LCG Application Areas

- Introduction - Planning - Manpower — Torre Wenaus
- Simulation — Gabriele Cosmo
  - Generic simulation framework / participation in Geant 4
  - Fluka Integration / Physics validation of simulation
  - Monte Carlo generator services
- SEAL and PI — Pere Mato
  - Foundation Libraries / Math Libraries / Framework Services
  - Dictionary Services / Python Services
- Persistency Framework — Dirk Duellmann
  - Object Persistency (POOL) / Conditions Database
- SPI — Alberto Aimar
  - Software Libraries / QA & Testing / Web & Documentation

Applications Area Organisation

Torre Wenaus

Architects Forum

Decisions

Strategy

Apps Area Manager

Librarians Integration Meeting

User ~~Provider~~

Peer to Peer

ROOT

SPI Project

POOL Project

SEAL Project

PI Project

Simulation Project

Consultation

Applications Area Meeting

Communication

Upcoming Level 1 Milestones

- 2004/12/31 - Geant4 validation in LHC production
  - Post-DC assessment of Geant4 by experiments
- 2004/12/31 - Work and resource plan through 2008
  - Long term support and development program
  - Scoped to available manpower and informed by DCs
  - In view of still-uncertain phase 2 resources, leadership change and attendant reorganization, **we propose a new date: 2005/3/31**
    - Use Feb internal review to expose plans & iterate
- 2005/9/30 - Phase 1 software complete & deployed
  - Full functionality available and successfully used
  - POOL, SEAL, ROOT, G4, CondDB, GENSER, ...
  - Union of specific L2 technical, validation milestones defined in previous MS

Outstanding Manpower Needs during 2005

Outstanding manpower needs during 2005, including CERN positions and experiment participation for which we don't yet have firm commitments:

- SPI: Several staff leaving at end 2005. Steady state SPI support must be maintained. ~1 CERN FTE missing.
- SEAL: 1.5-2 CERN FTEs for remaining development, experiment integration support, ROOT convergence
- Persistency: 2 CERN FTEs needed for core development, support. 1 FTE per participating experiment needed, particularly on ROOT storage manager and conditions DB
- Simu: 1.5 physicist FTEs for generator services, Geant4 hadronic physics, simulation physics validation
- ROOT: 2 CERN FTEs for GUI development, documentation and user guide, PROOF and grid integration

These are included in the full manpower profiles that are the basis of our planning...

Total Staffing Requirements

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

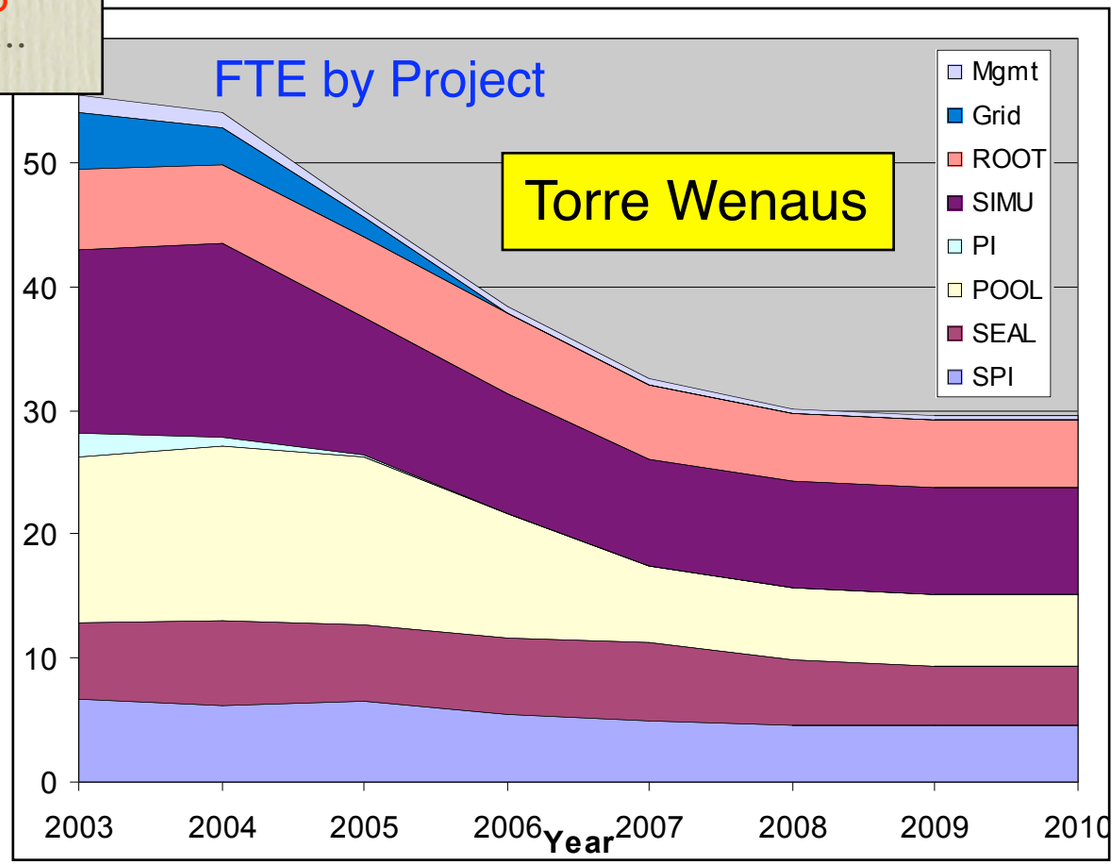
	2004	2005	2006	2007	2008
SPI	6.2	6.2	5.5	4.6	3.9
SEAL	6.8	5.9	6.2	5.7	4.7
Persistency	14.1	14.4	12.8	8.5	7.5
PI	0.7	0.1	0	0	0
Simulation	15.6	10.5	9.7	7.9	7.5
ROOT	6.5	6.2	6.4	5.5	4.8
Total	49.9	43.3	40.6	32.1	28.4
Contribution from Experiments:	10.6	7	6.5	6	5

