

DOCUMENTO DE BASES PARA UN PLAN COMÚN IBÉRICO DE INFRAESTRUCTURA PARA COMPUTACIÓN DISTRIBUIDA

ÍNDICE

- 1.- Introducción**
- 2.- Breve descripción de las capacidades de Portugal y España en e-Ciencia**
 - 2.1 Situación en España**
 - 2.2 Situación en Portugal**
- 3.- Estado de la organización de las redes Grid de Portugal y España**
 - 3.1 La organización de la red Grid en Portugal**
 - 3.2 La organización de la red Grid en España**
- 4.- Colaboraciones entre España y Portugal en materia de e-ciencia.**
- 5.- Criterios generales para la colaboración en Ibergrid**
 - 5.1 Infraestructura común de computación Grid**
 - 5.2 Comunicaciones**
 - 5.3 Aplicaciones**
 - 5.4 Supercomputación**
 - 5.5 Formación e Información**
 - 5.6 Movilidad de investigadores**
- 6.- Hoja de ruta**
- 7.- Organización**
- 8.- Presupuestos**

1.- Introducción

El 19 de noviembre de 2005 el Ministro de Ciencia, Tecnología y Enseñanza Superior de Portugal y la Ministra de Educación y Ciencia de España firmaron un Memorando de Entendimiento [1] bajo el Acuerdo de Cooperación Científica y Tecnológica de ambos países rubricado el 8 de noviembre de 2003.

El citado Memorando tenía como objetivo principal “desarrollar un programa de cooperación en materia de computación distribuida Grid y de supercomputación, destinado a facilitar la cooperación y acceso común de las dos Partes en el área Grid y áreas relacionadas”, contemplando las siguientes formas de cooperación:

- a) El acceso común a recursos y la coordinación de actividades
- b) El intercambio de información
- c) La movilidad de los investigadores portugueses y españoles en el marco Ibérico
- d) La participación de los investigadores en los programas de formación respectivos, incluyendo los programas de doctorado
- e) Proyectos de investigación y desarrollo comunes

Para el seguimiento institucional del Memorando se designaron a la Agencia para la Sociedad del Conocimiento (UMIC) de Portugal y a las Direcciones Generales de Investigación y de Política Tecnológica por parte de España. En cuanto al seguimiento científico se creó una Comisión Científica con el mandato de reunirse una vez al año.

En su reunión del 2 de octubre de 2006 en Madrid la Comisión Científica creó un grupo de trabajo con el mandato de elaborar un Plan Común de Infraestructura Ibérica de Computación Distribuida, en adelante el PLAN, el cual se reunió el 12 de noviembre de 2006 en Braga para perfilar los criterios generales para elaborar dicho plan, acordando además organizar, alternando entre Portugal y España, la Conferencia Ibérica Ibergrid con periodicidad anual, decidiendo para la primera vez la sede de Santiago de Compostela y las fechas de 14 a 16 de mayo de 2007.

En la reunión del 12 de noviembre, considerando las incertidumbres en los elementos de infraestructura Grid, se acordó plantear el PLAN como una hoja de ruta que se irá concretando a medida que se consoliden los diferentes temas.

Entre los criterios a considerar en el PLAN, además de la obvia conexión europea, debe darse especial atención a la colaboración con Latinoamérica, el mediterráneo y África, en consonancia con la política exterior de Portugal y España.

En la primera Conferencia Ibergrid deberá presentarse el PLAN objeto de este documento, previa aprobación del mismo por la Comisión Mixta.

2.- Breve descripción de las capacidades de Portugal y España en e-Ciencia

2.1 Situación en España

Los grupos españoles más representativos por su excelencia y la calidad del trabajo realizado en los proyectos internacionales son: Puerto de Información de Científica (PIC), Universidad Politécnica de Valencia (UPV), Instituto de Física Corpuscular del CSIC-UV (IFIC), Universidad de Zaragoza (BIFI), Centro de Supercomputación de Galicia (CESGA), Universidad Complutense de Madrid (UCM), Instituto de Física de Altas Energías de la UAB (IFAE), Instituto de Física de Cantabria del CSIC-UC (IFCA), Universidad de Barcelona (UB), Universidad de Santiago de Compostela (USC), Universidad Autónoma de Madrid (UAM), RedIRIS/Red.es y CIEMAT.

Estos grupos han participado en proyectos europeos ya finalizados como EU DataGrid, LHC Computing Grid en su fase I, EU CrossGrid o EGEE.

En cuanto a los proyectos europeos en marcha en los cuales participan grupos españoles cabe citar: EGEE II, EELA, EUMEDGRID, LCG en su fase II, DEISA, int.eu.grid, además de la participación en el foro e-IRG.

A lo largo de los últimos años, se ha ido generando una potente infraestructura sobre la base del middleware empleado en los proyectos de la UE, potenciada por la Red Temática de Middleware coordinada por la UCM, y de una creciente capacidad del hardware puesto a disposición de los usuarios en los centros de recursos, tanto de Clusteres en arquitectura Grid, como de Centros de Supercomputación.

En el área de aplicaciones, los grupos españoles han venido trabajando en diferentes campos, estableciéndose Organizaciones Virtuales (Virtual Organizations, VO). Destacan las de Biomedicina (Proteómica, Física Médica e Imagen Médica), Ingeniería (Fotónica, Materiales), Energía (Fusión, Fisión, Energías Renovables), Ciencias (Física de Altas Energías, Cosmología, Astronomía de Rayos Gamma, Medio Ambiente y Química).

En el ámbito de los recursos, deben destacarse los centros de supercomputación estatales, autonómicos y universitarios que proporcionan servicios de computación de altas prestaciones. Entre ellos, destaca el Centro Nacional de Supercomputación de Barcelona (BSC) que ocupa el primer lugar en Europa y una posición muy destacada en el mundo en capacidad de cálculo. Otros centros importantes son: CESGA, CESC, CIEMAT, CICA, INM, EPPE, CEPBA, etc. Recientemente se ha creado por el MEC la red nacional de supercomputación, formada inicialmente por el CBS, CESVIMA (UPM, CIEMAT), Universidades de Cantabria, Zaragoza, Valencia, Málaga y el Instituto de Astrofísica de Canarias.

