lvs+keepalived+mha+mysql 高可用架构配置说明

第一部分 MHA 介绍

MHA（Master High Availability）目前在 MySQL 高可用方面是一个相对成熟的解决方案，它由日本人 youshimaton 开发，是一套优秀的作为 MySQL 高可用性环境下故障切换和主从提升的高可用软件。在 MySQL 故障切换过程中，MHA 能做到 0~30 秒之内自动完成数据库的故障切换操作，并且在进行故障切换的过程中，MHA 能最大程度上保证数据库的一致性，以达到真正意义上的高可用。

MHA 由两部分组成：MHA Manager（管理节点）和 MHA Node（数据节点）。MHA Manager 可以独立部署在一台独立的机器上管理多个 Master-Slave 集群，也可以部署在一台 Slave 上。当 Master 出现故障是，它可以自动将最新数据的 Slave 提升为新的 Master，然后将所有其他的 Slave 重新指向新的 Master。整个故障转移过程对应用程序是完全透明的。

1.1 存在隐患

在 MHA 自动故障切换的过程中，MHA 试图从宕掉的主服务器上保存二进制日志，最大程度保证数据的不丢失，但这并不总是可行的。

例如，如果主服务器硬件故障或无法通过 SSH 访问，MHA 没有办法保存二进制日志，只能进行故障转移而丢失了最新数据。

拓：MySQL 服务挂了，但是可以从服务器拷贝二进制。但如果硬件宕机或者 SSH 不能连接，不能获取到最新的 binlog 日志，如果复制出现延迟，会丢失数据。
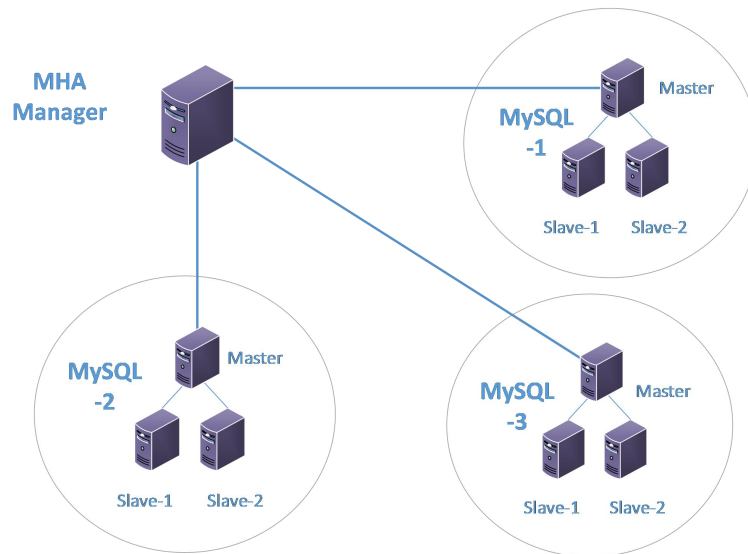
使用 MySQL5.5 的半同步复制，可以大大降低数据丢失的风险。MHA 可以和半同步复制结合起来。如果只有一个 Slave 已经收到了最新的二进制日志，MHA 可以将最新的二进制日志应用于其他所有 Slave 服务器上，保持数据一致性。

最新版 0.56 版本，增加了支持 GTID 的功能，建议在 MySQL5.6 及之后版本使用。MySQL5.5 建议使用管理节点版本 0.55，数据节点 0.54。

1.2 适用场景

目前 MHA 主要支持一主多从的架构，要搭建 MHA，要求一个复制集群必须最少有 3 台数据库服务器，一主二从，即一台充当 Master，一台充当备用 Master，另一台充当从库。出于成本考虑，淘宝在此基础上进行了改造，目前淘宝开发的 TMHA 已经支持一主一从。

1.3 MHA 工作原理

1. 从宕机崩溃的 Master 保存二进制日志事件（binlog event）；

2. 识别含有最新更新的 Slave；

3. 应用差异的中继日志（relay log）到其他 Slave；

4. 应用从 Master 保存的二进制日志事件；

5. 提升一个 Slave 为新的 Master；

6. 使其他的 Slave 连接新的 Master 进行复制；

1.4 MHA 的组成

MHA 软件由两部分组成，Manager 工具包和 Node 工具包，具体如下。

Manager 工具包情况如下：

- masterha_check_ssh:检查 MHA 的 SSH 配置情况。

- masterha_check_repl:检查 MySQL 复制状况。

- masterha_manager:启动 MHA。

- masterha_check_status:检测当前 MHA 运行状态。

- masterha_master_monitor:检测 Master 是否宕机。

- masterha_master_switch:控制故障转移（自动或手动）。

- masterha_conf_host:添加或删除配置的 server 信息。

Node 工具包（通常由 MHA Manager 的脚本触发，无需人工操作）情况如下：

- save_binary_logs:保存和复制 Master 的 binlog 日志。
- apply_diff_relay_logs:识别差异的中级日志时间并将其应用到其他 Slave。
- filter_mysqlbinlog:去除不必要的 ROOLBACK 事件（已经废弃）
- purge_relay_logs:清除中继日志（不阻塞 SQL 线程）

注：为了尽可能的减少因为主库硬件损坏宕机造成的数据丢失，因此在配置 MHA 的同时建议配置 MySQL5.5 半同步复制。如果对性能要求较高，允许丢失一小部分数据，可以不做半同步复制。

拓展思想：为了保证数据一致性，MySQL 复制中，常常会在 Master 上使用 sync_binlog 参数保证 binlog 持久化，保证数据一致性。但这种方式对磁盘 I/O 会造成 10~20%的影响。但是还有另外一个思路，就是使用 MySQL 半同步复制来保证数据一致性，MySQL 半同步复制是在从服务器的内存中处理数据并进行发馈，虽然也会造成性能影响，但是相对于对 Master 造成的磁盘 I/O 的影响来说，反而是个更好的方法。据《高性能 MySQL》 第三版中 10.9 的测试，写入远程的内存（一台从库的反馈）比写入本地的磁盘（写入并刷新）要更快。使用半同步复制相比主在主库上进行强持久化的性能有两倍的改善。

（以上内容感谢同学 BrandynX 编辑）

第二部分 环境部署信息

2.1 软件部署表

lvs 版本:ipvsadm-1.26

keepalived 版本:keepalived-1.1.19

mysql 版本:mysql-5.5.32

mha-manger 版本:mha4mysql-manager-0.55-0.el6.noarch

mha-node 版本:mha4mysql-node-0.54-0.el6.noarch

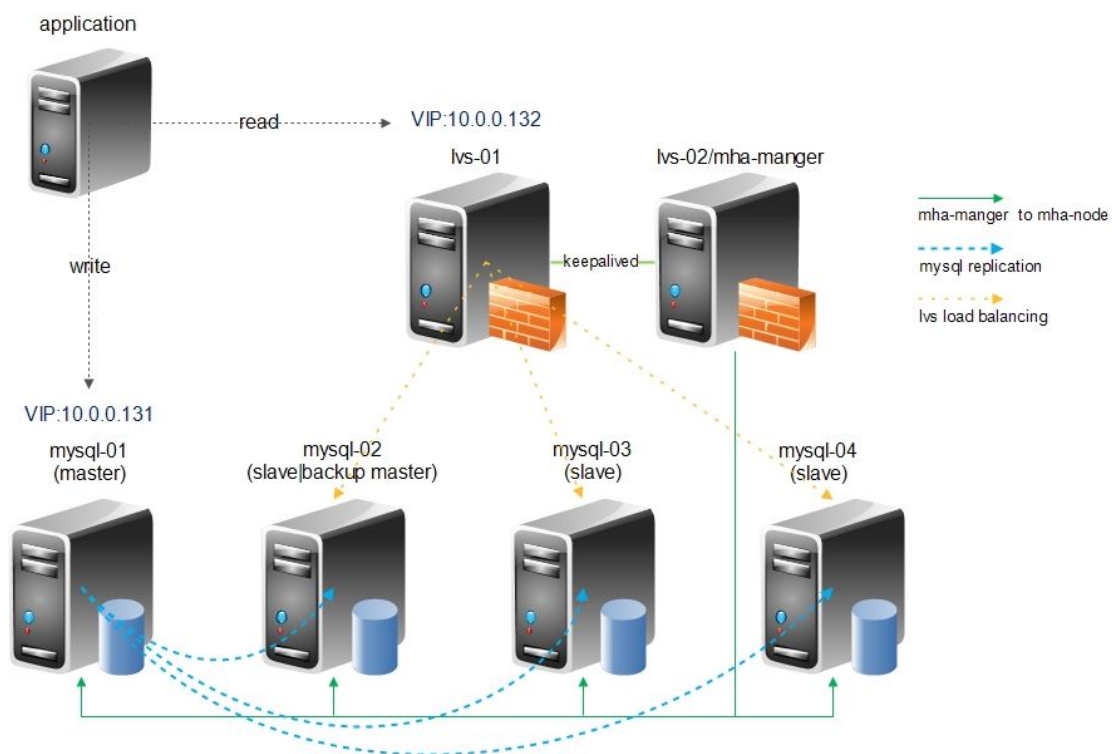| IP 地址及主机名	系统版本	部署服务
10.0.0.121 lvs-01	CentOS release 6.4 (Final) 2.6.32-358.el6.x86_64	lvs,keepalived
10.0.0.122 lvs-02	CentOS release 6.4 (Final) 2.6.32-358.el6.x86_64	lvs,keepalived,mha-manger

10.0.0.123 mysql-01	CentOS release 6.4 (Final) 2.6.32-358.el6.x86_64	mysql,mha-node
10.0.0.124 mysql-02	CentOS release 6.4 (Final) 2.6.32-358.el6.x86_64	mysql,mha-node
10.0.0.125 mysql-03	CentOS release 6.4 (Final) 2.6.32-358.el6.x86_64	mysql,mha-node
10.0.0.126 mysql-04	CentOS release 6.4 (Final) 2.6.32-358.el6.x86_64	mysql,mha-node
Mysql write vip：10.0.0.131		
Mysql read vip：10.0.0.132		

2.2 角色分配

主机名	角色	Server id
Lvs-01	Lvs 主节点	-
Lvs-02	Lvs 备节点，mha-manger 管理	-
Mysql-01	Mysql 主库	123
Mysql-02	Mysql 从库,备主	124
Mysql-03	Mysql 从库	125
Mysql-04	Mysql 从库	126

2.3 架构拓扑图

mysql master 故障后转移

application

read ┄┄┄┄┄┄┄┄▶ VIP:10.0.0.132

lvs-01   lvs-02/mha-manger

write

keepalived

mha-manger to mha-node
mysql replication
lvs load balancing

VIP:10.0.0.131

mysql-01   mysql-02   mysql-03   mysql-04
(fail over)   (master)   (slave)   (slave)
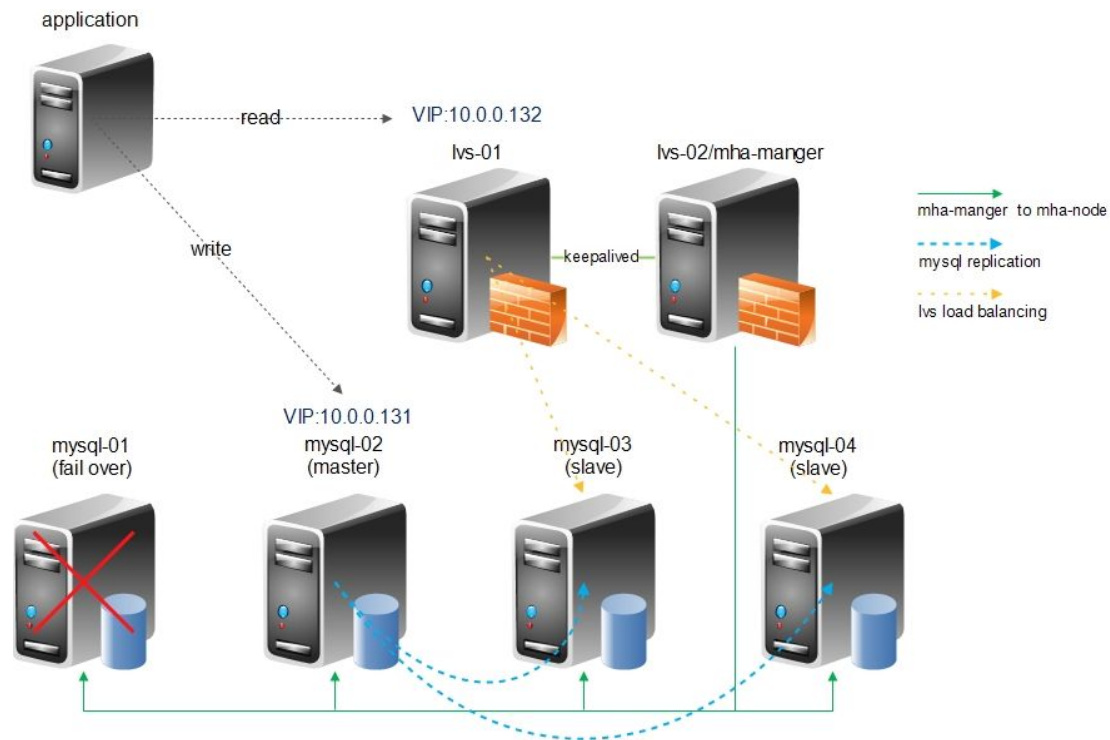
2.4 架构实现原理

