

Requirements for a Regional Data and Computing Centre in Germany (RDCCG)

Peter Malzacher, GSI, ALICE
Andres Sandoval, GSI, ALICE
Lutz Köpke, Universität Mainz, ATLAS
Alois Putzer, Universität Heidelberg, ATLAS
Dankfried Lanske, RWTH Aachen, CMS
Günter Quast, Universität Karlsruhe, CMS
Volker Lindenstruth, Universität Heidelberg, LHCb
Michael Schmelling, MPI Heidelberg, LHCb

Marcel Kunze, Universität Bochum, BABAR
Klaus R. Schubert, TU Dresden, BABAR
Thomas Müller, Universität Karlsruhe, CDF
Lars Schmitt, TU München, COMPASS
Peter Mättig, Universität Wuppertal, D0

Peter Braun-Munzinger, GSI, for KHK
Rolf Heuer, Universität Hamburg, for KET
Rainer Mankel, DESY, for DESY

Abstract

A Regional Data and Computing Centre in Germany (RDCCG) is proposed by the German Particle and Nuclear Physics Community. The steps towards the production phase are described, including hardware requirements, expected manpower support from the experiments, deliverables, and milestones. A proposal is made for the management structure of this centre.

1	Executive Summary	1
2	Introduction	2
3	General Requirements of the LHC Experiments	3
3.1	The LHC Experiments	3
3.2	The LHC Computing Challenge	3
3.3	The Steps towards a Production Environment	4
3.4	Hard- and Software Contingency Requirements	5
4	Specific Needs of the German LHC Groups	6
4.1	The German Participation	6
4.2	ALICE	6
4.3	ATLAS	8
4.4	CMS	9
4.5	LHCb	10
4.6	LHC summary	11
5	Requirements of the Non-LHC Experiments	13
5.1	BABAR	13
5.2	CDF and D0 Experiments at the Tevatron Collider	14
5.3	COMPASS	15
6	Overview and advisory structure	17
7	Requirements for the Operation of the RDCCG	18
7.1	Memorandum of Understanding	18
7.2	Accountability	18
7.3	Hardware Compatibility	18
7.4	Software Compatibility	18
8	Steps towards the RDCCG	21
8.1	Phase 1: Development and Prototype Construction, 2001-2004	21
8.2	Phase 2: Installation of the LHC Computing Production Facility, 2005 – 2007	23
8.3	Phase 3: Maintenance and Operation, 2007 – end of LHC data analysis	23

1 Executive Summary

The main aim is to provide a computing environment for the German Particle and Nuclear Physics Communities capable of meeting the challenges posed by the unprecedented computing requirements of the LHC experiments¹. It is proposed to establish a Regional Data and Computing Centre in Germany (RDCCG) to serve as Tier1 centre for the four LHC experiments, thus assuring competitive analysis resources for the German research groups involved. This RDCCG should also serve other major data intensive non-LHC experiments of both communities. A later extension to other sciences should be envisaged.

The extreme requirements in terms of computing, data storage and networking that the LHC experiments will need, requires the development and certification of a new computing paradigm that is called Grid computing². The Grid technologies will be developed in coordination with the LHC Tier0 and other Tier1 centres. Close collaboration with ongoing (e.g. DataGRID, PPDG and GriPhyN) and future Gridactivities is important to ensure the developments of a common set of tools, principles, protocols and standards that can support a widest range of applications. The centre should both have centralised hardware resources and act as a focal point for software development. Advantage should be taken of competence in computer science and Grid computing already available in Germany. The centre should encourage other branches in natural sciences to join activities in this promising field. It should take a leading role in education of young people in the rich area of applied computing and thus constitute a visible investment into the future of research.

The RDCCG should be developed in three stages. Between 2001 and 2004 it should serve as test bed for developing the infrastructure of LHC computing. At this first stage much experience could be gained from serving active non-LHC experiments. By the year 2004 a complete appraisal of the developments should be made before establishing the final RDCCG data infrastructure. The years 2005 to 2007 will be used to essentially finalize the computing infrastructure for the LHC experiments. From 2007 onwards the centre has to adapt to the increased needs of the ongoing LHC experiments. The centre has to remain fully functional for the whole duration of the LHC experiments, of at least 15 years.

The RDCCG should be integrated in the world-wide computing scheme of the different experiments. A management structure is needed with clear lines of responsibility and reporting. Also well defined milestones and deliverables have to be established to allow the monitoring of the progress from prototype to final implementation for the RDCCG. The management structure as well as milestones and deliverables will have to be established in agreement with all participating parties. A proposal is given in this document.

¹ For a recent account see: CERN/LHCC/2001-004, Report of the Steering Group of the LHC Computing Review

² See e.g.: I. Foster and C. Kesselman (ed.), The GRID, Blueprint for a New Computing Infrastructure, San Francisco 1999

2 Introduction

The key aim for setting up a German Regional Data and Computing Centre (RDCCG) is to provide a computing environment for the German Particle and Nuclear Physics Communities capable of meeting the challenges posed by the unprecedented requirements of the LHC experiments. It is a central assumption that these needs can best be met by harnessing internationally distributed computer power and storage capacity efficiently and transparently.

The emerging Grid technology is a natural framework within which to develop this capability. It is therefore an important objective that this development should provide the functionality required for full exploitation of the LHC. The requirements for Particle Physics should drive the development of a generic framework, which also will have a potential benefit to other sciences and the wider community. The tools and infrastructure will certainly have more general applications, some with definite commercial implications.

The process of creating and testing this computing environment for the LHC will naturally embrace the needs of the current generation of highly data-intensive experiments (in particular the BABAR experiments at SLAC, CDF and D0 at the Tevatron and COMPASS at CERN). These provide well-defined and well-focussed intermediate scale problems that can be used to “field test” the performance under actual running experiment conditions. This both complements the development of prototype systems for the LHC and deepens the experience in Germany in the issues to be tackled in delivering the robustness that will be essential for the LHC and other future applications. DESY and GSI support the foundation of a German regional centre as an important prerequisite for the German LHC groups to fully exploit the LHC physics program.

The project is a first step into a new field of general computing. The experiments at CERN combine for the first time two branches of physics at a single site in a common computing environment. In addition experiments situated at research centres in the US at Stanford and Chicago intend to share resources with their software environments. The ambitious goals of DataGRID, opening world-wide computing resources to a heterogeneous world of clients, will experience a basic test of realizability.

The nature of the Grid requires the development of a common set of tools, principles, protocols and standards that can support the wide range of applications envisaged. Collaboration with initiatives such as the DataGRID, PPDG, and GriPhyN is taken to be essential.

A vital element of the success of this project is a cohort of young scientists, who will be acquiring the skills needed to develop the Grid technology not only for Particle Physics but also for broader applications.

Although the primary focus of this project is the development of the computing infrastructure required to support Particle Physics, particularly the LHC experiments, it is an important aim to be open for collaboration with other sciences, industry and commerce. Close collaboration with Computing Science groups with the skills to deliver the necessary products are a feature, as are close links with other disciplines whose requirements could eventually match or exceed those of the LHC.

In the following sections the details of the requirements for the German Regional Data and Computing Centre are outlined.

