LCG Applications Area Status

Torre Wenaus, BNL/CERN
LCG Applications Area Manager

http://lcgapp.cern.ch

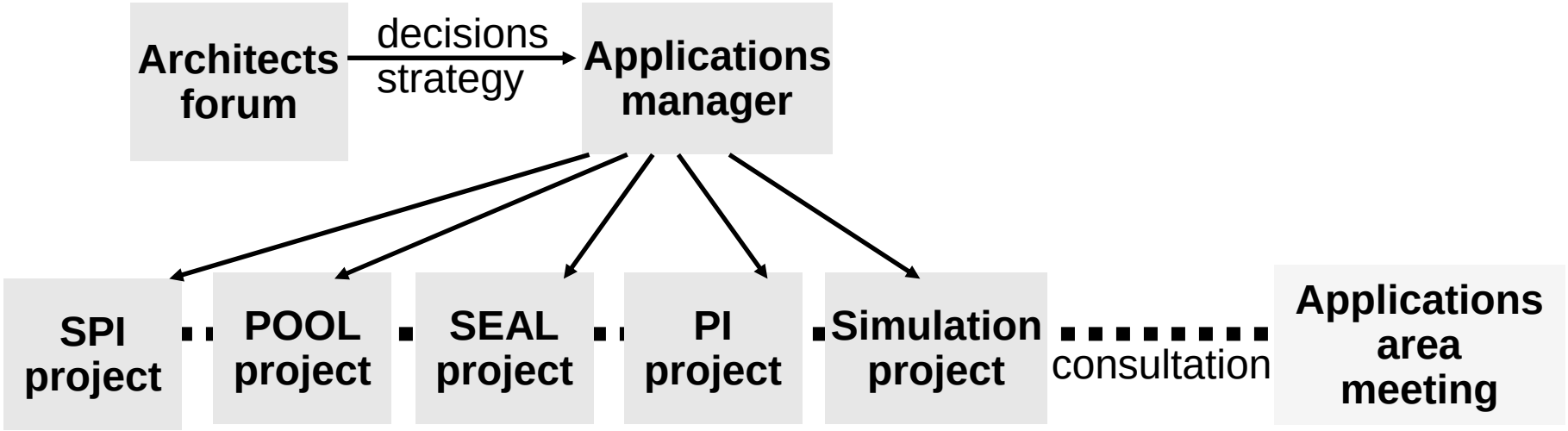
US ATLAS Physics and Computing Meeting
August 28, 2003

BROOKHAVEN
NATIONAL LABORATORY

Outline

- Applications area organization and overview
- Implementing the Architecture Blueprint
- Project planning
- Applications area projects
  - POOL, SEAL, PI, Simulation, SPI
- Personnel resources
- Non-CERN participation, collaboration
- Concluding remarks

Applications Area Organisation



Focus on Experiment Need

- Project structured and managed to ensure a focus on real experiment needs
  - ✓ **SC2/RTAG** process to identify, define (need-driven requirements), initiate and monitor common project activities in a way guided by the experiments themselves
  - ✓ **Architects Forum** to involve experiment architects in day to day project management and execution
  - ✓ **Open** information flow and decision making
    - ✓ Applications area meeting ~weekly
  - ✓ **Direct participation** of experiment developers in the projects
  - ✓ Tight **iterative feedback** loop to gather user feedback from frequent releases
  - **Early deployment and evaluation** of LCG software in experiment contexts
  - **Success defined by experiment adoption** and production deployment

