Applications Area
Introduction – Manpower – Planning

Torre Wenaus, BNL/CERN
LCG Applications Area Manager
http://lcgapp.cern.ch

LHCC Comprehensive Review of LCG
November 23, 2004

Applications Area Organization

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

Applications Area Meeting

Consultation ← → Communication

Torre Wenaus, BNL/CERN
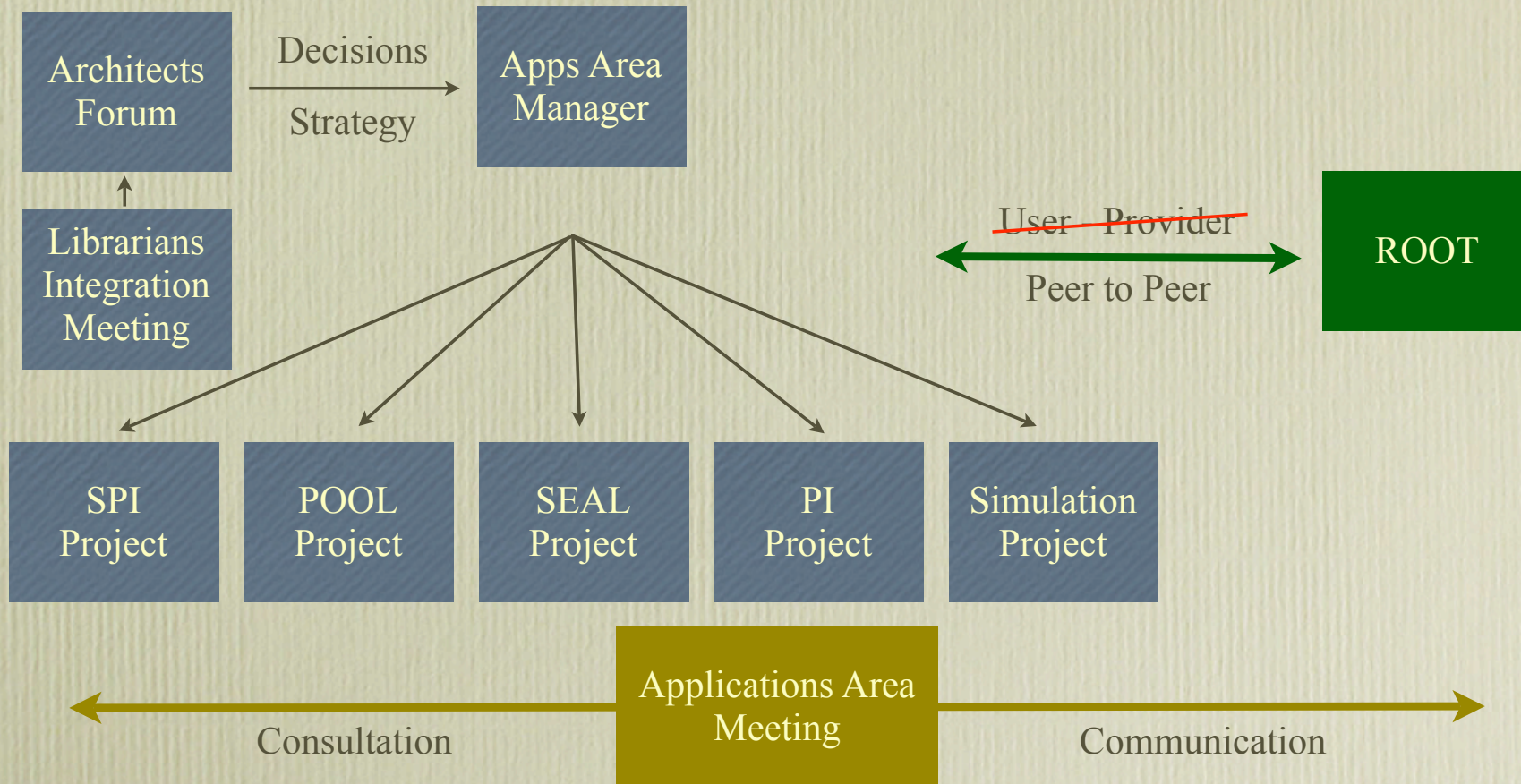
Project Focus

- Deliver the common physics applications software
- Organized to ensure focus on real experiment needs
  - SC2/RTAG/PEB: experiment-driven requirements and monitoring
  - **AF: 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

