ALICE: Offline Planning and personnel resources

LHCC Manpower Review of Computing

September 3, 2003

Questions to be answered

- Profile of available and required manpower at CERN / Regional Centers / Institutes
- Other resources existing and potential
- Measures of progress in producing necessary software
  - Management tools to track the progress
  - Verification of the quality of the LCG software
- Computing elements which will not be provided in case the required manpower and resources are not available

Foreword

- Lack of personnel in LHC computing (experiment & common HW/SW infrastructure) has been emphasized by LHC Computing Review (2001) and judged "*extremely worrying*"
- "*CERN and the Collaborations together must do all that they can to provide the HR that are needed* for Core Software development"
- The personnel shortage in common HW/SW has been alleviated by creating the LCG project and sending computing professionals funded by member countries
- No such mechanism exists for experiments where the personnel shortage remains a problem
- ALICE has re-profiled the planning and adapted its strategy to guarantee readiness for data processing

Menu : Planning & Resources

- ALICE Offline organization & management
- Strategy for the Offline project, DC & milestones
- Personnel resources : available and requests
- Conclusions
    - Securing two additional permanent staff positions at CERN for 2 area coordinators, now on medium-term position, is crucial for the project success.
    - Replacing at CERN 2-3 short term personnel with medium term staff (computing physicist) would alleviate the unstable situation of Core Offline team by making it less dependent on temporary personnel.
    - The needed personnel in detector projects must be recruited in the institutes of the collaboration.

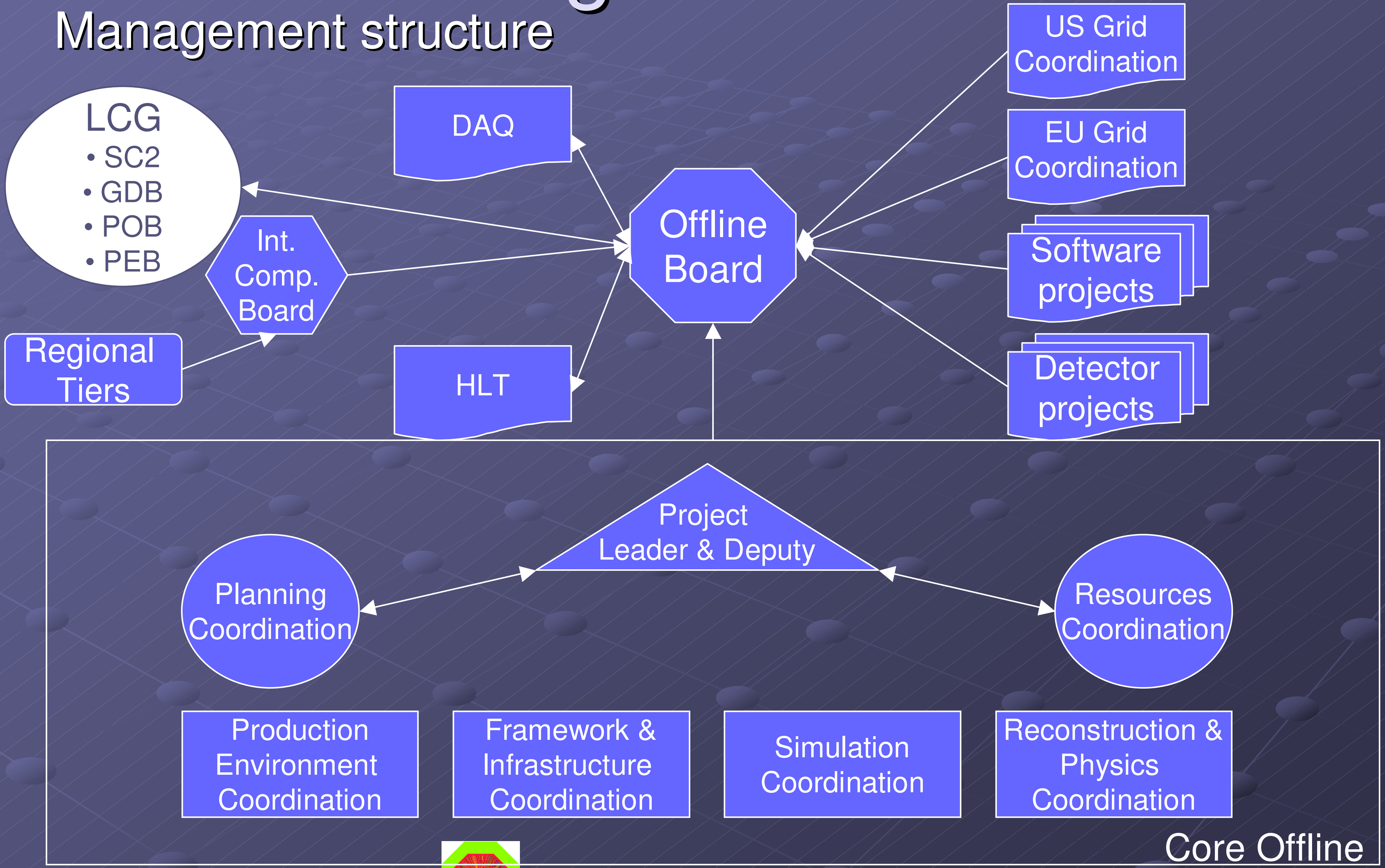
Organization

- Offline project mandate :
  - Prepare software and computing infrastructure for experiment’s data processing (+DAQ, +HLT projects);
  - Provide and maintain a complete infrastructure for simulation, reconstruction and analysis already during construction phase;
- Offline personnel for software developments:
  - Core Offline team : minority, full time, located at CERN;
  - SW projects : outside CERN;
  - Detector projects : most of the personnel, part time, located in collaboration institutes;
- LCG project provides common hardware and software infrastructure for LHC computing.

Strict coordination required to make the best usage of the personnel available.

