



INICIATIVA NACIONAL. GRID



29 de Abril de 2006

www.gridcomputing.pt

LigarPortugal

FCT
Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

 **UMIC**
Agência para
a Sociedade
do Conhecimento
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

INICIATIVA NACIONAL GRID

Partilha de recursos computacionais de elevado desempenho

www.gridcomputing.pt

SUMÁRIO

COMPUTAÇÃO GRID: O que é e para que serve?

O que é a Computação GRID?

A Computação GRID é uma tecnologia de computação distribuída que nasceu da ideia de replicar para o processamento computacional a filosofia e os princípios de funcionamento da World Wide Web (WWW) para a disponibilização de informação à escala mundial. Na verdade, a WWW tornou possível disponibilizar de forma distribuída e a partir de milhares de computadores usuais uma quantidade gigantesca de informação que seria impossível disponibilizar com um pequeno número de computadores, mesmo que estes tivessem capacidades extremamente elevadas. De forma análoga, a Computação GRID consegue disponibilizar elevadas capacidades computacionais à custa de distribuir tarefas de processamento por vários computadores de forma coordenada e eficiente.

Esta forma de computação tornou-se viável devido ao grande avanço das tecnologias de ligação em rede. Na verdade, para tornar claro o elevado potencial da Computação GRID basta notar que a capacidade de aumento das velocidades de transmissão de dados em redes de fibra óptica nos últimos anos tem sido aproximadamente o dobro do aumento das velocidades dos microprocessadores. Portanto, tornaram-se possíveis crescimentos mais significativos, e também mais eficientes em termos de custos, das capacidades computacionais através de explorar melhor as capacidades acrescidas das ligações em rede do que as capacidades acrescidas dos microprocessadores, aplicando-se uma observação semelhante para o armazenamento de grandes quantidades de dados.

A Computação GRID disponibiliza às mais diversas comunidades de utilizadores, de forma eficiente, transparente e segura, recursos computacionais partilhados de processamento e armazenamento, inclusivamente os organizados em “clusters” de processadores interligados em GRID sobre redes de alto débito, tornando possível o acesso a elevadas capacidades de processamento a quem se ligue à GRID através de banda larga.

A Computação GRID surge assim como uma alternativa aos super-computadores para o processamento de tarefas computacionais que requerem elevado desempenho, permitindo o avanço da ciência, da engenharia e de diversas áreas de interesse social ou económico que necessitam de

acesso a enormes quantidades de recursos computacionais, e tornando possível rentabilizar recursos que frequentemente se encontram sub-utilizados por serem usados aquém das suas capacidades ou até por ficarem totalmente disponíveis durante certos períodos.

Para que serve a Computação GRID?

A Computação GRID permite revolucionar a forma de trabalhar em ciência e tecnologia em várias áreas. Numa era em que o trabalho científico de qualquer domínio do conhecimento produz cada vez mais informação, como por exemplo a descodificação do genoma humano, a cura de doenças infecto-contagiosas, imunológicas ou cancerígenas, ou o estudo do comportamento da matéria e da energia, em experiências com aceleradores de partículas cujos dados são analisados e partilhados por investigadores dispersos por todo o mundo, a Computação GRID vem oferecer respostas às enormes exigências ao nível da capacidade computacional e de armazenamento que o processamento de grandes quantidades de dados coloca. Também na simulação computacional de sistemas de grande dimensão, como por exemplo em meteorologia, oceanografia, genómica e proteómica, a Computação GRID permite a conjugação de recursos computacionais distribuídos para a resolução de problemas que requerem elevado desempenho computacional.

Por exemplo, no âmbito do projecto *CrossGrid*, cuja infra-estrutura tem os seus sistemas centrais de gestão instalados em Portugal sob a responsabilidade operacional do LIP – Laboratório de Instrumentação e Física Experimental de Partículas, são consideradas aplicações de controlo, previsão e simulação de cheias em bacias fluviais, aplicações de modelação de poluição atmosférica e outros serviços de apoio à meteorologia, bem como uma aplicação de apoio à cirurgia vascular. Nesta última aplicação, partindo das imagens de TAC ou Ressonância Magnética de um paciente, o cirurgião pode simular o efeito de um *by-pass* e visualizar o fluxo sanguíneo, decidindo por tentativas sucessivas a melhor solução para a intervenção. Esta aplicação permite encurtar para poucas horas o tempo de intervenção que medeia entre a recolha da imagem e a intervenção cirúrgica otimizada pela simulação prévia.

A Computação GRID também abre perspectivas à criação de novos produtos e serviços, afectando a forma como as empresas podem conduzir os seus negócios. Grandes empresas financeiras recorrem a Computação GRID para efectuar análises de risco cada vez mais detalhadas que, por sua vez, levam à criação de novos serviços financeiros para os clientes. Outro exemplo é o das empresas farmacêuticas a quem esta tecnologia facilita o acesso a uma capacidade computacional cada vez maior que pode permitir a descoberta de novos compostos e analisar em mais detalhe os riscos associados aos medicamentos antes destes serem comercializados. Entre as empresas pioneiras na adopção de sistemas de Computação GRID encontram-se muitos bancos. Relatórios internacionais recentes estimam uma utilização crescente de sistemas de Computação GRID pelas empresas do sector financeiro.

OBJECTIVOS DA INICIATIVA NACIONAL GRID

- **Reforçar as competências e capacidades nacionais em Computação GRID devido à sua especial importância estratégica**

A *Iniciativa Nacional GRID* foi prevista no programa *Ligar Portugal* com o objectivo de promover em Portugal o desenvolvimento da computação GRID, assegurando o reforço das competências e capacidades nacionais necessárias para aproveitar as oportunidades científicas e económicas abertas pela evolução recente deste tipo de computação distribuída que permite a partilha de recursos na resolução de problemas complexos com necessidade de processamento intensivo de dados e, por isso, assume uma especial importância estratégica.

- **Prosseguir a integração de Portugal na rede internacional de Computação GRID**

A iniciativa baseia-se em trabalho anterior realizado no país em vários grupos de investigação, inclusivamente no âmbito de projectos internacionais, insere-se em iniciativas europeias em curso e acompanha iniciativas nacionais congéneres que decorrem em vários países. Portanto, um dos objectivos essenciais desta iniciativa é integrar plenamente Portugal na rede internacional de Computação GRID.

- **Melhorar as condições para as actividades científicas e para aplicações de interesse económico e social que envolvem computações complexas ou com elevadas quantidades de dados**

Em particular, pretende-se melhorar as condições para as actividades científicas que envolvem o tratamento computacional de elevadas quantidades de dados, como é o caso das grandes experiências com aceleradores de partículas, ou que envolvem a modelação e simulação de sistemas de grande dimensão nas mais variadas áreas do conhecimento.

Pretende-se também melhorar o desempenho de sistemas de suporte a actividades que têm impacto directo na qualidade de vida dos cidadãos, como por exemplo a previsão meteorológica, a modelação climática, a oceanografia, a geofísica e sismologia, a observação da terra com fins económicos ou de prevenção de riscos, a biodiversidade, a imageologia médica, o estudo de doenças complexas como as doenças infecto-contagiosas, imunológicas e diversas formas de cancro.

- **Reforçar a multidisciplinaridade e a colaboração entre as comunidades de investigadores e utilizadores de meios computacionais de elevado desempenho**

Reforçar a multidisciplinaridade e a colaboração entre as comunidades de investigadores e utilizadores de meios computacionais de elevado desempenho através da disponibilização aberta de infra-estruturas computacionais GRID para a resolução de problemas complexos de áreas de aplicação muito diversas.