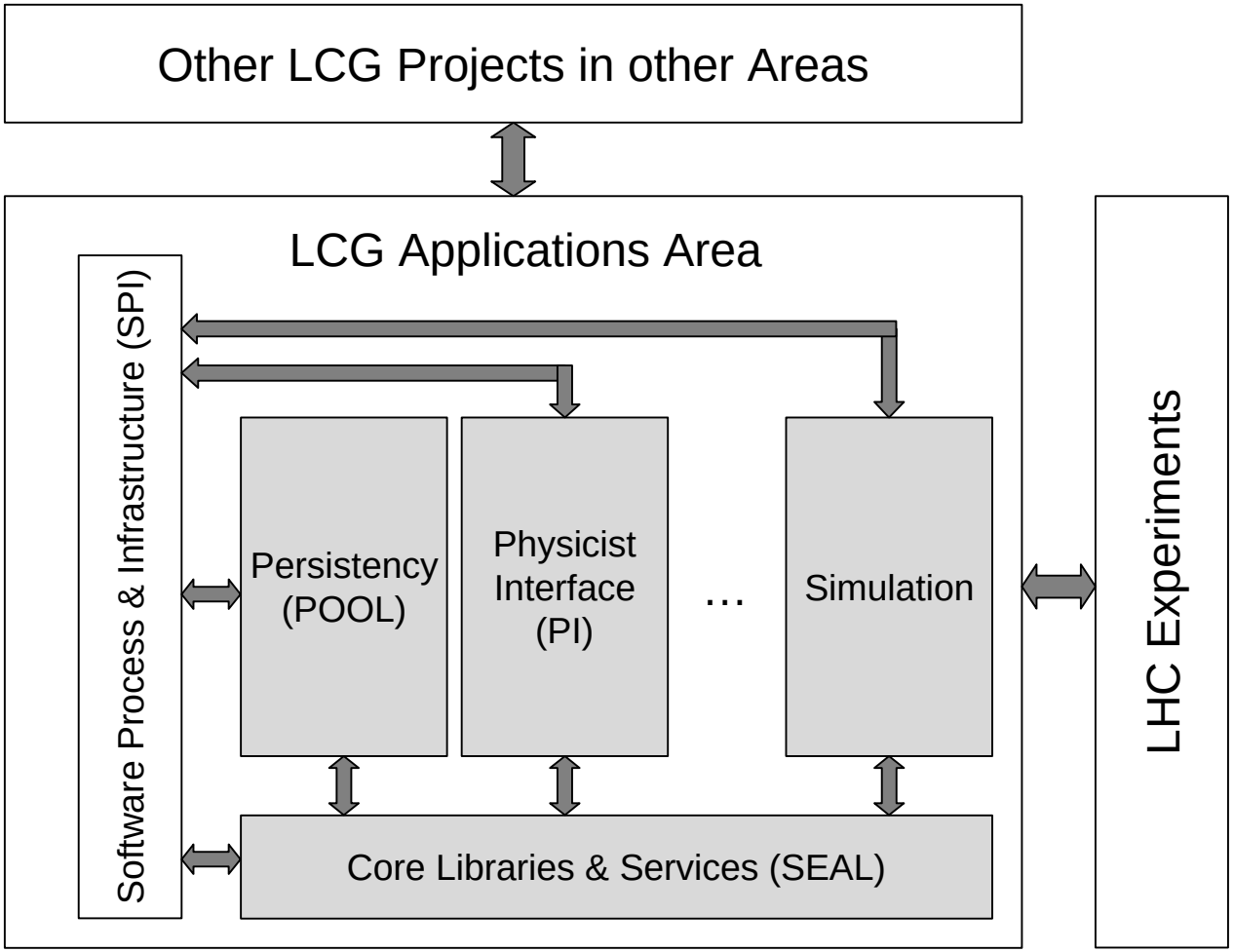
Substantive evaluation and feedback from experiment integration/validation efforts now in progress

Applications Area Projects

- **Software Process and Infrastructure (SPI)** *(operating – A.Aimar)*
  - Librarian, QA, testing, developer tools, documentation, training, …
- **Persistency Framework (POOL)** (*operating – D.Duellmann)*
  - POOL hybrid ROOT/relational data store
- **Core Tools and Services (SEAL)** *(operating – P.Mato)*
  - Foundation and utility libraries, basic framework services, object dictionary and whiteboard, math libraries, (grid enabled services)
- **Physicist Interface (PI)** *(operating – V.Innocente)*
  - Interfaces and tools by which physicists directly use the software. Interactive analysis, visualization, (distributed analysis & grid portals)
- **Simulation** *(operating – T.Wenaus et al)*
  - Generic framework, Geant4, FLUKA integration, physics validation, generator services

The set of projects is complete unless/until a distributed analysis project is opened

Project Relationships



Implementing the Architecture Blueprint

- **Use what exists:** almost all work leverages existing software
  - ROOT, Gaudi/Athena components, Iguana components, CLHEP, Aida, HepUtilities, SCRAM, Oval, NICOS, Savannah, Boost, MySQL, GSL, Minuit, gcc-xml, RLS, …
- **Component-ware:** followed, and working well
  - e.g. rapidity of integration of SEAL components into POOL
- **Object dictionary:** In place
  - Meeting POOL needs for persistency
  - Application now expanding to interactivity
  - ROOT and LCG agree on convergence on common dictionary; should see activity in this over the next year

Implementing the Architecture Blueprint (2)

- **Component bus/scripting environment:** both Python environment and its integration with ROOT/CINT progressing well
- **Object whiteboard:** Still to come
  - Serious design discussions should start in next month
- **[add one for analysis; full access to ROOT as analysis tool]**
- **Distributed operation:** almost no activity, not in scope
  - Awaits 'ARDA' RTAG outcome in October
- **ROOT:** The 'user/provider' relation is working
  - Good ROOT/POOL cooperation – POOL gets needed modifications, ROOT gets debugging/development input

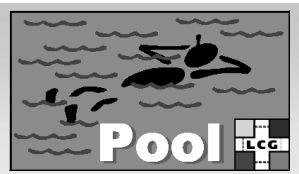
Applications Area Project Planning

- Planning page linked from applications area page
  - Project plans for the various projects
  - WBS, schedules
    - WBS mirrors (is) the project/subproject/work [illegible]e structure
    - Schedule defines milestones and deliverables [illegible]
      - Three le[illegible]:
        - Lev[illegible]
        - [illegible]
        - [illegible]
      - [illegible] take-up
  - [illegible]
  - [illegible]
    - [illegible]
  - [illegible] overall project plan
    - [illegible]dule, personnel resource requirements
    - Prepared as part of the blueprint (October 2002)
  - Risk analysis