If CERN and experiment commitments materialize, these needs will be met...

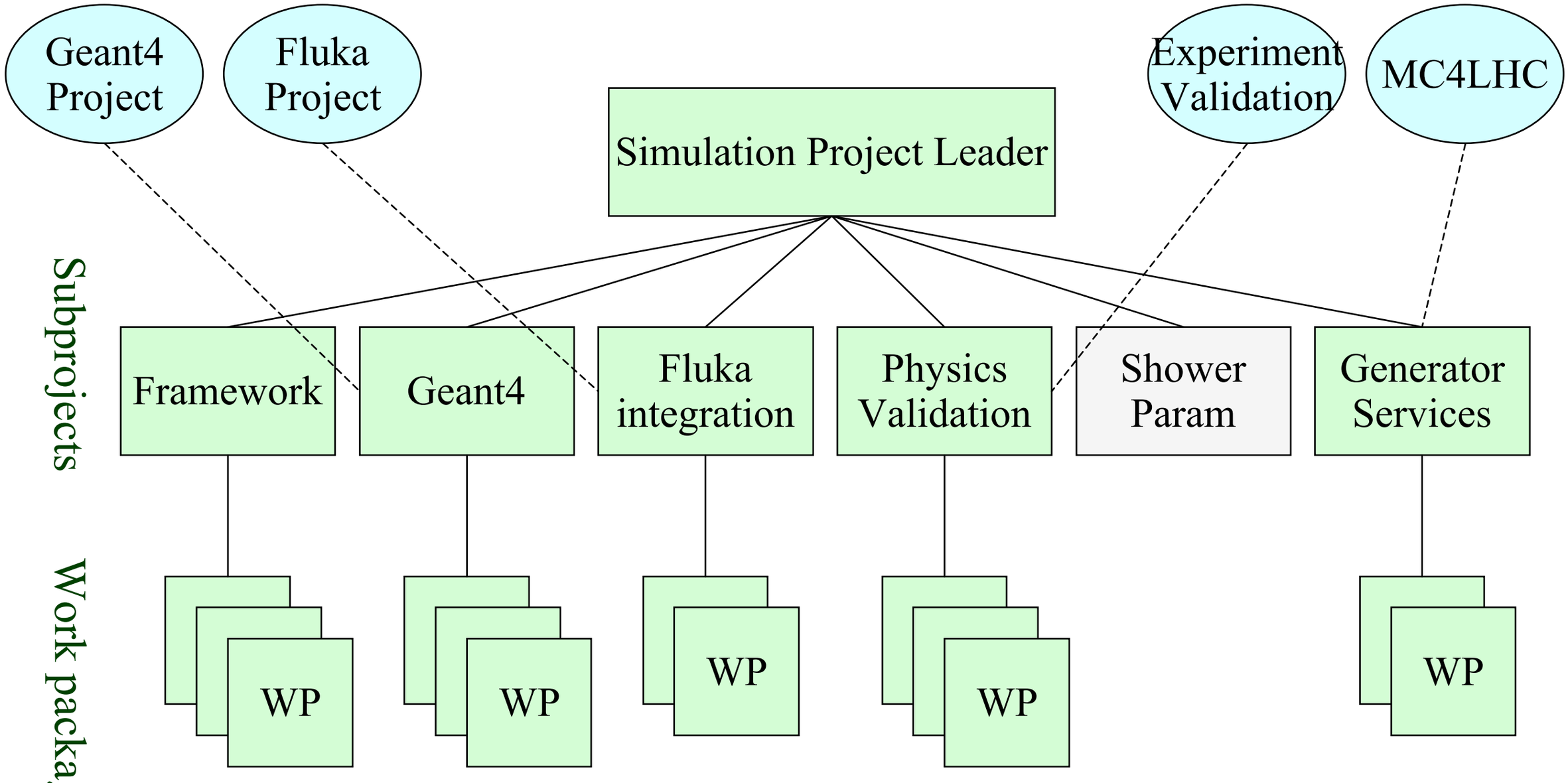
CERN-C-RRB-2004-114:
outlines need for additional CERN manpower after June 05

