

RVision: uma ferramenta aberta e configurável para monitoração de clusters

Tiago C. Ferreto, César A. F. De Rose

Pontifícia Universidade Católica do Rio Grande do Sul
Av. Ipiranga, 6681 - Prédio 30 (Bloco 4) - Porto Alegre - RS/Brasil - CEP: 90619-900
Telefone: 51 33203558 ramal: 4463, Fax: 51 33203758
{ferreto, derose}@cpad.pucrs.br

Introdução

A utilização de clusters para obtenção de alto desempenho é muito comum atualmente, devido principalmente a sua boa relação custo *versus* desempenho. Para obtenção de alto desempenho em clusters é necessário um eficiente sistema de gerência, a fim de descobrir possíveis problemas e gargalos na máquina. Para descobrir e resolver estes problemas de forma eficaz é necessária a monitoração do cluster, isto é, a obtenção periódica de informações sobre o seu estado.

A monitoração de sistemas [BAK 2000] consiste no ato de coletar parâmetros de performance do sistema como utilização de CPU, utilização de memória, taxa de I/O e interrupções, e apresentá-los de forma que seja facilmente compreendido pelo administrador de sistemas. Este serviço é importante para a operação estável de clusters com um grande número de nós, pois auxilia o administrador na descoberta de possíveis problemas antecipadamente. Além disso, outras partes do sistema podem ser beneficiadas com a informação provida. Por exemplo, a informação pode ser utilizada para modificar o escalonamento de processos, de forma que auxilie no balanceamento de carga. A seguir são descritas algumas das principais características desejadas em um sistema de monitoração de clusters.

Baixa Intrusividade: A intrusividade gerada pelo sistema de monitoração deve ser mínima, isto é, o sistema de monitoração não deve utilizar os recursos disponíveis de forma excessiva, a fim de que as aplicações que sejam executadas no cluster sofram a mínima alteração possível em relação a sua performance.

Tráfego de Rede Reduzido: Devido a alta utilização da rede por grande parte das aplicações paralelas e distribuídas, é desejado que a utilização deste recurso pelo sistema de monitoração seja reduzida consideravelmente, a fim de que a disputa pelo recurso seja minimizada. Para isto são utilizadas principalmente duas abordagens: reduzir a frequência em que os dados são colocados na rede, e diminuir a quantidade de dados que são passados pela rede, através da seleção das informações a serem monitoradas.

Suporte à Web: Devido ao surgimento da Internet diversos sistemas oferecem suporte à Web. Os sistemas de monitoração que possuem este suporte possibilitam o acompanhamento sobre o estado de clusters de forma virtual.

Atualmente existem diversos sistemas de monitoração para clusters, os mais utilizados são o BWatch [BWA 2001], o Parmon [BUY 2000] e o SCMS [UTH 2000]. Estes sistemas possuem em comum o fato de que definem uma lista fixa de informações que podem ser monitoradas pela ferramenta, além disto, todas as informações são monitoradas em todos os nós do cluster da mesma forma. Este tipo de abordagem restringe a flexibilidade da ferramenta, sendo necessária a utilização de outra quando se deseja monitorar outra informação ou se necessita diferenciar a informação e a forma como a mesma é monitorada em cada nó do cluster. Esta flexibilidade é necessária principalmente em clusters heterogêneos e cluster de clusters.

RVision

Motivação

A principal motivação no desenvolvimento do RVision foi a inflexibilidade encontrada nos sistemas de monitoração existentes. A arquitetura aberta do RVision provê monitoração genérica, a qual habilita a utilização de bibliotecas de monitoração próprias do usuário nos recursos monitorados, assim como, a criação de clientes de monitoração que conectam-se com o núcleo através da Internet. Diversas formas de aquisição e mecanismos de transporte são suportados.

Conceitos e Características

Durante o projeto do RVision alguns novos conceitos de monitoração foram introduzidos, a fim de aprimorar a funcionalidade do sistema. Estes conceitos são:

Sessão de Monitoração A sessão de monitoração consiste em um ambiente de monitoração dedicado, o qual é configurado a partir de um arquivo próprio de configuração. A configuração deste ambiente tem como objetivo principal definir a lista de nós monitorados, as informações a serem monitoradas em cada nó e a frequência de aquisição para cada informação.

Biblioteca de Monitoração A biblioteca de monitoração consiste em uma coleção de rotinas responsáveis pela captura da informação. Estas bibliotecas podem ser implementadas pelo usuário a fim de adicionar funcionalidades ao monitor durante a execução do mesmo. A principal vantagem obtida por esta abordagem é a separação dos mecanismos de monitoração como transporte das informações e frequência de monitoração, e a aquisição da informação.

As principais características apresentadas pelo RVision são:

- **Múltiplas Sessões de Monitoração:** o RVision pode ser utilizado por vários usuários de forma concorrente, cada um com sua própria sessão de monitoração, as quais podem possuir distintas configurações.