Not addressed here

http://lcgapp.cern.ch/project/mgmt/

Persistency Framework Project



◆Dirk Duellmann, CERN IT/DB

- To deliver the physics data store (POOL) for ATLAS, CMS, LHCb
- Scope now includes conditions DB as well as POOL
  - Carrying forward a long-standing common project
- Production POOL release delivered on schedule in June
  - A level 1 LHCC milestone, meeting a date set a year earlier
  - Contains all functionality requested by the experiments for initial production usage in their data challenges
  - Leverages proven technologies: ROOT I/O, relational databases
  - Provides stably supported (1 year) format for data files
- Focus is now on responding to feedback, debugging, performance, documentation, process and infrastructure
  - Two such releases since the June production release
- Project manpower is OK
  - IT/DB, LCG, and experiments are all contributing substantial and vital manpower
- Recently got a good QA report from SPI

POOL Provides

- Bulk event data storage – An object store based on ROOT I/O
  - Full support for persistent references automatically resolvable to objects anywhere on the grid
  - Recently extended to support updateable metadata as well, with some limitations
- File cataloging – Implementations using grid middleware (*EDG version* of RLS), relational DB (MySQL), and local files (XML)
- Event metadata – Event collections with queryable metadata (physics tags etc.) with implementations using MySQL, ROOT, and POOL itself
- Transient data cache – Optional component by which POOL can manage transient instances of persistent objects

POOL Integration and Validation

- With the POOL production release milestone met, the next vital step which will really measure POOL's success is successful take-up by the experiments
  - The next months will tell: how close are we to a good product
- CMS and ATLAS heavily active in POOL integration and validation
- CMS successfully integrated POOL/SEAL and validated it to begin production applications
  - The first integration/validation milestones of the project were met in July:
    - POOL persistency of CMS event
    - CMS acceptance of POOL for pre-challenge production (PCP)
- Similar milestones for ATLAS should be completed in September
- LHCb integration is beginning now
- POOL team member assigned to each experiment to assist/liaison
- Ready to deploy POOL on LCG-1 as soon as we are given access
- Should see TB data volumes stored with POOL this fall
  - CMS PCP
  - ATLAS DC1 data migration to POOL

Core Libraries and Services (SEAL)

◆Pere Mato, CERN EP/SFT

- Provide foundation and utility libraries and tools, basic framework services, object dictionary, component infrastructure
  - Facilitate coherence of LCG software and integration with non-LCG software
- Development uses/builds on existing software from experiments (e.g. Gaudi, Iguana elements) and C++, HEP communities (e.g. Boost)
- On schedule
  - Has successfully delivered POOL's needs, the top priority
  - New blueprint-driven component model and basic framework services, directed also at experiments, just released
- Doesn't always come easily: "Combining existing designs into a 'common' one is not trivial"
- Manpower is OK – team consists of LCG and experiment people
- CLHEP accepted our proposal to 'host' the project
  - Also reflects appeal of SPI-supported services
  - Highest priority LHC request – splitting CLHEP – now done
- Math libraries: Agreement that GSL can replace NAG (from a functionality viewpoint), but public report still to be done

SEAL Release Road Map

Release	Date	Status	Description (goals)
V 0.1.0	14/02/03	ternal	◆Establish dependency between POOL and SEAL ◆Dictionary support & generation from header files
V	/03	public	◆Essential functionality sufficient for the other existing LCG projects (POOL) ◆Foundation library, system abstraction, etc. ◆Plugin management
V 0.3.0	16/	nternal	◆Improve functionality required by POOL ◆Basic framework base classes
V 1.	30/06/03	ublic	◆Essential functionality sufficient to be adopted by experiments ◆Collection of basic framework services ◆Scripting support

Released 14&26/02/03

Released 04/04/03

Released 23/05/03

Released 18/07/03

SEAL Next Steps

- Handle feedback on SEAL issues from POOL integration
- Get feedback (from experiments+POOL+…) about new component model and framework services
  - Corrections and re-designs are still possible
- Documentation, training
- Support new platforms: icc, ecc, Windows
- Adapt to more SPI tools and policies
- Develop new requested functionality
  - Object whiteboard (transient data store)
  - Improvements to scripting: LCG dictionary integration, ROOT integration
  - Complete support for C++ types in the LCG dictionary

Physicist Interface (PI) Project

◆Vincenzo Innocente, CERN EP/SFT

- Responsible for the interfaces and tools by which a physicist (particularly a physicist doing analysis) will directly use the software
  - Interactivity (the "physicist's desktop"), analysis tools, visualization, distributed analysis, ‘grid portals’
- Currently a small effort (<2 FTEs) focused on the limited scope opened so far
- Analysis Services
  - Simplified, implementation-independent AIDA interface to histograms, tuples developed and offered to users for evaluation (little feedback!)
  - Principal initial mandate, a full ROOT implementation of AIDA histograms, recently completed
  - Integration of analysis services with POOL, SEAL
- Analysis Environment
  - Interactivity and bridge to/from ROOT – joint work with SEAL
    - pyROOT Python interface to ROOT – full ROOT access from Python command line
  - Interoperability via abstract interfaces to fitters and other analysis components
  - End-user interfaces for tuples/event collections
- Other identified work is on hold by SC2
  - Distributed analysis, general analysis environment, event and detector visualization
- Future planning depends on what comes out of the ARDA RTAG

