
LCG Applications Area Overview

Applications Area Internal Review
18-20 September 2006

Pere Mato/CERN



Outline

- ◆ Charge of the Review
- ◆ Applications Area organization and Overview
- ◆ Highlights of last 18 months
- ◆ Selected issues to be reviewed
- ◆ Personnel resources, participation
- ◆ Milestones

Charge of the AA internal Review

- ◆ Technical review of the software and infrastructure of the Applications Area
- ◆ Follow-up of the AA internal review that took place in April 2005
 - Progress that has been made since the last review
 - Level of adoption of the recommendations
- ◆ Assess the readiness of the Common Application Software for the LHC startup
 - Any change in the structure needed?
 - What functionality still needs to be developed?
 - What requires better consolidation?
 - Any other major service needed?

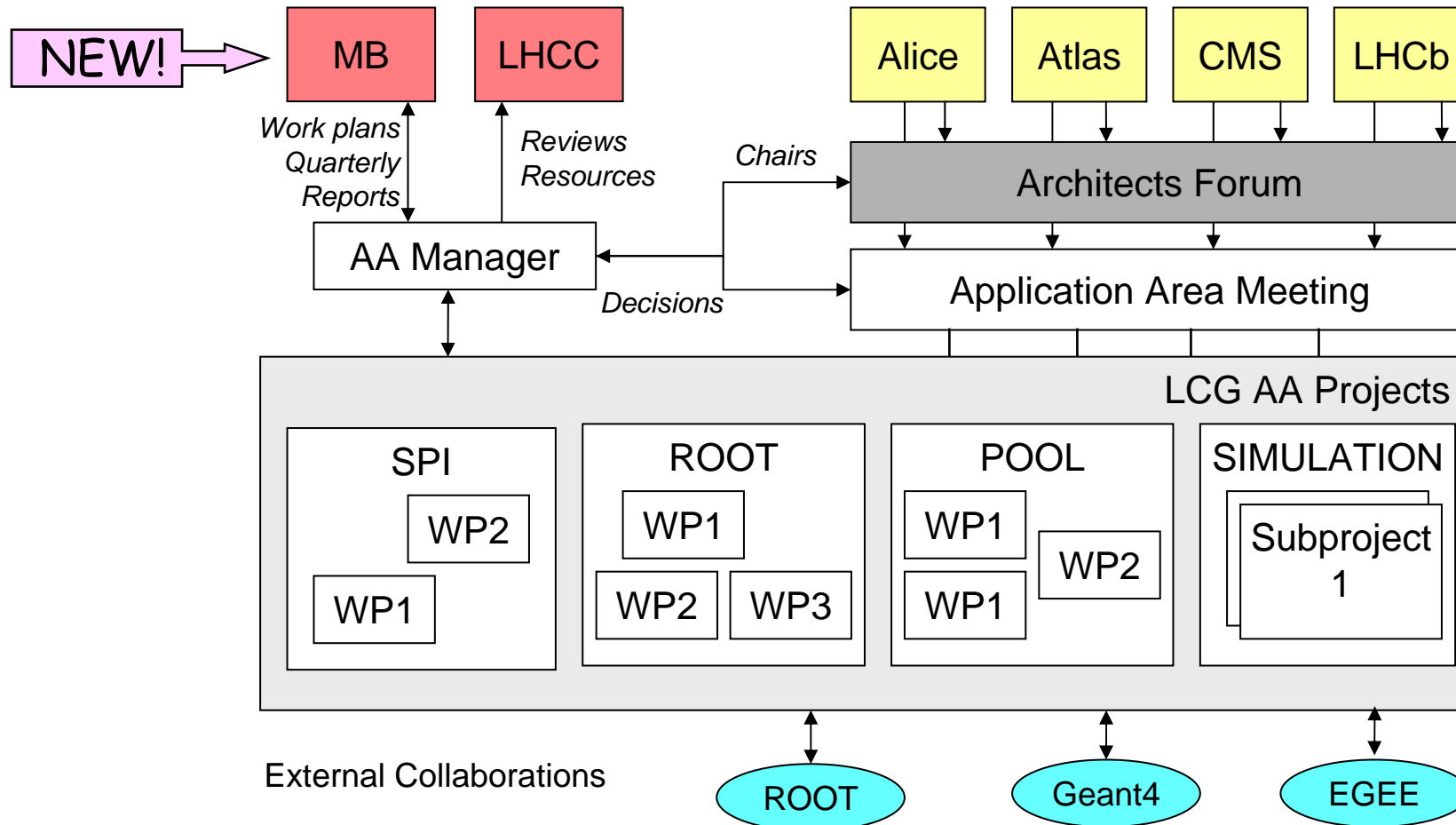
Applications Area Focus (a reminder)

