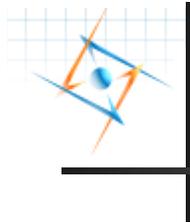


Planejamento para Implantação do GridUnesp

II Workshop GridUnesp
Outubro de 2006



Roteiro

- O que é Grid
 - Porque Grid?
 - Conceitos gerais.
 - Componentes de um Grid.
 - Exemplo: o SPRACE (São Paulo Regional Analysis Center)
- Implantação dos Componentes Físicos.
 - Compra e instalação das máquinas.
 - Criação dos Clusters locais.
 - Criação de um Cluster de Clusters.
- Implantação do Middleware
 - Inclusão dos Clusters no Open Science Grid.
 - Criação do Grid da Unesp, o GridUnesp
- Aplicativos
 - Adaptação dos aplicativos para uso em Grids.
 - Treinamento de usuários.
 - Formação de pessoal de apoio.

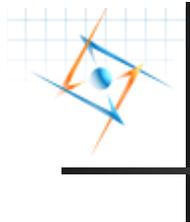
Porque Grid?

CERN

**Colisões Proton – Proton
Acelerador com 27 km de circunferência
Localizado em Genebra**



The Large Hadron Collider



Alguns números do LHC

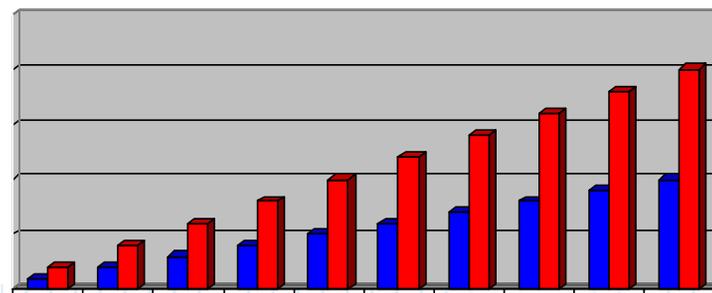
- Quantidade de dados produzido no LHC:
 - Observação de 10 bilhões de colisões (eventos) por ano.
 - Tamanho de cada evento:
 - 1 MB (CMS), 25 MB (Alice).
 - Total de 15 Petabytes por ano.
- Número de CPU's para analisar os dados:
 - 70.000 processadores no total
 - 5.000 processadores no CERN
- Transmissão dos dados (10 PB) :
 - 46 dias @ 2.5 Gbps
 - 746 dias @ 155 Mbps
 - 158 anos @ 2 Mbps

Potência X Conexão

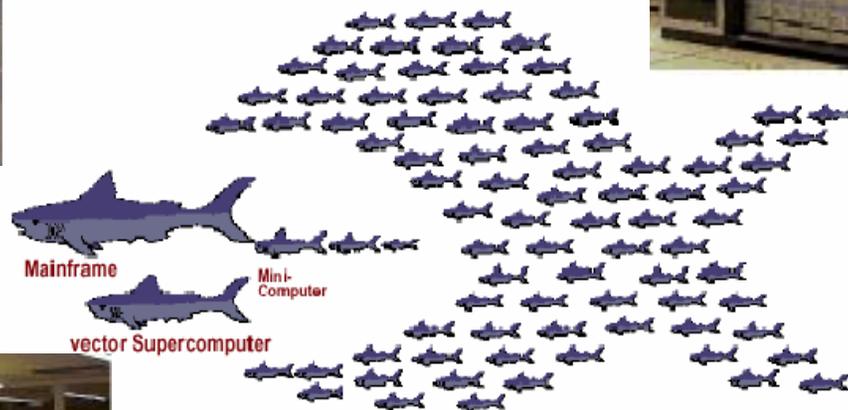
Processor	Year	Transistor
4004	1971	2,250
8008	1972	2,500
8080	1974	5,000
8086	1978	29,000
286	1982	120,000
386™	1985	275,000
486™ DX	1989	1,180,000
Pentium®	1993	3,100,000
Pentium II	1997	7,500,000
Pentium III	1999	24,000,000
Pentium 4	2000	42,000,000

Tecnology	Year	Conection (bps)
Analogue	1985	9.600
Digital	1989-1994	256.000
Shared	1990-1993	1.500.000
	1996-1998	4.000.000
	1999-2000	20.000.000
	2001-2002	310.000.000
	2002-2003	622.000.000
Lambda	2003-2004	2.500.000.000
	2005	10.000.000.000

- Número de transistors dobra a cada 18 meses
- Bandwidth dobra a cada 9 meses
- Diferença de 2 ordens de grandeza em 10 anos