Finalmente, debe destacarse la labor de RedIRIS dentro de la e-Ciencia española. Desde 1988 en que comienza la iniciativa de la red académica y de investigación, ha venido prestando un servicio en continua mejora, que ha permitido conectar varios centenares de instituciones de investigación entre sí y con el mundo exterior. Esta labor, junto con la de las Comunidades Autónomas que han creado sus redes regionales (Cataluña,

Galicia, Andalucía, Valencia, País Vasco y Madrid) o están en camino de crearlas, como es el caso de Extremadura y otras regiones, constituyen uno de los elementos clave para explicar el indudable avance científico y tecnológico que ha experimentado el país en los últimos años.

2.2 Situación en Portugal

El laboratorio de Instrumentación y Física Experimental de Partículas (LIP) es la institución portuguesa más involucrada en la infraestructura Grid.

LIP ha sido participante en los proyectos europeos CrossGrid y EGEE-1 y ha contribuido también al proyecto Data Grid.

Actualmente es el único participante en los proyectos EGEE-II, int.eu.grid y EELA de la Unión Europea y es miembro del proyecto WLCG (World Wide LHC Computing Grid) del CERN, la mayor infraestructura Grid del planeta orientada a la e-Ciencia.

LIP es operador de clusters Grid en preproducción y producción integrados en los proyectos antes mencionados y está también involucrado activamente en el EGEE ROC (Regional Operation Centre) de la federación suroeste de EGEE, que incluye a Portugal y España. Asimismo, LIP opera servicios Grid básicos de los proyectos EGEE e int.eu.grid y es la Autoridad Portuguesa de Certificación para Computación Grid (LIP.CA).

Existen, además en Portugal varias instituciones que participan o están interesadas en Grid:

- La Universidad de Aveiro, a través de IEETA, está participando en la infraestructura de preproducción de EGEE, como un centro de recursos, y está ensayando componentes de middleware gLite relacionados con los requisitos de las aplicaciones biomédicas.
- La Universidad de Lusiada en Famalicão opera un cluster de producción integrado en la infraestructura EGEE que soporta aplicaciones en biomedicina
- La Universidad de Minho en Braga y la Universidad de Porto están finalizando la implantación en la infraestructura de producción de EGEE.
- Un consorcio compuesto por el Laboratorio Portugués de Ingeniería Civil (LNEC), la Universidad de Porto (FEUP) y la Universidad de Minho en Guimarães tienen la intención de integrar sus clusters en una infraestructura Grid privada.
- La Universidad de Minho en Braga y el LNEC están interesados en explorar las funcionalidades de paralelismo e interactividad ofrecidas por el proyecto int.eu.grid.
- La Universidad de Coimbra participa en la red de excelencia COREGRID.

- La iniciativa nacional de Portugal en Grid (INGRID) está animando a otras muchas organizaciones del país a incorporarse a la Grid.

En cuanto a aplicaciones, la física de altas energías es claramente el área de mayor interés del país en Grid, siendo el LIP el líder de la implantación y operación del centro Tier2 que aporta recursos al proyecto GWLCG. Este centro soportará las Organizaciones Virtuales de los experimentos ATLAS y CMS. Otras áreas de interés incluyen las aplicaciones médicas, y el procesado de imagen (IEETA en Aveiro, Universidad de Porto, Universidad de Minho), simulaciones marítimas (LNEC), fusión nuclear (IST) y protección civil y simulación de incendios (Universidad de Minho y Universidad de Porto).

En el marco de la iniciativa INGRID se están identificando otras áreas y aplicaciones de interés para Portugal.

También debe resaltarse el trabajo que desarrollan la Universidad de Coimbra y FCT/UNL en el desarrollo de middleware y servicios.

La red nacional Portuguesa (FCCN), dentro del marco de INGRID, está encargada de suministrar servicios avanzados de red para computación Grid, operando una infraestructura de red dedicada, que interconectará los centros de computación Grid del país.

3.- Estado de la organización de las redes Grid de Portugal y España

El PLAN tiene su base en la experiencia previa de los dos países a través de su participación en proyectos de e-Ciencia, en los marcos nacionales respectivos y en el marco de la UE. Este esfuerzo se ha completado con la articulación, promovida por los Gobiernos respectivos, de las redes Grid nacionales para impulsar esta tecnología, con objeto de aprovechar de forma óptima las oportunidades del Programa Marco Europeo.

3.1 La organización de la red Grid en Portugal

El Gobierno lanzó en abril de 2006 la iniciativa Nacional de la Grid Portuguesa (INGRID) para desarrollar y soportar las infraestructuras de computación distribuida con el fin de compartir los recursos del país, resolver problemas complejos de cálculo científico y asegurar el desarrollo de competencias estratégicas a nivel nacional, que permitan la expansión de la computación Grid en el país (<http://www.gridcomputing.pt>).

La iniciativa INGRID es coordinada por FCT (Fundación para la Ciencia y la Tecnología) en cooperación con UMIC (Agencia portuguesa para la sociedad del conocimiento), perteneciente al Ministerio de Ciencia, Tecnología y Educación Superior.

La primera convocatoria para soportar el desarrollo de aplicaciones y middleware se lanzó a finales de 2006. Durante el año 2007 están previstas importantes inversiones para la red de comunicaciones e infraestructuras Grid. Estas inversiones se dedicarán a construir un núcleo inicial de infraestructura basada en la experiencia de los centros de

recursos ya existentes y con experiencia en el despliegue y operación de recursos computacionales y centros Grid. También en este contexto, el NREN portugués desplegará una arquitectura de red dedicada exclusivamente a soportar la integración de los centros de computación Grid.