- **Reforçar as condições para as empresas encontrarem em Portugal instituições científicas e recursos humanos com conhecimentos e experiência de Computação GRID**

O desenvolvimento de capacidades e competências em Computação GRID tem também objectivos estratégicos de ordem económica ao reforçar as condições para que empresas possam encontrar em Portugal instituições científicas e recursos humanos com conhecimentos e experiência deste tipo de computação, nomeadamente para gestão e exploração de bases de dados de grandes dimensões, na análise de informação geográfica, no processamento de vídeo, nos sistemas de projecto assistido por computador para aplicações de grande dimensão, nos projectos de desenho de componentes de microelectrónica e na análise de riscos de investimentos financeiros em bancos, entre outros.

ENTIDADE EXECUTORA

A entidade responsável pela execução da *Iniciativa Nacional GRID* é a Fundação para a Ciência e a Tecnologia (FCT). A Agência para a Sociedade do Conhecimento (UMIC) terá funções de observação e acompanhamento da iniciativa.

LINHAS DE ACÇÃO

As principais linhas de acção da *Iniciativa Nacional GRID* são:

- i) **Infra-estruturas:** Apoiar a implementação e o funcionamento de infra-estruturas GRID bem como a participação portuguesa em projectos GRID internacionais, nomeadamente ao nível Europeu;
- ii) **Conectividade:** Expandir e reforçar a conectividade de alta velocidade entre nós de GRID em Portugal e entre estes e outras infra-estruturas internacionais;
- iii) **Projectos de I&D:** Promover a I&D em Computação GRID e nas suas aplicações, com ênfase no desenvolvimento de software de suporte à infra-estrutura GRID e à optimização da utilização de recursos computacionais partilhados, e no desenvolvimento de projectos de aplicação da Computação GRID às várias áreas científicas;
- iv) **Projectos de demonstração e aplicação:** Apoiar projectos de demonstração e aplicação da tecnologia GRID, em áreas como, entre outras, a meteorologia, a oceanografia, a geofísica e sismologia, a física de altas energias;
- v) **Formação:** Apoiar a formação avançada de recursos humanos em Computação GRID, no país e no estrangeiro, nomeadamente no âmbito de projectos de Computação GRID e através de bolsas de doutoramentos e pós-doutoramento. Estimular a oferta de cursos superiores em computação distribuída que incluam o ensino de Computação GRID.
- vi) **Avaliação internacional:** Promover a avaliação internacional de acções e projectos;

- vii) **Observação, acompanhamento e disseminação de informação:** Assegurar a observação e o acompanhamento da iniciativa, e a disseminação de informação e conhecimento na comunidade de Computação GRID.

Os apoios acima descritos, que poderão ser utilizados tanto por instituições de ensino superior e de investigação bem como por empresas, deverão, em termos gerais, propiciar:

- a) A disponibilização de infra-estruturas de Computação GRID para a resolução de problemas complexos de diversas comunidades de utilizadores, promovendo-se assim a multidisciplinaridade e a proximidade entre as comunidades de investigadores e utilizadores;
- b) A participação de investigadores portugueses em consórcios e projectos internacionais de I&D de Computação GRID, o que por sua vez deverá permitir que “clusters” de processadores portugueses sejam parte integrante de infra-estruturas pan-europeias de Computação GRID.

ACÇÕES

A *Iniciativa Nacional GRID* envolve as acções seguintes:

- i) **Infra-estruturas:** No âmbito da concretização de projectos de Re-equipamento Científico, será promovida a implementação de infra-estruturas de Computação GRID e Computação Avançada e o seu funcionamento como infra-estruturas de uso comum, bem como o encorajamento à sua integração em infra-estruturas nacionais.
- ii) **Conectividade:** A Fundação para Computação Científica Nacional (FCCN) assegurará, através da RCTS – Rede Ciência Tecnologia e Sociedade, a conectividade de alta velocidade entre nós de GRID em Portugal e entre estes e outras infra-estruturas internacionais, promovendo a ligação das instalações GRID ao *backbone* desta rede e a ligação em fibra proprietária à rede GÉANT em Espanha, através da Galiza e da Extremadura.
- iii) **Projectos de I&D:**
 - a) Abertura de concursos para projectos na área da Computação GRID que contemplem o apoio ao financiamento de investigadores, bolseiros em formação no âmbito do projecto e assistência técnica à utilização de Computação GRID;
 - b) Abertura de concursos para projectos de I&D na área da Computação GRID em consórcio entre empresas e instituições científicas com objectivos de inovação na indústria ou serviços.
- iv) **Projectos de demonstração e aplicação:** Abertura de concursos para projectos de demonstração e aplicação da Computação GRID em áreas como meteorologia, oceanografia, geofísica e sismologia, física de altas energias e outras, com fortes ligações internacionais e a que possam concorrer consórcios que envolvam Laboratórios de Estado e Laboratórios Associados ou outras Unidades de I&D.

v) **Formação:**

- a) Abertura de concursos para bolsas de doutoramento e de pós-doutoramento em Computação GRID e nas suas aplicações, no país e no estrangeiro;
- b) Abertura de concursos para bolsas de estágios de engenheiros em instituições de Computação GRID de destaque internacional, à semelhança dos actuais estágios no CERN;
- c) Apoio à abertura de cursos superiores em computação distribuída que incluam o ensino de Computação GRID.

vi) **Avaliação internacional:** Constituição de uma Comissão Científica e Técnica Internacional para acompanhamento regular e avaliação das acções e projectos da *Iniciativa Nacional GRID*, com reuniões pelo menos anuais;

vii) **Observação, acompanhamento e disseminação de informação:**

- a) Observação e acompanhamento da *Iniciativa Nacional GRID* pela UMIC promovendo o cumprimento dos objectivos e das acções previstas e manutenção de um sítio na Internet, em www.gridcomputing.pt, onde será disponibilizada informação relevante sobre a iniciativa e sobre os projectos nacionais e internacionais em curso na área;
- b) Organização de reuniões anuais sobre os desenvolvimentos da Computação GRID, em âmbito nacional e internacional, para fomentar a partilha de experiências e o progresso da *Iniciativa Nacional GRID*.

1. Introdução

A Computação GRID é uma tecnologia de computação distribuída que permite partilhar e coordenar recursos para a resolução de problemas complexos, seja porque são de grande escala ou porque envolvem o processamento intensivo de dados, conjugando a contribuição de recursos computacionais dispersos na execução de tarefas computacionais comuns. Esta tecnologia permite a organizações geograficamente dispersas funcionarem como organizações virtuais que partilham recursos computacionais de armazenamento e aplicações, assim como o aproveitamento de recursos computacionais usuais e/ou pouco utilizados na resolução de problemas computacionais que requerem alto desempenho.

A Computação GRID permite revolucionar a forma de trabalhar em ciência e tecnologia em várias áreas. Numa era em que o trabalho científico de qualquer domínio do conhecimento produz cada vez mais informação, como por exemplo a descodificação do genoma humano, a cura de doenças terminais, ou o estudo do comportamento da matéria e da energia, em experiências com aceleradores de partículas cujos dados são analisados e partilhados por investigadores dispersos por todo o mundo, a Computação GRID vem oferecer respostas às enormes exigências ao nível da capacidade computacional e de armazenamento que o processamento de grandes quantidades de dados coloca. Também na simulação computacional de sistemas de grande dimensão, como por exemplo em meteorologia, oceanografia, genómica, a Computação GRID permite a conjugação de recursos computacionais distribuídos para a resolução de problemas que requerem elevado desempenho computacional.

