

Gerenciamento de Agregados

OpenPBS x CCS

¹Rafael von Hoonholtz Magrin, ¹Augusto D. P. dos Santos, Rafael B.

Ávila, Philippe O. A. Navaux

Universidade Federal do Rio Grande do Sul - Instituto de Informática
{magrin, adpsantos, avila, navaux}@inf.ufrgs.br

Introdução

Com a necessidade de se obter cada vez maior desempenho computacional, os agregados têm se mostrado uma boa opção, pois têm um custo baixo em relação aos supercomputadores. Mas o acesso aleatório aos nodos, pelos usuários, tem se mostrado um obstáculo para o bom funcionamento destes, pois muitas vezes alguns nodos ficam livres enquanto outros estão executando processos concorrentes.

Tendo em vista este problema foram desenvolvidos, nos últimos anos, diversos *softwares* para gerenciamento de agregados, possibilitando uma melhor utilização dos seus recursos, facilitando as tarefas do administrador e simplificando a utilização destes pelos usuários.

O trabalho tem por objetivo mostrar as principais características de dois dos principais *softwares* para gerenciamento de agregados disponíveis para plataforma Linux x86: o OpenPBS (*Portable Batch System*), desenvolvido inicialmente pela NASA, e o CCS (*Computing Center Software*), desenvolvido pela Universidade de Paderborn (Alemanha).

OpenPBS

O OpenPBS foi desenvolvido pela NASA em 1993 para substituir o NQS, seu, até então, software de gerenciamento, que se tornou inadequado para as suas necessidades. A partir de 1998 a Veridian Systems começou a distribuir e dar suporte ao OpenPBS, criando sua versão comercial, o PBS Pro [PBS 02].

Atualmente, o PBS tornou-se quase que um padrão entre os softwares para gerenciamento de agregados e sua versão livre, OpenPBS, vem contribuindo para o seu desenvolvimento.

Arquitetura

Os componentes do OpenPBS podem ser divididos em dois grandes grupos: os comandos de usuário e os *daemons* de sistema.

Os comandos de usuário podem ser executados tanto via *shell* quanto via interface gráfica. E são utilizados para enviar, monitorar, modificar e apagar trabalhos, e para configurar os *daemons*. Os comandos de usuário podem ser instalados em qualquer sistema suportado, não precisando estar no mesmo sistema que os *daemons*.

¹ Bolsistas do Projeto LabTeC DELL - UFRGS (Laboratório de Tecnologia de Cluster)

Os *daemons* de sistema são três [PBS 00]:

- O servidor de trabalhos: é o principal componente do OpenPBS, e para ele são enviados todos os comandos de usuário. Ele é o responsável por receber, alterar e enviar para execução os trabalhos. Normalmente existe somente um servidor para cada agregado e ele está localizado no *frontend*.
- O executor de trabalhos: é o responsável por monitorar os nodos, preparar o sistema para execução de trabalhos e executar os trabalhos. Cada nodo deve ter seu próprio executor.
- O agendador de trabalhos: é o responsável por aplicar as políticas do agregado e informar ao servidor qual trabalho deve ser enviado para execução. Para saber quais os recursos estão disponíveis para uso o agendador se comunica com o executor. Assim como o servidor, o agendador normalmente está localizado no *frontend*.

CCS

O CCS foi desenvolvido a partir de 1992 como um projeto do grupo PC² (Paderborn Center for Parallel Computing) da Universidade de Paderborn (Alemanha). Atualmente ele está sendo adaptado para ser distribuído livremente com o nome de OpenCCS.

O CCS, ao contrário do OpenPBS, é utilizado em muito poucos lugares, pois apesar de não ser distribuído livremente, também não é um *software* comercial, fazendo com que seja utilizado quase que somente em instituições acadêmicas. Este fato fez com que o CCS tivesse seu desenvolvimento mais lento que o OpenPBS e mais voltado para arquiteturas de hardware específicas.

Arquitetura

O CCS é composto por seis módulos, os quais podem ser compostos por um ou mais *daemons*. Estes são executados assincronamente em diferentes *hosts* para melhorar o tempo de resposta do CCS [CCS 02].

- User Interface (UI): oferece acesso ao sistema de gerenciamento através do modo gráfico ou de interface ASCII pelo *shell*.
- Access Manager (AM): gerencia a interface do usuário e é responsável pela autenticação, autorização e contabilização dos recursos utilizados pelo usuário.
- Queue Manager (QM): agenda as requisições dos usuários no gerenciador.
- Machine Manager (MM): prove funções específicas para as máquinas, como alocação de partições, controle dos trabalhos, etc. Pode ser dividido em duas partes: uma para a administração do agregado, executada no servidor, e outra para execução dos trabalhos, executada nos nodos.
- Island Manager (IM): fornece serviço interno de nomes e ferramentas para vigilância do sistema.
- Operator Shell (OS): é a principal interface de controle para o administrador do sistema.