Torre Wenaus, BNL/CERN

Validation Highlights

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

Torre Wenaus, BNL/CERN

Works In Progress

- ROOT analysis environment integration: full POOL file access from ROOT, dictionary unification, ...
  - Production focus up to now is shifting to analysis support
- Reflex as the basis for a merged LCG/ROOT dictionary
- Expanded POOL relational DB capability, from basic access services to 'object relational POOL', with conditions data (in particular) in mind
  - Event data support 'done'; conditions/relational now the focus
- Conditions DB - finally moving ahead after long inertial period, drawing on the POOL relational work
- Scalable distributed data access/management
  - Not in POOL scope; 'the grid will take care of it'
  - e.g. via collaboration with the 3D project (GDA)
- Planning for AA leadership change actively underway
  - Torre finishes end Feb

Torre Wenaus, BNL/CERN

Vital Collaborations for AA: ROOT and Grid

- ROOT provides:
  - Bulk event store for POOL
    - Object streaming, schema evolution, file (or grid) level I/O
  - Analysis environment
    - Statistical analysis & fitting, C++ interactivity, math, analysis-optimized I/O tools & data structures, graphics,...
  - etc!
- Grid provides:
  - All aspects of distributed computing; AA scope only involves interfaces, principally in POOL
  - Integrating AA software into a distributed computing infrastructure is an experiment task
    - cf. ARDA organization
- *Comments on these collaborations follow...*

ROOT Collaboration

- Most critical area lately: POOL/ROOT4 integration
  - Close and effective (if longer than expected!) collaborative effort between POOL and ROOT teams
  - Deliverables reflected in both POOL and ROOT work plans
- Slowly-developing area, now rapidly moving towards critical path: ROOT analysis environment support
  - POOL file access, in particular POOL reference support
  - Reflex/CINT integration: 'LCG class' access in ROOT
  - Common C++ math library
  - Currently in these areas most work is being done on the LCG side; objectives are not reflected in ROOT plans and work commitments

