Applications Area Overview

## LCG Comprehensive Review
25-26 September 2006

Pere Mato/CERN

Outline

- Applications Area Overview
- Current project status
- Milestones and Manpower
- AA Internal Review
- Conclusions

Applications Area Focus (a reminder)

- ◆ **Deliver the common physics applications software**
  - – The definition of "common software" is software that is used by at least two experiments
  - – Always allowed some level of speculation
- ◆ **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
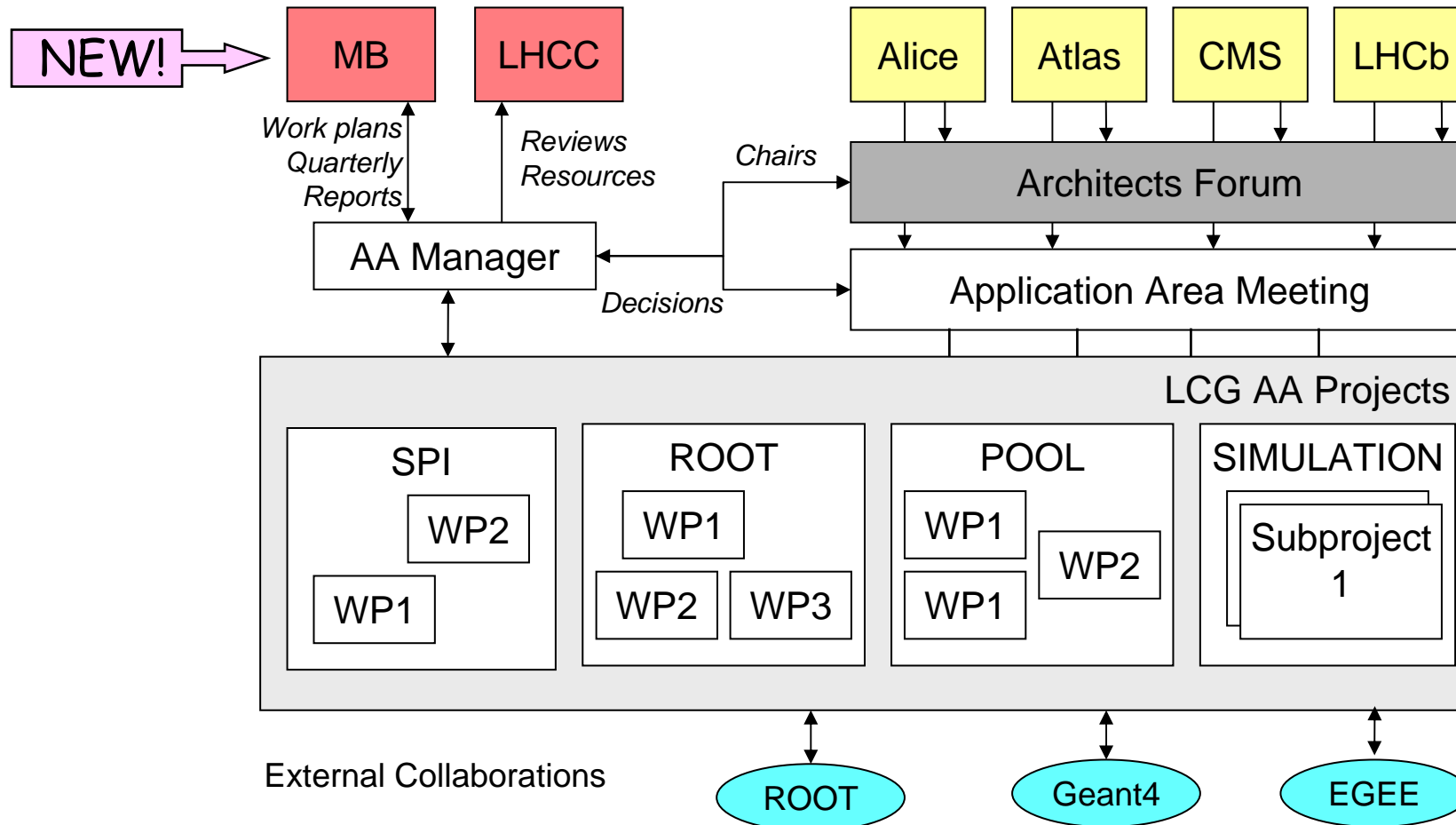
AA Projects

- **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 12 months

- Merge of SEAL and ROOT activities into a single project
- 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 (C++ MC Event format) 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

