

# SIRC 2012

XI SIMPÓSIO DE INFORMÁTICA DA UNIFRA

[www.sirc.unifra.br](http://www.sirc.unifra.br)

23 a 25 de Outubro de 2012

Realização:

Apoio:

Patrocínio:



## Aumento da Capacidade de Processamento de um Cluster de Computadores

Eduardo C. Biassus<sup>1</sup>, Henrique M. Gasparotto<sup>2</sup>, Adriano Petry<sup>3</sup>, Haroldo Fraga de Campos Velho<sup>4</sup>

<sup>1</sup>Curso de Ciência da Computação – Centro Universitário Franciscano (UNIFRA)  
Rua dos Andradas, 1614 – Santa Maria – RS – Brasil

<sup>2</sup>Curso de Engenharia de Computação – Universidade Federal de Santa Maria (UFSM)  
Av. Roraima, s/n – Santa Maria – RS – Brasil

<sup>3</sup>Centro Regional Sul de Pesquisas Espaciais (CRS/INPE)  
Av. Roraima, s/n, Prédio do INPE – Santa Maria – RS – Brasil

<sup>4</sup>Laboratório Associado de Computação e Matemática Aplicada (LAC/INPE)  
Av. dos Astronautas, 1758, São José dos Campos – SP – Brasil

eduardo.bi@hotmail.com, hmgasparotto@hotmail.com,  
adriano.petry@crs.inpe.br, haroldo@lac.inpe.br

**Abstract.** *This work presents the activities performed to increase the processing capacity of an existing computer cluster. Several processing nodes were available to this task, which consisted of hardware analysis, software installation, internal network configuration, and resources inclusion in resource manager database. At last, functioning tests were run in the new cluster, and the jobs execution capacity improvement was verified.*

**Resumo.** *Esse trabalho apresenta as atividades desenvolvidas para o aumento da capacidade de processamento de um cluster de computadores existente. Diversos nós de processamento foram disponibilizados para essa tarefa, que consistiu na análise do hardware, instalação de software, configuração da rede interna do cluster, e inclusão dos recursos no banco de dados do gerenciador de recursos utilizado. Ao final, testes de funcionamento foram executados no novo cluster, que demonstraram o aumento em sua capacidade de execução de tarefas.*

### 1. Introdução

A Computação de Alta Performance (HPC) [Cheung, A.L 1992][Kuck, D. J. 1996] abrange, dentre outras coisas, o trabalho cooperativo de diversos computadores conectados em um cluster, a fim de que a execução paralelizada de programas aconteça no menor tempo possível. O Laboratório de Computação para Clima Espacial (LCCE) do Centro Regional Sul de Pesquisas Espaciais (CRS/INPE) realiza diversas simulações que exigem grande poder de processamento para serem executadas em tempo hábil. Este trabalho relata o aumento na capacidade de processamento de um cluster de computadores existente, a partir de equipamentos ociosos recebidos pelo grupo de trabalho.

## **2. Desenvolvimento**

O LCCE dispõe de um cluster em pleno funcionamento, usado para fazer a previsão da dinâmica da ionosfera terrestre [Petry et al. 2012]. Recentemente foram recebidos novos nós de processamento para serem adicionados a esse cluster. Esses equipamentos passaram por verificações de integridade, de desempenho e configurações de hardware. A partir dessa análise do hardware recebido, foram feitas as instalações e configurações da rede e dos diversos softwares necessários ao bom funcionamento dos novos nós. A última etapa realizada foi a integração desses nós ao gerenciador de recursos OAR [Nicolas e Joseph 2011][Machado, Cargnelutti e Petry, 2012] para utilização na previsão realizada.

### **2.1. Cluster disponível**

O cluster inicialmente contava com 7 nós de processamento e 2 nós para armazenamento de dados. Esses nós de processamento podem ser divididos em dois grupos, de acordo com a sua configuração de hardware: no primeiro grupo, existem 2 nós de processamento, cada um contando com 2 processadores Intel Xeon 2.93 GHz, 24 núcleos de processamento e 32 GB de memória RAM; o segundo grupo consiste nos 5 nós de processamento restantes, onde cada um apresenta 2 processadores Intel Xeon 2.53 GHz, 16 núcleos de processamento e 16 GB de memória RAM. Os dois nós restantes possuem 3.6 TB para armazenamento de dados.