Organization

Management structure



LCG
- SC2
- GDB
- POB
- PEB

DAQ

US Grid Coordination

EU Grid Coordination

Int. Comp. Board

Offline Board

Software projects

Regional Tiers

HLT

Detector projects

Project Leader & Deputy

Planning Coordination

Resources Coordination

Production Environment Coordination

Framework & Infrastructure Coordination

Simulation Coordination

Reconstruction & Physics Coordination

Core Offline

Organization

Management structure

- Light weighted, single structure
  - Efficient use of available personnel
  - High adaptability to rapid changing technology
- Merge framework developers (providers) & physics algorithms developers (consumers)
  - Maximize communication
  - Rapid feedback to users requirements
  - Economy of personnel (polymorphism of software experts)

Single line of development & complete transition to OO/C++ in 1999

Planning Strategy

- Dynamic management of the work schedule
  - Develop a long term software infrastructure
  - Maintain the infrastructure in working state during detector construction
- Constraints
  - Depend on the planning of external projects
  - Most developers refer to detector projects
  - Take advantage of latest developments in fast evolving technology
  - No personnel available for in depth planning activity or traditional SW engeenering
  - Majority of personnel in Core Offline project is temporary
  - New people come with diverse skills and profiles

Light weighted and opportunistic strategy with flexible data challenges as high level milestones

Core Offline team @ CERN

- The best solution for ALICE
  - Need for a strong and centralized team of experts
    - To facilitate coordination in detector projects and regional centers
  - CERN, more than other ALICE groups, has the critical mass of people with the right skills
  - Benefit from co-habitation with ALICE and LCG management
  - Benefit from the attraction CERN exercises on young people with the right profile

Development strategy

- Minimize the effective amount of development
  - Chose mature and well tested products
    - ROOT : Common HEP solution for data persistency at the file level, interface to various libraries, visualization, graphical user interface, virtual Monte-Carlo, geometrical modeler
    - AliEn : The ALICE distributed computing environment all made with Open Source components based on Open Standards; 2 FTE for development, 0.5 for operation, in production since 2002
- Forced to rely on temporary personnel
  - However a minimum of long-term personnel needed
- Delegate well identified and modular packages to teams outside Core Offline group
  - Detector data base
  - EDG/EGEE test bed
  - AliEn components
  - More to come …

Data Challenges

- Stress-test the ALICE data model, DAQ/HLT hardware and software infrastructure with prototypes of increasing complexity
  - Computing DC: record HI data at 1.2 Gbytes/s and export a fraction of them for quasi online processing outside CERN
  - Physics DC: provide the infrastructure for organized distributed Monte-Carlo production and chaotic world-wide data-analysis

Computing Data Challenge

- ALICE in collaboration with IT division :
  - Assess the MS requirements and evaluate available products ;
  - Evaluate functions of DAQ, Offline, HLT projects ;
  - Large-scale high-throughput distributed DC (4) to :
    - Prototype the DAQ, Offline, HLT computing systems
    - Verify their integration
    - Assess technologies and computing models
    - Test hardware and software components in realistic environment
    - Achieve an early integration of the overall computing infrastructure

Milestones

When	Band Width Mbytes/s	To MSS Tbytes	# nodes with Gbit ethernet	Milestone
10/2002	200	200		Rootification of raw data (ITS+TPC)
09/2003	300	300	150	Integration of single detector HLT / Quasi on-line reconstruction @ CERN / Partial data replication to remote centers
05/2004	450	450	200	HLT for all detectors / Remote reconstruction of partial data streams / Raw digits for barrel & MUON
05/2005	750	750	250	Prototype of final HLT software / Prototype of final remote data replication / Raw digits for all detectors
05/2006	750 (1250 if possible)	750 (1250 if possible)	300	Final test on final system

Physics Data Challenge

- Objectives :
  - Prototype and test scalability of the components needed to simulate, reconstruct, and analyze data on distributed computing resources
- Three interlinked components :
  - ROOT
  - AliRoot
  - AliEn

Milestones

When	Fraction of final* capacity (%)	Physics Objective
06/01 – 12/01	1	pp studies, reconstruction of TPC and ITS
06/02 – 12/02	5	First test of the complete chain from simulation to reconstruction for the PPR / Simple analysis tools / Digits in ROOT format
01/04 – 06/04	10	Complete chain used for trigger studies / Prototype of the analysis tools / Comparison with parametrized Monte-Carlo / Simulated raw data
01/06 – 06/06	20	Test of the final system for reconstruction and analysis

* Fraction of events simulated in one year of standard data talking

Tracking progress

- Milestones set by the needs of PPR & DC
  - Full and fast simulation
  - Detector reconstruction
  - Global reconstruction
- Progress monitored by Physics DC
- Central coordination at CERN (architect, librarian, multi-platform compatibility)
- Offline board takes the decision on framework evolution and review progress
- Common working session during Offline week
- Code reviewed by experts

Tracking progress : verification of LCG software quality

Grid technology area

Title	When	Rationale
AliEn production interface to LCG1 readiness review (prototype and functionality)	11/2003	Verify that the design and functionality of the AliEn LCG-1 interface corresponds to the requirements of ALICE for using the LCG-1 resources for simulation and reconstruction during the PDC 3 to be executed in the first half of 2004
AliEn analysis interface to LCG-1 readiness review (prototype design and functionality)	04/2004	Verify that the design and functionality of the AliEn LCG-1 interface corresponds to the requirements of ALICE for using the LCG-1 resources for analysis of the data produced during the PDC 3 to be started in the second quarter of 2004
AliEn production interface to LCG-1 performance evaluation	07/2004	Verify on the basis of the PDC 3 results that the AliEn LCG-1 interface has the robustness and functionality necessary for a production environment
LCG-1 middleware evaluation	07/2004	Provide an evaluation of the LCG-1 middleware as used via AliEn based on the experience of the PDC 3
AliEn analysis interface to LCG-1 performance evaluation	10/2004	Verify on the basis of the PDC 3 results that the AliEn LCG-1 interface has the robustness and functionality necessary for performing analysis