3 General Requirements of the LHC Experiments

3.1 The LHC Experiments

Starting in 2006 the LHC accelerator will produce proton-proton collisions at 14 000 GeV (centre of mass) with an eventual rate of 10^9 events/s at design luminosity, which, for a multipurpose detector, corresponds to more than 10^{11} particles/s to be recorded. Events of fundamental scientific interest such as the production of Higgs particles decaying into detectable decay modes are predicted to occur at an approximate rate of 1 in 10^{12} of the average proton-proton collisions. Operating as a Heavy Ion accelerator, LHC will produce collisions at 5700 GeV per nucleon (centre of mass). The charged multiplicities for a single event may reach 8000 particles per unit of rapidity. The study of hot and dense QCD plasma matter, and the search for parton deconfinement, symmetry restoration and QCD phase transitions are central to the Heavy Ion program.

The LHC collider will serve 4 experiments with a wide range of physics programs. The ALICE experiment makes use of the Heavy Ion mode of the LHC and shall explore the fundamental properties of partonic matter beyond a new threshold of centre of mass energy. The two experiments ATLAS and CMS shall provide a new general insight into the appearance of particle physics and the fundamental forces at a new scale of energies. Both experiments will push the frontier in energies in Elementary Particle Physics by more than one order of magnitude compared to Tevatron Experiments at FNAL in the US. The LHCb experiment is dedicated to the investigation of b-quark particles. These particles are copiously produced at LHC energies and will reveal key information to open questions on the broken symmetries in nature. Each experiment will thus record about 10^9 events per year of operation, corresponding to a yearly mass-storage accumulation of around 7 Petabytes for all experiments. To this must be added the requirements for reconstructed and simulated data, and all calibration information, leading to a total mass storage at CERN in excess of 11 Petabytes/year for all LHC experiments.

The LHC community comprises more than 5000 physicists, residing in about 300 institutes in 50 countries. Providing transparent and efficient access to the very large amounts of data for all scientists will be a major challenge for networking.

To get an adequate chance to contribute to the physics program in a way proportional to their hardware investments the German LHC groups apply for a Regional Data and Computing facility in Germany fitting into the computing challenge of the LHC experiments.

3.2 The LHC Computing Challenge

A suitable coordinated world-wide computing fabric must be created to manage the data, from the first recording and near-line reduction down to the desktop of the physicist doing analysis. This fabric must be able to cope with the experiments' full analysis chains and with all of the very large data sets accumulated over many years.

This world-wide fabric, its organization and use by a multitude of distributed user groups, pose very novel demands on scalability, authentication, management of distributed resources, and coordinated collaboration amongst many computer centres. The effective use of the distributed system requires the application of a new combination of local and global policies for prioritising use, as well as new methods and strategies for cost/performance optimisation. In terms of recent computer science terminology, this would be the first and ambitious realization of a world-wide data-, compute- and user-intensive Grid structure.

A multi-Tier hierarchical model similar to that developed by the MONARC³ project is the key element of the LHC computing model. In this model, for each experiment, raw data storage, first level reconstruction and production of event summary data will be carried out at the Tier0 centre, based at CERN. Analysis and

³ See: MONARC Phase 2 report CERN/LCB 2000-001, March 2000,
<http://monarc.web.cern.ch/MONARC/docs/phase2report/Phase2Report.pdf>

Monte-Carlo data generation, partial data storage of event summary data and user support will mainly be the task of several Regional Tier1 centres. The Tier2 centres possibly located in a national centre of lower capacity - will participate in the Monte-Carlo production and mainly serve to produce and store Analysis Object Data and sub datasets for Derived Physics Data (DPD) and selectively tagged physics analyses (TAG), whereas the Tier3 centres for a university group and the end-user workstations (Tier4) take the load of the individual physicist's all day analysis. Quality enhancements due to calibrations and subsequent re-reconstructions are intended to be shared between the Tier0 and the Tier1 centres according to available computing and data transfer capacities.

Tier0 and basically also Tier1 centres are open to all members of a Collaboration (up to 500 active users per experiment) under conditions to be specified in Computing MoUs (memorandum of understanding). This automatically implies that the Tier1 centres must be available for the lifetime of the LHC, and that the Tier0 and all Tier1's together will be managed coherently.

Grid technology will be explored to contribute solutions to this model that provide a combination of efficient resource utilization and rapid turn around time.

Estimates of the required bandwidth of the wide area network between Tier0 and Tier1 centres arrive at 1.5 to 3 Gb/s for a single experiment. The traffic between the Tier1 centres will be comparable. While technology will certainly be able to deliver such rates, it is vital that a well-supported Research Network infrastructure with sufficient bandwidth is available, at affordable costs, by the year 2006.

Tier1 centres serving all LHC experiments are presently foreseen in the following European countries:

- France (Lyon, IN2P3)
- Italy (INFN)
- UK (RAL; Rutherford Appleton Lab)

In the US the Tier1 centres are separated with respect to the experiments:

- ATLAS (BNL; Brookhaven Nat. Lab.)
- CMS (FNAL; Fermi Nat. Lab.; Chicago)

Other regional centres (of smaller size) are currently under discussion in e.g. in Canada, China, India, Japan, Russia, Scandinavia, Switzerland.

3.3 The Steps towards a Production Environment

The LHC experiments do not expect data taking to start before 2006. Nevertheless, the multiple challenges of data recording at such high rates, between 50 Mbytes and 1250 Mbytes per second and experiment and their subsequent handling, storage and analysis, require planned and controlled steps from today into a working production computing environment, when data recording starts. Besides real data reconstruction simulated particle reactions being complete look-alikes of real data are needed at the same amount as real data to correct for background and detector effects and to optimise detector parameters and behaviour. The development of detector specific simulation software and its test and application are the main computing effort, before the experiments come into action. From now until 2004 a prototype phase is planned to develop, install and test the general computing model and the necessary general and experiment dependent tools. This will be done in a reduced configuration (approximately 50% in complexity of a large experiment in 2007). This prototype of a distributed computing system of resources will be fed with simulated data and will be used to prototype via so called production challenges of defined sizes aimed at proving its functionality for the production phase.

All LHC experiments plan "Data Challenges", addressing different critical points of their computing model with samples of simulated data and to study the physics, trigger and detector performance. These Data Challenges will include, besides CERN resources, preferentially all of the Tier1s and the available Tier2s to solve the deployment issues of the entire system, from DAQ flow and on-line farm to the physics analysis. Coordination amongst the experiments is needed to obtain resource-sharing benefits.

At the same time these developments and tests serve as the input to the planning of the basic first production environment, thus paving the road to an optimised investment at the latest possible time for financial and technical effectiveness.

When the functionality is proven, the final hardware configuration must be set up and brought into full operation.

3.4 Hard- and Software Contingency Requirements

The choices to be made concerning the architecture of the whole system can strongly influence the hardware needs and the human resources required to support production installations. Distributed computing models and the emergence of a hierarchical structure of Tier centres may constrain the hardware and software choices of the centres. For a given level of human resources, each centre can provide more and more efficient services if attention is given to the question of standardisation.

In what concerns operating systems, there is already a well-established trend towards one single Unix-like system, namely LINUX, or at most two, consisting of LINUX with the RedHat dialect and one commercial UNIX OS. There are in fact good arguments to have a second platform for code and result verification.

As far as the data persistency solution is concerned, recent experience shows that a large number of different solutions, especially commercial ones, has a strong impact on resources and should be kept to a minimum. It was agreed that a maximum of two choices for object persistency systems should be supported, one commercial and one Open Source.

