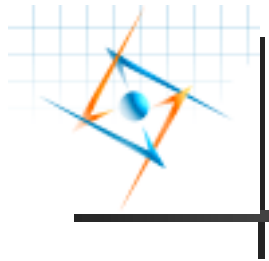


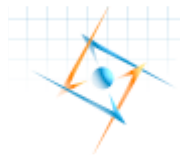
Grid Computing: Processamento de Alto Desempenho em Rede



José Roberto B. Gimenez

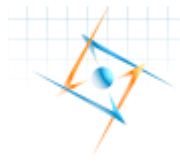
Coordenador de TI do GridUNESP

jr@unesp.br



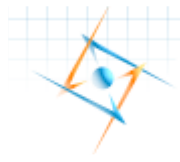
Estrutura da apresentação

- Motivação para uso de Grid
- Discussão sobre problemas em Grid
- Situação Geral da Tecnologia Grid
- Apresentação do SPRACE
- Descrição do Projeto GridUNESP



Por que grid?

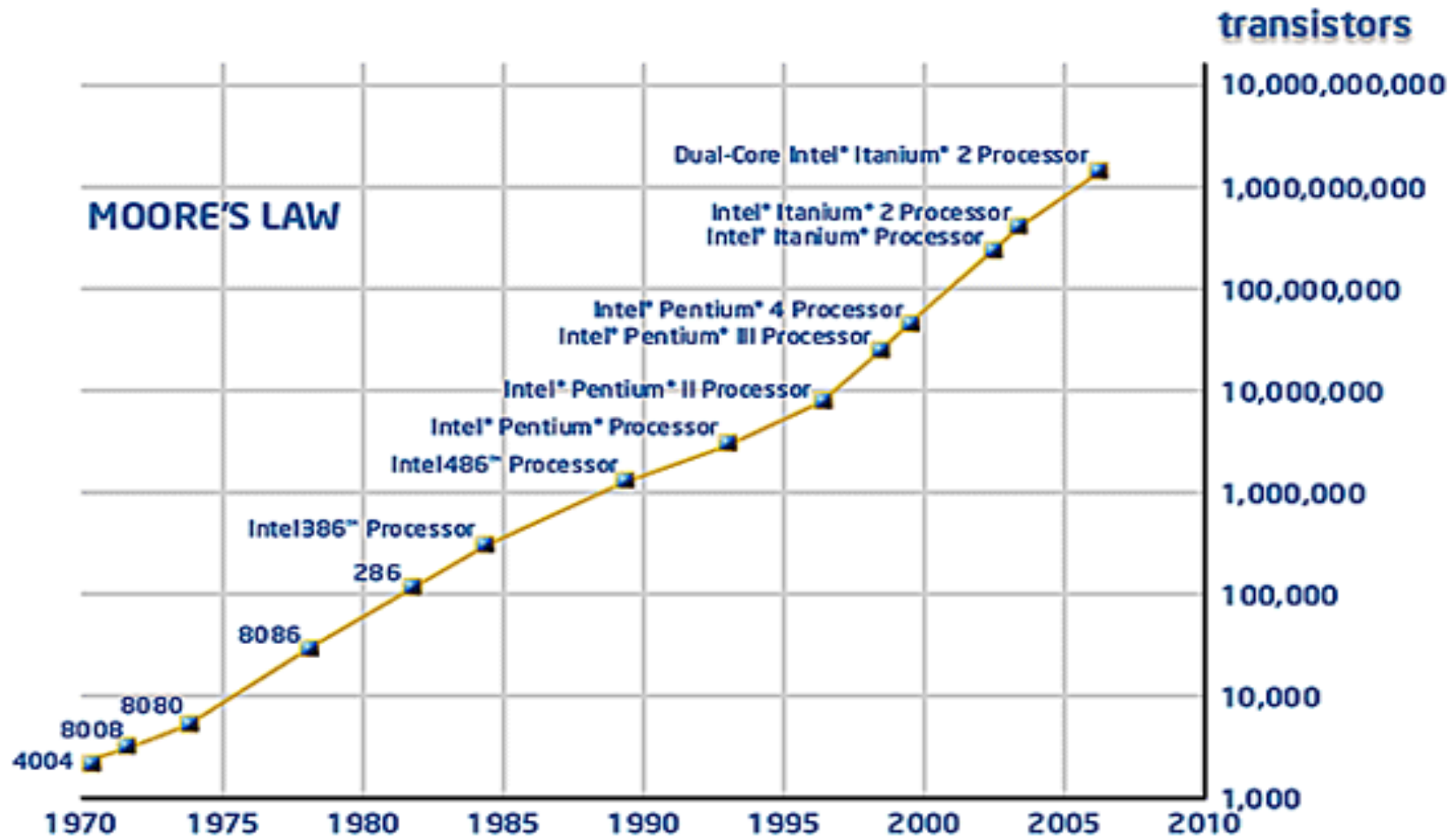
- Existência de problemas complexos
- Motivação para solução breve
 - Ninguém inicia a solução de um problema que termina na próxima vida
 - Previsões precisam ser finalizadas antes do evento ocorrer

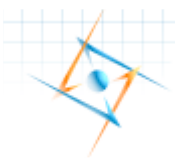


O que limita a velocidade das máquinas?