### **2.2. Avaliação do hardware disponível**

Inicialmente foi feito um levantamento das peças, equipamentos, e cabos disponíveis para atualização do cluster existente. Com isso, foi constatada a existência de outros 23 nós de processamento, além de vários cabos de redes e de energia. Tratam-se de nós HP Proaliant dl 145 G2, que possuem 2 processadores dual core AMD Opteron 275 trabalhando a frequência de 2.2 Ghz, contando com 1 MB de memória cache, 2 GB de memória principal e 2 discos rígidos (HDs) com capacidade de 80 GB cada.

Como o equipamento não era novo, ele foi inicialmente limpo e a integridade das peças foi verificada. Constataram-se vários problemas com as placas de rede deles, sendo tais problemas solucionados com uma limpeza utilizando uma simples borracha. Um dos nós apresentou problemas na fonte de alimentação e outro na placa mãe. Então foi aproveitada a fonte do segundo no primeiro. Com isso, 22 nós completos foram disponibilizados, e 4 placas de memória de 512 MB, 2 HDs e 2 processadores ficaram de reserva para eventuais substituições que se façam necessárias no futuro.

### **2.3. Instalações de software e configurações de rede**

Foi realizada a instalação do sistema operacional CentOS versão 6.3 64bits [Morimoto 2008] nos nós utilizados na atualização do cluster, o qual se adapta muito bem para as necessidades, além de ser o sistema operacional livre e já em uso no cluster. Os HDs foram configurados utilizando redundância tipo RAID 10 [Alecrim 2012] via software, com um espaço subdividido em 10Gb (/), 300Mb (/boot), 20Gb (/var), 8Gb (swap) e o restante disponível em /home.

Para a comunicação na rede interna e externa do cluster (ver figura 1), foram utilizadas as 2 interfaces de rede existentes em um dos nós (storage). Uma delas é foi

configurada para acesso externo, e outra para comunicação na rede interna do cluster. Assim, qualquer acesso aos recursos do cluster deverá, necessariamente, passar por esse nó, onde as configurações de segurança devem ser cuidadosamente definidas, a fim de evitar problemas de acesso indevido. Cada nó de processamento tem seu número IP variando de 192.168.100.50 a 192.168.100.80. A pasta /home de um dos nós (node000) está mapeada em todos os demais via Network File System (NFS) [Jones 2010], de forma que a execução paralela de rotinas possa acessar o mesmo espaço de memória em disco.

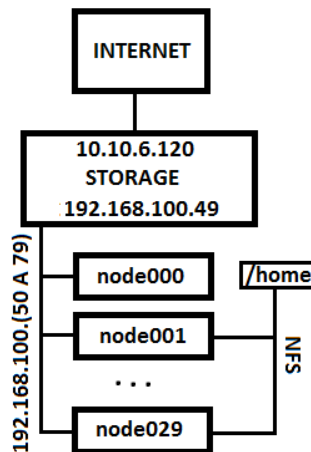


Figura 1. Conexão de rede dos computadores constituintes do cluster

### 2.3. Integração ao gerenciador de recursos

O OAR [Nicolas e Joseph 2011] é um gerenciador de recursos e tarefas para clusters e outras infraestruturas computacionais que vem sendo utilizado pelo grupo. Nessa etapa do projeto, foi feita a integração dos novos nós com o OAR. Para isso, foi realizada a instalação e configuração do gerenciador nos novos nós, além da criação do usuário “oar” necessário ao funcionamento desse sistema. Por fim, foi efetuada a atualização do banco de dados do OAR, atentando ao fracionamento dos recursos em função da memória disponível, da seguinte forma: foram adicionados os campos “*mem*” e “*core*” à tabela “*resources*” do banco de dados. Para que o hardware pudesse ser alocado tanto em função de um número de núcleos de processamento, quanto em função de uma quantidade de memória principal disponível, foi estabelecida uma unidade de memória comum entre os grupos de nós (1/24 GB) de forma que todos os nós pudessem dividir a memória disponível entre os núcleos de processamento utilizando unidades inteiras (e não fracionárias) de memória. Os detalhes desse tipo de abordagem são mostrados em [Machado, Cargnelutti e Petry, 2012]. Com isso, o primeiro grupo de nós antigos conterá 32 unidades de memória por *core*, o segundo, 24 registros por *core*, e o grupo de nós novos, 12 registros por *core*.