Computação Distribuída: Clusters e Grids



SAMGrid: O Grid do Fermilab

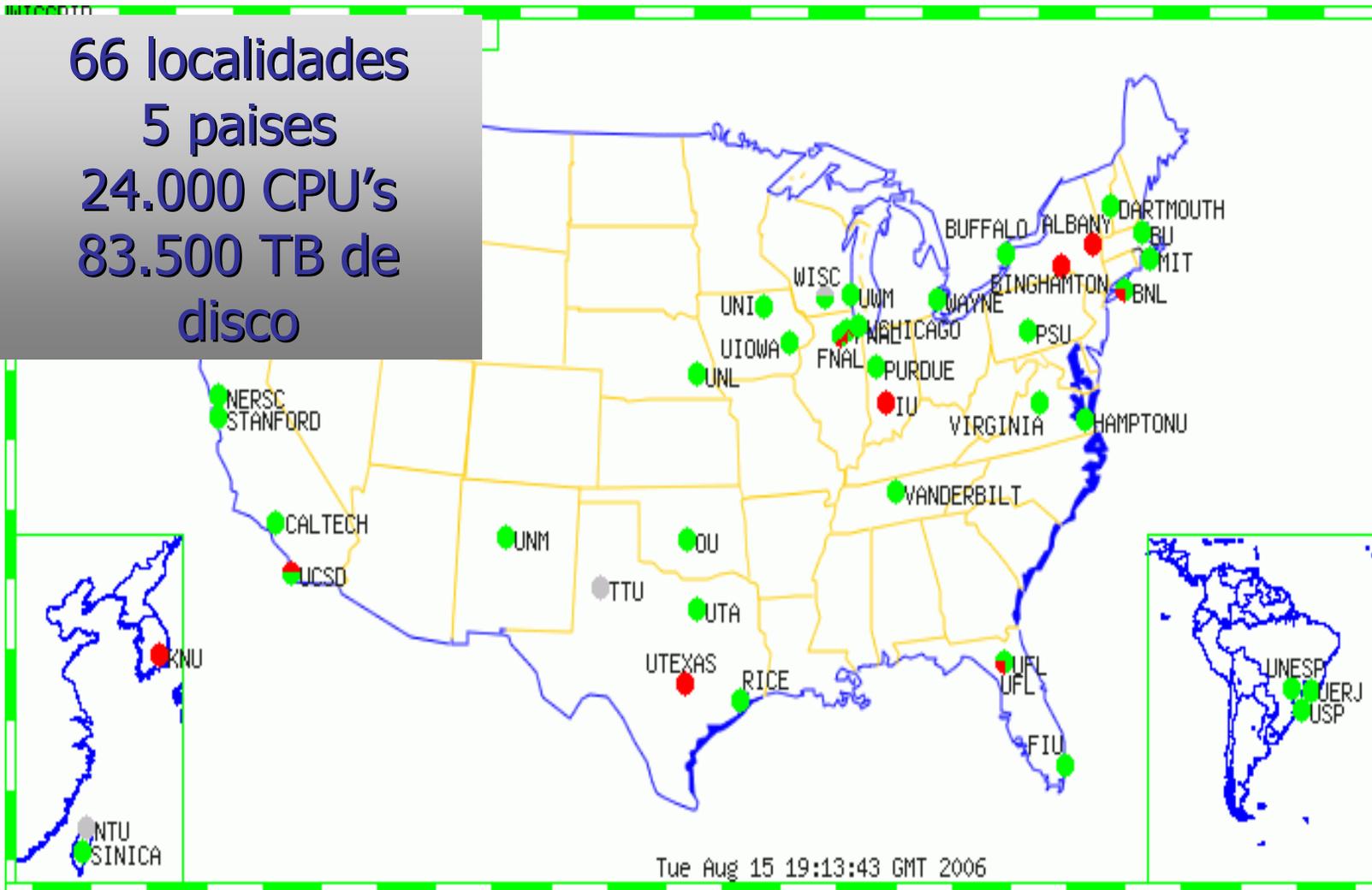


Participating Experiments:

