

# Avaliação de Desempenho de Sistemas de Arquivos Paralelos com o Andrew Benchmark Modificado

Caciano dos Santos Machado\*, Rodrigo Virote Kassick†

Rafael Bohrer Ávila‡

Philippe Olivier Alexandre Navaux§

GPPD - Grupo de Processamento Paralelo e Distribuído  
Instituto de Informática - Universidade Federal do Rio Grande do Sul  
{caciano, rvkassick, avila, navaux}@inf.ufrgs.br

## Introdução

Aplicações de alto desempenho com comunicação massiva de dados requerem o gerenciamento eficiente e transferência de gigabytes, terabytes ou até mesmo de petabytes de informação em ambientes paralelos. Em virtude dessas necessidades surgiram projetos sobre sistemas de arquivos paralelos cujos objetivos primordiais são:

- minimizar o gargalo da entrada e saída no acesso a unidades de disco;
- permitir uma maior escalabilidade da vazão de dados e capacidade de armazenamento com o aumento das máquinas servidoras do sistema através da distribuição e/ou fracionamento de arquivos grandes.

O presente trabalho consiste numa análise de desempenho de diferentes sistemas de arquivos paralelos. Para a realização dos testes contamos com uma variante do conhecido *Andrew Benchmark* adaptado para a nossas arquiteturas de clusters.

## Sistemas de Arquivos Paralelos

A diferença de desempenho entre processador/memória e disco vem aumentando constantemente. Esse fator acaba tornando o disco um gargalo para o processamento de aplicações que demandam alto desempenho e utilizam grande quantidade de dados. Uma solução imediata encontrada para isso foi a utilização de sistemas de arquivos paralelos que visam, principalmente, aumentar a vazão de dados transferidos.

Os Sistemas de Arquivos Paralelos que foram abordados na nossa análise foram o PVFS e o dNFSp. Ambos são sistemas de arquivos em nível de usuário que fornecem transparência de nomes, localização, acesso e replicação. Para fins de comparação também será considerado o desempenho do NFS em nível de usuário.

---

\*Graduando em Ciência da Computação - UFRGS

†Graduando em Ciência da Computação - UFRGS

‡Professor, Doutorando em Ciência da Computação - UFRGS

§Professor, Doutor (Institut National Polytechnique de Grenoble, França, 1979)

Segue uma ligeira introdução sobre cada um desses sistemas de arquivos:

**UNFS** O sistema de arquivos NFS [NFS] (*Network File System*) nível de usuário do linux foi utilizado para fins de comparação com os sistemas paralelos.

**PVFS** O PVFS [CAR 00] (*Parallel Virtual File System*) se baseia na distribuição dos arquivos em servidores de I/O . Um metaseruidor é responsável pelo gerenciamento dos arquivos de metadados que armazenam os atributos e as informações sobre como os arquivos estão distribuídos nos servidores de I/O.

**dNFSp** Analogamente ao PVFS, o dNFSp [AVI 04] (*Distributed Network File System Parallel*) é baseado no uso de metaservidores e servidores de I/O, com o diferencial de poder utilizar mais de um metaseruidor. A proposta da utilização de mais de um metaseruidor visa evitar que este se torne um gargalo para um grande número de clientes. Desta maneira cada metaseruidor fica responsável por atender as requisições de um conjunto limitado de clientes.

O dNFSp está sendo desenvolvido a partir do NFSP [LOM 02] que é desenvolvido no Laboratoire ID de Grenoble. Tanto o NFSP como o dNFSp objetivam, além do melhor desempenho e escalabilidade, manter compatibilidade do cliente com o NFS normal, facilitando o gerenciamento.

## DAB

O DAB (*Distributed Andrew Benchmark*) é uma variante do conhecido *Andrew Benchmark* para testes em ambientes distribuídos.

O *Andrew Benchmark* é bastante utilizado para avaliação de desempenho de sistemas de arquivos. O benchmark visa simular as operações rotineiras de um ambiente de desenvolvimento, como cópia, listagem e leitura de arquivos e compilação de códigos fonte. O processo do benchmark se dá em 5 fases, cada qual correspondendo a uma característica distinta do sistema de arquivos:

1. **mkdir** criação de diretórios
2. **cp** cópia de arquivos
3. **ls** *stat* nos arquivos
4. **grep+wc** leitura de arquivos
5. **make** compilação dos arquivos fonte copiados

As modificações implementadas no benchmark servem para avaliar o comportamento do sistema de arquivos com todos os clientes realizando o benchmark em paralelo. Nas fases 1 e 2 os clientes fazem a criação dos seus diretórios individuais e cópia de arquivos para esses diretórios. Nas fases 3 e 4 os arquivos dos outros clientes são acessados, gerando concorrência no acesso aos arquivos. A fase 5 faz a listagem de todos os arquivos

