

Portal Web da Grade Computacional do LNCC

C.G.Ribeiro¹, J.C. de Oliveira¹, F.S.G. de Oliveira¹, B.Schulze¹

Grupo COMCIDIS – Coordenação de Ciência da Computação
Laboratório Nacional de Computação Científica (LNCC)
Av. Getúlio Vargas, 333, Quitandinha – 21651-070 – Petrópolis – RJ – Brazil
{const, fgomes, jauvane, schulze}@lncc.br

***Abstract.** This work describes a computational grid infrastructure, the web portal to this infrastructure, and the implementation of a tool for the submission of tasks, with the objective to facilitate the use and the access to the computational resources available in the computational grid.*

***Resumo.** O trabalho descreve uma infraestrutura de grade computacional, o portal web esta infraestrutura e a implementação de uma ferramenta para a submissão de tarefas, objetivando facilitar o uso e acessibilidade dos recursos computacionais disponíveis na grade computacional*

1. Introdução

A Internet (e seus sucessores) no topo das redes de comunicação, com flexibilidade e largura de banda crescentes, resultará numa infra-estrutura de rede maciça composta de objetos interconectados altamente diversificados que devem suportar o projeto e uso de sistemas com comportamento previsível e desejável.

Com o desenvolvimento das tecnologias de informação e de comunicação espera-se permitir o projeto de sistemas computacionais e de informação baseados inteiramente na abundância de recursos da infra-estrutura disponível e de potencial de processamento geograficamente distribuído. Estes desenvolvimentos conduzirão ao grande desafio de definir e de explorar sistemas dinamicamente configurados composto de entidades que interagem de novas formas com seu ambiente para realizar ou controlar suas tarefas computacionais.

Os sistemas vislumbrados são altamente dinâmicos: dispositivos físicos são móveis, conectividade e largura de banda estão mudando, processos e dados computacionais podem migrar, e os aplicativos vão e vêm. A disponibilidade e a resposta dos recursos que estão ativos em uma aplicação em um dado instante de tempo são não previsíveis, sendo difíceis de controlar. Espera-se que a escala dos sistemas seja extremamente grande, em número de componentes e, em alguns casos, em área geográfica. Os aspectos de projeto envolvidos na construção de sistemas que podem ser configurados em tais ambientes coloca desafios enormes à computação, demandando pesquisa focada em problemas fundamentais que necessitam ser resolvidos.

Para aproveitar a flexibilidade e a potencialidade de sistemas interativos evoluindo rapidamente, são requeridas idéias e princípios fundamentais radicalmente novos. É preciso desenvolver os fundamentos para permitir o projeto destes sistemas no futuro e lidar com várias questões complexas resultantes da sua construção.

O objetivo como um todo é obter técnicas (modelos, estruturas, métodos, algoritmos) para construir sistemas que são flexíveis, confiáveis, seguros, robustos e eficientes. O interesse dominante não é simplesmente representar e manipular eficientemente dados mas no entanto tratar coordenação e interação, segurança, confiabilidade, robustez, modos de falha, controle de risco das entidade no sistema e no projeto, descrição e desempenho global do próprio sistema. Paradigmas completamente diferentes de computação podem ter que ser desenvolvidos para tratar eficazmente estas questões, devendo tratar sistemas como segue:

- A configuração varia com o tempo. Por exemplo, o sistema está aberto à entrada de entidades computacionais novas e do mesmo modo a sua remoção. O comportamento das entidades pode variar com o tempo.
- Os sistemas são compostos de entidades computacionais autônomas onde a atividade não é controlada centralmente, seja porque o controle global é impossível ou pouco prático, ou porque as entidades são criadas ou controladas por proprietários diferentes.
- As entidades computacionais são móveis, devido ao movimento das plataformas físicas ou pelo movimento da entidade de uma plataforma a outra.
- Os sistemas operam com informação incompleta sobre o ambiente. Por exemplo, a informação torna-se rapidamente desatualizada e a mobilidade requer informação sobre o ambiente a ser explorado.