A Computação GRID também abre novas perspectivas à criação de novos produtos e serviços, afectando a forma como as empresas podem conduzir os seus negócios. Grandes empresas financeiras recorrem a Computação GRID para efectuar análises de risco cada vez mais detalhadas que, por sua vez, levam à criação de novos serviços financeiros para os clientes. Outro exemplo é o das empresas farmacêuticas a quem esta tecnologia facilita o acesso a uma capacidade computacional cada vez maior que pode permitir a descoberta de novos compostos e analisar cada vez em mais detalhe os riscos associados aos medicamentos antes destes serem comercializados. Entre as empresas pioneiras na adopção de sistemas de Computação GRID encontram-se vários bancos (e.g., Bank of America, Bank of Montreal, BNP Paribas, Bowne, Citigroup, Genworth, HSBC, JPMorgan Chase, Markit, MassMutual, Nationwide, Royal Bank of Scotland, Societe Generale, TD Bank Financial Group, UBS, Wachovia and WestLB) e relatórios recentes estimam uma utilização crescente de sistemas de Computação GRID pelas empresas do sector financeiro.

A Computação GRID suporta de uma forma natural projectos multidisciplinares, tendo um enorme potencial para acelerar o ritmo de investigação e desenvolvimento em diferentes áreas científicas e tecnológicas e melhorar as práticas de partilha e acesso ao conhecimento, permitindo aumentar a qualidade do trabalho científico das instituições e a competitividade das empresas.

São elementos centrais da Computação GRID os “clusters” de processadores que, interligados de forma inteligente, oferecem enormes capacidades de processamento a baixo custo, e aplicações de

gestão de recursos partilhados que possibilitam o acesso a comunidades virtuais de utilizadores externos das mais diversas áreas científicas. Estes “clusters” estão normalmente espalhados, em termos geográficos, pelo que dependem da existência de redes de comunicação de alto débito para poderem trocar informação de forma eficiente e coordenar as suas tarefas de modo a serem utilizados como infra-estruturas unificadas de grande dimensão em tempo real.

A *Iniciativa Nacional GRID* foi prevista no programa Ligar Portugal com o objectivo de, a exemplo de outros países europeus, promover em Portugal o desenvolvimento da computação GRID e assegurar a constituição de competências e capacidades nacionais necessárias para aproveitar as oportunidades científicas e económicas abertas pela evolução deste tipo de computação distribuída. Os projectos a desenvolver devem ser concebidos de forma a aumentar a interligação das iniciativas portuguesas aos projectos internacionais que têm ultimamente proliferado nesta área, conforme se descreve no capítulo seguinte.

2. Projectos e Aplicações

2.1 Projectos Internacionais Emblemáticos

Existem vários projectos emblemáticos de Computação GRID ao nível Europeu suportados em infra-estruturas de grandes dimensões. O projecto *Enabling Grids for E-Science* (EGEE), e o seu sucessor *EGEE-II*, pretendem interligar vários projectos de investigação e desenvolvimento baseados em colaborações distribuídas suportadas pela Internet, possibilitando assim o acesso simplificado a uma infra-estrutura unificada composta pela união de várias infra-estruturas de países membros da União Europeia. Os aspectos inovadores deste projecto consistem na sua natureza global e no objectivo de agregar hardware, software e recursos humanos de um vasto conjunto de instituições. Unindo estes esforços criou-se uma infra-estrutura única com uma capacidade agregada de cerca de 20.000 processadores que potencia o aparecimento de novas aplicações e novos métodos de organização de recursos partilhados. O *EGEE* suporta comunidades diversificadas que vão desde a nanotecnologia à bio-informática.

O projecto *Diligent*, financiado também pela União Europeia, combina tecnologias de Computação GRID e bibliotecas digitais para criar um repositório de informação distribuído que facilite o acesso partilhado de organizações dinâmicas e virtuais a grandes quantidades de informação de forma segura e coordenada. O projecto pretende promover o acesso a esta informação a custos reduzidos, inclusivamente para pequenas instituições que de outra forma não teriam possibilidade de a obter. Estes projectos incluem o desenvolvimento de funções de procura, anotação e visualização de documentos e de instrumentos que facilitem, por exemplo, a organização de materiais de suporte à preparação de conferências, de relatórios de projectos e de cursos de formação avançada de recursos humanos.

O *WorldWide LHC Computing Grid Project* surge com a construção do *Large Hadron Collider* (LHC) no *European Organization for Nuclear Research* (CERN), um acelerador de partículas que será o maior instrumento científico do planeta. Quando terminado em 2007, produzirá cerca de 15 Petabytes

de informação por ano durante 15 anos (cerca de 10 biliões de páginas de texto), a qual será disponibilizada livremente à comunidade científica internacional para análise. O objectivo central do projecto *LHC Computing Grid Project* é a construção de uma infra-estrutura de armazenamento e análise que suporte a actividade de toda a comunidade científica que irá usufruir dos resultados das experiências a realizar. Estas experiências, que se focam na análise das propriedades das partículas mais elementares da matéria, pretendem recriar condições semelhantes às que se verificaram instantes depois do *Big-Bang* no intuito de se poder vir a saber mais sobre a origem do universo. Com base nas estimativas de 2004, serão necessários cerca de 100.000 processadores para analisar toda a informação que será produzida e compará-la com os resultados de simulações teóricas. A informação será espalhada numa GRID hierárquica por vários centros de computação instalados em todo o mundo. Presentemente, estão implementados mais de uma centena e meia destes centros que dispõem da capacidade agregada de mais de 10.000 processadores.

O projecto *E-Infrastructure shared between Europe and Latin América* (EELA), financiado pela Comissão Europeia, tem por objectivo estender o projecto *EGEE* à América Latina. Trata-se de um consórcio europeu e latino-americano que tem por objectivo a implementação de uma infra-estrutura piloto de Computação GRID transcontinental, explorando, entre outros, os desenvolvimentos do projecto *EGEE* e a rede *RedCLARA* para a interligação com as redes nacionais na América Latina. O sucesso desta iniciativa depende, em larga medida, da capacidade de interoperabilidade entre os vários sistemas e da largura de banda entre os dois continentes. O objectivo último deste projecto é contribuir para a sustentabilidade dos projectos de *e-science* na América do Sul através do estabelecimento de redes dedicadas à Computação GRID e de projectos-piloto em diversas áreas

Ao nível da infra-estrutura de conectividade destaca-se o contínuo desenvolvimento da rede de alto débito *GEANT* que interliga mais de 30 redes nacionais de investigação e ensino na Europa. Todas as regiões do mundo estão presentemente ligadas a esta rede que é a maior do seu tipo no planeta. A existência desta rede é fundamental e possibilita a partilha rápida de grandes quantidades de dados entre diferentes instituições, o que abre portas a um novo conjunto de aplicações. Por exemplo, esta rede pode fazer chegar a comunidades remotas consultas médicas pela Internet, permitindo acesso aos conhecimentos mais recentes nas áreas relevantes às patologias encontradas. As instituições portuguesas obtêm interligação à rede *GEANT* através da Rede Ciência, Tecnologia e Sociedade (RCTS), gerida pela Fundação para Computação Científica Nacional (FCCN). Actualmente, o projecto *GEANT2* procede à melhoria desta rede para um débito de múltiplos de 10Gbps no seu *backbone* central.