- expert 2 FTE
- experienced 2 FTE
- junior 5 FTE

- significant decrease of manpower
  - development → maintenance
- anticipated profile can be met only if
  - "soft statements" from experiments turn into hard committments
  - unfunded CERN positions are filled



Simulation



quaterly reviews

Simulation: Project Leaders

- **Simulation framework**
  - Interface to multiple simulation engines (Geant4, Fluka) and geometry models exchange

Witold Pokorski

- **Geant4** team participating
  - Aligned with and responding to needs from LHC experiments, physics validation, simulation framework

John Apostolakis

- **Fluka** team participating
  - Framework integration, physics validation

Alfredo Ferrari

- Simulation **physics validation** subproject
  - Assess adequacy of simulation and physics environment for LHC and provide the feedback to drive needed improvements

Alberto Ribon

- **Generator services** subproject
  - Generator librarian; common event files; validation/test suite; development when needed (HEPMC, etc.)

Paolo Bartalini

Simulation: Geant4

- Three LHC experiments (ATLAS, CMS, LHCb) now using it successfully in production
  - OSCAR (CMS), Gauss (LHCb) and ATLAS's Geant4-based simulation programs are the production tools
    - OSCAR and Gauss have replaced G3 based simulation
  - Substantial productions (numbers from Oct 20th Application Area Meeting presentations)
    - ATLAS DC2 (summer 2004) produced 12M events
    - Oscar (CMS) : 35 M pp interaction events, and first 100 Pb-Pb events
    - Gauss (LHCb): Over 200 M events simulated
  - In production use demonstrated low crash rate (and decreasing with new releases)
    - Rate decreasing from 1/10K events (5.2, CMS) to 1 per Million events (6.1, LHCb)
      - G4 team addressed issues found in test productions
- The Geant4 LCG/SI sub-project and the Geant4 Collaboration
  - LCG/SI/G4 responsible for CERN/LHC participation in Geant4
  - Work plan integrated with overall Geant4 plan
    - Geometry and tracking in field, Physics: hadronic and electromagnetic, testing and release, coordination
- Collaborating closely on validation, infrastructure and robustness
  - With Physics Valid. on validation, and with Framework on geometry exchange
  - Worked on shared infrastructure (testing, portal) with SPI: bonsai, Savannah
  - With experiment simulation and physics teams on robustness, integration and validation

Simulation: Geant4 Potential Future Goals

- Provide continued support and maintenance
  - Address queries on physics, geometry, tracking issues
    - Solve problems and respond to other issues / reports
  - Sustain existing physics use cases and enable emerging LHC-related uses
- Enhance physics
  - Address requirements for sub-1% EM stability and precision
  - Address precision needs for combined calorimetry in full detector simul. / LHC exp.
  - Address issues of pre-calibration using Monte Carlo (ATLAS)
  - Further enable use of radiation studies, addressing needs for data and physics lists
- Improve and extend 'automated' release validation
  - Extend and refine comparisons
    - Within available computing resources
- Address requirements for geometry and improved persistency
  - Improvements to Boolean solids, to address robustness issues
  - Enable exchange format (GDML) and 'direct' persistency (POOL / ROOT) for Geant4
- Address new requirements for radiation applications
  - Improve 'standard' tallying and extend with additional observables
- Continue to improve CPU performance
  - Address hot spots identified in 2004 and bring new tools into use

Simulation: Fluka

- Fluka is an official joint CERN-INFN project since January 2004
  - ☞ Fluka development is not an LCG activity

Current developments

- Physics
  - Heavy ions transport and interactions
    - DPMJET-III interface — done
    - Electromagnetic Dissociation — done
    - Special effects in heavy ion Coulomb interactions — in progress
    - Development of QMD models — in progress
  - Fragmentation
- Code structure and user interface (pre-condition for the full release) — in progress
  - Integration of PEMF functionalities in the run time code
  - Name-oriented advanced input interface
  - General clean-up
  - Publication of code documentation as CERN yellow report

Man power issues

- Physics developments: covered by A.Ferrari and non-CERN collaborators OK
- "Complex benchmarks" (test-beams): expertise support from FLUKA team, manpower must come from LCG ?
- Code structure and user interface (release oriented) : CERN-AB manpower absorbed by urgent LHC tasks. CERN-PH ending soon critical

Struggling to keep the December-January deadline for the public β release of the full code

Simulation: Physics Validation Status

- Geant4 electromagnetic physics validated at percent level
- Simulation physics requirements revisited
- First round of hadronic physics validation has been completed, with good results
  - For the observables, in the case of the simple benchmarks (pixels, neutron double differential, pion absorption) there is a reasonable agreement between data and both Geant4 and Fluka
  - For the calorimeter test-beams, Geant4 describes well the pion energy resolution, σ/E, and the ratio e/π
  - The shape of hadronic showers needs some improvement
- LCG notes:
  1. F.Gianotti et al., CERN-LCG-APP-2004-02
  2. A.Ribon, CERN-LCG-APP-2004-09
  3. F.Gianotti et al., CERN-LCG-APP-2004-10
  4. W.Pokorski, to be released very soon
- Monthly meetings presenting and coordinating experiment and project work
- Information, results gathering on web page

