Convocatoria de ayudas de Proyectos de Investigación (2005)

MEMORIA CIENTÍFICO-TÉCNICA DEL PROYECTO

1 **RESUMEN DE LA PROPUESTA** (Debe rellenarse también en inglés)

INVESTIGADOR PRINCIPAL: José Francisco SALT CAIROLS

TITULO DEL PROYECTO: **" Development of a Distributed TIER-2 Infrastructure for the ATLAS experiment of LHC"**

RESUMEN (debe ser breve y preciso, exponiendo sólo los aspectos más relevantes y los objetivos propuestos):

Este proyecto propone la creación de una infraestructura de cálculo TIER-2 para el análisis de datos experimentales y generación de sucesos Monte Carlo para el experimento ATLAS del acelerador LHC. El tamaño de esta infraestructura es el 5% de lo previsto en ATLAS, en proporción a la participación global de los grupos españoles participantes en ATLAS.
Este TIER-2 se construye de manera distribuida entre los 3 centros españoles participantes en el experimento ATLAS: el IFIC en Valencia, el IFAE en Barcelona y la UAM en Madrid. Con esta distribución se garantiza un servicio permanente y una óptima utilización de los recursos humanos de cada centro.
Esta infraestructura está integrada en el ´LHC Computing GRID´ de acuerdo a las directrices del ´Modelo de Computación´ acordado por la colaboración ATLAS. Por lo tanto, este proyecto tiene una fuerte interacción con los proyectos de desarrollo de Sistemas Distribuidos GRID. Concretamente, dentro del proyecto LCG-ES, del cual este proyecto es su continuación, estamos desarrollando la implantación de un sistema de Análisis Distribuido, Servicios de Gestión automática de Clusters y Servicios de Almacenamiento y Gestión de Datos.
Al mismo tiempo, este proyecto tiene una fuerte interacción con el sistema de proceso de datos de ATLAS. Son tareas a realizar por un TIER-2, la generación de sucesos de Monte Carlo y la Reconstrucción de los mismos y el filtrado y Análisis de Datos.
Finalmente es también objetivo del proyecto garantizar que los usuarios de los centros españoles participantes en ATLAS puedan realizar los análisis de los datos en cualquier canal de física de su interés. Por ello se incluye en este TIER-2 el Soporte al Usuario, la infraestructura del análisis final de Ntuplas (TIER-3) y la posibilidad de generación de Monte Carlo y selección de sucesos local por parte del usuario

PROJECT TITLE: 'Development of a distributed TIER-2 Infrastructure for the ATLAS Experiment of the LHC'

SUMMARY:

This Project proposes the creation of a TIER-2 computing infrastructure for experimental data analysis and Monte Carlo event generation for the ATLAS experiment of the LHC collider. The size of this infrastructure is 5% of the expected ATLAS TIER-2, in accordance to the overall participation of the spanish groups.

The TIER-2 will be carried out in a distributed manner shared by the 3 spanish institutes, members of the ATLAS Collaboration, namely: IFIC (Valencia), IFAE(Barcelona) and UAM (Madrid). This distribution guarantees a permanent service and an optimal use of the institute human resources.

The infrastructure is integrated in the LHC Computing GRID following the 'Computing Model' of the ATLAS Collaboration. Hence, this project has a strong link with the running GRID projects that develop tools of general use. In particular, this project is a continuation of the present LCG-ES project, where we are developing the implementation of a Distributed Analysis System, services for Automatic Cluster Management and services for Data Management.

In addition, this project has a strong link with the Data Processing system of ATLAS. Tasks to be done in a TIER-2 center are: Generation and Reconstruction of MonteCarlo events, Event Filter and Data Analysis.
Finally, it is also a goal of the project to guarantee that the users of the Spanish institutes of the ATLAS collaboration are able to perform data analysis in any physics channel they consider of interest. Therefore, the following tasks are also responsibility of the TIER-2: User Support, Infrastructure for Ntuple Analysis (TIER-3) and the possibility for any user to generate Monte Carlo events and make their own sets of data events.

2. INTRODUCCIÓN

(máximo **cinco** páginas)

- Deben tratarse aquí: la finalidad del proyecto; los antecedentes y estado actual de los conocimientos científico-técnicos, incluyendo la bibliografía más relevante; los grupos nacionales o internacionales que trabajan en la misma materia específica del proyecto, o en materias afines.-

Purpose of the Project:

A) The Data Challenge in the LHC

ATLAS is one of the experiments that will begin to work at the end of 2007 within the program of LHC (Large Hadron Collider) [1]. The proton beams that will collide in ATLAS will be provided by the LHC accelerator and they will have a 7 TeV energy , with which they will provide a total of 14 TeV in the Center of Masses. The initial Luminosity will be of 0.5 x 10^{33} cm^{-2} s^{-1} being able to arrive at 10^{34} cm^{-2} s^{-1} as of 2010. The collision frequency will reach 10^{9} Hertz and trigger rate will be of 200 Hertz. This implies a production of real data - coming from the taking of data of the experiment of the order of the PetaByte (1 Million Gigabytes) per year. The scientific potential is very considerable and includes enough aspects, one of most important is the search of the boson of Higgs and the discovery of New Physics [2].

Such volume of data and the necessity of their processing has been giving rise to have to focus the attention in the last years in the establishment of a Computing Model for the LHC experiments and, in concrete, for ATLAS. The Model extends from the storage of primary events to the possibility of making analysis in a remote university department. During 2005 they are going away to draft two documents of great importance: the so-called Computing TDR (Technical Design Report) and the Computing MOU that will allow to establish the responsibilities of the centers of different countries in the correct operation of computation infrastructures. It is necessary that the access to the raw data improves for calibration, development of algorithms, etc. Of great importance it is the sizing and organization of the resources in the so-called solution for the steady-state period during the data taking.

The collected data are going to undergo transformations whose aim is the downsizing and the extraction of relevant information being formed a chain of data types. There is an almost definitive list of them which it is important to give for clarification of the text that follows: RAW, they are the raw data, which are the data as they are collected from the detector; SIM, they are the simulated data; DRD, they are the reconstruction data, ESD are the data summarize of the events, AOD, that they are the Analyses Object Data, DPD, they are the data of Physics of Ntuple type, TAG, that they are the summaries of events for its selection.

The coming of the technologies GRID and their development and deployment have given rise to new ideas of how the ATLAS system is organized as an ` Installation of Distributed Computing'. It implies a high degree of commitment of the participant ATLAS Institutes to contribute with computation resources. In the ATLAS experiment 3 spanish groups of High Energies Physics participate: the group of the IFAE (Institut de Física d'Altes Energies) of Barcelona, the group of the UAM (Universidad Autónoma of Madrid) and the group of the IFIC (Instituto de Física Corpuscular) of Valencia. These groups began their collaborations with ATLAS more than ten years ago and took commitments from development, prototype and construction of detectors. Also they have become to be involved in other tasks, for instance, studies oriented to physics analysis. These activities can be seen in section 5 (Historial de los grupos).

B) Model of Computation and infrastructures:

The purpose of the project is to develop and to start up an infrastructure of Distributed TIER-2 for ATLAS experiment in Spain and for it the three groups previously mentioned (UAM, IFAE and IFIC) have coordinated ourselves to ask for a project to obtain this aim. In the three groups activities and infrastructures exist that make the execution feasible and there is a great interest on the part of the physicists who work in ATLAS in which the Distributed TIER-2 of ATLAS will be a reality.

Denomination TIER-2 comes from computation models that already settled down some years ago and that involved an hierarchical organization of geographically scattered centers; concretely it was begun with the TIER-0 of the CERN, it continued with centers TIER-1 in several countries, the denominated TIER-2 that agglutinate many day to day activities of the LHC physicists; and finally they are the TIER-3 associated to the activity of the physicists of institutes or departments with limited computing resources and services.

As it is specified in annex 4.3 of the TIER-2 Services of reference [3] , the responsibilities of a TIER-2 are the following ones (see [4],[5] and [6]):

1) Services of disk storage providing permanent storage and/or temporary of data for files and data bases;
2) To provide analysis capacity for the Physics Working Groups. To conduct the operation of an Installation of an ` End-User' Analysis of Data that gives service to an approximated total of 20 physics analyses subjects.
3) To provide other services, as the simulation, according to the agreement specified by the requirements of the experiment.
4) To provide Network Services for Interchange of data with TIER-1 centers.

Focusing in the data, an 'average' TIER-2 will store[4]:

- 1/4 of AOD and all TAG samples,
- some of DPD samples of the Physics groups according to the local interest,
- simulated data typically migrated to the TIER-1 unless the access under request in the center will be ensured;
- agreements on host primary for the data. Modest samples of Raw Data and ESD will be stored to make code development
- some TIER-2 will take a significant paper in calibration according to their interests and implications from the detectors.

