

Instruções para configuração de um Cluster Linux com a distribuição Debian GNU/Linux 3.0

Marcelo Linder, Paulo Teixeira, Ernesto Alquati

Fundação Universidade Federal do Rio Grande - FURG
Avenida Itália, km 8 – Campus Carreiros – CEP 96201-900 – Rio Grande/RS, Tel. (53)2336620
ec7msl@furg.br, teixeira@dmc.furg.br, alquati@dmc.furg.br

Introdução

A necessidade de executar simulações numéricas usando computação com alto desempenho é crescente nesses últimos anos. Os supercomputadores, que possuem esta característica, estão fora do alcance da maioria das empresas e instituições de ensino devido ao seu alto custo. Pode-se conseguir poder computacional semelhante agrupando computadores pessoais, denominado de cluster, que possuem um custo benefício bem mais baixo que o dos supercomputadores.

A computação agrupada, em seu nível mais básico, envolve dois ou mais computadores, servindo em único recurso. Baseado nisso, a área de simulação numérica de escoamentos de fluidos do Departamento de Materiais e Construção (DMC) da FURG optou por montar um cluster para computação paralela de alto desempenho, visto que os códigos e os problemas envolvidos exigem muito tempo de processamento. Este documento é resultado de estudos na busca por um cluster adequado às necessidades e aos recursos disponíveis, contendo as instruções necessárias para se configurar um Cluster Linux com a Distribuição Debian GNU/Linux 3.0.

Basicamente, um cluster é constituído por uma máquina denominada de servidor e pelos nodos clientes. O servidor é responsável pela interface com o operador, por receber os dados, pela paralelização do processamento entre os nodos clientes e pelo recolhimento dos resultados obtidos pelos nodos clientes, organizando-os para obter-se a resposta final do processamento. Os nodos clientes são máquinas utilizadas para realizar parte do processamento exigido pelo código.

A sincronização da operação será obtida utilizando métodos como Parallel Virtual Machine (PVM – máquina virtual paralela) ou Message Passing Interface (MPI – interface para passagem de mensagens). O primeiro capacita a execução de um código através de vários nós, permitindo que um grupo heterogêneo de máquinas execute C, C++ e Fortran através do cluster. O segundo é um protocolo que utiliza a passagem de mensagens para sincronização de processos.

A máquina servidora é composta por uma placa mãe, um processador (Athlon XP 2000+), uma HD (disco rígido de 40 GB), memória RAM (1 GB), duas placas de rede (uma para a rede privada do cluster e outra para a interface com o meio externo), drives de CD e disquete, o servidor possui periféricos de entrada e saída, como, teclado monitor, etc. Atualmente tem-se apenas um nodo cliente, composto por uma placa mãe, um processador (Athlon XP 2000+), uma placa de rede (com suporte a boot remoto - PXE) e memória RAM (1 GB), os nodos clientes não possuem HD (são clientes diskless).

Configurando o servidor

Os nodos clientes necessitam inicializar pela rede, uma vez que não possuem HD. Por esta razão, a máquina servidora é um servidor de DHCP (Protocolo de Configuração Dinâmica de Hosts), TFTP (Protocolo Trivial de Transferência de Arquivos) e NFS (Sistema de Arquivos de Rede) para os nodos clientes. O servidor DHCP é usado para informar ao nodo cliente o seu endereço IP, outras informações da rede (máscara de subrede, etc.) e o local do gerenciador de inicialização. Já o servidor TFTP é usado para enviar a imagem do kernel e por sua vez o diretório raiz (root) do nodo cliente é montado no servidor via NFS. A máquina servidora também é um servidor NIS para que os usuários e senhas sejam os mesmos tanto no servidor quanto nos nodos clientes.

Instalando a distribuição

Deve-se tomar cuidado durante o processo de instalação da distribuição, pois será indagado sobre a instalação do Md5 passwords, devendo ser dada uma resposta negativa a essa pergunta, pois esse pacote pode conflitar com o pacote "NIS" que será instalado para gerenciar as senhas dos usuários do cluster.

Inserindo novos pacotes

Para inserir novos pacotes deve-se, primeiramente, incluir os CDs da distribuição como fontes para o "apt-get". Para isso, deve-se inserir cada CD no drive de cdrom e, como super usuário, executar no Shell a linha "#apt-cdrom add". Logo após, basta apenas executar no Shell a linha "#apt-get install pacote_desejado".

Os pacotes que devem ser inseridos são:

- dhcp – protocolo de configuração dinâmica de hosts;
- atftpd – protocolo trivial avançado de transferência de arquivos;
- nis – protocolo controlador de senhas de usuários;

Configurando o servidor DHCP

Para configurar o DHCP deve-se editar o arquivo "/etc/dhcpd.conf". Segue abaixo um arquivo exemplo quando a inicialização se dá através de uma placa com suporte a PXE (Pre-boot Execution Environment):

```
# dhcpd.conf
# option definitions common to all supported networks...
option domain-name "cluster";
option domain-name-servers 192.168.1.1;
use-host-decl-names on;
option subnet-mask 255.255.255.0;
option broadcast-address 192.168.1.255;
option routers 192.168.1.1;
default-lease-time 600;
max-lease-time 7200;
subnet 192.168.1.0 netmask 255.255.255.0 {
    filename "/tftpboot/pxelinux.0";
    host nol {
        hardware ethernet 00:0A:01:D4:CA:30;
        fixed-address 192.168.1.2;
    }
}
```