1. 读操作

1）LVS 实现读操作的负载均衡；

2）Keepalived 在上层管理 LVS，并对两台从库进行健康检测（通过定义 Check 脚本）；

3）一台从库出现故障后，Keepalived 将其剔除出负载均衡集群；

2. 写操作

1）在 Master 上绑定写 VIP（MHA 启动后会通过脚本进行操作）；

2）MHA 监控 Master 状态，当 Master 出现故障后（宕机、复制暂停）时；

3）通过 Failover 脚本，卸载 Master 上的 WVIP；

4）通过 Failover 脚本在 Backup Master 上绑定 WVIP，提升其为主库；

5）同步并应用差异日志，并将从库指向新主库；

问题：当 MHA 把 Master 切换到了 Backup Master 上后，LVS 如何处理分发在 Backup Master 上的读操作？

解释：由于 Keepalived 会通过脚本定期监控 Backup Master 的状态，包括同步、SQL 线程、I/O 线程，所以当 Backup Master 升级为主库后，这些状态都将消失，Keepalived 将自动将 Backup Master 剔除出负载均衡集群。

第三部分 安装配置 mysql、mha 服务

3.1 安装 mysql 及配置主从

安装 mysql 可以用编译安装、二进制包安装、rpm 安装。我这为了方便，也准备了二进制包，我这就以二进制包方式安装了。

mysql-01 master 操作：

```
#解压二进制包
[root@mysql-01 tmp]# ll
总用量 182352
-rw-r--r-- 1 root root 186722932 6 月   20 2013 mysql-5.5.32-linux2.6-x86_64.tar.gz
[root@mysql-01 tmp]# tar zxf mysql-5.5.32-linux2.6-x86_64.tar.gz
[root@mysql-01 tmp]# ll
总用量 182356
drwxr-xr-x 13 root root          4096 3 月    6 14:09 mysql-5.5.32-linux2.6-x86_64
-rw-r--r--   1 root root 186722932 6 月   20 2013 mysql-5.5.32-linux2.6-x86_64.tar.gz
[root@mysql-01 tmp]# mv mysql-5.5.32-linux2.6-x86_64 /usr/local/mysql
[root@mysql-01 tmp]# mkdir /dbdata
[root@mysql-01 tmp]# cp /usr/local/mysql/support-files/my-small.cnf   /dbdata/my.cnf

#注意:
这里 my.cnf 由于机器为虚拟机，配置也就最小化了，我是将数据目录和程序分开的，所以配置如下
[client]
#password             = your_password
port                  = 3306
socket                = /tmp/mysql.sock
datadir               = /dbdata/data

# Here follows entries for some specific programs

# The MySQL server
[mysqld]
port                  = 3306
socket                = /tmp/mysql.sock
```

==datadir   = /dbdata/data==

...

<mark>嗨~~不要忘了如下配置</mark>

server-id   = 123

log-bin=mysql-bin

#初始化

```
[root@mysql-01 mysql]# useradd mysql -s /sbin/nologin –M
[root@mysql-01 dbdata]# cd /usr/local/mysql/
[root@mysql-01 mysql]# ./scripts/mysql_install_db --defaults-file=/dbdata/my.cnf --user=mysql
Installing MySQL system tables...
OK
Filling help tables...
OK

To start mysqld at boot time you have to copy
support-files/mysql.server to the right place for your system

...
```

#启动 mysql

```
[root@mysql-01 mysql]# ./bin/mysqld_safe --defaults-file=/dbdata/my.cnf --user=mysql&
[1] 4008
[root@mysql-01 mysql]# 150306 14:47:18 mysqld_safe Logging to '/dbdata/data/mysql-01.err'.
150306 14:47:18 mysqld_safe Starting mysqld daemon with databases from /dbdata/data
```

#创建同步账号及删除一些无用账号

```
[root@mysql-01 mysql]# ./bin/mysqladmin    password '123456'
[root@mysql-01 mysql]# history -c
[root@mysql-01 mysql]# ./bin/mysql -uroot -p
Enter password:
Welcome to the MySQL monitor.    Commands end with ; or \g.
Your MySQL connection id is 3
Server version: 5.5.32-log MySQL Community Server (GPL)

Copyright (c) 2000, 2013, Oracle and/or its affiliates. All rights reserved.

Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql> select user,host from mysql.user;
+------+-----------+
| user | host      |
```

```
+------+-----------+
| root | 127.0.0.1 |
| root | ::1       |
|      | localhost |
| root | localhost |
|      | mysql-01  |
| root | mysql-01  |
+------+-----------+
6 rows in set (0.00 sec)
mysql> delete from mysql.user where user = '' or host != 'localhost';
Query OK, 5 rows affected (0.00 sec)
mysql> select user,host from mysql.user;
+------+-----------+
| user | host      |
+------+-----------+
| root | localhost |
+------+-----------+
1 row in set (0.00 sec)
mysql> grant replication slave,replication client on *.* to 'rep'@'10.0.0.%' identified by
'reppasswd';
Query OK, 0 rows affected (0.00 sec)

mysql> select user,host from mysql.user;
+------+-----------+
| user | host      |
+------+-----------+
| rep  | 10.0.0.%  |
| root | localhost |
+------+-----------+
2 rows in set (0.00 sec)

mysql> Bye
[root@mysql-01 mysql]#
```

#关闭 mysql，打包/dbdata 目录，分发到其他 3 台 mysql 主机上（这里有很多方法，可以用逻辑备份等等，看个人爱好）

```
[root@mysql-01 mysql]# ./bin/mysqladmin shutdown -uroot -p
Enter password:
150306 15:19:42 mysqld_safe mysqld from pid file /dbdata/data/mysql-01.pid ended
[1]+    Done                              ./bin/mysqld_safe  --defaults-file=/dbdata/my.cnf
--user=mysql
[root@mysql-01 /]# cd / && tar zcf /opt/mysqldata.tar.gz ./dbdata
[root@mysql-01 /]# ll /opt/
总用量 772
-rw-r--r--   1 root root 785095 3 月     6 15:20 mysqldata.tar.gz
```

版权所有:www.chocolee.cn

drwxr-xr-x. 2 root root   4096 2 月   22 2013 rh

[root@mysql-01 /]# scp /opt/mysqldata.tar.gz   10.0.0.124:/opt

root@10.0.0.124's password:

mysqldata.tar.gz

100%   767KB 766.7KB/s   00:00

[root@mysql-01 /]# scp /opt/mysqldata.tar.gz   10.0.0.125:/opt

root@10.0.0.125's password:

mysqldata.tar.gz

100%   767KB 766.7KB/s   00:00

[root@mysql-01 /]# scp /opt/mysqldata.tar.gz   10.0.0.126:/opt

root@10.0.0.126's password:

mysqldata.tar.gz

100%   767KB 766.7KB/s   00:00

#再次开启 mysql

[root@mysql-01 /]# cd /usr/local/mysql/

[root@mysql-01 mysql]# ./bin/mysqld_safe --defaults-file=/dbdata/my.cnf --user=mysql&

[1] 4360

[root@mysql-01 mysql]# 150306 15:27:16 mysqld_safe Logging to '/dbdata/data/mysql-01.err'.

150306 15:27:16 mysqld_safe Starting mysqld daemon with databases from /dbdata/data

[root@mysql-01 mysql]# ./bin/mysql -uroot -p123456 -e "show master status"

+-----------------+----------+--------------+------------------+

| File | Position | Binlog_Do_DB | Binlog_Ignore_DB |

+-----------------+----------+--------------+------------------+

| mysql-bin.000004 | 107 | | |

+-----------------+----------+--------------+------------------+

mysql-02 slave 操作：

#解压二进制包，解压数据目录，创建 mysql 用户

[root@mysql-02 tmp]# ll

总用量 182352

-rw-r--r-- 1 root root 186722932 3 月   4 14:53 mysql-5.5.32-linux2.6-x86_64.tar.gz

[root@mysql-02 tmp]# tar zxf mysql-5.5.32-linux2.6-x86_64.tar.gz

[root@mysql-02 tmp]#

[root@mysql-02 tmp]# ll

总用量 182356

drwxr-xr-x 13 root root   4096 3 月   6 15:36 mysql-5.5.32-linux2.6-x86_64

-rw-r--r--   1 root root 186722932 3 月   4 14:53 mysql-5.5.32-linux2.6-x86_64.tar.gz

[root@mysql-02 tmp]# mv mysql-5.5.32-linux2.6-x86_64 /usr/local/mysql

[root@mysql-02 tmp]# tar zxf /opt/mysqldata.tar.gz -C /

[root@mysql-02 tmp]# ll -d /dbdata/

drwxr-xr-x 3 root root 4096 3 月   6 14:47 /dbdata/

[root@mysql-02 tmp]# ll /dbdata/

总用量 8

```
drwx------ 5    501 root 4096 3 月      6 15:19 data
-rw-r--r-- 1 root root 2901 3 月      6 14:46 my.cnf
[root@mysql-02 tmp]# useradd mysql -s /sbin/nologin    -M
[root@mysql-02 tmp]# ll /dbdata/
总用量 8
drwx------ 5 mysql root 4096 3 月      6 15:19 data
-rw-r--r-- 1 root    root 2901 3 月      6 14:46 my.cnf
```
#修改 my.cnf 中 server id，由于是备主，需要将 log-bin 打开
```
[root@mysql-02 tmp]# sed -i '/server-id/ s#123#124#g' /dbdata/my.cnf
[root@mysql-02 tmp]# grep server-id /dbdata/my.cnf
server-id  = 124
[root@mysql-02 tmp]# grep log-bin /dbdata/my.cnf
log-bin=mysql-bin
```
#启动 mysql,配置主从
```
[root@mysql-02 tmp]# cd /usr/local/mysql/
[root@mysql-02 mysql]# ./bin/mysqld_safe --defaults-file=/dbdata/my.cnf --user=mysql&
[1] 3536
[root@mysql-02 mysql]# 150306 16:11:55 mysqld_safe Logging to '/dbdata/data/mysql-02.err'.
150306 16:11:55 mysqld_safe Starting mysqld daemon with databases from /dbdata/data

[root@mysql-02 mysql]# ./bin/mysql -uroot -p
Enter password:
Welcome to the MySQL monitor.    Commands end with ; or \g.
Your MySQL connection id is 1
Server version: 5.5.32-log MySQL Community Server (GPL)

Copyright (c) 2000, 2013, Oracle and/or its affiliates. All rights reserved.

Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql> change master to
master_host='10.0.0.123',master_port=3306,master_user='rep',master_password='reppasswd',mast
er_log_file='mysql-bin.000004',master_log_pos=107;
Query OK, 0 rows affected (0.08 sec)

mysql> start slave;
Query OK, 0 rows affected (0.09 sec)

mysql> show slave status\G
*************************** 1. row ***************************
```

```
                   Slave_IO_State: Waiting for master to send event
                      Master_Host: 10.0.0.123
                      Master_User: rep
                      Master_Port: 3306
                    Connect_Retry: 60
                  Master_Log_File: mysql-bin.000004
              Read_Master_Log_Pos: 107
                   Relay_Log_File: mysql-02-relay-bin.000002
                    Relay_Log_Pos: 253
            Relay_Master_Log_File: mysql-bin.000004
                 Slave_IO_Running: Yes
                Slave_SQL_Running: Yes
                  Replicate_Do_DB:
              Replicate_Ignore_DB:
               Replicate_Do_Table:
           Replicate_Ignore_Table:
          Replicate_Wild_Do_Table:
      Replicate_Wild_Ignore_Table:
                       Last_Errno: 0
                       Last_Error:
                     Skip_Counter: 0
              Exec_Master_Log_Pos: 107
                  Relay_Log_Space: 412
                  Until_Condition: None
                   Until_Log_File:
                    Until_Log_Pos: 0
               Master_SSL_Allowed: No
               Master_SSL_CA_File:
               Master_SSL_CA_Path:
                  Master_SSL_Cert:
                Master_SSL_Cipher:
                   Master_SSL_Key:
            Seconds_Behind_Master: 0
Master_SSL_Verify_Server_Cert: No
                    Last_IO_Errno: 0
                    Last_IO_Error:
                   Last_SQL_Errno: 0
                   Last_SQL_Error:
      Replicate_Ignore_Server_Ids:
                 Master_Server_Id: 123
1 row in set (0.00 sec)
```

mysql-03,mysql-04 slave 操作与 mysql-02 操作相同，由于这两台只做从库，只需要修改 server id,不需要开启 log-bin。

