**Freeswitch主备实现**

1. 实现主备需要解决的几个问题？
2. 数据共享，主备服务器需要共享相同的数据，这样切换之后才能继续工作。
3. 心跳探测，探测服务器是否正常运行。
4. 切换策略，合适的时机进行主备服务切换。
5. 功能接续，主备切换之后，不影响正在使用服务的的用户。（热备）
6. 终端设备无感知，主备服务器需要对外暴露同一个IP地址和端口，并且功能相同
7. freeswitch数据存储与共享

众所周知，freeswitch默认使用的数据库是sqlite。Sqlite的设计目标是嵌入式的，占有资源非常的低，可能只需要几百K的内存就够了。Sqlite引擎不是一个程序与之通信的独立进程，而是连接到程序中成为它的一个主要部分，所以主要的通信协议是在编程语言内的直接API调用，相比与mysql、postgresql，它的处理速度比他们都快。

基于上述原因，主备服务数据共享无法使用sqlite。结合我们当前现有的服务，选用mysql可能是最佳选择了。

2.1 freeswitch使用mysql存储数据

2.1.1 需要依赖包

unixODBC-devel

Mysql-connector-odbc

2.1.2 配置要连接的数据库信息

编辑/etc/odbc.ini文件

[freeswitch]

Driver = MySQL

SERVER = 192.168.43.145

PORT = 3306

DATABASE = test

OPTION = 67108864

Sokcet = /var/lib/mysql/mysql.sock

2.1.3 配置正确的驱动包

编辑/etc/odbcinst.ini（一般情况下不用修改，除非系统找不到驱动，需要指定驱动路径）

[MySQL]

Description = ODBC for MySQL

Driver = /usr/lib/libmyodbc5.so

Setup = /usr/lib/libodbcmy5.so

Driver64 = /usr/lib64/libmyodbc5.so

Setup64 = /usr/lib64/libodbcmys.so

FileUsage = 1

2.1.4 连接mysql

需要在各个模块中引入sqlite访问。

* 编辑db.conf.xml

Settings节点内添加<param name=”odbc-dsn” value=”freeswitch:root:123456”/>

* 编辑switch.conf.xml

Settings节点内添加

<param name=”core-db-dsn” value=”freeswitch:root:123456”/>

<param name=”odbc-dsn” value=”freeswitch:root:123456”/>

* 编辑internal.conf.xml

Settings节点内添加<param name=”odbc-dsn” value=”freeswitch:root:123456”/>

* 编辑external.conf.xml

Settings节点内添加<param name=”odbc-dsn” value=”freeswitch:root:123456”/>

* 编辑voicemail.conf.xml

profile节点内添加<param name=”odbc-dsn” value=”freeswitch:root:123456”/>

如果其他模块也想通过这个方式访问mysql数据库，需要手动编写sql执行命令。

而freeswitch:root:123456中的freeswitch是/etc/odbc.ini中访问mysql的节点标签，

root是访问mysql的用户名，123456是访问密码。

1. 主备切换

下面介绍使用Keepalived实现主备切换功能，包含详细的主备心跳探测过程和切换策略描述。

3.0 keepalived

Keepalived主要模块：

Core：是keepalived的核心，负责主进程的启动、维护以及全局配置文件的加载和解析等。

Check：负责healthchecker（健康检查），包括了各种健康检查方式。

VRRP：VRRPD子进程，VRRPD子进程就是来实现VRRP协议的。

Keepalived只有一个配置文件，默认安装在/etc/keepalived/keepalived.conf，这里解释下常用的配置参数。

global\_defs {

vrrp\_script check\_fs {

script "/etc/keepalived/script/check\_fs.sh" //健康检查脚本

interval 1 //检查周期

}

vrrp\_instance VI\_1 { //VRRP路由器实例

state BACKUP //状态

interface enp0s3 //网卡标识

virtual\_router\_id 51 //VRRP\_ID

priority 100 //优先级

nopreempt //抢占模式设置

advert\_int 1 //通报报文间隔

authentication {

auth\_type PASS

auth\_pass 1111

}

track\_script {

check\_fs

}

virtual\_ipaddress {

192.168.43.16/24 //vrrp\_ip

}

notify\_master "/etc/keepalived/script/fs\_recover.sh" //切换成master要执行的

}

要知道keepalived如何实现主备切换的。必须了解VRRP，只要了解了VRRP协议才能更好的应用并分析问题。下面重点介绍下VRRP。

3.1 VRRP（虚拟路由冗余协议）简介

虚拟路由冗余协议通过把几台路由设备联合组成一台虚拟路由设备（虚拟路由组），将虚拟路由设备的ip地址作为用户的默认网关地址实现与外部的通讯。当网关设备发生故障时，VRRP机制能够选举新的网关设备承担数据流量，从而保证网络的可靠通信。

