

SEARCH Cluster

Universidade do Minho
Novembro, 2007

Ricardo Marques - pgi.0898
ricmarques@gmail.com

SEARCH – Vista Geral

- Alguns números
 - 46 nós de computação com 1U de altura
 - 92 processadores Xeon a 3.2 Ghz
 - 92 GB de Ram
 - 8 processadores Nvidia 7900GT
 - 920 Gbps Myrinet 10Gb
 - 92 Gbps GbE Ethernet
 - 3.0 TB de armazenamento em SAN
 - 3.7 TB de armazenamento em disco local
 - 17.5 kW de consumo de energia eléctrica
 - 59.5 kBTU/h de calor gerado
 - 1.5 toneladas

Universidade do Minho
Novembro, 2007

Ricardo Marques - pgi.0898
ricmarques@gmail.com

www.di.uminho.pt/search/equipamento.htm

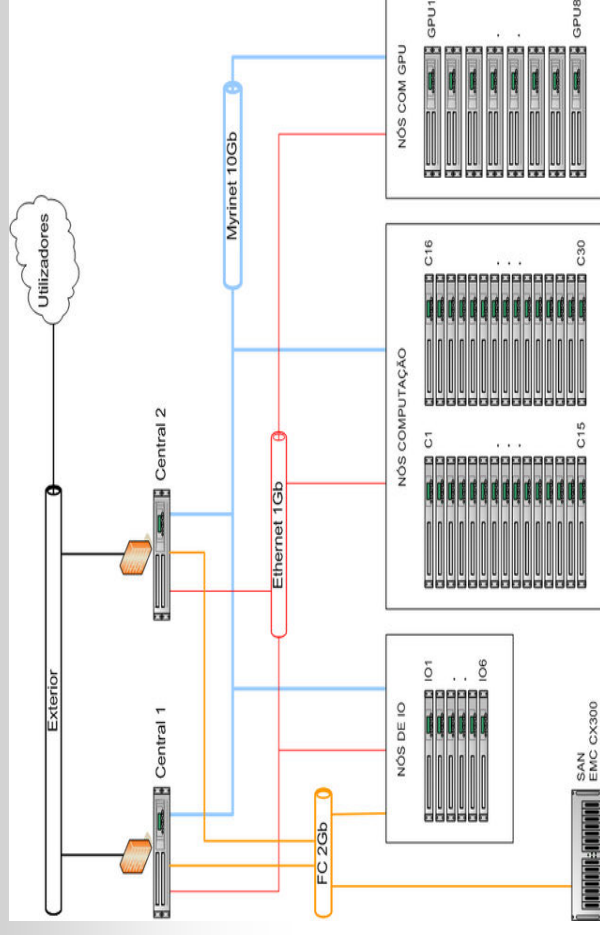
SEARCH – Vista Geral

- Composto por:
 - 46 nós
 - 2 centrais
 - 30 genéricos
 - 8 com GPU
 - 6 com I/O
 - sistema de armazenamento (SAN) ligado com um switch fibre channel
 - dois sistemas de inteligência
 - Myrinet 10 Gb
 - Ethernet 1 Gb

Universidade do Minho
Novembro, 2007

Ricardo Marques - pgi.0898
ricmarques@gmail.com

SEARCH – Vista Geral



Universidade do Minho
Novembro, 2007

Ricardo Marques - pgi.0898
ricmarques@gmail.com

Nós – Vista Geral

- assentam em plataformas com 1U de altura
- são compostos por (excepto os nós com GPU)
 - motherboard baseada no chipset Intel E7320
 - dois processadores Intel Xeon de 64 bits
 - 2 GB de memória RAM DDR2-400
 - discos locais em todos os nós
 - 160 GB em RAID-1 nos nós centrais
 - 80 GB nos restantes nós

Universidade do Minho
Novembro, 2007

Ricardo Marques - pgl.0898
ricmarques@gmail.com

Nós - Processador

- Pentium Xeon (baseado no pentium 4)
 - 3.2 GHz
 - 2 MB cache (L2)
 - 800 MHz bus speed (FSB)

Universidade do Minho
Novembro, 2007

Ricardo Marques - pgl.0898
ricmarques@gmail.com

Nós - Motherboard

- Motherboard de Servidor Intel SE7501WV2
 - baseada no chipset Intel E7320
 - suporta um ou dois processadores Intel Xeon
 - até 12 GB de suporte a memória DDR – no caso do SEARCH, estão equipadas com 2 GB
 - duplo 800 MHz system bus

Universidade do Minho
Novembro, 2007

Ricardo Marques - pgl.0898
ricmarques@gmail.com

Nós - Motherboard



Processor, server board, memory, and hard drives
not included with chassis

Meticulous design, and thorough validation, and extensive testing can mean less development work, higher quality, and faster time-to-market.