A IBM lançou o *World Community Grid*, que pretende apoiar causas de carácter filantrópico através da doação, para fins humanitários, do tempo inactivo de computadores à escala mundial. O primeiro projecto a beneficiar desta iniciativa, baseado em tecnologia GRID, foi o *Human Proteome Folding*, um projecto de investigação que visa determinar as estruturas tridimensionais moleculares das proteínas humanas, conhecimento que poderá ajudar na descoberta da cura e tratamento de diversas doenças, nomeadamente da doença de Alzheimer e da Esclerose Lateral Amiotrófica. Mais recentemente, e em parceria com o *Scripps Research Institute*, a IBM lançou, também sobre o *World*

Community Grid, o programa *FightAIDS@home* para investigação de cura da SIDA O *World Community Grid* conta hoje com cerca de 240.000 computadores, o que o torna um dos mais poderosos aglomerados GRID do mundo.

A SUN Microsystems lançou recentemente o projecto *SUNGRID* que permite aos utilizadores registarem-se no website da SUN desenvolvido para o efeito e submeterem, via Internet, processos que irão utilizar de forma transparente a capacidade de processamento livre da rede que suporta este projecto. Por cada hora de processamento num CPU a SUN cobra 1 dólar (US) sem a necessidade de contratos de longo prazo ou a compra de blocos de horas pré-pagos. A SUN Microsystems disponibiliza também capacidade de armazenamento para registar os resultados dos processamentos efectuados que são para serem transferidos para o cliente por banda larga. Todas as transacções são pagas via *PayPal*. O contacto físico directo com os sítios onde se encontram os processadores da *SUNGRID* não é permitido. A *SUNGRID* tem sido utilizada por empresas das áreas financeira, energia, entretenimento, educação, saúde e ciências da informação. Com este projecto, a SUN contribui de forma ímpar e pioneira para criar o primeiro grande mercado aberto de capacidade de processamento, tornando esta numa verdadeira “*commodity*”, embora disponível apenas para utilizadores registados dos Estados Unidos da América.

Ao nível das ferramentas destaca-se a aplicação *Globus toolkit* que foi produzida e é mantida pela aliança *Globus*, uma comunidade de organizações que desenvolvem tecnologias de suporte à Computação GRID, constituída principalmente por instituições de investigação europeias e norte-americanas. O *Globus toolkit* inclui um conjunto de serviços e software que suportam as actividades necessárias à partilha de recursos informáticos, nomeadamente a identificação, atribuição e monitorização dos mesmos, a transferência de dados, a autenticação e a segurança informática. Este conjunto de ferramentas constitui um dos softwares mais utilizados em todo o mundo no suporte à construção de aplicações GRID nas empresas e em centros de investigação.

Finalmente, destaca-se a infra-estrutura *PlanetLab* que consiste num conjunto de várias centenas de máquinas espalhadas por instituições académicas, governamentais e empresas em todos os continentes. Cada entidade participante neste consórcio fornece um conjunto relativamente pequeno de máquinas, e obtém acesso a todas as máquinas da rede *PlanetLab*. O acesso a esta rede permite que investigadores testem novas tecnologias distribuídas em escalas geográficas que se aproximam da diversidade actual da Internet. Periodicamente, realizam-se encontros para discutir os aspectos técnicos deste sistema e da investigação que sobre ele é conduzida.

2.2 Iniciativas Emblemáticas de Estados Membros da União Europeia

Além dos projectos transeuropeus, vários países da União Europeia têm vindo a desenvolver projectos de âmbito nacional na área da Computação GRID.

No Reino Unido, foi iniciado em 2000 um programa nacional em *e-science*, no âmbito do qual inúmeras actividades científicas se desenvolvem com base em colaborações distribuídas suportadas por uma rede de alto débito e Computação GRID. Este programa financiou projectos em diferentes

áreas do conhecimento, tais como astronomia, física das partículas, biologia, engenharia, ciências do ambiente, finanças, ciências da saúde, e informática. Entre os principais desenvolvimentos ao nível da infra-estrutura destaca-se o desenvolvimento de "middleware", isto é, de plataformas de programas de suporte a aplicações distribuídas, nomeadamente por actividade em consórcio com empresas.

Na Irlanda a *Higher Education Authority* criou uma infra-estrutura denominada *National Computational Grid for Ireland* que consiste na interligação de 17 "clusters" espalhados pelo país. As máquinas de acesso aos "clusters" e a rede de interligação são geridas a partir de um centro de operações localizado no *Trinity College*, em Dublin, que fornece serviços às comunidades de utilizadores como gestão de recursos, armazenamento de dados e repositório de software. Esta rede está também associada à rede *CosmoGrid* que se dedica à investigação de fenómenos naturais como tremores de terra, mudanças climáticas e explosões de supernovas.

Na Alemanha decorre desde Setembro 2005 o projecto *D-Grid* que envolve 5 parceiros na implementação de uma infra-estrutura GRID para suportar as actividades de *e-science* da comunidade académica alemã. Actualmente, destacam-se os seguintes projectos: *D-GRID Integration Project* visa aprovisionar a infra-estrutura central de Computação GRID – a *Kern-D-Grid* – através do desenvolvimento de modelos de partilha de recursos, operação de serviços de suporte e de "middleware"; *AstroGrid-D* que envolve a ligação de todos os institutos na área de investigação astronómica, incluindo os mais pequenos, para a partilha de recursos de informação, processamento de dados de observações astronómicas e software específico; *Collaborative Climate Community Data and Processing Grid* (C3-Grid) para investigação e desenvolvimento sobre a Terra nomeadamente nas vertentes de oceanos, atmosfera, biosfera, preservação do habitat humano e estudo das mudanças climáticas; o projecto *HEP Community Grid* que visa apoiar as experiências nas áreas de altas energias, nuclear, astro-física e física de partículas; *InGrid* que gere uma infra-estrutura GRID para aplicações em engenharia, nomeadamente de modelação, simulação e optimização em áreas como a dinâmica da interacção entre fluidos e estruturas sólidas e fluxos no sub-solo; e *TextGrid* que visa o desenvolvimento de ferramentas para processamento distribuído de texto e edição científica.

A República Checa concentra a sua actividade de Computação GRID no *METACentrum* do *CESNET Institute*, o qual estende a infra-estrutura académica de alto débito para suportar aplicações que requerem elevado poder computacional. Actualmente composto por três centros de super-computação, este centro detêm os seguintes projectos: *CoreGrid* que suporta I&D em seis áreas fundamentais para o desenvolvimento da Computação GRID: gestão de informação e conhecimento, modelos de programação, arquitectura de sistemas, monitorização de infra-estruturas GRID, calendarização no acesso a recursos partilhados, ambientes de desenvolvimento e ferramentas para sistemas GRID; *MediGrid* para a utilização distribuída de módulos de informação com aplicação na área da saúde; *GridLab* para o desenvolvimento de "middleware" para a gestão de recursos GRID que permita uma utilização transparente da infra-estrutura. A infra-estrutura unificada de Computação GRID da República Checa inclui cerca de 21 servidores para computação e 4 servidores de ficheiros,

espalhados por todo o país, com ligações de pelo menos 1 Gbps e utilizados por cerca de 40 centros de investigação.