Migration of SEAL Functionality to ROOT

- ◆ 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 made 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

Software Process & Infrastructure

- ◆ **Shifted emphasis to serve experiments more than AA projects**
  - – Recommendation from AA internal review 2005
- ◆ **Andreas Pfeiffer replaced Alberto Aimar as SPI project leader**
- ◆ **Participation of people from experiments and projects**
  - – SPI provides a "frame" to do complex work (guidelines, AFS space, web space, etc.)
- ◆ **Coping with strong (50%) reduction of resources**
  - – Optimization opportunities
  - – Automation of procedures
  - – "Outsourcing" some activities

SPI Services Overview

- ◆ External Software
  - – **Common libraries from HEP, IT, …**
- ◆ Savannah Project Portal
  - – **http://savannah.cern.ch**
- ◆ Build and Distribution
  - – **LCG Software configuration management**
  - – **Builds and releases**
  - – **Software Distribution**
- ◆ Quality Assurance Service
  - – **Testing frameworks**
  - – **QA checklists and reports**
- ◆ Documentation and Training
  - – **Development of LCG policies, templates**
  - – **Code Documentation (doxygen, lxr)**
  - – **Documentation and LCG Workbook**

Evolution of Configurations

Configuration	LCG_40a	LCG_46e
Released on	Jan 2006	Aug 2006
# projects	5	5
# externals	35	39
# externals changed (new or new versions)	-	18 (6 new, 2 dropped)
Volume size (in GB, all platforms)	15.2	20.8

Savannah Project Portal

- ◆ **New export facility available since fall 2005**
  - – Enables project members to collect all the information related to tracker items of their choice

- ◆ **The Cookbook**
  - – Enables project members to build structured series of 'how to's related to the use of savannah by their project

- ◆ **Consolidation**
  - – 76 Bugs fixed
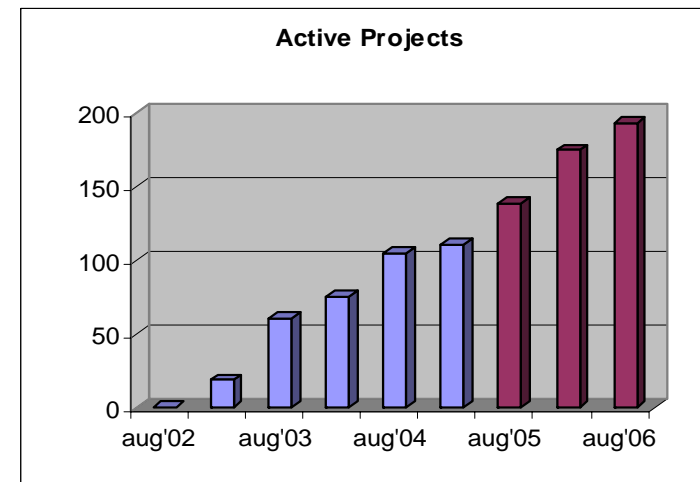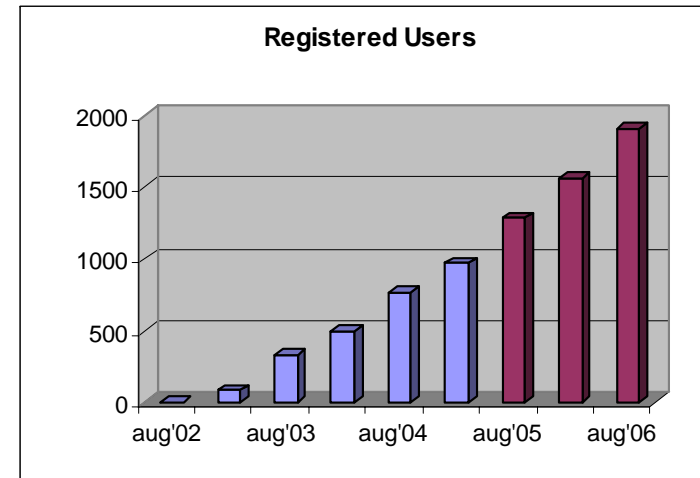 - – 70 Support requests handled

Savannah Status (users and projects)

1944 users

- 1083 at last review
- + 80%

196 projects

- 122 at last review
- + 60 %

Registered Users

Active Projects

Savannah Status (project activity)