Estrutura deste trabalho: Na seção 2 temos uma descrição de trabalhos e atividades relacionadas. Na seção 3 é apresentada a atividade de integração de uma grade computacional numa intranet institucional.

2. Trabalhos Relacionados

Uma grade computacional agrega recursos de alto desempenho computacional, assim como permite a formação de Organizações Virtuais para a realização de projetos temáticos baseados em ambientes de computação distribuída. Organizações Virtuais são grupos de atuação em um determinado tema podendo estar geográfica e institucionalmente dispersos. Para a operacionalidade de tais organizações precisamos de ambientes colaborativos, seja na elaboração e desenvolvimentos de aplicações associados, seja na utilização das aplicações desenvolvidas ou de aplicações de terceiros. Implica também em uma grade de computadores e redes, e respectivas configurações.

No exterior se verifica a organização de grades computacionais a partir da integração de centros de (super)computação e suas aplicações. Como exemplo podemos citar o projeto TeraGrid da NSF (National Science Foundation), que cria o 1º ambiente de computação distribuída em escala de Teraflops unindo centros numa estrutura de escala nacional (NCSA, SDSC, Argonne e Caltech). Este projeto resulta de iniciativas destes centros em projetos da NSF desde 1997 em infraestrutura de informação (ex. software de grade, ambientes escaláveis em clusters, ferramentas de softwares, gerenciamento de dados e visualização, e códigos de aplicação). O centro de operações do TeraGrid (TOC) estabelece um conjunto de políticas que guiam o TeraGrid em termos de operação, uso e transferência de tecnologia. Operacionalmente o TOC oferece

suporte online 24x7, disponibilizando ferramentas de monitoração para verificação de desempenho e coordenação de atualizações de hardware e software distribuídos.

Como plataforma de software de grade computacional encontramos diversos projetos e podemos definir 3 principais tendências: a clássica, a orientada a objetos e a orientada a serviços (de web). A 1ª trata principalmente os aspectos de alocação de recursos e monitoração de filas de processos (a serem) submetidos. A 2ª considera um ambiente onde as unidades de processamento são objetos distribuídos que se intercomunicam e dependem de suporte a localização (transparente ou não), a migração, a transação, entre outros. A 3ª (e mais atual) soma as características das duas primeiras com as características encontradas em sistemas de informação baseados em infraestrutura de web, e segue na linha de componentes e serviços ativos. Em todas as três a segurança é parte fundamental e imprescindível, pois atualmente não é mais viável tratar computação distribuída sem mecanismos e políticas mínimas de segurança

Na seção a seguir é descrita resumidamente a implementação de uma grade a nível institucional (LNCC).

3. Grade Institucional (do LNCC)

3.1 Hardware e software usados

Para a implementação da GRADE LNCC inicial foram usados 3 nós Sun Enterprise 220R rodando o sistema operacional Solaris versão 8, cada nó com 2 processadores UltraSPARC-II e memória de 1,5 GBytes.

O middleware adotado nesta fase inicial foi o SUN Grid Engine Enterprise Edition na sua última versão (SGE-EE versão 5.3). Conforme especificações foram instalados os seguintes deamons: administração, execução, submissão e de escalonamento. Um dos nós é configurado como administrador do middleware de gerência da grade, resultando então em uma grade homogênea com todas as máquinas de mesma arquitetura, sistema operacional e middleware de grade.

Seguindo a tendência adotada pelos principais sites de e-science, os serviços de grade do LNCC adotam o uso de um portal web para o acesso aos serviços de grade. No nosso caso uma plataforma Sun Netra foi utilizada para tal. Nela o *daemon* de submissão de tarefas do middleware SGE-EE foi instalado, ficando a mesma dedicada somente a esta tarefa de portal de acesso via Web.

A Figura 1 ilustra a situação atual da GRADE em funcionamento no LNCC.