Configurando o servidor ATFTP

Primeiro, deve-se criar o diretório "tftpboot". Em seguida, copiar para este diretório o arquivo pxelinux.0, que é o gerenciador de inicialização, e o arquivo imagem do kernel do nó. Depois, criar o subdiretório "tftpboot/pxelinux.cfg" e nele criar o arquivo "default". Segue um exemplo do seu conteúdo.

```
#default
timeout 3
default linux
```

```
prompt 1
label linux
    kernel kernel-apm
    append ip=:::dhcp nfsroot=192.168.1.1:/tftpboot/nol
```

Onde `kernel-apm` é o arquivo imagem do kernel do nó e `192.168.1.1:/tftpboot/nol` é o alvo para a montagem por nfs do diretório “root” do nó.

Deve-se também editar os arquivos:

- “`/etc/exports`”. Exemplo:

```
# /etc/exports: the access control list for filesystems which may be
exported
#           to NFS clients.  See exports(5).
/tftpboot  *(rw,no_root_squash)
/home      *(rw,no_root_squash)
```

- “`/etc/hosts`”. Exemplo:

```
127.0.0.1    localhost
192.168.1.1 servidor.cluster  servidor
192.168.1.2 nol.cluster      nol
# The following lines are desirable for IPv6 capable hosts
# (added automatically by netbase upgrade)
::1        ip6-localhost ip6-loopback
fe00::0    ip6-localnet
ff00::0    ip6-mcastprefix
ff02::1    ip6-allnodes
ff02::2    ip6-allrouters
ff02::3    ip6-allhosts
```

Configurando o servidor NFS

Para configurar o servidor de NFS deve-se compilar o kernel com suporte a NFS (não esquecer também de dar suporte a DHCP). Antes da compilação do kernel deve-se, primeiro, instalar com o ***apt-get*** o pacote com as fontes do kernel, executando no Shell a linha “`#apt-get install kernel-source-2.4.18`”. Após, descompactar o arquivo com as fontes do kernel indo para o diretório “`/usr/src`” e executar no Shell, como super usuário, a linha “`#bunzip2 kernel-source-2.4.18.tar.bz2`” e a linha “`#tar xvf kernel-source-2.4.18.tar`”. Depois, no diretório “`/usr/src/kernel-sources-2.4.18`”, executar as seguintes linhas no Shell como super usuário:

1. `#make mrproper;`
2. `#make xconfig`, (estabelecer as configurações de sua máquina, como processador, placa de som e etc). Este comando deve ser executado no modo gráfico;
3. `#make dep;`
4. `#make bzImage;`
5. `#make install;`
6. `#make modules;`
7. `#make modules_install.`

Configurando o servidor NIS

Editar os arquivos “`/etc/init.d/nis`” e “`/etc/default/nis`” alterando a linha “`NISSERVER=false`” para “`NISSERVER=master`”. Depois, no diretório “`/var/yp`”, executar “`make`” como super usuário (o `make` deve ser executado com os nodos clientes ligados).

Configurando os nodos clientes

A configuração de um nodo cliente é simples, já que ele é apenas um cliente dos serviços fornecidos pela máquina servidora.

Gerando a imagem que será importada via TFTP

Para gerar a imagem do kernel de uma máquina cliente, deve-se compilar um kernel (como foi explicado anteriormente) de acordo com as características da máquina cliente (processador, placa de rede e etc). Também se deve dar suporte a cliente DHCP e NFS.

Construindo o diretório raiz (root)

Na máquina servidora, deve ser construído o diretório raiz de cada nodo cliente o qual será montado pelo nodo cliente via NFS (durante a sua inicialização). Neste diretório, devem constar todos os arquivos e subdiretórios necessários para a execução/funcionamento do nodo cliente (exemplo: /etc/init.d, /etc/dhcpd.conf, /var/ e etc.). Deve-se configurar o arquivo “etc/fstab” do nodo cliente (indicando que o diretório raiz será montado via NFS). Abaixo segue um arquivo exemplo:

```
# /etc/fstab: static file system information.
# <file system> <mount point> <type> <options>
# <dump> <pass>
192.168.1.1:/tftpboot/nol / nfs defaults 0 1
proc /proc proc defaults 0 0
servidor:/tftpboot/g_all/usr /usr nfs defaults 0 0
servidor:/home /home nfs defaults,soft,bg 0 0
```

Conclusão

Neste trabalho apresentou-se as instruções necessárias para se configurar um Cluster Linux com a Distribuição Debian GNU/Linux 3.0, após os processos citados, o cluster possuirá as máquinas inicializáveis e com a rede configurada. A partir daí basta instalar uma ou mais bibliotecas de paralelização e sincronização de processos como, por exemplo: o PVM ou o MPI. Tornando o cluster funcional, ou seja, capaz de paralelizar processos.

Referências

- [BOO 02] BOOKMAN, C. **Linux Clustering: Building and Maintaining**, ed. New Riders, 2002.
- [PIT 02] PITANGA, M. **Construindo Supercomputadores com Linux**, ed. Brasport, 2002.
- [RUD 02] RUDOLPH, S., et al. **Instalando Debian GNU/Linux 3.0 para Intel x86**, versão 3.0.24, 18 December, 2002; disponível on line em <http://www.br.debian.org/releases/stable/i386/install.pt.txt> .