In order to achieve this goal, the objectives are going to be distributed according to the fact that this project is requested for two years, concretely the period would be from the 1st October of 2005 to the 30th September of 2007. For it is last date it is anticipated that already the first collisions in the accelerator LHC have taken place and that begins with calibrations and to understand as the detectors work. For this moment the complete infrastructure of the TIER-2 should be deployed and it should begin to be 100%. It is necessary to consider that towards first half of the 2008 is when it is anticipated that first trustworthy Physics events will be available.

Estimation of the necessities of hardware has been performed taking into account the ATLAS Computing Model [4], [7] .In the following table it is shown the hardware capacity required for the total set of TIER-2 centers:

Resources	Todos los TiER-2 ATLAS
Tape (PB)	0.0 PB
Disk (PB)	8.5 PB
CPU(MSI2K)	20 MSI2K

In the following table the time evolution of the necessary hardware of an average TIER-2 is given [7]:

	2007	2008	1009	2010	2011	2012
Disk (TB)	107.1	336.8	566.3	887.4	1315.5	1866.4
CPU (KSI2K)	244	704	1064	1983	3013	3944

The estimations of 2007 and 2008 are the used for the request of the equipment.

On the other hand, we considered that this project must include the hardware resources necessary for the analyses of the people participating in the physics teams of the centers. It is necessary to add, therefore, around 1,5 TB and 1,5 KSI2K by user (it is what would correspond with the TIER-3) that also are considered in the request. For the cost estimations that we have expressed in the request we have used information like the references given in [8].

Considering both concepts, TIER-2 more TIER-3, we located ourselves in a proportion near 5% of the total of all the TIER-2 of ATLAS.

Antecedents and present state of the scientist-technical knowledge:

During the period between the 1st October of 2002 until the 30th of September of 2005 we are developing the project LCG-ES – which has the title ` Development of an Infrastructure of DATAGRID for the analysis of LHC data "-. Spanish institutes pertaining to the 3 experiments of the LHC (ATLAS, CMS, LHCb) are participating in this project in a coordinated way. It has been the starting to form an GRID infrastructure in Spain and to contribute to the development of the International LHC Computing GRID (LCG) . The present project is the natural continuation of LCG-ES. Moreover, we have participated in European projects of development of networks of GRID Computing like the European DataGrid (EDG), CrossGrid and the present EGEE project.

Some of the profits obtained by the groups within these projects are enumerated and, in individual, in the LCG-ES are given:

1) Computer science Infrastructure and of Computation of the centers:

A very important section is the starting point in our computer infrastructures. In some centers it there has been financing coming from other different sources and not only from the Programa Nacional de Altas Energías (from institutions on which the centers depend, or from Comunidad Autónoma, etc). Let us see center by center:

a) IFAE: The different investigating groups from the IFAE have independent systems of computation in the same building of the IFAE but in the case of the Grid projects (in special LCG-ES) all the equipment has been located in the central room of computers of the building of computer science services of the UAB (Universitat Autonoma de Barcelona), where it counts on an uninterrupted electrical infrastructure with independent generator, direct connection gigabit to the external network (Anella Cientifica), system of mass storage in robotic tape of 3 Petabytes and operators who cover 12h/day 5 days/week in charge of the budgets.

At this moment all activities GRID the IFAE for ATLAS are performed in the infrastructure installed in the PIC according to the plan described in coordinated project LCG-ES financed by the Programa Nacional de Altas Energías that will end in autumn 2005. As of this date, from the conclusions of the meetings of the manager of the program with the different participant groups in the LHC, the PIC will be in charge of the construction of the Tier-1 LCG project and the ATLAS group of IFAE will participate in the Distributed TIER-2.

b) UAM: the group of the UAM has a production cluster integrated in LCG GRID, that it consists of 40 nodes of work (P4 2,6 GHz, 1 GB the necessary ram) and servants for its integration in Grid, i.e. a Computing Element (CE), 1 Storage Element (SE) and monitoring servers. The capacity of storage of the server is now 1,5 TB. In addition, it has a Monte Carlo Virtual Cluster (GVM). This system takes advantage of the computing resources 'external' to project LCG. At the moment, it is in development phase to implement the new advances of GRID software. Cluster GVM consists of an own ` Grid Service Broker' to free it from the LCG GRID production system and it has as computing nodes the PCs of computer science classrooms of the UAM .

c) IFIC: the IFIC has the GOG. The GOG (Group of Computers for the GRID) is a farm of 200 computers distributed geographically in 2 centers, the IFIC and the ICMOL (Institute of Molecular Science). It was financed with FEDER funds assigned to the Universitat of Valencia. Concretely, at IFIC, there are 138 PCs (half to 1.2 GH and 1.4 GH, 40 GB of hard disk) disposed in racks specially prepared to the adaptation of communications equipment. It has Fast Ehernet adding Gbit Ethernet. In addition we have 4 disk servers with a capacity of 6 TB and a tape robot STK L700e700 with a potential capacity of 134 TB. The GOG is operated by the Computing Service of the IFIC and acts as High Availability ` Monte Carlo Factory' within the ATLAS collaboration

2) I+D of technologies GRID.

It has been always a preoccupation and a motivation of the 3 groups participating in parallel in projects of GRID technologies I+D and GRID infrastructure operation (see section 6).

The IFIC been has involved in activities of middleware, in its integration, in the organization of resources for the deployment of testbed and to provide aid of infrastructure subjects within the European Project CrossGrid. In the present time we are participating in Project EGEE of the Sixth Frame Program with responsibilities that can be seen in section 3.2.

The group of the UAM, financed by project LCG-ES, it has participated in the developments of software for the automatic management of clusters, initially within the work group WP4 of the European project `` European Data Grid' ' and later within the project `` Quattor''. The centers integrated in grid LCG have seen the necessity of a tool as QUATTOR to maintain the structure of Grid Computing with a consistent and free configuration of errors, prepared thus for the production with real data. The number of centers implementing such tool is increasing. The group of the UAM has played a fundamental role in the development of the Quattor system and the appearance of one first stable version.

3) Tests "Date Challenge" from ATLAS

When project LCG-ES began, the GRID was not sufficiently developed as to make productions of Monte Carlo event and to analyse them. In this stage, year 2003, Data Challenge was made first (DC1) to which contributed the IFIC strongly [9]. It is with second Data Challenge (DC2), from summer 2004 to today, when the structure of Grid calculation is used for all the phases of generation, reconstruction and analysis of Monte Carlo events. The set of the three Institutes, IFIC, IFAE and UAM have contributed with 9% of the total production of the ATLAS. Additionally, the IFIC has taken responsibilities within the Equipment from ATLAS Production in the job submission in LCG, one of 3 existing ATLAS GRID environments [10].

4) Preparation and study to establish a system of GRID Distributed Analysis in ATLAS:

The participation in this subject is very recent and has been made by personnel of the IFIC, who have collaborated actively in a part of software within the group ADA (ATLAS Distributed Analysis).

The Distributed Analysis of data is a high-priority subject for the suitable development of the computational model of the experiment ATLAS. It is hoped to obtain a great amount of data coming from the processing of the Raw Data (Raw data) coming from the ATLAS detectors. These data will be stored in multiple centers geographically distributed everywhere to be analysed by physicists working in different centers. GRID Technologies us allow to have the most suitable tools so that that distributed analysis of data becomes possible and improves the productivity of the scientists. Since we have seen previously the first activity that has benefited from this new model of computation has been the Production of Simulated Data and at the moment the following objective is to be able the access to distributed resources to allow the so-called Data Distributed Analysis. ADA Project is being carried out to achieve the previous goal [11].

The Project ADA tries to identify those software components that will facilitate the task of analysis of data of the physicists. The components of ADA already have been identified considering elements of other projects that develop the required functionalities. Between these projects they are GANGA[12], DIAL [13], the ATLAS production system [14],] ARDA [15], AMI, as well as other more generic projects.

.

A first step has been the implementation in a library of routines written in C++ language of the different ingredients when a 'job' in the computer science sense is defined. Between these element, we can mention the application (or program) to execute, the task in concrete that is due to carry out from the application, the input files that are transformed into output files by means of job submission and processing to a queue system, etc. This library is called DIAL and consists of all these elements like classes with the appropriate methods to allow the last processing of the 'job'.

References and Bibliography:

[1] http://www.CERN.ch/ATLAS.

[2] "Detector and Physics Performance ". Technical Design Report. CERNLHCC/99-15 ATLAS TDR 15, 1999.

[3] Memorandum of for Understanding Collaboration in the Deployment and Exploitation of the LHC Computing GRID (LCG), Version of September 2004.