◆ 17819
submitted bugs

◆ 3539
submitted tasks

Submitted Bugs

Bug average lifetime
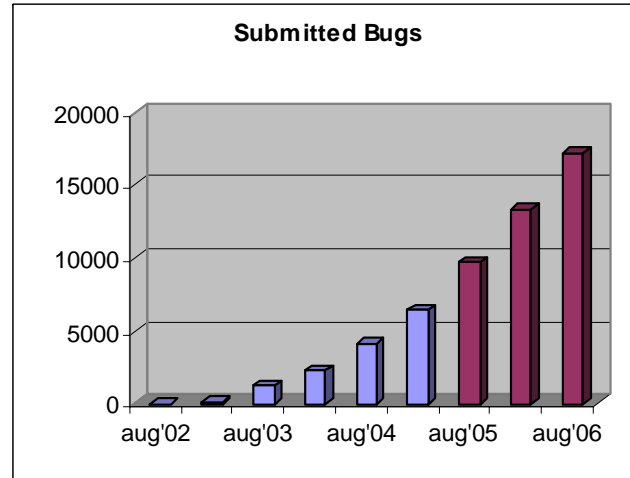
Submitted Tasks
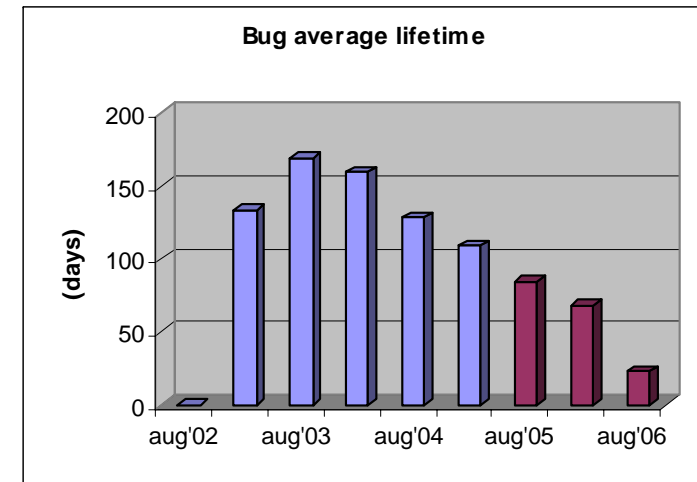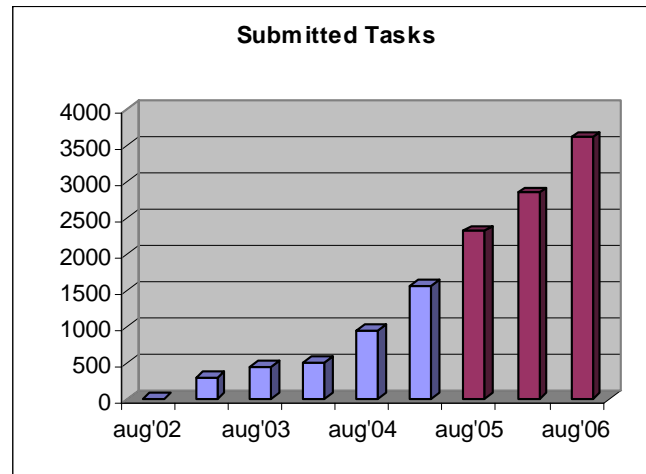
Frequent Releases Enabling Iterative Feedback

◆ 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)

Application

Exp Framework

COOL

POOL	CORAL

ROOT	G4
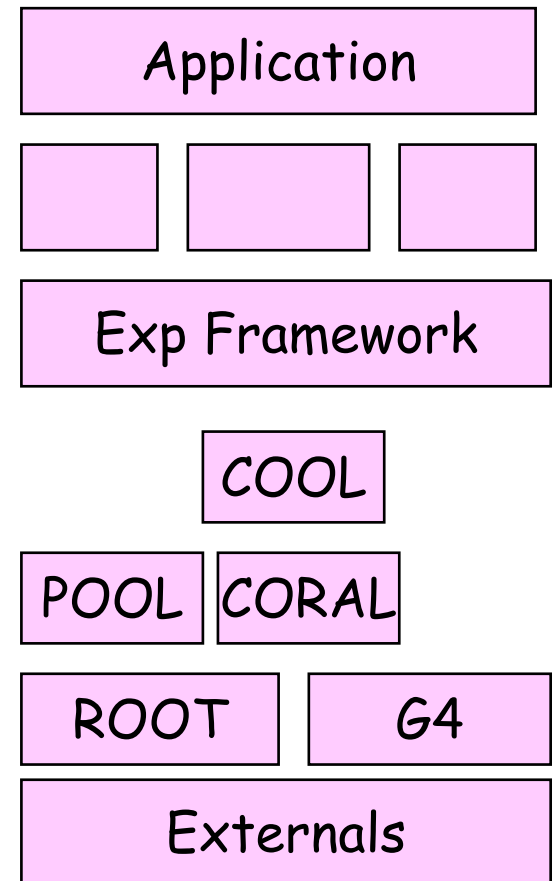
Externals

Frequent Releases

- What counts if the TOTAL time between the bug being fixed in any of the AA products or externals 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