Torre Wenaus, BNL/CERN

Committee Question

- 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.
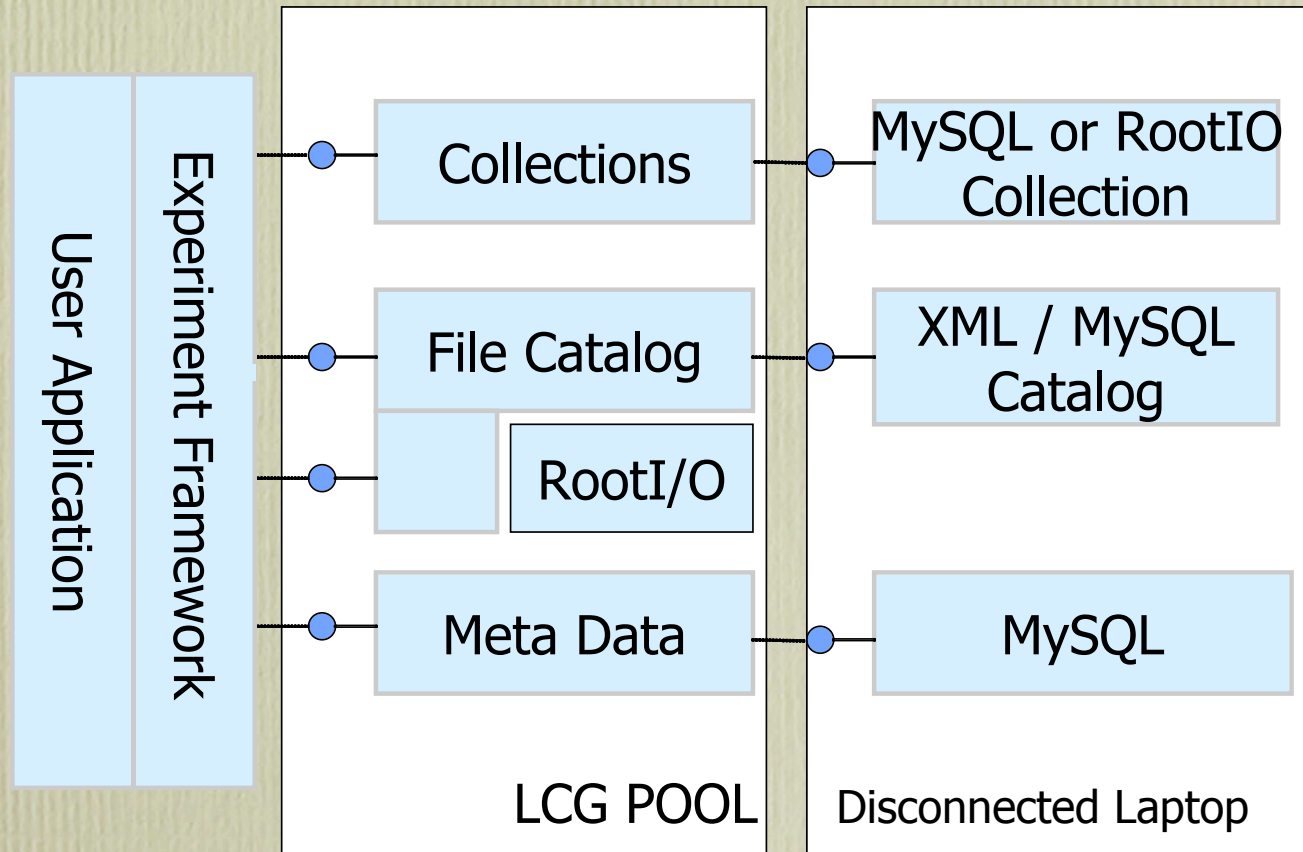
Torre Wenaus, BNL/CERN