- Limitação dos sistemas devido a componentes de hardware
 - Evolução da Tecnologia
- Limitação dos sistemas devido a elementos da natureza
 - Velocidade da luz

Tecnologia x Tempo





Tecnologia x Tempo

| | Ano 1994 | Ano 2008 | n° de vezes | Tempo p/ dobrar |
|---------------------|----------|----------|-------------|-----------------|
| Velocidade process. | 50 MHz | 4 GHz | 80 x | 22 meses |
| Espaço de memória | 16 MB | 4 GB | 250 x | 18 meses |
| Memória em disco | 300 MB | 300 GB | 1000 x | 14 meses |
| Conexões LAN | 10 Mb/s | 1 Gb/s | 1000 x | 14 meses |
| Conexões WAN | 64 Kb/s | 34 Mb/s | 500 x | 16 meses |



O Primeiro Disco Rígido

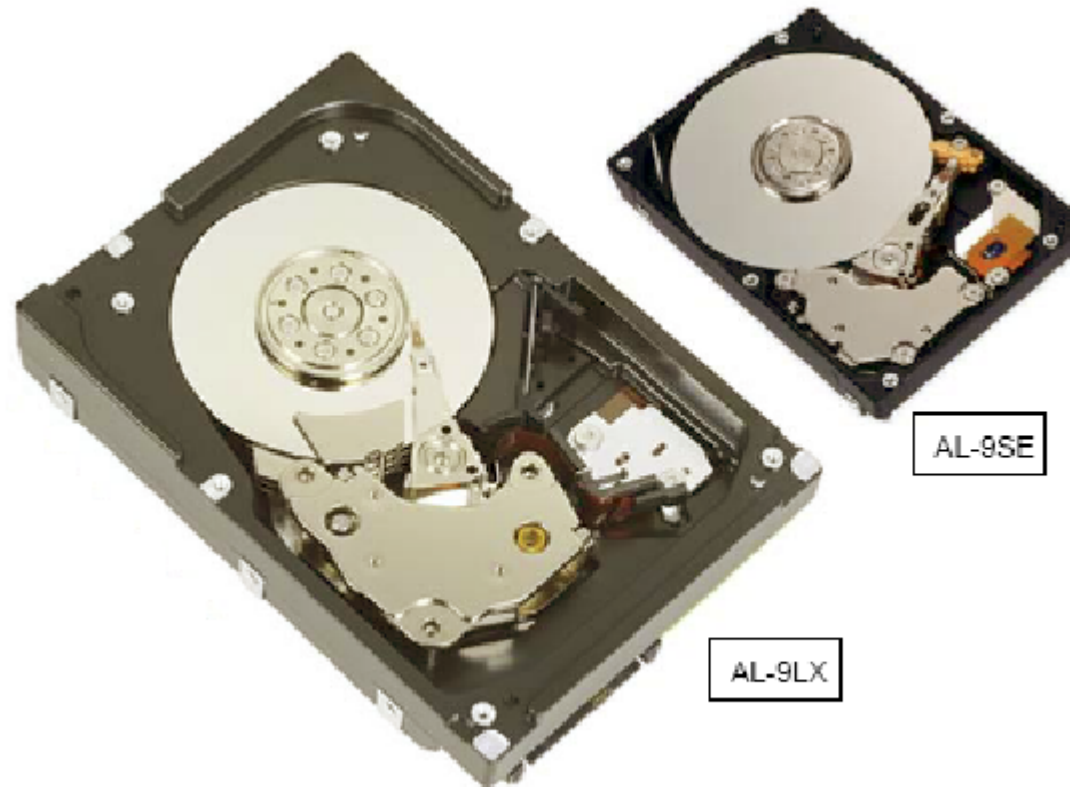
- O disco rígido já tem mais de 50 anos!
- O primeiro disco rígido foi o IBM 305 RAMAC (Random Access Method of Accounting and Control). lançado em 1956.
- Tinha uma capacidade de armazenamento de 5 MB.
- Custava cerca de USD 50,000.