Simulation Project

◆Torre Wenaus et al

- Principal development activity: **generic simulation framework**
  - Expect to build on existing ALICE work; currently setting the priorities and approach among the experiments

Andrea Dell'Acqua

- Incorporates longstanding CERN/LHC **Geant4** work
  - Aligned with and responding to needs from LHC experiments, physics validation, generic framework

John Apostolakis

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

Alfredo Ferrari

- Simulation **physics validation** subproject very active
  - Physics requirements; hadronic, em physics validation of G4, FLUKA; framework validation; monitoring non-LHC activity

Fabiola Gianotti

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

Paolo Bartalini

Physics Validation

◆Fabiola Gianotti, EP/SFT (ATLAS)

- To assess the adequacy of the simulation and physics environment for LHC physics
  - And provide the feedback to drive needed improvements
  - Ultimately, certify packages, frameworks as OK
- Validation based mainly on
  - Comparisons with LHC detector test beam data
  - Simulations of complete LHC detectors
  - “Simple benchmarks”: thin targets, simple geometries
- Coordinates a lot of work being done in the experiments, G4, FLUKA
  - Supplemented with a small amount of direct LCG effort to “fill in the cracks”
  - Foster cooperation, coherence, completeness
- So far:
  - Physics validation studies made by experiments revisited
  - Monitor and assess progress with G4 hadronic physics
    - E.g. improved pion shower profiles in the ATLAS HEC
  - First results from simple benchmarks
  - Information, results gathering on web page

http://lcgapp.cern.ch/project/simu/validation

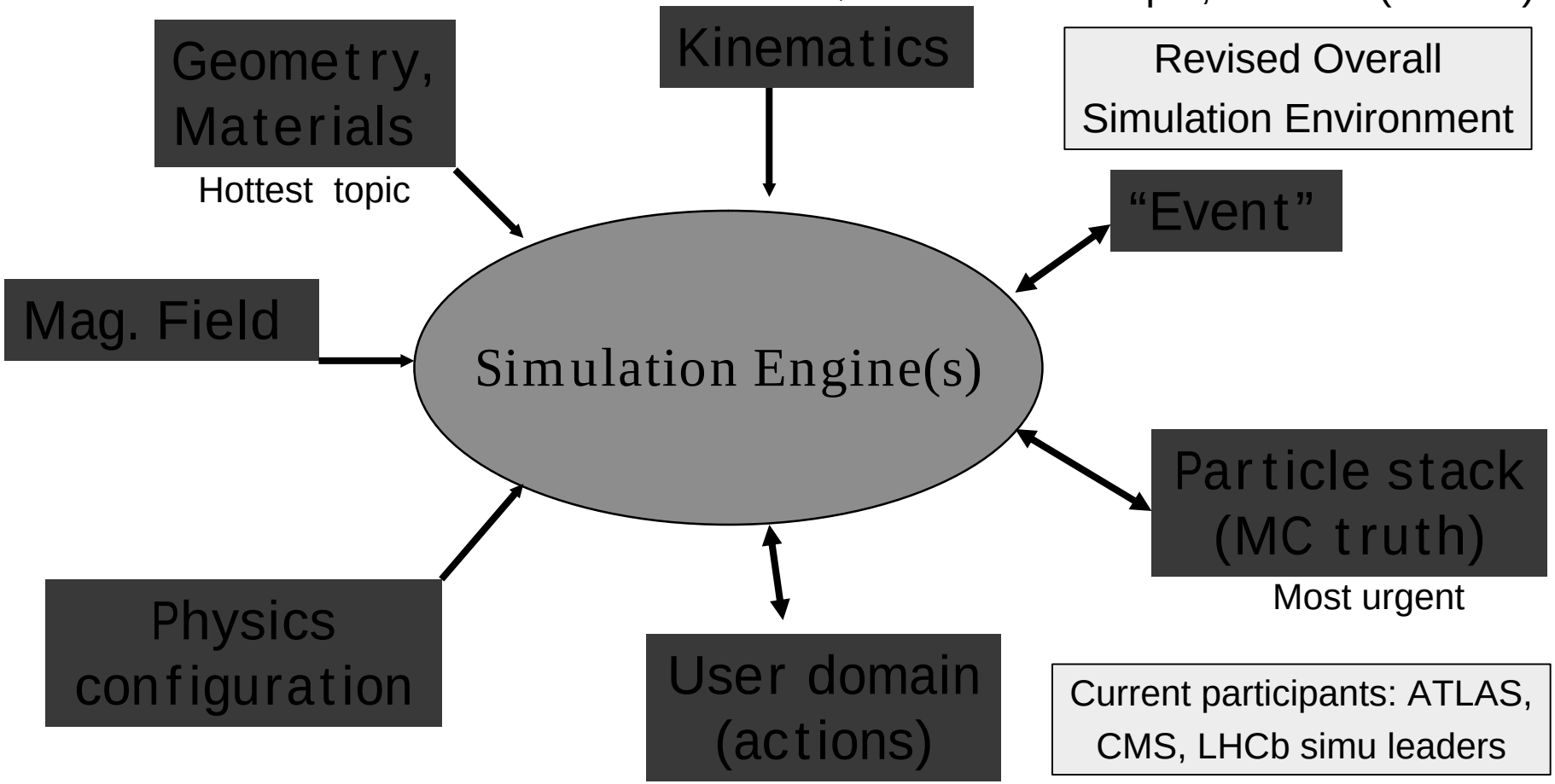
Generator Services

◆Paolo Bartalini, CERN EP/CMT (CMS)

