

Supercomputador Pleiades

Arquitetura de Computadores – ACH 2055



Pleiades

SGI Altix ICE 8200EX

51,200 total cores -

Intel Xeon processors (quad-core)



- ❑ Sexto computador mais rápido do planeta, conforme o TOP500
- ❑ É um projeto em andamento, há perspectivas de incremento, devido à disponibilidade de espaço para a adição de novos gabinetes
- ❑ Provê recursos computacionais para apoiar as missões críticas da NASA

Melina Brilhadori Alves

Murilo Galvão Honório

Junho de 2010

CARACTERÍSTICAS BÁSICAS

Nome	Pleiades
Localização	NASA / Centro de Pesquisa Ames / NAS
Família	SGI Altix
Modelo	SGI Altix ICE 8200EX/8400EX
Computadores	SGI Altix ICE 8200EX/8400EX, Xeon HT QuadCore 3.0 GHz / Xeon Westmere 2.93 GHz
Vendedor	SGI
Área de aplicação	Pesquisa
Memória principal	126970 GB
Ano de instalação	2010
Sistema Operacional	Novell SuSe Linux Enterprise Edition Pro Pack + SGI5
Interconexão	Infiniband
Processador	Intel Xeon EM64T E54xx (Harpertown) 3000 MHz (12 Glops)

ORGANIZAÇÃO E ARQUITETURA

- ❑ Sistema SGI Altix ICE, fabricado pela Silicon Graphics, dotado de processadores Intel rodando cerca de 544 trilhões de operações de ponto flutuante por segundo.
- ❑ A maior interconexão de rede InfiniBand do mundo, com cerca de 9 Km de cabos.
- ❑ Um dos supercomputadores mais eficientes em termos de consumo de energia.
- ❑ Vem sendo expandido de forma incremental.

Nro de Gabinetes	Processador	Total de Cores	Memória	Data da instalação
32	Intel 6-core Xeon 2.93 GHz X5670 Westmere	24.576	2GB por core	Maio/2010
20	Intel quad-core Xeon 2.93 GHz X5570 Nehalem-EP	10.240	3GB por core	Entre Novembro/2009 e Março/2010
92	Intel quad-core Xeon 3.0 GHz E5472 Harperstown	47.104	1GB por core	Setembro/2008



ORGANIZAÇÃO E ARQUITETURA

- ❑ Organizado na forma de cluster, implementando a arquitetura de servidores blade, configuração básica secundário ativo, pois todos os nós são usados para processamento.
- ❑ Organizado segundo a abordagem de discos compartilhados, de forma que cada nó tem acesso a todos os volumes de todos os discos.
- ❑ Os processadores são multicore, o que beneficia a velocidade de execução quando propicia o processamento paralelo de instruções.
- ❑ O acesso aos recursos do cluster é efetuado através de 8 nós frontais. O sistema possui um mecanismo de balanceamento de carga que, durante o login, seleciona automaticamente o nó frontal menos carregado.
- ❑ Há um robusto mecanismo de failover e restauração, que torna as falhas de servidores e reinicializações transparentes aos clientes. São possíveis atualizações e reinicializações enquanto jobs ativos continuam rodando, meramente havendo atraso. Para a realização de failover, as máquinas possuem o software Linux FailSafe, que provê redundância de processadores e controladoras de I/O.
- ❑ A interconexão é através de rede Infiniband com conexão bidirecional ponto-a-ponto, suportando taxas de transferência duplas e quádruplas, sendo que há duas topologias distintas. O gerenciamento de rede é Gigabit Ethernet.

ORGANIZAÇÃO E ARQUITETURA

- ❑ Sistema de arquivos local Nexis 9000, trabalhando sobre Infiniband escala até 1 GB/s na escrita e 3 GB/s na leitura.
- ❑ A infraestrutura de armazenamento é DDN9900 (combinação de SSD, SAS e SATA), com 3 PB ao todo, paridade RAID6.
- ❑ Há cinco sistemas de arquivos Lustre montados, com tamanho total de 2400 TB.
- ❑ Aplicações clientes enxergam um único sistema de arquivos unificado, embora ele seja composto por milhares de servidores individuais e uma combinação sistemas de arquivos MDT/OST.
- ❑ A distribuição base de sistema operacional instalada é o SUSE Linux Enterprise Server 10 (SLES10), da empresa Novell.
- ❑ Sistema desenvolvido para ser capaz de executar facilmente qualquer código convencional linux com pequenas modificações.
- ❑ Login através de 8 nós frontais, permite também operações remotas mediante obtenção de chave de acesso (SUP – Secure Unattended Proxy).



FERRAMENTAS COMPUTACIONAIS

- As aplicações podem rodar de forma independente umas das outras, pois os servidores do cluster fazem numerosas transações e execuções em paralelo.
- Middleware SGI ProPack6 SP3 que provê recursos para administração dos servidores, aplicações paralelas e mecanismos de balanceamento de carga.
- Os jobs submetidos ao sistema são gerenciados pelo escalonador de jobs PBS (Portable Batch System).
- Há sistema de contabilização de horas utilizadas.
- Possui diversos compiladores (Intel C, fortran, gcc), pacotes de otimização para operações concorrentes cooperativas, sincronização, troca de mensagens e computação paralela (MPT, MPI, SHMEM) e bibliotecas de funções matemáticas complexas (BLAS, Linpack).
- Executa aplicações para dinâmica de fluídos, modelos climáticos complexos, microgravidade, modelagem matemática, design aerodinâmico, estimativas de trajetórias, entre outras.