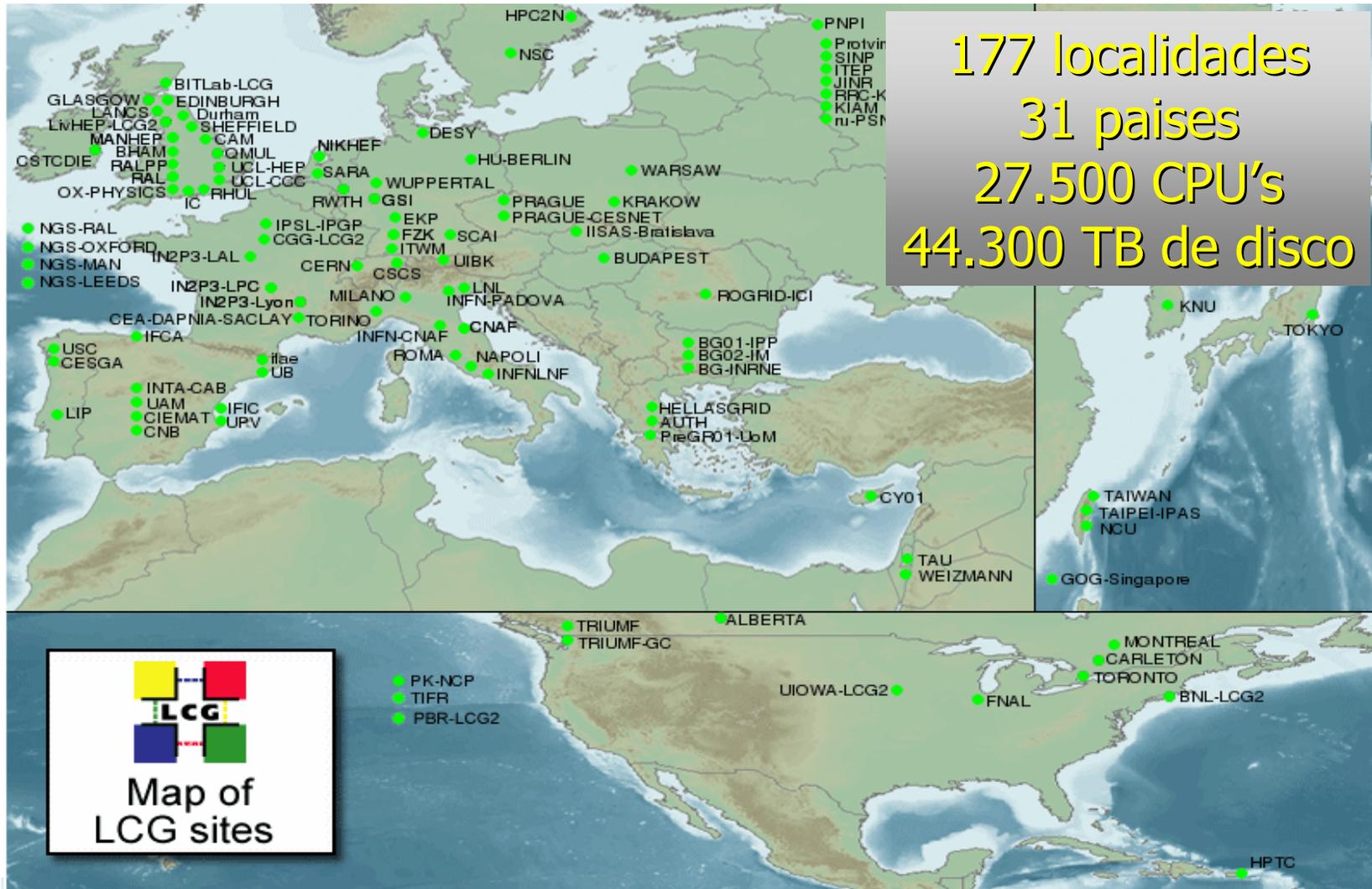
- DO
- CDF

Open Science Grid (OSG)

66 localidades
5 paises
24.000 CPU's
83.500 TB de
disco



LHC Computing Grid (LCG)