3.2 配置 ssh 免密码登陆

配置 manager 到所有 node

```
[root@lvs-02 ~]# ssh-keygen -t rsa
Generating public/private rsa key pair.
Enter file in which to save the key (/root/.ssh/id_rsa):
Created directory '/root/.ssh'.
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in /root/.ssh/id_rsa.
Your public key has been saved in /root/.ssh/id_rsa.pub.
The key fingerprint is:
dd:73:77:b5:91:2e:6f:d5:8c:b0:3c:10:e6:41:00:6a root@lvs-02
The key's randomart image is:
+--[ RSA 2048]----+
|        ...o=       |
|      .   o o    .|
|    E       o .   o.|
|    .       . + o.o=|
|          S . *.oo*|
|                +o.o|
|                  o |
|                .   |
|                    |
+----------------+
[root@lvs-02 ~]# cd .ssh/
[root@lvs-02 .ssh]# ssh-copy-id -i id_rsa.pub 10.0.0.123
The authenticity of host '10.0.0.123 (10.0.0.123)' can't be established.
RSA key fingerprint is 62:44:e9:a4:3b:35:a6:80:81:b4:bc:24:05:27:e7:8a.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added '10.0.0.123' (RSA) to the list of known hosts.
Address 10.0.0.123 maps to localhost, but this does not map back to the address - POSSIBLE BREAK-IN ATTEMPT!
root@10.0.0.123's password:
Now try logging into the machine, with "ssh '10.0.0.123'", and check in:

  .ssh/authorized_keys

to make sure we haven't added extra keys that you weren't expecting.
```

```
[root@lvs-02 .ssh]# ssh-copy-id -i id_rsa.pub 10.0.0.124
[root@lvs-02 .ssh]# ssh-copy-id -i id_rsa.pub 10.0.0.125
[root@lvs-02 .ssh]# ssh-copy-id -i id_rsa.pub 10.0.0.126
```

配置 master 到所有 node

```
[root@mysql-01 ~]# ssh-keygen -t rsa
[root@mysql-01 .ssh]# ssh-copy-id -i 10.0.0.123
[root@mysql-01 .ssh]# ssh-copy-id -i 10.0.0.124
[root@mysql-01 .ssh]# ssh-copy-id -i 10.0.0.125
[root@mysql-01 .ssh]# ssh-copy-id -i 10.0.0.126
```

配置 backup master 到所有 node

```
[root@mysql-02 ~]# ssh-keygen -t rsa
[root@mysql-02 .ssh]# ssh-copy-id -i 10.0.0.123
[root@mysql-02 .ssh]# ssh-copy-id -i 10.0.0.124
[root@mysql-02 .ssh]# ssh-copy-id -i 10.0.0.125
[root@mysql-02 .ssh]# ssh-copy-id -i 10.0.0.126
```

配置 slave 到所有 node

```
[root@mysql-03 ~]# ssh-keygen -t rsa
[root@mysql-03 .ssh]# ssh-copy-id -i 10.0.0.123
[root@mysql-03 .ssh]# ssh-copy-id -i 10.0.0.124
[root@mysql-03 .ssh]# ssh-copy-id -i 10.0.0.125
[root@mysql-03 .ssh]# ssh-copy-id -i 10.0.0.126


[root@mysql-04 ~]# ssh-keygen -t rsa
[root@mysql-04 .ssh]# ssh-copy-id -i 10.0.0.123
[root@mysql-04 .ssh]# ssh-copy-id -i 10.0.0.124
[root@mysql-04 .ssh]# ssh-copy-id -i 10.0.0.125
[root@mysql-04 .ssh]# ssh-copy-id -i 10.0.0.126
```

3.3 在数据库中创建mha 管理用户

在 master 上创建管理用户

```
mysql> grant all on *.* to 'mha'@'10.0.0.%' identified by 'mhapwd';
Query OK, 0 rows affected (0.16 sec)

mysql> flush privileges;
Query OK, 0 rows affected (0.17 sec)
```

在从库检查是否同步

```
mysql> select user,host from mysql.user;
+------+-----------+
| user | host      |
+------+-----------+
| mha  | 10.0.0.%  |
```

14

```
| rep  | 10.0.0.%  |
| root | localhost |
+------+-----------+
3 rows in set (0.03 sec)
```

3.4 配置 mysql 环境变量

mysql-01,mysql-02,mysql-03,mysql-04 配置

```
[root@mysql-01 ~]# echo 'PATH=/usr/local/mysql/bin:$PATH' >>/etc/profile
[root@mysql-01 ~]# source /etc/profile
[root@mysql-01 ~]# which mysql
/usr/local/mysql/bin/mysql
#建立 mysql,mysqlbinlog 软链接
ln -s /usr/local/mysql/bin/* /usr/bin/
```

3.5 安装 MHA

安装 mha

mha 下载及文档地址：https://code.google.com/p/mysql-master-ha/

（需要翻墙）

安装包可在本站下载，下载地址 http://www.chocolee.cn/download/mha/

lvs-02 安装 manager

```
#安装 epel 源
wget http://mirrors.zju.edu.cn/epel/6/i386/epel-release-6-8.noarch.rpm
rpm -ivh epel-release-6-8.noarch.rpm
sed -i '/#baseurl=/ s/#base/base/g' /etc/yum.repos.d/epel.repo
sed -i '/mirrorlist=/ s/^/#/g'   /etc/yum.repos.d/epel.repo
yum repolist
#安装 node(manager 依赖)
yum localinstall mha4mysql-node-0.54-0.el6.noarch.rpm    -y
#安装 manager
yum localinstall mha4mysql-manager-0.55-0.el6.noarch.rpm    -y
```

mysql-01,mysql-02,mysql-03,mysql-04 安装 node

```
#安装 epel 源
wget http://mirrors.zju.edu.cn/epel/6/i386/epel-release-6-8.noarch.rpm
rpm -ivh epel-release-6-8.noarch.rpm
sed -i '/#baseurl=/ s/#base/base/g' /etc/yum.repos.d/epel.repo
sed -i '/mirrorlist=/ s/^/#/g'   /etc/yum.repos.d/epel.repo
yum repolist
#安装 node
yum localinstall mha4mysql-node-0.54-0.el6.noarch.rpm    -y
```

3.6 配置 MHA

配置文件参数参考地址：http://www.chocolee.cn/books/mha/MHA_MySQL.html

（转载 from http://www.mysqlsupport.cn/mha-parameters/）

manager MHA 配置文件路径：/etc/mha

```
[root@lvs-02 app1]# cat app1.conf
[server default]
manager_workdir=/etc/mha/app1
manager_log=/etc/mha/app1/manager.log
master_binlog_dir= /dbdata/data
ssh_user=root
user=mha
password=mhapwd
repl_user=rep
repl_password=reppasswd
secondary_check_script= masterha_secondary_check -s 10.0.0.126 -s 10.0.0.123
ping_interval=3


master_ip_failover_script= /etc/mha/app1/master_ip_failover
#shutdown_script= /script/masterha/power_manager
#report_script= /script/masterha/send_report
#master_ip_online_change_script= /etc/mha/master_ip_failover



[server1]
hostname=10.0.0.123
port=3306
candidate_master=1

[server2]
hostname=10.0.0.124
port=3306
candidate_master=1
#check_repl_delay=0

[server3]
hostname=10.0.0.125
port=3306
no_master=1

[server4]
hostname=10.0.0.126
```

```
port=3306
no_master=1
```

3.7 故障转移脚本

```perl
[root@lvs-02 app1]# cat master_ip_failover
#!/usr/bin/env perl

use strict;
use warnings FATAL => 'all';

use Getopt::Long;

my (
    $command,          $ssh_user,         $orig_master_host, $orig_master_ip,
    $orig_master_port, $new_master_host, $new_master_ip,     $new_master_port
);

my $vip = '10.0.0.131/24';   # Virtual IP
my $key = "1";
my $ssh_start_vip = "/sbin/ifconfig eth0:$key $vip";
my $ssh_stop_vip = "/sbin/ifconfig eth0:$key down";
$ssh_user = "root";

GetOptions(
    'command=s'         => \$command,
    'ssh_user=s'        => \$ssh_user,
    'orig_master_host=s' => \$orig_master_host,
    'orig_master_ip=s'   => \$orig_master_ip,
    'orig_master_port=i' => \$orig_master_port,
    'new_master_host=s'  => \$new_master_host,
    'new_master_ip=s'    => \$new_master_ip,
    'new_master_port=i'  => \$new_master_port,
);

exit &main();

sub main {

    print "\n\nIN SCRIPT TEST====$ssh_stop_vip==$ssh_start_vip===\n\n";

    if ( $command eq "stop" || $command eq "stopssh" ) {

        # $orig_master_host, $orig_master_ip, $orig_master_port are passed.
```

```
        # If you manage master ip address at global catalog database,
        # invalidate orig_master_ip here.
        my $exit_code = 1;

        #eval {
        #      print "Disabling the VIP on old master: $orig_master_host \n";
        #      &stop_vip();
        #      $exit_code = 0;
        #};


        eval {
                print "Disabling the VIP on old master: $orig_master_host \n";
                #my $ping=`ping -c 1 10.0.0.13 | grep "packet loss" | awk -F',' '{print $3}' |
awk '{print $1}'`;
                #if ( $ping le "90.0%" && $ping gt "0.0%" ){
                #$exit_code = 0;
                #}
                #else {

                &stop_vip();

                # updating global catalog, etc
                $exit_code = 0;

                #}
        };


        if ($@) {
            warn "Got Error: $@\n";
            exit $exit_code;
        }
        exit $exit_code;
    }
    elsif ( $command eq "start" ) {

        # all arguments are passed.
        # If you manage master ip address at global catalog database,
        # activate new_master_ip here.
        # You can also grant write access (create user, set read_only=0, etc) here.
        my $exit_code = 10;
        eval {
            print "Enabling the VIP - $vip on the new master - $new_master_host \n";
```

```perl
                    &start_vip();
                    $exit_code = 0;
              };
              if ($@) {
                    warn $@;
                    exit $exit_code;
              }
              exit $exit_code;
        }
    elsif ( $command eq "status" ) {
          print "Checking the Status of the script.. OK \n";
          `ssh $ssh_user\@$orig_master_ip \" $ssh_start_vip \"`;
          exit 0;
        }
    else {
          &usage();
          exit 1;
        }
}

# A simple system call that enable the VIP on the new master
sub start_vip() {
      `ssh $ssh_user\@$new_master_host \" $ssh_start_vip \"`;
}

# A simple system call that disable the VIP on the old_master
sub stop_vip() {
      `ssh $ssh_user\@$orig_master_host \" $ssh_stop_vip \"`;
}

sub usage {
    print
    "Usage:  master_ip_failover  --command=start|stop|stopssh|status  --orig_master_host=host
--orig_master_ip=ip   --orig_master_port=port   --new_master_host=host   --new_master_ip=ip
--new_master_port=port\n";
}

# the end.
```

3.8 backup master & slave 设置 read_only 防止被写

```
mysql> select @@read_only;
+-------------+
| @@read_only |
```

```
+------------+
|          0 |
+------------+
1 row in set (0.00 sec)


mysql> set global read_only=1;
Query OK, 0 rows affected (0.00 sec)


mysql> select @@read_only;
+------------+
| @@read_only |
+------------+
|          1 |
+------------+
1 row in set (0.00 sec)
```

3.9 检查并启动 mha

检查 SSH 情况：masterha_check_ssh --conf=/etc/mha/app1/app1.conf

```
[root@lvs-02 ~]# masterha_check_ssh --conf=/etc/mha/app1/app1.conf
Mon Apr   6 13:08:08 2015 - [warning] Global configuration file /etc/masterha_default.cnf not found. Skipping.
Mon Apr   6 13:08:08 2015 - [info] Reading application default configurations from /etc/mha/app1/app1.conf..
Mon Apr   6 13:08:08 2015 - [info] Reading server configurations from /etc/mha/app1/app1.conf..
Mon Apr   6 13:08:08 2015 - [info] Starting SSH connection tests..
Mon Apr   6 13:08:09 2015 - [debug]
Mon Apr   6 13:08:08 2015 - [debug]   Connecting via SSH from root@10.0.0.123(10.0.0.123:22) to root@10.0.0.124(10.0.0.124:22)..
Mon Apr   6 13:08:08 2015 - [debug]    ok.
Mon Apr   6 13:08:08 2015 - [debug]   Connecting via SSH from root@10.0.0.123(10.0.0.123:22) to root@10.0.0.125(10.0.0.125:22)..
Mon Apr   6 13:08:08 2015 - [debug]    ok.
Mon Apr   6 13:08:08 2015 - [debug]   Connecting via SSH from root@10.0.0.123(10.0.0.123:22) to root@10.0.0.126(10.0.0.126:22)..
Mon Apr   6 13:08:08 2015 - [debug]    ok.
Mon Apr   6 13:08:09 2015 - [debug]
Mon Apr   6 13:08:08 2015 - [debug]   Connecting via SSH from root@10.0.0.124(10.0.0.124:22) to root@10.0.0.123(10.0.0.123:22)..
Mon Apr   6 13:08:09 2015 - [debug]    ok.
Mon Apr   6 13:08:09 2015 - [debug]   Connecting via SSH from root@10.0.0.124(10.0.0.124:22) to root@10.0.0.125(10.0.0.125:22)..
Mon Apr   6 13:08:09 2015 - [debug]    ok.
Mon Apr   6 13:08:09 2015 - [debug]   Connecting via SSH from root@10.0.0.124(10.0.0.124:22)
```

to root@10.0.0.126(10.0.0.126:22)..

Mon Apr   6 13:08:09 2015 - [debug]   ok.

Mon Apr   6 13:08:10 2015 - [debug]

Mon Apr   6 13:08:09 2015 - [debug]   Connecting via SSH from root@10.0.0.125(10.0.0.125:22) to root@10.0.0.123(10.0.0.123:22)..

Mon Apr   6 13:08:09 2015 - [debug]   ok.