- ◆ 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

Common Physics Applications Software

- ◆ The definition of "common software" is software that is used by at least two Experiments
- ◆ We have always allowed some level of speculation
 - Software "requested" by the experiments
 - Software that is going to be used by the experiments
 - Software that "we think" is going to be used/requested by experiments
- ◆ At the end of the day, the speculation needs to be confirmed
 - Seeking for communality
 - Dropping or transferring non-common software

Applications Area Organization



AA Projects



- ◆ SPI - Software process infrastructure (A. Pfeiffer)
 - Software and development services: external libraries, savannah, software distribution, support for build, test, QA, etc.
- ◆ ROOT - Core Libraries and Services (R. Brun)
 - Foundation class libraries, math libraries, framework services, dictionaries, scripting, GUI, graphics, SEAL libraries, etc.
- ◆ POOL - Persistency Framework (D. Duellmann)
 - Storage manager, file catalogs, event collections, relational access layer, conditions database, etc.
- ◆ SIMU - Simulation project (G. Cosmo)
 - Simulation framework, physics validation studies, MC event generators, Garfield, participation in Geant4, Fluka.



Application Area Meetings

- ◆ 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.
- ◆ Every two weeks on Wednesdays @ 16:30
- ◆ Encourage presentations from the projects and experiments
- ◆ Each meeting should have a "theme" defined in advance
 - Recurrent item in the AF agenda

Architects Forum Meetings

- ◆ 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
- ◆ Every two weeks
- ◆ Public minutes after internal circulation
 - <http://lcgapp.cern.ch/project/mgmt/af.html>
 - <https://twiki.cern.ch/twiki/bin/view/LCG/LcgBulletins>
- ◆ Good atmosphere, effective, agreement generally comes easily. No problems so far.

Workplans and Quarterly Reports

◆ Project milestones

- AA projects propose a number of milestones
 - » Discussed in the AF
 - » Added into quarterly reports for approval

◆ Quarterly reports

- To monitor the progress of the projects
- Basically limited to an executive summary and the status of the current milestones together with new proposals
- Reviewed by LCG Management Board

◆ Reviews

- Internal AA reviews
- LHCC Comprehensive Review



Main highlights during last 18 months

- ◆ Merge of the SEAL and ROOT teams into a single project team lead by Rene Brun
- ◆ AA Plan for LCG phase II
 - Many discussions in the context of the AF
 - The plan document was finally endorsed by the PEB committee on September 27th
- ◆ FLUKA released with new license schema
- ◆ Andreas Pfeiffer replaced Alberto Aimar as SPI project leader
 - Implied some changes in the scope and organization of the SPI activities
- ◆ ROOT workshop September 05
 - An opportunity to present to the world the latest developments in ROOT and get the feedback from the experiments other than LHC
- ◆ Migration to the new Reflex library by POOL and Python scripting has been completed by December 05
 - Experiments started to migrate to this new version (ROOT 5)

Main highlights during last 18 months (2)

- ◆ The first public release of the new re-engineered version of the relational database API package (CORAL) was made available in December 05
- ◆ End of the year production versions of ROOT and Geant4, which included a long list of new functionality required by the experiments
- ◆ The HepMC package installed in the LCG external area and maintained by FNAL effort
- ◆ Software adapted for the AMD64 architecture and certification for Linux SLC4
- ◆ New procedures for testing and building the software put in place to optimize the time that takes to integrate by the experiments the changes and bug fixes in libraries provided by the AA
- ◆ Many developments done in the PROOF system as the result of the serious testing done by ALICE
- ◆ Summer production releases of ROOT and Geant4