Simulation: Physics Validation Future

- Validate precision of Geant4 electromagnetic physics at the *permil* level
- More detailed studies of hadronic shower profiles, both at the simulation and experimental (test-beam data) level.
- Evaluate the possibility to study another simple benchmark, relevant to LHC, to validate both Geant4 and Fluka.
  - Man power could be a problem here!
- Complete validation for Fluka in the calorimeter test-beam validations (as it has been done for Geant4)
  - The adoption of FLUGG and strong interaction with the Simulation Framework project is required here
- Geant4 studies of background radiation in the LHC caverns in progress
  - Will be soon compared with Fluka.
- Man-power:
  - People from the experiments are busy with the new test-beam data
  - From LCG: M. Gallas, W. Pokorski, A. Ribon
    - all of them involved in other activities!
- First version of simulation test and benchmark suite delayed to end 2004
- Physics validation document delayed to June 2005

Simulation: Summary

- Striking success of Geant4 in the LHC experiments
  - ATLAS, CMS and LHCb now using it in full production
    - Crash-rate close to zero !
  - Strict collaboration with Geant4 team for support and new requirements
- Very active program in physics validation delivering results and conclusions
  - Shift of manpower to experiment-specific validation activities (ATLAS, LHCb)
  - First cycle of hadronic physics validation completed
  - New program of work for test-beam validation and simple benchmarks to be defined, also according to the available manpower in the project
- Outstanding progress for generator services
  - First production release of GENSER in December
    - Already used in production in ATLAS, under validation in CMS and LHCb
    - Impressive collection of validated generator packages with common event DB
  - Manpower issues partially addressed
- Simulation framework program scoped to make maximal use of existing software and meet the reality of minimal available manpower
  - Prototype setup based on Flugg for physics validation
  - Extension of GDML and implementation of interfaces for Geant4/Root to import and export geometry descriptions
  - Investigations and prototyping on Python-based interfaces
- Fluka integration/validation feeling effects of slow progress due to low manpower (physics validation, Flugg framework)
  - Expected public source-code release of Fluka soon

SEAL: Core Libraries and Services Project

The SEAL project provides the software infrastructure, basic frameworks, libraries and tools that are common among the LHC experiments. The project addresses the selection, integration, development and support of foundation and utility class libraries. The project is subdivided into the following tasks:

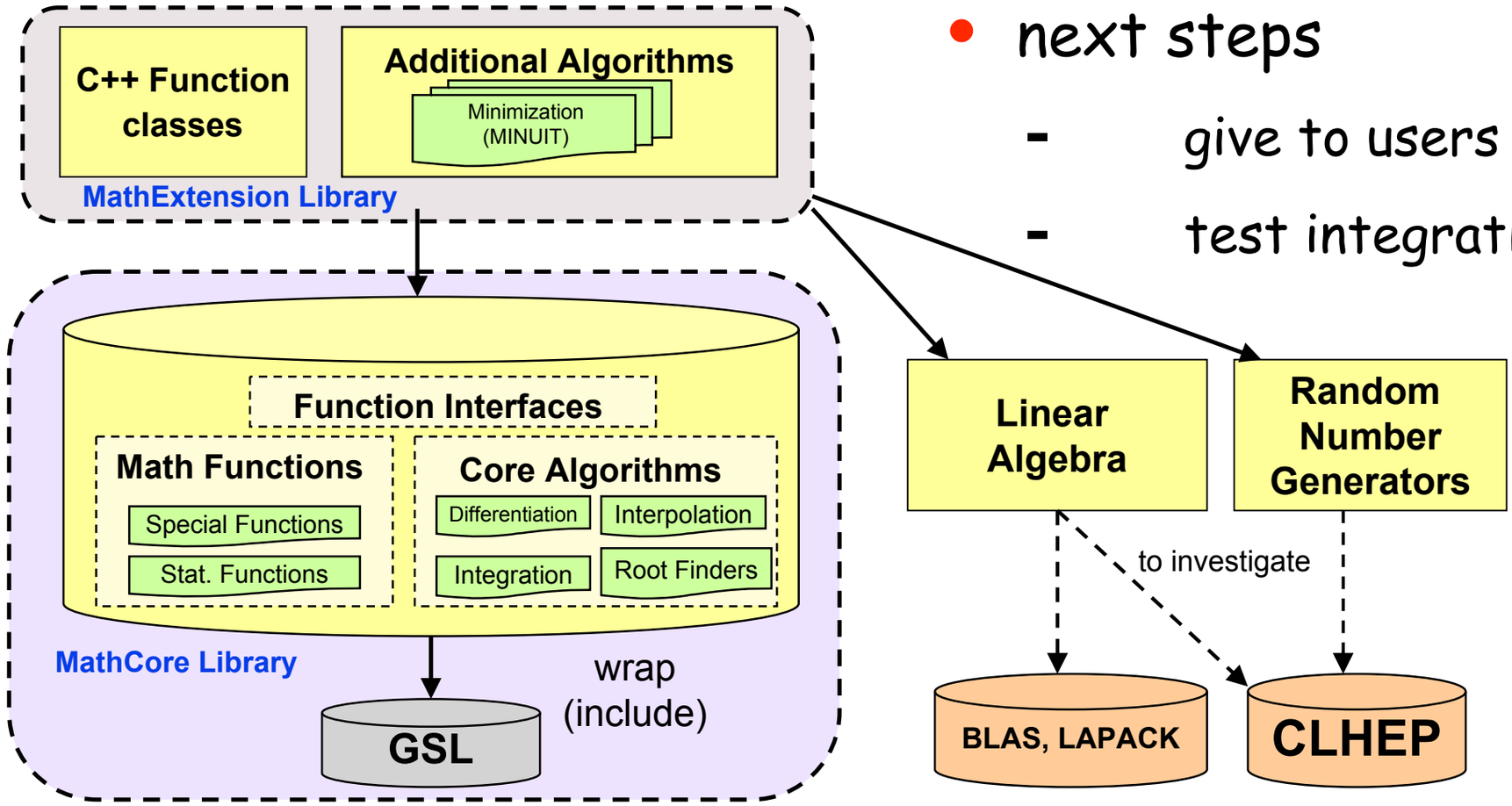
Foundation	Foundation and Utility Libraries and Plug-in Manager
MathLibs	Math Libraries Support and Coordination
Dictionary	LCG Object Dictionary
Framework	Component Model and Basic Framework services
Scripting	Scripting Services
Documentation	Education and Documentation

