Open team/project nature
  - Encourage external participation

Known difficulties

- Cultural merge
  - Team members with different backgrounds
  - Different ways of developing software
  - Need to establish a "common" culture

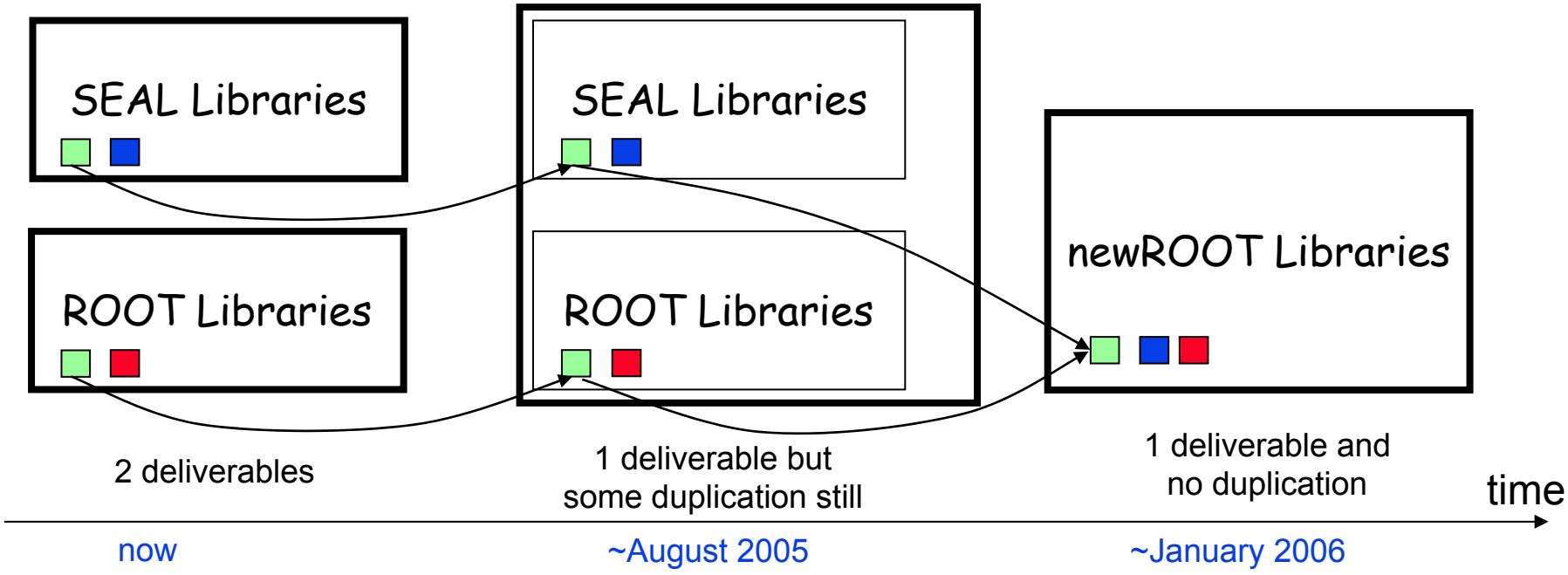
    ➔ Compromises from both sides
- Client merge
  - Possible conflicts between LHC requirements and other ROOT users requirements
  - Client adaptability to changes
  - Platforms support, licensing issues, etc.

    ➔ Study case by case
- The evolution of each functionality needs to be plan carefully
  - Detailed plans should be developed by the new Team and discussed together to the experiments

    ➔ mainly in AF

SEAL + ROOT Migration

- Adiabatic changes towards experiments
  - Experiments need to see libraries they use currently will evolve from current usage today towards a unique set
- Some details given during this review
  - More will be needed in the Programme of Work document
  - Will be extra tasks in order to complete migration