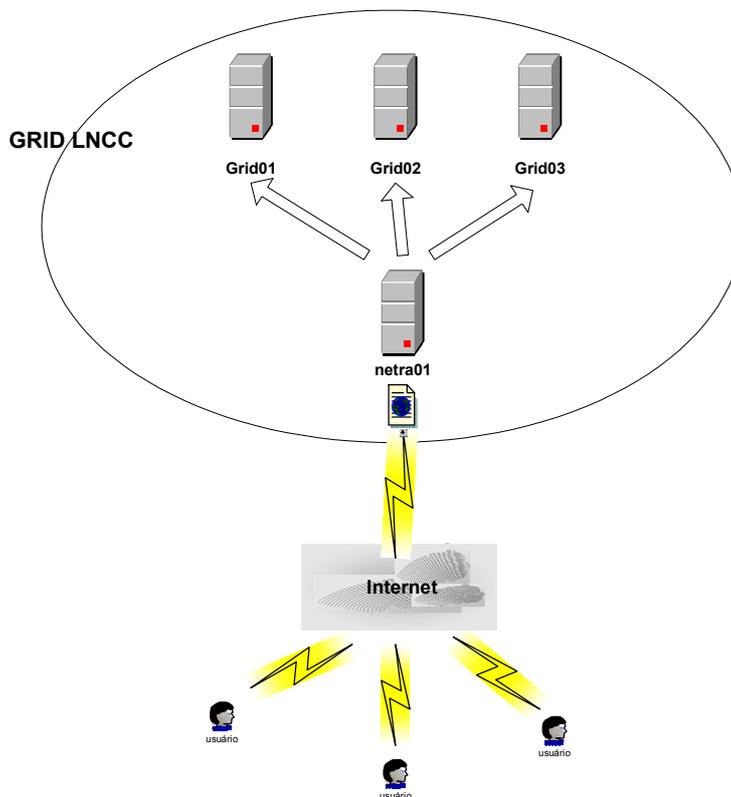


Figura 1. Grade do LNCC

3.2 Operação da GRADE LNCC:

O acesso a GRADE LNCC pode ser feito das seguintes maneiras: para usuários internos, via acesso ssh ou pelo portal web; e para usuários externos via portal web. Em ambos os casos são fornecidas conexões seguras. No caso do acesso web esta conexão é feita por um applet em Java, o qual fornece diversas funcionalidades úteis, tais como: transferência de arquivos, janela de terminal para a execução de comandos de linha, e outros. Além disso o próprio portal é uma fonte de informações básicas sobre computação em grade. Nele podemos encontrar artigos sobre a filosofia, componentes, middleware e softwares de computação em grade bem como instruções e exemplos de scripts de execução de tarefas no GRADE LNCC. A figura abaixo ilustra a página inicial do portal de acesso ao GRADE LNCC. Para acessar a GRADE LNCC devemos selecionar o botão “como utilizar o grid” como mostrado na Figura 3.

Após a validação do usuário pela tela mostrada na Figura 4, podemos escolher entre usar uma applet de conexão segura que fornece uma conexão segura ao estilo de uma term de telnet (figura 5), ou usar o gerador de scripts de execução como mostrado na figura 6. Dentre os serviços disponíveis podemos citar a execução de programas sequenciais em: C, C++, FORTRAN e Java, e a execução de programas paralelos/distribuídos em: C, C++ e FORTRAN via biblioteca de passagem de mensagem MPI.