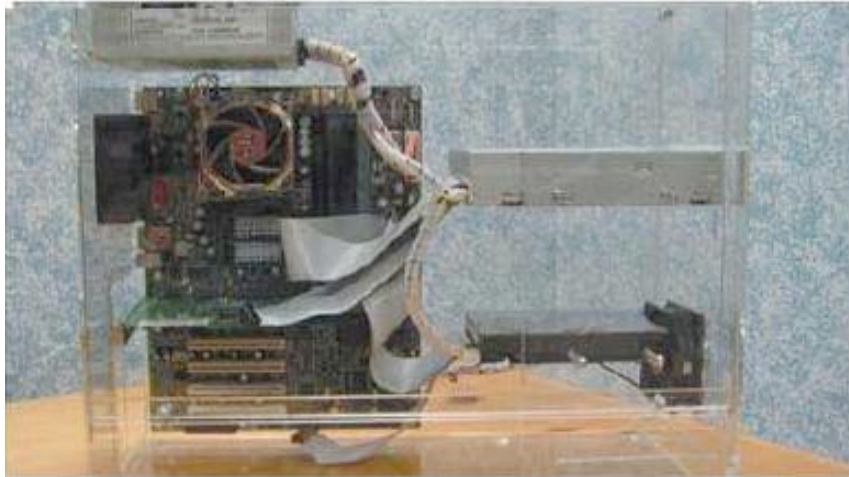
modelos de disco atuais

Formatos de 3,5" e 2,5"





cabo serial e cabo paralelo



Cabo Paralelo



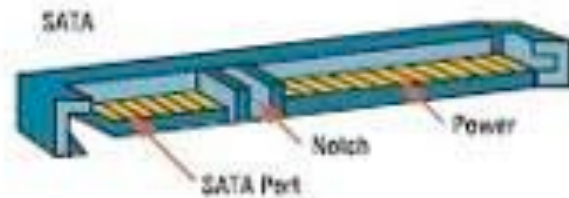
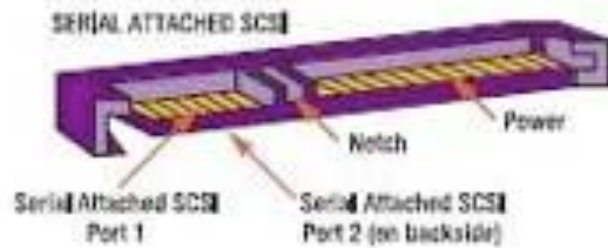
Cabo serial

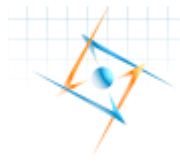




Conectores - SAS

SERIAL ATTACHED SCSI BACKPLANE ACCEPTS SERIAL ATTACHED SCSI AND SATA DISK DRIVES





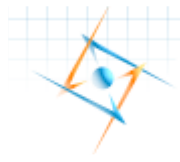
Algumas continhas...

Capacidade armazenada em uma trilha (externa)

$$84 \text{ mm} \times \pi \times 28.500 \text{ b/mm} \\ = 940 \text{ KB}$$

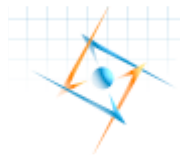
Taxa de transferência para esta mesma trilha

$$84 \text{ mm} \times \pi \times 1025/60 \times 28.500 \text{ b/mm} \\ = 1,25 \text{ Gb/s}$$



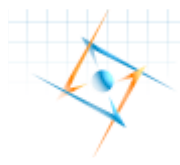
Velocidade da Luz

- Tempo de um ciclo de máquina para um hardware com clock de 4 GHz:
 - 250 pS (pico segundos)
- Distância percorrida por um sinal elétrico em uma trilha de circuito impresso em 250 pS:
 - 5 cm



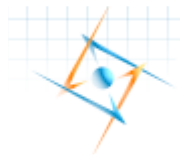
Portanto:

- A solução de problemas de grande complexidade exige a associação de muitas máquinas em paralelo.



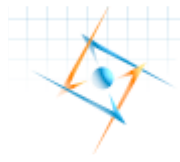
Entretanto:

- Nove grávidas não geram uma criança em um mês.



Requisitos para Processamento em Grid

- Problema paralelizável
- Fluxos de dados independentes
- Altas taxas de transferência
- Necessidade de uma instância de gerenciamento: o middleware.