Na Bélgica, a Iniciativa *BELNET Grid* promove a Computação GRID para a comunidade de investigação e serve como entidade de certificação para infra-estruturas GRID neste país. *BEGrid* é a infra-estrutura que resulta desta iniciativa, distribuída por seis locais. O “*middleware*” produzido é o utilizado no projecto *EGEE*, disponível a partir do CERN em versão *open-source*. A utilização da *BEGrid* é aberta a todas as instituições ligadas à *BELNET*, as quais podem contribuir com infra-estruturas e/ou recursos humanos. Está a ser concebido um esquema de compensação financeira para os utilizadores que apenas desejam utilizar a infra-estruturas alheias e, entretanto cada utilizador não deve utilizar mais do que 25% dos recursos disponíveis.

Na Itália, o Ministério da Educação, Universidade e Investigação lançou em 2001 um programa nacional de Computação GRID. Trata-se de um projecto de pendor fortemente interdisciplinar, cujos objectivos são estudar e utilizar ferramentas de software inovadoras para a gestão de recursos partilhados e demonstrar as potencialidades destas através de aplicações específicas nas áreas de geofísica, astronomia, biologia, genética e química computacional. Esta rede GRID interliga 6 centros de investigação que organizaram um projecto dividido em 14 tarefas: 9 visam o desenvolvimento da infra-estrutura GRID e as restantes visam a utilização desta infra-estrutura em projectos-piloto de demonstração das potencialidades da Computação GRID.

Nos países Escandinavos assiste-se a uma iniciativa conjunta denominada *NorduGrid*. Esta iniciativa, estabelecida em Maio de 2001, assenta na colaboração entre institutos académicos e de investigação nórdicos com o objectivo de construir uma infra-estrutura GRID capaz de dar resposta às necessidades das tarefas de investigação em curso nestes países.

Na Holanda foi estabelecida a organização *Grid Forum* que se dedica ao desenvolvimento e ao suporte de tecnologias GRID através da união de esforços e do estabelecimento de ligações entre peritos de diferentes instituições. Esta organização sem fins lucrativos é financiada por empresas, centros de investigação, universidades e organismos públicos. A sua principal tarefa é gerir a *DutchGrid* que é uma plataforma aberta para o desenvolvimento da Computação GRID na Holanda.

Em Espanha, a iniciativa nacional de Computação GRID, denominada *IRISGrid*, tem como objectivo coordenar, a nível académico e científico, os grupos de investigação interessados nestas tecnologias bem como criar uma infra-estrutura nacional que possa ser utilizada para o desenvolvimento de “*middleware*”.

A Grécia estabeleceu o grupo de trabalho *HellasGrid* que tem desenvolvido a infra-estrutura de Computação GRID neste país. A *HellasGrid* prevê interligar seis “*clusters*” de computadores, três em Atenas e outros três em Thessaloniki, Patras e Heraclion-Crete, com uma capacidade total de 768 processadores. Esta infra-estrutura estará aberta à utilização de centros de investigação e outras entidades. Até ao momento, a maior parte das utilizações são em física de altas energias e engenharia biomédica.

2.3 Projectos Internacionais com Participação Portuguesa

A participação portuguesa em projectos internacionais de Computação GRID começou em 2001, em resultado da participação portuguesa no CERN, o Laboratório Europeu de Física de Altas Energias. De facto, Portugal é membro, através do LIP – Laboratório de Instrumentação e Física Experimental de Partículas, de duas grandes colaborações científicas (*ATLAS* e *CMS*) que iniciarão a aquisição de dados no novo acelerador do CERN, o *Large Hadron Collider (LHC)*, em construção desde 2001 e que entrará em operação no Verão de 2007. Assim, o primeiro projecto de Computação GRID a que Portugal se associou foi o *Worldwide LHC Computing Grid*. A arquitectura de Computação GRID foi considerada pelo CERN a única capaz de recolher, armazenar, distribuir e tratar os 15 Petabytes (15 milhões de Gigabytes) de dados a produzir anualmente pelo *LHC*. Para o tratamento destes dados foi estimado ser necessária uma infra-estrutura de processamento equivalente a 50.000 processadores distribuídos por todos os continentes. A Portugal cabe desenvolver uma infra-estrutura nacional de segundo nível (Tier2) que deverá contar, em fins de 2009, com 1.000 processadores. A infra-estrutura global deverá operar e evoluir ininterruptamente até 2020.

A União Europeia considerou no 5º Programa Quadro de Investigação a importância estratégica deste projecto e decidiu apoiá-lo promovendo o Projecto *EU-DATAGRID*, de que Portugal foi parceiro. O projecto *DATAGRID*, terminado em 2004, visou partir à descoberta de uma nova geração de exploração científica que requer computação intensiva e a análise de gigantescas bases de dados, por uma comunidade de milhares de investigadores distribuídos mundialmente.

Dado o sucesso do projecto *DATAGRID*, muitos dos seus produtos (tecnologias, software, infra-estruturas, etc.) transitaram, no âmbito do 6º Programa Quadro, para um novo projecto de iniciativa europeia, o *EGEE – Enabling Grids for E-science* que foi concluído em 2006. Este projecto concentra-se nas seguintes tarefas principais: construção de uma rede GRID robusta e consistente, melhoria e manutenção contínua de *middleware*, atracção de novas comunidades de utilizadores, tanto investigadores como de empresas.

Em 1 de Abril de 2006 foi iniciado o projecto *EGEE-II* que corresponde a uma segunda fase do projecto anterior. Este projecto, que conta presentemente com a participação de mais de 30 países e mais de 90 instituições, visa instalar uma infra-estrutura de serviços GRID permanentemente operacional. Esta infra-estrutura conta presentemente com mais de 20.000 processadores activos, mais de 15 Petabytes de espaço em disco, e cerca de 200 bases de computação distribuídas globalmente.

Na transição do 5º para o 6º Programa Quadro Portugal participou no projecto *CrossGrid*. Este projecto, iniciado em 2002 e terminado em 2005, visou desenvolver, implementar e explorar novas componentes GRID para computação interactiva, implementando uma infra-estrutura de computação distribuída com acesso a recursos de computação avançada existentes na Europa. Os domínios de aplicação privilegiados pelo *CrossGrid* foram, entre outros, ciências da terra, apoio a cirurgia vascular, protecção ambiental, acompanhamento dos regimes de cheias fluviais. Os sistemas centrais de *CrossGrid* estão instalados em Portugal sob a responsabilidade do LIP e a infra-estrutura permanece

activa, mesmo depois da conclusão do projecto, nomeadamente a aplicação de apoio a cirurgia vascular.

O êxito do projecto *CrossGrid* levou a União Europeia a apoiar um novo de projecto de exploração e alargamento do seu âmbito aplicacional, o projecto *INT.EU.GRID – Interactive European Grid Project*, onde Portugal de novo assumiu responsabilidades. O seu objectivo é criar uma infra-estrutura de produção para aplicações interactivas de elevada complexidade capazes de se reflectirem no trabalho quotidiano dos investigadores: computação paralela distribuída interactiva (MPI) e armazenamento ao nível de Terabytes, desenvolvimento das interfaces de utilizador e de visualização, apoio a organizações virtuais em todas as vertentes (ambiente de colaboração, aplicações, ferramentas e intercomunicação entre operadores).

Em 1 de Março de 2006, Portugal associou-se ao projecto *EU-EELA – E-Infrastructure Shared Between Europe and Latin America*. Este projecto fará uso dos recursos disponíveis em Portugal, Espanha e Itália, já integrados em *EGEE*, disponibilizando-os à América Latina (Argentina, Brasil, Chile, Cuba, México, Peru e Venezuela) e potenciando o uso da *RedClara*. Será acompanhado de uma forte componente de formação e mobilidade.