Internal Review 2005 Concerns and Recommendations



Main Conclusions by 2005 Review

- ◆ The proposed evolution plan for the Application Area technically reasonable and supported by all experiments
 - It should allow integrating ROOT activity in LCG organization
 - Coherent set of products to users
 - Facilitate the long term support
- ◆ A number of recommendations on how to perform the merge of SEAL and ROOT
 - Backward compatibility maintained, lightweight packaging, minimizing dependencies, etc.
- ◆ The technical details of the plan should continue to be discussed and approved by the Architects Forum
 - LHC experiments should set schedule and priorities

Recommendations to SPI

- ◆ Direct participation in projects strongly encouraged
- ◆ Tools developed for SPI should be packaged for general use
- ◆ Doxygen and Savannah
 - Automation, cross-references, common practices, ...
 - No CERN/LCG resources dedicated to alternative systems
- ◆ External software
 - Clear procedures for selecting, providing, supporting external software
- ◆ Training
 - Should be responsibility of SPI
 - Highly successful Python course must be continued

Recommendations to POOL

- ◆ General concern about impact of SEAL+ROOT merge
 - Will generate additional workload which must be anticipated in the planning
- ◆ Documentation greatly improved but problems remain
- ◆ Welcome split of RAL and POOL release cycles
- ◆ Error handling and reporting need to be improved
- ◆ Confusion in the collections, not clear what is really needed
- ◆ Ability to follow a `pool::Ref<T>` from a interactive ROOT session is needed
- ◆ COOL: Experiments interested in COOL are encouraged to commit more manpower
- ◆ POOL and security: reuse solutions developed in Grid community

Recommendations for SIMULATION

- ◆ GENSER
 - Recommended more granular packaging
 - Concerns about HepMC. Make sure that HepMC is well supported.
- ◆ Physics Validation
 - The coming decrease of manpower is very worrying
 - The active participation from experiment should continue
 - Re-visit with the experiments the effects of present simulation uncertainty
- ◆ Geant4
 - Any concerns on the validity of the physics results should be brought forward in the validation subproject
- ◆ Fluka
 - Should be directly usable via geometry conversion mechanisms
 - Installation in SPI external software area
- ◆ Simulation Framework
 - Further development of GDML is encouraged
 - Encourage exchange of experience with Python interfaces

Selected Issues to be Reviewed

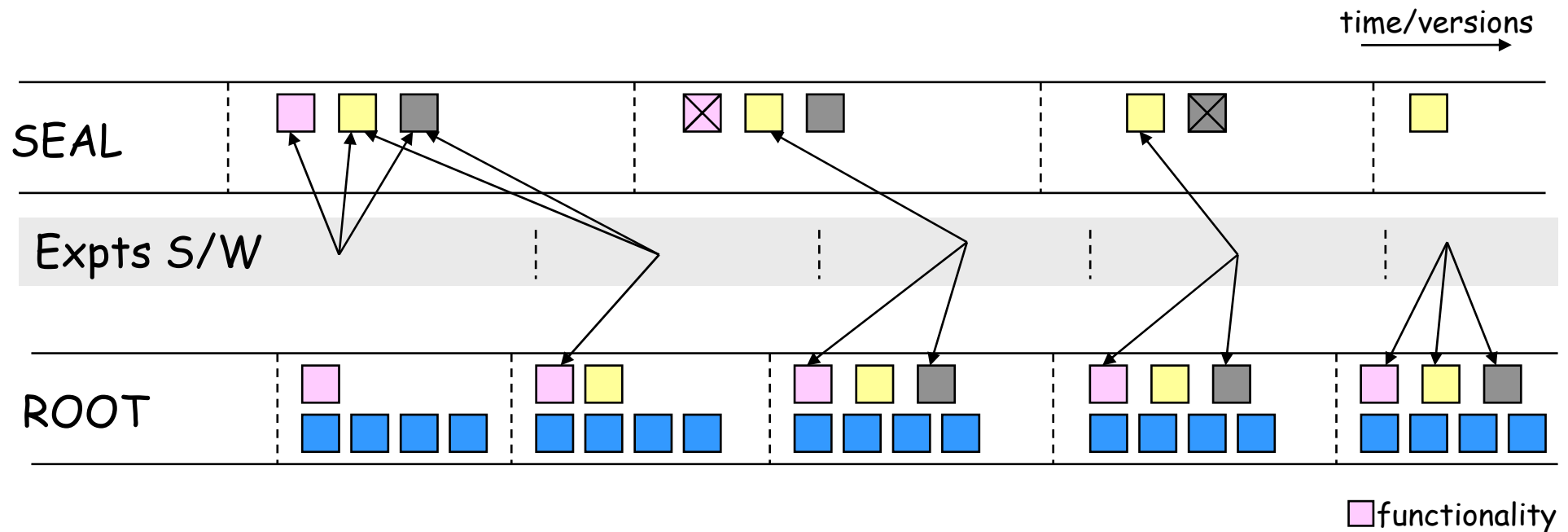
- ◆ SEAL and ROOT merge
- ◆ Frequent Releases
- ◆ Build System Evolution
- ◆ Interaction with Middleware/Fabric software
- ◆ Generator Service
- ◆ Level of coherency of AA projects

