# 1 网络知识回顾

先来回顾一下网络知识。

1 什么是ARP？

ARP🡪地址解析协议，每个主机都设有一个ARP地址高速缓存，里面记录了本局域网 所有主机的IP和MAC地址的映射表。

当A🡪B发送报文时，分两种情况，

A的映射表中有B的MAC地址，A就会通过MAC地址向B发送报文。

A的映射表中没有，就会产生以下步骤

1 A广播ARP请求分组

2 同网段的所有主机都能通过ARP进程接收

3 B向A以单播方式发送ARP响应分组，其他主机对A不理睬 **arp\_ignore**

4 A缓存B的IP🡪MAC关系

当新接入一个新的主机时，该主机会向本网段所有主机通告自己的IP和MAC关 --系，**arp\_announce**

**同一网段内的主机转发报文，仅仅修改mac首部。不会修改报文的ip首部。**

2 Linux的网络功能属于内核，但ip地址和网卡还是有联系的。

2.2

1.3

switch

switch

2.1

1.1

1.2

2.3

1.3是可以ping通2.1的，即使中部的主机**没有开启ip-forward**功能，因为2.1使用1.1的mac地址回应。但1.3无法直接和2.1通信，需要将网关设置为1.1后才可以。

# 2 DR模型详解

**1 RIP和VIP同一网段**

三

Switch

（2.4—2.6）/2.3

二

一

2.1

2.3/2.7

2.2

**注释：realserver：RIP/VIP**

**VIP以别名的方式设定在lo上，**

**RIP设定在eth0上**

**director：VIP/DIP**

**1** **为了完成步骤一，也就是所有目的ip为2.3的报文都去往director，而不是去往cluster，需要进行特别的设置方法，一共有三种**

**1 在路由设备上绑定VIP和Director的MAC。**

但路由设备一般为ISP公司提供，无法直接使用

**2 使用arptables命令对realserver进行设置**

**3 较新的内核引入了两个参数kernel parameter:**

**arp\_ignore：**接收到arp请求时的响应级别

0：只要本地配置的有相应地址，就给予响应；默认

1：仅在请求的目标地址配置请求到达的接口上的时候，才给予响应；

**arp\_announce：**定义主动将自己地址向外通告时的通告级别；

0：将本地任何接口上的任何地址向外通告；默认

1：试图仅向目标网络通告与其网络匹配的地址；

2：仅向与本地接口上地址匹配的网络进行通告；

如何配制？

方法一：

sysctl -w net.ipv4.conf.all.arp\_announce=2

sysctl -w net.ipv4.conf.all.arp\_ignore=1

sysctl -w net.ipv4.conf.eth0.arp\_announce=2

sysctl -w net.ipv4.conf.eth0.arp\_ignore=1

方法二：这两种都需要配置all目录的，lo和eth0选择一个就可以

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

**通过这样的设置，虽然realserver有VIP地址，但他们也不会对arp做出回应。**

**2** 第二步中，director类似于路由器，将mac地址修改后，转送至realserver。

**3** 第三步，因为VIP以别名的方式设定在lo上，RIP设定在eth0上。

而linux的默认路由规则则是：响应报文的源地址一般为流出接口的地址，若接口上的地址与网关不在同一网段中，就用接口上别名地址（别名地址此时必须和网关处于同一网段）。为了让响应数据报文的源地址为VIP，需要设定特殊的route规则。

如何设定呢？

director： route add -host VIP dev eth0:0

realserver：route add -host VIP dev lo:0

# 3 脚本

Director脚本:

#!/bin/bash

#

# LVS script for VS/DR

#

. /etc/rc.d/init.d/functions

**#所有设备处于同一网络中**

VIP=192.168.0.210

RIP1=192.168.0.221

RIP2=192.168.0.222

PORT=80

#

case "$1" in

start)

**#使vip只能广播给自己，同一网络其他设备无法收到其广播信息**

/sbin/ifconfig eth0:1 $VIP broadcast $VIP netmask 255.255.255.255 up

**#发送至vip地址的报文必须由eth0:1发送. 经过测试可以省略。**

/sbin/route add -host $VIP dev eth0:1

# Since this is the Director we must be able to forward packets

**#开启路由转发功能。**

echo 1 > /proc/sys/net/ipv4/ip\_forward

# Clear all iptables rules.

/sbin/iptables -F

# Reset iptables counters.

/sbin/iptables -Z

# Clear all ipvsadm rules/services.

/sbin/ipvsadm -C

# Add an IP virtual service for VIP 192.168.0.219 port 80

# In this recipe, we will use the round-robin scheduling method.

# In production, however, you should use a weighted, dynamic scheduling method.

/sbin/ipvsadm -A -t $VIP:80 -s wlc

# Now direct packets for this VIP to

# the real server IP (RIP) inside the cluster

/sbin/ipvsadm -a -t $VIP:80 -r $RIP1 -g -w 1

/sbin/ipvsadm -a -t $VIP:80 -r $RIP2 -g -w 2

/bin/touch /var/lock/subsys/ipvsadm &> /dev/null

;;

stop)

# Stop forwarding packets

echo 0 > /proc/sys/net/ipv4/ip\_forward

# Reset ipvsadm

/sbin/ipvsadm -C

# Bring down the VIP interface

/sbin/ifconfig eth0:1 down

/sbin/route del $VIP

/bin/rm -f /var/lock/subsys/ipvsadm

echo "ipvs is stopped..."

;;

status)

if [ ! -e /var/lock/subsys/ipvsadm ]; then

echo "ipvsadm is stopped ..."

else

echo "ipvs is running ..."

ipvsadm -L -n

fi

;;

\*)

echo "Usage: $0 {start|stop|status}"

;;

esac

RealServer脚本:

#!/bin/bash

#

# Script to start LVS DR real server.

# description: LVS DR real server

#

. /etc/rc.d/init.d/functions

VIP=192.168.0.219

host=`/bin/hostname`

case "$1" in

start)

# Start LVS-DR real server on this machine.

/sbin/ifconfig lo down

/sbin/ifconfig lo up

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

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

**#将vip别名，且设置为vip只能广播到自己。个人认为该语句与上语句起双重保险作用**

/sbin/ifconfig lo:0 $VIP broadcast $VIP netmask 255.255.255.255 up

**#目的地址为vip的仅仅由lo:0接收,由谁接收，就由谁发送。也就是发送时由lo:0发送**

**该句可以省略，可能是因为Linux默认策略就是所收到的包的目的地址将作为响应报文的源地址。**

/sbin/route add -host $VIP dev lo:0

;;

stop)

# Stop LVS-DR real server loopback device(s).

/sbin/ifconfig lo:0 down

echo 0 > /proc/sys/net/ipv4/conf/lo/arp\_ignore

echo 0 > /proc/sys/net/ipv4/conf/lo/arp\_announce

echo 0 > /proc/sys/net/ipv4/conf/all/arp\_ignore

echo 0 > /proc/sys/net/ipv4/conf/all/arp\_announce

;;

status)

# Status of LVS-DR real server.

islothere=`/sbin/ifconfig lo:0 | grep $VIP`

isrothere=`netstat -rn | grep "lo:0" | grep $VIP`

if [ ! "$islothere" -o ! "isrothere" ];then

# Either the route or the lo:0 device

# not found.

echo "LVS-DR real server Stopped."

else

echo "LVS-DR real server Running."

fi

;;

\*)

# Invalid entry.

echo "$0: Usage: $0 {start|status|stop}"

exit 1

;;

esac