These general considerations are overall guidelines in the development of the planned centre. It is self-evident that in the beginning it can only be useful to the experiments, if the environmental conditions are entirely compatible with the corresponding Tier0 centre, serving as the core of software distribution. This can only be achieved by a type of cloning strategy, duplicating user administration tools and procedures, operating system, compilers, libraries, key software tools and key hardware components - assuring the practical ignition conditions for the individual experimental software environments. The more the Grid tools get developed, to create look-alike fail-safe structural elements at all Tier levels, the less the elements – especially individual hardware choices – will be visible to the users and therefore important. The ignition conditions will have to be verified by the experiments. In general the process of adapting the experimental software to the software environment of the centre is understood to be an iterative one, but should stay as simple and modular as possible for the experiments, also in prevision of a possible later extensions of the participation to the centre.

4 Specific Needs of the German LHC Groups

4.1 The German Participation

At present, 25 German research and university groups participate in the four LHC experiments. The ALICE experiment, using the LHC as a Heavy Ion collider, represents a community of more than 150 physicists in the field of traditional Nuclear Physics. Among the Particle Physics experiments, ATLAS is joined presently by 10 German university research groups, the other experiments (CMS and LHCb) being somewhat smaller as far as their German contribution is concerned, but not at all in their overall size and technical challenge. Today all experiments are predominantly involved in the development and the construction of hardware elements of the four detectors and in the design and development of their software infrastructure. Each group is aiming to deliver a considerable contribution to the physics analysis as well, to get an appropriate harvest compared to the hardware and software investment provided.

The ALICE collaboration is planning for 10^6 sec/year Pb-Pb running with a trigger rate of 50 Hz and an average event size of 25 MB. ATLAS and CMS consider recording rates around 100 Hz, with raw event size in the 1 MB range. However, these values still have basic uncertainties. For instance, studies in ATLAS with fully simulated events have shown that recording rates can reach up to 270 Hz, if one wishes to perform b-physics. If zero suppression cannot be done on-line, the raw event size may reach 2 MB in the first month of data taking. LHCb foresees a 200 Hz recording rate, but with a smaller event size (0.125 MB). The ratios of the computing facilities installed at CERN and outside will be close to the 1/3 - 2/3 ratio traditionally adopted by the CERN policy makers. The four experiments plan to keep all their raw data at the Tier0 centre (CERN) and perform there the first reconstruction.

The requirements of the 4 LHC experiments in the different phases are detailed in the following.

4.2 ALICE

The German ALICE groups are at present:

GSI Gesellschaft für Schwerionenforschung, Darmstadt
Universität Frankfurt
Universität Heidelberg (2 Institutes)
Universität Münster

4.2.1 Program and status

The ALICE experiment at the CERN LHC has been optimised for the study of collisions of Heavy Ions at a centre of mass energy of 5700 GeV per nucleon that can produce in a single central interaction up to 60000 charged particles of which 20000 would be detected in the Central Barrel detectors. The experiment differs substantially from the three other LHC experiments in its physics program and has obviously specific offline needs, as well as in its software environment. Although the experiment is dedicated to fundamental Nuclear Physics, it will also participate in the data taking in the proton-proton collider mode of the LHC together with the ATLAS, CMS and LHCb experiments.

The German groups within ALICE are responsible for two of the main detectors in the Central Barrel: the Time Projection Chamber (TPC) and the Transition Radiation Detector (TRD), as well as for the Level 3 Trigger Farm.

4.2.2 Computing and Storage needs

The German groups in ALICE plan to join a Tier1 Regional Data and Computing Centre (RDCCG) to participate in the analysis of the ALICE data. This centre will be shared with the other three LHC experiments in order to benefit from the synergy of common approaches and to optimise the use of personnel, hardware and software. It is also planned to establish a Tier2 Regional Centre at GSI to complement the RDCCG in the area of interactive physics analysis based on the Derived Physics Data (DPD) and event Tags (TAG) produced at the Tier0 and Tier1.

The ALICE computing model currently assumes that the raw data (RAW) are collected and written to tape at CERN in the Tier0, where also the first pass of analysis is carried out. The created Event Summary Data (ESD) are then distributed to the various Tier1 Regional Centres for further processing during which the physics Analysis Object Data (AOD) and the DPD and TAG data are produced. The RDCCG Centre participates also in the production and analysis of Monte Carlo events of Simulated Raw Data (SIM) which is a crucial component for the evaluation of the acceptances and efficiencies of the different detectors in the ALICE experimental set-up. There will be also a small percentage of the RAW data that will be resident at the RDCCG Centre and which will be processed many times for detailed detector performance studies.

Between now and the start-up of data taking in 2006, ALICE will carry out a number of data challenges according to an aggressive plan leading in 2007 to an overall performance of 1.25GB/s. These are yearly exercises of one or two months in which the integrated DAQ and offline system and the computing technology relevant for ALICE is assessed. In 2002 it is planned that the future RDCCG actively participates in this efforts.

4.2.3 Manpower and Operations

The RDCCG computing requirements needed to fulfil its mandate are listed below. It is hard to extrapolate over seven years not knowing how the technology will evolve, particularly in the storage area of disks versus tapes and therefore these values are only indicative and will have to be reassessed every year.

The usage of the RDCCG can be divided into four different categories:

- Simulations
- Participation to Data Challenges
- Production of Analysis Object Data
- Physics Analysis

The work in the first three categories will be done mostly in a coordinated and centralised way.

From the collaboration there will be a coordinator in charge for any of these four tasks, each one of them supported by one or several coworkers to run the production day by day. For large production jobs they will install the certified version of the collaboration software, submit the jobs, verify and log the results, check for misbehaviour of any element, software or hardware of the computing fabric and inform the computing centre designed persons of it.

The general software environment foreseen at present, Intel PCs running RedHat Linux, does fit into the offline requirements of the ALICE experiment. In addition it will be necessary to install and support a batch submission system (at the moment preferably LSF), a data storage access system (at the moment preferably CASTOR), and Grid tools including the GLOBUS toolkit and the middleware developed by the DataGRID project. These are required at start-up to guarantee a working ignition phase in the first 2 years. The ALICE collaboration will install Root and AliRoot, its software framework.

During the prototyping and build-up phase of the RDCCG in the years 2001 up to 2004 it is necessary to have a minimum amount of resources available for Data Challenges and other studies during a peak time of 3 months every year. During the rest of the year only one third of these resources are needed and therefore the rest of the computing capacity can be shared with the other experiments utilizing the RDCCG. Of these 3 months a year 2 months have to coincide with the ALICE Data Challenges that are normally run in the spring of every year and the third month can be scheduled with more flexibility. In the table below the resources needed at peak time are listed:

Investment/year	2001	2002	2003	2004	2005	2006	2007
CPU (kSI95)	0.3	1	3	10	20	50	220
Disk Space (TB)	1	2	5	10	30	60	200
Tape (TB)	1	5	10	20	50	100	500
Act. users (cum)	2	5	10	20	30	60	150

4.3 ATLAS

The following nine German university groups and the MPI in Munich are members of the ATLAS collaboration:

Universität Bonn, Universität Dortmund, Universität Freiburg, Universität Heidelberg, Universität Mainz, Universität Mannheim, Universität München, MPI München, Universität Siegen, Universität Wuppertal

4.3.1 Program and Status

ATLAS is a large multipurpose detector under construction, designed to measure all accessible physics processes of interest in proton-proton collisions at the LHC. As of today, the German ATLAS groups represent about 125 physicists with essential hardware responsibilities for the inner tracker (pixel and silicon strip detectors), the liquid argon calorimeters, the chambers for the large toroidal muon spectrometer as well as for electronics and software of trigger levels 1,2 and 3.