[4] The ATLAS Computing Model, D. Adams, D. Barberis, et al , December 2004,

[5] ` TIER-2 Centres', GRIDPP Project Management Board'. GRIDPP2 Planning Document. GRIDPP-PMB-17-TIER2_v4.0. May 2003

[6]) MATC: Michigan ATLAS Tier-2 Center', To proposal to establish to Research and Development Oriented Prototype TIER-2 Computing Center at the University of Michigan. 28 of February of 2001, University of Michigan, Ann Arbor, MY 48109

[7] ` ATLAS Computing Model', presentation of R. Jones, Software ATLAS Web, 8th December 2004.

[8] "Price extrapolation parameters for the CERN LCG Phase II Computing Farm", B. Panzer-Steindel, CERN, 2004; Report PASTE, Cost of processors and Disk Storage.

[9] ` ATLAS Challenge Data 1', R Sturrock et al. (ATLAS DC1 Task Force); Note ATLAS - SOFT-2003-012.

[10] ` ATLAS Data Challenge 2: To Massive Monte Car it production on the GRID ", S. González, et al. Presented in EGC 2005 (European GRID Conference), Febr. 14-16, 2005, Ámsterdam (Holland)

[11] http://www.usatlas.bnl.gov/ADA

[12] http://ganga.Web.CERN.ch/ganga

[13] http://www.usatlas.bnl.gov/~dladams/dial/.
[14] http://www.nordugrid.org/applications/prodsys.
[15] http://lcg.Web.CERN.ch/LCG/peb/burn

National and International groups that work in the same subject:

Nationals: Group of High Energies of the University of Santiago de Compostela (U.S.), Group of High Energies of the Universitat of Barcelona (UB), Group of High Energies of the CIEMAT of Madrid, Group of High Energies of the IFCA of Santander, PIC of Barcelona.
International: FZK of Germany, RAL of United Kingdom, Italian groups of the INFN (Milan, Pisa, Rome, Bologna, etc), Computing Center of the IN2P3 of France, NIKHEF of Holland, etc

.

3. OBJETIVOS DEL PROYECTO

(máximo **dos** páginas)

◆ **3.1** Describir brevemente las razones por las cuales se considera pertinente plantear esta investigación y, en su caso, la **hipótesis de partida** en la que se sustentan los objetivos del proyecto (**máximo 20 líneas**)

The aim of this project is to provide an infrastructure of a computational GRID and the essential e-Science environment so that the physicists of the Spanish groups participating in ATLAS experiment can develop their works of Physics Analysis from the data collected in this experiment from the Spanish research centers and at the same level that our ATLAS colleagues. In addition to the analysis of real data also it will be necessary to have collections of simulated data, part of which they will be possible to be produced in the Farm of Simulation of the infrastructure which we want to deploy. If this effort is not made, it would be jeopardized the Spanish Contribution directed to obtain results on Physics of the highest level.

This project is scientifically complementary to the projects of the detector construction that the groups have been made during previous years in the LHC experiments (in particular, in ATLAS) ; and it is oriented towards basic scientific objectives of High Energy Physics. This project would also allow to continue the developments of software Grid, like Distributed Analysis, Automatic management of Clusters and Data, etc, that they have begun within present project LCG-ES and that they are of great importance for the correct operation of the system.

We join, between the three centers, the "critical mass" in human and in equipment resources that will allow to provide the infrastructure of Distributed TIER-2 for ATLAS, which that will represent the 5% of the total of all the TIER-2 anticipated for ATLAS.

◆ **3.2.** Indicar los **antecedentes y resultados previos,** del equipo solicitante o de otros, que avalan la validez de la hipótesis de partida

The scientific antecedents and previous results of the applicant team have been mentioned already in the introduction. It is also possible to see the trajectories of the investigating groups that we participated in this Coordinated Project in section 6. However, in this section we will be centered in the antecedents of viability as a project by paying attention to TIER-2 infrastructures which begin to exist, or are going to exist, in other participant countries in the experiment ATLAS, like for example Great Britain and the United States.

In Great Britain they have launched an organization program and a deployment of 4 Distributed TIER-2, they are the TIER-2 called ` London', ` NorthGrid', SouthGrid' and ` ScotGrid' where each one of them groups between 4 and 6 research HEP centers. All these centers are participating in some experiment of the LHC and other experiments of High Energies such as BaBar, CDF, Antares, etc. The association of centers in each TIER-2 has been made following the criterion of geographic proximity instead of a the property a concrete experiment. The organization of the TIER-2 becomes on the basis of the existing resources in each center and the perspective of financing coming from the e-Science program within the thematic Area of High Energy Physics (mainly for the analysis of data of the LHC) that receives the name of GridPP. They have developed a system distributed with a list of groups of activities that can be distributed between the centers. Another important goal of the project is to optimise the hardware and human resources, these last ones, by the complementary distribution of the tasks to be performed in a TIER-2.

On the other hand it is the model that presents the University of Michigan (the United States): one is a TIER-2 located in this University and dedicated only to ATLAS. Unlike the British model, they base the viability of its center on the grouping of different departments and centers (for example, they have coordinated the department of Radiological Sciences and Nuclear Engineering, the Parallel Computer center and the Department of Physics). Therefore, the activities are distributed following a scheme more based on the motivation of the centers which participate in the initiative than in terms of provision of a service.

In the Spanish case, we are 7 centers of High Energies with participation in the experiments of the LHC: 2 in CMS, 2 in LHCb and 3 in ATLAS. The plan is to direct us to a model where we will establish 3 distributed TIER-2: one for LHCb, another one for CMS and another one for ATLAS. In ATLAS we have adopted the solution, just as the British , to organize to us in distributed TIER-2 but taking as criterion the participation in the same experiment and not the geographic proximity. As far as the structure of operation of the TIER-2 raised in section 4 we have also been based initially on the used one by the British instead of the applied by the group of the University of Michigan.

- **3.3.** Enumerar brevemente y describir con claridad, precisión y de manera realista (es decir, acorde con la duración prevista del proyecto) los **objetivos concretos** que se persiguen, los cuales deben adecuarse a las líneas temáticas prioritarias del Programa Nacional al que se adscribe el proyecto (*ver Anexo de la convocatoria*).

  La novedad y relevancia de los objetivos (así como la precisión en la definición de los mismos) se mencionan explícitamente en los criterios de evaluación de las solicitudes (*ver apartado Noveno de la Convocatoria*)

The concrete objectives presented in the project fit completely into the thematic line

"Information Technologies for Particle Physics. Development of Technologies for Computing in the Network (LHC Computing GRID)":

The concrete objectives are:

1. Operation, organization and working setup of the Distributed TIER-2 of coordinated way:

   We have the challenge to deploy a distributed infrastructure of human and computational resources with an objective clearly oriented to the problem of the data coming from ATLAS

2. Establishment of a Stable Installation of Production of Simulated ATLAS Events in GRID in the distributed TIER-2 from the farms of centers (UAM;IFIC)

   There is a program of work coming from the coordination with the Data Challenges Program of ATLAS. To have this operative Installation is of extremely relevant due to the scientific yield that can be obtained.

3. Development and beginning of a System of Distributed Analysis of ATLAS Data based on GRID (UAM;IFAE;IFIC)

   Very novel since it would be a great achievement to deploy such complex and useful system

4.- System of Reconstruction of data and use for calibration and alignment of the ATLAS detectors
(UAM;IFIC;IFAE)

It is one of the functions that it has entrusted to a standard TIER-2 : to provide this service to the HEP physicists that construct and maintain detectors in our centers

5.- Put into operation of the Event Filter and the Calibration of Hadronic Calorimetry (IFAE.)

The Event Filter is an activity very related to the TIER-2 since its assignment is the filtrate of data coming from the detectors and is in the line of the interest of use of the TIER-2 in the works related to the detectors, in this case, with the Hadronic Calorimeter

◆ **3.4. En el caso de Proyectos Coordinados** (máximo **dos** páginas):

- el **coordinador** deberá indicar:

- los objetivos globales del proyecto coordinado, la necesidad de dicha coordinación y el valor añadido que se espera alcanzar con la misma
- los objetivos específicos de cada subproyecto
- la interacción entre los distintos objetivos, actividades y subproyectos
- los mecanismos de coordinación previstos para la eficaz ejecución del proyecto

---

Global objectives, need for coordination and added value:

In summary, one can say that the global objective of this coordinated project is to provide to the ATLAS experiment a Distributed TIER-2 Infrastructure that acts like a meta-centre of human resources and equipment in order to fulfil the TIER-2 assignments as defined by the ATLAS collaboration. Special emphasis is put on one of the most important aim, to provide the Spanish ATLAS physicist's community with access to the GRID Distributed Analysis System.

The following reasons have led the three Spanish groups participating in ATLAS to build jointly a distributed TIER-2 infrastructure and to present a Coordinated Project:

1) On the one hand one should guarantee a permanent service to the collaboration. For example, if one of the centres has a temporary operational problem, the other centres can compensate for it. Also the requests from the Collaboration can be attended more quickly if three centres are implied instead of one.

2) On the other hand, it is aimed at making optimal use of the human resources of each centre which implies a more efficient contribution to the development of software and collaborative systems.

