

XVI Workshop de Iniciação Científica da Embrapa Gado de Leite
Juiz de Fora – 13 de Agosto de 2015

Desenvolvimento e implementação de um sistema de *backup* para ambiente de computação científica com infraestrutura de baixo custo¹

Caio dos Santos Borsato de Carvalho², Katia C. Lage dos Santos³, Fernanda Nascimento Almeida⁴, Wagner Arbex^{5, 6}

¹ Parte do projeto: ZebuMelhor – Estratégias de melhoramento genético para as raças Gir Leiteiro, Guzerá, Sindi e Girolando em sistemas sustentáveis de produção de leite, liderado por João Cláudio do Carmo Panetto

² Acadêmico de Engenharia, Embrapa Gado de Leite, Juiz de Fora, MG – Bolsista IC, e-mail: caio.carvalho@colaborador.embrapa.br

³ Analista, Embrapa Gado de Leite e PGCC/UFJF, Juiz de Fora/MG. e-mail: katia.santos@embrapa.br

⁴ Pós-doutora, Programa de Pós-graduação em Ciência da Computação, PGCC/UFJF, Juiz de Fora, MG – Bolsista PNPd/CAPES, e-mail: fernanda.nascimento@colaborador.embrapa.br

⁵ Analista e Professor Adjunto, Embrapa Gado de Leite e PGCC/UFJF, Juiz de Fora, MG. e-mail: wagner.arbex@embrapa.br, wagner.arbex@ufjf.edu.br

⁶ Orientador

Resumo: Com o crescente volume de dados é imprescindível o uso de métodos eficientes para armazenamento e tratamento dos mesmos. Este trabalho apresenta um sistema de *backup* automático com infraestrutura de baixo custo, a partir da unificação de discos distribuídos em rede, utilizando o *Network File System* (NFS) e o *middleware* – ie, a camada de intermediação –, *Multi Hard Disk Distributed File System* (MHDDFS). Assim, foi possível transferir de dados entre um servidor, computadores de mesa obsoletos e praticamente sem finalidade e, ainda, dispositivos de armazenamento de dados em rede. Foi desenvolvido um conjunto de rotinas por meio de *scripts* e/ou procedimentos implementados no sistema operacional (SO) dos equipamentos para automatizar os processos de cópia de dados e de armazenamento dos dados. Tais *scripts* e procedimentos fizeram uso de utilitários do SO para localização (*find*), cópia de diretórios (*rsync*), acesso remoto (*ssh* e *sshpass*) e agendamento de execução (*cron*)



Gado de Leite

XIV Workshop de Iniciação Científica da Embrapa Gado de Leite
Juiz de Fora – 28 de Julho de 2014

no servidor. Com o uso combinado desses utilitários, foi desenvolvido um sistema de *backup* de baixo custo, facilmente adaptável a diferentes níveis de usabilidade e sem que houvesse a necessidade de aquisição de novos recursos. Além disso, configuração utilizada para a implementação do sistema tem caráter inovador, pois não haviam registros da utilização integrada de recursos do NFS com o MHDDFS. O NFS é amplamente conhecido e permite o compartilhamento de diferentes discos distribuídos em rede, enquanto o MHDDFS possibilita que vários discos, localizados em um só equipamento, fossem visto como um único disco. A configuração criada fez com que diferentes discos, em diferentes equipamentos em rede, fossem vistos como um único disco. Ressalta-se que o sistema apresenta características de “dependabilidade” – neologismo de *dependability*, ie, “confiança” no conceito de tolerância a falhas –, mantendo-se em funcionamento mesmo com parte de seus recursos desligados. Seus serviços são interrompidos somente quando o servidor ou os demais computadores ficassem inativos ao mesmo tempo.

Palavras-chave: armazenamento, *backup*, computação científica, custo, e-Science, segurança de dados