Mon Apr   6 13:08:09 2015 - [debug]   Connecting via SSH from root@10.0.0.125(10.0.0.125:22) to root@10.0.0.124(10.0.0.124:22)..

Mon Apr   6 13:08:09 2015 - [debug]   ok.

Mon Apr   6 13:08:09 2015 - [debug]   Connecting via SSH from root@10.0.0.125(10.0.0.125:22) to root@10.0.0.126(10.0.0.126:22)..

Mon Apr   6 13:08:10 2015 - [debug]   ok.

Mon Apr   6 13:08:10 2015 - [debug]

Mon Apr   6 13:08:09 2015 - [debug]   Connecting via SSH from root@10.0.0.126(10.0.0.126:22) to root@10.0.0.123(10.0.0.123:22)..

Mon Apr   6 13:08:10 2015 - [debug]   ok.

Mon Apr   6 13:08:10 2015 - [debug]   Connecting via SSH from root@10.0.0.126(10.0.0.126:22) to root@10.0.0.124(10.0.0.124:22)..

Mon Apr   6 13:08:10 2015 - [debug]   ok.

Mon Apr   6 13:08:10 2015 - [debug]   Connecting via SSH from root@10.0.0.126(10.0.0.126:22) to root@10.0.0.125(10.0.0.125:22)..

Mon Apr   6 13:08:10 2015 - [debug]   ok.

Mon Apr   6 13:08:10 2015 - [info] All SSH connection tests passed successfully.

检查复制情况：`masterha_check_repl --conf=/etc/mha/app1/app1.conf`

[root@lvs-02 ~]# masterha_check_ssh --conf=/etc/mha/app1/app1.conf

Mon Apr   6 13:08:08 2015 - [warning] Global configuration file /etc/masterha_default.cnf not found. Skipping.

Mon Apr   6 13:08:08 2015 - [info] Reading application default configurations from /etc/mha/app1/app1.conf..

Mon Apr   6 13:08:08 2015 - [info] Reading server configurations from /etc/mha/app1/app1.conf..

Mon Apr   6 13:08:08 2015 - [info] Starting SSH connection tests..

Mon Apr   6 13:08:09 2015 - [debug]

Mon Apr   6 13:08:08 2015 - [debug]   Connecting via SSH from root@10.0.0.123(10.0.0.123:22) to root@10.0.0.124(10.0.0.124:22)..

Mon Apr   6 13:08:08 2015 - [debug]   ok.

Mon Apr   6 13:08:08 2015 - [debug]   Connecting via SSH from root@10.0.0.123(10.0.0.123:22) to root@10.0.0.125(10.0.0.125:22)..

Mon Apr   6 13:08:08 2015 - [debug]   ok.

Mon Apr   6 13:08:08 2015 - [debug]   Connecting via SSH from root@10.0.0.123(10.0.0.123:22) to root@10.0.0.126(10.0.0.126:22)..

Mon Apr   6 13:08:08 2015 - [debug]   ok.

Mon Apr   6 13:08:09 2015 - [debug]

Mon Apr   6 13:08:08 2015 - [debug]   Connecting via SSH from root@10.0.0.124(10.0.0.124:22) to root@10.0.0.123(10.0.0.123:22)..

Mon Apr   6 13:08:09 2015 - [debug]   ok.

Mon Apr   6 13:08:09 2015 - [debug]   Connecting via SSH from root@10.0.0.124(10.0.0.124:22) to root@10.0.0.125(10.0.0.125:22)..

Mon Apr   6 13:08:09 2015 - [debug]   ok.

Mon Apr   6 13:08:09 2015 - [debug]   Connecting via SSH from root@10.0.0.124(10.0.0.124:22) to root@10.0.0.126(10.0.0.126:22)..

Mon Apr   6 13:08:09 2015 - [debug]   ok.

Mon Apr   6 13:08:10 2015 - [debug]

Mon Apr   6 13:08:09 2015 - [debug]   Connecting via SSH from root@10.0.0.125(10.0.0.125:22) to root@10.0.0.123(10.0.0.123:22)..

Mon Apr   6 13:08:09 2015 - [debug]   ok.

Mon Apr   6 13:08:09 2015 - [debug]   Connecting via SSH from root@10.0.0.125(10.0.0.125:22) to root@10.0.0.124(10.0.0.124:22)..

Mon Apr   6 13:08:09 2015 - [debug]   ok.

Mon Apr   6 13:08:09 2015 - [debug]   Connecting via SSH from root@10.0.0.125(10.0.0.125:22) to root@10.0.0.126(10.0.0.126:22)..

Mon Apr   6 13:08:10 2015 - [debug]   ok.

Mon Apr   6 13:08:10 2015 - [debug]

Mon Apr   6 13:08:09 2015 - [debug]   Connecting via SSH from root@10.0.0.126(10.0.0.126:22) to root@10.0.0.123(10.0.0.123:22)..

Mon Apr   6 13:08:10 2015 - [debug]   ok.

Mon Apr   6 13:08:10 2015 - [debug]   Connecting via SSH from root@10.0.0.126(10.0.0.126:22) to root@10.0.0.124(10.0.0.124:22)..

Mon Apr   6 13:08:10 2015 - [debug]   ok.

Mon Apr   6 13:08:10 2015 - [debug]   Connecting via SSH from root@10.0.0.126(10.0.0.126:22) to root@10.0.0.125(10.0.0.125:22)..

Mon Apr   6 13:08:10 2015 - [debug]   ok.

Mon Apr   6 13:08:10 2015 - [info] All SSH connection tests passed successfully.

[root@lvs-02 ~]# masterha_check_repl --conf=/etc/mha/app1/app1.conf

Mon Apr   6 13:09:16 2015 - [warning] Global configuration file /etc/masterha_default.cnf not found. Skipping.

Mon Apr   6 13:09:16 2015 - [info] Reading application default configurations from /etc/mha/app1/app1.conf..

Mon Apr   6 13:09:16 2015 - [info] Reading server configurations from /etc/mha/app1/app1.conf..

Mon Apr   6 13:09:16 2015 - [info] MHA::MasterMonitor version 0.55.

Mon Apr   6 13:09:16 2015 - [info] Dead Servers:

Mon Apr   6 13:09:16 2015 - [info] Alive Servers:

Mon Apr   6 13:09:16 2015 - [info]   10.0.0.123(10.0.0.123:3306)

Mon Apr   6 13:09:16 2015 - [info]   10.0.0.124(10.0.0.124:3306)

Mon Apr   6 13:09:16 2015 - [info]   10.0.0.125(10.0.0.125:3306)

Mon Apr   6 13:09:16 2015 - [info]   10.0.0.126(10.0.0.126:3306)

Mon Apr   6 13:09:16 2015 - [info] Alive Slaves:

Mon Apr   6 13:09:16 2015 - [info]   10.0.0.124(10.0.0.124:3306)   Version=5.5.32-log (oldest major version between slaves) log-bin:enabled

Mon Apr   6 13:09:16 2015 - [info]   Replicating from 10.0.0.123(10.0.0.123:3306)

Mon Apr   6 13:09:16 2015 - [info]   Primary candidate for the new Master (candidate_master is set)

Mon Apr   6 13:09:16 2015 - [info]   10.0.0.125(10.0.0.125:3306)   Version=5.5.32 (oldest major version between slaves) log-bin:disabled

Mon Apr   6 13:09:16 2015 - [info]   Replicating from 10.0.0.123(10.0.0.123:3306)

Mon Apr   6 13:09:16 2015 - [info]   Not candidate for the new Master (no_master is set)

Mon Apr   6 13:09:16 2015 - [info]   10.0.0.126(10.0.0.126:3306)   Version=5.5.32 (oldest major version between slaves) log-bin:disabled

Mon Apr   6 13:09:16 2015 - [info]   Replicating from 10.0.0.123(10.0.0.123:3306)

Mon Apr   6 13:09:16 2015 - [info]   Not candidate for the new Master (no_master is set)

Mon Apr   6 13:09:16 2015 - [info] Current Alive Master: 10.0.0.123(10.0.0.123:3306)

Mon Apr   6 13:09:16 2015 - [info] Checking slave configurations..

Mon Apr   6 13:09:16 2015 - [warning]   relay_log_purge=0 is not set on slave 10.0.0.124(10.0.0.124:3306).

Mon Apr   6 13:09:16 2015 - [warning]   relay_log_purge=0 is not set on slave 10.0.0.125(10.0.0.125:3306).

Mon Apr   6 13:09:16 2015 - [warning]   log-bin is not set on slave 10.0.0.125(10.0.0.125:3306). This host can not be a master.

Mon Apr   6 13:09:16 2015 - [warning]   relay_log_purge=0 is not set on slave 10.0.0.126(10.0.0.126:3306).

Mon Apr   6 13:09:16 2015 - [warning]   log-bin is not set on slave 10.0.0.126(10.0.0.126:3306). This host can not be a master.

Mon Apr   6 13:09:16 2015 - [info] Checking replication filtering settings..

Mon Apr   6 13:09:16 2015 - [info]   binlog_do_db= , binlog_ignore_db=

Mon Apr   6 13:09:16 2015 - [info]   Replication filtering check ok.

Mon Apr   6 13:09:16 2015 - [info] Starting SSH connection tests..

Mon Apr   6 13:09:18 2015 - [info] All SSH connection tests passed successfully.

Mon Apr   6 13:09:18 2015 - [info] Checking MHA Node version..

Mon Apr   6 13:09:19 2015 - [info]   Version check ok.

Mon Apr   6 13:09:19 2015 - [info] Checking SSH publickey authentication settings on the current master..

Mon Apr   6 13:09:19 2015 - [info] HealthCheck: SSH to 10.0.0.123 is reachable.

Mon Apr   6 13:09:19 2015 - [info] Master MHA Node version is 0.54.

Mon Apr   6 13:09:19 2015 - [info] Checking recovery script configurations on the current master..

Mon Apr   6 13:09:19 2015 - [info]   Executing command: save_binary_logs --command=test --start_pos=4   --binlog_dir=/dbdata/data   --output_file=/var/tmp/save_binary_logs_test --manager_version=0.55 --start_file=mysql-bin.000004

Mon Apr   6 13:09:19 2015 - [info]   Connecting to root@10.0.0.123(10.0.0.123)..

  Creating /var/tmp if not exists..   ok.

  Checking output directory is accessible or not..

   ok.

  Binlog found at /dbdata/data, up to mysql-bin.000004

Mon Apr 6 13:09:19 2015 - [info] Master setting check done.

Mon Apr 6 13:09:19 2015 - [info] Checking SSH publickey authentication and checking recovery script configurations on all alive slave servers..

Mon Apr 6 13:09:19 2015 - [info]   Executing command : apply_diff_relay_logs --command=test   --slave_user='mha'   --slave_host=10.0.0.124   --slave_ip=10.0.0.124 --slave_port=3306   --workdir=/var/tmp   --target_version=5.5.32-log   --manager_version=0.55 --relay_log_info=/dbdata/data/relay-log.info   --relay_dir=/dbdata/data/   --slave_pass=xxx

Mon Apr 6 13:09:19 2015 - [info]   Connecting to root@10.0.0.124(10.0.0.124:22)..

  Checking slave recovery environment settings..

    Opening /dbdata/data/relay-log.info ... ok.

    Relay log found at /dbdata/data, up to mysql-02-relay-bin.000003

    Temporary relay log file is /dbdata/data/mysql-02-relay-bin.000003

    Testing mysql connection and privileges.. done.

    Testing mysqlbinlog output.. done.

    Cleaning up test file(s).. done.

Mon Apr 6 13:09:20 2015 - [info]   Executing command : apply_diff_relay_logs --command=test   --slave_user='mha'   --slave_host=10.0.0.125   --slave_ip=10.0.0.125 --slave_port=3306   --workdir=/var/tmp   --target_version=5.5.32   --manager_version=0.55 --relay_log_info=/dbdata/data/relay-log.info   --relay_dir=/dbdata/data/   --slave_pass=xxx

Mon Apr 6 13:09:20 2015 - [info]   Connecting to root@10.0.0.125(10.0.0.125:22)..

  Checking slave recovery environment settings..

    Opening /dbdata/data/relay-log.info ... ok.

    Relay log found at /dbdata/data, up to mysql-03-relay-bin.000002

    Temporary relay log file is /dbdata/data/mysql-03-relay-bin.000002

    Testing mysql connection and privileges.. done.

    Testing mysqlbinlog output.. done.

    Cleaning up test file(s).. done.

Mon Apr 6 13:09:20 2015 - [info]   Executing command : apply_diff_relay_logs --command=test   --slave_user='mha'   --slave_host=10.0.0.126   --slave_ip=10.0.0.126 --slave_port=3306   --workdir=/var/tmp   --target_version=5.5.32   --manager_version=0.55 --relay_log_info=/dbdata/data/relay-log.info   --relay_dir=/dbdata/data/   --slave_pass=xxx

Mon Apr 6 13:09:20 2015 - [info]   Connecting to root@10.0.0.126(10.0.0.126:22)..

  Checking slave recovery environment settings..

    Opening /dbdata/data/relay-log.info ... ok.

    Relay log found at /dbdata/data, up to mysql-04-relay-bin.000003

    Temporary relay log file is /dbdata/data/mysql-04-relay-bin.000003

    Testing mysql connection and privileges.. done.