3) the coordinated distribution of resources guarantees the users of each centre the necessary infrastructure to perform analyses of the data of all physics channel of ATLAS.

3) Another important aspect is the communication and relationship within the ATLAS Collaboration: it is important to have an interlocutor that will present the agreed work responsibilities.

The coordination of the Project will be the responsibility to the IFIC.

Specific objectives of each subproject:

The specific objectives are of great importance within this Project since they try to enhance the groups scientific productivity in ATLAS on the base of a consolidated TIER-2 infrastructure. Here are detailed the specific objectives by subprojects:

UAM

The specific objectives of the UAM ATLAS group can be summarised as follows:

- Contribution to the production of Monte Carlo events for the ATLAS experiment
- Contribution to the developments and support related with the administration and management of clusters in a GRID environment.

IFAE

The specific objectives of the IFAE ATLAS group, which are detailed in the methodology section, are related to the group responsibilities in the construction of the ATLAS detector:

- Commissioning of the Event Filter
- Calibration of the Hadronic Calorimetry

IFIC

The specific objectives of the IFIC ATLAS group can be summarised in the following way:

- Characterization, evaluation and optimization of the computing resources of a TIER-2 infrastructure;
- Management of Operations of the TIER-2, as the coordinating centre
- Contribution to the production of Monte Carlo events for the ATLAS experiment

Interaction between the objectives, activities and subprojects of each subproject:

The global objectives mentioned previously are also obvious objective of each subproject. Therefore a strong interaction between the subprojects exists. These global objectives can be characterised as "Provision of Infrastructure and Services". On the other hand, there are objectives associated to certain centres which are due to their experience in these subjects. We have, for example, the use of the Tier-2 infrastructure for the calibration of the Hadronic Tilecal Calorimeter in the case of the IFAE, and the development and support of the Cluster management in the Grid in the case of the UAM.

Coordination mechanism:

We will hold meetings of the principal investigators (PI) of the subprojects every 3 months to make an evaluation of the progress and evolution of the coordinated project.

We will hold weekly or bi-weekly meetings of the different sub-activities that we have proposed.

The meetings will be made by video-conferencing or audio-conferencing in most cases but we believe it is necessary to have at least one PI's meeting in person per year and one meeting per year of the full coordinated project in the form of a workshop with the purpose of sharing the progress and status of the different activities.

At the level of the LHC-Spain GRID Computing, we are going to propose a commission for the coordination of the projects.

4. METODOLOGÍA Y PLAN DE TRABAJO (en el caso de proyectos coordinados deberá abarcar a todos los subproyectos)

Se debe **detallar y justificar con precisión la metodología y el plan de trabajo** que se propone y debe exponerse la planificación temporal de las actividades, incluyendo cronograma (se adjunta un posible modelo a título meramente orientativo).

- El plan de trabajo debe desglosarse en actividades o tareas, fijando los hitos que se prevé alcanzar en cada una de ellas. En los proyectos que empleen el Hespérides o se desarrollen en la zona antártica, deberán también incluir el plan de campaña en su correspondiente impreso normalizado.

- En cada una de las tareas debe indicarse el centro ejecutor y las personas (ver apartados 2.1, 2.2 y 2.3 del formulario de solicitud) involucradas en la misma. Si en el proyecto participan investigadores de otras entidades no relacionados en el apartado 2.3 del formulario de solicitud, deberán exponerse los méritos científicos que avalan su participación en el proyecto.

- Si solicita ayuda para personal contratado justifique claramente su necesidad y las tareas que vaya a desarrollar.

La adecuación de la metodología, diseño de la investigación y plan de trabajo en relación con los objetivos del proyecto se mencionan explícitamente en los criterios de evaluación de las solicitudes (*ver apartado Noveno de la convocatoria*).

Outline of the work plan by activities or work packages:

We started from the base that each centre will require a team of system administrators for the local facilities and that additional specialists will be necessary for the developments associated to the project.

The IFIC will have the participation of two persons, at the level of 50% each and who are not part of the research team of the IFIC, from the Computer Science service of the IFIC for tasks of technical support and operation. These people are Amparo Lacruz Lacruz, who has the charge of Operator of the Computing Centre, and Alejandro Lamas Daviña, contracted technical (I3P) of the CSIC who will have an activity focused on technical support for the hardware installation and maintenance.

Let us review the various tasks that have been identified for the operation of a TIER-2 of ATLAS. For each activity an estimation of the human resources necessary to carry out these tasks is provided. It has been taken into account that the 3 centres must contribute their own personnel, in addition to the one asked in the project, when computing the amount of “Persons Months” (PM). We are going to use PM as a unit, such that one Full Time Equivalent (FTE) during one year is equal to 12 PM; the conversion of FTE to PM will be 1FTE = 24 PM taking into account that full time dedication will last during 2 years.

We can make the following classification of the activities:

a) Administrator of the operations of the TIER-2 centre:

This will be the person in charge of the global operation of the distributed TIER-2, who will interact with the personnel of the centres with the purpose of establishing the necessary technical coordination for the accomplishment of the different operations. In addition this person will act as the interface with the ATLAS TIER-1:

Distribution of manpower:

IFIC	**24PM (+12PM*)**
UAM	**0 PM**
IFAE	**0 PM**

b) Hardware support and system administration:

Since the hardware resources are going to be distributed between the centres involved in the TIER-2 it will be necessary to have one or two system administrators depending on the centre. Once defined the characteristics of the TIER-2 system these people will be in charge of the installation and verification of the acquired hardware and, later, of its administration. They will be coordinated by the Administrator of Operations of the TIER-2.

Distribution of manpower:

IFIC	**36 PM (+12*)**
UAM	**36 PM**
IFAE	**12 PM**

+12 PM* corresponds to the hired person #1 y 6 PM to the participation of A. Lamas (does not appear in the research team being support personnel)

c) Grid Services of Global Management :

The provision of Core services is crucial for the existence of a TIER2. While the services that are required will be decided by the TIER2, they must include information services such as hierarchical GIIS, users services such as the management of Virtual Organizations and Accounting services to insure that the centres respect the sharing policies.

Distribution of manpower:

IFIC	**30 PM**
UAM	**0 PM**
IFAE	**18 PM**

d) User support

The TIER-2 will be equipped with "User Help Service" that in order to be effective will require a structure that regroups the human resources dedicated by all the centres that form the distributed TIER-2. One should take into account that some centres have already initiatives and efforts directed to the user support t (this is the case of the IFIC and IFAE that are already providing the user support for the SW Federation of the EGEE). The fundamental idea is to take advantage of this trajectory to approach efficiently the task of user support for the TIER-2. We estimated that of the order of 0,5 FTE will be necessary in each centre to cover this activity.

A possible proposal is:

- to establish a ticketing system where the TIER-2 users will be able to direct their questions and where all the members of the support team will participate (it will be mandatory for them);
- at least one person in each institute will the role of liaison aid between the users and the global support structure. The arrival of questions and request for support is quite random and of chaotic character, this is the reason why the effectiveness of work of these persons in other tasks will be quite reduced. This person will have to solve a significant fraction of the user problems.

Distribution of manpower:

IFIC	**12 PM**
UAM	**12 PM**
IFAE	**12 PM**

e) Experiment software support

The role of a TIER 2 is to make possible that the physicists can perform efficiently their investigations in their own centres. It is of no use to develop a distributed GRID system if the software of the experiment is not maintained. The level of maintenance of software must be an objective of agreement between the ATLAS experiment and the TIER2.

Distribution of manpower:

IFIC	**6 PM**
UAM	**6 PM**
IFAE	**18 PM**

f) Production of simulated events using GRID technology .

The exercises leading to the validation of the Computing Model are going to continue during the period that will include this project. It is foreseen to perform a Data Challenge 3 (DC3) that will be carried out between the end of 2005 and the first half of 2006. DC3 will be the last occasion to make a test of the Computing Model of ATLAS on a large scale before the beginning of data taking.

Before the start of DC3, some tasks will have to be finalised like:

- Online operation of the ATHENA framework-
- Monitoring of the software of the algorithms and the data flow
- Selection and reconstruction algorithms (for Level 2 and Event Filter)
- Misalignment of the detector in the simulation
- Calibration and alignment procedures for the detector in the main data flow
- Condition data base
-Detector inefficiency

The work of integration will take place between end of 2004 and summer of 2005: the objective is to have a system as complete as possible for DC3.

Distribution of manpower:

IFIC	**48 PM**
UAM	**12 PM**
IFAE	**0 PM**

g) Fabric Management Support

The development of Grid fabric management will continue, as initiated within LCG-ES project. They are important questions still to be solved as they are, on the one hand, additional components to be developed for some Grid services and for the operating system and, on the other hand, there are some problems associated to the scalability and security of the system that have to be solved. A support for the other centres will be established about the configuration and the use of the Quattor system and errors arising due to the use in Grid production will be corrected.