Tracking progress : verification of LCG software quality

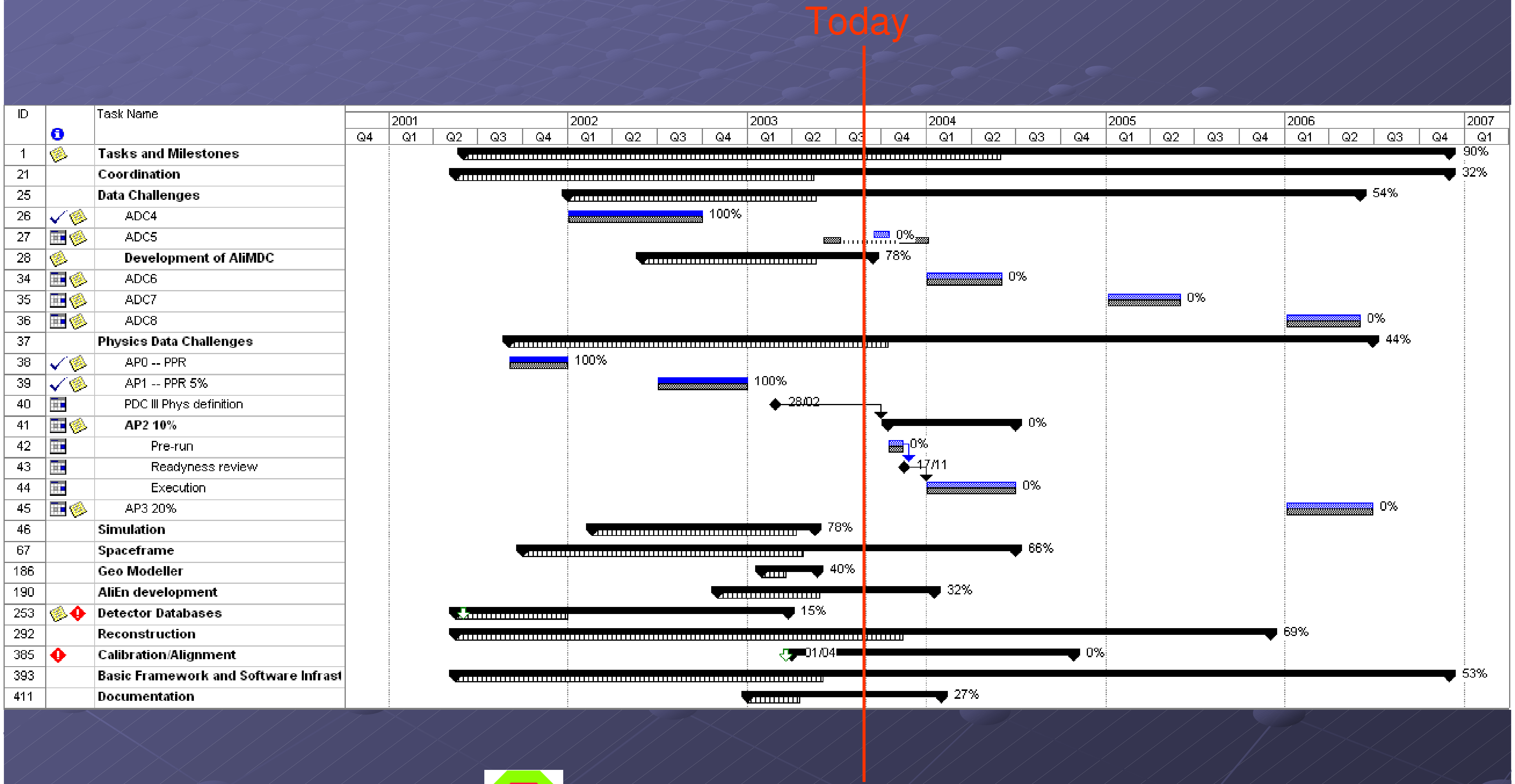
Grid deployment area

Title	When	Rationale
LCG-1 PDC3 production environment verification	07/2004	Provide an evaluation of the LCG-1 production environment and infrastructure for simulation and reconstruction based on the results of PDC 3
LCG-1 PDC3 analyses environment verification	10/2004	Provide an evaluation of the LCG-1 environment for analysis based on the experience of the analysis of the PDC 3 data
LCG-1 PDC4 production environment verification	07/2004	Verify on the basis of the PDC 3 results that the AliEn LCG-1 interface has the robustness and functionality necessary for a production environment
LCG-1 middleware evaluation	07/2006	Provide an evaluation of the LCG-1 production environment and infrastructure for simulation and reconstruction based on the results of PDC 3
LCG-1 PDC4 analysis environment verification	10/2006	Provide an evaluation of the LCG-1 environment for analysis based on the experience of the analysis of the PDC 3 data

Tracking progress : verification of LCG software quality
Fabric area

Title	When	Rationale
Verification of new release of CASTOR	11/2003	Evaluate the design and functionality of the new release of CASTOR in relation to the need of the ALICE Computing Data Challenges and DAQ system
Verification of target performance for ADC V	11/2003	Verify that the fabric infrastructure and MSS satisfy ALICE target performances for ADC V (tape bw 300Mb/s) and scale up to 150 nodes
Verification of target performance for ADC VI	11/2004	Verify that the fabric infrastructure and MSS satisfy ALICE target performances for ADC VI (tape bw 450Mb/s) and scale up to 200 nodes
Verification of target performance for ADC VII	11/2005	Verify that the fabric infrastructure and MSS satisfy ALICE target performances for ADC VII (tape bw 750Mb/s) and scale up to 250 nodes.

Tracking progress : ALICE Offline Planning


Today
2001
2002
2003
2004
2005
2006
2007
Q4 Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1 Q2 Q3 Q4 Q1
90%
32%
54%
100%
0%
78%
0%
0%
0%
44%
100%
100%
28/02
0%
0%
17/11
0%
0%
78%
66%
40%
32%
15%
69%
01/04
0%
53%
27%

ID		Task Name
1		**Tasks and Milestones**
21		**Coordination**
25		**Data Challenges**
26	✓	ADC4
27		ADC5
28		**Development of AliMDC**
34		ADC6
35		ADC7
36		ADC8
37		**Physics Data Challenges**
38	✓	AP0 -- PPR
39	✓	AP1 -- PPR 5%
40		PDC III Phys definition
41		**AP2 10%**
42		Pre-run
43		Readyness review
44		Execution
45		AP3 20%
46		**Simulation**
67		**Spaceframe**
186		**Geo Modeller**
190		**AliEn development**
253	!	**Detector Databases**
292		**Reconstruction**
385	!	**Calibration/Alignment**
393		**Basic Framework and Software Infrast**
411		**Documentation**

Personnel Profile (task oriented)

- Core Offline at CERN :
  - 4 permanent staff persons
  - Profile is build up with the assumptions that temporary personnel is NOT replaced*
  - Evolution reported since 1998