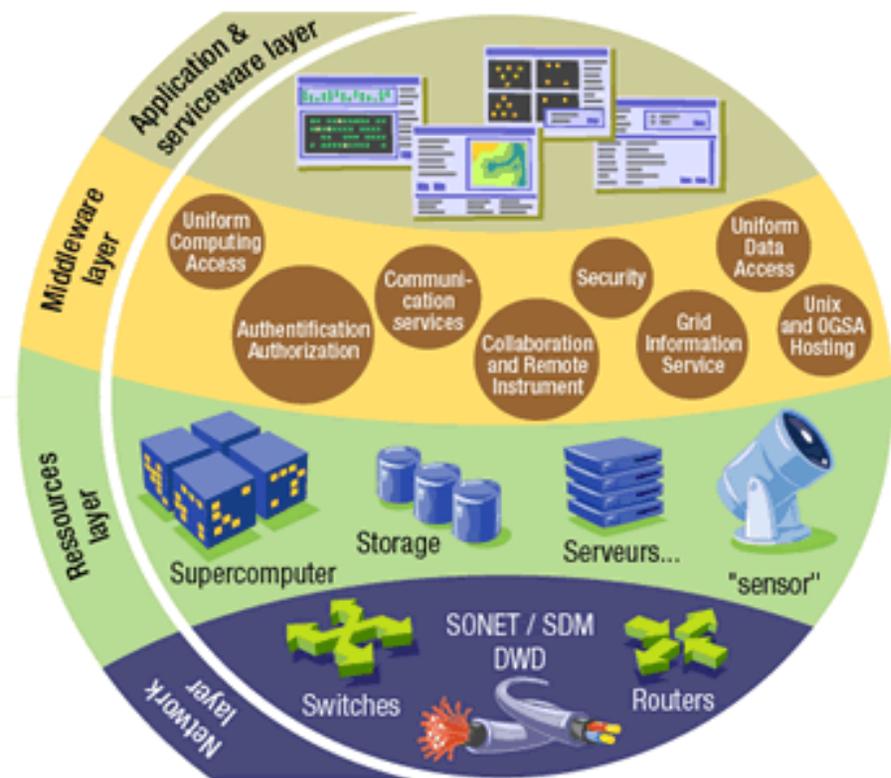


177 localidades
31 países
27.500 CPU's
44.300 TB de disco



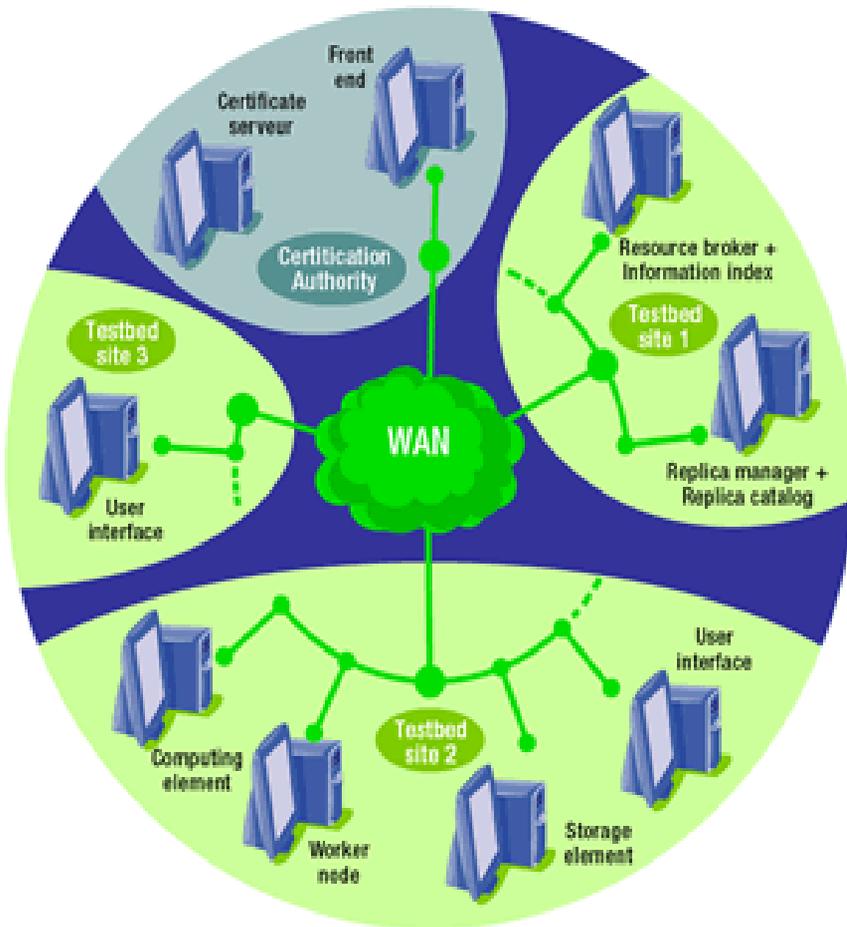
O que é Grid?

- Grid é um sistema que coordena recursos sem um sistema de controle central.
- Usa protocolos padrão e código aberto.
- Possui diferentes níveis de funcionalidade.

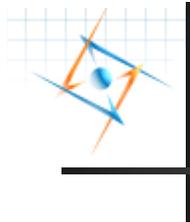


- **Aplicativos e Serviços**
 - Programas dos Usuários
 - Interfaces de Uso
 - Interfaces de Monitoramento.
- **Middleware**
 - Acesso aos processadores
 - Acesso aos dados
 - Serviços de Informação
- **Recursos**
 - Servidores, Discos, etc..
- **Infraestrutura de Rede**
 - Conexões, Roteadores, etc...