Grid Collaboration

- Not a success so far in terms of deliverables or effective collaboration
- POOL's catalog/metadata components are only as good as the services behind them, which are out of POOL's control except for the in-house implementations
  - Overall POOL+catalog functionality negatively impacted by RLS problems, and unclear whether lessons are fully learned
- AA/GDA/middleware connections too weak; AF has very recently put forward proposals to strengthen them (to be discussed in PEB)
- Some promising 'distributed POOL' directions do not involve Grid
  - Collaborative work with FNAL/CMS on FroNtier/POOL integration for scalable distributed DB access (3D project)
  - The demonstration via POOL's in-house implementations of capable, simple catalogs directly based on RDBMSs
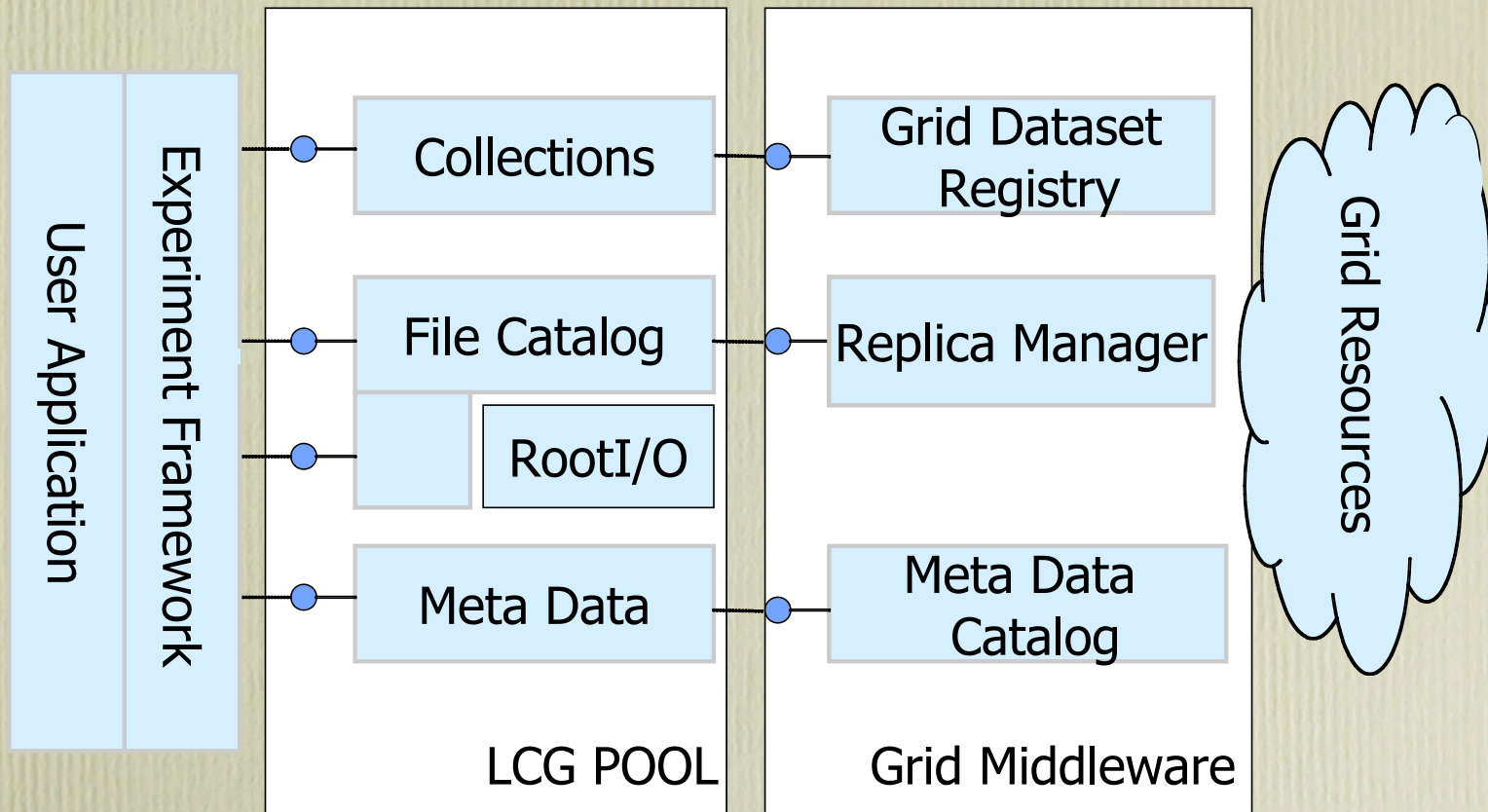
Torre Wenaus, BNL/CERN

Two slides from Dirk, from Fall 2002:

POOL Off The Grid

User Application | **Experiment Framework**

LCG POOL:
- Collections
- File Catalog
- RootI/O
- Meta Data

Disconnected Laptop:
- MySQL or RootIO Collection
- XML / MySQL Catalog
- MySQL

All implemented and in routine use

POOL On The Grid

Two years later, **none** of the expected middleware is available as stable, adequate, routinely used products