Os projectos referidos acima são, na parte da colaboração portuguesa, da responsabilidade do LIP – Laboratório de Instrumentação e Física Experimental de Partículas.

Finalmente, Portugal está envolvido, através do Centro de Investigação em Informática e Tecnologias de Informação (CITI) da Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa, no *Asia-Link Programme – Curriculum Development for High-Performance and Grid Computing (2005-2007)*, uma iniciativa da União Europeia com o objectivo de promover a conexão em rede das instituições de ensino superior da Europa, Ásia do Sul e Sudeste e China e, através da *Interactive European Grid Mission*, potenciar aplicações avançadas, particularmente no domínio do ensino.

3. A Iniciativa Nacional GRID

OBJECTIVOS

- **Reforçar as competências e capacidades nacionais em Computação GRID devido à sua especial importância estratégica**

A *Iniciativa Nacional GRID* foi prevista no programa *Ligar Portugal* com o objectivo de promover em Portugal o desenvolvimento da computação GRID, assegurando o reforço das competências e capacidades nacionais necessárias para aproveitar as oportunidades científicas e económicas abertas pela evolução recente deste tipo de computação distribuída que permite a partilha de recursos na resolução de problemas complexos com necessidade de processamento intensivo de dados e, por isso, assume uma especial importância estratégica.

- **Prosseguir a integração de Portugal na rede internacional de Computação GRID**

A iniciativa baseia-se em trabalho anterior realizado no país em vários grupos de investigação, inclusivamente no âmbito de projectos internacionais, insere-se em iniciativas europeias em curso

e acompanha iniciativas nacionais congéneres que decorrem em vários países. Portanto, um dos objectivos essenciais desta iniciativa é integrar plenamente Portugal na rede internacional de Computação GRID.

- **Melhorar as condições para as actividades científicas e para aplicações de interesse económico e social que envolvem computações complexas ou com elevadas quantidades de dados**

Em particular, pretende-se melhorar as condições para as actividades científicas que envolvem o tratamento computacional de elevadas quantidades de dados, como é o caso das grandes experiências com aceleradores de partículas, ou que envolvem a modelação e simulação de sistemas de grande dimensão nas mais variadas áreas do conhecimento.

Pretende-se também melhorar o desempenho de sistemas de suporte a actividades que têm impacto directo na qualidade de vida dos cidadãos, como por exemplo a previsão meteorológica, a modelação climática, a oceanografia, a geofísica e sismologia, a observação da terra com fins económicos ou de prevenção de riscos, a biodiversidade, a imageologia médica, o estudo de doenças complexas como as doenças infecto-contagiosas, imunológicas e diversas formas de cancro.

- **Reforçar a multidisciplinaridade e a colaboração entre as comunidades de investigadores e utilizadores de meios computacionais de elevado desempenho**

Reforçar a multidisciplinaridade e a colaboração entre as comunidades de investigadores e utilizadores de meios computacionais de elevado desempenho através da disponibilização aberta de infra-estruturas computacionais GRID para a resolução de problemas complexos de áreas de aplicação muito diversas.

- **Reforçar as condições para as empresas encontrarem em Portugal instituições científicas e recursos humanos com conhecimentos e experiência de Computação GRID**

O desenvolvimento de capacidades e competências em Computação GRID tem também objectivos estratégicos de ordem económica ao reforçar as condições para que empresas possam encontrar em Portugal instituições científicas e recursos humanos com conhecimentos e experiência deste tipo de computação, nomeadamente para gestão e exploração de bases de dados de grandes dimensões, na análise de informação geográfica, no processamento de vídeo, nos sistemas de projecto assistido por computador para aplicações de grande dimensão, nos projectos de desenho de componentes de microelectrónica e na análise de riscos de investimentos financeiros em bancos, entre outros.

ENTIDADE EXECUTORA

A entidade responsável pela execução da *Iniciativa Nacional GRID* é a Fundação para a Ciência e a Tecnologia (FCT). A Agência para a Sociedade do Conhecimento (UMIC) terá funções de observação e acompanhamento da iniciativa.

LINHAS DE ACÇÃO

As principais linhas de acção da *Iniciativa Nacional GRID* são:

- i) **Infra-estruturas:** Apoiar a implementação e o funcionamento de infra-estruturas GRID bem como a participação portuguesa em projectos GRID internacionais, nomeadamente ao nível Europeu;
- ii) **Conectividade:** Expandir e reforçar a conectividade de alta velocidade entre nós de GRID em Portugal e entre estes e outras infra-estruturas internacionais;
- iii) **Projectos de I&D:** Promover a I&D em Computação GRID e nas suas aplicações, com ênfase no desenvolvimento de software de suporte à infra-estrutura GRID e à optimização da utilização de recursos computacionais partilhados, e no desenvolvimento de projectos de aplicação da Computação GRID às várias áreas científicas;
- iv) **Projectos de demonstração e aplicação:** Apoiar projectos de demonstração e aplicação da tecnologia GRID, em áreas como, entre outras, a meteorologia, a oceanografia, a geofísica e sismologia, a física de altas energias;
- v) **Formação:** Apoiar a formação avançada de recursos humanos em Computação GRID, no país e no estrangeiro, nomeadamente no âmbito de projectos de Computação GRID e através de bolsas de doutoramentos e pós-doutoramento. Estimular a oferta de cursos superiores em computação distribuída que incluam o ensino de Computação GRID.
- vi) **Avaliação internacional:** Promover a avaliação internacional de acções e projectos;
- vii) **Observação, acompanhamento e disseminação de informação:** Assegurar a observação e o acompanhamento da iniciativa, e a disseminação de informação e conhecimento na comunidade de Computação GRID.

Os apoios acima descritos, que poderão ser utilizados tanto por instituições de ensino superior e de investigação bem como por empresas, deverão, em termos gerais, propiciar:

- a) A disponibilização de infra-estruturas de Computação GRID para a resolução de problemas complexos de diversas comunidades de utilizadores, promovendo-se assim a multidisciplinaridade e a proximidade entre as comunidades de investigadores e utilizadores;
- b) A participação de investigadores portugueses em consórcios e projectos internacionais de I&D de Computação GRID, o que por sua vez deverá permitir que “clusters” de processadores portugueses sejam parte integrante de infra-estruturas pan-europeias de Computação GRID.

ACÇÕES

A *Iniciativa Nacional GRID* envolve as acções seguintes:

- i) **Infra-estruturas:** No âmbito da concretização de projectos de Re-equipamento Científico, será promovida a implementação de infra-estruturas de Computação GRID e Computação Avançada e

o seu funcionamento como infra-estruturas de uso comum, bem como o encorajamento à sua integração em infra-estruturas nacionais.