SEAL: Foundation

- Provide support for foundation, utility and operating system isolation libraries
- Inventory of existing libraries
  - http://seal.cern.ch/components.html
- Main external library: Boost
  - Being adopted by experiments and LCG AA projects
- Developed SEAL utility and system libraries complementary to Boost and STL from existing code
  - SealBase - library containing a large variety of utility classes
  - SealIOTools - library containing utility classes for stream oriented I/O
  - SealZip - library for compression I/O and producing archive files
- Plugin Manager
  - Basic concept: advanced object factory, dynamic loading of plugins
  - Two simple interfaces: object instantiation, plug-in provider
- Not foreseen major new developments in this work package
  - The work needed is mainly to educate users by developing and setting up tutorials, user-guides and coaching developers
- Entering maintenance phase
  - Additional functionality only on direct demand
  - Further reduction of unnecessary dependencies in external packages

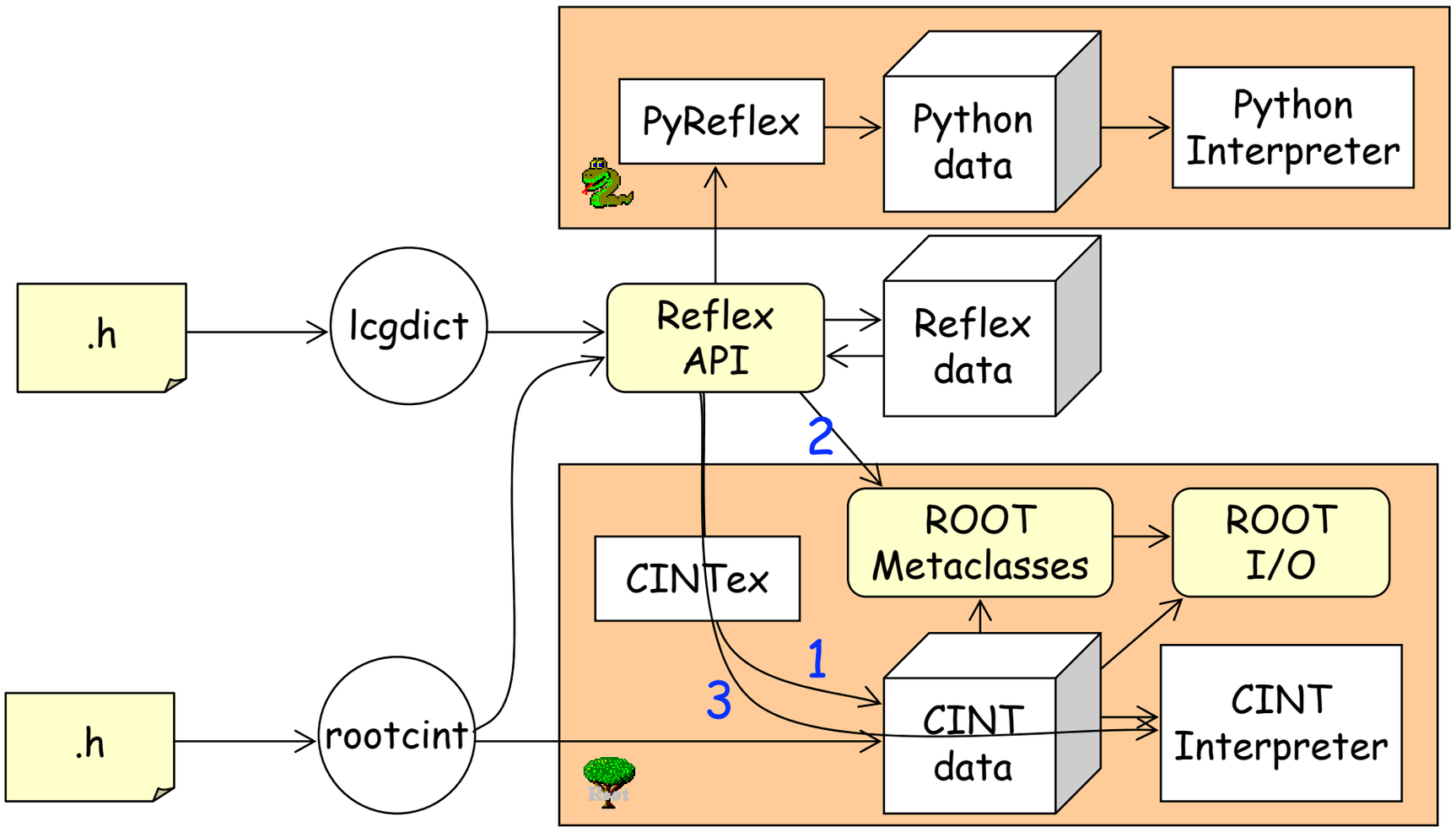
SEAL: Math Libs

- provide coherent set to end-users