- Improving the time to release is one of the main goals for this year
  - "Automation" is an important element for the solution
  - Adoption of "nightly builds" have been recommended

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.

MC 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 Ribon presentation at Internal Review
  - 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

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 accomodate 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.	Some 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 Table

Sum of FTE(2006)	Afiliation	ALICE	ATLAS	BNL	CERN	CMS	FNAL	LCG-Russia	LCG-Spain	LHCb	Other	LCG-Italy	Grand Total
Project	Sub-Project												
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			0.7		0.1						1
	Dictionary		0.2		1.7		0.3				0.2		2.4
	Geom	0.8											0.8
	Graf	0.1			1						0.8		1.9
	Gui			0.1	0.9								1
	I/O						1.3			0.1			1.4
	Math				1.4								1.4
	Mgr				1								1
	Proof				4.1						0.2		4.3
	Seal				0.2								0.2
ROOT Total		1.1	0.2	0.1	11		1.7			0.1	1.2		15.4
SIMU	Framework				1.5								1.5
	Garfield				1								1
	Geant4				7.2								7.2
	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					12.95	0.3	0.1	1.8	0.25			0.25	15.65
SPI	(blank)				2.6								2.6
SPI Total					2.6								2.6
Grand Total		1.1	1.2	0.1	31.05	0.4	1.8	1.8	0.25	0.4	1.2	0.25	39.55

Estimated Resource Needs

Resource Estimates per Project
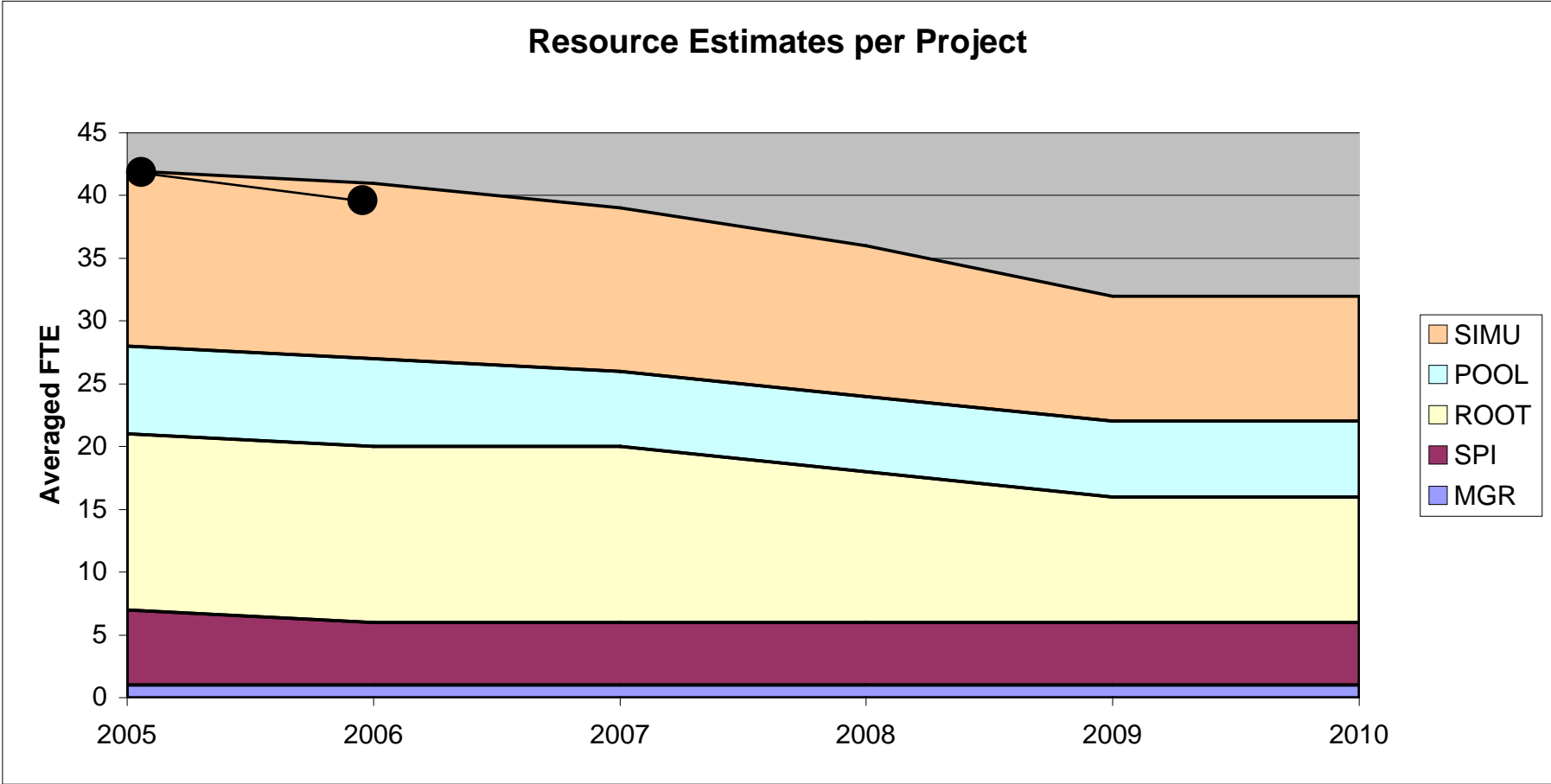
Staffing Remarks