- Responsible for
  - Generator librarian services
  - Tuning and validation of event generators
  - Common event files, event database
  - Event storage, interface and particle services
- Guided and overseen by the LHC-wide MC4LHC group
- GENSER generator repository on schedule
  - Alpha version released in June for feedback (substantial)
  - Beta version due mid-Sep; pre-release already out
    - Includes PYTHIA, HERWIG, ISAJET; HIJING coming
- Active program of broad monthly meetings
- Lots of useful input from the large MC generator workshop in July
- Resources (1-2 FTEs) from LCG (Russia)

Generic Simulation Framework - ROSE

◆Andrea Dell'Acqua, EP/SFT (ATLAS)



Aim of the subproject: To provide services implementing the red boxes
Integrated with SEAL, and building on existing software (particularly ALICE VMC)

Geant4

◆John Apostolakis, EP/SFT

- Responsible for CERN/LHC participation in Geant4
  - Focusing the CERN/LHC effort on LHC priorities
  - While supporting CERN's long-standing and valuable role as the international 'home' of Geant4 with a leading role in management and infrastructure
- Has developed a workplan for the subproject integrated with the overall Geant4 plan and with LCG simulation subproject activities and priorities:
  - Management, coordination, infrastructure
  - Hadronic physics
  - Geometry, tracking, em physics
- Employs substantial personnel resources (addressed later)
  - ~3:3:1 distribution among Management/infrastructure, hadronic physics, and other development (geometry, tracking, em physics)
  - Strong cooperation with Physics Validation (to which ~1.5 FTEs were transferred
  - Working with the generic framework team on architecture and integration
  - Working on improving the synergy and cooperation of the infrastructure effort with SPI

Fluka Integration

- Alfredo Ferrari, CERN AB

- Fluka development proper is not a project activity, though it has recently received strengthened support as a CERN activity
  - CERN effort supplied to Fluka team
  - As part of this agreement, Fluka source code will be opened in ~12 months
- Project activity involves
  - Integration of Fluka as a simulation engine in the generic framework
    - Generic framework design not advanced enough for this to be an active program yet
    - Indications at this point are that the work already done by ALICE-FLUKA will be usable; only modest additional work needed
  - Physics validation of Fluka
    - Already working with the physics validation subproject
- Activity is led by the CERN-resident Fluka project leader

Simulation Project Major Milestones

- 2003/6: **Generator librarian** and first library version in place
- 2003/7: Simulation **physics requirements** revisited
  - Will complete in Sep
- 2003/7: Decide **generic framework high level design**, implementation approach, software to be reused
  - Might complete in Sep
- 2003/9: 1st cycle of **EM physics** validation complete
- 2003/12: Generic **framework prototype** with G4, FLUKA engines
- 2004/1: 1st cycle of **hadronic physics** validation complete
- 2004/3: Simulation **test and benchmark suite** available
- 2004/9: First generic **simulation framework production** release
- 2004/12: Final **physics validation document** complete

Software Process and Infrastructure (SPI)

◆Alberto Aimar, CERN IT

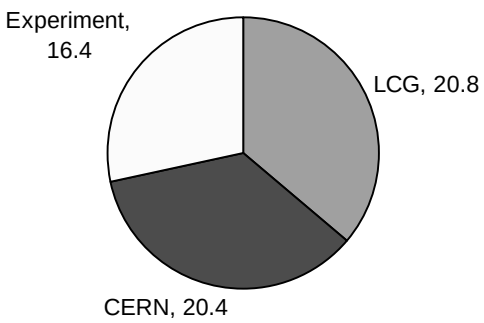
- Full suite of tools and services in place, supporting developers and users
  - 'Best of breed' adopted from experiments, open source community
- Policies on code standards and organization, testing, documentation in place
- Strong QA activity: (positive) POOL QA report just released
- Software distribution/remote installation service just deployed
- Currently navigating personnel transitions, partially offset by new arrivals
  - Transitions to other projects (keeping 'maintenance' level SPI participation)
    - New LCG applications area arrivals contribute to SPI (a policy)
  - Manpower level (just) enough to support the essentials
    - 1-2 more would improve handling new requests and reduce 'firefighting mode'
  - Team made up of LCG, IT, EP, experiment participants
- Savannah development portal a great success with 63 projects, 351 users recently
  - Used by LCG-App, -GD, -Fabric, 3 experiments, CLHEP
- Training program: ROOT; SCRAM; POOL and SEAL coming soon
- Additional platform porting underway: icc (new Intel); ecc (64-bit), Windows

Foster software quality "as least as good as and preferably better than that of any experiment"