SEAL and ROOT Merge

The merge consisted of two different parts:

1. Merge of the development teams into a single team
 - done immediately at the time of the review
2. Evolution of the software products into a single set of core software libraries
 - originally foreseen to be completed within a very aggressive timescale but later adapted to allow more time for the experiments to migrate their software
 - SEAL work package added to the ROOT project (Summer 05)

Smooth transition for Experiments



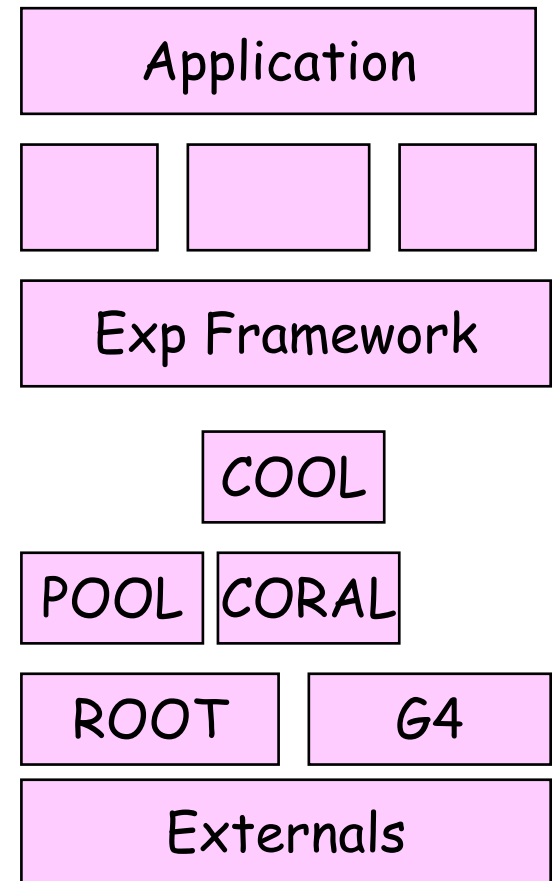
- ◆ Adiabatic changes towards the experiments
- ◆ SEAL functionality will be maintained as long as the experiments require

Migrating the SEAL Functionality

- ◆ LHCC recommended (Nov 2005)
 - "...it is important that agreement is reached for the contentious issues and that a strict set of milestones are set for the final conversion in the near future"
- ◆ It is also in the interest of the SEAL maintainers to shorten the time of the full migration
- ◆ Made a selection last year of what was more urgent for the experiments
 - Higher priority: Dictionaries, Math libraries, and Scripting
 - New functionality that didn't make it in SEAL (linear algebra)
- ◆ More than 50% of the functionality has migrated already
 - ROOT release in December (5.08) contained the new functionality
 - SEAL release in January (1.8.0) the duplicated functionality was removed
 - Experiments are basing their current software of these releases

Frequent releases enabling iterative feedback

- ◆ The AA/Experiments software stack is quite large and complex
- ◆ Only 2-3 production quality releases per year are affordable
 - Complete documentation, complete platform set, complete regression tests, test coverage, etc.
- ◆ Feedback is also required before the production release
 - No clear solution on how to achieve this
 - Semi-successful attempt last time
- ◆ As often as needed bug fix releases
 - Quick reaction time and minimal time to user (TTU)



Frequent Releases

- ◆ What counts is the TOTAL time between the bug being fixed in any of the AA products until the experiment application is running with the fix
 - Overall optimization is what is required
 - Releasing very fast one of the software packages does not help if the rest takes very long
 - In some cases "shortcuts" would be possible (overriding one or few libraries). "Service packs" à la Windows.
- ◆ Improving the time to release is one of the main goals for this year
 - "Automation" is a big part of the solution
 - Nightly builds?