- **Reduction of manpower as anticipated for 2006**
  - 42 FTE -> 39 FTE (~3FTE)
  - But perhaps not in the projects or work packages expected

- **The projects most affected by the reduction**
  - SPI: 5 FTE -> 2.5 FTE (EGEE contribution + IT)
  - COOL: 2.6 FTE -> 1.7 FTE

- **Project affected positively**
  - PROOF: 2.6 FTE -> 4.3 FTE (DOCT, MC Fellow)

- **Major reduction is expected in 2008**
  - End LD contracts + retirement. No solution yet

Applications Area Internal Review

◆ **The AA interal review took place last week**
- http://agenda.cern.ch/fullAgenda.php?ida=a063165
- Status and Plans for all projects and LHC experiment's feedback

◆ **Mandate**

- Technical review of the software and infrastructure of the Applications Area

- Follow-up of the AA internal review that took place in April 2005

- Assess the readiness of the Common Application Software for the LHC startup

◆ **Reviewers**

- Michael Doser (CERN, chair), Predrag Buncic (CERN), Paolo Calafiura (LBNL), Marco Clemencic (CERN), Sunanda Banerjee (TIFR), Gerhard Raven (NIKHEF), Maya Stavrianakou (FNAL), Liz Sexton-Kennedy (FNAL), Vakho Tsulaia (U. Pittsburgh)

Preliminary Conclusions (Communication)

◆ Generally very good communication between AA and experiments via AF; however, several stake holders are missing (middleware/fabric, Grid developers in particular) and should be involved

◆ Similarly, coordination of deployment of site-specific code (dCache, Castor) should take place at the same forum

◆ Good release schedules

◆ Stability of the environment, and predictability are important for the experiments; going to nightly builds (requested by several experiments) could further improve this

◆ Input from the experiments in strategic choices (prioritization of (major) bug fixes and requested features) should be more actively pursued

Preliminary Conclusions (Manpower)

- There is a general decrease of manpower for the whole AA, with more severe decreases in individual sub-projects, some of which have reached a critical status (COOL, SPI)

- At the same time, several projects are approaching maturity (CORAL), while others have not yet reached full production status (PROOF, CINT/Reflex)

- Optimization within sub-projects, between projects and with external collaborators will be essential to deal with this reduction

- Experiments should be implicated (via AF) in the discussion about the choices to be made, also in case de-scoping of certain features would be required, at as early a time as possible

Summary

- ◆ **The organization of the Applications Area is mature and works reasonable well**
  - – Detailed planning (e.g. scheduling bug fix releases, configuration changes, etc.) discussed and agreed at Architects Forum meetings
- ◆ **Need to coordinate software releases with other areas (middleware, fabric)**
- ◆ **The projects in AA has made substantial progress in many aspects**
  - – See next presentations
- ◆ **Reduction of manpower as anticipated for this year**
  - – However major reduction expected by 2008
- ◆ **The AA Internal Review took place last week**
  - – An opportunity take stock of the status of all projects and the future plans
  - – Expected final report in couple of weeks