La arquitectura y estructura de la infraestructura Grid que debe resultar de la integración de los diferentes centros está aún pendiente de definición.

3.2 La organización de la red Grid en España

Tras una amplia participación de grupos españoles de varias universidades y Organismos Públicos de Investigación (OPIs) en proyectos del Programa Marco de la UE y del Plan Nacional de I+D, a instancias del MEC, se elaboró en el año 2004 el “Libro Blanco de la e-Ciencia en España”, en el que se incluyó una propuesta para organizar e impulsar las tecnologías Grid y la supercomputación.

A lo anterior, también contribuyó notablemente IRISGrid, una red temática desarrollada por la comunidad investigadora, e impulsada por RedIRIS desde el 2002. IRISGrid utiliza la tecnología de Globus Toolkit y cuenta con la participación de 40 grupos de investigación.

Recientemente, el MEC, en base a la recomendación del Libro Blanco, está promoviendo la creación de la Red Española de e-Ciencia, a través de una Acción Especial, para la que se dispone ya de un documento básico consensuado por todos los grupos que actualmente trabajan en e-Ciencia, con vocación de promover nuevos grupos de investigación que se adhieran a la iniciativa.

La organización de esta red cubre todos los elementos necesarios para el progreso de la e-Ciencia: infraestructura, comunicaciones, middleware, aplicaciones y elementos de gestión general (calidad, formación, información, etc.)

4.- Colaboraciones entre España y Portugal en materia de e-ciencia.

España y Portugal tienen un amplio historial de colaboración en materia de e-Ciencia dentro del contexto bilateral y europeo. A continuación se describen brevemente los proyectos e iniciativas más significativas en que se ha tenido colaboración.

EGEE (Enabling Grid for EScienE)

El proyecto Enabling Grids for E-science (EGEE) está financiado por la Unión Europea dentro del 6º Programa Marco y tiene como objetivo principal desplegar una infraestructura de servicios Grid que esté disponible 24 horas al día para tareas de investigación científica.

La estructura de este proyecto es de dos fases, de dos años de duración cada una. La primera fase finalizó el 31 de Marzo de 2006, y en ella se lograron cumplir dos importantes objetivos. En primer lugar, se consiguió desplegar una infraestructura Grid a gran escala que conecta centros en más de 30 países de Europa y otras partes del mundo y que está operada, monitorizada y mantenida de forma profesional, al mismo

tiempo que crece continuamente. Por otro lado, se fusionaron las distribuciones de middleware de LCG y gLite, hecho que proporcionó una base sólida para la futura evolución del middleware EGEE.

La segunda fase de EGEE comenzó el 1 de Abril de 2006 y tiene como objetivo principal el extender y consolidar la infraestructura construida durante la primera fase, así como también establecer enlaces con las iniciativas Grid regionales, nacionales y temáticas. De este modo, se enfatiza la interoperabilidad con otros Grids que puedan existir en otras partes del mundo.

Esta segunda fase del proyecto incluye a más de 90 socios en 32 países, que están organizados en 13 federaciones. Una de estas federaciones, la llamada del “Suroeste de Europa”, está formada por España y Portugal y en la actualidad contiene 14 centros de recursos.

Dentro del proyecto EGEE los centros españoles y portugueses participan en las actividades de operación de la infraestructura (SA1), soporte a las aplicaciones (NA4) y certificación e integración de middleware (SA3) así como también en las de formación y divulgación (NA2 y NA3).

Cada una de las 13 federaciones de EGEE tiene un Centro de Operaciones Regional (ROC) encargado de coordinar las tareas de operaciones como soporte a usuarios y a centros de recursos, gestión de incidencias de seguridad, monitorización y contabilidad. En estos momentos, los 14 centros de recursos de la federación España-Portugal proporcionan a la infraestructura EGEE un total de más de 1000 CPUs y alrededor de 300 TB de almacenamiento. El Centro de Operaciones ROC de la federación España-Portugal está distribuido entre centros de ambos países.

WLCG

El “Worldwide LHC Computing Grid” es un proyecto del CERN cuyo objetivo es el despliegue de la infraestructura de computación para procesar los datos experimentales del LHC.

La comunidad de física de partículas, y en particular los cuatro experimentos del futuro colisionador LHC del CERN, es el mayor usuario de la infraestructura EGEE en estos momentos ya que genera una carga superior a los 20000 trabajos diarios, y varios cientos de TB de datos al año. Desde el inicio del proyecto, EGEE formó una alianza estratégica con el proyecto de computación para el LHC, WLCG. Esta alianza proporcionó una gran cantidad de centros de recursos a EGEE, y al mismo tiempo benefició a LCG al permitirle hacer uso de toda la infraestructura de operaciones desarrollada por EGEE para dar soporte a usuarios y a centros de recursos, monitorización, contabilidad, gestión de incidencias de seguridad, etc.

Debido a la enorme cantidad de datos que generarán los experimentos de LHC, y a la relación extremadamente pequeña señal/ruido de los mismos, la infraestructura de computación para la gestión de datos se implementará organizando los centros de recursos en un modelo de capas. En la primera capa (Tier-0) estará el CERN, como centro productor de los datos. A continuación, unos pocos centros con gran capacidad y fiabilidad, llamados Tier-1, se usarán para el almacenamiento a largo plazo y

reprocesamiento masivo de los datos. Finalmente, un gran número de centros más pequeños, llamados Tier-2, dedicarán sus recursos fundamentalmente a la producción de datos simulados y al análisis de los usuarios finales.