- avoid duplication
- use same core library in all environments
- selected functionality (math only)
- have developed C++ MathCore prototype
  - based on GSL
- next steps
  - give to users to get feedback
  - test integration with ROOT



SEAL: Dictionary

- provide C++ reflection system
  - essential for providing generic object persistency and interactivity
- aim: converge with single dictionary service for all s/w systems
  - in particular LCG and ROOT
- implementation foreseen in 3 steps



SEAL: Framework

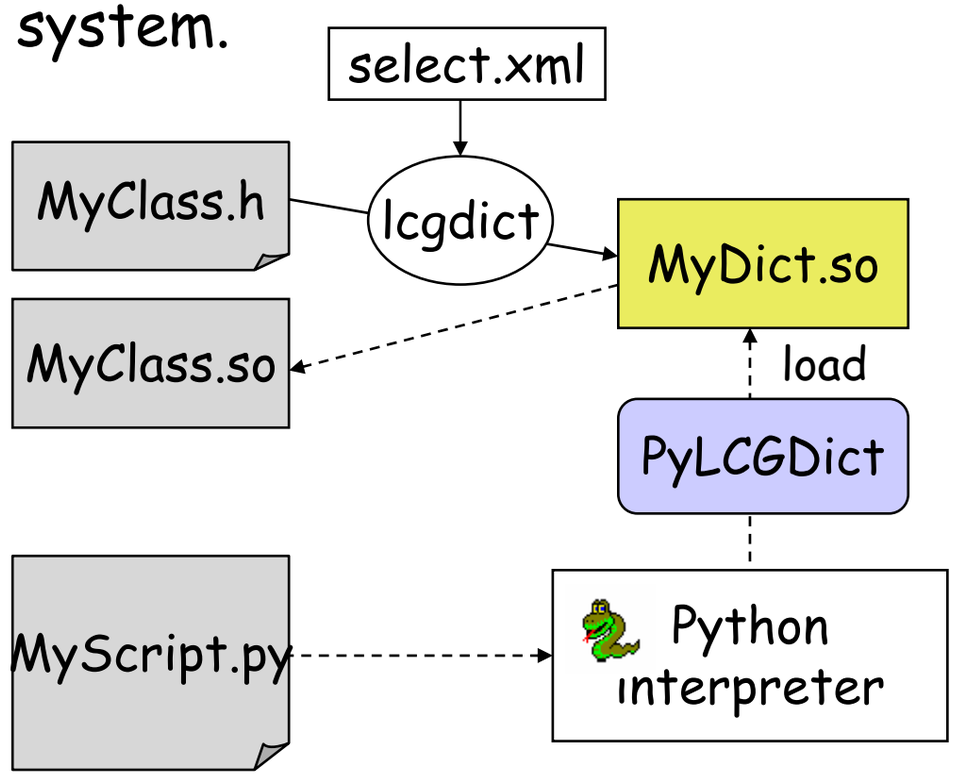
- Aims to provide a component model and a number of basic services
  - A component model is a basic set of mechanisms and base classes
- First set of basic services came with new component model
  - Application / Message Service / Configuration Service
- Framework Status:
  - The objective has been to integrate SEAL component model into the existing Gaudi/Athena framework and evaluate its costs and benefits
    - Not yet done due to lack of manpower (LHCb & ATLAS)
    - The development of new services was put on hold until there is a firm commitment from at least 2 experiments

still in workplan !