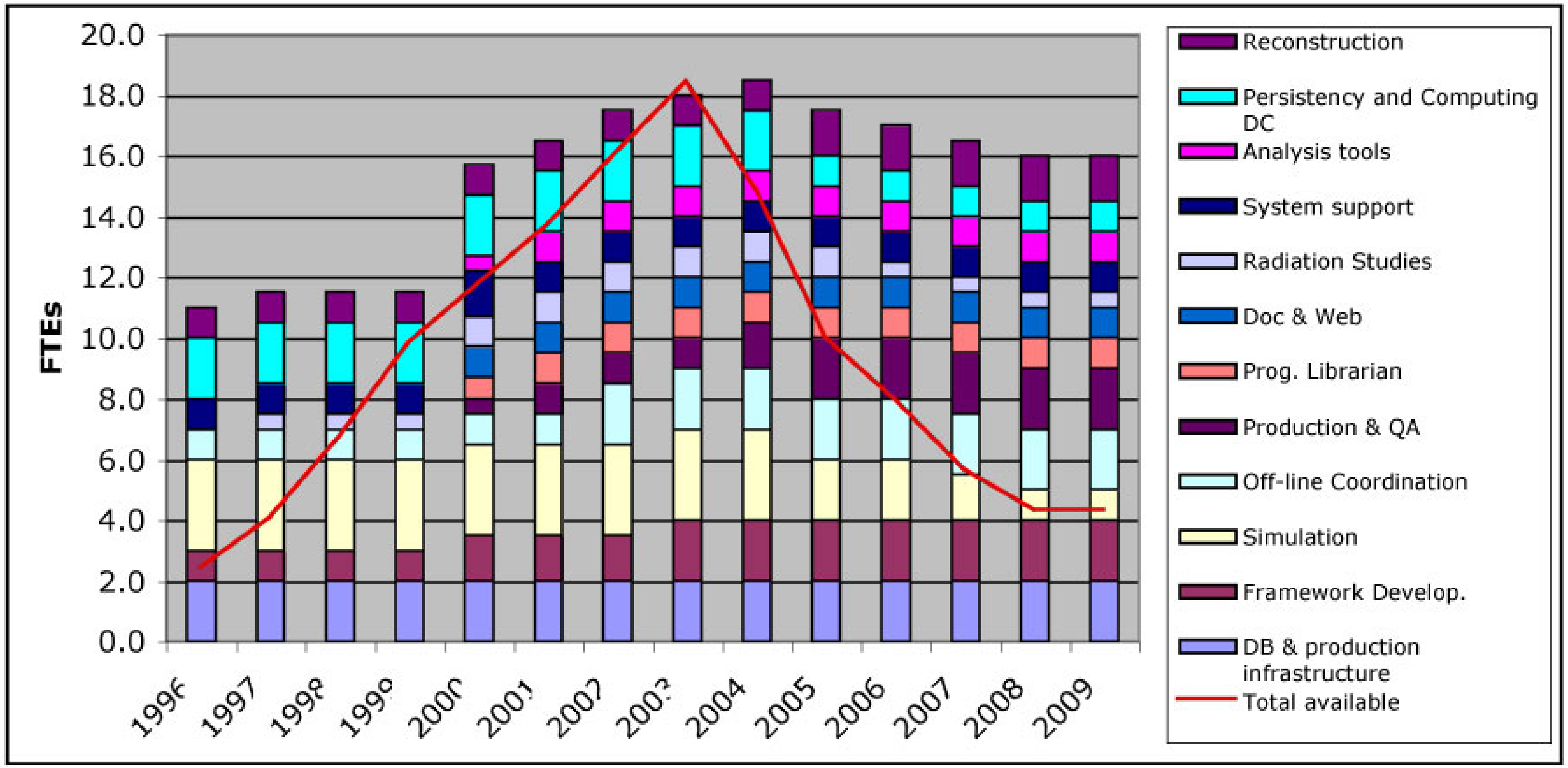
* Pessimistic scenario to emphasize fragility of the structure

Personnel Profile (task oriented) - 1/2

Summary Core Offline team

	98	**99**	**00**	**01**	**02**	**03**	**04**	**05**	**06**	**07**	**08**	**09**
Avail.	6.8	9.8	11.8	13.7	16.1	18.4	14.9	10.0	8.0	5.6	4.3	4.3
Needed	11.5	11.5	15.7	16.5	17.5	18.0	18.5	17.5	17.0	16.5	16.0	16.0
Missing	4.8	1.7	3.9	2.8	1.4	0.4	3.7	7.5	9.0	10.9	11.7	11.7

Personnel Profile (task oriented) - 2/2



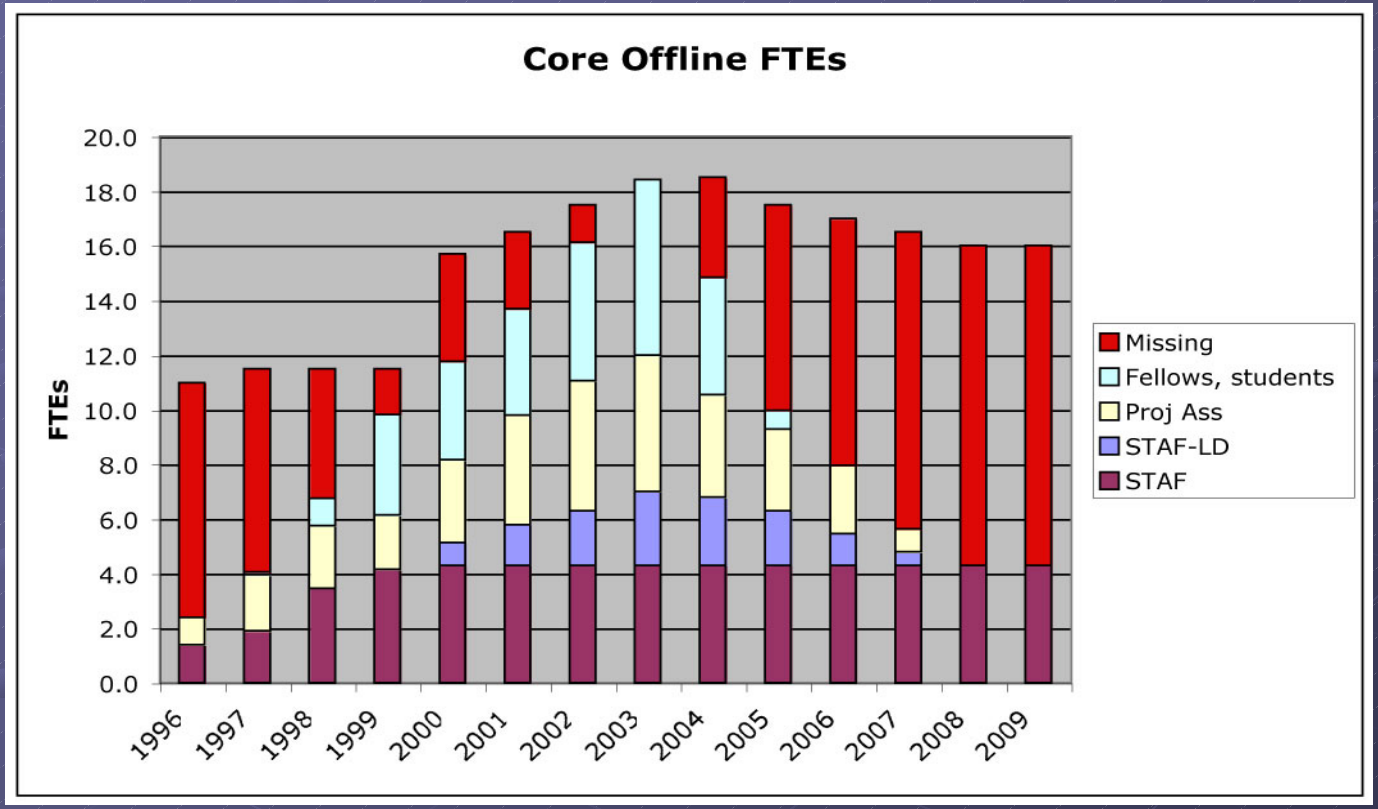
- Long build-up time
- Must sustain plateau after 2003

