

Sistemas de Arquivos Paralelos: Alternativas para a redução do gargalo no acesso ao sistema de arquivos

Alfredo Goldman¹, Roberto Pires de Carvalho¹

¹Instituto de Matemática e Estatística
Universidade de São Paulo (USP) – São Paulo, SP – Brasil

{gold, carvalho}@ime.usp.br

***Abstract.** During the last years, the grown on the speed to access a hard disk drive have not increased at the same ratio found when talking about processors and networks. Therefore, many applications can't reach the full use of the processors, because they have to wait the arrival of needed data before processing. A popular way to solve this is the adoption of parallel file systems, which use several disks to store data, which reduces the bottleneck caused by constantly accessing only one disk.*

In this study, we will show that a parallel file system, PVFS2 in the case, can be faster than a local file system, even having just one machine as a client.

***Resumo.** No decorrer dos últimos anos o aumento das velocidades de acesso aos discos rígidos não aumentou na mesma proporção que os aumentos das velocidades de processamento e transmissão de dados em redes. Com isso, muitas aplicações não atingem o pleno uso dos processadores, pois têm que esperar os dados chegarem do disco para serem processados. Uma forma popular para resolver isso é a adoção de sistemas de arquivos paralelos, que por utilizarem vários discos para armazenar os dados, acabam reduzindo o gargalo provocado pelo acesso constante a apenas um disco.*

Neste estudo, mostraremos que um sistema de arquivos paralelo, no caso o PVFS2, pode ser mais rápido do que um sistema de arquivos local, mesmo no caso de termos como cliente apenas uma máquina.

1 Introdução

A velocidade de acesso aos arquivos a partir de seus meios físicos de armazenamento (como dispositivos magnéticos, óticos, etc) não evoluiu na mesma proporção que a velocidade dos processadores ou da transmissão de dados, sejam estes internamente à máquina (como no acesso à memória, registradores, etc) ou externamente (como no acesso às outras máquinas via rede).

Isso torna o acesso aos arquivos um gargalo para muitas aplicações, especialmente para aquelas que buscam informações armazenadas remotamente. Para reduzir esse problema, no lugar de um sistema de arquivos remoto mais simples, costumam-se usar sistemas de arquivos paralelos, que são especializados em fornecer alto desempenho no acesso aos dados para um número crescente de clientes dentro de uma rede local. Essa solução também melhora o desempenho de aplicações distribuídas.

Neste trabalho mostramos que esta mesma solução pode ser usada por aplicações que manipulam uma grande quantidade de dados em uma só máquina, sejam elas paralelas ou não, para resolver o problema do gargalo provocado pelo sistema de arquivos local.

Esse artigo está organizado da seguinte forma: na seção 2 apresentaremos o PVFS2, sistema de arquivos paralelo usado para validar nossa proposta. Em seguida, na seção 3, comentaremos sobre a motivação de se comparar o PVFS2 com o sistema de arquivos local Ext3. Na seção 4 comentaremos sobre os resultados encontrados. Por fim, na seção 5, apresentaremos as conclusões obtidas.

2 PVFS2

O *Parallel Virtual File System* [Carns et al. 2000, Ramachandran 2002, PVFS2 Development Team 2003, Haddad 2000] é um sistema de arquivos distribuído desenvolvido para prover alto desempenho e escalabilidade paralela para aglomerados de PCs linux. Em geral, o PVFS2 promete 4 características:

- Um espaço de nomes consistente para todo o aglomerado;
- Acesso transparente para programas e aplicações já existentes, sem a necessidade de recompilá-los;
- Distribuição física de dados dos arquivos em múltiplos discos e múltiplos nós;
- Alto desempenho no acesso em modo usuário.

Quando uma aplicação acessa o PVFS2 para obter os dados de um determinado arquivo, várias máquinas serão acessadas, de forma transparente. Por acessar várias máquinas, utiliza-se de vários caminhos pela rede para chegar aos respectivos discos em que os dados estão armazenados. Isso elimina o gargalo de entrada e saída de dados quando se tem toda a informação armazenada em uma só máquina, distribuindo a carga e aumentando o potencial total da banda para múltiplos clientes.