11

Other Issues, Concerns

- Still don't really have the 'better understanding of ARDA' that was to guide our future plans relating to PI, distributed analysis, and distributed aspects of POOL
    - No AA role in (distributed) analysis tools affects focus and priorities; harder to focus on analysis needs which is now vital
    - AA/POOL's experience in successfully delivering the 'off-grid' system should be leveraged to successfully complete the rest
- Still surprisingly little (ie. virtually no) experiment involvement in simulation physics validation. Effort is small and getting smaller.
- Concerns expressed in March internal review of generator services have been addressed -- continuity of coverage in the librarian & leader roles, and author communication
    - GENSER now in production in ATLAS

Torre Wenaus, BNL/CERN

AA Internal Reviews

- Oct 2003 internal review was very useful, setting directions and providing recommendations for 2004
  - Status of main issues identified by the committee:
    - Build/config tool: XML-based config description now in development will support SCRAM, CMT and config/make
    - Schema evolution: addressed by ROOT4 migration
    - POOL collections: not resolved, cf. ARDA/analysis comments
    - SEAL libraries management: dependency management/ reduction ongoing, e.g. SEAL partitioning work
    - Dictionary convergence: ongoing as a priority
    - Generic simu framework: dropped due to ATLAS, CMS, LHCb disinterest
    - Redefine PI in light of ARDA: PI in maintenance mode
- Next review ~Feb 2005, inviting the same committee plus 'SC2 AA godparents'; organized by new AA leader

Torre Wenaus, BNL/CERN

Key Recent/Upcoming Milestones

- 2004/5/31 - POOL relational abstraction layer done
- 2004/6/30 - New dictionary API & prototype
- 2004/8/31 - POOL relational storage manager release (this month)
- 2004/10/1 - First mathlib release (before holiday)
- 2004/10/15 - Consolidated Geant4 acceptance suite (before holiday)
- 2004/10/31 - POOL integrates ROOT4 (completion imminent)
- 2004/12/1 - Generator production framework (on time)
- **2004/12/31 - Geant4 validation in LHC production (on time)**
- **2004/12/31 - Development, support and resource plan**
- 2005/3/31 - Conditions DB production release
- 2005/6/30 - Final physics validation document
- **2005/9/30 - Phase 1 software complete and deployed**

Torre Wenaus, BNL/CERN

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

Torre Wenaus, BNL/CERN

Personnel Changes

- Recent personnel changes include: (with FTE gain/loss)
    - Andreas Pfeiffer began as Software Librarian (0)
    - Fabiola Gianotti left as Simulation Physics Subproject leader (-)
    - Alberto Ribon took this job over (0)
    - Torre Wenaus reduced participation and left Simu Leader (-)
    - Gabriele Cosmo took over as Simu Leader (0)
    - POOL manpower shrinkage due to unbalanced transfer of FTEs to DB services (which IT-DB agrees needs to be corrected), and experiment participation needs strengthening, most critically in the ROOT storage manager (which expts promise to address) (-)
    - Declines in experiment participation in SEAL (-)
    - EGEE supplied two FTEs to SPI, partially offsetting departures (-)

Torre Wenaus, BNL/CERN

Manpower Needs

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...

Torre Wenaus, BNL/CERN

Staffing Requirements Estimates

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

	2004	2005	2006	2007	2008
SPI	6.2	6.2	5.5	4.6	3.9
SEAL	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