Personnel Profile (post oriented)

- Core Offline at CERN :
  - 4 permanent CERN staff
  - Temporary CERN personnel (no replacement assumed*)
    - Staff LD
    - Technical and Doctoral students
    - CERN Fellows
  - Temporary CERN Project Associates (direct contribution from collaboration institutes + ALICE CERN exploitation budget ; no replacement assumed*)

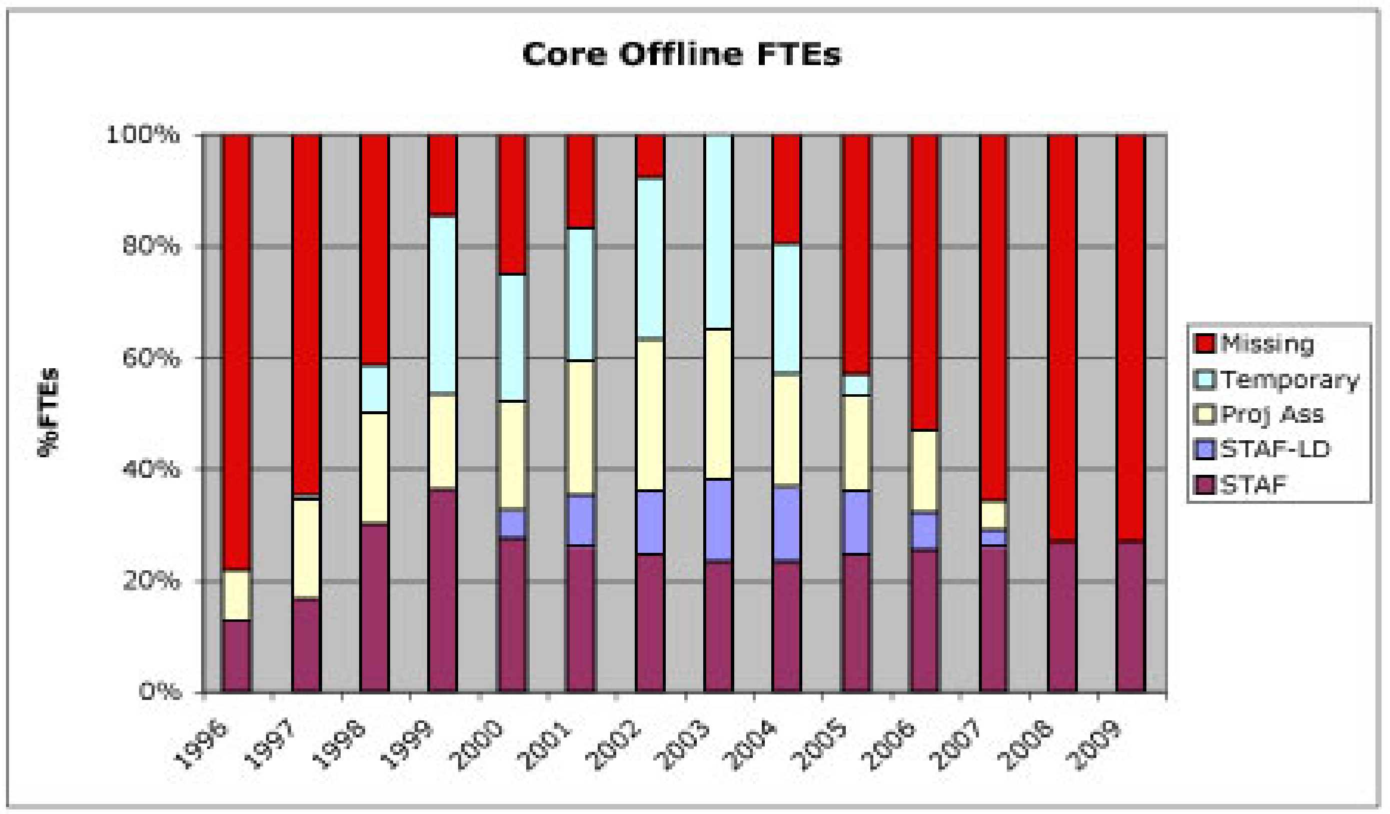
* Pessimistic scenario to emphasize fragility of the structure

Personnel Profile (profile oriented) - 1/2



- Mostly temporary personnel
- Substantial contribution from collaboration institutes
- ROOT effect in 1999, AliEn effect in 2003

Personnel Profile (profile oriented) - 2/2



- Only 25% permanent personnel
- More than 60% are short term personnel

Out-sourced projects - 1/3

- Detector DB by Physics Department and Computer Science Department @ Warsaw University : a single DB (economy of personnel) common to all detectors in the experiment

[redacted]	**2003**	**2004**	**2005**	**2006**
Total	2.9	2.9	2.9	2.9

Out-sourced projects - 2/3

- EDG TestBed validation and participation in various GRID projects by ALICE/Italy, ALICE/US, and the EDG/DataTAG project; to be continued with EGEE

	2003	2004	2005	2006
Total	4.1	4.1	3.1	3.1

Out-sourced projects - 3/3

- AliEN: basis of the ALICE distributed computing infrastructure : Coordination and main development by Core Offline group but several specific sub-tasks delegated to individuals at remote places