- ii) **Conectividade:** A FCCN assegurará, através da RCTS – Rede Ciência Tecnologia e Sociedade, a conectividade de alta velocidade entre nós de GRID em Portugal e entre estes e outras infra-estruturas internacionais, promovendo a ligação das instalações GRID ao *backbone* desta rede e a ligação em fibra proprietária à rede GÉANT em Espanha, através da Galiza e da Extremadura.
- iii) **Projectos de I&D:**
 - a) Abertura de concursos para projectos na área da Computação GRID que contemplem o apoio ao financiamento de investigadores, bolseiros em formação no âmbito do projecto e assistência técnica à utilização de Computação GRID;
 - b) Abertura de concursos para projectos de I&D na área da Computação GRID em consórcio entre empresas e instituições científicas com objectivos de inovação na indústria ou serviços.
- iv) **Projectos de demonstração e aplicação:** Abertura de concursos para projectos de demonstração e aplicação da Computação GRID em áreas como meteorologia, oceanografia, geofísica e sismologia, física de altas energias e outras, com fortes ligações internacionais e a que possam concorrer consórcios que envolvam Laboratórios de Estado e Laboratórios Associados ou outras Unidades de I&D.
- v) **Formação:**
 - a) Abertura de concursos para bolsas de doutoramento e de pós-doutoramento em Computação GRID e nas suas aplicações, no país e no estrangeiro;
 - b) Abertura de concursos para bolsas de estágios de engenheiros em instituições de Computação GRID de destaque internacional, à semelhança dos actuais estágios no CERN;
 - c) Apoio à abertura de cursos superiores em computação distribuída que incluam o ensino de Computação GRID.
- vi) **Avaliação internacional:** Constituição de uma Comissão Científica e Técnica Internacional para acompanhamento regular e avaliação das acções e projectos da *Iniciativa Nacional GRID*, com reuniões pelo menos anuais;
- vii) **Observação, acompanhamento e disseminação de informação:**
 - a) Observação e acompanhamento da *Iniciativa Nacional GRID* pela UMIC promovendo o cumprimento dos objectivos e das acções previstas e manutenção de um sítio na Internet, em www.gridcomputing.pt, onde será disponibilizada informação relevante sobre a iniciativa e sobre os projectos nacionais e internacionais em curso na área;
 - b) Organização de reuniões anuais sobre os desenvolvimentos da Computação GRID, em âmbito nacional e internacional, para fomentar a partilha de experiências e o progresso da *Iniciativa Nacional GRID*.

ANEXO

LEVANTAMENTO PRELIMINAR DE INFRA-ESTRUTURAS E ACTIVIDADES EM PORTUGAL DE COMPUTAÇÃO GRID E COMPUTAÇÃO AVANÇADA

O levantamento que se segue, relativo a infra-estruturas e actividades de Computação GRID em Portugal é um levantamento preliminar, a ser complementado e actualizado, que se baseia nas informações prestadas pelas instituições referidas.

A Universidade do Porto, através do seu Instituto de Recursos e Iniciativas Comuns (IRICUP), lançou em 2005 o projecto *GridUP* para interligar com tecnologia GRID três “clusters” de processadores adquiridos pelo projecto. O *Campus Grid* está operacional e a sua gestão reparte-se pela Faculdade de Engenharia (FEUP), pela Faculdade de Ciências (FCUP) e pelo IRICUP. Adicionalmente, a FEUP desenvolveu a infra-estrutura de computação *NxPy*, projecto que envolve 13 grupos das diversas áreas de ensino e investigação existentes na Faculdade com o objectivo de construir uma infra-estrutura de elevado desempenho, sólida, actualizada e para a qual exista uma equipa de apoio que disponibilize o suporte técnico necessário. A FEUP tem esta infra-estrutura em plena produção desde Janeiro de 2006, actualmente com um agrupamento de 66 processadores, para apoio a projectos em dinâmica dos fluidos, simulação em mecânica estrutural, *data mining*, computação paralela, multimédia e computação gráfica. Está previsto o alargamento a outras áreas da ciência e estão a ser desenvolvidos os mecanismos para interligar em GRID esta infra-estrutura com outros nós de computação, nomeadamente os dos projectos *GridUP* (três nós: FEUP, FCUP e IRICUP) e da Rede Nacional de Computação Avançada (quatro nós: FEUP-IDMEC, IST, LNEC e Universidade do Minho).

A Universidade do Minho desenvolve actividades de I&D relacionadas com *High Performace Computing* (HPC) em ciências (Física e Matemática), engenharia (Mecânica e Informática) e computação paralela e distribuída. Está em curso um projecto inter-disciplinar – *Services and Advanced Research Computing with HTC/HPC Clusters* (SeARCH) – que assenta num agrupamento de processadores Linux de 850 GFlops que se interligam em GRID aos restantes. Este projecto também inclui I&D em estratégias de gestão de carga, dados e objectos em nós partilhados, e aplica, avalia e integra os resultados com actividades de I&D que requerem computação de elevado débito e/ou desempenho com I/O intensivo, nomeadamente modelação e simulação de propriedades dos materiais, sistemas de navegação 3D, sistemas de pesquisa de informação e *data/web mining*.

Na Universidade de Coimbra, o Centro de Física Computacional (CFC) gere um agrupamento de processadores com 108 computadores, chamado *Centopeia*, que é utilizado para actividades de investigação que envolvem super-computação, em particular em Física, nas áreas das interacções fortes (modelos de hádrões, física nuclear e cromodinâmica quântica), física da matéria condensada, geofísica e astrofísica. Outras áreas de investigação apoiadas pela *Centopeia* incluem a bioquímica (simulação de dobragem de proteínas), matemática aplicada (algoritmos de optimização de Geometria Molecular) e engenharia (simulações de Monte-Carlo). O sistema computacional tem tido

utilizadores de outros departamentos da Universidade de Coimbra e de outras universidades portuguesas, entre as quais as universidades Técnica de Lisboa, de Évora, do Porto, do Minho e de Aveiro. Já foi também utilizado pela empresa Critical Software para teste de software. Dentro de alguns meses deverá entrar em funcionamento um novo agrupamento de 520 processadores, que será mantido aberto à utilização da comunidade científica portuguesa. O CFC organizou já várias reuniões sobre super-computação e cursos sobre o protocolo de paralelização *Message Passing Interface* (MPI), e tem mantido actividades com investigadores do Departamento de Engenharia Informática da Universidade de Coimbra com vista à participação em projectos de Computação GRID. Além disso, vai ser assinado um protocolo com o Laboratório de Instrumentação e Partículas (LIP) com vista à utilização da *Centopeia* no projecto *GRID-LHC*, relacionado com a participação portuguesa no tratamento de dados das experiências no novo acelerador do CERN.

No Instituto Superior Técnico (IST), o *Laboratory and Simulation of Energy and Fluids* (LASEF) participa no projecto *GRIDS* para investigação em mecânica dos fluidos utilizando redes computacionais de alto desempenho. O Centro de Física dos Plasmas (CFP) do IST, membro do Centro de Fusão Nuclear – Laboratório Associado, está empenhado na criação de um programa de *e-science* multidisciplinar apoiado em Computação GRID que deverá contemplar o desenvolvimento de aplicações para este novo paradigma de computação, programas de acesso às infra-estruturas partilhadas e a formação avançada de recursos humanos em computação avançada, Computação GRID e visualização científica, nomeadamente através da formação de segundo e terceiro ciclos em *e-science*. Aproveitando o *know-how* do Centro, estas actividades apoiam-se ainda em colaborações com projectos internacionais e no acesso a recursos computacionais de grande escala. O CFP participa no consórcio *Osiris*, que visa o desenvolvimento de software para simulação numérica de plasmas. Este consórcio, liderado pela UCLA e pelo IST, envolve mais sete instituições de I&D nos Estados Unidos e na Europa. O CFP também iniciou uma colaboração com o Centro de Supercomputação de Barcelona para *benchmarking* do agrupamento de processadores *MareNostrum*, e está envolvido na implementação do agrupamento de processadores do IST que, numa primeira fase, vai ter cerca de 256 processadores.