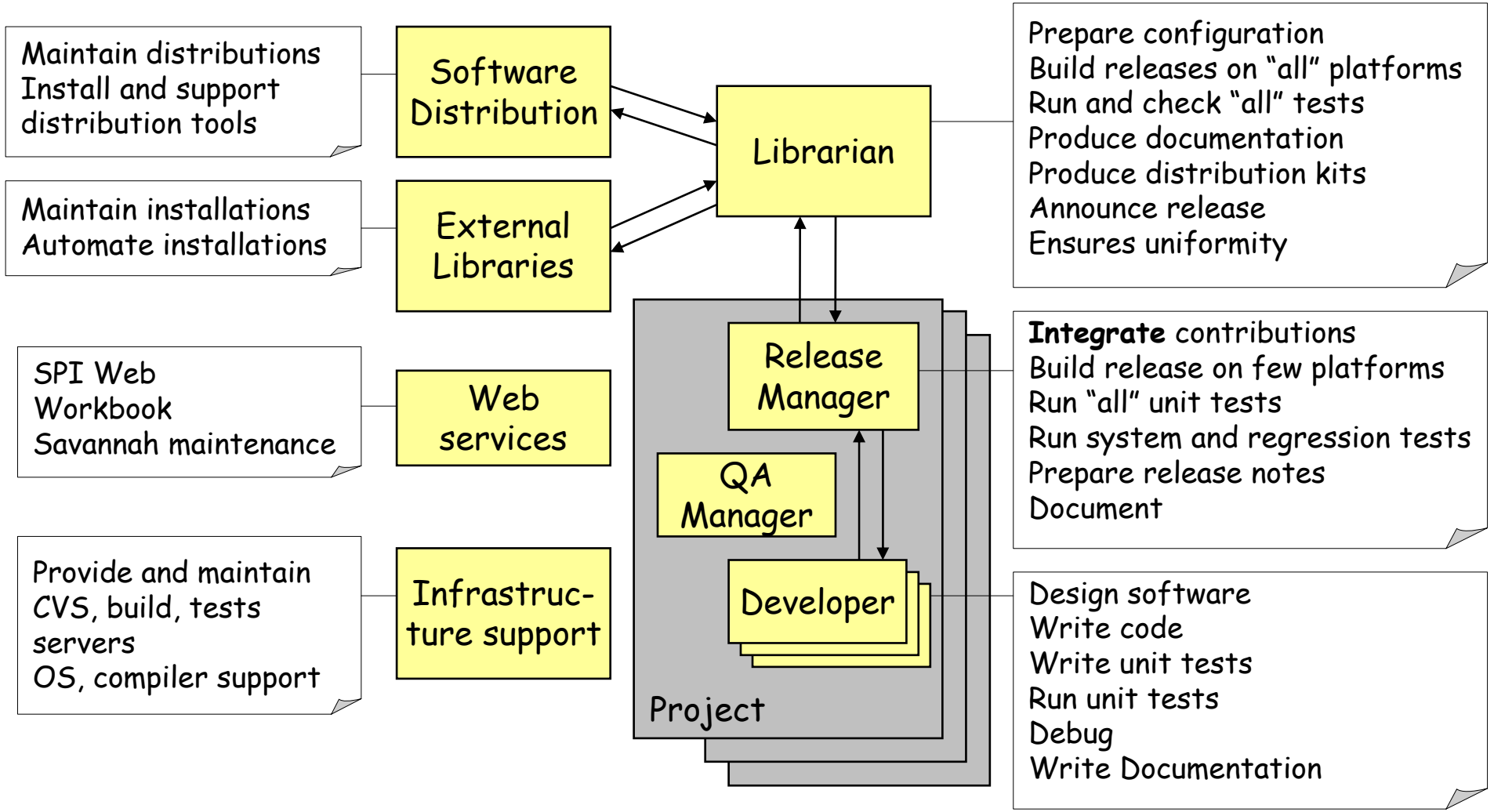
SEAL+ROOT

- Foundation libraries
  - Workout a proposal to merge SEAL/Foundations with ROOT base
- Dictionary
  - The goal is a single dictionary independent of its clients (I/O, interpreters, etc.)
  - Complete the already started convergence work
- I/O system
  - Migrate ROOTStorageSvc from POOL.
  - Possible convergence between POOL Object-Relational storage and ROOT TTreeSQL
- Mathematical libraries
  - Preparing a strategy with TMath, GSL, Fitting, MathCore, … and future of CLHEP
- Component-ware
  - Convergence in Components, Plugin management, abstract interfaces, etc.
- Scripting
  - Uniformity between CINT and Python