	2003	2004	2005	2006
Total	5.1	4.2	3.9	1.3

Resources summary

- Distribution of personnel for common offline activities

Activity	**2003**		**2004**		**2005**		**2006**	
	FTE	%	FTE	%	FTE	%	FTE	%
Core Offline	18.0	59.8	18.5	62.3	17.5	63.9	17.0	70.0
GRID testing	4.1	13.6	4.1	13.8	3.1	11.3	3.1	12.8
Distributed comp inf	5.1	16.9	4.2	14.1	3.9	14.2	1.3	5.3
DCDB	2.9	9.6	2.9	9.8	2.9	10.6	2.9	11.9
Total	30.1	100.0	29.7	100.0	27.4	100.0	24.3	100.0

- About 40% of the work is distributed outside CERN

HLT Software

- Only personnel working on algorithms and simulation in collaboration with Offline project

Activity	2003	2004	2005	2006
Needed	4.5	4.5	4.5	4.5
Available	4.1	3.9	2.9	2.9
Missing	0.4	0.6	1.6	1.6

- Part of missing personnel should come from PhD students

LCG projects in application area

- ALICE has already made most of choices, after evaluation and assessment, for critical issues and demonstrated them (persistency, data DB, tracking, geometry descriptor, distributed computing, etc…)
- Does not need to rely on common LCG applications

- However ALICE contributes to common developments (substantial effort on G4, Objty before 2003):

Item	Name	Task	2003	2004	2005	2006
ROOT						
	F.Rademakers	Root support and Proof development	0.3	0.3	0.3	0.3
	A.Gheata	Virtual MonteCarlo & Geometry Modeller	1.0	1.0	1.0	1.0
	M.Gheata	Virtual MonteCarlo & Geometry Modeller	0.5	0.5	0.5	0.5
	R.Brun	Root development	0.1	0.1	0.1	0.1
	G.Ganis	Proof development	1.0	1.0	1.0	1.0
Virtual MonteCarlo						
	E.Futo	Integration of FLUKA	1.0	0.5	0.0	0.0
	A.Fasso	Integration of FLUKA, FLUKA support	0.5	0.0	0.0	0.0
	I.Gonzalez	Verification of G4	0.1			
	A.Morsch	Integration of FLUKA & G4, Virtual MonteCarlo development	0.1	0.1	0.1	0.1
	F.Carminati	Virtual MonteCarlo	0.1	0.1	0.1	0.1
	I.Hrivnacova	Integration of G4 & Virtual MonteCarlo	0.2	0.2	0.2	0.2
Total			4.9	3.8	3.3	3.3

- To come : AliEn coupled with PROOF as generic architecture for LCG interactive analysis

Other resources

- EU project : EDG , DataTag for ALICE
- External collaborations :
  - IRST Trento : Code checker
  - Ericson : AliEn (monitoring & software distribution)
  - Nasa : Virtual Monte-Carlo, Geometrical Modeler, FLUKA physics developments

Offline in detector projects - 1/3

- AliRoot: An object Oriented framework which directly uses ROOT and provides:
  - Many event generators
  - Tracking using Virtual Monte-Carlo
  - IO infrastructure
  - Steering functionalities
  - Global reconstruction

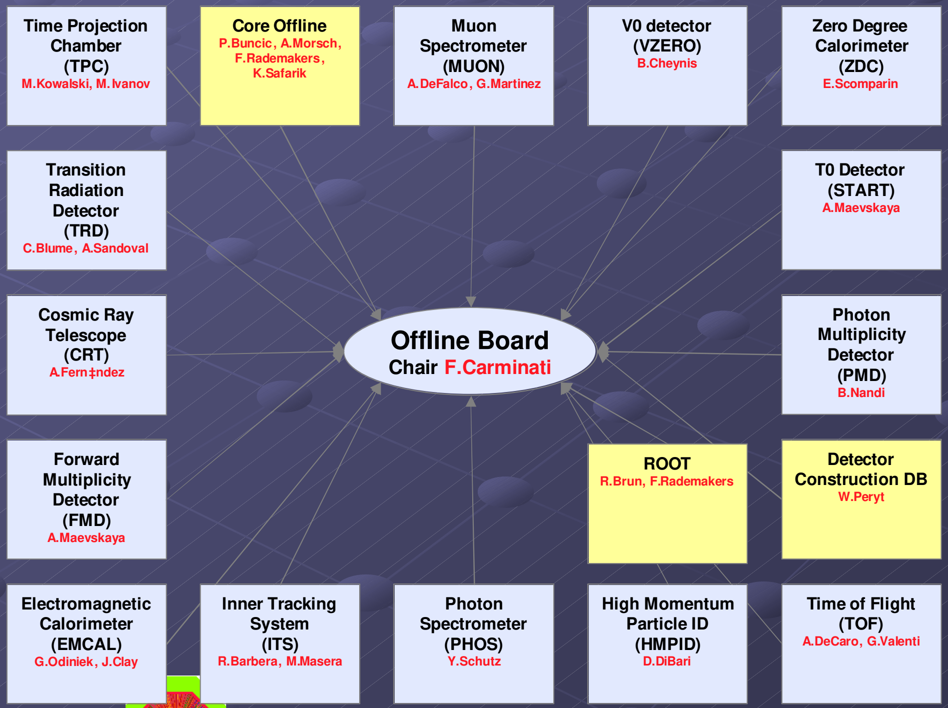
Core Offline team

  - Detector (13) tracking and reconstruction
  - Analysis

Detector Groups

Offline in detector projects - 2/3

- No full-time dedicated developers
- Schedule defined by global milestones (DC)
- Planning is task oriented rather than personnel oriented