O INESC-ID efectua investigação em aspectos de Computação GRID, nomeadamente em autenticação, segurança, escalonamento, métodos para paralelização e distribuição de carga e aplicações, em particular para processamento de fala, processamento de imagem e biologia computacional. O INESC-ID participa no projecto *BioGrid*, desenvolvido em consórcio com o Instituto Gulbenkian Ciência (IGC), para o desenvolvimento de algoritmos paralelos em biologia computacional, e está também envolvido na proposta de projecto europeu *Magma*, a qual envolve a construção de uma nova infra-estrutura GRID com ênfase na partilha de recursos de rede, nomeadamente largura de banda. O INESC-ID participa neste projecto através da definição e implementação de modelos inovadores de políticas de controlo de acesso e utilização de recursos.

No Departamento de Informática da Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa, o Centro de Investigação em Informática e Tecnologias de Informação (CITI), participa nos projectos: *TransGrid*, sobre aplicações, modelos e ambientes computacionais para Computação

GRID; *Linux Cluster and Grid Computing*, que se insere no programa *Equinox* da IBM (Education & Research); e *High-Performance and Grid Computing* que se insere no programa *EU AsiaLink* iniciado em 2005 com a Universidade de Delft, a Escola de Graduação da Academia Chinesa das Ciências e a Universidade de Tsinghua, sobre o desenvolvimento conjunto de currículos na área de Computação GRID e o intercâmbio de professores e estudantes. A Computação GRID é também leccionada a nível do curso de Mestrado em Engenharia de Informática da FCT/UNL, havendo actualmente diversas teses de doutoramento em curso sobre o tema.

O Laboratório de Instrumentação e Física Experimental de Partículas (LIP) é responsável por quase toda a actividade portuguesa no CERN, na área da Física Experimental de Partículas. É membro do projecto *Worldwide LHC Computing Grid*, tendo por responsabilidade central a instalação, manutenção e operação de uma infra-estrutura de nível TIER2 com uma capacidade equivalente a 1.000 unidades de processamento. O envolvimento do LIP em actividades relacionadas com a Computação GRID começou em 2001 e desde então foi ou é parceiro nos seguintes projectos europeus: *DATAGRID*, *CROSSGRID*, *EGEE-I*, *EGEE-II*, *ELLA*, *INT.EU.GRID*. No âmbito destes projectos, as responsabilidades do LIP cobrem áreas tão diversas como o desenvolvimento e validação de software, a instalação, operação e manutenção de infra-estruturas centrais dos projectos, a segurança e a formação avançada. Adicionalmente, o LIP opera a Autoridade Nacional de Certificação GRID para emissão e gestão de certificados de assinaturas electrónicas para autenticação das máquinas GRID, válida para todo o território nacional.

O Laboratório Associado ITQB-IGC-IBET (Instituto de Tecnologia Química e Biológica – Instituto Gulbenkian de Ciência – Instituto de Biologia Experimental e Tecnológica) tem em desenvolvimento uma infra-estrutura computacional de alto desempenho para gerar e analisar informação biológica em grandes quantidades. Em termos de armazenamento e análise, esta está particularmente vocacionada para informação de Microarrays, e de outras experiências de alto débito em Proteómica, ou da sequenciação em larga escala típica de grandes projectos de genotipagem. A infra-estrutura de análise, baseada no IGC, inclui um agrupamento de 120 processadores com uma capacidade agregada de armazenamento superior a 4TB, e utilizando software no âmbito do programa académico da IBM (Academic Initiative). Este permite também ensaiar e disponibilizar software desenvolvido localmente para uma arquitectura paralela, oferecendo o seu uso como serviço adicional. Tal é o caso, por exemplo, de software já desenvolvido no INESC-ID no âmbito do projecto BioGRID. Permite igualmente fortalecer a ligação da comunidade de utentes bioinformáticos portugueses a redes internacionais, como a Rede Europeia de Biologia Molecular (EMBnet), da qual o IGC é membro.

No Laboratório Nacional de Engenharia Civil (LNEC), os estudos apoiados por modelação matemática exigem elevada capacidade computacional. Entre os domínios de intervenção computacionalmente mais exigentes estão a hidrodinâmica e transporte em estuários, a modelação de ondas de cheias, a agitação marítima, o risco sísmico, o comportamento de barragens, a dinâmica dos fluidos aplicada à ventilação em edifícios, a detecção e prospecção de informação. No âmbito da actividade de investigação programada do LNEC, os principais projectos em curso são a

implementação de um sistema computacional de elevado desempenho e a distribuição e paralelização de modelos numéricos em hidráulica e em ambiente.

A Chipidea é uma empresa de desenvolvimento de circuitos integrados de sinal misto (analógico-digital) de larga escala. O desenho de um *chip* pode necessitar de centenas de simulações. Cada simulação poderá ter de utilizar 3 GB de memória RAM para ser processada e os seus resultados podem ocupar mais do que 5 GB de memória em disco. A complexidade crescente dos chips que a Chipidea desenha, para continuar a actuar de forma competitiva no mercado mundial, levou a que a empresa implementasse uma infra-estrutura GRID que lhe permite utilizar o tempo livre dos processadores dos seus computadores de forma eficiente para as simulações do desenho de *chips*. A arquitectura desta GRID compreende uma *Technical Computational Farm* e toda a rede de computadores dos utilizadores em plataforma Linux. Os utilizadores apenas têm de submeter, a partir dos seus postos de trabalho, pedidos de simulações à GRID. Esta, de acordo com a configuração estabelecida, ocupa-se de encontrar os processadores livres necessários (e, se possível, os mais adequados) para a simulação desejada, processá-la e disponibilizar os seus resultados em local conhecido para posterior análise. Com esta organização em GRID dos seus recursos computacionais, a Chipidea foi capaz de reduzir o tempo de desenho e verificação de *chips*, proceder a verificações mais completas dos mesmos, aumentar a complexidade dos *chips* desenhados e ainda reduzir alguns custos em equipamento informático.

A Critical Software é uma empresa de desenvolvimento de soluções software para sistemas críticos que tem vindo crescentemente a actuar na área dos sistemas distribuídos. Recentemente desenvolveu o WMPI II, uma versão avançada do *Message Passing Interface* (MPI) que apoia a paralelização do código para utilização em “*clusters*” de elevado desempenho. O WMPI II acrescenta novas potencialidades face às versões anteriores, nomeadamente ao nível da gestão dinâmica de processos, da simultaneidade de acesso a disco e da segurança. O WMPI II é utilizado em aplicações industriais, em particular nas áreas da computação dinâmica de fluidos e análises de elementos finitos, e educacionais. O WMPI II é a única implementação completa do standard MPI-2 para “*clusters*” heterogéneos com máquinas Windows e Linux. Esta aplicação, que a Critical Software licenciou em todo o mundo, é utilizada por empresas, universidades e laboratórios de I&D de ponta como MIT San Diego Super Computer Center, Deutsche Telekom, Chevron-Texaco e Honeywell.

A empresa YDreams desenvolveu para o Instituto da Água (INAG) um sistema de *Data Mining* que usa redes neuronais para apoio à gestão de um vasto conjunto de redes de monitorização automáticas e não-automáticas essenciais para, por exemplo, sistemas de alerta de cheias, que alimentam o Sistema Nacional de Informação de Recursos Hídricos (SNIRH). Para superar a elevada necessidade de capacidade computacional envolvida o *SNIRH Data Mining* utiliza um conjunto de computadores dedicados e de secretária do INAG. A tecnologia usada neste sistema foi totalmente desenvolvida pela YDreams e está a ser aplicada em projectos de áreas distintas, incluindo sistemas para desporto de alta competição, sistemas automáticos de alerta e ferramentas analíticas na área biomédica.