- **Monitoração Genérica:** através da utilização de bibliotecas de monitoração o RVision pode ser utilizado para monitorar qualquer tipo de informação, sendo necessário somente a criação de uma biblioteca de monitoração com uma função que obtenha a informação desejada.
- **Instante de Análise de Informações:** as informações podem ser analisadas durante a monitoração, isto é, de forma ONLINE, ou após o encerramento da monitoração, chamada de análise POST-MORTEM, através de um arquivo com as informações monitoradas criado durante a monitoração.
- **Frequência de Monitoração:** as frequências de monitoração suportadas pelo RVision são: cíclica, por alteração e por demanda. A frequência cíclica captura as informações periodicamente, a frequência por alteração também captura as informações periodicamente, mas ela só é enviada quando uma condição é satisfeita, a frequência por demanda consiste na aquisição das informações devido a uma explícita requisição do cliente.

Arquitetura e Implementação

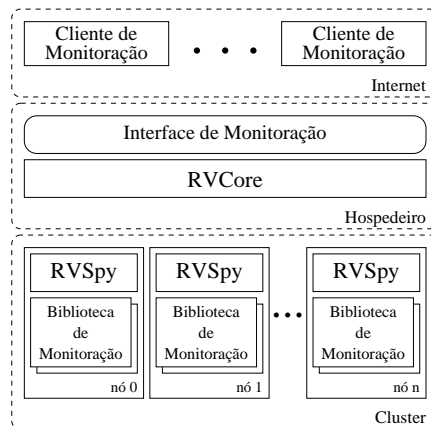
A arquitetura do RVision é baseada no modelo mestre-escravo, utilizando uma abordagem centralizada. Esta abordagem foi escolhida pela sua menor complexidade em comparação com a distribuída e, de certa forma, mais eficiente em clusters com até um número médio de nós. A implementação foi realizada no GNU/Linux usando a linguagem C. A arquitetura é composta por cinco módulos, divididos em Cliente de Monitoração, Interface de Monitoração, RVCore, RVSpy, e Biblioteca de Monitoração. O RVision é Open Source e utiliza a licença GPL.

Funcionamento

A figura 1 apresenta a arquitetura do RVision, mostrando os seus cinco módulos. O *Cliente de Monitoração* utiliza-se das funções providas pela *Interface de Monitoração* para se comunicar com o *RVCore*. As funções existentes na *Interface de Monitoração* provêm os mecanismos básicos para gerenciamento da monitoração, como a configuração da sessão de monitoração, inicialização e encerramento da monitoração, entre outros. O *RVCore* cria para cada conexão recebida uma sessão de monitoração distinta, porém, antes é realizada a autenticação do usuário no sistema. Após a criação da sessão de monitoração, a configuração desta pode ser realizada e a monitoração inicializada.

A etapa de inicialização informa ao módulo *RVSpy* existente em todas as máquinas que estão definidas na sessão de monitoração sobre as *Bibliotecas de Monitoração*, informações e frequências de monitoração utilizadas em cada nó monitorado. Após o recebimento da configuração, o *RVSpy* realiza a ligação dinâmica das *Bibliotecas de Monitoração* utilizadas e dispara submódulos de coleta de informação, os quais agem de acordo com o tipo de frequência e análise definidos e acessam as funções definidas nas *Bibliotecas de Monitoração* para obter as informações. A *Biblioteca de Monitoração* consiste em uma *shared library* do GNU/Linux. Ela deve possuir algumas definições em sua estrutura interna, para que possa ser ligada dinamicamente ao *RVSpy*. No término

Figura 1: Arquitetura do RVision.



da monitoração os módulos *RVSpy* finalizam os submódulos de coleta de informação e limpam a área de memória utilizada.

Conclusão e Trabalhos Futuros

O RVision está sendo utilizado no CPAD (<http://www.cpad.pucrs.br>), onde já foram realizados diversos testes com a ferramenta. Os resultados demonstraram que a ferramenta apresenta uma baixa intrusividade (média de 2%) para monitorações típicas. Alguns clientes de monitoração para o RVision já foram criados (monitor gráfico e textual para performance dos nós) e outros estão em desenvolvimento (monitor da rede Myrinet, monitor para arquivos POST-MORTEM). Também encontram-se em desenvolvimento algumas bibliotecas de monitoração. Maiores informações sobre o RVision podem ser obtidas em <http://www.cpad.pucrs.br/rvision>.

Referências

- [BAK 2000] BAKER, M. **Cluster computing white paper**.
- [BUY 2000] BUYYA, R. PARMON: a portable and scalable monitoring system for clusters. **SP&E**, v.30, n.7, p.723–739, 2000.
- [BWA 2001] BWATCH 1.0.3. <http://www.sci.usq.edu.au/staff/jacek/bWatch/>.
- [UTH 2000] UTHAYOPAS, P.; MANEESILP, J.; INGONGNAM, P. SCMS: an integrated cluster management tool for beowulf cluster system. In: PDPTA 2000, 2000, Las Vegas, Nevada, USA. **Proceedings...** [S.l.: s.n.], 2000.