El centro Tier-1 español está en el PIC (Barcelona) y, como tal, está asociado y provee servicios a los Tier-2 de España y Portugal. En España existen tres Tier-2 federados, formados por siete centros de recursos, y en Portugal existe un centro federal Tier-2 distribuido en dos centros. Según el modelo de computación de cada experimento, el PIC actuará por un lado como repositorio a largo plazo de los datos simulados que se produzcan en los Tier-2 españoles y portugueses, así como de centro distribuidor de aquella parte de los datos reales que se vayan a analizar en los Tier-2. Durante los últimos meses todos estos flujos de datos se están ensayando sobre la infraestructura de producción, con el fin de comprobar que todo está listo para el inicio de la toma de datos a finales de este año. El hecho de que la parte de la jerarquía de capas LCG que cuelga del PIC esté contenida dentro de la federación España-Portugal de EGEE tiene múltiples ventajas a nivel operativo, puesto que permite beneficiarse de la estructura organizativa desarrollada por EGEE durante los últimos dos años.

EELA (E-infraestructura shared between Europe and Latin America)

El proyecto EELA, iniciado en 2006 y con dos años de duración, dentro del 6º Programa Marco, pretende establecer y dar soporte a una colaboración sostenible entre las iniciativas de infraestructura Grid de Europa y América latina, impulsando el desarrollo de esta nueva tecnología en los países latinoamericanos. Este proyecto se apoya y beneficia de la madurez del proyecto ALICE y de la red CLARA, sin los cuales no sería factible.

El proyecto está coordinado por el CIEMAT y cuenta con la colaboración del LIP en Portugal, participando asimismo las siguientes instituciones: INFN (Italia), UFRJ (Brasil), CNEN (Brasil), CEDERJ (Brasil), UFSM (Chile), CEADEN (Cuba), UNAM (México), ULA (Venezuela), UNLP (Argentina), CERN, CLARA, CSIC (España), UPV (España), PIC (España), NRENs (Retina, Cudi, RNP), RedIris (España).

EELA cuenta con una potente infraestructura basada en la arquitectura de EGEE con 10 centros de recursos que aportan más de 500 CPUs, sobre los cuales se han implementado aplicaciones en las áreas de educación, biomedicina, física de altas energías y clima.

El proyecto ha cubierto satisfactoriamente su primer año demostrando una gran capacidad para implantar la infraestructura y las aplicaciones y se espera que, tras su finalización en 2007, tenga continuidad con EELA II dentro del 7º Programa Marco. Ha tenido, por otra parte una amplia difusión a través de Congresos y Seminarios en Europa y Latinoamérica.

CORE Grid

El Core Grid, red de excelencia (NoE), pretende reforzar y alcanzar la excelencia en el área de Grid. Para alcanzar este objetivo, la red reúne una masa crítica de investigadores

expertos (119 investigadores y 165 estudiantes de doctorado) de 42 instituciones, que han construido un ambicioso programa común de actividades. Este programa está estructurado en torno a 6 áreas de investigación complementarias que se han seleccionado por su importancia estratégica, sus retos científicos y la reconocida experiencia europea en ellas, y ello con objeto de desarrollar la próxima generación de middleware Grid:

- Gestión del conocimiento y de la información
- Modelos de programación
- Arquitectura de sistemas
- Información Grid, recursos y servicios de monitorización del flujo de trabajo
- Gestión de recursos y planificación
- Sistemas Grid, herramientas y entorno

La red, financiada por la UE por 4 años, es operada como un laboratorio Europeo de Investigación (conocido como CoreGrid Research Laboratory), teniendo 6 institutos relacionados con las áreas identificadas en el programa conjunto de actividades.

En este proyecto participan la Universidad de Coimbra y la Universidad Politécnica de Cataluña.

TORGA.net

El proyecto TORGA.net permitió crear una red de comunicaciones avanzadas de última generación que interconecta los campus de la Universidad de Vigo (Vigo, Ourense y Pontevedra) con los de La Universidad de Uminho (Braga e Guimaraes) mediante una línea de fibra óptica transfronteriza de una velocidad de 155 Mbps, al tiempo que los diferentes campus se interconectaron a 1 Gbps. A través de la Red de Ciencia y Tecnología de Galicia (RECETGA) gestionada por el CESGA, se interconecta también con el resto de las Universidades y Centros Tecnológicos gallegos y portugueses.

Además de la infraestructura de comunicaciones, el proyecto ha desplegado una red de 7 salas colaborativas de tecnología AccessGrid que ha estrechado los lazos de comunicación entre las comunidades de Galicia y norte de Portugal.

En esta infraestructura, se realizaron numerosas actividades como cursos de formación y postgrado conjuntos a distancia, participación remota en congresos científicos con base en EEUU..., se difundieron más de treinta eventos regionales y nacionales (Jornadas de e-Ciencia, HPCN, Jornadas Técnicas de RedIris...), se publicaron numerosos artículos en medios técnicos, así como otras actividades de promoción tales como visitas a las dependencias del CESGA, jornadas de puertas abiertas o incluso un encuentro infantil transfronterizo en la Semana de la Ciencia.

Asimismo, el CESGA desarrolló con la colaboración de las Universidades de Vigo y do Minho, una plataforma de computación Grid que permite realizar cálculos científicos de forma remota en ordenadores distribuidos por la red Torga.net.

Ambos países colaboraron activamente tanto en la implantación de la infraestructura (cada uno en su lado) como en el desarrollo de las actividades.

CYTED Grid

El proyecto CYTED Grid propone trabajar en la creación de una infraestructura tecnológica y humana entre diversos grupos iberoamericanos dedicados al campo de la informática y con experiencia en el área de aplicaciones científicas potencialmente usuarias de la tecnología Grid.

El objetivo principal de este proyecto es construir una red Grid como soporte de diferentes tipos de aplicaciones, analizando hasta qué punto las soluciones disponibles actualmente son válidas, y proponiendo desarrollos nuevos que proporcionen la funcionalidad necesaria. Al final del proyecto se dispondrá de un conjunto significativo de aplicaciones piloto en temas de Biocomputación, Ciencias de la Salud y Medio Ambiente, que podrán ser utilizadas como demostradores de la tecnología Grid para su difusión en otros ámbitos. En él participan 5 instituciones españolas (UCM, UAB, UPV, UDC y CESGA), una portuguesa (UP) y hasta un total de 23 instituciones provenientes de países latinoamericanos (Brasil, Venezuela, Uruguay, Cuba, Argentina, Paraguay, Colombia, México, Perú, Chile y Ecuador).