  - The new developments in POOL are exploiting the use of SEAL component model to implement true components and services together with the component loader
    - POOL Relational Abstraction Layer set of components

SEAL: Scripting & Python

- Scripting allows rapid application development to produce quick prototypes, which are essential in physics analysis and the ability to integrate heterogeneous software components produced independently into a coherent application ("component bus").
- Two scripting languages under consideration:
  - Python and C++ (CINT).
- Evaluation of the existing technologies for developing Python bindings (Python extension modules) of C++ classes.
- Development of a module that enables interoperability between CIPython bindings using the C++ reflection system.
  - PyROOT
    - Provides access to ROOT functionality from Python
    - Uses ROOT/CINT dictionary
    - Removed any external dependency and performance improvement
    - Integrated and distributed with ROOT
  - PyLCGDict
    - Access to C++ libraries from Python
    - Uses LCG dictionary. Automatically generates Python proxies for C++ objects dynamically.
    - Mapping C++ constructs to Python natural constructs



SEAL: Products and their Usage

		ATLAS	Alice	CMS	LHCb	Other
Foundation	SealBase					
	SealZip					
	SealIOTools					
	PluginManager					
Framework	Component Model					
	Basic Services					
Dictionary	Reflection					
	Lcgdict tool					
	Specific Dictionaries					
Scripting	PyROOT					
	PyLCGDict					
	PyBus					
MathLibs	Minuit					

In use directly In use indirectly Planned use

SEAL: 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
- 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

SEAL & PI: Summary

- SEAL has delivered a number of components that constitutes the basic foundation and utility libraries and object dictionary
  - Most of the delivered components are already in use or being tested and planned to be in use by LHC experiments
- Ongoing development in two areas mainly:
  - New C++ Reflection system with the goal to achieve a single dictionary between ROOT and LCG applications
  - Coherent set of common Mathematical Libraries
- PI development is basically finished, entering maintenance phase
- Advocating scripting based on Python
  - Peaceful coexistence between CINT and Python
  - Powerful tools have been developed (PyROOT, PyLCGDict)
  - Feedback from early adopters is encouraging (LHCb, ATLAS)

LCG Persistency Framework

- The LCG persistency framework project consists of two parts
  - Common project with CERN IT and strong experiment involvement
- POOL
  - Hybrid object persistency integration object streaming (using ROOT I/O for event data) with Relational Database technology (for meta data and collections)
  - Established baseline for three LHC experiments
  - Has been successfully integrated into the software frameworks of ATLAS, CMS and LHCb
  - Successfully deployed in three large scale data challenges

in total 370 TB of POOL data produced

- Conditions Database
  - Conditions DB was moved into the scope of the LCG project
    - To consolidate different independent developments
    - and integrate with other LCG components (SEAL, POOL)
  - Should share storage of complex objects into Root I/O and RDBMS backend with POOL

Persistency: POOL Developments 2004

- Move to ROOT4 (POOL2.0 Line)
  - To take advantage of automatic schema evolution and simplified streaming of STL containers
    - Need to insure backward compatibility for POOL 1.x files
  - Currently undergoing validation by the experiments
    - Will release two branches until POOL 2 is fully certified
- File Catalog deployment issues
  - DC productions showed some weaknesses of grid catalog implementations
    - Several new/enhanced catalogs coming up
    - Changes in the experiment computing models need to be taken into account
  - POOL tries to generalise from specific implementations and provides an open interface to accommodate upcoming components
- Collections
  - Several implementations of POOL collections exist
  - Collection cataloguing has been added in response to experiment requests
    - Similar to file catalogs
    - re-use of catalog implementation and commandline tools
  - Experiment analysis models are still being concretized
  - Expect experience from concrete analysis challenges

Persistency: Conditions Database

- **Project launched in summer 2003 (within LCG Persistency Framework)**
  - Background in 2000-2003:
    - C++ API definition and Objectivity implementation (Application Programming Interface)
    - Oracle implementation of the original API ("BLOB" data payload)
    - API extensions and MySQL implementation (user-defined relational data payload)
  - Two goals for the common project:
    - Integrate the existing Oracle and MySQL packages into LCG Application Area
    - Coordinate new development of API, software and tools
- **Status overview**
  - Kick-off workshop at CERN in December 2003
  - Activity along two directions in parallel
    - Integrate the existing software into LCG Application Area
    - Review two APIs and implementations, coordinate discussion about new developments
  - *Main problem so far: lack of committed manpower*
    - *New developments also slowed down by the divergence in the two APIs*

