

Alta disponibilidade utilizando Heartbeat e DRBD

Alexandre Mulatinho <alex@mulatinho.net> 2012/2

Abstract. This article have as principal objective describe the high availability function on linux using tools like heartbeat and drbd also show some characteristics and specific configurations of this solution in an virtual environment. **Resumo.** Este artigo tem como principal objetivo descrever o funcionamento da alta disponibilidade no linux utilizando as ferramentas heartbeat e o drbd, além de mostrar algumas características e configurações específicas desta solução num ambiente virtual.

1. Conceitos Básicos

A alta disponibilidade nos dias de hoje torna-se um fator crucial tanto para pequenos, médios e grandes negócios que envolvem sistemas essenciais, sistemas como os de saúde, bancos e do governo em geral por exemplo tem necessidade de estarem online 24 (vinte e quatro) horas durante os 7 (sete) dias da semana devido sua criticidade e por isso cada minuto em que ficam de fora eles são responsabilizados de forma severa.

Para que não haja downtime nesses serviços é necessário a implementação de ambientes descentralizados contendo uma coleção de computadores independentes que apresentem ao usuário um sistema único e independente [Tanenbaum 2002]. Neste artigo iremos mostrar um exemplo de como implementar este conceito utilizando como exemplo uma máquina linux que tenha como serviço critico um banco de dados mysql e um software de monitoração de desempenho chamado cacti.

1.1 Definições Principais

O heartbeat é a principal ferramenta do conceito de alta disponibilidade no linux, ele age como gestor do cluster e de seus serviços enviando periodicamente sinais (pacotes ICMP) através do protocolo UDP para os nós do cluster (máquinas e/ou serviços), estes sinais são chamados de hearbeats (batidas de coração), e tem como principal objetivo identificar um sinal positivo ou a ausência do sinal o que faz com que o heartbeat marque o servidor como parado e tome as atitudes para que os serviços não fiquem indisponíveis para os usuários, nós iremos ver isso mais detalhadamente depois.

O DRBD é um acrônimo em inglês que traduzido significa dispositivo de blocos distribuídos e replicados, ele consiste principalmente de um módulo desenvolvido para o kernel do linux e alguns scripts que tem a função de replicar diretórios e arquivos (sistema de arquivos) entre várias máquinas distintas, isto é compartilhando seus recursos.2. Cenário Para exemplificar melhor esta solução vamos fornecer um passo a passo de instalação e configuração dessa solução para que possamos discutir melhor as técnicas e as tecnologias utilizadas nesta solução. Utilizaremos duas máquinas:

Máquina-01: 192.168.5.1 (CentOS 5.7) com o MySQL 5 e o Cacti instalados.

Máquina-02: 192.168.5.2 (CentOS 5.7) com o MySQL 5 e o Cacti instalados.

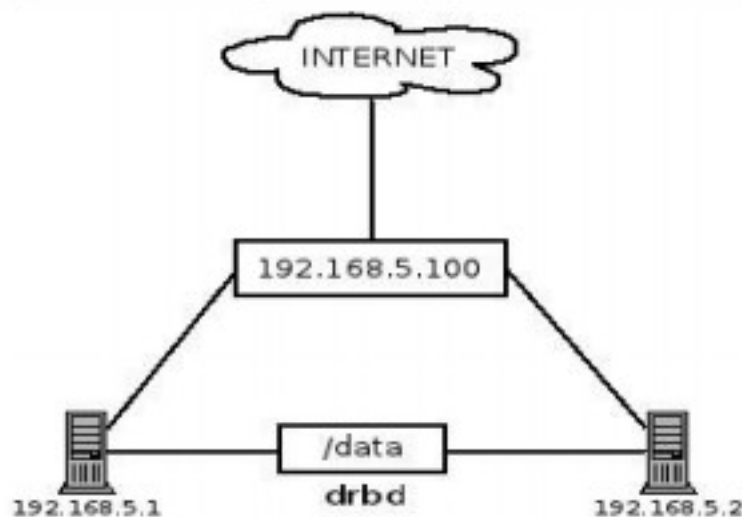


Figura 1: Cluster utilizando heartbeat e o drbd compartilhando o diretório /data

Como pode ser visto na figura 1, para o usuário que está na internet toda a solução de alta disponibilidade é transparente pois ele sempre irá acessar, enviar e receber os dados através do IP 192.168.5.100. O que ele não sabe é que internamente existem dois servidores que tem a mesma função: fornecer o serviço do cacti e o banco de dados mesmo que uma destas máquinas parem, isso se torna crucial para sistemas críticos pois permite que caso aconteça algum sinistro com um destes servidores (problemas de hardware, software, energia, etc.) o outro assuma.