CROSSGRID

El proyecto europeo **CrossGrid** se inició en el año 2002 con el objetivo de desarrollar nuevos componentes para integrar en un entorno Grid aplicaciones interactivas con elevadas necesidades de cálculo y de uso de datos. En este proyecto se puso a punto una infraestructura para el desarrollo y prueba de los nuevos servicios y herramientas de programación, así como de aplicaciones interactivas, con un entorno sencillo y amigable para el usuario. Uno de los objetivos alcanzados fue asegurar la completa interoperabilidad con el middleware del proyecto EDG (European Data Grid) para beneficiarse de sus resultados y experiencia y extender dicho entorno básico a 11 países, incluyendo España y Portugal. Esta coordinación hizo también posible la contribución eficaz de muchos de los participantes en CrossGrid a la iniciativa EGEE.

España y Portugal jugaron un papel esencial en este proyecto: los centros del CSIC (IFCA, IFIC) coordinaron la instalación del entorno de pruebas, en estrecho contacto con LIP en Portugal. El equipo de la UAB desarrolló el “CrossGridBroker”, una extensión del utilizado en EDG que permite la ejecución paralela de aplicaciones usando el protocolo MPI, y la interactividad de diferentes servicios y herramientas. El equipo de la USC, en colaboración con el CESGA, desarrolló nuevas herramientas de predicción de rendimiento e implementó una aplicación de contaminación atmosférica.

Esta infraestructura Grid, desplegada bajo coordinación ibérica, compartía recursos de 16 centros europeos, ofreciendo un entorno ideal de ensayo al incluir centros universitarios con capacidades relativamente pequeñas junto a otros con recursos del orden de más de 100 procesadores.

Tras tres años de trabajo, el consorcio CrossGrid presentó el Grid Interactivo desarrollado en la European Grid Conference en Amsterdam, permitiendo la ejecución de aplicaciones interactivas en medicina, medio ambiente y física. Estas aplicaciones se ejecutan desde un entorno sencillo y amigable que integra gráficos avanzados, pero permitiendo utilizar en modo paralelo los potentes recursos de distribuidos de la infraestructura Grid.

INTERACTIVE EUROPEAN GRID

El proyecto “Interactive European Grid (int.eu.grid)” ha arrancado en mayo de 2006 con el objetivo de proporcionar una e-Infraestructura europea común basada en potentes clusters conectados a través de redes avanzadas, para soportar el trabajo de colaboración diario de los investigadores con aplicaciones interactivas muy potentes. El proyecto garantiza la interoperabilidad con EGEE pero extiende su funcionalidad al uso de aplicaciones interactivas con mayor prioridad, y a la ejecución en paralelo mediante el protocolo MPI.

La participación de España y Portugal es muy relevante: el proyecto está coordinado por IFCA-CSIC y por LIP, participando la UAB, cómo responsable de la arquitectura, CESGA, a cargo de la infraestructura y BIFI, cómo encargado de la aplicación de Fusión. En el proyecto los grupos españoles están implementando además aplicaciones cómo el estudio de aneurismas cerebrales, el análisis del fondo cósmico de microondas en astrofísica, o el modelado integral de embalses en el área de medio ambiente.

La infraestructura del Interactive European Grid incluye recursos en 12 centros europeos de 7 países, entre los que están Portugal y España: clusters en los que se integran entre 10 y 180 procesadores con diferentes arquitecturas, y concedes de altas prestaciones cómo Infiniband o Myrinet. En total más de 300 cpus están disponibles para el uso interactivo.

Tras 9 meses de trabajo se puede decir que el proyecto está apoyando la investigación en Física, Medicina y Medio Ambiente, ofreciendo un soporte completo para la ejecución de aplicaciones paralelas basadas en MPI, desde un interfase amigable y transparente, el “Migrating Desktop”, que integra una potente visualización.

COMUNICACIONES

Las redes de comunicaciones son una infraestructura esencial para el desarrollo de la e-Ciencia. Las redes basadas en fibra óptica oscura pueden ofrecer diferentes soluciones tecnológicas sobre los mismos enlaces de telecomunicaciones. Esto representa un esfuerzo tecnológico y presupuestario muy importante, pero permite ofrece nuevas posibilidades de servicios avanzados, que son esenciales para el desarrollo de servicios para la e-Ciencia.

Por eso, como fruto de las estrechas relaciones de colaboración entre España y Portugal se acordó en noviembre de 2005 y posteriormente se ratificó en octubre de 2006, realizar la unión directa de las redes electrónicas de investigación y de enseñanza entre España y Portugal. Por ello, se constituyó una comisión técnica mixta para fijar los puntos de conexión de ambas redes y estudiar los servicios comunes para ambas comunidades.

La comisión mixta recomienda la adopción de una estrategia de integración de las infraestructuras ibéricas de Redes de Investigación y Educación con vistas a la preparación de la transición de redes de topología desconectada en una topología coherente de redes integradas.

La comisión mixta reconoce que tal integración es consecuencia lógica de las conexiones de fibra oscura en curso entre España y Portugal, en las fronteras de Alentejo-Extremadura y Miño-Galicia.

La comisión mixta recomienda que sobre las redes integradas de los dos países se reconozcan mutuamente amplios derechos de paso, potenciando de este modo las redundancias que permite la nueva topología.



Los enlaces de fibra óptica de Portugal hacia España estarán localizados en Valença (lado Portugués), en la frontera con Galicia, y en Elvas (lado Portugués) en la frontera con Extremadura.