Elementos do Grid



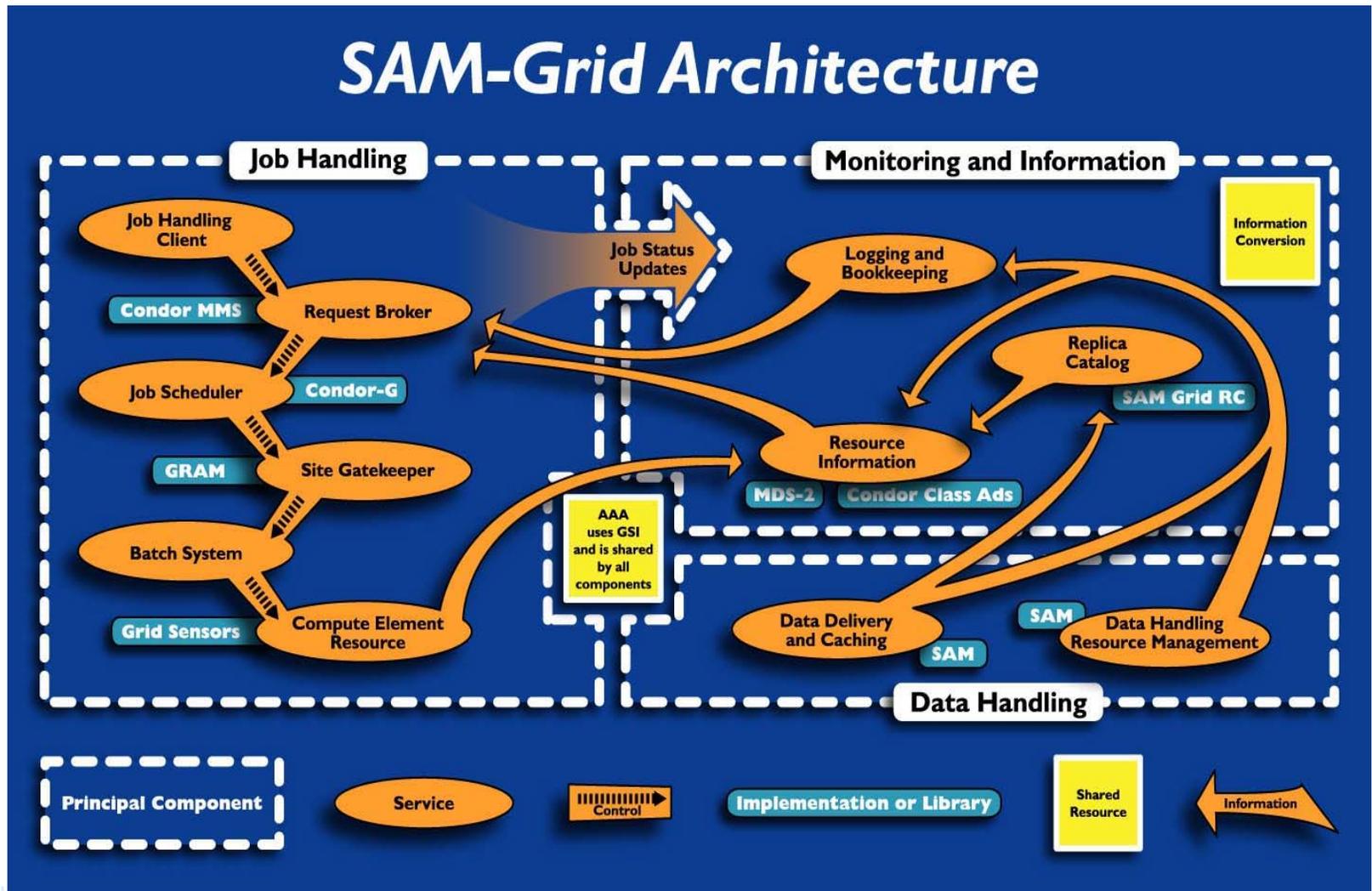
- **Alocador de Recursos**
 - Recebe os pedidos dos usuários
 - Encontra os recursos necessários
 - Pergunta ao Índice de Informações
- **Índice de Informações**
 - Mantém a informação sobre os recursos disponíveis
 - Pode residir na mesma máquina que o Alocador de Recursos
- **Gerenciador de Réplicas**
 - Coordena o replicação dos dados
 - Providencia a redundância dos dados
 - Move os dados para a proximidade do local onde será realizado o processamento
- **Catálogo de Réplicas**
 - Pode residir na mesma máquina que o Catálogo de Réplicas
 - Mantém a informação sobre as réplicas
 - Associa o nome lógico do dado a um ou mais nomes físicos
- **Elemento de Processamento**
 - Recebe os pedidos de trabalho e os distribui aos trabalhadores
 - Provê a interface entre o Grid e os sistema de ordenamento de trabalhos
 - Pode gerenciar um ou mais trabalhadores.
- **Nó de Trabalho**
 - É a máquina que faz o processamento dos dados.
- **Elemento de Armazenamento**
 - Provê espaço de armazenamento de dados
 - Provê uma interface homogênea aos diversos Sistemas de Armazenamento.
- **Interface do Usuário**
 - Habilita o acesso ao Grid.



O Middleware: Globus Toolkit

- **GRAM** (Globus Resource Allocation Manager): Converte um pedido de uso de recursos em comandos que o sistema local compreenda.
- **GSI** (Grid Security Infrastructure): Provê a identificação e autenticação do requerente de uso dos recursos
- **MDS** (Monitoring and Discovery Service): Realiza a coleta de informações sobre os recursos disponíveis
- **GRIS** (Grid Resource Information Service): Questiona os recursos sobre sua configuração, situação e capacidade
- **GIIS** (Grid Index Information Service): Coordena os diversos serviços de GRIS
- **GridFTP**: Realiza a transferência dos arquivos.

Arquitetura Básica de um Grid



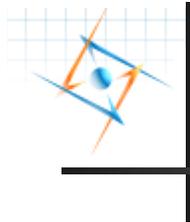
Exemplo: o SPRACE



Fase 1

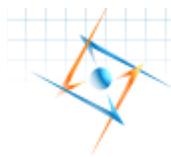
Fase 2





Infraestrutura do SPRACE

- Componentes do Cluster:
 - 20 Nós de Trabalho dual Xeon 2.4 GHz 1GB desde Março de 2004.
 - 32 Nós de Trabalho dual Xeon 3.0 GHz 2 GB desde Junho 2005.
 - 32 Nós dual Xeon dual Core (Woodcrest) 2.4 GHz 4GB sendo entregues.
 - 2 Head Nodes, 1 Servidor de Disco, 1 Servidor de dCache.
 - 12 TB em 4 módulos RAID (SCSI Ultra 360 10K).
 - 232 Condor batch slots com 320 kSpecInt2k de poder de processamento.
 - Mais 16 TB em discos locais (em breve) perfazendo um total de 28 TB.
- Conectividade:
 - Conexão interna Gigabit entre todos os elementos do cluster
 - Conexão externa Gigabit compartilhada com a USP até a PTT em Barueri.
 - Conexão internacional gigabit via WHREN-LILA até rede ABILENE (US)
 - Conexão Gigabit exclusiva nas próximas semanas até PTT
- Serviços instalados:
 - 2 Clusters separados
 - Cluster do Dzero/SamGrid
 - Cluster do CMS/OSG