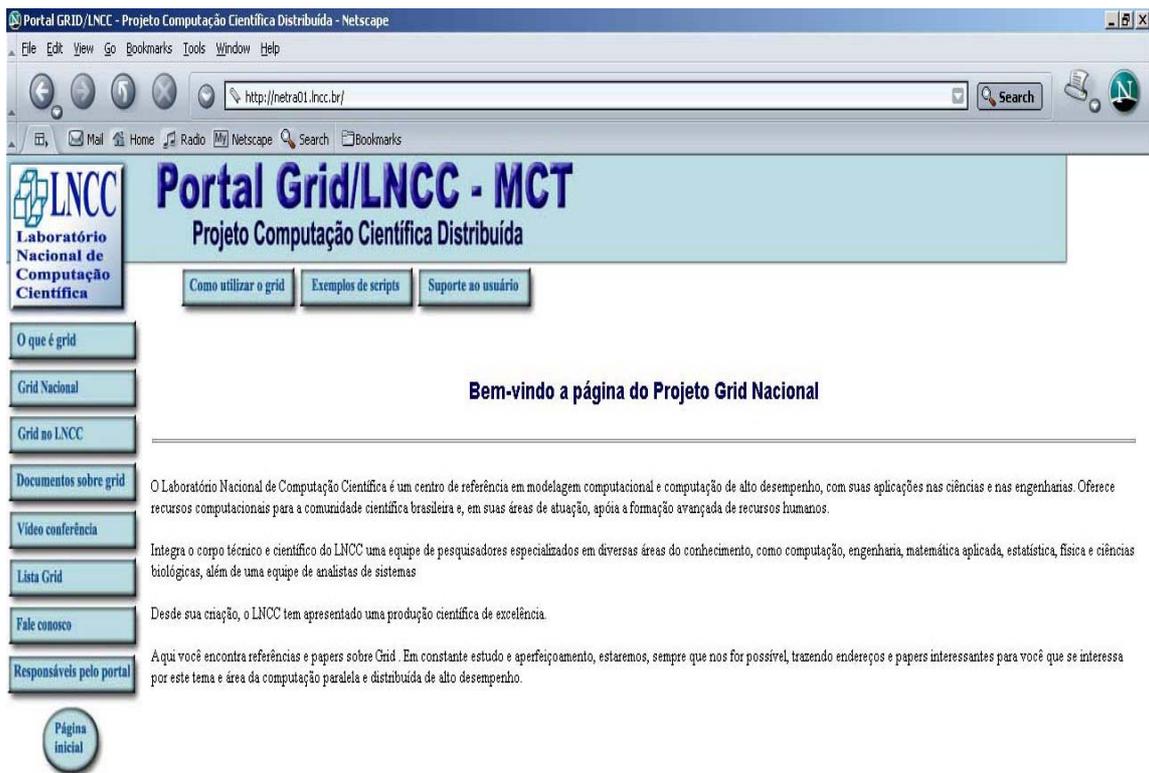


Figura 2. Portal GRID

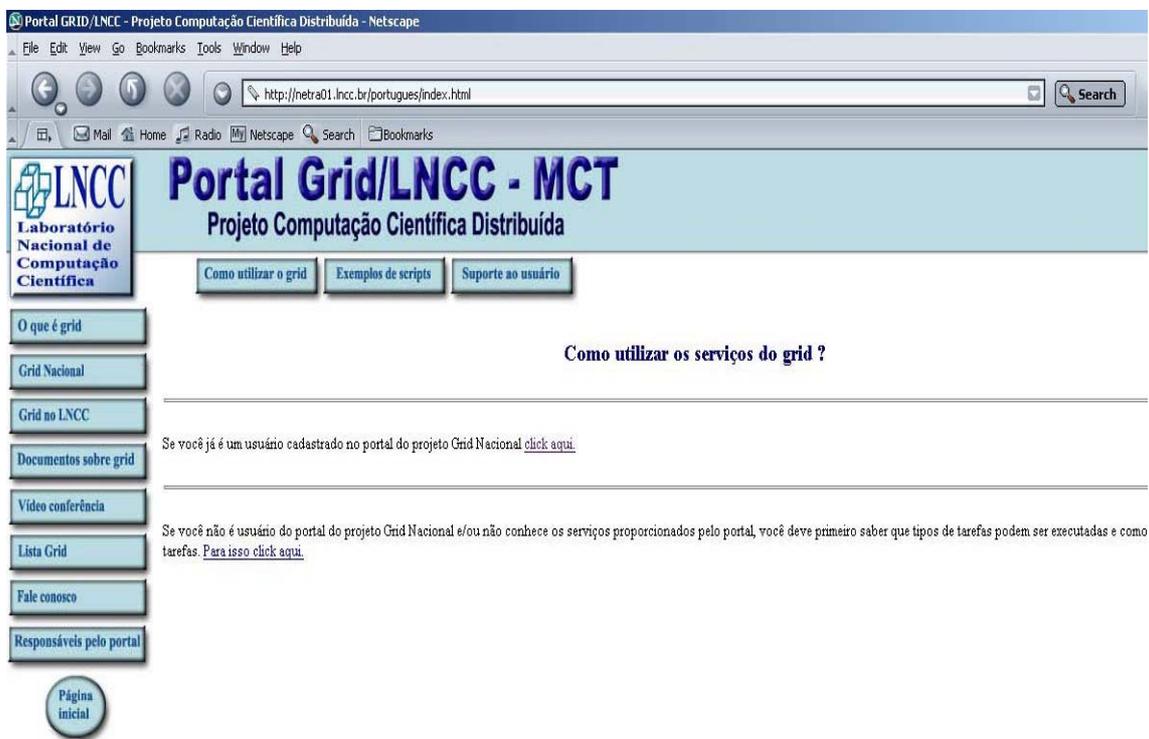


Figura 3. Acesso ao Portal GRID

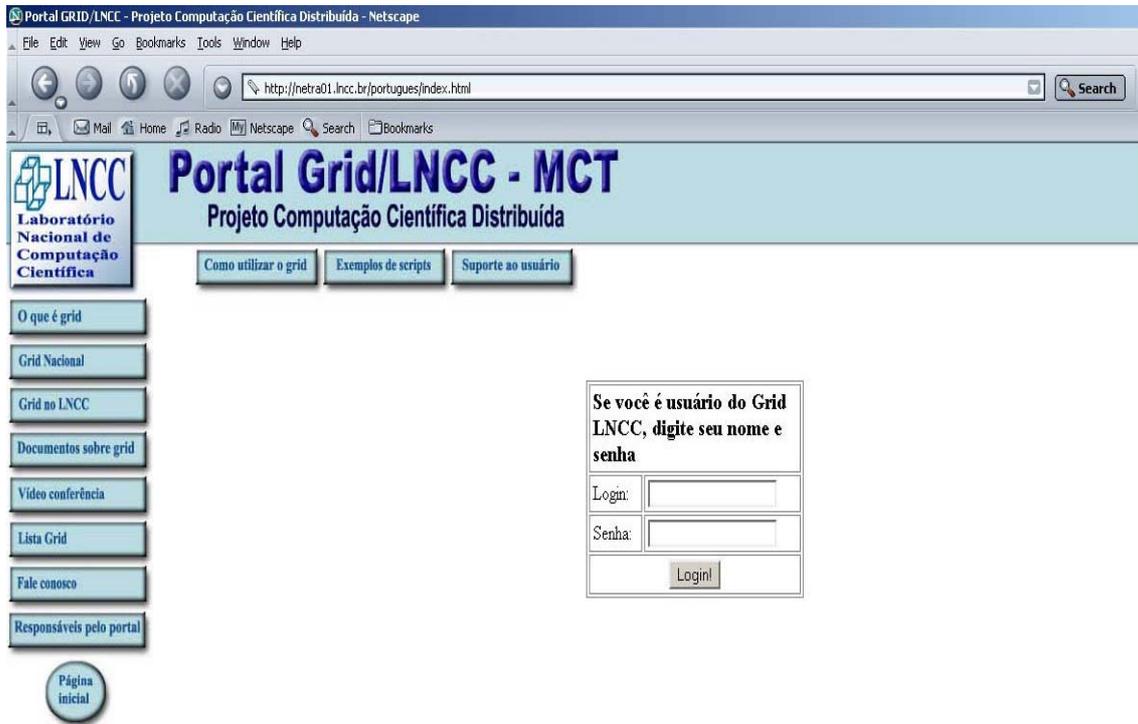


Figura 4. validação no acesso ao Portal GRID

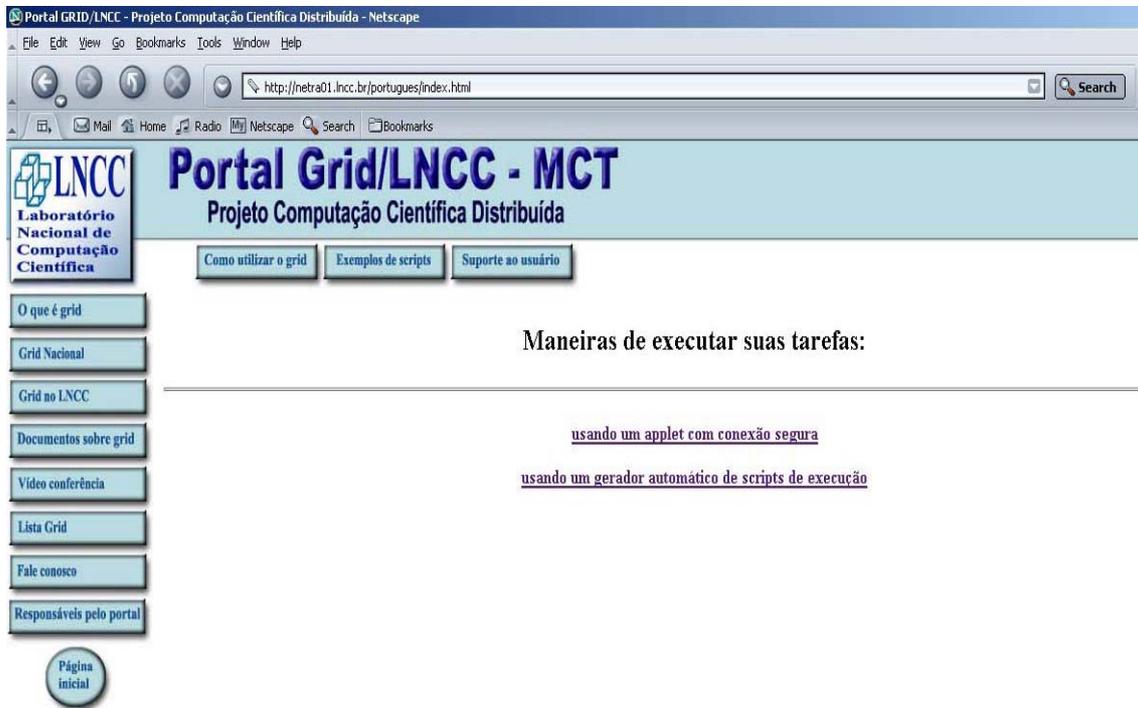


Figura 5. opções de uso do Portal GRID

Portal GRID/LNCC - Projeto Computação Científica Distribuída - Netscape

File Edit View Go Bookmarks Tools Window Help

http://netra01.lncc.br/portugues/index.html

LNCC Laboratório Nacional de Computação Científica

Portal Grid/LNCC - MCT
Projeto Computação Científica Distribuída

Como utilizar o grid Exemplos de scripts Suporte ao usuário

O que é grid
Grid Nacional
Grid no LNCC