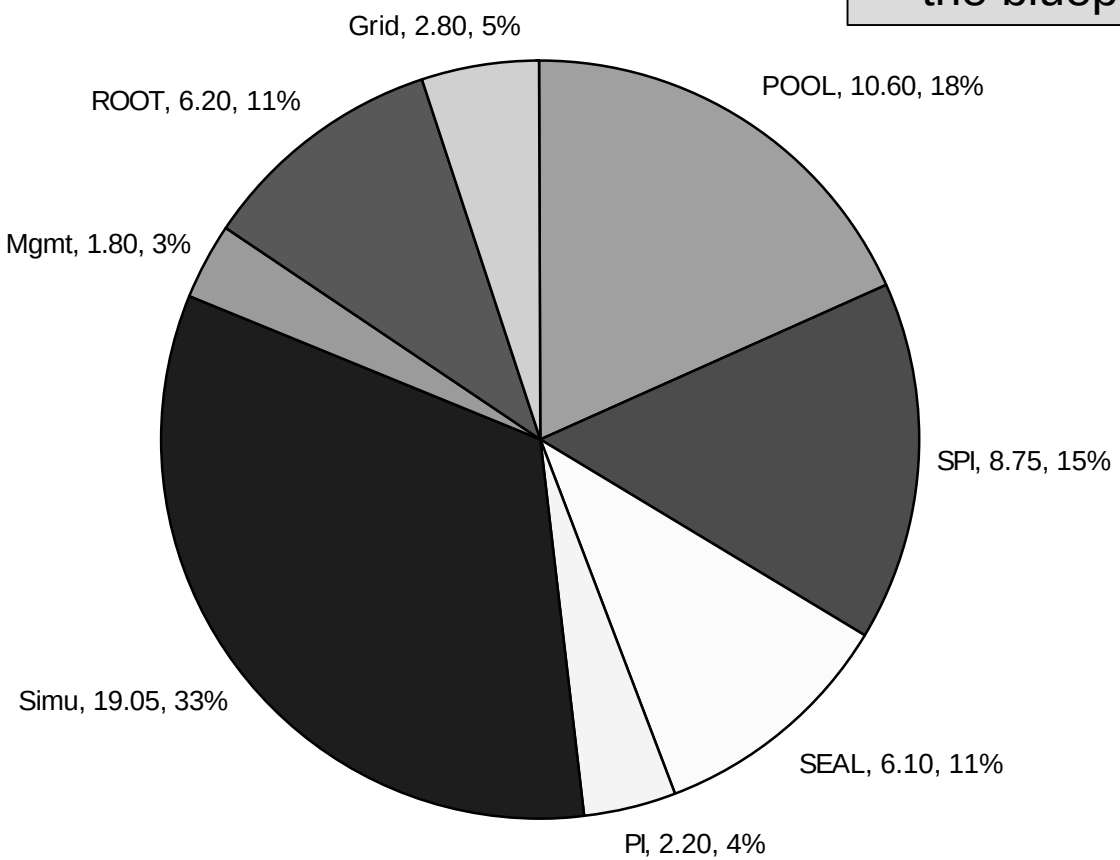
SPI Services and Tools

- CVS service and servers administration
- External software service
- LCG librarian
- SCRAM configuration/build manager
- Savannah portal
- Quality assurance and policies
- Code documentation tools
- Software distribution and installation
- Automatic build/test system
- Testing frameworks
- Web management
- Workbook, documentation, templates
- Training

Personnel Distribution



LCG AA : FTEs by project



FTE levels match the estimates of the blueprint

Non-CERN Participation

- Examples:
  - POOL collections (US)
  - POOL RDBMS data storage back end (India)
  - POOL tests (UK)
  - POOL-driven ROOT I/O development & debugging (US)
  - SEAL scripting tools (US)
  - Generator services (Russia)
  - SPI tools (France, US)
  - Math libraries (India)
- Many more opportunities:
  - Throughout the simulation project
  - Several PI work packages
  - Unit and integration tests
  - End-user examples

Concluding Remarks

- POOL, SEAL, and PI software is out there
  - Take-up is underway for POOL and SEAL
    - The real measure of success, and signs are good so far
    - First round of CMS POOL-SEAL validation milestones met
- Our most important milestone, POOL production version, was met on time and with the required functionality
  - Required effective collaborative work among POOL, SEAL, SPI, experiments
- Doesn't always come easily: "Combining existing designs into a 'common' one is not trivial"
  - But it is being done
- Manpower is appropriate
  - is at the level the experiments themselves estimated is required
  - is being used effectively and efficiently for the common good
  - is delivering what we are mandated to deliver
- Collaborative development of common software with a strong focus on reuse and tight coupling to experiment need – is working so far

For more information

- Applications area web
  - http://lcgapp.cern.ch
- Applications area work plan Version 2 draft (8/03)
  - http://lcgapp.cern.ch/project/mgmt/AppPlanV2.doc
- PI status (7/03)
  - http://agenda.cern.ch/fullAgenda.php?ida=a031778
- POOL status (7/03)
  - http://agenda.cern.ch/fullAgenda.php?ida=a031778
- SEAL status (7/03)
  - http://agenda.cern.ch/fullAgenda.php?ida=a031780
- SPI status (7/03)
  - http://agenda.cern.ch/fullAgenda.php?ida=a031779
- Simulation project
  - Overview (5/03)
    - http://lcgapp.cern.ch/project/simu/SimuProjectOverview200303.ppt
  - Generic simulation project status (7/03)
    - http://agenda.cern.ch/fullAgenda.php?ida=a031780
  - Physics validation subproject status (7/03)
    - http://agenda.cern.ch/fullAgenda.php?ida=a031780
  - Geant4 status (7/03)
    - http://agenda.cern.ch/fullAgenda.php?ida=a031780
  - Generator services subproject status (8/03)
    - http://agenda.cern.ch/fullAgenda.php?ida=a032131
  - Recent FLUKA developments (8/03)
    - http://agenda.cern.ch/fullAgenda.php?ida=a032130