Distribution of manpower:

IFIC	**0 PM**
UAM	**24 PM**
IFAE	**0 PM**

h) Development, Installation and Startup of an GRID-based distributed analysis

As the current requirements for the ADA software functionality will be fully implemented at project startup, due to the IFIC involvement, we propose to create a prototype to be used as test bed for a reduced number of users. Once the functionality would be as required, we propose to open the service to the whole ATLAS community.

Today we have two models based in the DIAL library, still in a pre-release phase. In one case, DIAL library is used to transfer the job to an scheduler based on Condor or PBS using a web service. In the second case, an interface from the DIAL library to the EGEE Middleware called gLite has been made. Both cases are in a preliminary phase; however the first one is more robust with a mature user interface and it is operated from the ATLAS group at the BNL (Brookhaven National Laboratory, USA). On the other hand the second system has a more clear connection with the Grid interfaces as he is using the grid job scheduler and can be used in a more generic way. A test infrastructure is being built at CERN with a limited number of hosts. From the two models, we think that the second option is the choice for our project.

It is important to open as soon as possible the system to the user community to get advantage from their experience for new releases. The ultimate goal is to get an distributed data analysis facility to help the Spanish physicists working in ATLAS.

Distribution of manpower:

IFIC	**54 PM**
UAM	**30 PM**
IFAE	**12 PM**

i) Event Filter Commissioning:

The Event Filter project has a lot of similarity with the Tier-2. The Event Filter is a grid-based computer farm running online applications to filter the flow of events. However, the software to select the events is based in the offline reconstruction program called ATHENA. The selection algorithms of the Event Filter will be implemented directly into ATHENA. The code will be exactly the same in the online event filtering and the offline simulation with Monte Carlo data. In this way all the aspects of ATHENA can be shared with the offline environment (the Tier-2), the reconstruction algorithms, and database of calibration, between others. The IFAE group in ATLAS is in charge of the ATHENA aspects in the Event Filter. It is planned to have regular tests: to configure partially the Tier-2 facility at IFAE as Event Filter farm, installing the offline and online software and to run the trigger using simulated data. The goal is to monitor the whole system, in order to identify and to improve issues on CPU, disk, network and databases. All this work will be useful to find the optimal configuration for the event filtering as well for the data analysis.

Another aspect relevant to the Tier-2 from the Event Filter is the trigger efficiency studies. From simulated data we will evaluate the different trigger algorithms and the impact from the available calibration from the DAQ system or hardware failures from the detector such as dead channels. To make such studies we will need to make a specific reconstruction of the data with ATHENA and its corresponding analysis.

Distribution of manpower:

IFIC	**0 PM**
UAM	**0 PM**
IFAE	**12 PM**

j) Calibration of the hadronic calorimeter:

The IFAE responsibilities in the Jet calibration requires specific data processing with specific versions of the reconstruction program in order to obtain the calibration constants, the monitoring of the jet reconstruction algorithms in the calorimeter with a selection of specific data. The interesting data will come from the main data stream or from preselected data coming from an stream created at the Event Filter level which can be sent directly to the IFAE through the network for its exclusive use.

The responsibilities of the IFAE in the Tilecal calorimeter include the calibration of the detector at electromagnetic scale using muons. During the start-up phase of the detector, it is foreseen to take cosmic runs and with only a beam from the end of 2005 and 2006 respectively. The IFAE group will analyze this data as soon as it is being available.

Distribution of manpower:

IFIC	**0 PM**
UAM	**0 PM**
IFAE	**12 PM**

k) Characterization, Evaluation and optimization of Tier-2 computer resources

The continuous development of new technologies in the processor's design, data storage systems and networking will provide a significative improvement of performances. The profit of the new performances comes from a permanent effort to evaluate new hardware in different configurations. It is needed to characterize the performance by building prototypes before a major purchase.

In computer servers the architecture of the main computer board, the CPU technology and the memory speed will contribute directly to the performance of the simulation, reconstruction and analysis programs. The collection of speed time for the programs using different platforms with a different number of CPUs, different memory sizes and types of local hard disk and networking connections as a function of the cost and user requirements is basic for the correct construction of the Tier-2 facility.

In disk servers not only the storage capacity is important, the input/output data rate must be adequate to the processing needs, while preserving services in case of hardware failures. The configuration and choice of adequate technologies in disks, controllers, networking interfaces and drivers is critical to achieve the fastest analysis as possible.

As a result of the characterization and validation of the hardware, one or more choices will be proposed to the data processing, storage and networking systems in order to achieve an adequate performance and help to implement and manage the hardware of the Tier-2 facility.

Distribution of manpower:

IFIC	**18 PM**
UAM	**0 PM**
IFAE	**0 PM**

l) Organization and management of the Project: Representation in ATLAS

The organization and global management of the project is responsibility of the "Investigador Coordinador del Proyecto" with the following tasks:

- Organization of coordination and technical meetings inside the project in line with the lines developed in the section 3.4;
- He will represent the Tier-2 distributed facility in the ATLAS collaboration.
- He will establish the coordination links between the Spanish projects related with the LCG project.

Distribution of manpower:

IFIC	**12 PM**
UAM	**0 PM**
IFAE	**0 PM**

Justification of the Personnel to be hired with Project funds:

The groups will provide a total of 10 FTE from their own personnel budget (as staff, postdoctoral students …): 6 from IFIC, 2 from IFAE and 2.5 from UAM. For the adequate development of the tasks we propose to hire 8 additional people with the following distribution: 4 for the IFIC, 2 for the IFAE and 2 for the UAM. In the section 4.1 the distribution of these people are shown and in the forms submitted by the subprojects the profile of this personnel is detailed. This project will use a total of 18.5 FTE.

4.1 MODELO DE CRONOGRAMA (ORIENTATIVO)

En este cronograma deben figurar la totalidad del personal investigador incluido en el formulario de solicitud y, en su caso, el personal contratado que se solicite con cargo al proyecto.
Debe subrayarse el nombre de la persona responsable, en cada tarea.

.

Actividades/Tareas	Centro Ejecutor	Persona responsable y otras involucradas	Primer año (*)	Segundo año (*)	Tercer año (*)
TIER-2 Operations Manager	IFIC	J.Sánchez (A. Lacruz)	x x| x| x |x x|x x| x |x	x | x | x | x |x |x |x |	| | | | | | | | | |
Hardware maintenance and Tier-2 system administration	IFIC UAM IFAE	Contract # 1 Contract # 2 (A. Lamas) Contract 1 Contract 2 Contract # 1	x x| x| x |x x|x x| x |x	x x| x| x |x x|x x| x |x x	| | | | | | | | | |
GRID services of Global Management (TIER2)	IFIC IFAE	A. Fernández Contract #2 A. Pacheco Contract #1	x x| x| x |x x|x x| x |x	x x| x| x |x x|x x| x |x x	| | | | | | | | | |
User Support	IFIC UAM IFAE	M. Kaci Contract # 2 Contract #2	x x| x| x |x x|x x| x |x	x x| x| x |x x|x x| x |x x	| | | | | | | | | |

	UAM	M. Peez			
	IFIC	J. Valls	x x| x| x |x x|x x| x |x	x x| x| x |x x|x x| x |x x	| | | | | | | | | | |
Atlas software support	IFAE	A. Pacheco Contract #2			
	IFIC	J.Salt, E.Ros S. González, L.March (B) Contract #3			
			x x| x| x |x x|x x| x |x	x x| x| x |x x|x x| x |x x	| | | | | | | | | | |
	UAM	C. Glasman			
Monte Carlo production of simulated data					
	IFIC	F.Fassi, S. González, Contract #4			
		D. Jordán (B)	x x| x| x |x x|x x| x |x	x x| x| x |x x|x x| x |x x	| | | | | | | | | | |
Grid distributed analysis	UAM	J.d.Peso, M Peez			
	IFAE	M. Dosil			

(*) Colocar una X en el número de casillas (meses) que corresponda

Actividades/Tareas	Centro Ejecutor	Persona responsable y otras involucradas	Primer año (*)	Segundo año (*)	Tercer año (*)
Event Filter Commisioning	IFAE	H.Gaitaonandia	x x| x| x |x x|x x| x |x	x | x |x | x |x | x | x | x |	| | | | | | | | | | |
Calibration of the Hadronic Calorimeter	IFAE	D.Sushkov	x x| x| x |x x|x x| x |x	x x| x| x |x x|x x| x |x x	| | | | | | | | | | |

Customization and optimization of TIER-2 resources	IFI C.	Contract #4 L. March (B)	x x| x| x |x x|x x| x |x	x x| x| x |x x|x x| x |x x	| | | | | | | | | | |
Fabric Management Support	UAM	R. Garcia	x x| x| x |x x|x x| x |x	x x| x| x |x x|x x| x |x x	| | | | | | | | | | |
Project Management	IFIC	J. Salt	x x| x| x |x x|x x| x |x	x x| x| x |x x|x x| x |x x	| | | | | | | | | | |
			| | | | | | | | | | |	| | | | | | | | | | |	| | | | | | | | | | |
			| | | | | | | | | | |	| | | | | | | | | | |	| | | | | | | | | | |

(*) Colocar una X en el número de casillas (meses) que corresponda

(B) : it indicates predoctoral fellowship working in his/her Research Work, D.E.A (Diploma de Estudios Avanzados) or Ph.D. Thesis.