Build and Configuration Tool

- ◆ Not all the projects in AA build their software using a common tool
 - Initial homogeneity not continued with the incorporation of existing projects (e.g. ROOT, Garfield,...)
 - In some cases difficulties in adapting the software to the required structure (e.g. some MC generators)
- ◆ The maintenance of SCRAM is uncertain
 - The migration to SCRAM v1 has been put on hold
- ◆ Do we need a single mechanism?
 - Towards the experiments is not important the tool used to build as long as is reliable, adaptable, fast, easy to use, etc.
- ◆ Proposed to migrate the projects using SCRAM to use CMT
 - Maintenance ensured (ATLAS and LHCb rely on it)
 - Quite a lot of experience in the projects already

Interfacing to Middleware/Fabric

- ◆ Application software has a number of contact points with middleware or fabric software
 - File catalogues, mass storage interface, etc.
 - Client libraries of these packages are needed for building the AA software
- ◆ A number of problems has been observed recently
 - These libraries needed to be part of the configuration
 - » E.g. Castor libraries installed in the external area and "controlled" by the configuration
 - Insufficient testing from our side
 - » It is not sufficient just to build our plugins, they need to be tested in all combinations
 - Information about what version to use not always clear
 - » Coordination body non existing, heterogeneity among computing centers, etc.

Generator Services

- ◆ Some of the leading authors of general purpose MC generators have expressed some strong concerns about the generator services project during the MC4LHC 2006 workshop
 - See more details in Alberto's presentation
 - This has triggered the resignation of Paolo Bartalini as project leader
- ◆ Received a strong message from the LHC experiments that there is a clear need for the project to continue under its original mandate
 - to provide services for well-maintained repositories of Monte Carlo generators on LCG-supported platforms
- ◆ The new project leader, Witek Pokorski, will prepare a new plan taking into account their concerns

Level of coherency of AA projects

- ◆ There are recent examples of work done in an AA project re-used or adopted in another
 - ROOT adopted savannah for bug tracking
 - GDML readers and writers integrated in ROOT
 - G4 example of storing and retrieving hits using ROOT
- ◆ --> but the integration remains slow
- ◆ I think projects should put more effort on this
 - Maintaining modularity and minimal dependencies while at the same time increasing the coherency
 - Minimizing overall duplication
 - Lowering cultural barriers