4.3.2 Computing and Storage needs

The ATLAS computing model assumes that Raw data are only kept at the CERN Tier0 centre, which will also provide a first pass processing to produce event summary data (ESD). These data, which are reduced by a factor of 8 with respect to the raw data, are distributed to the various Tier1 centres where the data are further analysed in several passes. Analysis object data (AOD) and even more reduced samples will then be made available. Monte Carlo production of simulated raw data will also be done at the Tier1 centres, which will in addition keep small samples of detector RAW data for in-detail investigations of the detectors. A specific model for ATLAS data flow in a German Tier1 centre has been simulated.

Levels of integration have been defined as milestones and corresponding tests have been scheduled within ATLAS. Mock data challenges at the Tier1 centres are foreseen for 2001 – 2003 to test the full software chain, analysis procedure and the Grid infrastructure for 10^8 simulated events. For the German Tier1 centre to participate in this endeavour, the elements of the production software, right now being installed at the various other Tier1 sites, will have to be installed and tested with priority. They will later serve as reference to the Tier3 installations at the institutes. Otherwise parallel human efforts will thus get concentrated. In the long run, the existence of a Tier1 centre in Germany at compatible strength to the prominent European collaborators is a precondition to provide adequate chances for the German groups to contribute to the physics analysis according to their detector investments.

4.3.3 Manpower and Operations

Maintenance and installation of ATLAS specific software will be shared among the various German ATLAS institutions in an organized way. While residency at the site of the RDCCG will be required at the beginning it is hoped that the work can mostly be done remotely in the long run. Cooperation in the development of Grid middleware tools is also foreseen with the possibility to dedicate up to two FTE positions to these activities. While most groups are presently concentrating their efforts on construction, testing and assembly, there is a growing contribution to software, simulation and analysis preparation, corresponding to about 8 FTEs. ATLAS has started to organize development and long term support for software by agreements, which are precursors of Computing MoUs. Some detector subsystems have already called for software expressions of interest based on a detailed work breakdown structure. In the case of the Liquid Argon calorimeter, e.g., the University of Mainz and the MPI Munich will contribute with 3.5 FTE to various reconstruction and simulation efforts.

The organizational structure of ATLAS computing builds on a matrix of sub-detector and general activities, with an emphasis on close liaison between 'physics' and computing. The supervisory structure contains a National Computing Board, which is at present chaired by the German representative. It acts as a forum for regional and institutional interests. The German ATLAS groups meet regularly to discuss outstanding

issues and to decide on a common policy. Software and data handling issues will become major discussion items in the future. Software groups, organized along detector lines, have already formed.

The overall requirements as seen today by the German ATLAS groups read as follows:

Investment/year	2001	2002	2003	2004	2005	2006	2007
CPU (kSI95)	0	2	3	10	20	45	220
Disk Space.(TB)	0	3	8	5	35	130	120
Tape (TB)	0	5	5	10	80	500	400
Active users (cum)	0	15	35	50	70	80	100

4.4 CMS

The 4 German university groups involved in CMS are concentrated at Aachen and Karlsruhe:

I. Phys. Institut B, RWTH Aachen

III. Phys. Institut A, RWTH Aachen

III. Phys. Institut B, RWTH Aachen

Universität Karlsruhe

representing 45 physicists at present.

4.4.1 Program and Status

The CMS experiment is the other of the two general purpose detectors and covers the same full physics program, accessible in proton-proton collisions. The German CMS groups have taken essential responsibilities in elements of the hardware of the tracking system and the construction of the barrel muon chambers.

4.4.2 Computing and Storage

The overall software concept in CMS foresees a development period for functional prototypes and a reality check with 1 % of the data size (Data challenge) by end of 2001, a fully functional system including a data challenge with 5 % of the data by 2003, a pre-production system with a data challenge with 20% of the data in 2004, and a production system ready in 2005. On the road towards this goal there are major core software milestones planned to calibrate the development timing. The milestones cover different levels of interaction with the general object oriented database management system with simulated data, calibrated data, reconstructed data and physics objects with increasing rates and integration levels. The detailed planning of the test of the distributed software environment together with the Grid technology development will be done together with the prototype phase until end of 2003.

The groups are predominantly involved in detector development right now, but already today about 10 persons, three of them post-docs, are actively working on the development and test of event simulation software. It is expected that this number will increase by a factor of 3 over the next three years. Computing resources used right now are local PC farms and partially the resources at CERN. Experience shows that even using all resources available it is impossible to generate the key number of 10 millions of simulated and reconstructed events needed for a single physics channel in a reasonable amount of time, i.e. one month, not mentioning that more than one physics topic shall be worked at. Since the computing costs are expected to decrease, data simulation has to be restricted practically to just the development and test of the functionality, until adequate resources become available for a reasonable price. This is expected at the start of the implementation phase. It is intended to bridge this gap in capacity by extending the local resources at the university level. Active participation in the maintenance of the CMS related software will be supported by the groups in parallel to their local software, since here the synergy effect is expected to set free human resources. Participation in the tests and development of Grid tools has been foreseen in the near future, the latest at the start of the next fiscal period.

4.4.3 Manpower and Operations

The Karlsruhe group is prepared to take responsibility in installing the CMS software and to participate in the Grid middleware development by dedicating 1-2 post-doc and 1-2 PhD student positions to this field. The Aachen groups are presently using their home computing equipment for their simulations in addition to CERN resources. They hope to enlarge the local capacity to fulfil usual simulation needs for the coming years, until the prototype phase of the RDCCG has ended successfully. It is foreseen to dedicate 1 FTE position to software development with the start of the next financing period.

Investment/year	2001	2002	2003	2004	2005	2006	2007
CPU (kSI95)	0	4	2	8	22	30	78
Disk Space (TB)	0	3	8	5	35	70	120
Tape (TB)	0	10	20	15	135	260	420
Act. users (cum)	0	5	10	15	20	30	45

The computing needs specified above take the cost effectiveness into account as mentioned before. They restrict to the basic functionality requirements for the prototype phase, turning into real needs at the implementation phase, starting in 2005. The final capacity listed is lower than the average Tier1 capacity estimated by the CMS collaboration. The German groups expect to get their share in simulated data reduced according to the number of scientists participating. The CPU power in 2002 is reflecting the expected additional simulation calculations needed to produce the physics TDR of the CMS experiment. The tape requirements consist of so-called active (fast response) and archive tape (slow response) needs, not distinguishing the obvious different access times. Development in technology and cost effectiveness may split the active tape need between the disk space and the archive tape needs and thus will have to be adjusted accordingly. The software to be installed and maintained will in addition to the standard CERN compatible elements, Linux RedHat, AFS, a batch submission system and a data storage/retrieve system also comprise an object oriented database package, Objectivity as of today's expectation, besides the experiment dependent environment.

4.5 LHCb

At present four German groups are members of the LHCb collaboration:

Technische Universität Dresden, Universität Heidelberg (2 Groups), and the Max Planck Institute for Nuclear Physics Heidelberg,

representing about 25 physicists in total.

4.5.1 Program and Status

LHCb is a dedicated Forward Spectrometer for the study of Heavy Flavour- and in particular B-meson decays. It exploits the large $b - \bar{b}$ production cross section at the LHC for precision measurements of CP-violation in the B-system. The goal is to perform stringent tests of the standard model through redundant determinations of the parameters of the unitary triangle, and possibly find evidence for new physics.