5. BENEFICIOS DEL PROYECTO, DIFUSIÓN Y EXPLOTACIÓN EN SU CASO DE LOS RESULTADOS
(máximo **una** página)

Deben destacarse, entre otros, los siguientes extremos:

- Contribuciones científico-técnicas esperables del proyecto, beneficios esperables para el avance del conocimiento y la tecnología y, en su caso, resultados esperables con posibilidad de transferencia ya sea a corto, medio o largo plazo.

- Adecuación del proyecto a las prioridades de la convocatoria y, en su caso, del Programa Nacional correspondiente.

- Plan de difusión y, en su caso, de explotación, de los resultados del proyecto, el cual se valorará en el proceso de evaluación de la propuesta (ver apartado Noveno de la convocatoria) y en el de seguimiento del proyecto.

- Expected scientific and technical contributions from this Project:

Taking into account the main goal of this Project it becomes clear that a basic contribution will be provision of a medium sized scientific infrastructure which will allow the work of the physicists in ATLAS data analysis. However, during the development of the prototypes and the construction of such infrastructure some computing issues have to be solved. These issues will be reported as contributions to conferences and papers in specialized scientific reviews in the same way as has been done in the LCG-ES project. As it can be verified in the EPO letters of interest attached, the technological development at IFIC is being of interest for the private and public sectors

- This project represents a very good match with the priorities written in the MEC call for proposals, we detail the following points:

    1. We think that this is a quality contribution to the research and to the regional development of Spain and in particular of the regions where are based the centers.
    2. We present this project as a Coordinated project with several subprojects with a main goal which will benefit all the groups to taking into account their complementarity's which respect to their competences.
    3. The main goal of this project is focused on high energy physics but some partial goals o the development of the project is of interest to other fields. We can include this project in the field of e-science.
    4. With respect to the "Programa Nacional de Física de Partículas" (PNFAE), this project is strongly related with the CERN LHC project. In particular, with the LHC Computing Grid (LCG). This permits to classify this project as "Tecnologías de la Información para la Física de Partículas. Desarrollo de tecnologías para el cálculo en red" (page 42109 of the BOE)

- Dissemination Plan

    The main method we use to disseminate our advancements is by publishing our work in papers of specialized reviews, oral presentations and participation in workshops, internal meetings between others. In addition to the presentations in project conferences we will use postgraduate and doctoral programs at CSIC and at the Universities. In the other hand we will disseminate our achievements using the press, project reports between others.

6. HISTORIAL DEL EQUIPO SOLICITANTE EN EL TEMA PROPUESTO (En caso de Proyecto Coordinado, los apartados 6. y 6.1. deberán rellenarse para cada uno de los equipos participantes)
(máximo **dos** páginas)

- **Indicar las actividades previas del equipo y los logros alcanzados <u>en el tema propuesto</u>:**

  Si el proyecto es continuación de otro previamente financiado, deben indicarse con claridad los objetivos ya logrados y los resultados alcanzados.

  Si el proyecto aborda una nueva temática, deben indicarse los antecedentes y contribuciones previas del equipo, con el fin de justificar su capacidad para llevar a cabo el nuevo proyecto.

Este apartado, junto con el 3, tiene como finalidad determinar la adecuación y capacidad del equipo en el tema (y en consecuencia, la viabilidad de la actividad propuesta).

Although the groups are involved in several experiments, we are going to restrict ourselves of the ATLAS project only along this record. The three groups IFAE, IFIC and UAM started their research in the ATLAS Collaboration about 10 years ago. The ATLAS detector inside the Large Hadron Collider (LHC) at CERN, which will reach a new energy and luminosity regime, will provide measurements at the frontier of Particle Physics. The ATLAS detector is 22 m high and 44 m long. It uses new technologies in High Energy Physics like the superconducting toroids, the silicon pixel detector (with tens of million electronic channels) and the electromagnetic calorimeter (with the novel accordion shape). The overall cost is about 600 MCHF. The ATLAS Collaboration includes 150 institutes all over the world that makes about 1800 physicists and engineers. The 3 groups IFAE, IFIC and UAM have contributed to a wide spectrum of topics related with the design, construction and preparation of the ATLAS detector.

Concerning computing, the 3 groups participate in the LHC Computing Grid project (LCG) through the Spanish project LCG-ES already described in a previous section. The LCG project started in 2002 coordinated by CERN. Its main goal is to deploy a Grid computing system to allow physicists from all over the world to analyze the huge amount of data (12-14 Petabytes/year) which will be produced when the LHC begins to operate (in 2007). Simultaneously, in October 2002, started the LCG-ES project with participation of the 7 Spanish institutes involved in the LHC experiments, namely: IFCA (Santander), IFIC (Valencia), CIEMAT (Madrid), UAM (Madrid), UB (Barcelona), USC (Santiago de Compostela) and IFAE (Barcelona). The center IFAE is the coordinator of this project through PIC Computing Center.

In what follows we summaries the record of ATLAS related activities in the last 5 years of the 3 groups.

A) ATLAS group at IFIC.

- Contribution to the construction of ATLAS sub-detectors.
  There are two projects: Hadronic Tile Calorimeter (TileCal) and Barrel Silicon Tracker (SCT).
  a) The TileCal group is composed by 4 permanent physicists, 1 post-doc, 2 engineers and 5 fellows. The group collaborates with the TileCal group of the IFAE from the beginning of the project. The contribution to the TileCal has been the mechanics and test of the photomultipliers and production and test of a set of Read Out Drivers (ROD). The main researcher of the project is Prof. Emilio Higon.

  b) The SCT group is composed by 5 senior physicists, 1 post-doc and 3 students. The group commitment has been the production of 220 modules and their integration in the detector. In parallel other tasks related

to software are being performed, namely: the analysis (including simulation) of the Combined Beam Test (CBT) data and the development of algorithms for the SCT alignment. These algorithms will also be used for the cosmic test planned before the SCT is inserted in ATLAS. Additionally some analysis programs are in preparation which will use the tracks and secondary vertexes from reconstruction in an optimal way. The main researcher of the group is Carmen Garcia.

- Contribution to the ATLAS physics

From 1996 the IFIC has participated in different topics of LHC physics. From 1996 till 1999 the improvement of the b tagging algorithms for event selection. More recently the IFIC is working in Monte Carlo studies of different processes beyond the Standard Model like "Little Higgs" and "Extra Dimension" models. These models predict the existence of heavy resonances like Z_H, W_H, Z*, W*, g*,.... with same properties as the corresponding Standard Model particles (Z,W,g,..) but masses of several TeV. The IFIC is studying the decay of these resonances in the $b\bar{b}$ and $t\bar{t}$ channels using both the ATLAS fast simulation (ATLFAST) and the complete simulation with GEANT3.The following people are working in the topic: E. Ros, J. Valls and L. March.

- Computing:

3.1) Projects inside the High Energy Physics National Program.

a) "Accion Especial" (2002) with title "Preparación de la Infraestructura Local para la Participación en el Proyecto de Cálculo en la Red (GRID) del experimento ATLAS (LHC) del CERN" (see section 6.1 for details). This allowed the basic knowledge to start participating in the computing of the LHC projects, in particular of ATLAS.

b) LCG-ES project (2002-2005), running at present (see section 6.1 and the Introduction). The IFIC has contributed mainly to the ATLAS Data Challenges, both in the Monte Carlo event generation and the System for Distributed Analysis. The IFIC is coordinating the task of "Operations" of LCG-ES (job done by Dr. Santiago González de la Hoz). The members of the LCG-ES project are: J. Sánchez, E. Ros, S. González, F. Fassi, L. March, J. Lozano y J. Salt. More details of the LCG-ES are given in the record of the IFAE group.

3.2) CrossGrid project. Contribution of the IFIC: a) development of a Resource Broker adapted to the interactive applications of the project; b) integration of the middleware developed by the CrossGrid Collaboration; c) deployment of a testbed system; d) development of a system for user support. High level responsibilities: J. Salt is coordinator of the task "Infrastructure Support" and S. González has been coordinator of the working group "Integration of CrossGrid Middleware". A. Fernández participates in Scheduling Agents and M. Kaci in the deployment of Helpdesk for user support.