La fibra óptica de Valença estará instalada y probada a mediados de junio de 2007. En cuanto a la fibra óptica de Elvas se prevé que esté disponible a finales de octubre de 2007.

Extremadura lleva muy avanzado el proyecto de despliegue de fibra y pronto estará en condiciones de dar servicio dentro de la comunidad autónoma y podrá conectar con la red portuguesa cerca de Badajoz.

Galicia ha desarrollado un plan que permitirá desplegar, en el momento que se considere más adecuado, fibra entre el punto de conexión con la red portuguesa y el punto de conexión con RedIRIS en Santiago de Compostela.

RedIRIS ha iniciado los trabajos de preparación del proyecto de despliegue de fibra desde Extremadura y Galicia al nodo nacional en Madrid que permite la conexión de esta red de fibra con el resto de la red académica española y la red europea GÉANT. El proyecto está previsto que se pueda aprobar y empezar a ejecutar entre diciembre de 2007 y enero de 2008, completándose la ejecución en 2008.

Merece también citarse la red Torga.net, desarrollada por el CESGA con las Universidades de Vigo y do Minho y el Centro de Computación Gráfica de Guimaraes, que permitió un incremento notable en la colaboración Galicia-Norte de Portugal mediante la creación de una red de salas de tecnología Access Grid y sirvió de plataforma de computación científica compartida por investigadores portugueses y

españoles. La existencia de esta red permitió la participación de la Universidad de Uminho en la experiencia Operta Oberta.



5.- Criterios generales para la colaboración en Ibergrid

Ibergrid debe ser una marca que simbolice una cooperación amplia y abierta entre España y Portugal en materia de e-Ciencia, así como una alianza estratégica para defender en común los intereses de ambos países en la UE, particularmente los relacionados con la cooperación con Latinoamérica, Área Mediterránea y África para crear una Comunidad Grid en el ámbito de la colaboración Ibérica.

A continuación se describen brevemente los elementos que se consideran más sustanciales para la colaboración.

5.1 Infraestructura común de computación Grid

Dada la situación de las infraestructuras Grid, particularmente de la dificultad para apostar por un middleware y unos estándares definitivos, es conveniente dejar abiertas las decisiones sobre las arquitecturas a compartir, definiendo provisionalmente aquellas que, por su implantación actual en España y Portugal, así como en la UE, se consideran más adecuadas por ambas partes.

Debido a la experiencia acumulada en los proyectos de computación Grid antes descritos y entendiendo como infraestructura de computación Grid un sistema que ante todo debe ser utilizable, se hace aconsejable partir de infraestructuras ya existentes y que han nacido con vocación de apertura e integración con iniciativas de carácter científico, al tiempo que han comprometido recursos para articular procedimientos operativos para el mantenimiento de la propia infraestructura. Por tanto la infraestructura común de computación entre los países de España y Portugal nacería como la continuación coordinada de los diferentes proyectos de infraestructura existentes, particularmente EGEE, aunque lógicamente evolucionaría con el paso del tiempo hacia otras arquitecturas. Es destacable la posible evolución del middleware a utilizar por los centros, dado que sin duda se fortalecerían las colaboraciones hispano-

lusas en este ámbito si existiesen mecanismos para potenciar su intervención en la generación del mismo.

La infraestructura común estaría basada en la existencia de:

- Red de comunicaciones que interconecte los recursos.
- Centros de recursos (RC): Los que aportan recursos computacionales al Grid. Para ello, los diferentes clusters de los grupos de investigación y de las diferentes instituciones establecen acuerdos con las organizaciones virtuales para permitir la utilización de dichos recursos en las condiciones que se determinen.
- Centro de operación Ibérico (COI). Este centro englobaría las funciones proporcionadas por el ROC (regional Operation Centre) y aquellas del CIC (Core Infrastructure Centre) necesarias para adaptarse a las circunstancias particulares en el ámbito ibérico. Al igual que está estructurado actualmente en la federación Sudoeste dentro del proyecto EGEE, es de carácter distribuido. En él estarán definidas las funciones de:
 - Certificación de nuevos centros de recursos para ser incorporados al Grid
 - Soporte a los centros de recursos, tanto previamente como posteriormente a su incorporación.
 - Coordinación de los aspectos de seguridad y la respuesta ante incidentes.
 - Formación específica sobre la infraestructura y de cómo mantenerla.
 - Monitorización de los diferentes servicios operativos.
 - Contabilidad de la utilización de los diferentes recursos por parte de los usuarios y Organizaciones Virtuales.
 - Soporte a la creación de Organizaciones Virtuales.

Es interesante destacar que en el modelo propuesto, el Grid no es por sí mismo el que dispone de una ingente cantidad de recursos computacionales, a disposición de los usuarios, sino más bien, la estructura que permite que los recursos de los que ya disponen las diferentes instituciones puedan ser combinados y utilizables de una manera más racional.

En este modelo, un usuario incluido en una Organización Virtual podrá utilizar los recursos a los que su VO tenga acceso, los cuales a su vez habrán sido obtenidos mediante acuerdos con los Centros de Recursos. En el ámbito de la infraestructura Ibergrid, además de las VO que ya existen de carácter internacional y que también deben soportarse, es de esperar que surjan nuevas VO específicas.

5.2 Comunicaciones

Los planes actuales de conexión entre los países a través de Extremadura y Galicia y las tecnologías asociadas se consideran de gran interés para poder llevar a cabo la iniciativa Ibergrid, por lo que es conveniente que se active en lo posible su ejecución.

Se espera que Portugal complete la instalación de fibra óptica hasta los puntos de conexión con España a mediados de junio de 2007 en la frontera con Galicia y en octubre de 2007 en la frontera con Extremadura. España deberá completar el Proyecto de conexión y ejecutarlo lo antes posible en 2008.

