9º FEPEG FÓRUM DE ENSINO, PESQUISA, EXTENSÃO E GESTÃO
23 A 26 SETEMBRO DE 2015
Campus Universitário Professor Darcy Ribeiro
TRABALHOS CIENTÍFICOS APRESENTAÇÕES ARTÍSTICAS E CULTURAIS DEBATES MINICURSOS E PALESTRAS
ISSN 1806-549X
A HUMANIZAÇÃO NA CIÊNCIA, TECNOLOGIA E INOVAÇÃO

Implementação de Alta Disponibilidade em Network file System com DRBD e Heartbeat

Hugo Luciano de Souza Leal, Gabriela Santos Silva, Raimundo Sebastião Teodoro Santana, Rodrigo Everton Soares Oliveira, Michael Jonnes Lopes Souza

Introdução

Network File Systems (NFS), ou sistemas de arquivos em rede, provêem acesso remoto a diretórios de arquivos de modo transparente, como se o acesso fosse dado localmente. O serviço de Network File System normalmente é usado como solução central para backups, armazenamento e compartilhamento de arquivos entre máquinas de uma rede, podendo armazenar até todos os arquivos de diferentes usuários de uma organização. Percebe-se a necessidade de segurança e disponibilidade dos dados em Network File Systems, demandas que podem ser melhor atendidas com a implementação de um Sistema de Arquivos em Rede com alta disponibilidade.

Este trabalho visa documentar uma implementação de alta disponibilidade em Network File System na plataforma Linux, servindo como base para implementações futuras que seguem o mesmo gênero. Foi desenvolvido como requisito parcial para aprovação na disciplina de Sistemas Distribuídos II e utiliza as ferramentas Linux NFS, Distributed Replicated Block Device (DRBD) e Heartbeat.

Material e métodos

A. Alta disponibilidade

Um sistema de alta disponibilidade possui mecanismos para detecção, mascaramento e recuperação de falhas. O intuito de um sistema de alta disponibilidade é manter o maior tempo possível em funcionamento, de maneira que se os serviços oferecidos pelo sistema possam ser acessados pelos usuários sem que estes percebam falhas ocorridas e sem que os serviços estejam indisponíveis.

B. Network File System

O Network File System (NFS), ou Sistema de Arquivos em Rede, é uma maneira de compartilhar arquivos entre máquinas de uma rede, como se estes arquivos estivessem localizados no disco rígido local do cliente. O cliente Network File System tem por finalidade tornar o acesso remoto transparente para o usuário do computador, e esta interface cliente e servidor, executada pelo Network File System através da arquitetura Cliente-Servidor, fica bem definida quando o usuário ao chamar um arquivo no servidor, lhe parece estar acessando localmente, sendo que está trabalhando com arquivos remotos. Portanto, os diretórios são acessados em máquinas remotas.

Para que os clientes tenham acesso aos arquivos, é feita uma requisição ao servidor que, dependendo das permissões do cliente, responde confirmando a requisição. A partir desse ponto a hierarquia de diretórios e arquivos remotos passa a fazer parte do sistema de arquivos local da máquina. O Network File System é comumente dito como um servidor de arquivos. Ele é também uma ferramenta poderosa, e a mais indicada para montar uma rede de compartilhamento de arquivos em redes com máquinas Linux. Ele segue o modelo computacional Cliente/Servidor. O servidor implementa o sistema de arquivos e o armazenamento compartilhado aos quais os clientes se conectam. Os clientes implementam a interface com o usuário para o sistema de arquivo compartilhado, disposto no espaço no arquivo do cliente.

No Linux, o comutador do sistema de arquivo virtual fornece os meios para suportar vários sistemas de arquivos simultaneamente em um host. Ele determina para qual armazenamento uma solicitação é destinada e qual sistema de arquivos deve ser usado para satisfazer a solicitação. Quando se sabe que a solicitação é destinada para o Network File System, o sistema de arquivo virtual a passa para a instância do Network File System no kernel. Ele interpreta a solicitação de entrada e saída e a converte para um procedimento do Network File System. Quando um procedimento é selecionado de uma solicitação de entrada e saída, ele é realizado na camada de chamada de procedimento remoto - Remote Procedure Call (RPC). Como o nome sugere, ele fornece os meios para a realização de chamadas de procedimentos entre sistemas.

Neste trabalho o cliente Network File System foi instalado na máquina cliente com o comando “apt-get install nfs-common”. O Servidor NFS foi instalado nas máquinas server1 e server2 com o comando “apt-get install nfs-kernel-server”.