    Testing mysqlbinlog output.. done.

    Cleaning up test file(s).. done.

Mon Apr 6 13:09:21 2015 - [info] Slaves settings check done.

Mon Apr 6 13:09:21 2015 - [info]

10.0.0.123 (current master)

 +--10.0.0.124

 +--10.0.0.125

```
 +--10.0.0.126

Mon Apr    6 13:09:21 2015 - [info] Checking replication health on 10.0.0.124..
Mon Apr    6 13:09:21 2015 - [info]    ok.
Mon Apr    6 13:09:21 2015 - [info] Checking replication health on 10.0.0.125..
Mon Apr    6 13:09:21 2015 - [info]    ok.
Mon Apr    6 13:09:21 2015 - [info] Checking replication health on 10.0.0.126..
Mon Apr    6 13:09:21 2015 - [info]    ok.
Mon Apr    6 13:09:21 2015 - [info] Checking master_ip_failover_script status:
Mon Apr    6 13:09:21 2015 - [info]    /etc/mha/app1/master_ip_failover --command=status
--ssh_user=root          --orig_master_host=10.0.0.123          --orig_master_ip=10.0.0.123
--orig_master_port=3306


IN SCRIPT TEST====/sbin/ifconfig eth0:1 down==/sbin/ifconfig eth0:1 10.0.0.131/24===

Checking the Status of the script.. OK
Mon Apr    6 13:09:21 2015 - [info]    OK.
Mon Apr    6 13:09:21 2015 - [warning] shutdown_script is not defined.
Mon Apr    6 13:09:21 2015 - [info] Got exit code 0 (Not master dead).

MySQL Replication Health is OK.
```

启动 mha

当有 slave 节点宕掉的情况是启动不了的，加上--ignore_fail_on_start 即使有节点宕掉也能启动 mha

```
nohup masterha_manager --conf=/etc/mha/app1/app1.conf
--ignore_fail_on_start > /etc/mha/app1/mha_manager.log < /dev/null 2>&1
&
```

检查 mysql-01 虚拟 IP

```
[root@mysql-01 ~]# ifconfig
eth0      Link encap:Ethernet    HWaddr 00:0C:29:14:7B:1E
          inet addr:10.0.0.123    Bcast:10.0.0.255    Mask:255.255.255.0
          inet6 addr: fe80::20c:29ff:fe14:7b1e/64 Scope:Link
          UP BROADCAST RUNNING MULTICAST    MTU:1500    Metric:1
          RX packets:427894 errors:0 dropped:0 overruns:0 frame:0
          TX packets:512683 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:236018271 (225.0 MiB)    TX bytes:598512311 (570.7 MiB)


eth0:1    Link encap:Ethernet    HWaddr 00:0C:29:14:7B:1E
          inet addr:10.0.0.131    Bcast:10.0.0.255    Mask:255.255.255.0
```

```
                 UP BROADCAST RUNNING MULTICAST    MTU:1500    Metric:1

lo              Link encap:Local Loopback
                inet addr:127.0.0.1    Mask:255.0.0.0
                inet6 addr: ::1/128 Scope:Host
                UP LOOPBACK RUNNING    MTU:16436    Metric:1
                RX packets:288 errors:0 dropped:0 overruns:0 frame:0
                TX packets:288 errors:0 dropped:0 overruns:0 carrier:0
                collisions:0 txqueuelen:0
                RX bytes:36354 (35.5 KiB)    TX bytes:36354 (35.5 KiB)
```

第四部分 安装配置 lvs+keepalived

4.1 安装 lvs,keepalived

```
#安装 lvs
yum install libnl* popt* -y
ln -s /usr/src/kernels/2.6.32-358.el6.x86_64/ /usr/src/linux
ll /usr/src/
mkdir -p /application/tools
cd /application/tools
wget    http://www.linuxvirtualserver.org/software/kernel-2.6/ipvsadm-1.26.tar.gz
tar zxf ipvsadm-1.26.tar.gz
cd ipvsadm-1.26
make
make install
ipvsadm
lsmod |grep ip_vs

#安装 keepalived
yum install openssl* popt* -y
#ln -s /usr/src/kernels/2.6.32-358.el6.x86_64/ /usr/src/linux
#ll /usr/src/
mkdir -p /application/tools
cd /application/tools
wget http://www.keepalived.org/software/keepalived-1.1.19.tar.gz
tar zxf keepalived-1.1.19.tar.gz
cd keepalived-1.1.19
./configure --sysconf=/etc
make
make install
cp /usr/local/sbin/keepalived    /usr/sbin/keepalived
```

```
/etc/init.d/keepalived start
ps -ef|grep keep
/etc/init.d/keepalived stop
```

注意：

1.ln 命令的链接路径要和 uname -r 输出结果内核版本对应，工作中如果安装虚拟化可以有多个内核路径。如果 kernels 目录下没有 uname -r 输出信息名的目录，需要安装 yum install kernel-devel -y

2.ipvsadm-1.26 适用于内核 2.6.28 及之后的内核版本。

3.yum install libnl* popt* -y 是安装依赖包

4.lsmod |grep ip_vs 出现了 ip_vs 等信息，证明安装成功了。

5.CentOS5.X 安装 LVS，使用 1.2.4 版本，不要用 1.2.6.

4.2 backup master & slave 配置 arp 抑制及绑定 vip

一键脚本：

```
#!/bin/bash
vip=10.0.0.132
open() {
        ifconfig lo:Rvip ${vip}/32 up
        sysctl -w net.ipv4.conf.lo.arp_announce=2
        sysctl -w net.ipv4.conf.lo.arp_ignore=1
        sysctl -w net.ipv4.conf.all.arp_announce=2
        sysctl -w net.ipv4.conf.all.arp_ignore=1
}
close() {
        ifconfig lo:Rvip down
        sysctl -w net.ipv4.conf.lo.arp_announce=0
        sysctl -w net.ipv4.conf.lo.arp_ignore=0
        sysctl -w net.ipv4.conf.all.arp_announce=0
        sysctl -w net.ipv4.conf.all.arp_ignore=0
}
case $1 in
start)
        open
..
,,
stop)
        close
..
,,
*)
        echo "Usage: $0 need argument    [start|stop]"
```

```
;;
esac
```

4.3 配置 keepalived

lvs-01 keepalived 配置文件

```
[root@lvs-01 keepalived]# cat keepalived.conf
! Configuration File for keepalived

global_defs {
    notification_email {
      test@gmail.com
    }
    notification_email_from    alert-noreply@test.com.cn
    smtp_server 127.0.0.1
    smtp_connect_timeout 30
    router_id blade1
}


# db master server.
vrrp_instance VI_1 {
    state MASTER
    interface eth0
    virtual_router_id 51
    priority 200
    advert_int 5
    authentication {
        auth_type PASS
        auth_pass 123qwe
    }
    virtual_ipaddress {
    10.0.0.132/24
    }
}


# VIP 10.0.0.132
virtual_server 10.0.0.132 3306 {
    delay_loop 10
    lb_algo rr
    lb_kind DR
    nat_mask 255.255.255.0
```

```
protocol TCP

#sorry_server 10.0.0.124 3306

real_server 10.0.0.124 3306 {
    weight 1
    TCP_CHECK {
       connect_port 3306
       connect_timeout 10
       nb_get_retry 3
       delay_before_retry 5
    }

    MISC_CHECK {
       misc_path "/etc/keepalived/check_slave.py 10.0.0.124 3306"
       misc_dynamic
    }

}


real_server 10.0.0.125 3306 {
    weight 1
    TCP_CHECK {
    connect_port 3306
    connect_timeout 10
    nb_get_retry 3
    delay_before_retry 5
    }
    MISC_CHECK {
       misc_path "/etc/keepalived/check_slave.py 10.0.0.125 3306"
       misc_dynamic
    }
}

real_server 10.0.0.126 3306 {
    weight 1
    TCP_CHECK {
    connect_port 3306
    connect_timeout 10
    nb_get_retry 3
    delay_before_retry 5
    }
    MISC_CHECK {
```

```
                misc_path "/etc/keepalived/check_slave.py 10.0.0.126 3306"
                misc_dynamic
            }
        }

}
```

lvs-02 keepalived 配置文件

```
[root@lvs-02 keepalived]# cat keepalived.conf
! Configuration File for keepalived

global_defs {
    notification_email {
        test@gmail.com
    }
    notification_email_from    alert-noreply@test.com.cn
    smtp_server 127.0.0.1
    smtp_connect_timeout 30
    router_id blade1
}


# db master server.
vrrp_instance VI_1 {
    state BACKUP
    interface eth0
    virtual_router_id 51
    priority 100
    advert_int 5
    authentication {
        auth_type PASS
        auth_pass 123qwe
    }
    virtual_ipaddress {
    10.0.0.132/24
    }
}




# VIP 10.0.0.132
virtual_server 10.0.0.132 3306 {
    delay_loop 10
    lb_algo rr
    lb_kind DR
```

```
nat_mask 255.255.255.0
protocol TCP

#sorry_server 10.0.0.124 3306

real_server 10.0.0.124 3306 {
    weight 1
    TCP_CHECK {
       connect_port 3306
       connect_timeout 10
       nb_get_retry 3
       delay_before_retry 5
    }

    MISC_CHECK {
       misc_path "/etc/keepalived/check_slave.py 10.0.0.124 3306"
       misc_dynamic
    }

}


real_server 10.0.0.125 3306 {
    weight 1
    TCP_CHECK {
    connect_port 3306
    connect_timeout 10
    nb_get_retry 3
    delay_before_retry 5
    }
    MISC_CHECK {
       misc_path "/etc/keepalived/check_slave.py 10.0.0.125 3306"
       misc_dynamic
    }
}


real_server 10.0.0.126 3306 {
    weight 1
    TCP_CHECK {
    connect_port 3306
    connect_timeout 10
    nb_get_retry 3
```

```
        delay_before_retry 5
        }
        MISC_CHECK {
            misc_path "/etc/keepalived/check_slave.py 10.0.0.126 3306"
            misc_dynamic
        }
    }

}
```

check_slave.py 文件

```python
#!/usr/bin/env python
#encoding:utf-8
import MySQLdb
import sys
ip=sys.argv[1]
user='rep'
pwd='reppasswd'
port=int(sys.argv[2])
sbm=200

Slave_IO_Running = ''
Slave_SQL_Running = ''
Seconds_Behind_Master = ''
e=''

try:
    conn = MySQLdb.connect(host=ip,user=user,passwd=pwd,port=port,charset='utf8')
    cur = conn.cursor()
    cur.execute('show slave status')
    db_info = cur.fetchall()
    for n in db_info:
        Slave_IO_Running = n[10]
        Slave_SQL_Running = n[11]
        Seconds_Behind_Master = n[32]
    cur.close()
    conn.close()
except MySQLdb.Error,e:
    print "MySQLdb Error",e

if e == "":
    if db_info != ():
        if Slave_IO_Running == "No" or Slave_SQL_Running == "No":
            #print 'thread err'
```

```
        exit(1)
    else:
        if Seconds_Behind_Master > sbm:
            #print 'timeout err'
            exit(1)
        else:
            #print 'OK'
            exit(0)
    else:
        #print 'slave err'
        exit(1)
else:
    #print 'db err'
    exit(1)
```

4.4 启动 keepalived 并检查 vip

```
[root@lvs-01 keepalived]# /etc/init.d/keepalived start
[root@lvs-01 keepalived]# ip addr|grep 10.0.0.132
    inet 10.0.0.132/24 scope global secondary eth0
[root@lvs-01 keepalived]# ipvsadm -Ln
IP Virtual Server version 1.2.1 (size=4096)
Prot LocalAddress:Port Scheduler Flags
  -> RemoteAddress:Port              Forward Weight ActiveConn InActConn
TCP   10.0.0.132:3306 rr
  -> 10.0.0.124:3306                 Route   1       0            0
  -> 10.0.0.125:3306                 Route   1       0            0
  -> 10.0.0.126:3306                 Route   1       0            0
[root@lvs-01 keepalived]#
```

第五部分 测试

5.1 测试 read vip 负载均衡

分别在从库 mysql-02、mysql-03、mysql-04 创建可区分的库，库名分别为

read_one、read_two、read_three

```
#mysql-02
mysql> create database read_one;
Query OK, 1 row affected (0.00 sec)
```

```
mysql> show databases like 'read_one';
+--------------------+
| Database (read_one) |
+--------------------+
| read_one           |
+--------------------+
1 row in set (0.00 sec)

#mysql-03,mysql-04 操作略
```

在 lvs-02 上用 mysql 客户端连接测试 read 的 vip

```
[root@lvs-02 keepalived]# mysql -umha -pmhapwd -h 10.0.0.132 -P3306 -e "show
databases"|egrep -v "*schema|mysql|test|Database"
read_one
[root@lvs-02 keepalived]# mysql -umha -pmhapwd -h 10.0.0.132 -P3306 -e "show
databases"|egrep -v "*schema|mysql|test|Database"
read_three
[root@lvs-02 keepalived]# mysql -umha -pmhapwd -h 10.0.0.132 -P3306 -e "show
databases"|egrep -v "*schema|mysql|test|Database"
read_two
[root@lvs-02 keepalived]# mysql -umha -pmhapwd -h 10.0.0.132 -P3306 -e "show
databases"|egrep -v "*schema|mysql|test|Database"
read_one
[root@lvs-02 keepalived]# mysql -umha -pmhapwd -h 10.0.0.132 -P3306 -e "show
databases"|egrep -v "*schema|mysql|test|Database"
read_three
[root@lvs-02 keepalived]# mysql -umha -pmhapwd -h 10.0.0.132 -P3306 -e "show
databases"|egrep -v "*schema|mysql|test|Database"
read_two
```