List of Milestones

SPI			
SPI-1	31.12.05	Provide the tools for generating CMT and SCRAM configurations from a common generic configuration description based on XML description files. Be able to update the web and distribution's kits from the same description.	Partial completion
SPI-2	28.02.06	Provide a web based "user discussion forum" service interfaced with Savannah. This new service should allow projects and experiments to easily setup and manage discussion subjects.	Done
SPI-3	31.03.06	Provide the interconnection/interoperability between the Savannah and HyperNews services.	Done
SPI-4	31.03.06	Generate CMT configuration and distribution kits from the common (XML based) configuration description.	Rescheduled 30.09.06
SPI-5	31.06.06	Provide the tools for the pre- and post-build procedures for the AA projects and externals through a web interface, such that it can be done or triggered by the project release managers.	Done
SPI-6	30.09.06	Move the build infrastructure of the LCG AA projects from scram version 0 to scram version 1. Provide the initial setup for the projects in collaboration with the experts on scram v1 from CMS.	New
SPI-7	30.09.06	Port the external packages and SEAL to the osx104_ppc_gcc401 platform (Mac OS X). Update the configurations to accommodate the new platform.	New
ROOT			
ROOT-1	30.09.05	Make available prototypes addressing different topics for the SEAL+ROOT merge (Math libraries, Dictionary libraries, etc.)	Completed
ROOT-2	30.09.05	Demonstration of the new the Parallel ROOT facility (PROOF) in a cluster of 32 CPU's provided by CERN/IT. This new version of the system should include asynchronous queries, GUI session controller, interactive batch mode.	Completed
ROOT-4	31.12.05	Finalization of the fitting and minimization application programming interfaces and integration of the new C++ implementation of Minuit in the ROOT release.	Completed
ROOT-5	31.03.06	The Python interface to ROOT (PyROOT) adapted to directly use the new C++ reflection library (Reflex).	Rescheduled to 31.03.07
ROOT-8	31.03.06	Have the rootcint dictionary code generator interfaced with the Reflex and gccxml options	Done
ROOT-6	30.04.06	The ROOT C++ interpreter (CINT) adapted to use the new C++ reflection library (Reflex). Applications will require a single dictionary with reflection information in memory. Backward compatibility will need to be provided to old ROOT and POOL applications.	Same as ROOT-9. Deleted
ROOT-9	30.06.06	First version of CINT running directly with the Reflex data structures as part of the ROOT June release	In progress. Rescheduled 31.12.06
ROOT-10	30.06.06	The new Fit GUI released as part of the ROOT June release.	In progress. Rescheduled 31.10.06
ROOT-7	31.10.06	Complete the merge of SEAL and ROOT functional components into a single set of libraries.	
ROOT-11	15.12.06	PROOF demonstrated in production in at least one of the LHC collaborations.	
ROOT-12	31.12.06	Speed-up I/O performance with remote files by eliminating as much by minimizing the number of network transactions.	
ROOT-13	31.12.06	Improvements in the PROOF system to support a realistic analysis environment for an experiment.	

POOL			
POOL-1	31.10.05	Production quality release of the relational database API (RAL) package, which should include the new interface recently reviewed.	Completed
POOL-2	31.12.05	POOL framework based on new C++ reflection libraries (Reflex) available for the experiments to be used in production. Validation by the experiments completed.	Completed
POOL-3	30.08.31	Finalize the migration POOL/CORAL to the new platforms (MacOSX, SLC4_amd64) with regular builds, and full running of the functional and data regression tests. Migration to scram v1	New
POOL-4	30.09.06	Development and deployment of LFC-based lookup and DB authentication services of CORAL	New
POOL-5	30.10.06	Complete migration to CORAL (AttributeList) and the SEAL component model of all POOL components	New
POOL-6	31.12.06	Make all CORAL components thread-safe.	New
COOL			
COOL-1	30.11.05	Conditions Database (COOL) release based on the latest version of RAL including bulk insertion operations and extended tagging functionality.	Delayed to 31.03.06
COOL-2	31.12.05	First prototypes of API and command line tools for data extraction and cross-population of COOL databases.	Completed
COOL-3	31.03.06	COOL overall performance study and validation of the experiments requirements. This study should identify the areas that will require further work and optimization.	
COOL-4	30.06.06	Support for multi-channel bulk insertion operations	In progress. Rescheduled 30.09.06
SIMU			
SIMU-1	15.12.05	Apply the Fluka-Geant4 (Flugg) geometry interface to one of the LHC calorimeter test-beam simulation.	Delayed to 31.10.06
SIMU-2	15.12.05	Production quality release of the MC generator level production framework.	Delayed and changed to SIMU-8
SIMU-3	20.12.05	New Geant4 public release including positron annihilation and geometry voxelisation improvements in addition to the regular bug fixes and small improvements included on each release.	Achieved
SIMU-4	31.12.05	First results of the ATLAS combined and 2004 test-beams data comparisons.	Delayed to 31.10.06
SIMU-5	31.03.06	Monte Carlo event generator files database (MCDB) publicly available and able to deal with large files.	Achieved
SIMU-8	30.06.06	New generator level production framework: beta release (from SIMU-2, due 15.12.05)	Achieved
SIMU-9	30.06.06	Investigation of correction for test-beam data for validation of stand-alone simulation engines (VD617)	Delayed to 30.09.06
SIMU-6	31.10.06	First release of a common framework for handling MC truth information to be used by experiment's simulation programs.	
SIMU-7	31.10.06	Validation of shower parameterization packages completed. The results of the validation should be summarized in a document.	
SIMU-10	30.06.07	Application of corrections of test-beam data, for validation of stand-alone simulation, to the LHC calorimeter test-beams (VD703)	
SIMU-11	31.12.06	Report on the main physics effects responsible for the hadronic shower development in Geant4 simulations (G4615)	