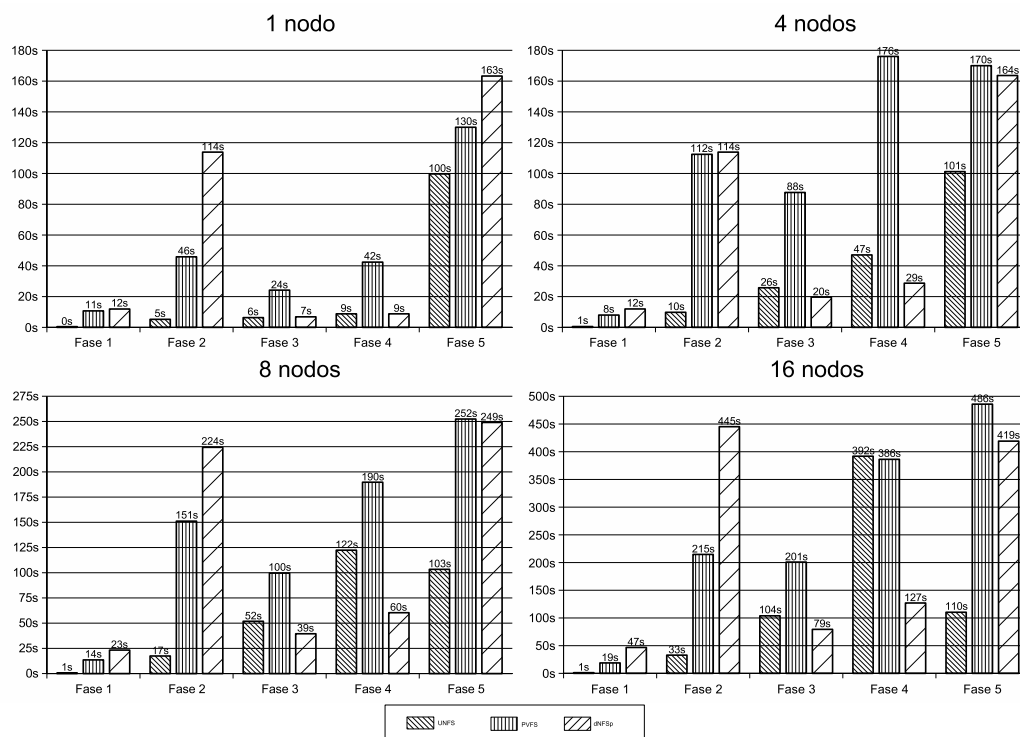


Figura 1: Tempos de execução do Andrew Benchmark Modificado

e a fase 4 faz a leitura de todos os arquivos com a utilização de grep e wc. Na fase 5 cada cliente compila um código fonte no diretório individual que foi criado na fase 1.

## Análise de Desempenho

Para os testes do benchmark utilizamos o cluster Labtec do Instituto de Informática da UFRGS. O cluster possui 20 nodos Dual Pentium III 1.1 MHz, 1GB de RAM e Disco SCSI de 18GB, conectados por um switch Fast Ethernet.

As configurações de cada sistema de arquivos para o teste do Andrew Benchmark Modificado foram as seguintes:

**UNFS** 1 nodo servidor e até 16 nodos clientes

**PVFS** 1 nodo metaservidor e servidor de I/O, outros 3 nodos exclusivamente servidores de I/O e até 16 nodos clientes

**dNFSp** 4 nodos metaservidores e servidores de I/O e até 16 nodos clientes

Para garantir a consistência do teste, foram feitas três execuções do benchmark para cada fase. Na Figura 1 são apresentadas as médias das três execuções para os diferentes sistemas de arquivos com 1, 4, 8 e 16 clientes executando o benchmark simultaneamente.

Pode-se notar que o dNFSp apresentou uma performance pior nas fases 1, e 2, caracterizadas pela criação de diretórios. Uma vez que o protótipo do dNFSp ainda não possui um mecanismo de validação de conteúdo de diretório, a criação do arquivo ou diretório acaba por exigir a sincronização dos metaservidores através de `rsh`, o que é pouco eficiente.

O dNFSp, no entanto, mostrou-se vantajoso nas fases 3 e 4, nas quais é feita a leitura dos dados. A sincronização dos metaservidores neste caso é menor, pois já foi realizada nas outras etapas. Assim, pode-se notar a vantagem da existência de vários metaservidores, que permite o acesso mais rápido aos dados por parte dos clientes devido à inexistência de gargalo em um único metaservidor, como no caso do PVFS.

Na fase 5, o dNFSp não apresenta vantagem quando há apenas um cliente. Para 4 ou mais clientes, no entanto, o dNFSp apresenta pequena vantagem com relação ao PVFS, mas ainda tem desempenho pior que o NFS. Ainda não foram realizados testes com mais de 16 clientes consecutivos para verificar se esta tendência se mantém.

## Conclusões

A utilização de múltiplos metaservidores de dados mostrou-se eficiente para a obtenção de um melhor desempenho no acesso aos dados. É necessário, no entanto, um aprimoramento do mecanismo de sincronização dos metaservidores, visto que a criação de arquivos e diretórios mostrou-se bastante ineficiente devido à esta sincronização.

Resta, também, realizar testes com um maior número de clientes para verificar a escalabilidade do dNFSp e do PVFS com relação ao NFS normal. Tais testes ainda não foram realizados devido à quantidade de nodos disponíveis nos clusters utilizados.

## Referências

- [NFS]      **The Network File System** <http://nfs.sourceforge.net/>
- [CAR 00]   P. H. Carns, W. B. Ligon III, R. B. Ross, and R. Thakur, PVFS: A Parallel File System For Linux Clusters. **Proceedings of the 4th Annual Linux Showcase and Conference**, Atlanta, p.317–327 (Best Paper Award), 2000
- [LOM 02]   Pierre Lombard, Yves Denneulin, nfsp: A Distributed NFS Server for Clusters of Workstations. **Proceedings of the 16th International Parallel and Distributed Processing Symposium**, Fort Lauderdale, 2002.
- [AVI 04]   Rafael B. Ávila, Philippe O. A. Navaux, Pierre Lombard, Adrien Lebre, Yves Denneulin, Performance Evaluation of a Prototype Distributed NFS Server. Admitido no **16º Simpósio em Arquitetura de Computadores e Processamento de Alto Desempenho, SBAC-PAD'04** para publicação em Outubro de 2004 em Foz do Iguaçu, Brasil.