Supplemental

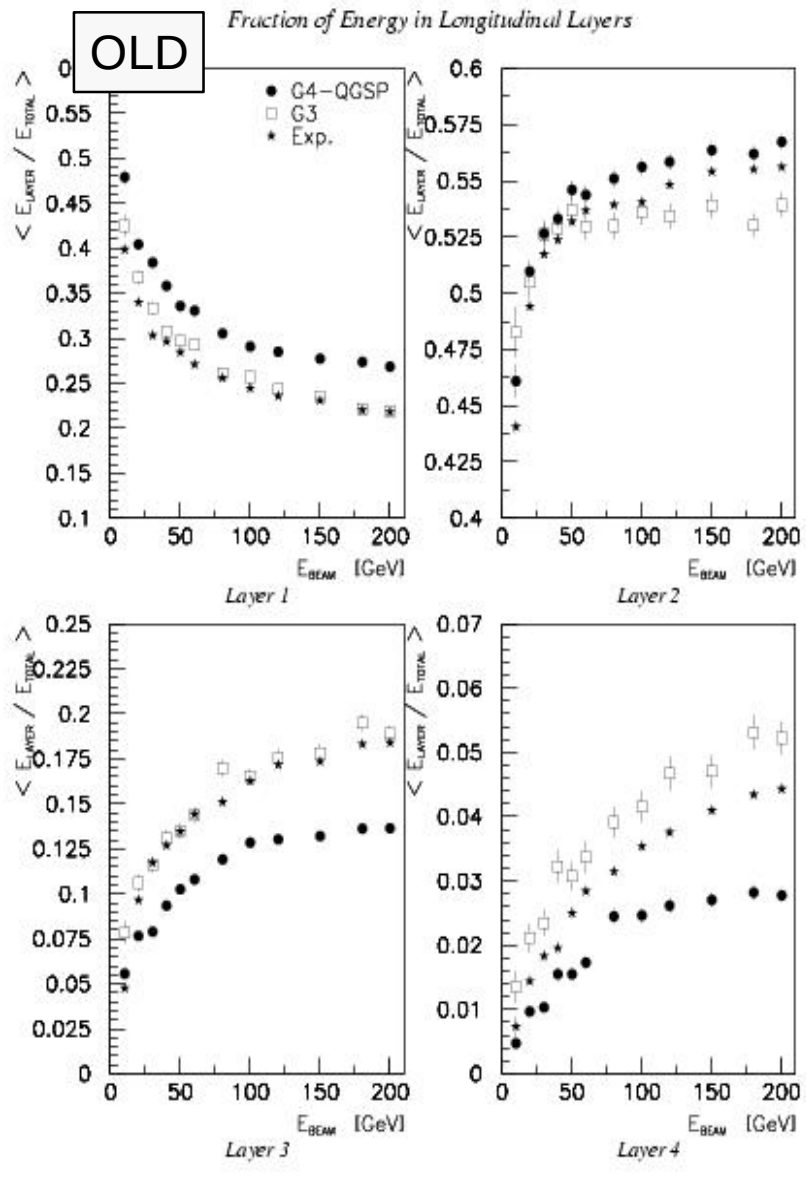
RTAG on
An Architectural Roadmap towards Distributed Analysis (ARDA)

- To **review** the current DA activities and to capture their **architectures** in a consistent way
- To confront these existing projects to the HEPCAL II **use cases** and the user's potential **work environments** in order to explore potential shortcomings.
- To consider the interfaces between Grid, LCG and experiment-specific services
  - Review the functionality of experiment-specific packages, state of advancement and role in the experiment.
  - Identify **similar functionalities** in the different packages
  - Identify functionalities and components that could be **integrated** in the generic GRID middleware
- To confront the current projects with critical GRID areas
- To develop a roadmap specifying wherever possible the **architecture**, the **components** and potential **sources of deliverables** to guide the medium term (2 year) work of the LCG and the DA planning in the experiments.
- Started a couple of days ago; to conclude in October
  - Membership from experiments, LCG apps area and grid tech area, outside experts