Milestones Remarks

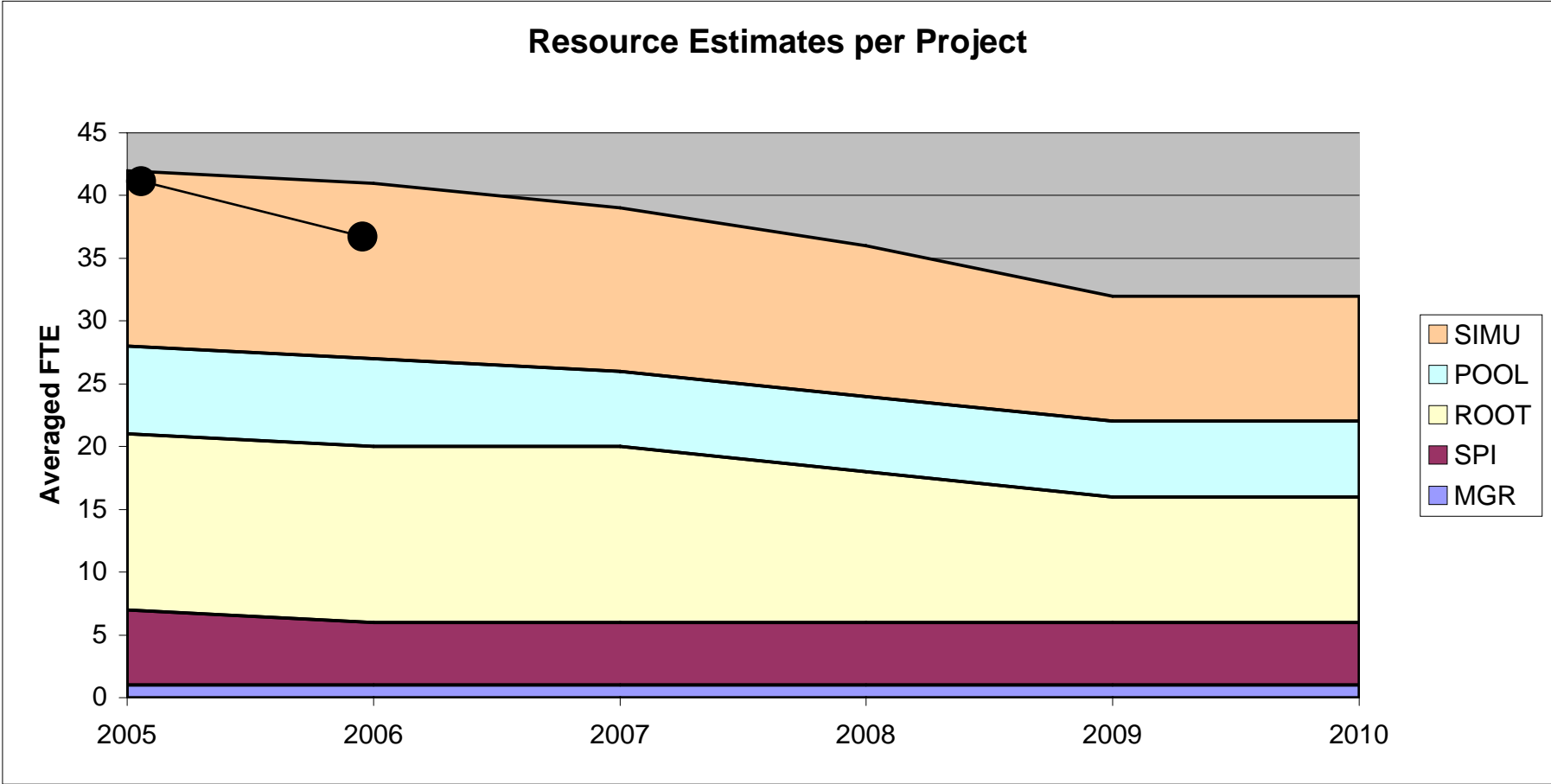
- ◆ A couple of high-level milestones per project per quarter is the ideal number of milestones
 - This is not always achieved
- ◆ For projects moving to a maintenance and consolidation phase it is difficult to define long term milestones
 - Bug fixes, improvements, evolutions, etc. are typically decided on the spot and not planned