At the core of the experiment is a sophisticated multi-level trigger scheme, which selects B-meson decays with high efficiency. The level-0 stage which looks for a high pt lepton candidate in the calorimeter or the muon system reduces the reaction rate from the bunch crossing frequency of 40 MHz to an acceptance-rate of 1 MHz. Then three levels of software triggers take over to bring down the event rate to a final logging rate of 200 Hz. Already at level-1 a secondary vertex search is performed using the data of the silicon vertex tracker in order to select B-meson decays.

4.5.2 Computing and Status

In the coming years the main computing effort will go into the optimisation of the various trigger levels. For the final commissioning of the analysis chain a massive Monte Carlo production is foreseen in 2005.

The computing model and task distribution follows the multi-tier structure developed by the MONARC study. The current planning assumes the following distribution of resources, where the first three levels contribute with roughly equal weight to the total computing power:

Tier category	0	1	2	3
No of centres	1	5	15	50

The Tier1 centres are viewed as collaboration wide resources for data analysis with a Grid access policy. After the start-up of the LHC the Tier Centres provide the support for physics analysis of the outside groups. In addition the Tier1 centres will carry the main load of Monte Carlo production, Tier2 centres are special facilities for restricted production work.

4.5.3 Manpower and Operations

Three of the German LHCb groups contribute to the tracking system of the experiment. The Kirchhoff institute is involved in the level 1 trigger and is also active in a Grid working group on fabric management. All activities would greatly benefit from a Regional Centre in Germany. In addition, the centre would focus the task of software maintenance for the LHCb experiment to one central institution, allowing to cover the requirements of all groups with an expected investment of 2 FTE positions. A participation in the LHCb data challenges at latest in 2003 would assure a timely integration of the German groups into the general LHCb computing environment.

Investment/year	2001	2002	2003	2004	2005	2006	2007
CPU (kSI95)	0	0.6	2.3	2.7	15	47	62
Disk Space (TB)	0	2	2	4	9	38	50
Tape (TB)	0	0	0	0	42	123	208
Active users (cum)	0	2	4	8	14	28	35

4.6 LHC summary

The German collaborators of the four LHC experiments contribute to the rich research program at the new hadron collider at CERN by providing valuable investments in the construction of essential detector elements. As a natural consequence they also will participate accordingly in sharing the benefits of these investments by asking for a compatible share of analysis possibilities of the rich scientific data in comparison to other European collaborators. Therefore they apply for a future oriented investment on computing power in a new, unprecedented structure, not only to serve their immediate needs, but also to serve as a test and development centre for the more heterogeneous needs of other different natural science research projects. A first step into this direction is the integrated service foreseen to satisfy the needs of experiments in different branches in physics.

In order to maximize the value of the total investments the LHC groups restrict their present needs to the minimum amount of resources needed to install and maintain the experiment dependent software and to participate in the tests planned to test and enable their software functionality.

The quality of data communication between the Tier0 and the Tier1 centres as well as among each other is an essential ingredient of the functionality of the whole concept. Therefore, the effective network bandwidth among the centres is a key issue of the project. The final bandwidth per experiment at production time is estimated to be 1.5 Gb/s between a Tier0 and each of the Tier1 centres and 622 Mb/s towards each Tier2 centre. In the mean time networking quality at the affordable upper limit, as of today's expectation, has to be implemented and its financing ensured. It should reach its final capacity at the latest at the end of 2005 to provide full functionality for the production systems.

The following table presents the sum of the requirements of the four experiments, not taking into account possible synergy effects during the prototype phase.

Investment/year	2001	2002	2003	2004	2005	2006	2007
CPU (kSI95)	0.3	7.6	10.3	30.7	77	172	580
Disk Space (TB)	1	10	23	24	109	298	490
Tape (TB)	1	20	35	45	307	983	1528
Active users (cum)	2	27	59	93	144	198	320

The configuration in 2007 comprises the two years before, defining the start of the production phase.

Initial Production Configuration in 2007	CPU (kSI95)	Disk (TB)	Tape (TB)	Act. users
	829	897	2818	320

In the years after production has started, the CPU, Disk and storage capacity is expected to increase at a fixed amount per year, including the replacement of older hardware components, following a rolling upgrade scheme to cope with the storage and analysis of the new and growing data samples. The rates expected are 33% for the CPU power, 50% for the disk storage space and 100% for the tape archiving capacity. A period of 15 years of running is to be envisaged, at least until the end of the LHC experiments.

Besides the operation as a computing and data centre all German LHC groups plan to use the centre as a place of communication and training for their experimental software environment as well as for modern tools and software techniques. E.g. training in object oriented programming and modern tools is considered of high importance, since many personal investments have gone into the move of value proven Fortran software to an object oriented approach in C++. Some of the respective courses are offered at CERN or are available as interactive media, but all groups would appreciate that the computing centre offers own courses on general tools (e.g. data base packages, Grid software etc.), and supports the experiments to organize training for experiment dependant software on the site.

5 Requirements of the Non-LHC Experiments

5.1 BABAR

Ruhr-Universität Bochum, TU Dresden, Universität Rostock

5.1.1 Status and long term plans

The luminosity delivered by PEP-II and recorded by BABAR at SLAC has been increasing throughout the experiment's operation. Current peak luminosity is over $3.10^{33} \text{ cm}^{-2} \text{ s}^{-1}$, and daily-integrated luminosity of over 100 pb^{-1} is routine. In addition, there are plans to upgrade PEP-II luminosity as well as the BABAR detector⁴. The corresponding increase in data requires distribution of data to computing centres in the participating countries, following the guidelines of the CERN MONARC study. BABAR is currently taking data, and will continue to take data for approximately 10 months every year until 2010. The projected integrated luminosity until 2005 is foreseen to reach 620 fb^{-1} .

5.1.2 Network, Storage and CPU Needs

The current hierarchical event format has proven to be the right approach to allow for an efficient iterative analysis. Since disk space is such an important issue, the size of the event data, content and packing, is a crucial parameter that has been carefully optimised. All data levels are available in BABAR's TIER-A centres (SLAC, IN2P3) in various processing levels, serving the collaborating institutes. It is envisaged to start up a TIER-B site in the new computer centre to allow the currently 25 German BABAR collaborators to do filtering and physics analyses on the micro/mini level and to modestly contribute to BABAR's Monte-Carlo production of simulated events. It is foreseen to accept all physics data of the most recent processing level on the tag/micro level, and in addition mini data for selected physics topics. A 10 nb physics cross section, a number that corresponds to roughly 15 million events per fb^{-1} , leads to the following **yearly** incremental requirements in disk space, network bandwidth for imports and CPU power:

Investment/year	2001	2002	2003	2004	2005
CPU (kSI95)	1.0	1.5	2	2.5	3
Disk Space (TB)	6	10	14	16	27
Tape (TB)	6	10	14	16	27
Network Bandwidth (Mb/s)	1	2	3	4	5

It should be stressed that the previous table is for beam data and the same amount of Monte-Carlo events and takes into account the increased use of the so-called mini format. As BABAR produces data in steady-state operation using a rolling calibration technique, it is reasonable to take averages of network bandwidth for processing of new data and re-processing of old data over the year. At the end the final analyses are conducted in the participating universities based on the resulting NTuple files that are typically a few GB in size. During the course of the actual year all cumulated data of the previous years are replaced with data at more recent processing levels. Elder processing levels are still available in the TIER-A sites and can be transparently accessed through the Grid: jobs requesting old data could migrate there and a TIER-B would need mass storage for passive backup of local data only; thus the tape archiving requirements are of the same order as the numbers given for disk space. Assuming a modest participation in MC-production in addition, installation of 50 state-of-the-art Pentium-like CPUs per year seems to be adequate to cover the needs of the German BABAR community (the SI95 requirements scale with Moore's law).