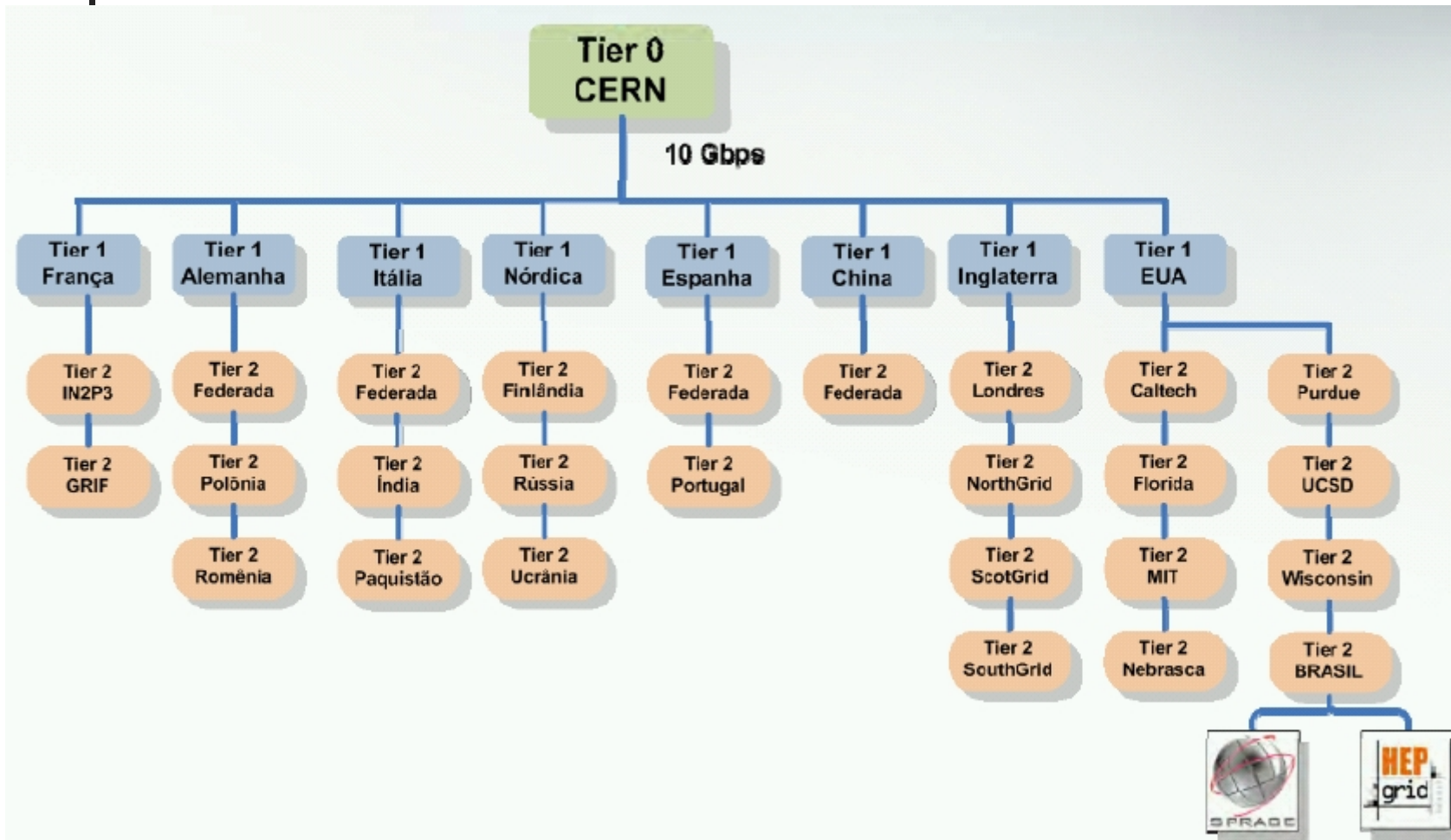


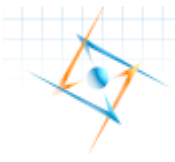
Paradoxos da TI

- Para caber mais informação os dispositivos precisam diminuir de tamanho.
- Canais seriais transferem mais informação que vias paralelas.
- Padrões: sempre existe uma imensidade deles para se escolher.
- Atrasar a decisão de se fazer uma compra resulta em uma compra melhor.