Para configurar o servidor Network File System apenas foi inserida nos dois nós do cluster a linha “/dados 10.0.0.247(rw,sync,no_root_squash)” no arquivo “/etc/exports” fornecendo permissão para a máquina de ip 10.0.0.247

para acesso de leitura, escrita e modificação. Após isso o serviço de NFS foi reiniciado como comando "/etc/init.d/nfs-kernel-server restart".

No cliente foi configurado um ponto de montagem com os comandos "mkdir /dados" e "mount -t nfs 10.0.0.10:/dados /dados". Com isso foi estabelecido o acesso do cliente ao diretório "/dados" no NFS. Após isso, foi configurada a montagem automática do diretório remoto no cliente inserindo a linha "10.0.0.10:/dados /dados/ nfs defaults 0 0" em "/etc/fstab".

C. Distributed Replicated Block Device

O Distributed Replicated Block Device (DRDB) se trata de um módulo para o núcleo do sistema Linux que executa a função de espelhamento dos dados de uma partição de disco entre servidores de uma rede.

As partições envolvidas no Distributed Replicated Block Device possuem um estado, primário ou secundário. As operações de escrita são realizadas no primário e replicadas para o secundário. Um protocolo replicação proporciona a sincronia e a integridade dos dados replicados. As operações de leitura são localmente.

O Distributed Replicated Block Device permitirá o espelhamento, replicação, de uma partição entre máquinas, portanto será instalado nas máquinas server1 e server2. O relógio das máquinas precisa estar sincronizado, para tanto foi instalado um cliente Network File Protocol (NTP) como o comando "aptitude install ntpdate tzdata". Em seguida foram utilizados os comandos "ntpdate a.ntp.br" e " hwclock –systohc" para sincronizar o relógio das máquinas com a servidor Network File Protocol em 'a.ntp.br'

Para instalar o Distributed Replicated Block Device em um sistema operacioanl Linux basta acessar o shell com permissão de usuário root e executar o comando: "aptitude install drbd8-utils". A configuração da replicação entre partições é relativamente simples pois se dá somente em dois arquivos o "/etc/drbd.d/global_common.conf" e o arquivo de recurso para cluster Distributed Replicated Block Device, como o "/etc/drbd.d/r0.res" que foi criado para esta implementação.

O arquivo "/etc/drbd.d/global_common.conf" recebe as configurações comuns e globais do cluster Distributed Replicated Block Device, como tempo limite do ping, tamanho máximo do buffer, e outras. O arquivo deve ser configurado igualmente em todos os nós do cluster. O arquivo para configurar o recurso de replicação "/etc/drbd.d/r0.res" consiste em informações sobre os nós do cluster, como endereço ip e porta de comunicação e qual partição será replicada. Este arquivo deve ser configurado igualmente em todos em nós do cluster. Tal arquivo foi configurado como mostra a figura 3. Após feitas as configurações, os recursos de replicação foram iniciados nos dois nós através dos comandos "drbdadm create-md r0", " modprobe drbd" e "drbdadm up r0". Com isso, foi feita a sincronização dos nós pela primeira fez através do comando "drbdadm -- --overwrite-data-of-peer primary r0" executado no server1.

Com os nós sincronizados foi feita a formatação da partição "mkfs.ext4 /dev/drbd0" e montagem do diretório que será acessado pelos clientes com "mkdir /dados" e "mount -t ext4 /dev/drbd0 /dados".

Para certificar se a replicação funcionou corretamente foi criado o arquivo "teste.txt" dentro do diretório "cd /dados", após isso o diretório foi desmontado com o comando "umount /dados" aplicado estando no diretório "cd /", o server1 foi definido como nó secundário do cluster com "drbdadm secondary all", o server2foi definido como nó primário com "drbdadm primary all" e, para verificar que arquivo criado foi replicado para a partição do server2 foi montado o diretório "cd /dados" no server2 com o comando e "mount -t ext4 /dev/drbd0 /dados" e listado o conteúdo do repositório com o comando "ls".

D. Heartbeat

O Heartbeat, considerado o núcleo do ambiente de alta disponibilidade, age sobre as falhas, monitora os servidores em produção e realiza automaticamente procedimentos para manter o sistema em funcionamento, de modo transparente.

Instâncias do Heartbeat são instaladas no servidor primário e secundário. O secundário monitora o funcionamento do primário e, caso este falha, o secundário assume o estado de primário e passa a ser acessado diretamente pelo sistema. Para tanto, o servidor secundário assume o endereço IP do serviço e montas as partições de disco para acesso.