Time Projection Chamber (TPC)
M.Kowalski, M.Ivanov

Core Offline
P.Buncic, A.Morsch, F.Rademakers, K.Safarik

Muon Spectrometer (MUON)
A.DeFalco, G.Martinez

V0 detector (VZERO)
B.Cheynis

Zero Degree Calorimeter (ZDC)
E.Scomparin

Transition Radiation Detector (TRD)
C.Blume, A.Sandoval

T0 Detector (START)
A.Maevskaya

Cosmic Ray Telescope (CRT)
A.Fern‡ndez

Offline Board
Chair F.Carminati

Photon Multiplicity Detector (PMD)
B.Nandi

Forward Multiplicity Detector (FMD)
A.Maevskaya

ROOT
R.Brun, F.Rademakers

Detector Construction DB
W.Peryt

Electromagnetic Calorimeter (EMCAL)
G.Odiniek, J.Clay

Inner Tracking System (ITS)
R.Barbera, M.Masera

Photon Spectrometer (PHOS)
Y.Schutz

High Momentum Particle ID (HMPID)
D.DiBari

Time of Flight (TOF)
A.DeCaro, G.Valenti

Offline in detector projects - 3/3

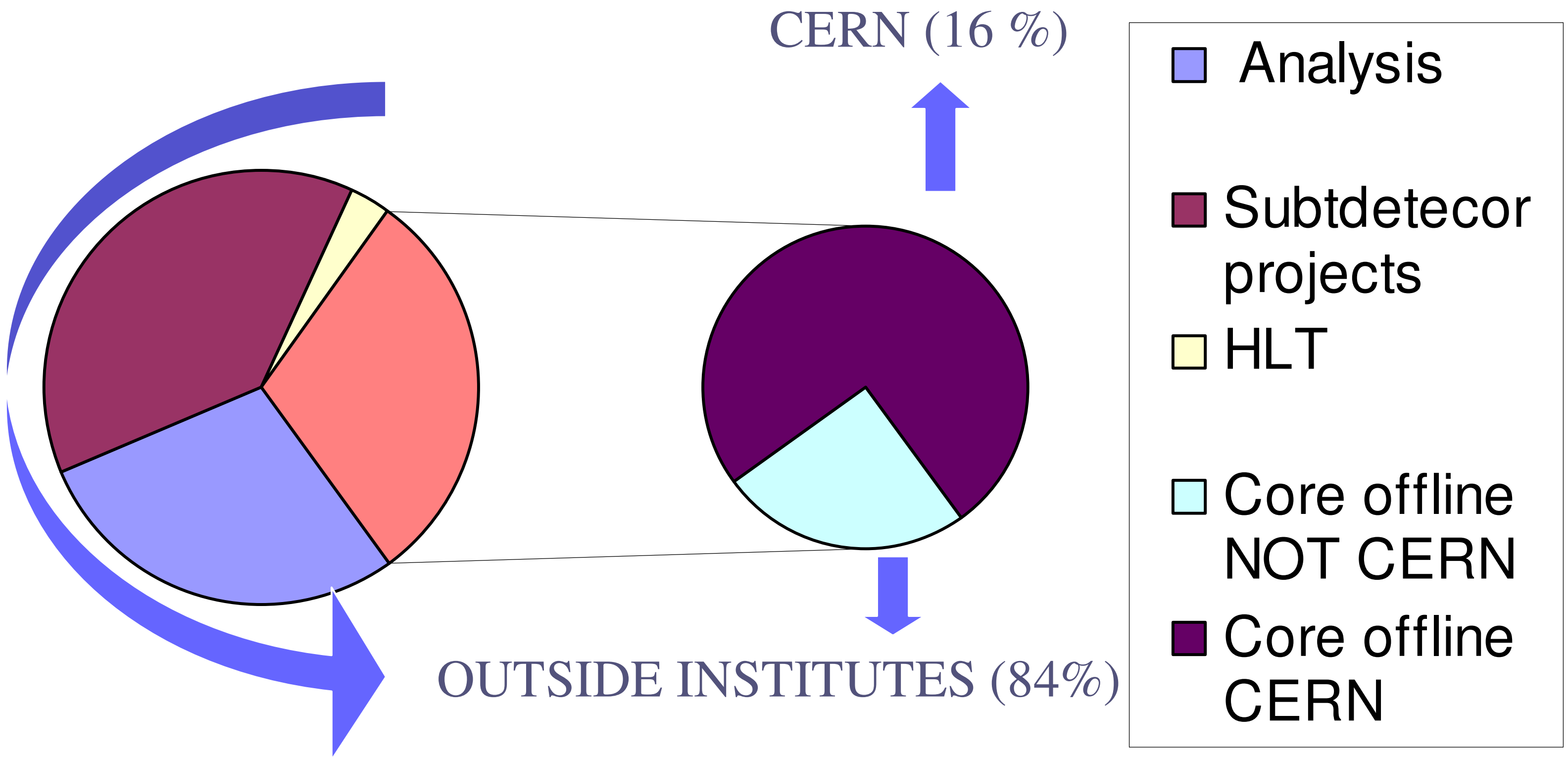
Activity	Name	2003	2004	2005	2006
TRD	**Available**	4.2	3.4	3.0	3.0
	Missing	0.3	1.3	1.5	1.5
PMD	**Available**	3.2	3.3	3.3	3.3
	Missing	0.5	0.5	1.0	1.0
T0/FMD	**Available**	1.4	1.4	1.4	1.4
	Missing	1.8	1.8	1.8	1.8
PHOS/EMCAL	**Available**	2.1	2.1	2.1	2.1
	Missing	1.6	1.6	1.6	1.6
ITS	**Available**	10.7	10.1	9.4	9.4
	Missing	0.5	3.0	3.0	3.0
ZDC	**Available**	1.0	1.0	1.0	1.0
	Missing	0.0	0.0	0.0	0.0
CRT	**Available**	1.6	1.6	1.6	1.6
	Missing	1.0	1.0	1.0	1.0
V0	**Available**	1.0	1.0	1.0	1.0
	Missing	0.0	0.0	0.0	0.0
Muon	**Available**	2.5	2.0	2.0	2.0
	Missing	1.0	1.5	1.5	1.5
TOF	**Available**	6.3	6.3	6.3	6.3
	Missing	0.0	0.0	0.0	0.0
TPC	**Available**	3.1	2.4	2.0	2.0
	Missing	0.9	1.6	2.0	2.0
RICH/HMPID	**Available**	2.7	2.7	2.7	2.7
	Missing	1.0	1.0	1.0	1.0
	Needed	**8.6**	**13.3**	**14.4**	**14.4**

Personnel resources in Offline project

Activity	Item	2003	2004	2005	2006
Common software infrastructure	Needed	30.1	29.7	27.4	24.3
	Available	30.3	26.1	19.9	15.3
	Missing	0.2	3.7	7.5	9.0
Subdetector Software	Needed	46.3	48.6	48.2	48.2
	Available	37.7	35.3	33.8	33.8
	Missing	8.6	13.3	14.4	14.4
Analysis	Needed	52.0	54.0	58.0	61.0
	Available	50.7	51.2	48.3	45.4
	Missing	1.3	2.8	9.7	15.6
HLT	Needed	4.5	4.5	4.5	4.5
	Available	4.1	3.9	2.9	2.9
	Missing	0.4	0.6	1.6	1.6
Total	Needed	132.9	136.8	138.1	138.0
	Available	122.8	116.5	104.9	97.4
	Missing	10.1	20.4	33.2	40.6

- About 16% of the personnel at CERN, the remainder in collaboration institutes, no experiment dedicated personnel at regional centers.