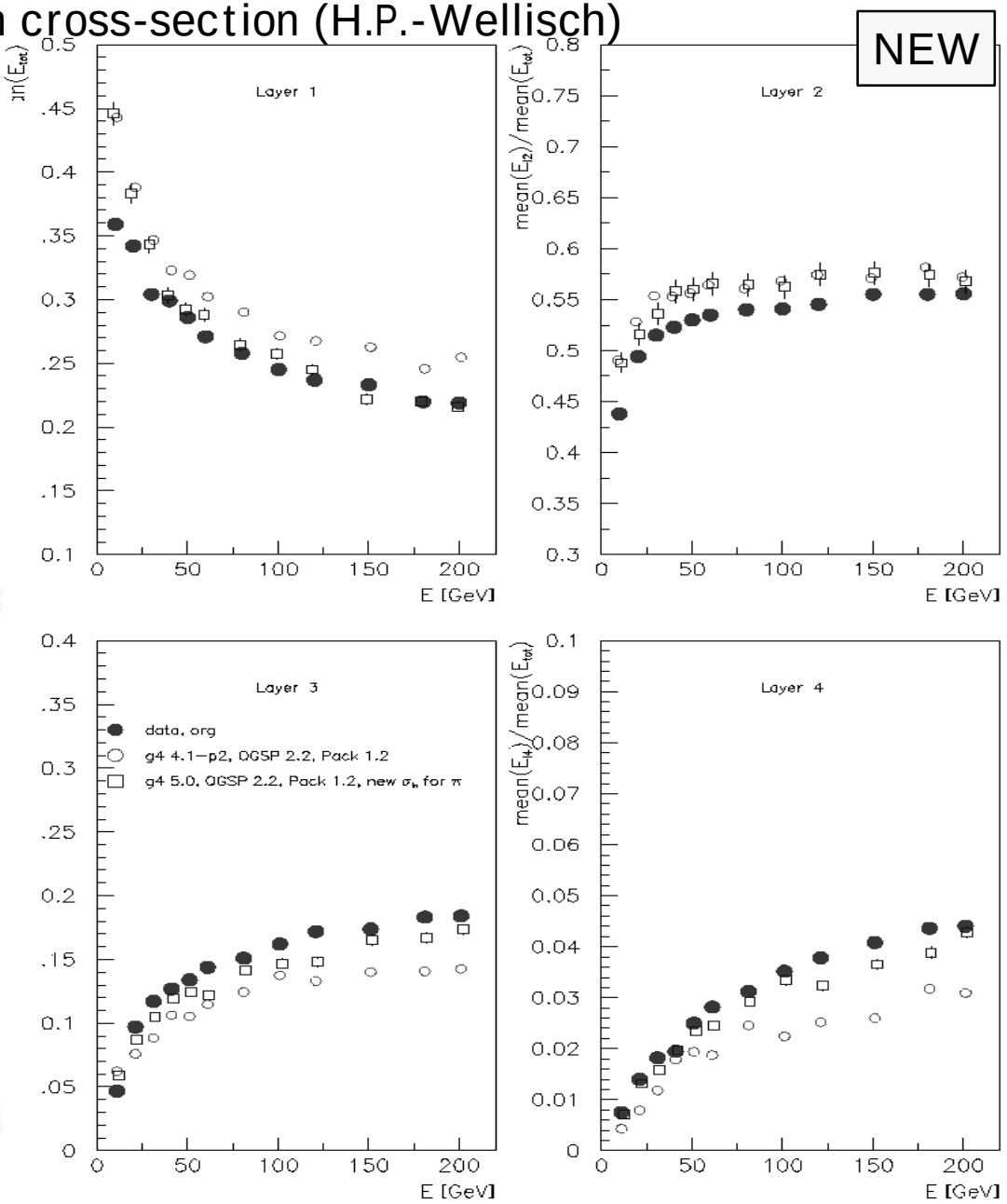
❷ Progress with G4 hadronic physics

Improved pion shower profile in the ATLAS hadronic end-cap calorimeter after fixing 10% mismatch in pion cross-section (H.P.-Wellisch)

OLD



NEW



Ongoing Physics Validation Work

- Most recent G4 hadronic physics lists which describe well ATLAS HEC and Tilecal data will be tested by LHCb and CMS.
- Documentation of hadronic physics lists for LHC being prepared (first version recently posted)
- All experiments are taking test-beam data with many sub-detectors this Summer

  → expect new extensive round of comparison results in Autumn
- Two FLUKA activities starting:

  -- update ATLAS Tilecal test-beam simulation

  -- simulate hadronic interactions in ATLAS pixel test-beam set-up
- Active program of broad monthly meetings
- 1-day meeting in November or December to discuss validation item by item (e.g. electron energy resolution, hadronic shower profile) across experiments

  → complete first cycle of physics validation

Non-CERN Participation

- Engaging external participation well is hard, but we are working at it
- Difficulties on both sides
  - Being remote is difficult
    - More than it needs to be… e.g. VRVS physical facilities issue – recently improved, but more improvements needed
  - Remote resources can be harder to control and fully leverage, and may be less reliably available
    - We work around it and live with it, because we must support and encourage remote participation

Collaborations

- Examples…
  - Apart from the obvious (the experiments, ROOT)…
  - GDA: Requirements from apps for LCG-1, RLS deployment, Savannah
  - Fabrics: POOL and SPI hardware, Savannah
  - GTA: Grid file access, Savannah
  - Grid projects: RLS, EDG testbed contribution, software packaging/distribution
  - Geant4: Extensive LHC-directed participation
  - FLUKA: Generic framework, simulation physics
  - MC4LHC: Generator services project direction
  - CLHEP: Project repository hosting