3.3) EGEE (Enabling Grids for e-Science in Europe) project. The main contributions of the IFIC are: development of a system with IFIC as a resource center, management of Virtual Organizations (VO) and the system of user support for southwest Europe. Dr. J. Salt is CSIC representative in the EGEE Collaboration Board. A. Fernández is responsible of the VO management and J. Sánchez is the system manager of the IFIC resource center. More details of the EGEE project are given in the record of the IFAE group.

3.4) In addition, the group has lead some initiatives to establish a Grid and e-Science in Spain, for example the “Acción Especial TIC 2002-11109-E” (known as IRISGRID). Moreover, the IFIC has been funded to improve the computing infrastructure in the past 4 years. Also the “ Conselleria de Industria de la Generalitat Valenciana” has funded the creation of a e-Science Metacenter composed by “CSIC” , “Universitat de Valencia-Estudi General” and “Universidad Politécnica de Valencia

B) ATLAS group at UAM.

A) Contribution to the construction of the ATLAS Electromagnetic Calorimeter.
The ATLAS group at UAM has been, together with the CPPM of Marseille, responsible for the design and construction of the ElectroMagnetic calorimeter of the EndCap region of ATLAS (EMEC). The EMEC is a sampling calorimeter with excellent energy and position resolutions, needed to the discovery and the study of the Higgs boson. The main commitments of the UAM group in this construction have been:
1.1) Design and production of absorbers. About 2000 pieces machined with an accordion shape, made of lead cladded with steel. The demanding energy resolution lead to strict tolerances in the dimensions of the pieces, namely 0.01 mm for the lead thickness and 0.1 mm for the overall (x,y) dimensions.
1.2) Design and production of some cold electronics. About 600 boards of 9 types to distribute the high voltage, 4000 Summing Boards of 32 types and 2000 test boards of 10 types for the quality control during the stacking of the modules.
1.3) Assembly of half of the EMEC modules (8 octants) in the clean room of our laboratory at UAM. An extra module (reference module) has been assembled as well, which can be used as a test module in the future if needed to understand an unexpected behavior in the EMEC response. The stacking or assembly of the modules finished in the summer 2004.
1.4) Beam tests with electrons in the North Area at CERN. There have been 4 tests, the first one with a prototype module and the rest with 3 final modules. (The EMEC consists of 16 modules in total). The UAM group has contributed to all the stages of beam tests, namely the setup, the data taking and the analysis of the data.

As an academic result of this contribution to the construction and test of the EMEC we remark the 3 PhD thesis presented in the last 5 years and one more being submitted by the end of 2005.

- Contribution to the ATLAS physics.
  Recently the group has begun Monte Carlo studies related with the Higgs production. Only two decay modes are tagged, namely the one in 2 photons and the one in 4 leptons. The choice of this topic is determined by the fact that the EMEC plays an important role in the energy and position measurement of the decay products, photons and electrons.
- Contribution to Computing projects.
  The LCG-ES project (2002-2005) already mentioned in the Introduction. The main contributions of the UAM group to this project have been:
  3.1) Development of a system for the installation, configuration and management of the software and operating system, called Quattor.
  3.2) Construction of a Grid node, initially for the testbeds of the EDG project and afterwards for the testbeds of the LCG project. In particular, Monte Carlo events of the ATLAS Data Challenge 2 have been generated in this node.
  3.3) System to exploit computing resources that are not considered in the LCG-ES project.

C) ATLAS group at IFAE.

Contribution to the construction of the ATLAS hadronic Tilecal Calorimeter

The ATLAS group at IFAE has participated to the construction of the hadronic Tile calorimeter from 1993. The tasks performed have been the construction of a prototype from 1993 till 1996 (RD34 project), the design of the subdetector, the electronics for calibration, the assembly and setup of the 64 barrel modules (2002 till now), the analysis of the data from tests with both beams and radioactive sources and the simulation of the subdetector response to several types of particles.

Contribution to ATLAS physics

The ATLAS group at IFAE has participated from 1994 in the Jets and Transverse Energy studies for the LHC physics, performing several studies on the characteristics of the whole ATLAS calorimetry, algorithms for jet reconstruction, development of calibration methods and the coordination of the working group in ATLAS. Moreover, some studies have been performed (PhD thesis of Dra. Mireia Dosil) on physics topics as the study of the discovery and measurement of the mass of the neutral Higgs Boson in the Standard Model ($H \rightarrow ZZ \rightarrow llnn$) and the charged Higgs Boson (supersymmetry) ($t \rightarrow H^{\pm} b$; $H^{\pm} \rightarrow tu$).

European DataGrid (EDG)

The IFAE has participated in the European DataGrid (EDG) project from 2001 till 2004 in the working group "Integration Testbed", for the deployment of an international infrastructure of distributed computing, which can be used to test scientific applications that manage large amount of data. This infrastructure has demonstrated the performance of a data grid to deal with such type of problems, operating in a high performance network.
During the three years of the project, the IFAE group has coordinated the deployment of the EDG infrastructure in several spanish centers: IFCA (Santander), IFIC (Valencia), CIEMAT (Madrid) and UAM (Madrid). A basic Certification Authority (CA) has been implemented, which allows to integrate in the Grid the resources of the spanish centers in a transparent way, as well as to allow the scientists of this centers to access the overall resources provided by other countries to the Grid.

The IFAE group has kept computing and storage resources connected to the EDG Infrastructure for the whole period of the EDG operation. Special effort has been put in order that these resources could have transparent access to massive storage systems, since this should be one of the fundamental characteristics of a Data Grid. In the beginning of 2003 the PIC (Puerto de Información Científica) is created as an IFAE division and as Grid node for intensive data processing of scientific and technological nature. The PIC has a Storagetek tape robot with a capacity up to 1 PetaByte which is managed by the CASTOR system to access the data in transparent manner. The PIC was one of the first centers to connect a massive storage system to the EDG infrastructure and it is at present one of the three suppliers of this service. The PIC has also collaborated in the process of integration of the Grid middleware, developing testbed and certification platforms, participating actively in thedevelopment of this software.
Dr. Andres Pacheco was member of the EDG project.
Participation in the construction of a cluster for the data filtering of the third level trigger (Event Filter) of the ALTAS experiment
The IFAE joined the Event Filter, the third level acquisition system and selection of ATLAS data. The Event Filter is a high performance cluster that has access to the complete information of the ATLAS detector and will use the same algorithms of the offline reconstruction when possible. The Event Filter will carry out other tasks: monitoring of the detector status and of the whole selection process; the on-line analysis of the events useful for calibration,

as well as fast identification of events of special interest for physics to give them high priority in the subsequent reconstruction. The IFAE group is interested in all aspects of the Event Filter related to physics and it contributes to the development of the infrastructure as well.

In 1999 A. Pacheco and the post-doc K. Karr made some contributions, in collaboration with the CPPM Marseille, on the implementation of the Java language of one of the most critical Event Filter components, the data flow. Such implementation inspired the design of the data flow, presented in the TDAQ Technical Proposal, that started a new phase of the project. The IFAE group took the commitment of following the development of the new Object Oriented offline Reconstruction code and make possible the use of the framework ATHENA/GAUDI in the Event Filter. This implied to develop many new software elements.The fellow M. Dosil, in collaboration with the Event Filter group of Marseille, wrote a program to convert from offline to DAQ data format and viceversa. It was a C++ program in the ATHENA framework and one of the few Object Oriented algorithms written at that time. A first implementation of the communication protocol between the data flow and ATHENA was performed as well as an estimation of the CPU power needed by the Event Filter. These studies were the basis for a new iteration in the Event Filter design. Doctors K. Karr, G. Merino, M.Dosil, C.Sánchez worked in this design. Dr. Sushkov is at present in charge of the task with the help of an engineer, H. Garitaonandia, and a student, E. Segura. At present, Doctors Sergei Sushkov, Mireia Dosil y Andreu Pacheco, and the engineer Hego Garitaonandia are members of the ATLAS Collaboration. Dr Andreu Pacheco is member of the International Computing Board (ICB) of the ATLAS Collaboration

LHC Computing Grid (LCG)

In July 2003 the first release of the LCG Grid infrastructure was deployed (LCG-1). It was built mostly by the middleware developed by the EDG project. Due to the experience acquired in the EDG project, the PIC could deploy LCG-1 very fast, being the first spanish center and one of the 4 centers in the world which could join the LCG-1 testbed. In September 2003, two working groups of the LCG-ES project started, to coordinate the increasing activity foreseen with the LCG-1 deployment. One of these groups, “Development and Implementation”, has been coordinated by the PIC. Tasks of this group are related with the implementation of new Grid services (architecture design, deployment, testbeds, etc) until they enter in the production phase.

The first coordinated action of the group was the incorporation of all spanish groups to LCG-1. This process started on October 17th and finished November 12th. The PIC provided support to the rest of the spanish centers and it was also the link with CERN. As a consequence, Spain was the first country incorporating a “large” number of centers to LCG-1 in a systematic way. On December 5th 2003, the Spanish Ministry for Science and Technology nominated PIC as a Regional Tier-1 Center for processing and storage of the LHC data.