A Hierarquia de Grid





Papel do SPRACE

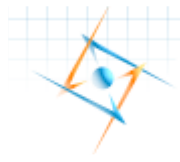
- Membro da Colaboração DØ e CMS,
- Desde março de 2004 o SPRACE (São Paulo Regional Analysis Centre) contribui na produção de Monte Carlo e reprocessamento de eventos da Colaboração DØ.
- Em agosto de 2005 ingressou no OSG (Open Science Grid). Também participa na iniciativa de grid do LHC, preparando-se para as atividades do CMS, que se iniciam este ano.
- <http://hep.ift.unesp.br/SPRACE/>

SPRACE



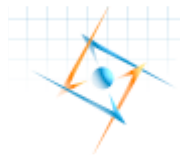
- Tier 2 do OSG / CMS
 - 240 cores (~400 kSI2K)
 - 12 TB de memória em disco
 - Totalmente dedicado às colaborações Dzero e CMS

| | fase 1 (2004) | Phase 2 (2005) | Phase 3 (2006) | Phase 4 (2008-10) planned |
|----------------------|--------------------------|---------------------------|---------------------------|--|
| CPUs (cores) | 50 | 116 | 240 | > 500 |
| Power (kSI2K) | 40 | 132 | 400 | > 1,200 |
| Memória (TB) | 4 | 12 | 12 | > 200 TB |



O Projeto GridUNESP

- Primeiro Campus Grid universitário em andamento no Brasil.
- Objetivo: atender à necessidade de equipes de pesquisa que demandam processamento de alto desempenho.



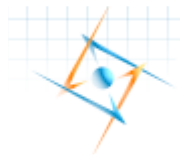
Linhas de Pesquisa

- Física de Altas Energias
- Biologia Computacional
- Química Quântica
- Fluidodinâmica Computacional
- Genômica
- Supercondutividade
- Geologia
- Segurança de Redes



Filosofia do GridUNESP

- Clusters
 - Um Cluster Central
 - 7 Cluster Periféricos
- Baseado em tecnologia CISC X86
- Componentes de mercado (commodities)
- Rede Infiniband (10 Gbb/s full-duplex)
- Processadores com 4 núcleos (quad-core)
- Interfaces SAS e SATA e FC (3 – 4 Gb/s)



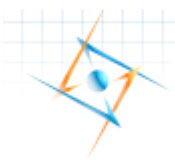
Unesp – Universidade Estadual Paulista

- Presente em 23 cidades do Estado de São Paulo.
- Cerca de 40 campi universitários
- Perfil apropriado para Grid Computing
- Distribuição balanceada de recursos
- Acesso estratégico aos avanços em TI