SPI: Redefinition of its role

- A number of "services" of common interest will continue to be run
  - Savannah, external libraries, etc.
- Helping projects and LHC experiments to provide/maintain the software development infrastructure
  - CVS servers, build server, etc.
- Direct participation in the software development projects
  - Librarians, release managers, documentation, toolsmiths, QA

Software Development Roles



Estimated resources

	Common (SPI)	CORE	POOL	SIMU
Software Builds	Librarian (0.5 FTE)	Release Manager (0.5 FTE)	Release Manager (0.5 FTE)	Release Manager (0.5 FTE)
External Libraries	External Lib Mgr (0.5 FTE)	(0.1 FTE)	(0.1 FTE)	(0.1 FTE)
Software Distribution	Distribution Mgr (0.5 FTE)			
Quality Assurance	QA support (0.2 FTE)	(0.5 FTE)	(0.5 FTE)	(0.5 FTE)
Documentation	Web master (0.4 FTE)	(0.2 FTE)	(0.2 FTE)	(0.2 FTE)
Project portals (savannah)	Savannah support (0.8 FTE)	(0.1 FTE)	(0.1 FTE)	(0.1 FTE)
Development Infrastructure	Infrastructure support (0.2 FTE)			
	~3.1 FTE	~1.3 FTE (*)	~1.3 FTE (*)	~1.3 FTE (*)

(*) do not need to be equally distributed among all projects

POOL: Some adaptations required

- Domain of expertise in data persistency, data management, deployment in the Grid and (relational) databases in general
- No major changes in the structure are proposed
- Two differentiated parts
  - POOL (object persistency)
  - COOL (conditions database)
- Proposed to move the ROOT storage manager implementation to new SEAL+ROOT project
- Study the collections and their relations with ROOT trees
- Started discussions about the Relation Access Layer (RAL) organization

PI: absorbed by client projects

- The proposal is that the project is discontinued as such
- Make the inventory of existing libraries and study their usage by LHC experiments
  - If not used then abandon library
  - If used by a single experiment (or single framework) move the library in question to the experiment (or framework)
  - Incorporate remaining parts to SEAL+ROOT project

SIMULATION: Basically unchanged

- Domain of expertise in event generators and detector simulation
- No changes in the structure are proposed
- Current subprojects
  - Simulation Framework
  - Geant4
  - Fluka
  - Physics Validation
  - Generator Services
- Added new subproject Garfield
- Encourage to (re)use the core software and software development infrastructure
  - Interactivity, persistency, analysis, etc.

Staffing Requirements Estimates

The total staffing levels estimated as required, and being planned:

	2004	2005	2006	2007	2008
SPI	6.2	6.2	5.5	4.6	3.9
SEAL+ROOT	6.8+6.5	5.9+6.2	6.2+6.4	5.7+5.5	4.7+4.8
POOL	14.1	14.4	12.8	8.5	7.5
PI	0.7	0.1	0	0	0
Simulation	15.6	10.5	9.7	7.9	7.5
Total	49.9	43.3	40.6	32.1	28.4

Staffing Sources

	2004	2005	2006	2007	2008
CERN base	21.2				
LCG special contributions	17.7				
Staff at CERN	38.9	37	34.5	27	24
Experiments	10.6	7	6.5	6	5
Total	49.5	44	41	33	29

Concluding Remarks

- Started to plan second phase of Applications Area
- The major proposed change for this new phase is the merge of ROOT and SEAL projects
- Adjustments proposed in SPI, POOL
- No major changes foreseen in SIMU
- Work plans are going to be presented during this review for each project
- The plans might be with insufficient detail at this time, so further discussions with experiments are going to be needed to polish them