3 PVFS2 vs. Ext3

O nosso objetivo é propor uma comparação entre o sistema de arquivos paralelo PVFS2 e o sistema de arquivos local Ext3, usando uma única máquina, como cliente. À primeira vista pode parecer estranho compará-los, já que têm propósitos distintos, porém, caso haja largura de banda suficiente na rede para o tráfego dos dados, um sistema de arquivos paralelo pode ser mais rápido que um sistema de arquivos local. Para deixar mais clara a idéia, a figura 1 ilustra uma rede muito rápida interligando servidores e clientes que possuem dispositivos de armazenamento notavelmente mais lentos.

Assim, muitas requisições enviadas ao PVFS2, mesmo que vindas de um único cliente, utilizam toda a banda disponível dos discos dos servidores, já que não têm a rede como gargalo, enviando o resultado para o cliente rapidamente. Comparando essa banda agregada com a banda disponível pelo sistema de arquivos local, percebemos a clara vantagem em se utilizar o PVFS2.

4 Resultados

Nosso estudo foi realizado em duas etapas: simulação [de Carvalho 2005] e verificação [Goldman and de Carvalho 2006]. Essas etapas foram necessárias por não possuímos um aglomerado conectado em uma rede de alta velocidade no início do estudo.

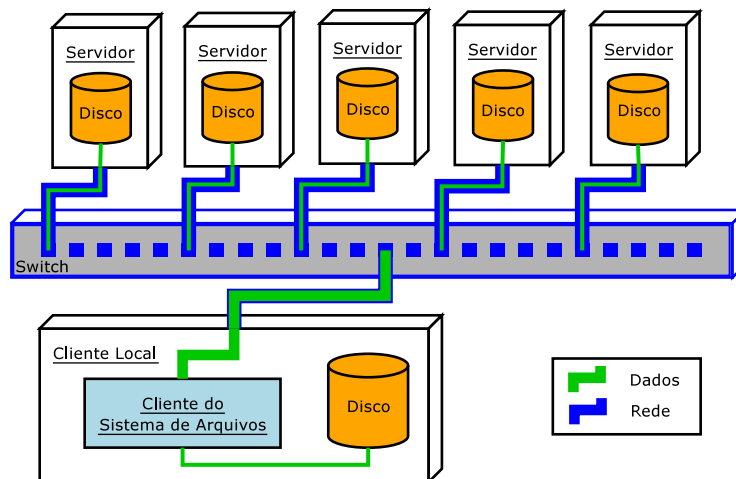


Figura 1. Sistemas de arquivos paralelos vs. locais

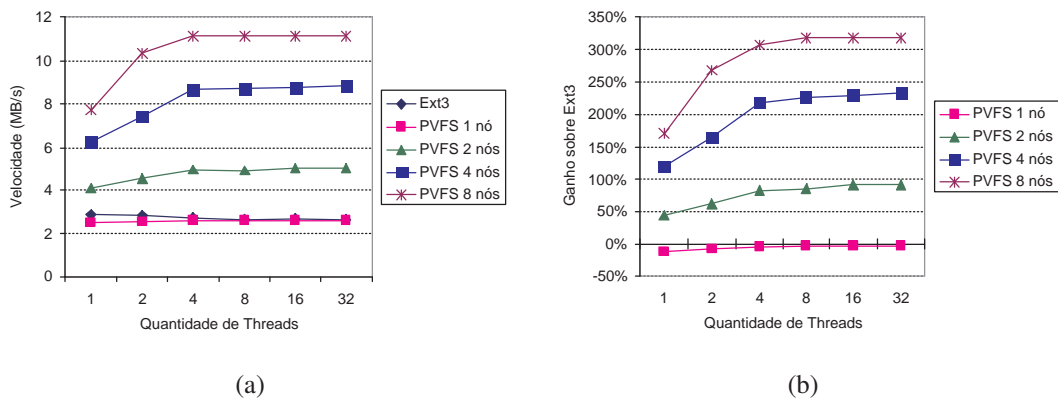


Figura 2. (a) Velocidade do PVFS2 e Ext3 e (b) ganho do PVFS2 sobre o Ext3 ao aumentar a concorrência no acesso, usando discos na velocidade 2,5MB/s

Os resultados da etapa inicial foram obtidos através da redução da velocidade dos discos locais de todas as máquinas envolvidas. Essa estratégia possibilitou simularmos uma rede de dados mais rápida que o sistema de armazenamento local. A figura 2 mostra que se a rede for realmente mais rápida que os discos, temos ganho considerável ao se usar o PVFS2 no lugar do Ext3, principalmente no acesso concorrente.