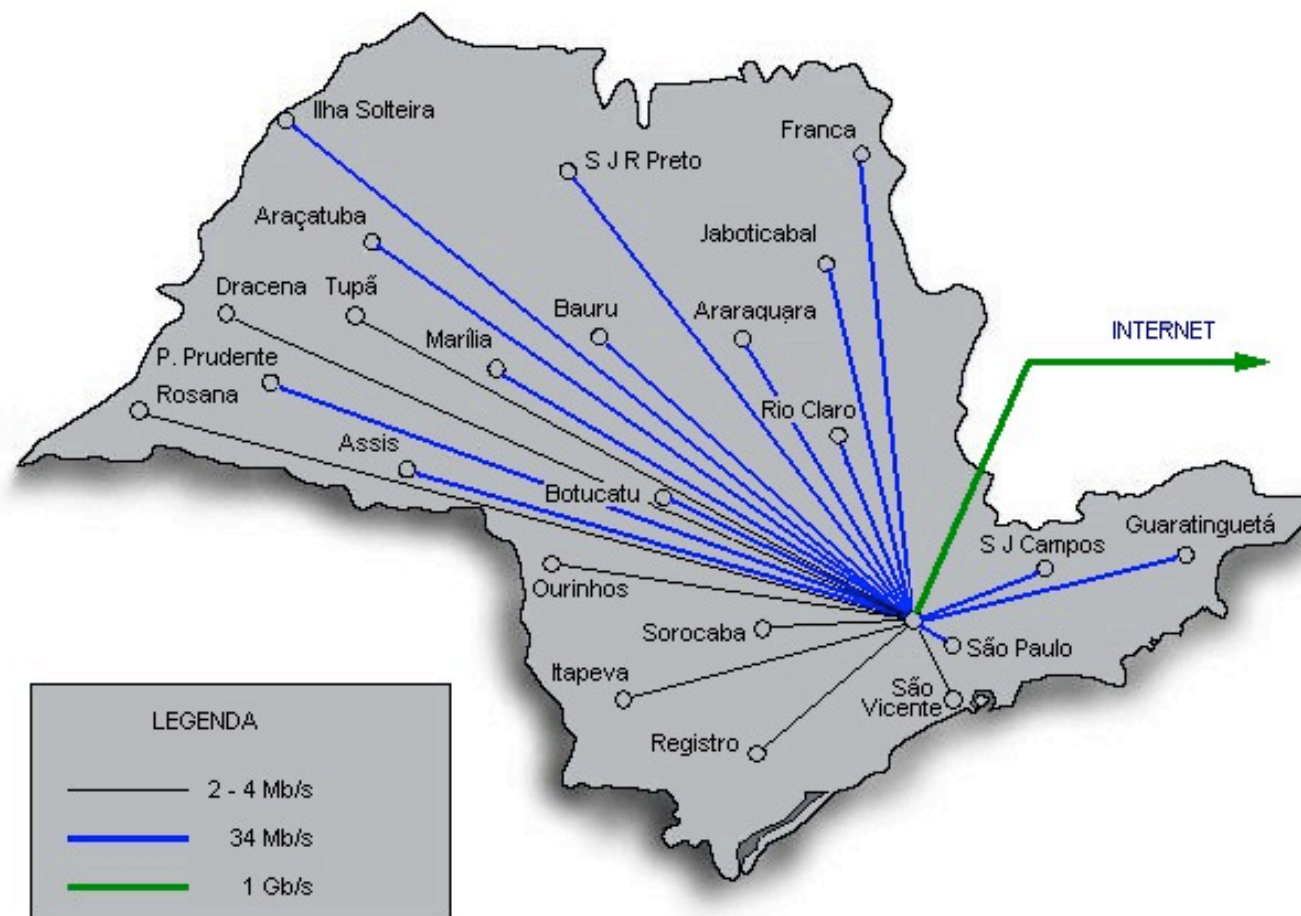


Localização dos Clusters





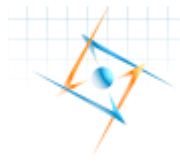
unespNET – Versão Atual – 2006/2007





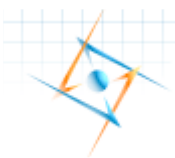
Rede KyaTera





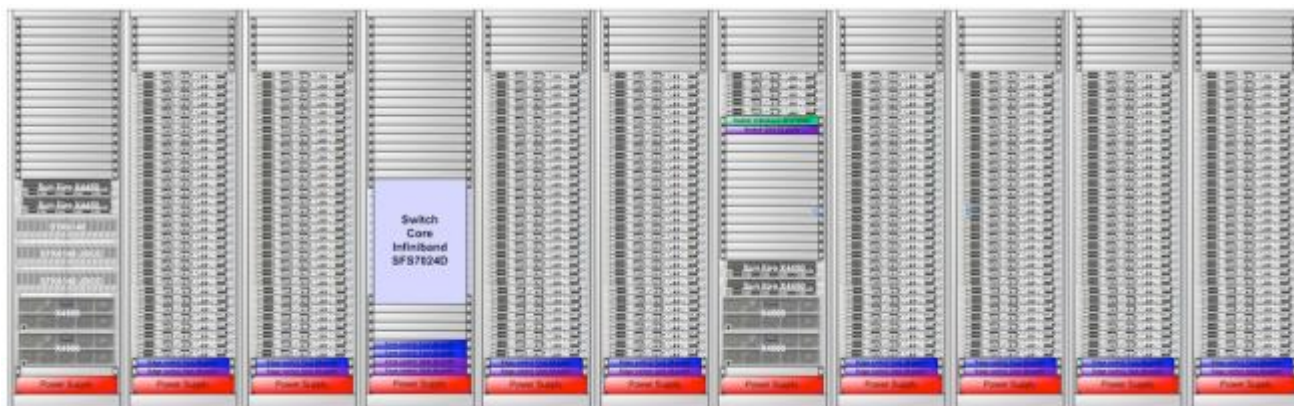
Equipe de Coordenação

- Sergio F. Novaes (Coordenador Geral)
- Ney Lemke (Coordenador Científico)
- Eduardo M. Gregores (Coordenador Técnico)
- José Roberto B. Gimenez (Coord. de TI)



