Gerenciamento Distribuído de Servidores de Dados

Danilo Fukuda Conrad, Rodrigo Virote Kassick,
Everton Herrman, Philippe O. A. Navaux

Universidade Federal do Rio Grande do Sul
Av Bento Gonçalves 9500 - Bloco IV, Fone:(51)3316 6168,Fax:(51)3316 7308
{dfconrad, rvkassick,ehermann,navaux}@inf.ufrgs.br

Introdução

Um dos grandes problemas na computação atualmente é a diferença entre o desempenho do processador e o das operações de entrada e saída (I/O). Essa diferença é ainda mais perceptível em aplicações que precisam processar uma grande quantidade de dados, como as executadas em arquiteturas de clusters. Uma alternativa para solucionar essa diferença são os sistemas de arquivos paralelos.

O NFSp[Lombard and Denneulin, 2002] é um sistema de arquivos paralelo compatível com o NFS tradicional, com modificações apenas na parte do servidor. Inspirado no PVFS[Carns et al., 2000], ele divide as tarefas do servidor entre gerenciamento de dados e metadados. No entanto, o NFSp possui apenas um servidor de metadados. Em consequência disso, as escritas têm seu desempenho afetado, pois o servidor de metadados torna-se um gargalo para o sistema. Para amenizar esse problema foi desenvolvido o dNFSp[Ávila, 2005], que possui múltiplos servidores de metadados.

Para simplificar a configuração dos servidores de dados (IODs) no NFSp, foi criado um gerenciador de servidores de dados, chamado de *iod_mgmt*. A ferramenta foi desenvolvida como uma alternativa à configuração estática de IODs, permitindo que o sistema seja alterado sem a interrupção dos seus serviços essenciais. Ela funciona também como uma camada de abstração na configuração desses servidores.

O *iod_mgmt* permite a execução de um conjunto de operações de configuração. Algumas delas são a adição e a remoção de IODs, verificação do estado atual do sistema e de detecção de falhas. No entanto, os comandos precisam ser executados no mesmo nó que o servidor de metadados, pois a comunicação é feita através de um canal interprocessos. Este artigo apresenta as modificações realizadas na ferramenta de modo a adaptá-la para o dNFSp.

Modelo Distribuído

Primeiramente, foi preciso estabelecer um modelo para adaptar a ferramenta, pois, diferentemente do NFSp, o dNFSp possui suporte a múltiplos servidores de metadados. A solução encontrada, que minimiza as modificações necessárias na ferramenta, foi aumentar em mais um nível a abstração da configuração. Criou-se um lançador de comandos, que pode ser executado em qualquer nó do sistema e envia as requisições das operações

aos gerenciadores de IODs. O *iod_mgmt* continua a ser executado como antigamente, no mesmo nó do servidor de metadados. Para cada servidor é executada uma instância do *iod_mgmt* no mesmo nó. Ao receber as requisições dos comandos a serem executados, cada instância executa o comando localmente, como no modelo original. Para enviar os comandos a todas as instâncias o lançador de comandos precisa utilizar um mecanismo de comunicação remota.

Implementação

Para realizar a comunicação entre o lançador de comandos e as múltiplas instâncias do *iod_mgmt* foi utilizado um canal TCP, por onde são enviados os comandos a serem executados. Ao receber um comando, a instância o executa localmente, diretamente no servidor de metadados que se encontra na mesma máquina. Foi escolhido o canal TCP pois, além de possibilitar uma implementação simplificada, seu *overhead* não representa um problema, pois as mensagens de configuração são enviadas com pouca frequência.

A verificação dos IODs funcionais, como na versão original da ferramenta, continua sendo feita por apenas uma das instâncias do gerenciador de IODS. Ao verificar quais IODs estão operantes a instância envia uma mensagem para todas as outras informando o estado dos IODs.

Conclusão

Com a ferramenta adaptada para o modelo distribuído é possível reconfigurar o sistema dinamicamente, sem a necessidade de se reiniciar todos os serviços essenciais. Desse modo, é possível incorporar novas funcionalidades que dependem do dinamismo de servidores, como tolerância a falhas nos servidores de dados. A adaptação do modelo permitiu também combinar as vantagens oferecidas pelo *iod_mgmt* com o alto desempenho do dNFSp.

Como trabalho futuro, pretende-se adaptar a ferramenta para que o tratamento da perda de nós seja automatizado. Assim, o sistema pode reagir às alterações na sua configuração de forma automática, sem nenhuma interferência do administrador.

Referências

[Ávila, 2005] Ávila, R. B. (2005). *Uma Proposta de Distribuição do Servidor de Arquivos em Clusters*. PhD thesis, PPGC-UFRGS/ID-IMAG, Porto Alegre/Grenoble.

[Carns et al., 2000] Carns, P. H., Ligon III, W. B., Ross, R. B., and Thakur, R. (2000). PVFS: a parallel file system for Linux clusters. In *Proc. of the 4th Annual Linux Showcase and Conference*, pages 317–327, Atlanta, GA. Best Paper Award.

[Lombard and Denneulin, 2002] Lombard, P. and Denneulin, Y. (2002). nfsp: a distributed NFS server for clusters of workstations. In *Proc. of the 16th International Parallel & Distributed Processing Symposium, IPDPS*, page 35, Ft. Lauderdale, Florida, USA. Los Alamitos, IEEE Computer Society. Abstract only, full paper available in CD-ROM.