1. Two 64-bit/100MHz PCI/PCI-X riser cards
2. Five instrumented system-cooling fans
3. Chassis-intrusion and server-management capabilities
4. Enhanced front panel with video and USB ports
5. One flexible drive bay for hot-swap SCSI hard drive or slim-line CD-ROM/floppy or DVD/floppy module
6. Two hot-swap SCSI or two cold-swap ATA hard-drive bays
7. Single 350W PFC power supply

Universidade do Minho
Novembro, 2007

Ricardo Marques - pgl.0898
ricmarques@gmail.com

Nós - Especificidades

- Nós com GPU
- motherboard
- processador
- Nós centrais
- Interface Fibre Channel
- DVD-ROM
- Disco
- Nós com I/O
- Interface Fibre Channel

Universidade do Minho
Novembro, 2007

Ricardo Marques - pgl0898
ricmarques@gmail.com

Nós com GPU - Motherboard

- Motherboard de Servidor Intel S5000PAL
- baseada no chipset Intel 5000
- suporta um ou dois processadores Intel Xeon Dual-Core
- até 32 GB de suporte a memória DDR – no caso do SEARCH, estão equipadas com 4 GB
- duplo 1333 MHz data bus (FSB)

Universidade do Minho
Novembro, 2007

Ricardo Marques - pgl0898
ricmarques@gmail.com

Nós com GPU - Motherboard



Universidade do Minho
Novembro, 2007

Ricardo Marques - pgl0898
ricmarques@gmail.com

Nós com GPU - Processador

- Intel Xeon Dual Core 5130
- 2 GHz
- 128 KB cache L1
- 4 MB cache L2
- 1333MHz FSB

Universidade do Minho
Novembro, 2007

Ricardo Marques - pgl0898
ricmarques@gmail.com

Nós com GPU - GPU!

- GPU Nvidia 7900GT
- 256MB DDR3
- bom desempenho em cálculo matricial

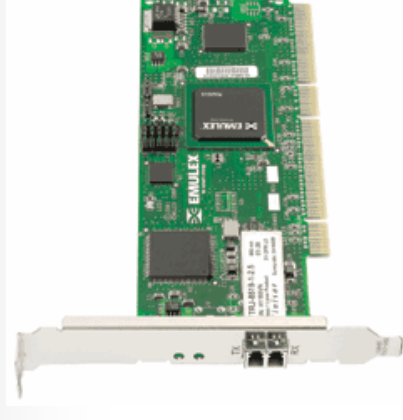


Universidade do Minho
Novembro, 2007

Ricardo Marques - pgi.0898
ricmarques@gmail.com

Nós com I/O

- Interface FibreChannel (HBA) Emulex LP-982E (PCI-X 66)



Universidade do Minho
Novembro, 2007

Ricardo Marques - pgi.0898
ricmarques@gmail.com

Sistema de Comunicações

- Ethernet 1 Gb
- Myrinet 10 Gb
- Fibre Channel – Storage Area Network (SAN)

Universidade do Minho
Novembro, 2007

Ricardo Marques - pgi.0898
ricmarques@gmail.com

Interligações - Ethernet

- Não satisfaz totalmente as necessidades de desempenho requeridas por clusters (HPC e HTC)
- No entanto:
 - Usa um protocolo bem estabelecido (ao contrário da Myrinet) que pode ser usado para comunicações mais leves
 - Útil como sistema de comunicação alternativo que sobre o qual operam sistemas de monitorização do sistema

Universidade do Minho
Novembro, 2007

Ricardo Marques - pgi.0898
ricmarques@gmail.com

Interligações - Ethernet

- Sobre esta conexão opera um sistema de monitorização remoto baseado no protocolo IPMI
 - Define mensagens e o interface do sistema para o controlo do hardware da plataforma
 - Permite o controlo da “saúde” do sistema
 - Temperatura, voltagem, necessidades de energia, etc
 - Possui mecanismos que permitem reduzir o custo de manutenção da plataforma
 - Automatic shutdown and restart, power control capabilities