GridUnesp - Hardware

Cluster Central



Clusters periféricos

7 X



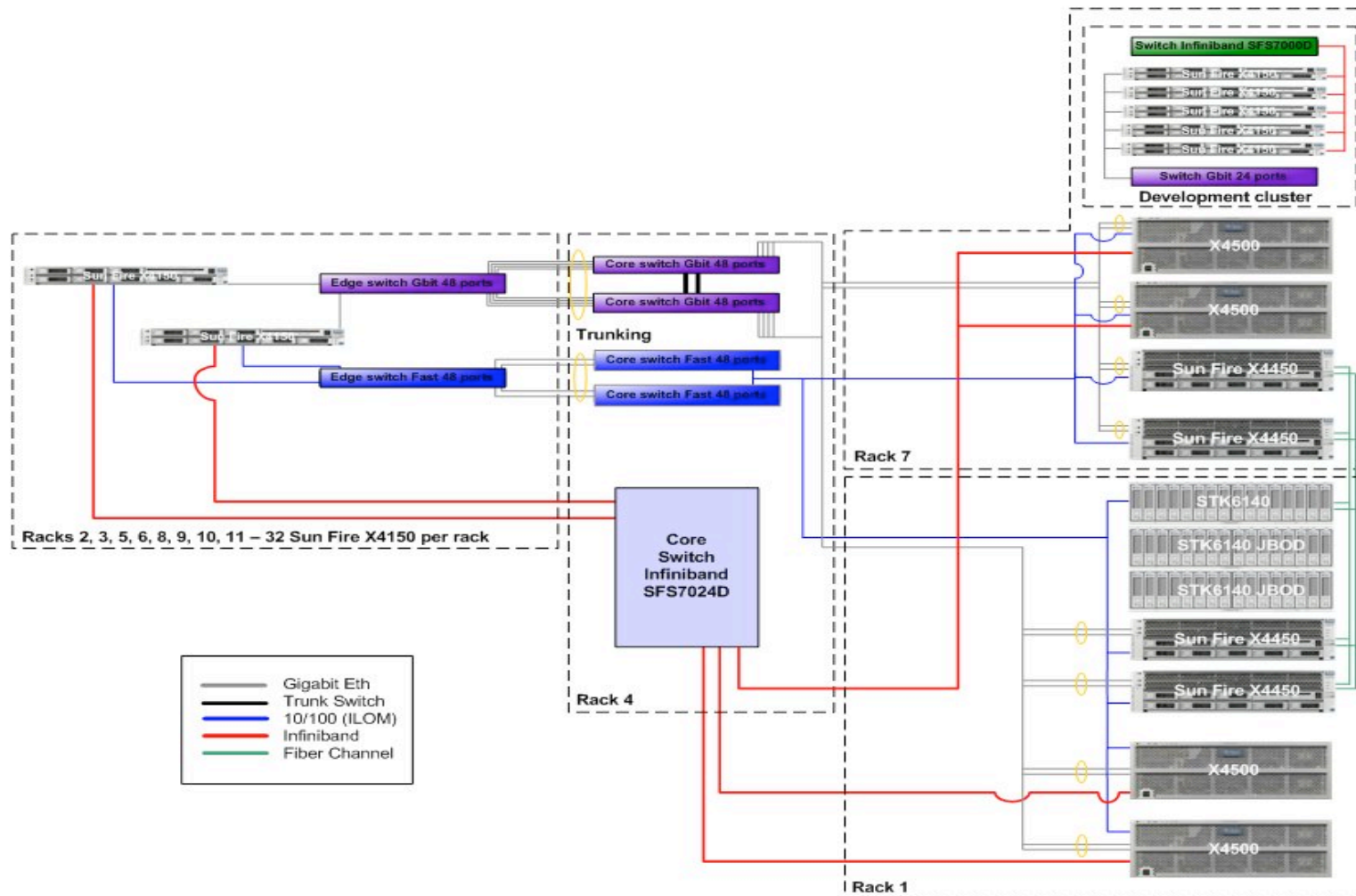
| | Cluster central | Clusters periféricos | Total |
|----------------------|------------------------|-----------------------------|--------------|
| CPUs (cores) | 2,048 | 896 | 2,944 |
| Power (kSI2K) | 6,100 | 2,600 | 8,700 |
| Storage (TB) | 150 | 92 | 242 |

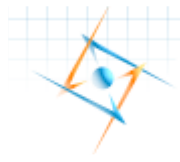


Cluster Central - Características

- 256 work nodes - dual quad-core Intel Xeon (E5440 Harpertown 2.83GHz, 12MB L2 Cache), 16 GB mem
- 4 head nodes: quad quad-core Intel Xeon (E5440 Harpertown 2.83GHz, 12MB L2 Cache), 32 GB mem
- 4 hybrid data servers (Sun Fire X4500), 24 TB each
- Storage area network (fibre-channel 4 Gbps), 36 TB
- Local area network: 10 x Cisco Catalyst 2960
- High speed interconnect: Cisco SFS 7024D 4 x DDR InfiniBand server switch (288 ports)
- Out-of-band management network
- 2 console/admin workstations (Sun Ultra 24)

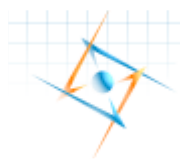
Cluster Central – Esquema de Conexões





Clusters Periféricos

- 16 nós: dual quad-core Intel Xeon processors (E5440 Harpertown 2.83GHz, 12MB L2 Cache), 16 GB mem
- 2 head-nodes: dual quad-core Intel Xeon processors (E5440 Harpertown 2.83GHz, 12MB L2 Cache), 16 GB mem
- SAN: (fibre-channel 4 Gbps), 12 TB
- LAN: GigabitEthernet (Cisco Catalyst 2960)