Monitoramento do SPRACE



Cluster Report for Wed, 26 Jul 2006 18:42:31 -0300

Get Fresh Data

Metric Last Sorted

Physical View

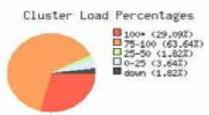
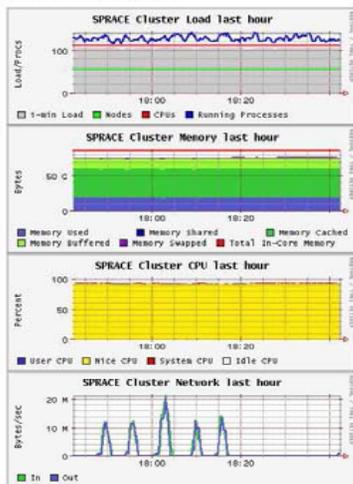
SPRACE Grid > SPRACE Cluster >

CPU's Total: 112
Hosts up: 54
Hosts down: 1

Avg Load (1.5, 5, 1m):
95%, 95%, 93%

Localtime:
2006-07-26 18:42

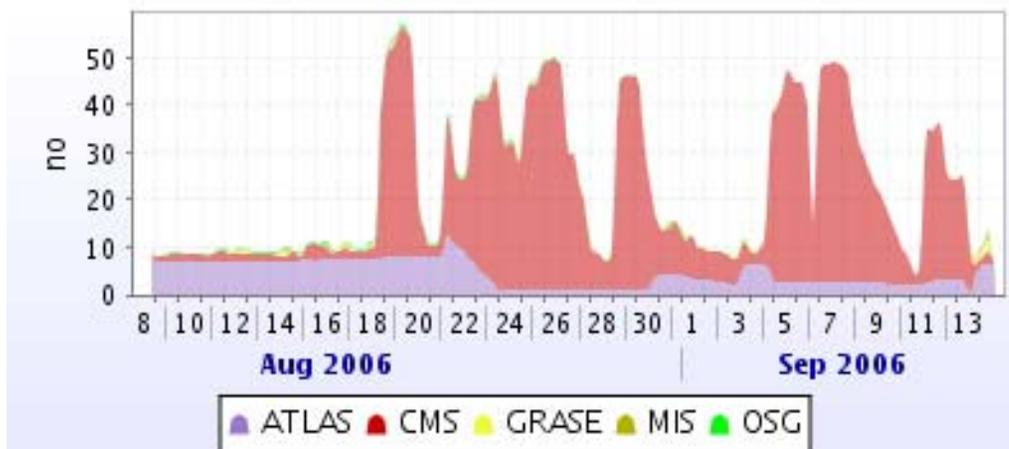
Overview of SPRACE Cluster



Show Hosts: yes no | SPRACE Cluster load_one last hour sorted by hostname | Columns



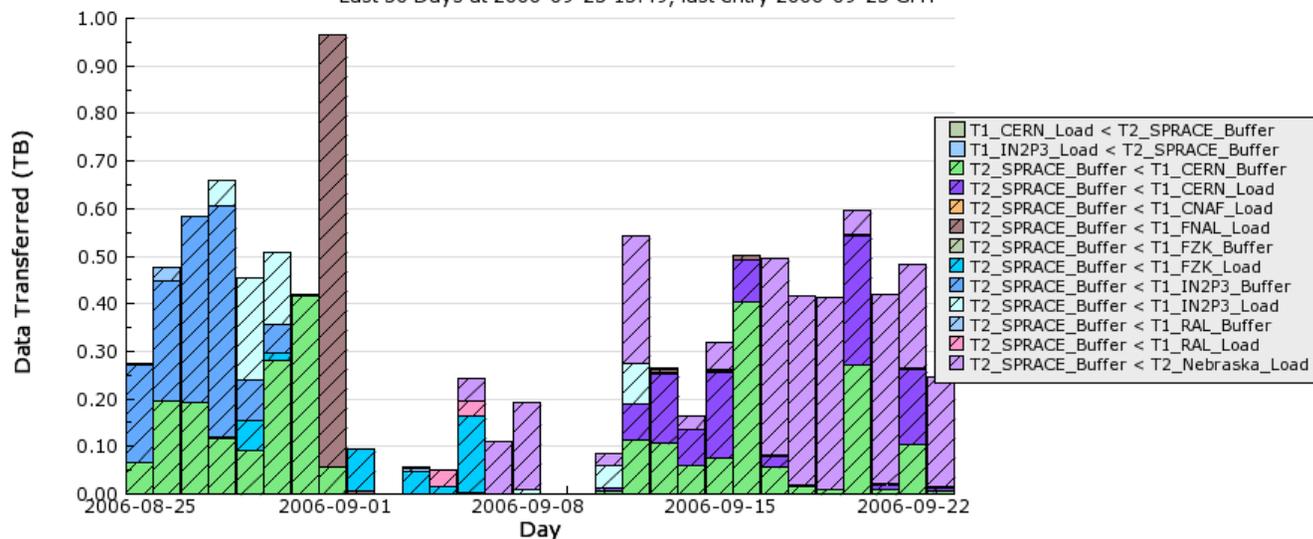
Trabalhos no OSG e Transferência de Dados



- Monitoramento do Monalisa
- Análise de Dados para o CMS
- Produção de Monte Carlo para o CMS.
- Uso por outras Organizações Virtuais

PhEDex SC4 Data Transfers By Links matching 'SPRACE'