5.2 测试从库故障被剔除，恢复被挂起

将 mysql-04 数据库关闭，观察 lvs 状态

```
#lvs 未关闭 mysql 时的状态
[root@lvs-01 keepalived]# ipvsadm -Ln
IP Virtual Server version 1.2.1 (size=4096)
Prot LocalAddress:Port Scheduler Flags
  -> RemoteAddress:Port           Forward Weight ActiveConn InActConn
TCP   10.0.0.132:3306 rr
  -> 10.0.0.124:3306              Route   1       0          4
  -> 10.0.0.125:3306              Route   1       0          5
  -> 10.0.0.126:3306              Route   1       0          5
```

34

#关闭 mysql-04

[root@mysql-04 mysql]# mysqladmin shutdown -uroot -p123456

150427 23:26:25 mysqld_safe mysqld from pid file /dbdata/data/mysql-04.pid ended

[1]+   Done   ./bin/mysqld_safe --defaults-file=/dbdata/my.cnf

--user=mysql

[root@mysql-04 mysql]# ps -fe |grep mysql

root   10254   9022   0 23:26 pts/2   00:00:00 grep mysql

#再次观察 lvs 状态

[root@lvs-01 keepalived]# ipvsadm -Ln

IP Virtual Server version 1.2.1 (size=4096)

Prot LocalAddress:Port Scheduler Flags

  -> RemoteAddress:Port   Forward Weight ActiveConn InActConn

TCP   10.0.0.132:3306 rr

  -> 10.0.0.124:3306   Route   1   0   0

  -> 10.0.0.125:3306   Route   1   0   0

#启动 mysql-04

[root@mysql-04 mysql]# ./bin/mysqld_safe --defaults-file=/dbdata/my.cnf --user=mysql&

[1] 10255

[root@mysql-04 mysql]# 150427 23:27:59 mysqld_safe Logging to '/dbdata/data/mysql-04.err'.

150427 23:27:59 mysqld_safe Starting mysqld daemon with databases from /dbdata/data

[root@mysql-04 mysql]# ps -ef |grep mysql

root   10255   9022   0 23:27 pts/2   00:00:00 /bin/sh ./bin/mysqld_safe

--defaults-file=/dbdata/my.cnf --user=mysql

mysql   10507 10255   1 23:27 pts/2   00:00:00 /usr/local/mysql/bin/mysqld

--defaults-file=/dbdata/my.cnf --basedir=/usr/local/mysql --datadir=/dbdata/data

--plugin-dir=/usr/local/mysql/lib/plugin --user=mysql --log-error=/dbdata/data/mysql-04.err

--pid-file=/dbdata/data/mysql-04.pid --socket=/tmp/mysql.sock --port=3306

root   10529   9022   0 23:28 pts/2   00:00:00 grep mysql

#再次观察 lvs 状态，看 mysql 是否被挂起

[root@lvs-01 keepalived]# ipvsadm -Ln

IP Virtual Server version 1.2.1 (size=4096)

Prot LocalAddress:Port Scheduler Flags

  -> RemoteAddress:Port   Forward Weight ActiveConn InActConn

TCP   10.0.0.132:3306 rr

  -> 10.0.0.124:3306   Route   1   0   0

  -> 10.0.0.125:3306   Route   1   0   0

  -> 10.0.0.126:3306   Route   1   0   0

5.3 测试 keepalived 高可用 vip 切换

关闭 lvs-01 的 keepalived,查看 vip 漂移情况

```
#lvs-01 关闭 keepalived
[root@lvs-01 keepalived]# ip addr |grep 10.0.0.132
    inet 10.0.0.132/24 scope global secondary eth0
[root@lvs-01 keepalived]# /etc/init.d/keepalived stop
停止 keepalived：                                    [确定]
[root@lvs-01 keepalived]# ip addr |grep 10.0.0.132


#lvs-02 检查是否获得 VIP
[root@lvs-02 keepalived]# ip addr |grep 10.0.0.132
    inet 10.0.0.132/24 scope global secondary eth0


#lvs-01 启动 keepalived
[root@lvs-01 keepalived]# ip addr |grep 10.0.0.132
[root@lvs-01 keepalived]# /etc/init.d/keepalived start
正在启动 keepalived：                                [确定]
[root@lvs-01 keepalived]# ip addr |grep 10.0.0.132
[root@lvs-01 keepalived]# ip addr |grep 10.0.0.132
[root@lvs-01 keepalived]# ip addr |grep 10.0.0.132
    inet 10.0.0.132/24 scope global secondary eth0
```

5.4 测试 write vip 切换，backup master 成为 master

关闭 mysql-01(master),查看 mha 的 manager.log,观察 write vip、slave 换新
主的过程

```
[root@lvs-02 app1]# tail -f manager.log
Mon Apr 27 23:42:51 2015 - [warning] Got error on MySQL select ping: 2006 (MySQL server
has gone away)
Mon Apr 27 23:42:51 2015 - [info] Executing secoanry network check script:
masterha_secondary_check -s 10.0.0.126 -s 10.0.0.123    --user=root    --master_host=10.0.0.123
--master_ip=10.0.0.123    --master_port=3306
Mon Apr 27 23:42:51 2015 - [info] Executing SSH check script: save_binary_logs
--command=test --start_pos=4 --binlog_dir=/dbdata/data
--output_file=/var/tmp/save_binary_logs_test --manager_version=0.55 --binlog_prefix=mysql-bin
Mon Apr 27 23:42:52 2015 - [info] HealthCheck: SSH to 10.0.0.123 is reachable.
Monitoring server 10.0.0.126 is reachable, Master is not reachable from 10.0.0.126. OK.
Monitoring server 10.0.0.123 is reachable, Master is not reachable from 10.0.0.123. OK.
Mon Apr 27 23:42:53 2015 - [info] Master is not reachable from all other monitoring servers.
Failover should start.
Mon Apr 27 23:42:54 2015 - [warning] Got error on MySQL connect: 2013 (Lost connection to
MySQL server at 'reading initial communication packet', system error: 111)
Mon Apr 27 23:42:54 2015 - [warning] Connection failed 1 time(s)..
Mon Apr 27 23:42:57 2015 - [warning] Got error on MySQL connect: 2013 (Lost connection to
MySQL server at 'reading initial communication packet', system error: 111)
```

Mon Apr 27 23:42:57 2015 - [warning] Connection failed 2 time(s)..

Mon Apr 27 23:43:00 2015 - [warning] Got error on MySQL connect: 2013 (Lost connection to MySQL server at 'reading initial communication packet', system error: 111)

Mon Apr 27 23:43:00 2015 - [warning] Connection failed 3 time(s)..

Mon Apr 27 23:43:00 2015 - [warning] Master is not reachable from health checker!

Mon Apr 27 23:43:00 2015 - [warning] Master 10.0.0.123(10.0.0.123:3306) is not reachable!

Mon Apr 27 23:43:00 2015 - [warning] SSH is reachable.

Mon Apr 27 23:43:00 2015 - [info] Connecting to a master server failed. Reading configuration file /etc/masterha_default.cnf and /etc/mha/app1/app1.conf again, and trying to connect to all servers to check server status..

Mon Apr 27 23:43:00 2015 - [warning] Global configuration file /etc/masterha_default.cnf not found. Skipping.

Mon Apr 27 23:43:00 2015 - [info] Reading application default configurations from /etc/mha/app1/app1.conf..

Mon Apr 27 23:43:00 2015 - [info] Reading server configurations from /etc/mha/app1/app1.conf..

Mon Apr 27 23:43:00 2015 - [info] Dead Servers:

Mon Apr 27 23:43:00 2015 - [info]   10.0.0.123(10.0.0.123:3306)

Mon Apr 27 23:43:00 2015 - [info] Alive Servers:

Mon Apr 27 23:43:00 2015 - [info]   10.0.0.124(10.0.0.124:3306)

Mon Apr 27 23:43:00 2015 - [info]   10.0.0.125(10.0.0.125:3306)

Mon Apr 27 23:43:00 2015 - [info]   10.0.0.126(10.0.0.126:3306)

Mon Apr 27 23:43:00 2015 - [info] Alive Slaves:

Mon Apr 27 23:43:00 2015 - [info]   10.0.0.124(10.0.0.124:3306)   Version=5.5.32-log (oldest major version between slaves) log-bin:enabled

Mon Apr 27 23:43:00 2015 - [info]   Replicating from 10.0.0.123(10.0.0.123:3306)

Mon Apr 27 23:43:00 2015 - [info]   Primary candidate for the new Master (candidate_master is set)

Mon Apr 27 23:43:00 2015 - [info]   10.0.0.125(10.0.0.125:3306)   Version=5.5.32 (oldest major version between slaves) log-bin:disabled

Mon Apr 27 23:43:00 2015 - [info]   Replicating from 10.0.0.123(10.0.0.123:3306)

Mon Apr 27 23:43:00 2015 - [info]   Not candidate for the new Master (no_master is set)

Mon Apr 27 23:43:00 2015 - [info]   10.0.0.126(10.0.0.126:3306)   Version=5.5.32 (oldest major version between slaves) log-bin:disabled

Mon Apr 27 23:43:00 2015 - [info]   Replicating from 10.0.0.123(10.0.0.123:3306)

Mon Apr 27 23:43:00 2015 - [info]   Not candidate for the new Master (no_master is set)

Mon Apr 27 23:43:00 2015 - [info] Checking slave configurations..

Mon Apr 27 23:43:00 2015 - [info]   read_only=1 is not set on slave 10.0.0.124(10.0.0.124:3306).

Mon Apr 27 23:43:00 2015 - [warning]   relay_log_purge=0 is not set on slave 10.0.0.124(10.0.0.124:3306).

Mon Apr 27 23:43:00 2015 - [info]   read_only=1 is not set on slave 10.0.0.125(10.0.0.125:3306).

Mon Apr 27 23:43:00 2015 - [warning]   relay_log_purge=0 is not set on slave 10.0.0.125(10.0.0.125:3306).

Mon Apr 27 23:43:00 2015 - [warning]   log-bin is not set on slave 10.0.0.125(10.0.0.125:3306). This host can not be a master.

Mon Apr 27 23:43:00 2015 - [info]   read_only=1 is not set on slave 10.0.0.126(10.0.0.126:3306).

Mon Apr 27 23:43:00 2015 - [warning]   relay_log_purge=0 is not set on slave 10.0.0.126(10.0.0.126:3306).

Mon Apr 27 23:43:00 2015 - [warning]   log-bin is not set on slave 10.0.0.126(10.0.0.126:3306). This host can not be a master.

Mon Apr 27 23:43:00 2015 - [info] Checking replication filtering settings..

Mon Apr 27 23:43:00 2015 - [info]   Replication filtering check ok.

Mon Apr 27 23:43:00 2015 - [info] Master is down!

Mon Apr 27 23:43:00 2015 - [info] Terminating monitoring script.

Mon Apr 27 23:43:00 2015 - [info] Got exit code 20 (Master dead).

Mon Apr 27 23:43:00 2015 - [info] MHA::MasterFailover version 0.55.

Mon Apr 27 23:43:00 2015 - [info] Starting master failover.

Mon Apr 27 23:43:00 2015 - [info]

Mon Apr 27 23:43:00 2015 - [info] * Phase 1: Configuration Check Phase..

Mon Apr 27 23:43:00 2015 - [info]

Mon Apr 27 23:43:01 2015 - [info] Dead Servers:

Mon Apr 27 23:43:01 2015 - [info]   10.0.0.123(10.0.0.123:3306)

Mon Apr 27 23:43:01 2015 - [info] Checking master reachability via mysql(double check)..

Mon Apr 27 23:43:01 2015 - [info]   ok.

Mon Apr 27 23:43:01 2015 - [info] Alive Servers:

Mon Apr 27 23:43:01 2015 - [info]   10.0.0.124(10.0.0.124:3306)

Mon Apr 27 23:43:01 2015 - [info]   10.0.0.125(10.0.0.125:3306)

Mon Apr 27 23:43:01 2015 - [info]   10.0.0.126(10.0.0.126:3306)

Mon Apr 27 23:43:01 2015 - [info] Alive Slaves:

Mon Apr 27 23:43:01 2015 - [info]   10.0.0.124(10.0.0.124:3306)   Version=5.5.32-log (oldest major version between slaves) log-bin:enabled

Mon Apr 27 23:43:01 2015 - [info]   Replicating from 10.0.0.123(10.0.0.123:3306)

Mon Apr 27 23:43:01 2015 - [info]   Primary candidate for the new Master (candidate_master is set)

Mon Apr 27 23:43:01 2015 - [info]   10.0.0.125(10.0.0.125:3306)   Version=5.5.32 (oldest major version between slaves) log-bin:disabled

Mon Apr 27 23:43:01 2015 - [info]   Replicating from 10.0.0.123(10.0.0.123:3306)

Mon Apr 27 23:43:01 2015 - [info]   Not candidate for the new Master (no_master is set)

Mon Apr 27 23:43:01 2015 - [info]   10.0.0.126(10.0.0.126:3306)   Version=5.5.32 (oldest major version between slaves) log-bin:disabled

Mon Apr 27 23:43:01 2015 - [info]   Replicating from 10.0.0.123(10.0.0.123:3306)

Mon Apr 27 23:43:01 2015 - [info]   Not candidate for the new Master (no_master is set)

Mon Apr 27 23:43:01 2015 - [info] ** Phase 1: Configuration Check Phase completed.