通常，同一网段内的所有主机上都设置一条相同的、以网关为下一跳的缺省路由。主机发往其他网段的报文将通过缺省路由发往网关，再由网关进行转发，从而实现主机与外部网络的通信。

当网关发生故障时，本网段内所有以网关为缺省路由的主机将无法与外部网络通信。增加出口网关是提高系统可靠性的常见方法，此时如何在多个出口之间进行选路就成为需要解决的问题。

VRRP的出现很好的解决了这个问题。VRRP能够在不改变组网的情况下，采用将多台路由设备组成一个虚拟路由器，通过配置虚拟路由器的IP地址为默认网关，实现默认网关的备份。当网关设备发生故障时，VRRP机制能够选举新的网关设备承担数据流量，从而保障网络的可靠通信。

3.3 VRRP协议报文

VRRP协议报文用来将Master设备的优先级和状态通告给同一备份组的所有Backup设备。

VRRP协议报文封装在IP报文中，发送到分配给VRRP的IP组播地址。在IP报文头中，源地址为发送报文接口的主IP地址（不是虚拟IP地址），目的地址是224.0.0.18，TTL是255，协议号是112。

主IP地址（Primary IP Address）：从接口的真实IP地址中选出来的一个主用IP地址，通常选择配置的第一个IP地址。

目前，VRRP协议包括两个版本：

VRRPv2和VRRPv3。VRRPv2仅适用于IPv4网络，VRRPv3适用于IPv4和IPv6两种网络。

基于不同的网络类型，VRRP可以分为VRRP for IPv4和VRRP for IPv6（简称VRRP6）。VRRP for IPv4支持VRRPv2和VRRPv3，而VRRP for IPv6仅支持VRRPv3。

3.2 VRRP报文结构

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 1 | 2 | 3 | 4 | | 5 | 6 | 7 | 8 | | 9 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | | | 7 | 8 | 9 | 0 | 1 | 2 | 3 | 4 | | 5 | 6 | 7 | 8 | 9 | 0 | 1 |
| Version | | | | | Type | | | | | Virtual Rtr ID | | | | | | | | | Priority | | | | | | | | | | Count IP Addrs | | | | | | | |
| Auth Type | | | | | | | | | | Adver Int | | | | | | | | | | Checksum | | | | | | | | | | | | | | | | |
| IP Address(1) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| ... | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| IP Address(n) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Authentication Data(1) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Authentication Data(2) | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

图一

Version：协议版本号（4bit）

Type：报文的类型（4bit），该字段的值总是1，标识是VRRP通告报文，若收到非1的类型值，会被丢弃

Virtual Rtr ID:一个VRID唯一地标识了一个虚拟路由器,取值范围是[1,255]。在同一个VRRP组内的路由器必须有相同的VRID。

Priority：Master设备在备份组中的优先级，取值范围是0~255。0表示设备停止参与VRRP备份组，用来使备份设备尽快成为Master设备，而不必等到计时器超时；255则保留给IP地址拥有者(Master),缺省值是100。