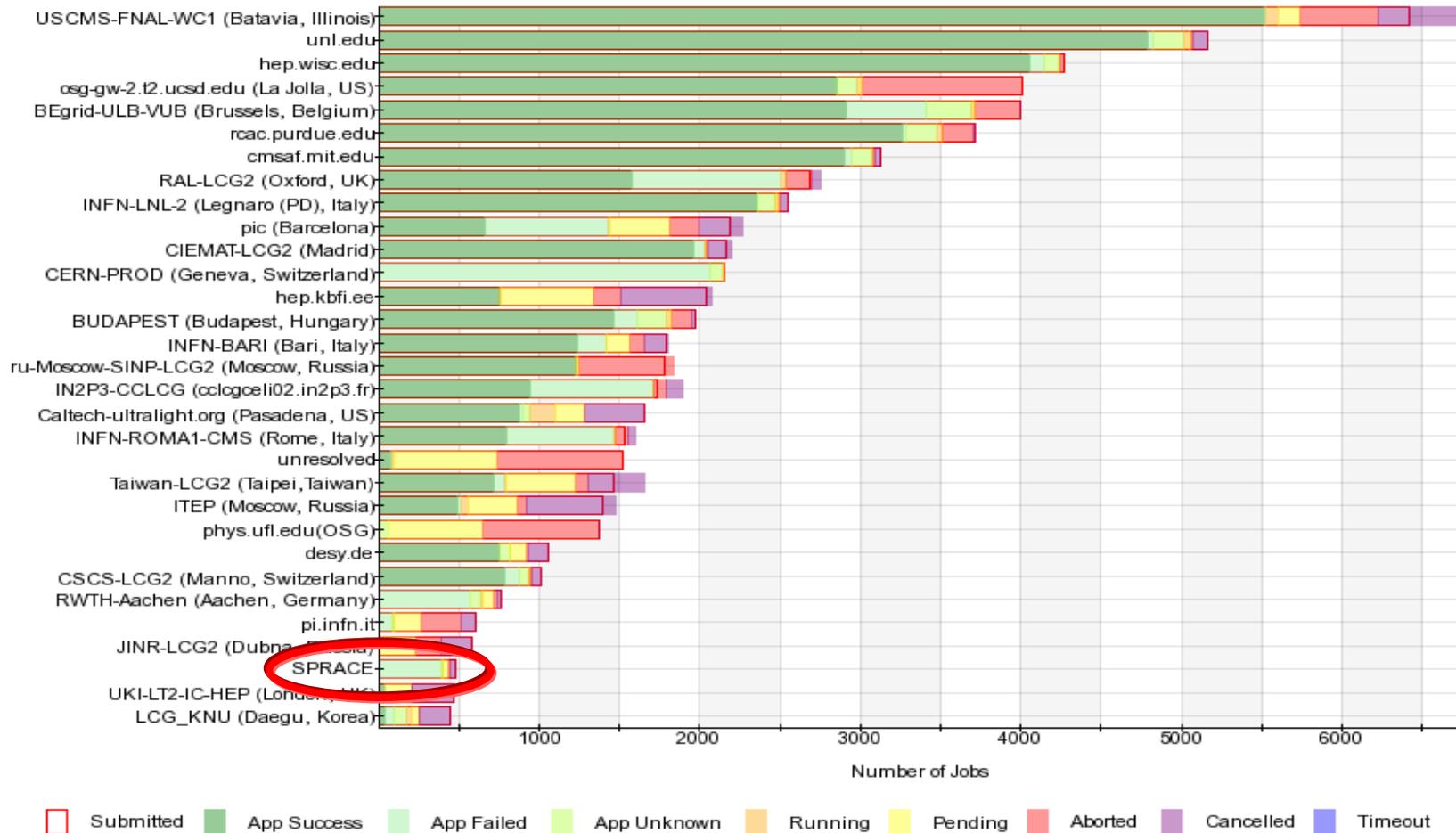
Last 30 Days at 2006-09-23 13:49, last entry 2006-09-23 GMT



- Fluxo de dados de/para SPRACE
- Média de 500 GB/dia
- Sistema de checagem de conexão do CMS