Análise Comparativa

Apesar de terem o mesmo objetivo os dois *softwares* apresentam características muito diferentes: a principal delas é a forma como é feito o agendamento dos trabalhos, que torna possível dividir os softwares de gerenciamento de agregados em duas categorias, os gerenciadores de filas de trabalhos, como OpenPBS e os agendadores de trabalhos, como o CCS [CCS 02]. Além disto eles apresentam várias outras diferenças como as que serão apresentadas a seguir.

Agendamento de trabalhos e políticas de agendamento

O OpenPBS trabalha com o conceito de filas, de modo que os trabalhos são enfileirados à medida que são enviados para o servidor, onde pode existir somente uma fila ou múltiplas filas com prioridades e propriedades diferentes, e agendados somente na hora de sua execução, impossibilitando em que horário seu trabalho será executado.

Já o CCS permite ao usuário escolher, através das requisições de tempo fixo, em qual horário deseja que sua aplicação seja executada, ou caso não tenha esta necessidade, saber quando seu trabalho terá encerrado sua execução fazendo uma requisição de tempo variável, a qual só pode ser adiantada nunca adiada.

O CCS tem três políticas disponíveis: *first-come-first-serve*, onde o primeiro trabalho enviado é o primeiro trabalho a ser executado, *shortest-job-first*, onde os menores trabalhos são executados primeiro, e *longest-job-first*, onde os trabalhos maiores são executados primeiro. Em adicional a estas estratégias o CCS tenta preencher os espaços vagos na agenda.

No OpenPBS a política do agregado pode ser configurada de duas maneiras, pode-se utilizar a política padrão do OpenPBS que pode ser configurada em alguns aspectos, ou pode-se criar uma política própria do agregado utilizando as linguagens de programação C, Tcl e BaSL, permitindo uma melhor personalização da política.

Como o OpenPBS não agenda os trabalhos no momento do seu envio para o servidor suas políticas se tornam mais efetivas do que no CCS, possibilitando também dar diferentes prioridades de execução aos usuários, além de possibilitar a criação de políticas muito mais complexas.

Tipos de nodos

No OpenPBS os nodos podem ser classificados de duas maneiras genéricas: nodos de alocação exclusiva e nodos de alocação compartilhada.

Os nodos de alocação exclusiva somente podem ser utilizados por um trabalho de cada vez, impedindo a concorrência entre trabalhos pelos recursos do mesmo nodo.

Os nodos de alocação compartilhada podem ser utilizados por vários trabalhos ao mesmo tempo, o número máximo de trabalhos que podem ser executados neste tipo de nodo pode ser um número fixo ou um número variável, levando em consideração que a utilização do processador nunca poderá ser maior que um dado limite estipulado.

Diferente do OpenPBS, o CCS somente trabalha com nodos de alocação exclusiva, mas está entre os objetivos de seus desenvolvedores possibilitar a alocação compartilhada em sua próxima versão.

Controle de acesso aos nodos

O CCS tem recursos de controle de acesso aos nodos nativo, sendo somente necessário configurar o sistema para bloquear o acesso de qualquer usuário aos nodos, a não ser que o usuário esteja listado no arquivo `/etc/security/access.conf`.

O OpenPBS, a princípio, não tem nenhum recurso de controle de acesso aos nodos, assim permitindo que um usuário se conecte a um nodo e execute uma aplicação fazendo com que a utilização do OpenPBS seja de pouco valor. Porém, como o OpenPBS possibilita a execução de um *script* antes e após a execução de cada trabalho, pode-se utilizar estes *scripts* para bloquear o acesso de qualquer usuário aos nodos de modo semelhante ao CCS.

Conclusão e Trabalhos Futuros

A escolha de um software para gerenciamento de agregados é muito específica para cada tipo de agregado, pois cada um tem suas necessidades próprias.

O OpenPBS se apresenta interessante para agregados que tem a intenção de implantar políticas complexas, com alguns usuários tendo prioridade de utilização sobre outros.

Já o CCS se apresenta interessante para agregados que não necessitam de políticas complexas nem de prioridades diferentes entre os usuários, mas sim que os usuários tenham divisão justa dos recursos disponíveis. Ele também se mostrou interessante para agregados onde é necessário fazer antecipadamente reservas de recursos, ou seja necessário reservas de recursos com horário fixo.

Como continuação deste trabalho pretendemos também estudar outros softwares para gerenciamento de agregados como Cronos, desenvolvido pela PUC-RS, Sun Grid Engine, desenvolvido pela Sun Microsystems, e Condor, desenvolvido pela University of Wisconsin (USA).

Referências

- [PBS 00] BAYUCAN, Albeaus, et al. **Portable Batch System Administration Guide**. Veridian Sytem, v2.3, 114p, Agosto. 2000.
- [CCS 02] KELLER, A. **OpenCCS Administrator Manual**. Alemanha: Paderborn Center for Paraller Computing, v0.8, 114p, Outubro. 2002.
- [PBS 02] JONES, James P., et al. **Portable Batch System User Guide**. Veridian Sytem, v5.2, 127p, Março. 2002.