La coordinación entre RedIris (Red.es) y RCTS (FCCN) y con los organismos políticos que sustentan a estas organizaciones, deberá mantener, e incluso reforzarse, para

garantizar la actualización de las tecnologías de comunicaciones en base al desarrollo tecnológico y a las necesidades que vayan surgiendo en la operación. En particular, deberá analizarse la posibilidad de disponer de una red común dedicada a Grid; Portugal está estudiando esta posibilidad y, si España decidiera lo mismo, sería muy conveniente coordinar estas iniciativas.

La Comisión Técnica Hispano-Lusa para tratar los temas técnicos de interconexión entre ambos países, creada en 2006, bien como tal o en un contexto más amplio junto a la Comisión Grid, es un elemento a conservar y potenciar para conseguir el objetivo anterior.

5.3 Aplicaciones

Deberá prestarse una atención especial a la creación de Organizaciones Virtuales (VO) y al impulso de aplicaciones científicas comunes en los campos de interés de ambos países.

Inicialmente conviene seleccionar un número reducido de VO conjuntas, entre las que pueden tener interés para ambos países y sobre las que se viene trabajando en estos. A este respecto, se identificaron en la reunión de Braga varias que pueden ser un buen punto de partida.

En cada una de las VO deberán identificarse los centros de recursos con su aportación y sus funciones, los usuarios científicos que se benefician de ellas, así como los objetivos y los resultados esperados, sobre los que será conveniente realizar una evaluación continua durante el periodo de explotación.

En las Aplicaciones se procurará hacer intervenir a todas las instituciones que puedan aportar calidad y utilidad, particularmente a las instituciones de servicios públicos y empresas que puedan tener interés.

Para conseguir los objetivos buscados en el área de Aplicaciones se considera conveniente articular un elemento específico dentro de la organización de la colaboración, cuya función se centre en las mismas.

5.4 Supercomputación

Se establecerán colaboraciones entre las redes de Supercomputación de ambos países, permitiendo el cálculo científico de usuarios de un país en los recursos del otro, dentro de los límites y condiciones que se delimiten en acuerdos específicos de uso de recursos.

Se promoverán proyectos en colaboración que permitan este acceso a través de convocatorias específicas en áreas de interés estratégico para España y Portugal. Dado que el estado actual de la tecnología no permite una integración clara entre el Grid y la Supercomputación, se propone inicialmente considerar a la Supercomputación como un elemento más disponible dentro de la infraestructura Ibergrid, aunque con mecanismos y procedimientos de acceso diferentes. A medida que la infraestructura Ibergrid vaya madurando, se deberá estudiar la problemática asociada a esa integración y se

desarrollarán métodos que permitiesen la utilización de la Supercomputación de manera coordinada con otros recursos de tipo Grid, con procedimientos comunes.

5.5 Formación e Información

Se establecerán planes coordinados de formación no académica en el campo de la e-Ciencia, abiertos a estudiantes y profesionales de cualquiera de los países.

Asimismo, se fomentará la coordinación de los programas de formación académica de Doctorado y Postdoctoral, permitiendo la colaboración mutua a nivel de profesorado y estudiantes.

Se establecerá una Web para Ibergrid conectada con las Web de cada uno de las redes nacionales y se cooperará en el objetivo de transmitir a la sociedad el interés de las tecnologías Grid y de supercomputación.

5.6 Movilidad de investigadores

A través de los mecanismos de los Planes Nacionales de Investigación y de cooperación bilateral entre España y Portugal, sería conveniente arbitrar mecanismos que favorezcan, de forma específica para Ibergrid, la movilidad de los investigadores dentro del ámbito ibérico. Asimismo, se procurará coadyuvar en este objetivo mediante relaciones bilaterales entre Centros de Investigación y Universidades de ambos países.

6.- Hoja de ruta

Para el cumplimiento progresivo de los criterios generales anteriores, es necesario ir realizando una serie de acciones en sucesivas fases, unas de carácter organizativo, otras de infraestructura de Grid y supercomputación y otras de gestión general.

A continuación se indican las que se consideran más importantes:

FASE 1

- Conexión de las redes telemáticas a través de Galicia (Valençã) y Extremadura (Badajoz).
- Aprobación e implantación de la organización necesaria para llevar a cabo el Plan.
- Arranque con infraestructuras existentes (con la última versión de glite).
- Identificación de centros que constituyen el COI de entre los que actualmente participan en proyectos de infraestructura Grid.
- Inventario de recursos disponibles (incluido personal de soporte y atención a usuarios)
- Establecimiento de condiciones y procedimientos de acceso a los recursos identificados.
- Utilización de EuGridPMA (pkIrisgrid + LIP CA).
- Mecanismos comunes de seguridad y respuesta ante incidentes.
- Selección de 4 Organizaciones Virtuales iniciales de entre las siguientes áreas:
 - control de riesgos ambientales
 - protección civil y respuesta a emergencias,
 - meteorología,

- detección remota por satélite,
- sismología
- investigación biomédica
- Identificación de aplicaciones a potenciar en las áreas anteriores.
- Definición de características de los centros de recursos.
- Coordinación de cursos de formación
- Estudio del modelo de sostenibilidad económica de la infraestructura Ibergrid, más allá del periodo de apoyo del Programa Marco.

FASE 2:

- Incorporación de centros de recursos adicionales.
- Incorporación de nuevas Organizaciones Virtuales y aplicaciones relacionadas
- Testbed de pruebas.
- Detección de necesidades de middleware de las nuevas aplicaciones y VOs y desarrollo del mismo.
- Establecimiento de mecanismos para la evolución e incorporación de nuevo middleware.
- Definir y poner en práctica un sistema de colaboración en materia de supercomputación, con acceso a los recursos de ambos países.
- Coordinación de cursos de formación

FASE 3:

- Incorporación de Supercomputación.
- Definición de mecanismos de acceso específicos en Ibergrid adicionales a los existentes para cualquier comunidad internacional a:
 - MareNostrum
 - RedEspañola de Supercomputación
 - FinisTerra
 - Otros Centros de Supercomputación
- Programas de formación específica.
- Integración de Middleware.
- Arbitrar mecanismos específicos de movilidad de personal

Se estima un tiempo de 6 a 9 meses para la Fase 1 y de 12 meses para la Fase 2. La Fase 3 empezaría dentro de un año y medio aproximadamente. Desde el principio se deberán mantener las actividades que den continuidad a las iniciativas que se vayan poniendo en marcha, evitando la natural y frecuente tendencia a la degradación que suelen tener estos temas. En este sentido, es mejor avanzar a un ritmo menor; pero mantenido y estabilizado.