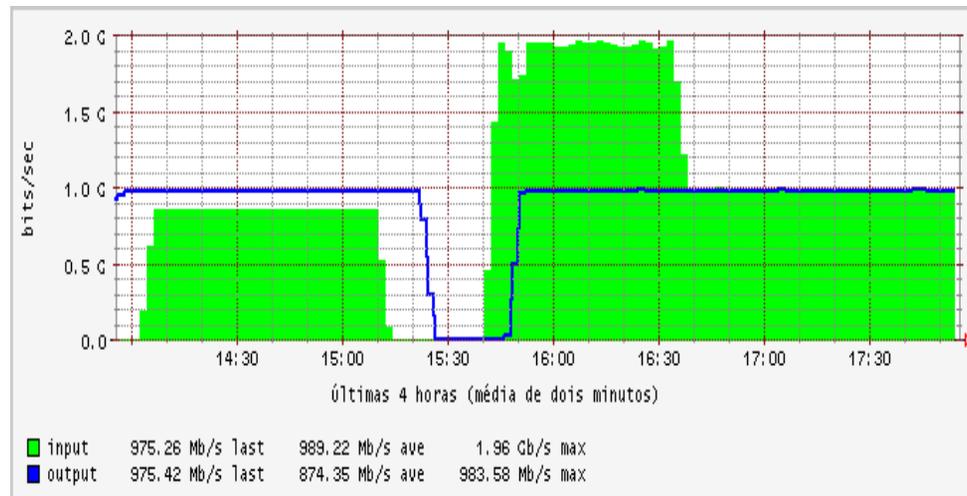
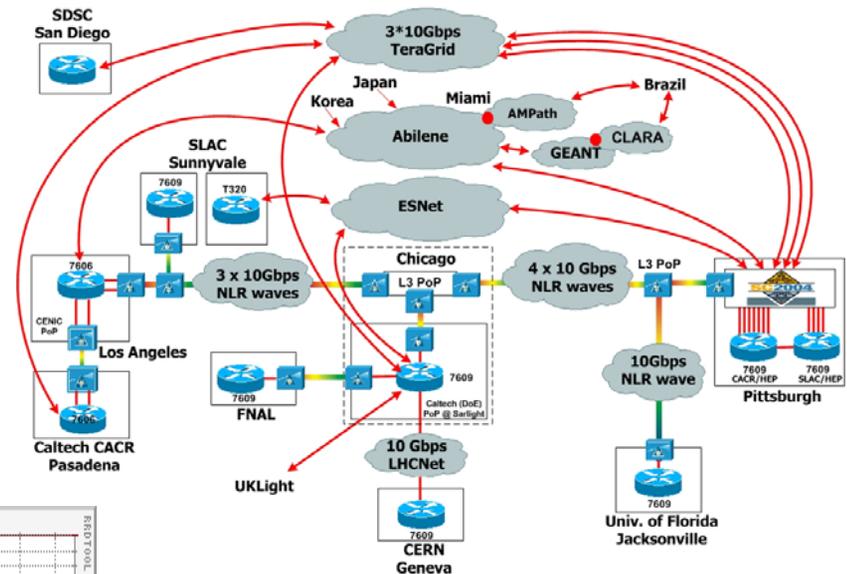
Trabalhos de Análise do CMS no OSG and EGEE

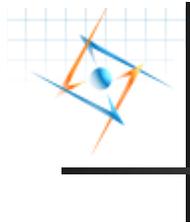
Site vs Number of Jobs for HajduCsaba/csaba for JobRobot



SPRACE na SC2004 Bandwidth Challenge

- Troca de dados entre São Paulo e Pittsburg.
- Sustentado 2.93 (1.95 + 0.98) Gbps durante 1 hora
- Recorde de transmissão de dados entre os hemisférios Norte e Sul.





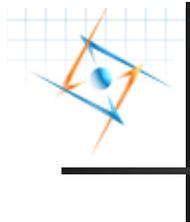
Implantação dos Componentes Físicos

- Estabelecimento de funcionalidades em cada etapa do processo
- Alocação e treinamento de pessoal para realização das etapas
- Compra e instalação das máquinas.
 - Compra do equipamento com sistema operacional instalado.
 - Instalação nos racks e cabeamento da rede.
 - Configuração do sistema de armazenamento em RAID5
- Criação dos Clusters locais.
 - Instalação dos serviços de rede local.
 - Configuração das interfaces de rede.
 - Configuração dos serviços de NFS, NIS, sshd e httpd.
 - Configuração de IPTABLES para acesso dos nós à WAN.
 - Instalação e configuração do Condor.
 - Instalação dos serviços de monitoramento (Ganglia)
 - Criação de usuários locais.
- Criação de um Cluster de Clusters.
 - Configuração do Condor para execução de “Flocking” dos clusters



Implantação do Middleware

- Inclusão do Cluster de Clusters no Open Science Grid.
 - Instalação do pacote do OSG no servidor central.
 - Obtenção de certificado de grid para o servidor central.
 - Criação do Serviço de Registro da Organização Virtual GridUnesp
 - Emissão de certificado de grid para os usuários qualificados.
 - Registro no OSG enquanto Elemento de Processamento.
 - Registro no OSG enquanto Elemento de Armazenamento.
 - Alocação de pessoal para administração do Grid.
 - Criação do Centro de Operações do GridUnesp.
- Criação do Grid da Unesp, o GridUnesp
 - Formação/treinamento de pessoal para criação do GridUnesp.
 - Implantação de uma Autoridade de Certificação
 - Registro dessa autoridade nos demais grids
 - Criação dos serviços de Alocação de Recursos
 - Criação dos serviços de Gerenciamento dos Dados



Aplicativos de Grid

- Adaptação dos aplicativos para uso em Grids.
 - Criação de envelopes de Grid para os aplicativos convencionais.
 - Criação de Portal de Grid para submissão e monitoramento dos processos
 - Criação de Portal para localização, acesso e informação dos dados.
- Treinamento de usuários.
 - Criação de centros de apoio aos usuários do Grid.
 - Criação de centros de treinamento.
- Formação de pessoal de apoio e desenvolvimento.
 - Criação de cursos de especialização e pós-graduação
 - Desenvolvimento de ferramentas de grid.
- Oferecimento de serviços de grid à sociedade.