2.1. Configurando o HeartBeat

O principal responsável por esta solução é o heartbeat, ele age como um gerenciador do cluster controlando os nós (máquinas clientes) e fornecendo ao usuário final a sensação de transparência e principalmente garantindo disponibilidade do serviço. O principal arquivo de configuração chama-se `"/etc/ha.d/ha.cf"` e pode ser visto abaixo:

1. debugfile /var/log/ha-debug
2. logfile /var/log/ha-log
3. logfacility local0
4. keepalive 2 (Segundos)
5. deadtime 30 (Segundos)
6. bcast eth0
7. udpport 694 # porta UDP utilizada para intra-cluster# communication
8. auto_failback yes # Retorna serviço para master
9. node maquina-01 # nome das máquinas do cluster
10. node maquina-02 # nome das máquinas do cluster

(1) define log de debug; (2) define arquivo de log; (3) define tipo de log; (4) Intervalo entre os heartbeats; (5) Define quando um nó está offline; (6) Qual interface os heartbeats serão enviados; (7) Utilizando porta UDP 694 para comunicação; (8) Sempre tenta retornar serviço para master; (9,10) define os nomes das máquinas que serão os nós do cluster e devem estar contidos no arquivo `"/etc/hosts"`.

O arquivo “/etc/ha.d/haresources” contém as informações dos recursos que serão compartilhados na alta disponibilidade, neste caso nós informamos a ele que precisamos compartilhar um disco via drbd e dois serviços: o httpd e o mysql, conforme pode ser visto na primeira linha abaixo:

```
maquina-01 192.168.5.3 drbddisk Filesystem::/dev/drbd0::/opt/ha::ext3 httpd mysql  
(nó principal do cluster; ip do cluster; sistema de arquivos; serviços compartilhados;)
```

Apartir daí basta inicializar o serviço do heartbeat que ele vai tentar se comunicar com os dois nós e tentar inicializar os recursos citados acima na máquina principal do cluster, mas antes precisamos configurar o DRBD.

2.2. Configurando o DRBD

O principal arquivo de configuração do DRBD é o “/etc/drbd.conf”, nele pode-se fazer diversas configurações alternativas e a maioria delas pode ser encontrada na página de manual do “drbd.conf”, no nosso caso nós iremos apenas adicionar dois nós que compartilham um mesmo sistema de arquivos através do dispositivo de blocos drbd0.

```
1. resource httpd {  
2.     protocol C;  
3.     on maquina-01 {  
4.         device /dev/drbd0;  
5.         disk /dev/sdb1;  
6.         address 192.168.5.1:7788;  
7.         meta-disk internal;  
8.     }  
9.     on maquina-02 {  
10.        device /dev/drbd0;  
11.        disk /dev/sdb1;  
12.        address 192.168.5.2:7788;  
13.        meta-disk internal;  
14.    }  
15. }
```

(1) informas um nome para categoria principal; (2) define que a entrada/saída de dados será replicada tanto localmente como remotamente; (3,9) define o nome da maquina a ser usada, deve estar no “/etc/hosts”; (4,10) informa o dispositivo de bloco a ser utilizado; (5) informa a partição usada; (6) registra o ip dos nós e a porta de comunicação do drbd; (7,13) diz ao drbd que a última parte do disco será para a escrita dos meta-dados;

2.2. Analisando uma parada e uma inicialização

Após inicializar os serviços corretamente nós podemos ver abaixo a saída do log “/var/log/ha-log” em duas situações diferentes, a primeira é onde nós simulamos a parada da maquina-01 semelhante ao que ocorreria em caso de uma queda de energia, problema de hardware e/ou software. O log então registra a máquina indo para baixo:

```
1. heartbeat[9952]: 2012/05/23_00:55:28 info: maquina-01 wants to go standby [all]  
2. heartbeat[9952]: 2012/05/23_00:55:28 info: standby: maquina-02 can take our all resources  
3. heartbeat[11307]: 2012/05/23_00:55:28 info: give up all HA resources (standby).4.
```

```
.....
5. IPAddr[11591]: 2012/05/23_00:55:36 INFO: ifconfig eth0:0 down
6. IPAddr[11574]: 2012/05/23_00:55:36 INFO: Success
7. heartbeat[11307]: 2012/05/23_00:55:36 info: all HA resource release completed (standby).
```