Doctors M. Dosil and A. Pacheco are members of the LCG project. Dr. A. Pacheco is member of the Grid Deployment Board of the LCG project.

Enabling Grids for E-Science in Europe (EGEE)

From beginning 2002 the IFAE group has been involved in the design and negotiations of the EGEE project. In October 2003 the project was approved and in 2004 started. A total of 70 european and russian institutions participate in EGEE and a coordination with similar projects (US, Japan, etc) is foreseen. The EGEE project deploys and operates in production the first international large scale Grid infrastructure based on a scheme of federated Regional Centers of Operation and Centers of Critical Infrastructure. The PIC manages the Operational Center of the Southwest-Europe (Spain and Portugal) and coordinates, with RedIris and MEC, the spanish participation in EGEE. Dr. A. Pacheco is the manager of this Regional Center for Operations in the South-west Europe.

Doctors M. Dosil and A. Pacheco are members of the EGEE Collaboration. Dr. A. Pacheco is member of the Collaboration Board.

6.1 FINANCIACIÓN PÚBLICA Y PRIVADA (PROYECTOS Y CONTRATOS DE I+D) DE LOS MIEMBROS DEL EQUIPO INVESTIGADOR

Debe indicarse únicamente lo financiado en los últimos cinco años (2000-2004), ya sea de ámbito autonómico, nacional o internacional.
Deben incluirse las solicitudes pendientes de resolución.

Título del proyecto o contrato	Relación con la solicitud que ahora se presenta (1)	Investigador Principal	Subvención concedida o solicitada EURO	Entidad financiadora y referencia del proyecto	Periodo de vigencia o fecha de la solicitud (2)
Enabling Grid for E-Science in Europe (EGEE)	1	M. Delfino Reznicek	C 248.000 €	EU Commission IST-2003-508833	2004-2005
Desarrollo de infraestructura DATAGRID para análisis de datos del LHC	1	M. Delfino Reznicek	C 956.952 €	MCYT FPA2002-04208-C07-04	2002-2005
Preparación infraestructura computación del IFAE para la participación en el experimento ATLAS	1	A. Pacheco	C 87.146 €	MCYT FPA2001-3969E	2002
DATAGRID	1	E. Fernández Sánchez	C 128.784 €	EU Comisión IST-2000-25182	2001-2003
Física Protón-Protón en el LHC con el Detector ATLAS	2	Martine Bosman	C 1.310.000 €	MCYT FPA2003-00407	2003-2006
Física Protón-Protón en el LHC con el Detector ATLAS	2	M. Cavalli-Sforza	C 1.078.800 €	CICyT FPA2000-1693	2000-2003

Table 1. Financial support of the IFAE staff participating in the project

(1) Escríbase 0, 1, 2 o 3 según la siguiente clave:
0 = Es el mismo proyecto 1 = está muy relacionado 2 = está algo relacionado 3 = sin relación
(2) Escríbase una C o una S según se trate de una concesión o de una solicitud.

Título del proyecto o contrato	Relación con la solicitud que ahora se presenta (1)	Investigador Principal	Subvención concedida o solicitada EURO	Entidad financiadora y referencia del proyecto	Periodo de vigencia o fecha de la solicitud (2)
``Desarrollo de una Infraestructura de DATAGRID para el análisis de datos del LHC" (ref. FPA 2002-04208-C07-05)	1	Dr. J. F. Salt Cairols	418.000'- Eur (concedida, con costes indeirectos)	CICyT, Ministerio de Ciencia y Tecnología España FPA2002-04208-C07-05	C 2002-2005
Development of GRID Environment for Interactive Applications (ref. IST-2001-32243).	1	Dr. Michal Turala	455.000'-Eur (CSIC)	Quinto Programa Marco, IST, Unión Europea IST-2001-32243	C 2002-2005
Enabling GRID Computing for e-Science in Europe (EGEE)	1	Dr. Fabrizio Gagliardi	128.000'- Eur	Sexto Programa Marco, I3, Unión Europea EU-508833	C 2004-2006
'Preparación de Proyectos GRID en el marco de las iniciativas de e-Ciencia en Europa'	1	Dr. J.F. Salt Cairols	21.000'- Eur	Acción Especial TIC, Ministerio Ciencia y Tecnología. España TIC2002-11109-E	C 2003-2004
'Construcción de un Detector de Silicio para ATLAS'	2	Dra.Carmen García	1.776.600' Eur	M.C.y T. FPA2003-03878-C02-01	C 2003-2006
'Contribuciones al Calorímetro Hadrónico TILE CAL de ATLAS'	2	Dr. Emilio Higón	1.081.600'-	M.C.y T FPA2003-09220-C02-01	C 2003-2006

Table 2. Financial support of the IFIC staff participating in the project

Título del proyecto o contrato	Relación con la solicitud que ahora se presenta (1)	Investigador Principal	Subvención concedida o solicitada EURO	Entidad financiadora y referencia del proyecto	Periodo de vigencia o fecha de la solicitud (2)
Desarrollo de infraestructura DATAGRID para análisis de datos del LHC	1	José del Peso Malagón	C 233.040 Eur.	MCYT FPA2002-04208-C07-03	2002-2005
Construcción del Calorímetro Electromagnético de Argón Líquido de ATLAS	2	Fernando Barreriro Alonso	C 863.376 Eur.	MCYT FPA2002-01008	2002-2005

Table 3. Financial support of the UAM staff participating in the project

7. CAPACIDAD FORMATIVA DEL PROYECTO Y DEL EQUIPO SOLICITANTE (En caso de Proyecto Coordinado deberá rellenarse para cada uno de los equipos participantes)

Este apartado sólo debe rellenarse si se ha respondido afirmativamente a la pregunta correspondiente en el cuestionario de solicitud.
Debe justificarse que el equipo solicitante está en condiciones de recibir becarios (del Programa de Formación de Investigadores) asociados a este proyecto y debe argumentarse la capacidad formativa del equipo. En caso de Proyecto Coordinado, debe rellenarse por cada subproyecto que solicite becarios de FPI.

A) Training ability of the IFIC group.

The research team of the IFIC has made training in the period 2000-2005 incorporating the following members:

- FPI fellow (Luis March Ruiz) who started in May 2003 in the LCG-ES project and has presented his research work at "Universidad de Valencia" in September 2004 in the Department for Atomic, Molecular and Nuclear Physics (FAMN). His is working in his PhD thesis at present.
- D. Alvaro Fernández Casani will present soon his work with title "Grid Architectures for Resource Management" at "Universidad Politécnica de Valencia"
- Fellow M. Dolores Jordán, has joined the LCG-ES project recently (in January 2005). She will work on ATLAS Distributed Analysis.

The IFIC team has shown its training ability and is planning to continue it. (Three FPI fellows are requested).

The IFIC team participate in the joined Postgraduate Program of three Departments of the Valencia University, namely: FAMN, Theoretical Physics and Astrophysics. In particular, there is a course on "Grids and e-Science" linked to the project. Doctors E. Ros and J. Salt teach in this Program.

B) Training ability of the IFAE group.

Since the IFAE joined the ATLAS Collaboration, the group has received many fellows, FPI and from the Generalitat de Catalunya. In total 12 postgraduate research works in physics, 2 in engineering and 2 PhD thesis have been presented. One more postgraduate work on "Software Design of the Event Filter Application" is being finalized at present, which will apply the framework of data reconstruction and analysis that will be used in an ATLAS TIER2.The group offers a PhD thesis work on the "Qualification and Optimal Configuration of the Processors and the Required Connectivity for the Data Reconstruction in the Event Filter". Hence, a FPI fellow is requested.
The project will allow all physicists, senior and students, to learn about Grid technologies and get familiar with new tools for data processing and interactive analysis during the Data Challenge 3. In this way, the physicists will be able to analyse the data in optimal conditions.
The IFAE has trained several generations of system managers, who have been in charge of the computing infrastructure of the IFAE. Some of them, like Marc Rodríguez, are doing doctorate studies on computing engineering and topics on computational grids.
The IFAE participates in the Postgraduate Program of the "Universidad Autónoma de Barcelona", and A. Pacheco is professor of the Program since the beginning in 1996.

B) Training ability of the members of the UAM group signing this project.

In the last three years of the ATLAS and LCG-ES projects, J. Del Peso has supervised a PhD thesis, presented in December 2003, and is supervising another one that will be presented in the end of 2005. He has also supervised research works of 4 graduate students of the "Collaboration Fellowships of the UAM". Two members of the UAM group, J. Del Peso and M. Peez, are professors at UAM and give regularly graduate lectures in the Physics Studies. A third member, Dr C. Glasman, teach a postgraduate course on

“Experimental Particle Physics” within the Postgraduate Program of the Theoretical Physics Department of the UAM.