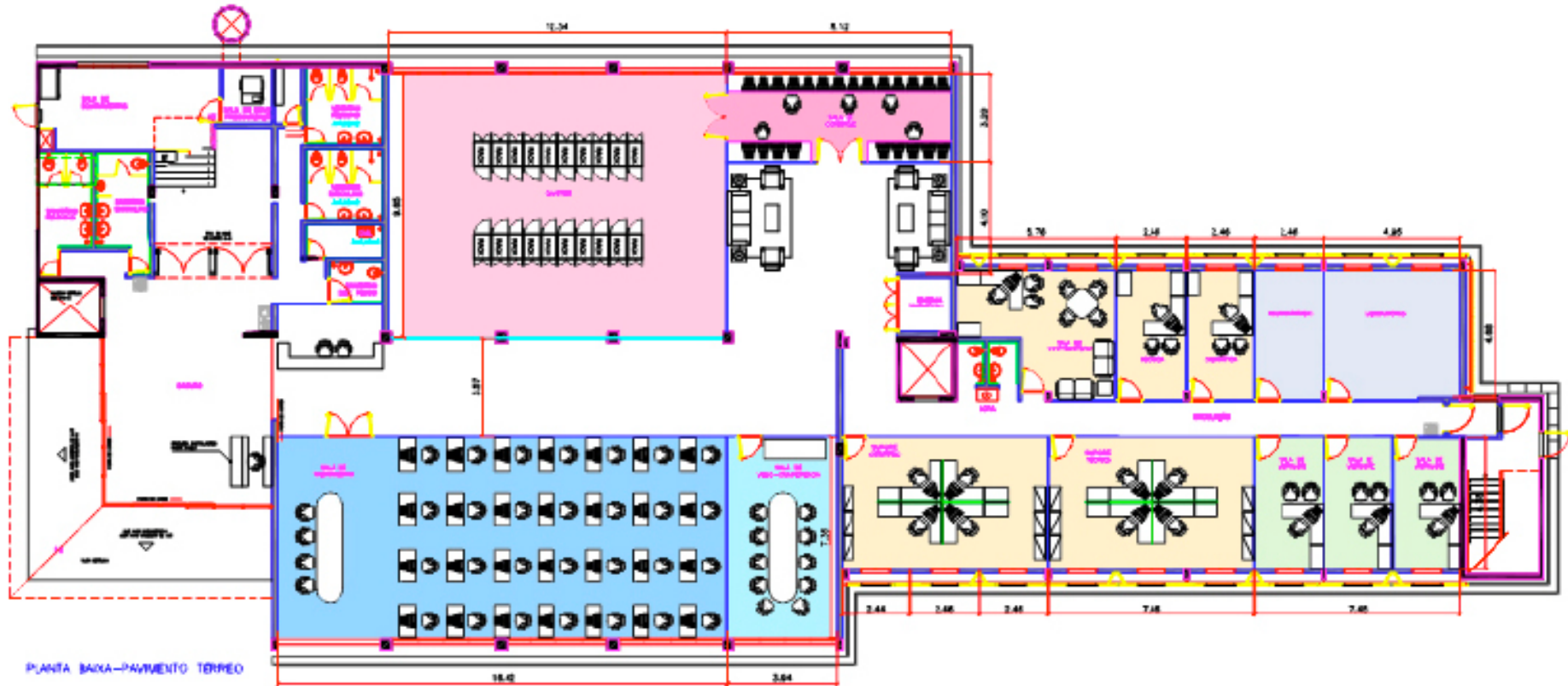


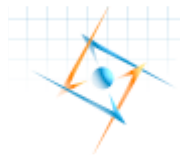
Treinamento

- Sala de Treinamento
 - 20 estações SUN - processadores Intel Core 2 quad (Q6600 Kentsfield 2.40 GHz, 4MB L2 Cache), 02 GB mem, 19" monitor de LCD
- Cluster de Desenvolvimento
 - 4 servers - dual quad-core Intel Xeon (E5335 2.00GHz, 12MB L2 Cache), 08 GB mem (32 cores)
 - 1 head-node dual quad-core Intel Xeon (E5335 2.00GHz, 12MB L2 Cache), 08 GB mem (8 cores)
 - Rede Infiniband: Switch Cisco SFS 7000D 4 x DDR 24 portas



Instalações do GridUNESP





Mais informações

- <http://unesp.br/gridunesp>
- <http://www.sprace.org.br>
- <http://www.sprace.org.br/twiki/bin/view/Main/Gridunesp>