Do outro lado a maquina-02 verificando a ausência de heartbeats da primeira maquina confirma que existe algum problema e começa a iniciar o procedimento de tomada de recursos, isso quer dizer que agora a máquina principal do cluster que recebe as requisições não é mais a maquina-01. Veja a máquina obtendo acesso:

```
1. harc[10080]: 2012/05/23_00:41:13 info: Running /etc/ha.d/rc.d/ip-request-resp ip-request-resp
2. ip-request-resp[10080]: 2012/05/23_00:41:13 received ip-request-resp 192.168.5.3 OK yes
3. ResourceManager[10101]: 2012/05/23_00:41:13 info: Acquiring resource group: maquina-01
192.168.5.3 drbddisk Filesystem::/dev/drbd0::/opt/ha::ext3 httpd mysqld
4. ResourceManager[10101]: 2012/05/23_00:41:14 info: Running /etc/ha.d/resource.d/IPAddr
192.168.5.3 start
5. IPAddr[10204]: 2012/05/23_00:41:15 INFO: Using calculated nic for 192.168.5.3: eth0
6. IPAddr[10204]: 2012/05/23_00:41:15 INFO: Using calculated netmask for 192.168.5.3:
255.255.255.0
7. IPAddr[10204]: 2012/05/23_00:41:15 INFO: eval ifconfig eth0:0 192.168.5.3 netmask 255.255.255.0
broadcast 192.168.5.255
8. IPAddr[10187]: 2012/05/23_00:41:16 INFO: Success
```

(2,4) Começa o procedimento de aquisição do IP virtual do cluster para a maquina local; (3) Também dá inicio a aquisição do recurso de disco do drbd na maquina-02; (8) Confirma o recebimento do IP virtual; O DRBD então age disponibilizando o disco remoto e o montando localmente:

```
1. ResourceManager[10101]: 2012/05/23_00:41:16 info: Running /etc/ha.d/resource.d/drbddisk start
2. ResourceManager[10101]: 2012/05/23_00:41:17 info: Running /etc/ha.d/resource.d/Filesystem
/dev/drbd0 /opt/ha ext3 start
3. Filesystem[10428]: 2012/05/23_00:41:17 INFO: Running start for /dev/drbd0 on /opt/ha
4. Filesystem[10417]: 2012/05/23_00:41:18 INFO: Success
```

(1) Inicia comando que toma o disco para maquina local; (2,3,4) Confirma a montagem do disco na máquina local utilizando um script dentro do hearbeat que se comunica com o DRBD. Os serviços são inicializados localmente,

```
1. ResourceManager[10101]: 2012/05/23_00:41:18 info: Running /etc/init.d/httpd start
2. ResourceManager[10101]: 2012/05/23_00:41:19 info: Running /etc/init.d/mysqld start
```

Por fim, inicia os serviços que nós listamos no arquivo de configuração “haresources”. Acontece então a mensagem final no log de confirmação de aquisição dos recursos,

```
1. heartbeat[11660]: 2012/05/23_00:57:00 info: local HA resource acquisition completed (standby).
2. heartbeat[9952]: 2012/05/23_00:57:00 info: Standby resource acquisition done [foreign].
3. heartbeat[9952]: 2012/05/23_00:57:00 info: remote resource transition completed.
```

E pronto!! A máquina confirma através dos logs do heartbeat que todos os recursos compartilhados agora pertencem a máquina local deixando não havendo portanto nenhum downtime nos serviços.

3. Conclusão

O conceito de Alta Disponibilidade (High Availability, ou simplesmente HA) não deve, e nem pode, ser uma novidade para os administradores de sistemas, principalmente em ambientes computacionais onde a disponibilidade é uma característica crítica. Esse conceito, que não é recente, está sendo amplamente disseminado, alcançando importância diretamente proporcional a influência que os sistemas de computação exercem nas empresas ou entidades que sustentam.

Em meio a algumas soluções proprietárias para construir um ambiente de alta disponibilidade, o Heartbeat e o DRBD conseguem juntos excelentes resultados para sistemas Linux, e sem absolutamente nenhum custo por se tratarem de softwares livres. A maturidade desses softwares atingiu um certo nível que venceu determinadas desconfianças presentes no rigoroso mercado corporativo.

Portanto torna-se essencial para uma empresa séria implementar recursos de alta disponibilidade afim de garantir que os seus serviços e servidores nunca fiquem for a do ar.

4. Bibliografia

[Tanenbaum 1995]
Distributed Operating Systems. A. S. Tanenbaum.
Prentice-Hall, 1995. ISBN 0-13-219908-4.

[Linux-HA]
The definitive reference guide for users of the Heartbeat
<http://www.linux-ha.org/doc/users-guide/users-guide.html>