Documentos sobre grid
Vídeo conferência
Lista Grid
Fale conosco
Responsáveis pelo portal

Página inicial

Criação de arquivos de script para submissão de tarefas no grid

Escolha os parâmetros de sua tarefa:

Tipo de programa:	- por favor escolha -
Se o programa for paralelo/distribuído, informar o número de processos do mesmo:	<input type="text"/>
Nome do arquivo fonte:	<input type="text"/>
Nome do executável a ser gerado:	<input type="text"/>
Nome para o arquivo de saída padrão de dados e erros:	<input type="text"/>
E-mail para contato	<input type="text"/>
Quando enviar e-mail:	<input type="radio"/> no início da execução do programa <input type="radio"/> ao final da execução do programa <input type="radio"/> caso o programa aborte
Compilador a ser usado:	- por favor escolha -

Verificar arquivo gerado Criar Arquivo

Faça o Upload dos arquivos com: o programa fonte, os dados e o script de execução para que você possa executar sua tarefa no grid

Browse... faça o upload dos arquivos

Figura 6. opções de uso do Portal GRID

Agradecimentos: Este trabalho conta com o apoio do CNPq e PCI/MCT e também equipamento disponibilizados por Sun Microsystems e CIMCORP.

Referências

The Globus Project, The Open Grid Service Architecture, www.globus.org/research/.
 Sun Microsystems, SunTM Grid Engine Software SGE, www.sun.com/software/gridware.
 OMG, Common Object Request Broker Architecture CORBA, www.corba.org.