⁴ See: <http://www.slac.stanford.edu/BFEROOT/www/Detector/Upgrades/FINALREPORT.pdf>

5.1.3 Manpower and Operations

The operation of a German data centre for BABAR requires two experts FTE of the experiment to take care of the specific software release environment, tools and local organization of the data. The universities will provide this personal.

BABAR supports the Linux/Intel, the SPARC/SOLARIS and the ALPHA/OSF platforms. The German data centre will focus on Linux/Intel only. System extensions to be installed by the computer centre are: AFS client, CVS, ROOT, Oracle and Objectivity DB. The experiment specific software will be installed and maintained by experts from the experiment.

5.2 CDF and D0 Experiments at the Tevatron Collider

5.2.1 Status and long term plans

The Tevatron Collider at the Fermi National Accelerator Laboratory (FNAL) has completed a major upgrade and started a new data-taking period on March 1, 2001. Also the two experiments CDF and D0 at the Tevatron have been considerably improved to handle the significantly increased event rates and to allow for a higher precision in data reconstruction. These upgrades will extend the sensitivity to physics beyond the Standard Model and particularly allow one to find a Higgs boson if it is lighter than 170 GeV. The Tevatron is expected to run at least until 2007 and to collect some 2 fb^{-1} until 2004 (Run 2a) and, after a further upgrade, some 15 fb^{-1} until 2007. The huge amount of data to be expected over the next years implies the need for an efficient computing environment. German groups are involved in both experiments. They are planning to collaborate closely to minimise potential overhead from implementing the software specific to the FNAL software.

5.2.2 CDF Requirements

Universität Karlsruhe

Karlsruhe's main physics interest pursued with DELPHI and CDF in the past deals with precision electro weak tests, heavy flavour physics and search for physics beyond the Standard Model. Because of the large amount of data substantial discovery potential at the Tevatron is expected in the coming years even though the long-term goal will be physics at the LHC. The CDF community has at present 17 participants (14 FTE) whose interest for Run II focuses on the following topics: search for the Higgs Boson, Beauty Physics, Top Physics, and possibly SUSY searches. Main ingredients of such events are jets with Beauty content, W and Z bosons in the leptonic channel. Whilst one raw event has the size of 250 kB, it is regarded sufficient to copy and store the subset of processed events (first 60 kB, later 20 kB) with these categories. Based on these numbers the following **yearly** resources are needed:

Investment/year	2001	2002	2003	2004	2005	2006	2007
CPU (kSI95)	-	2	1.5	1.5	5	2	2
Disk Space (TB)	-	8	9	8	28	28	28
Tape (TB)	-	17	13	13	27	30	30
Network Bandwidth (Mb/s)	-	1	1	1	2	2	2

5.2.3 CDF Manpower and Operations

The operation of a German data centre for CDF requires two experts FTE of the experiment to take care of the specific software environment, tools and local organization of the data. Karlsruhe University will provide these personnel.

CDF supports the RedHat Linux platform on Intel processors. System extensions to be installed by the computer centre are: AFS client, CVS, ROOT, KAI compiler and Oracle DB. The software specific to the experiment will be installed and maintained by the respective experts.

5.2.4 D0 Requirements

Universität Aachen, Universität Bonn, Universität Mainz, LMU München, Universität Wuppertal

The German D0 community of currently 25 physicists follows common interests that allow sharing resources effectively. Thus it is necessary to transfer and handle only 10% of the experiment's grand total reconstructed and 1% of the raw data volume within Germany. There are no needs for reconstruction of large data sets, the plan is to transfer reconstructed data and only run larger filter and analysis jobs on site. It is foreseen to generate about 10^7 MC events per year and to have about the same amount of data and MC events available for data analysis. Those with the most up-to-date reconstruction should have a fast access and should be stored on disk. The data and MC events with older reconstruction should be mostly stored on tape. The following table summarises the **yearly** incremental requirements.

Investment/year	2001	2002	2003	2004	2005	2006	2007
CPU (kSI95)	-	1	1	2	2	2	2
Disk Space (TB)	-	9	18	30	50	50	50
Tape (TB)	-	54	30	45	70	150	150
Network Bandwidth (Mb/s)	-	1	1	1	2	2	2

The average network load is about 30 GB per day, ramping up to 100GB-200GB per day in the case of re-processed data transfers. The final analyses are done on the basis of NTuples that are exported to the university institutes (a few GB per day).

5.2.5 D0 Manpower and Operations

D0 requires two experts FTE of the experiment to take care of the specific software release environment, tools and local organization of the data. The universities will provide this personnel.

D0 supports the RedHat Linux platform on Intel processors. System extensions to be installed by the computer centre are: AFS client, CVS, ROOT, KAI compiler and Oracle DB. The software specific to the experiment will be installed and maintained by the respective experts.

5.3 COMPASS

Universität Bielefeld, Universität Bonn, Ruhr-Universität Bochum, Universität Erlangen, Universität Freiburg, Universität Heidelberg, Universität Mainz, LMU & TU München

5.3.1 Status

COMPASS is currently under construction at CERN and plans to start data taking in 2001. The first run will be short with simple geometry and about 50 TB of data. 2002 will see a longer run with about 150 to 200 TB. In 2003 and 2004 nominal running will take place and COMPASS expects about 300 TB of raw data per year. The COMPASS collaboration is considering an extension beyond 2005 with an upgraded apparatus.

5.3.2 Network, Storage and CPU Needs

For data analysis a DST set of 60-80 TB is extracted from the whole data set. Several subsets with tight selection criteria can amount to a total of 30 TB, a minimum sample useful to be stored in a commonly accessible tier data centre for the currently 20 German COMPASS users. The origin of the data will be the COMPASS Computing Farm (CCF) at CERN, which will consist of 200 CPUs plus data servers. It is there where initial data reconstruction takes place and the DST's are created. Common reprocessing/filtering jobs as well as Monte Carlo production, however, could be shared by CPUs in the German data centre. The following are the experiment's **yearly** incremental requirements

Investment/year	2001	2002	2003	2004	2005	2006	2007
CPU (kSI95)	-	0.2	1	1	1	1	1
Disk Space (TB)	-	1	4	12	20	20	20
Tape (TB)	-	3	8	15	30	30	30
Network Bandwidth (Mb/s)	-	0.15	1	2	4	4	4

The table reflects average network bandwidth for data imports from CERN. A higher network bandwidth is needed for exports due to the idea that several of the universities with strong groups should be able to access sub-samples (< 1TB) to be analysed in their local farms of up to 100 CPUs. The layout of the institute farms as Tier2 or 3 centres might be considered. In particular where strong other groups (different from the COMPASS groups) participate in LHC experiments the centres may profit from this layout. It should be noted that e.g. INFN Trieste and Torino as well as CEA Saclay plan to join similar efforts in Italy and France. In the pre-LHC phase the Grid prototype shared by COMPASS and other running experiments may be seen as real life test for the larger installation for the LHC GRID infrastructure. The COMPASS institutes are willing to participate in GRID test-bed activities with their local farms.