Ao aumentarmos a velocidade dos discos, analisamos o quanto o desempenho do PVFS2 é afetado. A figura 3 mostra esse caso. A partir desses resultados, estimamos, através de uma simples regra de três, que o desempenho do PVFS2 em uma rede conectada a 1Gbit/s seria de aproximadamente 109MB/s, considerando que a velocidade máxima da rede seria de 125MB/s, que os discos não usam mais do que 92% da banda e que não haveria nenhum outro gargalo a não ser o próprio disco.

Na etapa seguinte (verificação) tivemos acesso a um aglomerado de 12 máquinas, todas conectadas entre si através de placas de rede Gigabit, ligadas a um switch de mesma velocidade. Realizamos testes comparativos de desempenho entre o PVFS2 e o Ext3 com os discos na sua velocidade máxima, onde concluímos quais condições são necessárias para que um sistema seja mais eficiente que o outro.

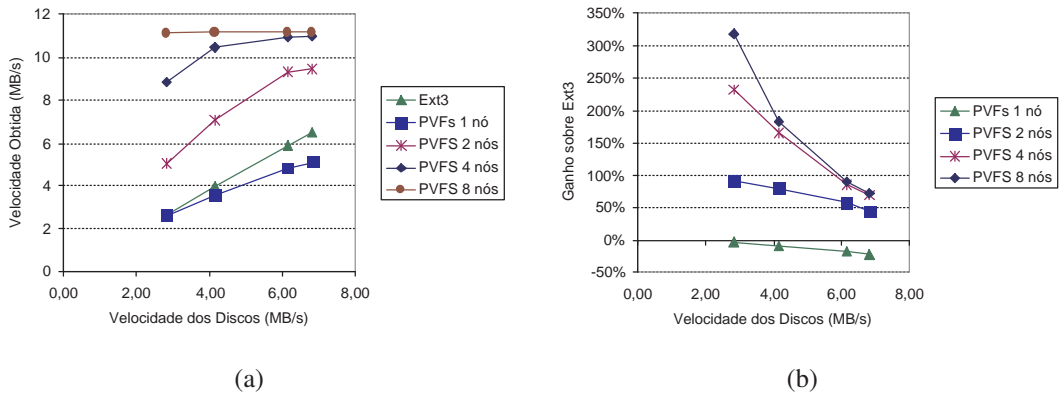


Figura 3. (a) Velocidade do PVFS2 e do Ext3 e (b) ganho do PVFS2 sobre o Ext3 ao variarmos a velocidade dos discos, usando 32 threads

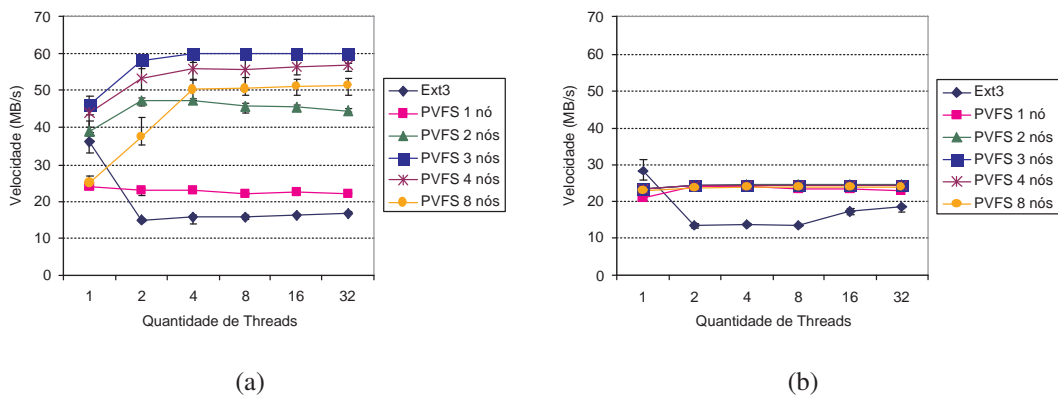


Figura 4. Velocidade do PVFS2 e do Ext3 em (a) um cliente com processador Pentium 4 2,8GHz e em (b) um cliente com processador Pentium 4 1,8GHz, ambos em rede Gigabit

Os resultados encontrados estão apresentados na figura 4. Nela notamos que o Ext3 perde muito desempenho quando se tem algum acesso concorrente, enquanto que o PVFS2 se mantém constante. Note também que o desempenho do PVFS2 com 8 servidores é pior do que com 4. Isso se deve ao fato de não termos uma rede homogênea nos nossos testes, sendo priorizadas as melhores máquinas para os testes com menos servidores.

