**集群与存储（**理论重听一遍11：00**）**

简介（集群是什么）  
 一组通过高速网络互联的计算机组，并以单一系统的模式加一管理

将很多服务器集中起来一起，提供一种服务，在客户端看起来就像是只有一个服务器

可以在付出较低成本的情况下获得在性能、可靠性、灵活性方面的相对较高的收益

任务调度是集群系统中的核心技术

集群目的：

提高性能：计算密集型应用，如：天气预报，核试验模拟

降低成本：相对百万美元的超级计算机，价格便宜

提高可扩展性：只要增加集群节点就可以

增强可靠性：多个节点完成相同功能，避免单点失败的故障

集群分类：

高可用计算集群HPC

通过以集群开发的并行应用程序，解决复杂的科学问题

应用：在天气、航空、航海等科学领域的科学计算

负载均衡（LB）集群

客户端负载在计算机集群中尽可能平均分摊

高可用（HA）集群

避免单点故障，当一个系统发生故障时，可以快速迁移

LB的实现：LVS软件 Haproxy软件 Nginx

HA的实现：keepalived软件

LVS项目介绍：

Linux虚拟服务器（LVS）是章文嵩在国防科技大学就读博士期间创建的

LVS可以实现高可用的、可伸缩的web、Mail、Cache和Media等网络服务

最终目标时利用Linux操作系统和LVS集群和LVS集群软件实现一个高可用、高性能、低成本的服务器应用集群

LVS集群组成

前端：负载均衡层，由一台或多台

中间：服务器群组层，由一组实际运行应用服务的服务器组成

地端：数据共享存储层，提供共享存储空间的存储区域

LVS术语

Director Server ： 度服务器

将负载分发到Real server的服务器

Real server：真实服务器

真正提供应用服务的服务器

VIP：虚拟IP地址

公布给用户访问的细腻IP地址

RIP：真实IP地址

集群节点上使用IP地址

DIP：调度器连接节点服务器的IP地址

CIP：客户端的IP地址

LVS工作模式

VS/NAT

通过网络地址转换实现的虚拟服务器

VS/DR

直接使用路由技术实现虚拟服务器

节点服务器需要配置VIP，注MAC地址广播

VS/TUN

通过隧道凡是实现虚拟服务器

不同地点的不同机房需要提供服务时，就需要将LVS设置为VS/TUN模式

LVS负载均衡调度算法，目前实现了10中调度算法

轮询（Round Robin）

加权轮询（Weighted Round Robin）

最少连接（Least Connections）

加权最少连接（Weighted Least Connections）

LVS实验（轮询）：

LVS是应用在第四层，传输层，基于端口的软件

1. 配置LVS/NAT集群

主机角色：

Real server ：主机52和53

分发器 ：主机54

客户端 ：主机50

存储服务 ：主机51

实验准备：

分发器： 主机54

Eth0:私由网络 IP 地址 192.168.4.54

Eth1:公有网络Ip 地址：192.168.2.54

开启内核的路由转发功能

real server

主机52 eth0 192.168.4.52

主机网关：192.168.4.54

主机：53 eth0 192.168.4.53

网关：192.168.4.54

打开主机54内核路由转发功能

sysctl -a | grep ip\_forward //查看路由转发是否打开，默认是打开了的

vim /etc/sysctl.conf

net.ipv4.ip\_forward = 1 //新添加

Real server:

route命令用来显示并设置Linux内核中的网络路由表，route命令设置的路由主要是静态路由。

主机52/53：添加网关

Route add default gw 192.168.4.54

主机50：添加网关

Route add default gw 192.168.2.54

注：route del default gw 192.168.4.54 //配置错误以这条命令删除

Route -n //查看路由表

配置分发器主机54

ipvsadm 用法：

|  |  |
| --- | --- |
| -A | 添加虚拟服务器 |
| -t | 设置集合群地址（VIP，Virtual IP） |
| -s | 指定负载调度算法 |
| -a | 添加真实服务器 |
| -d | 删除真实服务器 |
| -r | 指定真实服务器（real server） |
| -m | 使用NAT模式；-g、-i分别对应DR、TUN模式 |
| -w | 为节点服务器设置权重，默认为1 |
| -C | 清除ipvsadm所有配置 |
| -S | 永久保存配置 |
| -E | 更改调度算法 |
| -e |  |
| -Z | 清空（ipvsadm -Ln --stats）计数器 |
| -D | 删除一个虚拟服务 |

NAT ：客户端发起请求将包发给VIP，分发器把VIP转换称DIP，DIP再去请求内部的（RIP）服务器，回包的时候，RIP把响应包发给DIP，调度器把DIP转换成VIP，VIP再把响应包发给（CIP）客户端