Manpower (average 2006)

Sum of FTE(2006)		Affiliation												Grand Total	
Project	Sub-Project	ALICE	ATLAS	BNL	CERN	CMS	EGEE	FNAL	LCG-Russia	LCG-Spain	LHCb	Other	LCG-Italy		
MGR	(blank)				0.9										0.9
MGR Total					0.9										0.9
POOL	Catalog					0.1									0.1
	Collections		0.2												0.2
	Cool		0.6		0.9						0.2				1.7
	Mgr				0.5										0.5
	Ral				2.2										2.2
	StorageMgr		0.2								0.1				0.3
POOL Total			1		3.6	0.1					0.3				5
ROOT	Base	0.2			1			0.1							1.3
	Dictionary		0.2		1.7			0.4				0.2			2.5
	Geom	0.8													0.8
	Graf	0.1			1							0.8			1.9
	Gui			0.1	1.1										1.2
	I/O							0.5			0.1				0.6
	Math				1.4										1.4
	Mgr				1										1
	Proof				1.3							0.2			1.5
	Seal				0.2										0.2
ROOT Total		1.1	0.2	0.1	8.7			1			0.1	1.2			12.4
SIMU	Framework				1.5										1.5
	Garfield				1										1
	Geant4				6.45										6.45
	Genser				0.7	0.3		0.1	1.8	0.25			0.25		3.4
	Mgr				0.25										0.25
	Validation				2.3										2.3
SIMU Total					12.2	0.3		0.1	1.8	0.25			0.25		14.9
SPI	(blank)				2.6										2.6
SPI Total					2.6										2.6
Grand Total		1.1	1.2	0.1	28	0.4		1.1	1.8	0.25	0.4	1.2	0.25	35.8	



Resource Estimates



Staffing Remarks

- ◆ Reduction of manpower stronger than anticipated
 - 42 FTE -> 36 FTE (~6FTE)
- ◆ The projects most affected are:
 - SPI: 5 FTE -> 2.5 FTE (EGEE contribution + IT)
 - COOL: 2.6 FTE -> 1.7 FTE
 - ROOT: 14.5 FTE -> 12.5 FTE (1 Fellow + some fractions)
- ◆ Difficult to achieve again to the ideal level
 - De-scoping has been necessary in some cases (e.g. SPI)
 - Need to work out solutions case by case

Agenda for the Review

	Monday 18 September 2006		Tuesday 19 September 2006		Wednesday 20 September 2006	
AM	09:00	Monday Morning (32-1-A24) (until 12:30)	09:00	Tuesday morning (32-1-A24) (until 12:30)	09:00	Wednesday morning (32-1-A24) (until 13:30)
	09:00	Review committee discussion (Michael Doser)	09:00	Persistency Framework Status and Plans	09:00	Possible further questions from committee
	09:30	Applications Area Overview (Pere Mato) transparencies	11:00	— Break —	10:00	Review committee report writing
	10:30	— Break —	11:30	SIMULATION Status and Plans (part I)		
	11:00	SPI Status and Plans (Andreas Pfeiffer) transparencies				
PM	14:00	Monday afternoon (32-1-A24) (until 18:00)	14:00	Tuesday afternoon (32-1-A24) (until 18:00)	13:30	Wednesday afternoon (32-1-A24) (until 16:00)
	14:00	ROOT Status and Plans (Part I)	14:00	SIMULATION Status and Plans (part II)	14:00	Closeout. Committee feedback
	15:30	— Break —	15:00	Experiment Feedback		
	16:00	ROOT Status and Plans (part II)	16:00	— Break —		
			16:30	Experiment Feedback		
		17:30	Discussion			