Além disso, ao utilizarmos apenas um servidor notamos que a capacidade de processamento se torna um limitante para um melhor desempenho, que é resolvido conforme o número de servidores aumenta. Porém, ao aumentarmos a quantidade de servidores, o poder de processamento do cliente passa a ser um novo gargalo devido ao grande volume de dados sendo processados. Isso é verificado na comparação entre as figuras 4(a) e 4(b), onde temos um cliente com melhor poder de processamento contra um cliente com menor poder, respectivamente, ambos acessando os mesmos servidores, mas não ao mesmo tempo.

5 Conclusão

Dessa forma, concluímos que o gargalo no acesso aos dados do disco foi resolvido com o PVFS2, mesmo usando apenas um cliente. Como os servidores são acessados aleatoriamente, acabam recebendo as requisições de forma balanceada, o que melhora as respostas do PVFS2, não deixando-o se tornar um gargalo.

Já o Ext3 não possui um desempenho muito bom quando se tem acesso concorrente, ficando muito inferior ao PVFS2. Mesmo quando o acesso é mono-tarefa, o PVFS2 pode ser melhor, desde que a máquina cliente tenha capacidade suficiente de processamento e no acesso à rede de alta velocidade.

Assim, temos como gargalo principal o cliente, ou mais especificamente seu poder de processamento e transmissão de dados, conforme observa-se na figura 4. Nela, vemos claramente que um cliente mais potente consegue obter melhor desempenho do PVFS2 do que um cliente mais fraco, acessando o mesmo sistema de arquivos, o que mostra que tanto a rede como a CPU são mais importantes que a velocidade do disco quando o objetivo é atingir alto desempenho no acesso ao sistema de arquivos.

Entre os possíveis trabalhos futuros, temos o estudo de aplicações reais que usam uma grande quantidade de dados, como, por exemplo, as de bio-informática, assim como o estudo do impacto da redundância (através da duplicação) de informações.

Referências

- [Carns et al. 2000] Carns, P. H., III, W. B. L., Ross, R. B., and Thakur, R. (2000). PVFS: A parallel virtual file system for linux clusters. In *Proceedings of the 4th Annual Linux Showcase and Conference*, pages 317 – 327. <http://www.parl.clemson.edu/pvfs/el2000/extreme2000.ps>, (Best Paper Award).
- [de Carvalho 2005] de Carvalho, R. P. (2005). Sistema de Arquivos Paralelos: Alternativas para a redução do gargalo no acesso ao sistema de arquivos. Master's thesis, Instituto de Matemática e Estatística da Universidade de São Paulo. <http://www.ime.usp.br/~carvalho/defesa/mestrado.pdf>.
- [Goldman and de Carvalho 2006] Goldman, A. and de Carvalho, R. P. (2006). Sistemas de arquivos paralelos: Reduzindo o gargalo no acesso ao disco. <http://www.ime.usp.br/~carvalho/sbrc2006/artigo.pdf>.
- [Haddad 2000] Haddad, I. F. (2000). PVFS: A parallel virtual file system for linux clusters. *Linux Journal*, 80:74 – 82. <http://www.linuxjournal.com/article/4354>.
- [PVFS2 Development Team 2003] PVFS2 Development Team (2003). Parallel Virtual File System, Version 2. Technical report, Clemson University. <http://www.pvfs.org/pvfs2/pvfs2-guide.html>.
- [Ramachandran 2002] Ramachandran, H. (2002). Design and implementation of the system interface for PVFS2. Master's thesis, Clemson University. <ftp://ftp.parl.clemson.edu/pub/techreports/2002/PARL-2002-008.ps>.