Persistency: Summary

- The LCG POOL project provides a hybrid store integrating object streaming with RDBMS technology
  - POOL has been successfully integrated into three quite different LHC experiments software frameworks
  - Successfully deployed as baseline persistency mechanism for CMS, ATLAS and LHCb at the **scale of ~400TB**
- POOL continues the LCG component approach by abstracting database access in a vendor neutral way
  - A Relational Abstraction Layer has been released and is being picked up by several experiments
  - Minimised risk of vendor binding, simplified maintenance **and data distribution** are the main motivations
- POOL as a project is (slowly) migrating to a support and maintenance phase
  - Need keep remaining manpower focused in order to complete remaining developments and to provide adequate support to user community
  - Maintaining a significant experiment contribution is required insure the the tight feedback loop which made POOL an effective project
- The LCG Conditions DB project has produced several releases of the Oracle and MySQL based implementations within the LCG Application Area
  - After an interface and extension review a concrete plan to consolidate the implementations has been discussed
  - Manpower also from the experiments is now becoming available to the project allowing to re-factor the package based on the Relational Abstraction Layer
- New complementary technologies such as FroNtier are being integrated into the LCG persistency framework as distributed access to database data gets more interest

SPI Services (November 2004)

- External Software
- Savannah Project Portal
- Software Librarian, builds and releases
- Testing Frameworks
- QA checklists and reports
- Software Distribution
- LCG Software Configuration
- Development of LCG policies, templates
- Code Documentation (doxygen, lxr, viewcvs)
- Documentation and LCG Workbook

SPI: Summary

- **SPI provides a set of stable services that fully use the current resources**
  - Savannah, External Software, Testing Frameworks, Software Distribution, Build and Release, QA Activities
- **Resources will be reduced at the end of 2005 and we will not be able to add new activities**
  - We will maintain and improve the existing ones
  - No requests for major new services from LCG, need to adapt or interface to external tools
- **We refer to and follow the guidelines of the Architects Forum**
  - Representing the users and steering our priorities
- **Whenever possible we also do help/work for experiments specific needs**
  - Installations of additional software, interface to their build/distribution systems, add features

Applications

Les Robertson

- Concern over the long-term continuity of personnel, and the long-term support of products, in particular the maths library.
  - The problem of Phase1-Phase2 staffing at CERN has been taken up by the CERN management. It is hoped that this will be resolved during the next two months.
  - The long-term support plan, scheduled for the end of 2004, has been postponed to the end of the first quarter of 2005..
- Stresses the importance to support the Monte Carlo generator codes required by the LHC experiments. Such support appears to fit the scope of the Simulation project.
  - Improved support for MC generators implemented, largely following the conclusions of the Review of Generator Services Subgroup of LCG that took place on 25 March.

Applications (ii)

Les Robertson

- Requests further clarification of how proposals made in the Architect's Forum are to be incorporated in to the Applications Area.
  - For proposals in areas already within the scope of the applications area, the Architects Forum (AF) acts as the decision making body, subject to PEB endorsement of significant decisions, particularly those requiring future allocation of resources. Issues where agreement cannot be reached in the AF are taken to the PEB.
  - For proposals in areas outside the present scope of the applications area, applications area projects develop proposals, present them to the AF for modification and approval, and the AF-agreed proposal is presented to the PEB for a decision.
  - For proposals within AA scope but for which the AF is unable to reach agreement, a policy has been in place from the beginning to escalate the issue to higher management (now the PEB) for resolution and decision. To date, this course has not had to be invoked.
  - After each meeting of the Architects' Forum, a summary of the conclusions and decisions is presented to the PEB.

Answers to Committee Questions

- How does ROOT integrate with the AA projects, in particular with POOL and SEAL? Is there mechanism for joint planning? What is review process?
  - ROOT integration process, status addressed in POOL and SEAL talks
  - Process formalized in integration objectives, activities more on LCG side than on the ROOT side
  - Joint planning implies defined, agreed workplans in both projects in areas relevant to their common work; this we don't have
  - LCG/AA planning process yields (resourced) workplans that in areas of ROOT collaboration have ROOT's agreement with the AA plan
  - We do not have the equivalent from the ROOT project
  - Planning relies on the AA plan and relatively informal agreements with the ROOT team on ROOT deliverables and participation.
  - Geant4 is another case in which AA collaborates with a wide;y scoped project. Here joint planning is done in a more symmetrical way. CERN/AA Geant4 team develops within the context of Geant4 planning (and in consultation with AA) a detailed subplan for AA-related activities.
  - Similarly regarding the review process. AA activities are extensively reviewed, including of course the ROOT collaborative elements, but ROOT activities and plans are not directly reviewed.

Achievements & Concerns

- important steps made in production and development
- ATLAS, CMS and LHCb data challenges demonstrated production use of software from many application area projects
  - e.g. successful use of POOL (~400TB)
- very robust and stable operation of Geant4 (~10^{-6} crashes /evt)
- integration of ROOT4 into POOL underway
- reduction of manpower due to transition from development to support, deployment and maintenance had negative impact, on e.g.
  - support for POOL file access from ROOT environment
  - concrete workplan for conditions DB project
- sofar late deliverables not yet on experiments or project critical path
- to avoid this for next year need too maintain stable and focused manpower from CERN and experiments
- avoid duplication of work (ROOT vs AA projects) → managment issue