Mon Apr 27 23:43:01 2015 - [info]

Mon Apr 27 23:43:01 2015 - [info] * Phase 2: Dead Master Shutdown Phase..

Mon Apr 27 23:43:01 2015 - [info]

Mon Apr 27 23:43:01 2015 - [info] Forcing shutdown so that applications never connect to the current master..

Mon Apr 27 23:43:01 2015 - [info] Executing master IP deactivatation script:

Mon Apr 27 23:43:01 2015 - [info]   /etc/mha/app1/master_ip_failover
--orig_master_host=10.0.0.123 --orig_master_ip=10.0.0.123 --orig_master_port=3306
--command=stopssh --ssh_user=root

IN SCRIPT TEST====/sbin/ifconfig eth0:1 down==/sbin/ifconfig eth0:1 10.0.0.131/24===

Disabling the VIP on old master: 10.0.0.123
Mon Apr 27 23:43:01 2015 - [info]   done.
Mon Apr 27 23:43:01 2015 - [warning] shutdown_script is not set. Skipping explicit shutting
down of the dead master.
Mon Apr 27 23:43:01 2015 - [info] * Phase 2: Dead Master Shutdown Phase completed.
Mon Apr 27 23:43:01 2015 - [info]
Mon Apr 27 23:43:01 2015 - [info] * Phase 3: Master Recovery Phase..
Mon Apr 27 23:43:01 2015 - [info]
Mon Apr 27 23:43:01 2015 - [info] * Phase 3.1: Getting Latest Slaves Phase..
Mon Apr 27 23:43:01 2015 - [info]
Mon Apr 27 23:43:01 2015 - [info] The latest binary log file/position on all slaves is
mysql-bin.000004:316
Mon Apr 27 23:43:01 2015 - [info] Latest slaves (Slaves that received relay log files to the latest):
Mon Apr 27 23:43:01 2015 - [info]   10.0.0.124(10.0.0.124:3306)   Version=5.5.32-log (oldest
major version between slaves) log-bin:enabled
Mon Apr 27 23:43:01 2015 - [info]   Replicating from 10.0.0.123(10.0.0.123:3306)
Mon Apr 27 23:43:01 2015 - [info]   Primary candidate for the new Master (candidate_master
is set)
Mon Apr 27 23:43:01 2015 - [info]   10.0.0.125(10.0.0.125:3306)   Version=5.5.32 (oldest
major version between slaves) log-bin:disabled
Mon Apr 27 23:43:01 2015 - [info]   Replicating from 10.0.0.123(10.0.0.123:3306)
Mon Apr 27 23:43:01 2015 - [info]   Not candidate for the new Master (no_master is set)
Mon Apr 27 23:43:01 2015 - [info]   10.0.0.126(10.0.0.126:3306)   Version=5.5.32 (oldest
major version between slaves) log-bin:disabled
Mon Apr 27 23:43:01 2015 - [info]   Replicating from 10.0.0.123(10.0.0.123:3306)
Mon Apr 27 23:43:01 2015 - [info]   Not candidate for the new Master (no_master is set)
Mon Apr 27 23:43:01 2015 - [info] The oldest binary log file/position on all slaves is
mysql-bin.000004:316
Mon Apr 27 23:43:01 2015 - [info] Oldest slaves:
Mon Apr 27 23:43:01 2015 - [info]   10.0.0.124(10.0.0.124:3306)   Version=5.5.32-log (oldest
major version between slaves) log-bin:enabled
Mon Apr 27 23:43:01 2015 - [info]   Replicating from 10.0.0.123(10.0.0.123:3306)
Mon Apr 27 23:43:01 2015 - [info]   Primary candidate for the new Master (candidate_master
is set)
Mon Apr 27 23:43:01 2015 - [info]   10.0.0.125(10.0.0.125:3306)   Version=5.5.32 (oldest
major version between slaves) log-bin:disabled
Mon Apr 27 23:43:01 2015 - [info]   Replicating from 10.0.0.123(10.0.0.123:3306)

Mon Apr 27 23:43:01 2015 - [info]   Not candidate for the new Master (no_master is set)

Mon Apr 27 23:43:01 2015 - [info]   10.0.0.126(10.0.0.126:3306)   Version=5.5.32 (oldest major version between slaves) log-bin:disabled

Mon Apr 27 23:43:01 2015 - [info]   Replicating from 10.0.0.123(10.0.0.123:3306)

Mon Apr 27 23:43:01 2015 - [info]   Not candidate for the new Master (no_master is set)

Mon Apr 27 23:43:01 2015 - [info]

Mon Apr 27 23:43:01 2015 - [info] * Phase 3.2: Saving Dead Master's Binlog Phase..

Mon Apr 27 23:43:01 2015 - [info]

Mon Apr 27 23:43:01 2015 - [info] Fetching dead master's binary logs..

Mon Apr 27 23:43:01 2015 - [info] Executing command on the dead master 10.0.0.123(10.0.0.123:3306): save_binary_logs --command=save --start_file=mysql-bin.000004 --start_pos=316 --binlog_dir=/dbdata/data --output_file=/var/tmp/saved_master_binlog_from_10.0.0.123_3306_20150427234300.binlog --handle_raw_binlog=1 --disable_log_bin=0 --manager_version=0.55

  Creating /var/tmp if not exists..   ok.

 Concat binary/relay logs from mysql-bin.000004 pos 316 to mysql-bin.000004 EOF into /var/tmp/saved_master_binlog_from_10.0.0.123_3306_20150427234300.binlog ..

  Dumping binlog format description event, from position 0 to 107.. ok.

  Dumping effective binlog data from /dbdata/data/mysql-bin.000004 position 316 to tail(335).. ok.

 Concat succeeded.

Mon Apr 27 23:43:02 2015 - [info] scp from root@10.0.0.123:/var/tmp/saved_master_binlog_from_10.0.0.123_3306_20150427234300.binlog to local:/etc/mha/app1/saved_master_binlog_from_10.0.0.123_3306_20150427234300.binlog succeeded.

Mon Apr 27 23:43:02 2015 - [info] HealthCheck: SSH to 10.0.0.124 is reachable.

Mon Apr 27 23:43:03 2015 - [info] HealthCheck: SSH to 10.0.0.125 is reachable.

Mon Apr 27 23:43:04 2015 - [info] HealthCheck: SSH to 10.0.0.126 is reachable.

Mon Apr 27 23:43:05 2015 - [info]

Mon Apr 27 23:43:05 2015 - [info] * Phase 3.3: Determining New Master Phase..

Mon Apr 27 23:43:05 2015 - [info]

Mon Apr 27 23:43:05 2015 - [info] Finding the latest slave that has all relay logs for recovering other slaves..

Mon Apr 27 23:43:05 2015 - [info] All slaves received relay logs to the same position. No need to resync each other.

Mon Apr 27 23:43:05 2015 - [info] Searching new master from slaves..

Mon Apr 27 23:43:05 2015 - [info]   Candidate masters from the configuration file:

Mon Apr 27 23:43:05 2015 - [info]   10.0.0.124(10.0.0.124:3306)   Version=5.5.32-log (oldest major version between slaves) log-bin:enabled

Mon Apr 27 23:43:05 2015 - [info]   Replicating from 10.0.0.123(10.0.0.123:3306)

Mon Apr 27 23:43:05 2015 - [info]   Primary candidate for the new Master (candidate_master is set)

Mon Apr 27 23:43:05 2015 - [info]   Non-candidate masters:

Mon Apr 27 23:43:05 2015 - [info]   10.0.0.125(10.0.0.125:3306)   Version=5.5.32 (oldest

major version between slaves) log-bin:disabled

Mon Apr 27 23:43:05 2015 - [info]   Replicating from 10.0.0.123(10.0.0.123:3306)

Mon Apr 27 23:43:05 2015 - [info]   Not candidate for the new Master (no_master is set)

Mon Apr 27 23:43:05 2015 - [info]   10.0.0.126(10.0.0.126:3306)   Version=5.5.32 (oldest major version between slaves) log-bin:disabled

Mon Apr 27 23:43:05 2015 - [info]   Replicating from 10.0.0.123(10.0.0.123:3306)

Mon Apr 27 23:43:05 2015 - [info]   Not candidate for the new Master (no_master is set)

Mon Apr 27 23:43:05 2015 - [info]   Searching from candidate_master slaves which have received the latest relay log events..

Mon Apr 27 23:43:05 2015 - [info] New master is 10.0.0.124(10.0.0.124:3306)

Mon Apr 27 23:43:05 2015 - [info] Starting master failover..

Mon Apr 27 23:43:05 2015 - [info]

From:

10.0.0.123 (current master)

 +--10.0.0.124

 +--10.0.0.125

 +--10.0.0.126

To:

10.0.0.124 (new master)

 +--10.0.0.125

 +--10.0.0.126

Mon Apr 27 23:43:05 2015 - [info]

Mon Apr 27 23:43:05 2015 - [info] * Phase 3.3: New Master Diff Log Generation Phase..

Mon Apr 27 23:43:05 2015 - [info]

Mon Apr 27 23:43:05 2015 - [info]   This server has all relay logs. No need to generate diff files from the latest slave.

Mon Apr 27 23:43:05 2015 - [info] Sending binlog..

Mon Apr 27 23:43:06 2015 - [info] scp from local:/etc/mha/app1/saved_master_binlog_from_10.0.0.123_3306_20150427234300.binlog to root@10.0.0.124:/var/tmp/saved_master_binlog_from_10.0.0.123_3306_20150427234300.binlog succeeded.

Mon Apr 27 23:43:06 2015 - [info]

Mon Apr 27 23:43:06 2015 - [info] * Phase 3.4: Master Log Apply Phase..

Mon Apr 27 23:43:06 2015 - [info]

Mon Apr 27 23:43:06 2015 - [info] *NOTICE: If any error happens from this phase, manual recovery is needed.

Mon Apr 27 23:43:06 2015 - [info] Starting recovery on 10.0.0.124(10.0.0.124:3306)..

Mon Apr 27 23:43:06 2015 - [info]   Generating diffs succeeded.

Mon Apr 27 23:43:06 2015 - [info] Waiting until all relay logs are applied.

Mon Apr 27 23:43:06 2015 - [info]   done.

Mon Apr 27 23:43:06 2015 - [info] Getting slave status..

Mon Apr 27 23:43:06 2015 - [info] This slave(10.0.0.124)'s Exec_Master_Log_Pos equals to Read_Master_Log_Pos(mysql-bin.000004:316). No need to recover from Exec_Master_Log_Pos.

Mon Apr 27 23:43:06 2015 - [info] Connecting to the target slave host 10.0.0.124, running recover script..

Mon Apr 27 23:43:06 2015 - [info] Executing command: apply_diff_relay_logs --command=apply --slave_user='mha' --slave_host=10.0.0.124 --slave_ip=10.0.0.124 --slave_port=3306 --apply_files=/var/tmp/saved_master_binlog_from_10.0.0.123_3306_20150427234300.binlog --workdir=/var/tmp --target_version=5.5.32-log --timestamp=20150427234300 --handle_raw_binlog=1 --disable_log_bin=0 --manager_version=0.55 --slave_pass=xxx

Mon Apr 27 23:43:06 2015 - [info]

Applying differential binary/relay log files /var/tmp/saved_master_binlog_from_10.0.0.123_3306_20150427234300.binlog on 10.0.0.124:3306. This may take long time...

Applying log files succeeded.

Mon Apr 27 23:43:06 2015 - [info]   All relay logs were successfully applied.

Mon Apr 27 23:43:06 2015 - [info] Getting new master's binlog name and position..

Mon Apr 27 23:43:06 2015 - [info]   mysql-bin.000006:198

Mon Apr 27 23:43:06 2015 - [info]   All other slaves should start replication from here. Statement should be: CHANGE MASTER TO MASTER_HOST='10.0.0.124', MASTER_PORT=3306, MASTER_LOG_FILE='mysql-bin.000006', MASTER_LOG_POS=198, MASTER_USER='rep', MASTER_PASSWORD='xxx';

Mon Apr 27 23:43:06 2015 - [info] Executing master IP activate script:

Mon Apr 27 23:43:06 2015 - [info]   /etc/mha/app1/master_ip_failover --command=start --ssh_user=root --orig_master_host=10.0.0.123 --orig_master_ip=10.0.0.123 --orig_master_port=3306 --new_master_host=10.0.0.124 --new_master_ip=10.0.0.124 --new_master_port=3306 --new_master_user='mha' --new_master_password='mhapwd'

Unknown option: new_master_user

Unknown option: new_master_password

IN SCRIPT TEST====/sbin/ifconfig eth0:1 down==/sbin/ifconfig eth0:1 10.0.0.131/24===

Enabling the VIP - 10.0.0.131/24 on the new master - 10.0.0.124

Mon Apr 27 23:43:06 2015 - [info]   OK.