Count ip addrs：VRRP通告报文中包括的IP地址数量，事实上就是为VRRP虚拟路由器所分配的IP地址的数量。

Auth Type：VRRP报文的认证类型。

1. 表示无认证
2. 表示明文认证
3. 标识MD5认证

值得注意的是：VRRPv3不支持认证功能（RFC3768中已经去掉了全部的认证类型）；VRRPv2支持认证功能（RFC2338）。是的

Auth Int：通报报文的发送时间间隔，单位是秒，默认是1秒。

Checksum：VRRP报文的校验和，用于检测VRRP报文中的数据破坏情况。

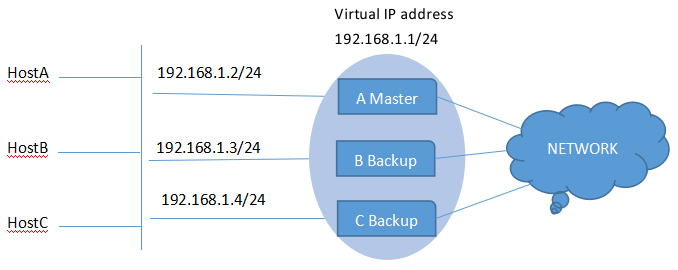
IP Address：VRRP备份组的虚拟ipv4地址。

Authentication Data：VRRP报文的认证部分，目前只有明文认证和MD5认证才用到该部

分，对于其他认证方式，一律填0。

**3.3 VRRP协议的几个概念**

|  |  |
| --- | --- |
| VRRP路由器 | 运行VRRP协议的设备 |
| 虚拟路由器 | 一个master设备和多个backup设备组成，被当做一个共享局域网内主机的缺省网关，也叫VRRP备份组 |
| Master路由器 | 承担转发报文任务的的VRRP设备 |
| Backup路由器 | 一组没有承担转发任务的VRrP设备，当Master设备出现故障时，他们将通过竞选成为新的Master设备。 |
| VRID | 虚拟路由器的标识。 |
| 虚拟IP地址 | 虚拟路由器的IP地址，一个虚拟路由器可以有一个或多个IP地址 |
| Ip地址拥有者 | Master VRRP设备，将虚拟路由器IP地址作为真是的接口地址的设备。 |
| 虚拟MAC地址 | 虚拟路由器根据虚拟路由器ID生成的MAC地址。一个虚拟路由器拥有一个虚拟MAC地址，格式为：00-00-5E-00-01-{VRID}{VRRP FOR IPV4} |

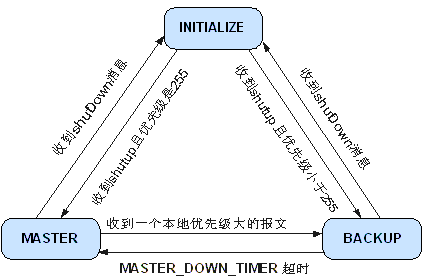


图二 虚拟路由器

如图二所示，HostA、HostB、HostC是三个VRRP路由器，组成了虚拟路由器（椭圆区），虚拟路由的ip地址是192.168.1.1。当前HOSTA是承担转发报文任务的VRRP设备,即它就是当前的IP地址拥有者。

**3.4 VRRP路由器的状态**

VRRP协议中定义了三种状态机：初始状态（Initialize）、活动状态（Master）、备份状态（Backup）。其中，只有处于活动状态的设备(Master)才可以转发VRRP报文。



图三 VRRP状态机转换

|  |  |
| --- | --- |
| **状态** | **说明** |
| Initialize | * 设备启动时进入此状态（设备检测到故障时也会进入Initialize状态） * 当收到接口Startup的消息，将转入Backup或Master状态（IP地址拥有者的接口优先级为255，直接转为Master；设备的优先级小于255，则先切换至Backup状态）。 * 在此状态时，不会对VRRP报文做任何处理。 |
| Master | * 定时发送VRRP通告报文。 * 以虚拟MAC地址响应对虚拟IP地址的ARP请求。 * 转发目的MAC地址为虚拟MAC地址的IP报文。 * 如果它是整个虚拟IP地址的拥有者，则接收目的IP地址为这个虚拟IP地址的IP报文。否则，丢弃这个IP报文。 * 如果收到到比自己优先级大的报文，立即转为Backup状态。 * 如果收到与自己优先级相等的VRRP报文，并且本地接口IP地址小于对端接口IP，立即转换为Backup状态。 * 当接收到接口的shutdown事件时，转为Initialize状态。 |
| Backup | * 接收Master设备发送的VRRP通告报文，判断Master设备的状态是否正常。 * 对虚拟IP地址的ARP请求，不做响应。 * 丢弃目的IP地址为虚拟IP地址的IP报文。 * 如果收到比自己优先级小的报文，报文优先级不是0，丢弃报文，立即成为Master。 * 如果收到比自己优先级小的报文，报文优先级等于0（priority为0表示停止参与VRRP备份组，也就是退出该虚拟路由器），定时器设置为skew\_time。 * 如果收到优先级和自己相同或者比自己的大的报文，则重置Master\_Down\_Interval定时器，不进一步比较IP地址。   Master\_Down\_Interval=(3\*Advertisement\_Invterval) + skew\_time。  Skew\_Time=(256-Priority)/256。  Master\_Down\_Interval是Backup用来判断master是否还正常工作的超时时间。Advertisement\_Invterval是Master发送报文的间隔时间，默认是1s。Skew\_Time是偏移时间，可以看出Skew\_Time小于1s。 |

**3.5 VRRP的工作工程**

* VRRP备份组中的设备根据优先级选举出Master。Master设备立即发送ARP报文，将虚拟MAC地址通知给与它连接的数倍或者主机，从而承担报文转发任务。
* Master设备周期性向备份组内所有Backup设备发送VRRP通告报文，以公布Master设备的配置信息（优先级等）和工作状况。
* Master设备出现故障时，VRRP备份组中的Backup设备将根据优先级重新选举新的Master.
* VRRP备份组状态切换时，Master设备由一台设备切换为另外一台设备，新的Master设备会立即发送携带虚拟路由器的虚拟MAC地址和虚拟IP地址信息的免费ARP报文，刷新与它连接的主机或设备中的MAC表项，从而把用户流量引到新的Master设备上来，整个过程对用户完全透明。
* 原来的Master设备故障恢复时，若该设备为IP地址拥护者（优先级为255），将直接切换至Master状态。若该设备优先级小于255，将首先切换至Backup状态，且其优先级恢复为故障前配置的优先级。