Personnel resources in Offline project



How to alleviate the lack of Personnel

- The ALICE off line project is committed to provide the collaboration with the adequate software to analyze data starting 2007.
- The project has already adapted its strategy to the shortage of personnel and aims toward a bare minimum which enables to fulfill its tasks.
- Any reduction in the Core Offline team personnel will put in danger the success of its goals.
- The lack of personnel in the detector projects will translate in lack of readiness in terms of accuracy in the algorithms and in lack of availability of categories of algorithms.
- Such a situation will have a negative impact on the quality of physics results.

ALICE priorities - 1/4

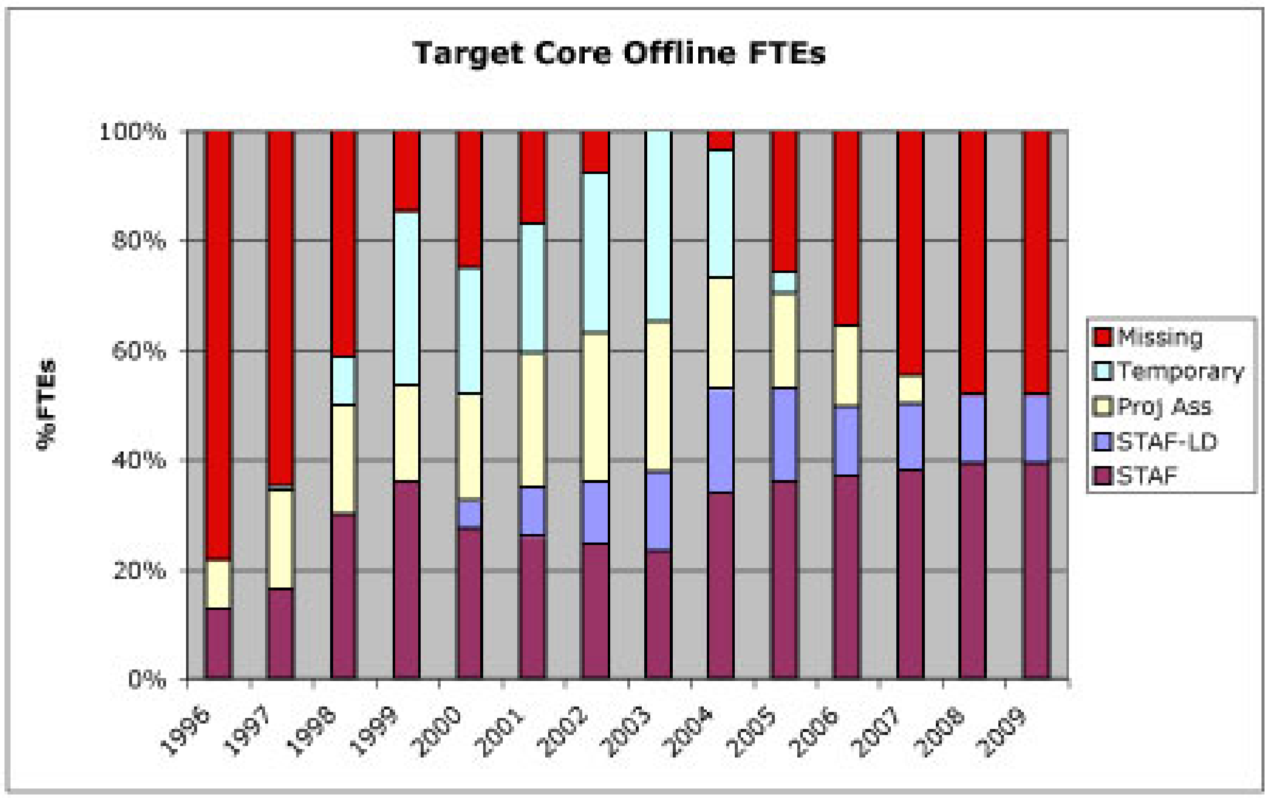
- Core Offline group at CERN :
  - Less than 1/4 of personnel in Core Offline group at CERN are permanent
  - More than 50% are short term personnel
  - Dependence on availability of short term CERN positions
    - Uncertainty on renewals
    - Loss of knowledge -- difficulty of knowledge transfer
    - Difficulty to cover key positions with people with the appropriate profile
    - Competition within ALICE in a fixed quota situation

ALICE priorities - 2/4

- Core Offline group at CERN :
  - Have at least 1/3 of long-term personnel, limit use of fellows and students to 1/2, without changing the target number of FTEs
  - Ensure the covering of key areas by converting two area coordinators (Production Environment, Framework & Infrastructure) now on temporary positions into CERN permanent staff
  - Alleviate the “volatility” of Core Offline Team with at least two long-term (6 years, LD-like) positions at CERN to replace short-term ones (Detaching LCG personnel to ALICE would be a natural solution)

ALICE priorities - 3/4

- Target for Core Offline group at CERN :



ALICE priorities - 4/4

- Detector Offline at collaboration institutes :
  - About 10 FTEs missing in the subdetector projects for software developments
  - This is a responsibility of the Institutes in charge of the subdetector projects
  - We are working hard to find these people
  - Additional resources from funding agencies will have to be discussed case-by-case

Conclusions - 1/2

- The core team of the offline project has adapted to the reduced personnel available and established its tasks and objectives accordingly.
- The edifice is fragile : any cut in (temporary) personnel might hinder the availability in due time of the software needed for data taking and analysis.
- Securing two additional permanent staff positions for 2 area coordinators, now on medium-term position, is crucial for the project success.

Conclusions - 2/2

- Replacing 2-3 short term personnel with medium term staff (computing physicist) would alleviate the unstable situation by making it less dependent on temporary personnel.
- The lack of personnel dedicated to software development in the detector projects is worrisome as the absence of indispensable algorithms might substantially delay first physics results from LHC.
- The needed personnel (not necessarily computer specialists) must be recruited by the institutes of the collaboration.