VIP地址：对外提供服务的地址

vip :192.168.2.54

rip :192.168.4.52

192.168.4.53

1.安装ipvsadm软件包

yum -y install ipvsadm.x86\_64

2.创建虚拟服务 采用rr调度算法，轮询算法；wrr

]# ipvsadm -A -t 192.168.2.54:80 -s rr

3.添加服务器节点（real server） 用-m采用NAT模式

]# ipvsadm -a -t 192.168.2.54:80 -r 192.168.4.52:80 -m

]# ipvsadm -a -t 192.168.2.54:80 -r 192.168.4.53:80 -m

4.保存ipvs配置

ipvsadm -S > /etc/sysconfig/ipvsadm-config

5.查看IPVS

]# ipvsadm -Ln

-> RemoteAddress:Port Forward Weight ActiveConn InActConn

TCP 192.168.4.54:80 rr

-> 192.168.4.52:80 Masq 1 0 0

-> 192.168.4.53:80 Masq 1 0 0

Forward：分发模式

ActiveConn：是活动连接数

Weight：权重，在负载均衡的时候，谁的值越大分发的次数越多

InActConn：

验证：

ipvsadm -Ln --stats

Prot LocalAddress:Port Conns InPkts OutPkts InBytes OutBytes

TCP 192.168.2.54:80 5 30 20 1980 2495

-> 192.168.4.52:80 2 12 8 792 998

-> 192.168.4.53:80 3 18 12 1188 1497

Conns：请求连接数

InPkts：进来的数据包

OutPkts：出去的数据包

InBytes：进来的数据字节数

OutBytes：出去进来的数据字节数

watch -n 1 ipvsadm -Ln --stats //watch -n 1 命令1秒刷新后面的命令

管理集群

nat模式的配置

删除

]# ipvsadm -d -t 192.168.2.54:80 -r 192.168.4.53

添加realserver（NAT）

]# ipvsadm -a -t 192.168.2.54:80 -r 192.168.4.53 -m

修改调度算法

]# ipvsadm -E -t 192.168.2.54:80 -s wrr

修改real server的权重

]# ipvsadm -e -t 192.168.2.54:80 -r 192.168.4.53 -w 2 -m

清空计数器（ipvsadm -Ln --stats的计数）

]# ipvsadm -Z

清空所有配置

]# ipvsadm -C

保存所以配置

]# ipvsadm -S > /etc/sysconfig/ipvsadm-config

重启ipvsadm服务

]# ipvsadm -S > /etc/sysconfig/ipvsadm

]# systemctl restart ipvsadm.service

配置LVS/DR模式集群

LVS的只要有配置，LVS这个服务不用启动

需要配置开机自启 systemctl enable ipvsadm

DR：CIP发起请求包，CIP发给VIP，调度器按照自己的分发策略将请求发给CIP，CIP把响应包直接回复给CIP

原理：请求流量由分发器下发给realserver，响应流量由服务器直接发给客户端

准备：

需求：客户端访问VIP地址：192.168.4.253访问网站集群

配置分发器 54

1.在本机的eth0接口上绑定vip地址192.168.4.253，这里命令行设置当 前有效重启后失效，将其夹入到/etc/rc.load文件中每次开机自动运行

]# ifconfig eth0:1 192.168.4.254/32

]# echo ‘ifconfig eth0:1 192.168.4.253/32’ >> /etc/rc.local

2.创建虚拟服务，以192.168.4.253为VIP 算法采用rr 轮询算法

]# ipvsadm -A -t 192.168.4.253:80 -s rr

3.添加real server节点，-g采用DR模式

]# ipvsadm -a -t 192.168.4.253:80 -r 192.168.4.52:80 -g

]# ipvsadm -a -t 192.168.4.253:80 -r 192.168.4.53:80 -g

4.保存ipvs配置

]# ipvsadm -S > /etc/sysconfig/ipvsadm-config

5.查看IPVS

]# ipvsadm -Ln

-> RemoteAddress:Port Forward Weight ActiveConn InActConn

TCP 192.168.4.253:80 rr

-> 192.168.4.52:80 Route 1 0 0

-> 192.168.4.53:80 Route 1 0 0

配置readlserver 52和53

1.修改网络接口的内核参数，主机52和53，为保以后每次开机都能生效，将以下命名追加至/etc/rc.load中,需要给/etc/rc.d/rc.local加x权限

echo 1 > /proc/sys/net/ipv4/conf/lo/arp\_ignore

网络接口可以和自己的兄弟接口接包和发包 //兄弟接口:同以主机的网络接口

echo 2 > /proc/sys/net/ipv4/conf/lo/arp\_announce

echo 1 > /proc/sys/net/ipv4/conf/all/arp\_ignore

echo 2 > /proc/sys/net/ipv4/conf/all/arp\_announce

1. 在本机的lo接口绑定vip地址 192.168.4.253/32,主机52和53都要加

ifconfig lo:1 192.168.4.253/32

echo ‘ifconfig lo:1 192.168.4.253/32’ > /etc/rc.local

1. 运行web服务

主机52：yum -y install httpd

echo xiao > /var/www/html/index.html

主机53：yum -y install httpd

echo 123 > /var/www/html/index.html

1. 验证：

在客户机上

curl http://192.168.4.253

**扩展LVS：**

LVS服务做LB集群时，不能够对realserver服务做健康行检查

解决办法，在分发器上定期执行监测脚本

vim /root/checklvs.sh

#!/bin/bash

如何知道realserver 80 是不是开放状态可以用nmap去检查

nmap -sS -n -p 80 192.168.4.53 在这个命令里去过滤open 如果是open就是打 开不是就是关闭的把他的值给一个变量

nmap -sS -n -p 80 192.168.4.53  ｜ grep open

当realserver 80 不是开放状态的时候就在本机执行删除realserver的命令

ipvsadm -d -t 192.168.4.253:80 -r 192.168.4.x

当realserver 80 是开放状态的时候在把删除realserver的命令添加到虚拟服务里

ipvsadm -a -t 192.168.4.253:80 -r 192.168.4.x:80

**扩展**

LVS扩展：LVS服务做LB集群时，不能对realserver服务做健康检查

当集群内有一台realserver挂掉的时候，调度器不能判断这台主机是否还能提供服务

就存在分发给down的主机的时候，会存在卡顿，不能访问

Vim /root/checklvs.sh

#!/bin/bash

定期扫描目标主机的80端口存活状态 用if语句

死亡：ipvsadm -d -t 192.168.2.54:80 -r 192.168.4.53

存活：ipvsadm -a -t 192.168.2.54:80 -r 192.168.4.53 -m

rontab -e //书写计划任务

5 \* \* \* \* /root/checklvs.sh //执行脚本