Com o cluster Distributed Replicated Block Device funcionando corretamente, foi instalado o Heartbeat atrvés do comando "aptitude install heartbeat". Uma vez instalado, foi necessário configurar os arquivos "/etc/ha.d/authkeys", "/etc/ha.d/ha.cf" e "/etc/ha.d/haresources" que foram copiados da pasta "/usr/share/doc/heartbeat/", pois não estão presentes no Xubuntu por padrão. O arquivo "/etc/ha.d/ha.cf" define a interface de broadcast, o modo de recuperação à falhas e os nós que serão analisados pelo heartbeat. Ficou configurado como mostra a figura 4.

O arquivo "/etc/ha.d/authkeys" define a senha para o serviço de alta disponibilidade e seu modo de criptografia o arquivo deve ter suas permissões alteradas como comando "chmod 600 /etc/ha.d/authkeys".

O arquivo "/etc/ha.d/haresources" define quais serviços serão inspecionados pelo Heartbeat e o ip virtual que responderá pelo serviço.

A montagem da partição foi colocada para ser feita de modo automático pela inserção da linha "/dev/drbd0 /dados ext4 _netdev,defaults 0 0" no arquivo "/etc/fstab".

Com as configurações finalizadas, o serviço do Heartbeat foi iniciado nas duas máquinas pelo comando "/etc/init.d/heartbeat start".

E. Ambiente Virtual

A implementação da alta disponibilidade em Network File System foi feita em ambiente virtual configurado com o Oracle Virtual Box. Foram utilizadas quatro máquinas virtuais sendo três com sistema operacional Linux Xubuntu e uma com o MikrotikOS.

As máquinas virtuais foram inseridas no modo rede interna do VirtualBox, que fornece um ambiente virtual de rede isolado, sem comunicação com a máquina física nem com a rede física onde esta se encontra. Cada máquina virtual possui uma interface de rede na rede interna criada.

O MikrotikOS foi configurado para fornecer acesso à Internet para as máquinas da rede interna. Para tanto, este foi configurado com duas interfaces de rede, sendo, uma na rede interna, atuando como DHCP Server, e outra em NAT com a rede física. O MikrotikOS obtém acesso à internet pela sua interface que faz NAT com a rede física e possibilita o acesso à Internet das máquinas que estão na rede interna atuando como Gateway e mascarando o acesso da rede interna para a placa de rede física.

Das três máquinas virtuais Linux, duas foram configuradas em cluster através do Distributed Replicated Block Device formando o Network File System. A terceira máquina virtual Linux foi configurada como cliente para o Network File System.

Para fins de melhor entendimento das configurações abordadas nas próximas sessões, deve-se adotar como 'hostname = server1 e ip=10.0.0.249' a máquina com O Linux que será configurada como servidor primary do cluster Distributed Replicated Block Device e 'hostname = server2 e ip=10.0.0.248' a máquina com OS Linux que será configurada como servidor secondary do cluster Distributed Replicated Block Device.

Resultados

Os resultados obtidos com a implementação, bem como uma demonstração prática da alta disponibilidade do cluster, são apresentadas no vídeo que acompanha este documento. Foi observado o correto funcionamento das ferramentas utilizadas e sua fácil configuração. A implementação realmente resultou em um sistema de arquivos de alta disponibilidade e transparência para os clientes.

Conclusão

Conclui-se que a implementação de um Network File System de alta disponibilidade utilizando o Linux NFS, o Distributed Replicated Block Device e o Heartbeat provê uma configuração simples e funciona da maneira esperada. Tais recursos podem ser utilizados para aumentar a segurança e disponibilidade dos arquivos de uma organização de modo simples e barato, uma vez que todas as ferramentas necessárias para o cluster são gratuitas e fáceis de serem configuradas.

Referências

[1] DANTAS, Jamilson Ramalho. **Cluster de alta disponibilidade com arquitetura Heartbeat**, 2008. Disponível em: <http://www.fasete.edu.br/revistarios/media/revistas/2008/cluster_de_alta_disponibilidade_com_arquitetura_heartbeat.pdf.> Acesso em : 10 dez. 2014

[2] **How to Configure NFS in Linux**. Disponível em: <http://linuxconfig.org/how-to-configure-nfs-on-linux> Acesso em: 11 dez. 2014

[3] IBM Developers- **Network File System.** Disponível em: <http://www.ibm.com/developerwors/br/library/1-network-filesystems/ > Acesso em: 11 dez. 2014

[4] SMITH, Christopher. **Linux NFS-HOWTO**, 2006. Disponível em: <nfs.sourceforge.net/nfs-howto/> Acesso em: 12 dez. 2014

[5] HASS, Florian. REINER Philipp. ELLENBERG, Lars. **The DRBD User's Guide**. Disponível em: <drbd.linbit.com/users-guide-8.3/> Acesso em: 11 dez. 2014