7.- Organización

De acuerdo con el Memorando de Entendimiento, existe una Comisión Científica de seguimiento constituida por autoridades de los Ministerios de Educación y Ciencia de Portugal y España, que se reunirá al menos una vez al año.

Sin embargo, para articular adecuadamente la colaboración, es conveniente crear un órgano de trabajo y gestión que defina actividades y objetivos y controle su realización, dando cuenta de sus actuaciones a la Comisión Científica.

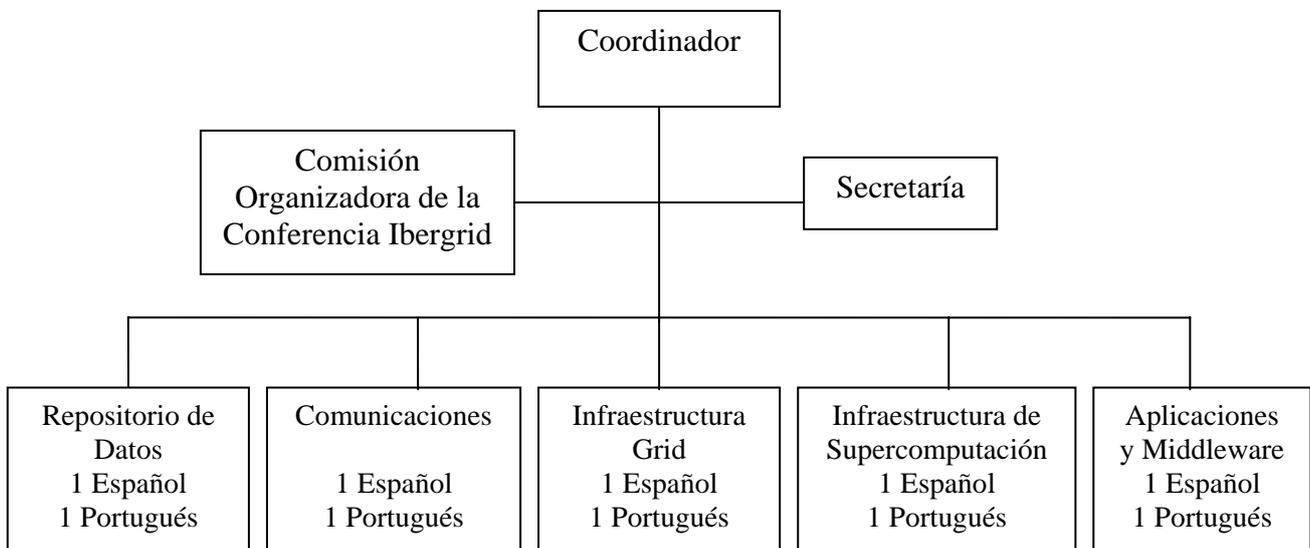
Para ello, se propone un Comité de Coordinación, presidido alternativamente por España y Portugal, en el que participen expertos de ambos países en las diferentes facetas de la e-Ciencia. Este Comité estará asistido por una secretaria que aportará el país encargado de la coordinación y se reunirá trimestralmente, alternativamente, en España y Portugal, para analizar conjuntamente la marcha de las actividades de la colaboración.

Esta organización, que deberá integrarse en la estructura de las redes nacionales Grid respectivas, deberá favorecer las relaciones entre los grupos de investigación de los dos países, fomentando la infraestructura Grid común, las comunicaciones y las aplicaciones.

La organización de la Conferencia Ibergrid anual se inscribirá en el Comité de Coordinación, sintonizándose la responsabilidad de dicha organización con la presidencia del Comité.

Inicialmente, Portugal presidirá el Comité de Coordinación al tiempo que se encargará de la Segunda Conferencia Ibergrid a celebrar en el año 2008.

La estructura del Comité de Coordinación podría ser la siguiente:



8.- Presupuestos

La financiación de las actividades de cada país en Ibergrid se realizará por las instituciones, o el gobierno, de cada uno de los países, según los casos.

Dentro de los presupuestos de Acciones Especiales de los Planes Nacionales de cada país, se atenderán los gastos de desplazamiento, formación, información y gestión general para la coordinación de Ibergrid, así como los gastos de su secretaría.

Una primera estimación de estos gastos generales de coordinación se cifra en 30k€ año por país, con los que se financiarían los desplazamientos de coordinación, la formación, la información común y la promoción de aplicaciones conjuntas.

Para llevar a cabo las actividades específicas de mejora de infraestructuras, desarrollo de middleware, desarrollo de aplicaciones, mejora de las comunicaciones, etc., los grupos de investigación obtendrán los fondos de los Planes Nacionales de I+D respectivos o del Programa Marco Europeo, en el régimen de competencia de estos. Sin embargo, en el caso particular de las comunicaciones, se considera conveniente una coordinación específica en materia presupuestaria a nivel de los Ministerios de Educación y Ciencia de ambos países, para realizar de forma coordinada las mejoras de tecnología, velocidad, etc. que se consideren necesarias.

No obstante, sería deseable una mayor coordinación de las convocatorias de I+D relacionadas con Grid de los dos países, y, aún más, si fuera posible, la creación de mecanismos conjuntos de subvención para los proyectos de e-Ciencia comunes.