Mon Apr 27 23:43:06 2015 - [info] ** Finished master recovery successfully.

Mon Apr 27 23:43:06 2015 - [info] * Phase 3: Master Recovery Phase completed.

Mon Apr 27 23:43:06 2015 - [info]

Mon Apr 27 23:43:06 2015 - [info] * Phase 4: Slaves Recovery Phase..

Mon Apr 27 23:43:06 2015 - [info]

Mon Apr 27 23:43:06 2015 - [info] * Phase 4.1: Starting Parallel Slave Diff Log Generation Phase..

Mon Apr 27 23:43:06 2015 - [info]

Mon Apr 27 23:43:06 2015 - [info] -- Slave diff file generation on host 10.0.0.125(10.0.0.125:3306) started, pid: 27507. Check tmp log /etc/mha/app1/10.0.0.125_3306_20150427234300.log if it takes time..

Mon Apr 27 23:43:06 2015 - [info] -- Slave diff file generation on host

10.0.0.126(10.0.0.126:3306) started, pid: 27508. Check tmp log
/etc/mha/app1/10.0.0.126_3306_20150427234300.log if it takes time..

Mon Apr 27 23:43:06 2015 - [info]

Mon Apr 27 23:43:06 2015 - [info] Log messages from 10.0.0.125 ...

Mon Apr 27 23:43:06 2015 - [info]

Mon Apr 27 23:43:06 2015 - [info]   This server has all relay logs. No need to generate diff files from the latest slave.

Mon Apr 27 23:43:06 2015 - [info] End of log messages from 10.0.0.125.

Mon Apr 27 23:43:06 2015 - [info] -- 10.0.0.125(10.0.0.125:3306) has the latest relay log events.

Mon Apr 27 23:43:06 2015 - [info]

Mon Apr 27 23:43:06 2015 - [info] Log messages from 10.0.0.126 ...

Mon Apr 27 23:43:06 2015 - [info]

Mon Apr 27 23:43:06 2015 - [info]   This server has all relay logs. No need to generate diff files from the latest slave.

Mon Apr 27 23:43:06 2015 - [info] End of log messages from 10.0.0.126.

Mon Apr 27 23:43:06 2015 - [info] -- 10.0.0.126(10.0.0.126:3306) has the latest relay log events.

Mon Apr 27 23:43:06 2015 - [info] Generating relay diff files from the latest slave succeeded.

Mon Apr 27 23:43:06 2015 - [info]

Mon Apr 27 23:43:06 2015 - [info] * Phase 4.2: Starting Parallel Slave Log Apply Phase..

Mon Apr 27 23:43:06 2015 - [info]

Mon Apr 27 23:43:06 2015 - [info] -- Slave recovery on host 10.0.0.125(10.0.0.125:3306) started, pid: 27511. Check tmp log /etc/mha/app1/10.0.0.125_3306_20150427234300.log if it takes time..

Mon Apr 27 23:43:06 2015 - [info] -- Slave recovery on host 10.0.0.126(10.0.0.126:3306) started, pid: 27513. Check tmp log /etc/mha/app1/10.0.0.126_3306_20150427234300.log if it takes time..

Mon Apr 27 23:43:08 2015 - [info]

Mon Apr 27 23:43:08 2015 - [info] Log messages from 10.0.0.125 ...

Mon Apr 27 23:43:08 2015 - [info]

Mon Apr 27 23:43:06 2015 - [info] Sending binlog..

Mon Apr 27 23:43:07 2015 - [info] scp from
local:/etc/mha/app1/saved_master_binlog_from_10.0.0.123_3306_20150427234300.binlog to
root@10.0.0.125:/var/tmp/saved_master_binlog_from_10.0.0.123_3306_20150427234300.binlog
succeeded.

Mon Apr 27 23:43:07 2015 - [info] Starting recovery on 10.0.0.125(10.0.0.125:3306)..

Mon Apr 27 23:43:07 2015 - [info]   Generating diffs succeeded.

Mon Apr 27 23:43:07 2015 - [info] Waiting until all relay logs are applied.

Mon Apr 27 23:43:07 2015 - [info]   done.

Mon Apr 27 23:43:07 2015 - [info] Getting slave status..

Mon Apr 27 23:43:07 2015 - [info] This slave(10.0.0.125)'s Exec_Master_Log_Pos equals to Read_Master_Log_Pos(mysql-bin.000004:316). No need to recover from Exec_Master_Log_Pos.

Mon Apr 27 23:43:07 2015 - [info] Connecting to the target slave host 10.0.0.125, running recover script..

Mon Apr 27 23:43:07 2015 - [info] Executing command: apply_diff_relay_logs --command=apply --slave_user='mha' --slave_host=10.0.0.125 --slave_ip=10.0.0.125   --slave_port=3306 --apply_files=/var/tmp/saved_master_binlog_from_10.0.0.123_3306_20150427234300.binlog

--workdir=/var/tmp --target_version=5.5.32 --timestamp=20150427234300
--handle_raw_binlog=1 --disable_log_bin=0 --manager_version=0.55 --slave_pass=xxx
Mon Apr 27 23:43:07 2015 - [info]
Applying differential binary/relay log files
/var/tmp/saved_master_binlog_from_10.0.0.123_3306_20150427234300.binlog on
10.0.0.125:3306. This may take long time...
Applying log files succeeded.
Mon Apr 27 23:43:07 2015 - [info]   All relay logs were successfully applied.
Mon Apr 27 23:43:07 2015 - [info]   Resetting slave 10.0.0.125(10.0.0.125:3306) and starting
replication from the new master 10.0.0.124(10.0.0.124:3306)..
Mon Apr 27 23:43:08 2015 - [info]   Executed CHANGE MASTER.
Mon Apr 27 23:43:08 2015 - [info]   Slave started.
Mon Apr 27 23:43:08 2015 - [info] End of log messages from 10.0.0.125.
Mon Apr 27 23:43:08 2015 - [info] -- Slave recovery on host 10.0.0.125(10.0.0.125:3306)
succeeded.
Mon Apr 27 23:43:08 2015 - [info]
Mon Apr 27 23:43:08 2015 - [info] Log messages from 10.0.0.126 ...
Mon Apr 27 23:43:08 2015 - [info]
Mon Apr 27 23:43:06 2015 - [info] Sending binlog..
Mon Apr 27 23:43:07 2015 - [info] scp from
local:/etc/mha/app1/saved_master_binlog_from_10.0.0.123_3306_20150427234300.binlog to
root@10.0.0.126:/var/tmp/saved_master_binlog_from_10.0.0.123_3306_20150427234300.binlog
succeeded.
Mon Apr 27 23:43:07 2015 - [info] Starting recovery on 10.0.0.126(10.0.0.126:3306)..
Mon Apr 27 23:43:07 2015 - [info]   Generating diffs succeeded.
Mon Apr 27 23:43:07 2015 - [info] Waiting until all relay logs are applied.
Mon Apr 27 23:43:07 2015 - [info]   done.
Mon Apr 27 23:43:07 2015 - [info] Getting slave status..
Mon Apr 27 23:43:07 2015 - [info] This slave(10.0.0.126)'s Exec_Master_Log_Pos equals to
Read_Master_Log_Pos(mysql-bin.000004:316). No need to recover from Exec_Master_Log_Pos.
Mon Apr 27 23:43:07 2015 - [info] Connecting to the target slave host 10.0.0.126, running recover
script..
Mon Apr 27 23:43:07 2015 - [info] Executing command: apply_diff_relay_logs --command=apply
--slave_user='mha' --slave_host=10.0.0.126 --slave_ip=10.0.0.126   --slave_port=3306
--apply_files=/var/tmp/saved_master_binlog_from_10.0.0.123_3306_20150427234300.binlog
--workdir=/var/tmp --target_version=5.5.32 --timestamp=20150427234300
--handle_raw_binlog=1 --disable_log_bin=0 --manager_version=0.55 --slave_pass=xxx
Mon Apr 27 23:43:08 2015 - [info]
Applying differential binary/relay log files
/var/tmp/saved_master_binlog_from_10.0.0.123_3306_20150427234300.binlog on
10.0.0.126:3306. This may take long time...
Applying log files succeeded.
Mon Apr 27 23:43:08 2015 - [info]   All relay logs were successfully applied.
Mon Apr 27 23:43:08 2015 - [info]   Resetting slave 10.0.0.126(10.0.0.126:3306) and starting

44

replication from the new master 10.0.0.124(10.0.0.124:3306)..

Mon Apr 27 23:43:08 2015 - [info]   Executed CHANGE MASTER.

Mon Apr 27 23:43:08 2015 - [info]   Slave started.

Mon Apr 27 23:43:08 2015 - [info] End of log messages from 10.0.0.126.

Mon Apr 27 23:43:08 2015 - [info] -- Slave recovery on host 10.0.0.126(10.0.0.126:3306) succeeded.

Mon Apr 27 23:43:08 2015 - [info] All new slave servers recovered successfully.

Mon Apr 27 23:43:08 2015 - [info]

Mon Apr 27 23:43:08 2015 - [info] * Phase 5: New master cleanup phase..

Mon Apr 27 23:43:08 2015 - [info]

Mon Apr 27 23:43:08 2015 - [info] Resetting slave info on the new master..

Mon Apr 27 23:43:08 2015 - [info]   10.0.0.124: Resetting slave info succeeded.

Mon Apr 27 23:43:08 2015 - [info] Master failover to 10.0.0.124(10.0.0.124:3306) completed successfully.

Mon Apr 27 23:43:08 2015 - [info]

----- Failover Report -----

app1: MySQL Master failover 10.0.0.123 to 10.0.0.124 succeeded

Master 10.0.0.123 is down!

Check MHA Manager logs at lvs-02:/etc/mha/app1/manager.log for details.

Started automated(non-interactive) failover.

Invalidated master IP address on 10.0.0.123.

The latest slave 10.0.0.124(10.0.0.124:3306) has all relay logs for recovery.

Selected 10.0.0.124 as a new master.

10.0.0.124: OK: Applying all logs succeeded.

10.0.0.124: OK: Activated master IP address.

10.0.0.125: This host has the latest relay log events.

10.0.0.126: This host has the latest relay log events.

Generating relay diff files from the latest slave succeeded.

10.0.0.125: OK: Applying all logs succeeded. Slave started, replicating from 10.0.0.124.

10.0.0.126: OK: Applying all logs succeeded. Slave started, replicating from 10.0.0.124.

10.0.0.124: Resetting slave info succeeded.

Master failover to 10.0.0.124(10.0.0.124:3306) completed successfully.

检查 write vip 是否漂移

```
[root@mysql-02 mysql]# ifconfig
eth0      Link encap:Ethernet    HWaddr 00:0C:29:62:39:F8
          inet addr:10.0.0.124    Bcast:10.0.0.255    Mask:255.255.255.0
          inet6 addr: fe80::20c:29ff:fe62:39f8/64 Scope:Link
          UP BROADCAST RUNNING MULTICAST    MTU:1500    Metric:1
```

```
        RX packets:236889 errors:0 dropped:0 overruns:0 frame:0
        TX packets:69213 errors:0 dropped:0 overruns:0 carrier:0
        collisions:0 txqueuelen:1000
        RX bytes:221270914 (211.0 MiB)   TX bytes:5949360 (5.6 MiB)


eth0:1      Link encap:Ethernet   HWaddr 00:0C:29:62:39:F8
        inet addr:10.0.0.131   Bcast:10.0.0.255   Mask:255.255.255.0
        UP BROADCAST RUNNING MULTICAST   MTU:1500   Metric:1


lo          Link encap:Local Loopback
        inet addr:127.0.0.1   Mask:255.0.0.0
        inet6 addr: ::1/128 Scope:Host
        UP LOOPBACK RUNNING   MTU:16436   Metric:1
        RX packets:244 errors:0 dropped:0 overruns:0 frame:0
        TX packets:244 errors:0 dropped:0 overruns:0 carrier:0
        collisions:0 txqueuelen:0
        RX bytes:22781 (22.2 KiB)   TX bytes:22781 (22.2 KiB)


lo:Rvip     Link encap:Local Loopback
        inet addr:10.0.0.132   Mask:0.0.0.0
        UP LOOPBACK RUNNING   MTU:16436   Metric:1
```

检查从库是否换了新主

```
mysql> show slave status\G
*************************** 1. row ***************************
                Slave_IO_State: Waiting for master to send event
                   Master_Host: 10.0.0.124
                   Master_User: rep
                   Master_Port: 3306
                 Connect_Retry: 60
               Master_Log_File: mysql-bin.000006
           Read_Master_Log_Pos: 198
                Relay_Log_File: mysql-03-relay-bin.000002
                 Relay_Log_Pos: 253
         Relay_Master_Log_File: mysql-bin.000006
              Slave_IO_Running: Yes
             Slave_SQL_Running: Yes
.....................
```

lvs 检查新主是否在 read 组中被剔除

```
[root@lvs-01 keepalived]# ipvsadm -Ln
IP Virtual Server version 1.2.1 (size=4096)
Prot LocalAddress:Port Scheduler Flags
  -> RemoteAddress:Port           Forward Weight ActiveConn InActConn
```

TCP	10.0.0.132:3306 rr				
-> 10.0.0.125:3306	Route	1	0	0	
-> 10.0.0.126:3306	Route	1	0	0	