5.3.3 Manpower and Operations

The operation of a German data centre for COMPASS requires one expert FTE of the experiment to take care of the specific software release environment, tools and local organization of the data. The universities will provide this personnel.

COMPASS supports the Linux platform. System extensions to be installed by the computer centre are: AFS client, CVS, ROOT and Objectivity DB. The experiment specific software will be installed and maintained by the experts.

6 Overview and advisory structure

The RDCCG has to operate in the context of the international LHC GRID Computing Project as well as in the computing scheme of the non-LHC experiments. The coordination and management of such an ambitious technical undertaking is a challenge in itself.

The German Particle and Nuclear Physics Communities propose two boards for a close interaction with the RDCCG, an Overview Board (OB), and an Advisory Board (AB). In addition RDCCG representatives have to cooperate and coordinate with the CERN based international LHC GRID Computing Project and in particular with the Software and Computing Committee (SC2) proposed by the LHC computing review. The centre should appoint a Project Leader (PL) in close cooperation with the OB.

The OB decides on general policy, resource and funding issues affecting the RDCCG and arbitrates in case of conflicts on technical issues. The membership should include

- computing director of the centre,
- PL and deputy,
- national spokespersons of the experiments or their representatives,
- members from BMBF, KET and KHK,
- chair of the AB and deputy.

The OB should meet at least once per year, or more frequently if required, to review and agree on major changes in human and material resources and the project plan.

The AB consists of:

- software and computing experts of the experiments,
- representatives from one or two other Regional Centres and/or the LHC Computing GRID project,
- observer from DESY,
- observers from KET and KHK,
- ex-officio: PL and deputy.

It may decide to invite other experts to its meetings. The AB helps to elaborate a detailed work programme, using inputs from the experiments. The primary mechanism to perform the advisory role are updates of the hardware and software requirements to the centre and periodic technical reviews of the achievements to fulfil these requirements. The PL informs the AB of all relevant technical issues and seeks agreement with the AB prior to installation. In case of disagreement the OB has to be involved. The AB should meet at least four times a year.

The centre should appoint a PL based on recommendation from the OB. The PL is the principal authority for the day-to-day management of the RDCCG, is in charge of developing a project plan in close relationship with the AB and the SC2, and manages the execution of the approved project plan. He/she reports to the director of computing of the centre and to the OB and is responsible for a close liaison to the physics community.

Each experiment nominates a dedicated technical contact person to the RDCCG and the RDCCG provides a corresponding contact expert.

7 Requirements for the Operation of the RDCCG

The experiments at the future hadron collider LHC at the European research centre CERN in Geneva enter new grounds of computing with their requirements of computational performance and data processing needs. The MONARC study proposes a network of computing centres to master a task of this order of magnitude in international cooperation. Apart from the general requirements on the hardware and software level additional requirements for the operation of a regional centre have to be met that allow for effective cooperation of the computing centres involved. Those extra requirements must combine two aspects: On the one hand the hardware set-up should be homogeneous in order to execute computer applications of the individual research projects transparently all over the world, on the other hand different fields of research should be able to share resources in common responsibility. The successful operation in such a heterogeneous environment of requests may lay grounds for integration of general research activities with their computing demands: The collaboration of data and competence centres in this form could be the first step towards the availability of world-wide computer resources for e-sciences.

7.1 Memorandum of Understanding

The cooperating regional data centres are located in various countries and are subject to different investment and operating conditions. The principle of collaboration is ensured by means of Computing MoUs between the experiments, CERN, and the data centres. The institute that hosts and operates a data centre has to sign and to adhere to the corresponding memorandum of understanding.

7.2 Accountability

The RDCCG should provide and operate a transparent scheme of quantitative assessment of all services per experiment.

The RDCCG has to provide an annual operational plan which has to be approved by the AB.

7.3 Hardware Compatibility

7.3.1 CPU Installation

Hardware compatibility of the CPU installation with respect to the other centres is an indispensable requirement regarding the internal number representation and arithmetic. The hardware compatibility issue should form the basis of all procurements in order to avoid divergences. Agreement in hardware architecture regarding memory and cache sizes is highly desirable. A dedicated reference system is indispensable in identical configuration at each centre, for verification and testing under identical conditions.

7.3.2 Data Store

Hardware compatibility of the storage media is extremely important in the area of data archives. In particular divergences are to be avoided with new procurements. Deviations of this requirement have to be discussed with the advisory board.

7.4 Software Compatibility

7.4.1 Operating Systems and Upgrades

The operating systems must be compatible between the collaborating centres. The OS software maintenance is the duty of the data centre. Upgrades of the operating systems are to be coordinated between the centres. In case of upgrades a backwards-compatible environment is to remain for a transition period. The advisory board has to decide in case of discrepancies.

7.4.2 Compiler, Libraries and other Software Packages of general Interest

A catalogue of supported software determines the compilers, libraries and further software packages of general interest and thus the software licences for storage management systems, standard libraries, graphics packages etc. Software packages of general interest are:

- GLOBUS, AFS, CVS, ORACLE, Objectivity, CERNLIB, GEANT4, ROOT, FLUKA

Software maintenance proceeds according to the principles as defined for the operation systems maintenance. The advisory board updates the supported software catalogue on a regular basis. Installation and support of user specific software is the task of the corresponding user communities.

7.4.3 GRID Software Development and Job Submission Systems

CPU-intensive applications will normally run in batch, though interactive access for testing and debugging must be possible. The centre should operate a suitable wide area network batch system like GLOBUS to manage remote job entry. Hardware resources and manpower for active development, integration and testing of GRID applications have to be provided with the idea to establish a GRID competence centre that closely collaborates with other institutes working in the field. The GRID middleware as well as the job submission systems are regarded as part of the operating system. As such the same compatibility issues apply. In case of alternative solutions or contradiction due to heterogeneous requests, the agreement of the advisory board must be achieved.

7.4.4 User Administration

The user administration should follow the organization of the corresponding Tier0 and Tier-A centres, hence the integration into a common GRID based authentication authority. In this context the centre should support the timely foundation of a national authentication authority for e-science and help with the registry and the administration of the GRID authentication certificates. This registry should foresee trust with corresponding registries in the participating institutes and make use of existing directory services. The centre has to provide home directories for all users with backups on a regular schedule.

7.4.5 Reference System

At least one reference system for compatibility tests is to be maintained that has to be generally accessible.

7.4.6 Networking

The quality of the network performance between the cooperating centres must comply with international standards. Funding of the network components and the network operation is to be guaranteed, an appropriate development plan is to be worked out with the cooperating centres.

7.4.7 Advisory Services and Training

The centre should provide general consultation for the software catalogue specified above. Training courses in the middleware, software and hardware environment as well as into individual areas of the user software packages have to be offered. The centre should prepare an appropriate lecturing infrastructure. Documentation of the available hardware and software as well as the appropriate manuals have to be provided, special local features are to be well documented. All sources of information should be made accessible through the WWW. A hypernews system has to be maintained in order to stimulate discussion of relevant technical topics and communication.

7.4.8 Consultation and Problem Tracking

The centre has to operate a user consultation office during the usual on-site working hours. In order to support out-of-house requests from external users a helpdesk system has to be installed to manage problem reports and allow for problem tracking. Detailed status information about the hardware and software should be distributed via the WWW to allow users to monitor the state and the performance of the installation.