Universidade do Minho
Novembro, 2007

Ricardo Marques - pgi.0898
ricmarques@gmail.com

Interligações Ethernet - Topologia

- Em termos físicos, os nós estão ligados em estrela
 - Todos os nós estão ligados entre si através de um nó central – switch HP ProCurve 2848, com 48 portas
- Em termos lógicos, trata-se de uma topologia de barramento
 - Cada nó limita-se a lançar o pacote na rede esperando que o receptor o recolha
 - A otimizar esta topologia está o switch, que encaminha o pacote para o seu destinatário, aumentando a disponibilidade do sistema



Universidade do Minho
Novembro, 2007

Ricardo Marques - pgi.0898
ricmarques@gmail.com

Interligações - Myrinet

- “Myrinet is a cost-effective, high-performance, packet-communication and switching technology that is widely used to interconnect **clusters** of workstations, PCs, servers, blade servers, or single-board computers.” -

www.myri.com

- Conexão 10 Gb full duplex
- Usa o protocolo MPICH-MX

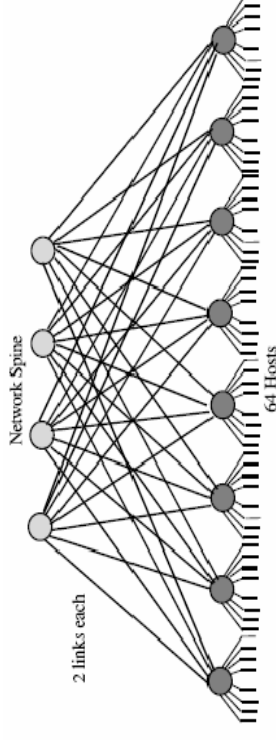


Universidade do Minho
Novembro, 2007

Ricardo Marques - pgi.0898
ricmarques@gmail.com

Myrinet - Topologia

- Conecta todos os nós através de uma topologia de rede CLOS (expansível a 128 portas)
 - 4 spine switches
 - 6 módulos de 8 portas (total 48 portas)

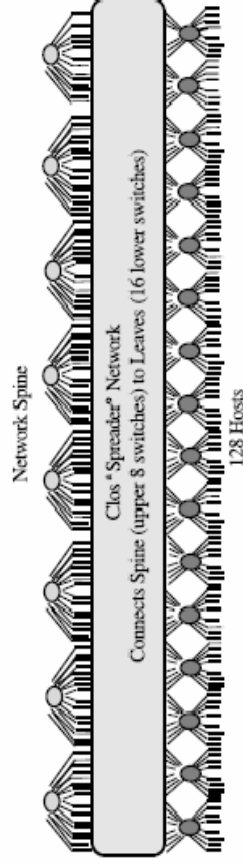


Universidade do Minho
Novembro, 2007

Ricardo Marques - pgi.0898
ricmarques@gmail.com

Myrinet - Topologia

- No caso de se concretizar a expansão a 128 portas
 - 16 módulos de 8 portas
 - 8 spine switches



Universidade do Minho
Novembro, 2007

Ricardo Marques - pgl.0898
ricmarques@gmail.com

Sistema de Armazenamento



- EMC CLARiON CX300
 - 50K IOPS
 - 698 MB/s transfer rate
 - 3 TB de espaço de armazenamento (estensível a 27 TB)
 - 12 discos
 - 2 Gb Fibre Channel technology
 - Ligado a switch fibre channel de 16 portas
 - “based on a common high-availability design”

Universidade do Minho
Novembro, 2007

Ricardo Marques - pgl.0898
ricmarques@gmail.com

Interligações – SAN

- SAN – 4 portas de I/O
- Switch Fibre Channel
 - Conecta os 6 nós de I/O a 2 portas da SAN
- Os 2 nós centrais estão conectados directamente à SAN

Universidade do Minho
Novembro, 2007

Ricardo Marques - pgl.0898
ricmarques@gmail.com

Referências

- www.di.uminho.pt/search
- www.intel.com
- Intel Server Board S5000PAL /S5000XAL Technical Product Specification
- www.myri.com
- www.emc.com
- www.siber.cankaya.edu.tr – Myrinet & CLOS

Universidade do Minho
Novembro, 2007

Ricardo Marques - pgl.0898
ricmarques@gmail.com

Fotos do instalação do cluster



Universidade do Minho
Novembro, 2007

[Ricardo Marques - pgi0898
ricimmarques@gmail.com](mailto:Ricardo.Marques - pgi0898@icimmarques@gmail.com)

Fotos do instalação do cluster



Universidade do Minho
Novembro, 2007

[Ricardo Marques - pgi0898
ricimmarques@gmail.com](mailto:Ricardo.Marques - pgi0898@icimmarques@gmail.com)