### 3. Resultados alcançados

O desempenho do cluster resultante (figura 2), foi comparado com o cluster original para a execução da previsão do conteúdo eletrônico da ionosfera na região da América do Sul [Petry et al. 2012], em termos de tempo de processamento requerido para rodar

56 instâncias do modelo Sheffield University Plasmasphere-ionosphere Model (SUPIM) [Bailey et al. 1997]. Com o aumento da capacidade do cluster, o tempo de processamento foi reduzido de 1812 segundos para 1630 segundos. Essa redução de pouco mais de 10% no tempo total se deu, basicamente, em razão da redução do número de processos simultaneamente executados em cada nó de processamento. No entanto, é importante mencionar que o cluster resultante não requereu a ocupação de todos os recursos de hardware disponíveis. Isso possibilitará, portanto, a previsão do conteúdo eletrônico da ionosfera para uma região maior, executando mais instâncias do SUPIM, e dentro do mesmo tempo de processamento.



Figura 2. Imagem do cluster após aumento da capacidade de processamento

#### 4. Conclusão

Este trabalho detalhou as atividades relacionadas ao aumento da capacidade de processamento de um cluster existente, utilizando recursos de hardware que foram avaliados, configurados e instalados para funcionarem de forma conjunta ao hardware existente. Após a atualização, o novo cluster pode contar com uma capacidade de processamento significativamente superior, abrangendo um total de 58 CPUs, 216 núcleos de processamentos (cores), 4512 unidades de memória equivalendo a uma quantidade total de 188 GB de memória RAM. Esses recursos podem ser livremente alocados para execução de tarefas que exijam muito esforço computacional ou alocação de memória. Para isso, está sendo usado o gerenciador de recursos OAR, que permite a submissão de tarefas tanto em função do número de CPUs, *cores* ou unidades de memória.

Como trabalhos futuros, espera-se integrar outro conjunto de equipamentos focados na questão do aumento da capacidade de armazenamento de dados. Tal melhoria visa possibilitar que as saídas processadas pelas simulações possam ser mantidas para acessos futuros.

**Agradecimentos:** Os autores agradecem o auxílio recebido via Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS), através das bolsas: PROBIC processo 0254-2551/12-4, e PROBIT processo 0256-2551/12-0.

## Referências

- Alecrim, E. . Sistemas RAID (Redundant Array of Independent Disks). Publicado em 25/01/2012. Disponível em: <<http://www.infowester.com/raid.php>>. Acesso em 08/09/2012.
- Bailey, G.J., Y.Z. Su, and N. Balan, The Sheffield University plasmasphere ionosphere model - a review, J. Atmos. Terr. Phys., 59, 1541-1552, 1997.
- Cheung, A.L. . High performance computing on a cluster of workstations. In: High-Performance Distributed Computing, 1992. (HPDC-1), Sep 1992, pp 152-160.
- Jones, T. . Network File System and Linux. 2010. Disponível em: <<http://www.ibm.com/developerworks/linux/library/l-network-file-systems/>>. Acesso em 10/9/2012.
- Kuck, D.J. . High Performance Computing: Challenges for Future Systems. Oxford University Press, Inc., New York, NY, USA, 1996.
- Machado, E.R.P., Cargnelutti, R. e Petry, A. . Otimização do Balanceamento de Carga Usando OAR. In: 12a Escola Regional de Alto Desempenho, 2012, Erechim, 2012. p. 185-188.
- Morimoto, C.E. . Tutorial do CentOS. 2008. Disponível em: <<http://www.hardware.com.br/tutoriais/centos/>>. Acesso em 10/9/2012.
- Nicolas, C. and Joseph, E. (2011) “Resource Management System for High Performance Computing”, <http://oar.imag.fr>, Novembro de 2011.
- Petry, A. et al. Operational Ionospheric Dynamics Prediction in Brazilian Space Weather Program. Submetido para Journal of Atmospheric and Solar-Terrestrial Physics, 2012.