Backup设备的优先级高于Master设备，由Backup设备的工作方式（抢占方式和非抢占方式）决定是否重新选举Master。

抢占模式：

在抢占模式下，如果Backup设备的优先级比当前Master设备的优先级高，则主动将自己切换成Master。

非抢占模式：

在非抢占模式下，只要Master设备没有出现故障，Backup设备即使随后被配置了更高的优先级也不会成为Master设备。

由此可见，为了保证Master设备和Backup设备能够协调工作，VRRP需要实现以下功能：

Master设备的选举

Master设备状态的通告

**3.6 Master设备选举**

VRRP根据优先级来确定虚拟路由器中每台设备的角色（Master设备或Backup设备）。优先级越高，则越有可能成为Master设备。

初始创建的VRRP设备工作在Initialize状态，收到接口Up的消息后，如果设备的优先级为255，则直接成为Master设备；如果设备的优先级小于255，则会先切换至Backup状态，待Master\_Down\_Interval定时器超时后再切换至Master状态。首先切换至Master状态的VRRP设备通过VRRP通告报文的交互获知虚拟设备中其他成员的优先级，进行Master的选举：

如果VRRP报文中Master设备的优先级高于或等于自己的优先级，则Backup设备保持Backup状态；

如果VRRP报文中Master设备的优先级低于自己的优先级，采用抢占方式的Backup设备将切换至Master状态，采用非抢占方式的Backup设备仍保持Backup状态。

如果多个VRRP设备同时切换到Master状态，通过VRRP通告报文的交互进行协商后，优先级较低的VRRP设备将切换成Backup状态，优先级最高的VRRP设备成为最终的Master设备；优先级相同时，VRRP设备上VRRP备份组所在接口主IP地址较大的成为Master设备。

如果创建的VRRP设备为IP地址拥有者，收到接口Up的消息后，将会直接切换至Master状态。

**3.6 Master设备状态通告**

Master设备周期性地发送VRRP通告报文，在VRRP备份组中公布其配置信息（优先级等）和工作状况。Backup设备通过接收到VRRP报文的情况来判断Master设备是否工作正常。

当Master设备主动放弃Master地位（如Master设备退出备份组）时，会发送优先级为0的通告报文，用来使Backup设备快速切换成Master设备，而不用等到Master\_Down\_Interval定时器超时。这个切换的时间称为Skew time，计算方式为：（256－Backup设备的优先级）/256，单位为秒。

当Master设备发生网络故障而不能发送通告报文的时候，Backup设备并不能立即知道其工作状况。等到Master\_Down\_Interval定时器超时后，才会认为Master设备无法正常工作，从而将状态切换为Master。其中，Master\_Down\_Interval定时器取值为：3×Advertisement\_Interval＋Skew\_time，单位为秒。

在性能不稳定的网络中，网络堵塞可能导致Backup设备在Master\_Down\_Interval期间没有收到Master设备的报文，Backup设备则会主动切换为Master。如果此时原Master设备的报文又到达了，新Master设备将再次切换回Backup。如此则会出现VRRP备份组成员状态频繁切换的现象。

为了缓解这种现象，可以配置抢占延时，使得Backup设备在等待了Master\_Down\_Interval后，再等待抢占延迟时间。如在此期间仍没有收到通告报文，Backup设备才会切换为Master设备。

* 将局域网的一组路由器构成一个备份组，相当于一台虚拟路由器。局域网内的主机只需要知道这个虚拟路由器的IP地址，并不需知道具体某台设备的IP地址，将网络内主机的缺省网关设置为该虚拟路由器的IP地址，主机就可以利用该虚拟网关与外部网络进行通信。

VRRP将该虚拟路由器动态关联到承担传输业务的物理路由器上，当该物理路由器出现故障时，再次选择新路由器来接替业务传输工作，整个过程对用户完全透明，实现内部网络和外部网络不间断通信。

虚拟路由组内设备之间通过VRRP协议之间通信，选举出一个MASTER设备，其他为BACKUP设备。

心跳是通过网络多播（组播）的通信方式实现的，MASTER路由设备会定期向224.0.0.18组播地址发送VRRP消息报文。BACKUP路由设备收到VRRP消息报文之后，会根据本地配置的策略反复执行选举动作。