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

Manpower Requirements by Project
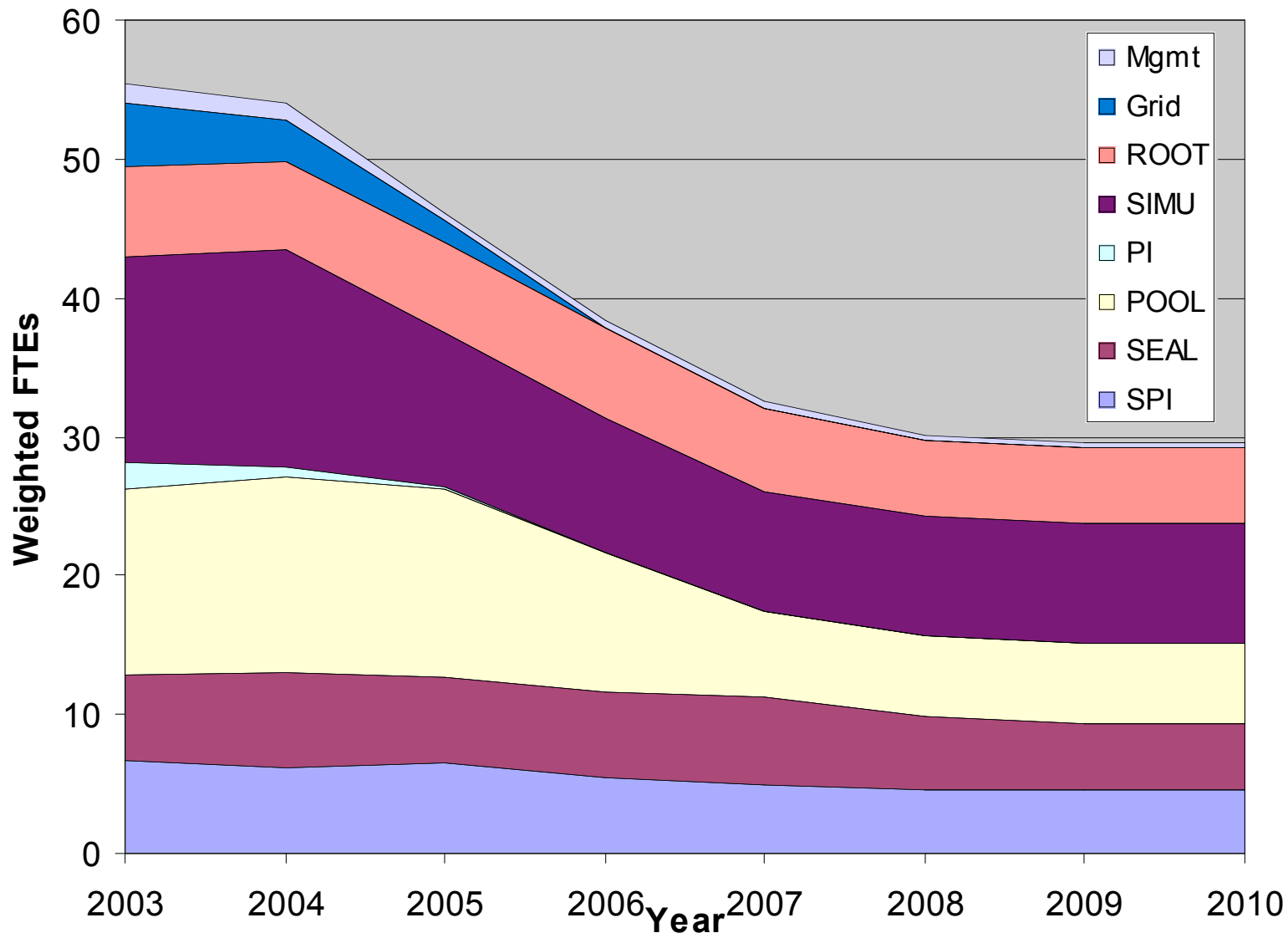
Staffing Sources

Distribution of CERN support between base and special uncertain from 05

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

Staffing and Long Term Planning

- LCG, CERN currently addressing obtaining the needed resources
- Detailed long term AA planning will proceed when
  - Available resources are understood
  - New AA management is named -- leadership change and phase 1/2 transition present an opportunity for examining and changing the program & organization
  - Any changes in experiment requirements and plans are accounted for -- consultation round will begin when new management is named
- As mentioned, L1 milestone for a program and resource plan through 2008 is proposed to be delayed 3 months to end March 2005

Torre Wenaus, BNL/CERN