7.4.9 Operating Conditions, Maintenance and Exchange Acquisition

The operation of the centre is to be guaranteed 24 hours a day and 7 days a week. Operators and experts have to be available during on-site working hours. In hours of non-supervised operation an on-call expert service has to be arranged. All systems should be set up for fail-save operation. The technical details of implementation have to be discussed with the advisory board. Hardware maintenance is to be ensured, whereby the form, e.g. servicing contracts or extended guarantee periods in the context of a rolling upgrade concept are to be coordinated with the advisory board. Modifications in operating conditions require the agreement of the advisory board if they have an impact on access and throughput.

8 Steps towards the RDCCG

In the following sections the proposed steps towards a German RDCCG will be outlined. This centre will have to collaborate closely with the other regional centres both for the LHC and the other experiments. The non-LHC experiments will need a stable production environment as soon as possible. Concerning the needs of LHC computing, the whole project can be divided into 3 phases:

- Phase 1: Development and Prototype Construction, 2001 - 2004
- Phase 2: Installation of the LHC Computing Production Facility, 2005 - 2007
- Phase 3: Maintenance and Operation, 2007 – end of LHC data analysis

Due to the current knowledge only the first phase can be envisioned in some detail. Decisions about and planning of the later phases will be part of the previous phase.

8.1 Phase 1: Development and Prototype Construction, 2001-2004

Beside detector simulations, the main goals of the LHC groups till 2004 are the development, deployment and test during Data Challenges of their production software, as well as the demonstration of the feasibility of their world computing model. The underlying hardware of this model consists of a multi-Tier hierarchical model of large computing fabrics. GRID technology will be used to provide a single image of this world-wide distributed computing system to every physicist, and to ensure a combination of efficient resource utilisation and fast turnaround time.

The development of the RDCCG should be organized around annual stable prototype systems with increasing size and functionality. By 2004 a production quality system of about 50% of the complexity of the 2007 version of one experiment is needed. This system is the basis for a complete appraisal of the developments before defining and establishing the final LHC regional data and computing infrastructure.

The years until 2004 can in addition serve to have regular meetings and workshops with the experts of the other European regional centres and CERN. Emerging tasks and problems in the field of the management of large fabrics, their interconnection via GRID middleware, and the support of a growing user community can be discussed with the RDCCG as a serious partner amongst the already well established competence centres at CERN, Lyon, INFN and Rutherford Lab.

All LHC experiments plan annual or biannual data challenges – large scale tests - of increasing size and complexity until LHC start-up to test their software systems and the computing facilities.

The non-LHC experiments will provide physics data beginning with 2001 and demand installation of hardware and software prior to the large scale LHC activities in order to allow for timely production of valuable physics results in the years of 2001-2004. They can thus be seen as a natural test bed for operation of a competitive particle physics regional computing centre, to ramp up operation and gain experience in both, commissioning and maintenance of large computing farms and the corresponding infrastructure, especially the development of GRID competence in the years of LHC preparation. BABAR is an experiment that routinely runs in factory mode and delivers a tremendous amount of data. It is actively investigating GRID technology to relocate and share databases between SLAC, Lyon, Rutherford and Rome. Germany as a further partner would seamlessly fit into this scenario. The FermiLab experiments CDF and D0 will follow up soon with large databases in 2002; especially D0 with its strong European user community promises a good chance to begin a common GRID activity with other European data centres in France and UK. COMPASS finally will produce very large databases at CERN and could be seen as a natural test bed for data relocation between CERN and the RDCCG; its computing model foresees a commitment for large computer farms in the German universities and thus fits into the GRID context very well.

As non-LHC investment is intended to be used in the scope of the LHC data challenges as well, a significant synergetic effect is to be expected, leading to a significant saving in overall cost for computing equipment in the German Particle and Nuclear Physics communities. The sharing of resources furthermore

leads to a decrease in operational costs and the timely development of GRID competence, a fact that will lay a very solid foundation for the upcoming LHC activities.

The following tables summarize the resources needed for the first phase:

LHC Investment/year	2001	2002	2003	2004
CPU (kSI95)	0.3	7.6	10.3	30.7
Disk Space (TB)	1	10	23	24
Tape (TB)	1	20	35	45

non-LHC Investment/year	2001	2002	2003	2004
CPU (kSI95)	1	4.7	5.5	7
Disk Space (TB)	6	28	45	66
Tape (TB)	6	84	65	89

Since the LHC CPU requests are peak values, which are needed only during 2 to 3 month per year for the data challenges or detector simulations, a reasonable equipment for both LHC and non-LHC experiments is:

Investment/year	2001	2002	2003	2004
CPU (kSI95)	1	8	10	27
Disk Space (TB)	7	38	68	90
Tape (TB)	7	104	100	134

The annual prototype systems and the data challenges form a natural set of high level milestones for phase 1. Prototype 0 should provide basic functionality for BABAR and first simulations for the ALICE physics performance report. The prototypes up to 2004 should follow the milestones of the LHC computing project. In addition it has to provide the functionality required by the non-LHC experiments.

Milestones	Date	Description	Who
1.1	1/7/2001	Requirements Document written	Experiments
1.2	15/8/2001	Final answer from the Centre Received	Centre
1.3	9/2001	Agreement for the Prototype Phase	Centre + Experiments
1.4	10/2001	Management Structure defined	Centre + Experiments
1.5	11/2001	Version 0 of the Prototype	Centre
1.6	4/2002	Version 1 of the Prototype	Centre
1.7	12/2002	Data Challenges completed	Experiments
1.8	4/2003	Version 2 of the Prototype	Centre
1.9	12/2003	Data Challenges completed	Experiments

1.10	4/2004	Version 3 of the Prototype	Centre
1.11	6/2004	Decision about Phase 2	OB
1.12	12/2004	Data Challenges completed	Experiments
1.13	12/2004	Production Prototype	Centre

The detailed definitions of the content of each prototype and the deliverables needs to be defined in close collaboration with the experiments, including the needs of their data challenges. Also the intermediate and final results of the various GRID projects especially the DataGRID project should be taken into account. But at least a major external project review should be foreseen at the transition to phase 2 of the project.

8.2 Phase 2: Installation of the LHC Computing Production Facility, 2005 – 2007

The main goal of phase 2 is the installation of the initial full production facility. The resources needed and the milestones are not yet known to a sufficient degree of precision. They will be defined according to the experience from the prototyping phase, and the experiments computing Technical Design Reports.

LHC Investment/year	2005	2006	2007
CPU (kSI95)	77	172	580
Disk Space (TB)	109	298	490
Tape (TB)	307	983	1528

The numbers given for the non-LHC experiments in the years 2006 to 2007 in the table below do not include the BABAR requirements, since no model for the estimated amount of data can be given.

non LHC Investment/year	2005	2006	2007
CPU (kSI95)	11	5	5
Disk Space (TB)	125	98	98
Tape (TB)	154	210	210

Estimates of the required WAN bandwidth arrive at 1.5 to 3 Gb/s per experiment by 2007.

8.3 Phase 3: Maintenance and Operation, 2007 – end of LHC data analysis

The LHC Computing Review estimates the maintenance and operation cost on the assumptions, that after the construction period, the capacity of the system is increased by a constant amount of the 2007 value every year:

- The amount of CPU is increased every year by 33%. The obsolete equipment is replaced after the 3-year maintenance period.
- The disk space is increased every year by 50%, to make the analysis of bigger data samples possible.
- The tape storage capacity is increased every year by 100%, to store the new data.

LHC experiments	CPU (kSI95)	Disk (TB)	Tape (TB)
Configuration in 2007	829	897	2818
Yearly upgrade(without repl.)	276	450	2818