通过日志分析 MHA 工作原理

老男孩 IT 教育（北京）
分析：陈桂林（50 期）
日期：2018－11－27
导师：郭导

背景：

角色	IP	软件
主节点	10.0.0.146	MYSQ，MHA NODE
从节点	10.0.0.145	MYSQL，MHA NODE
从节点	10.0.0.147	MYSQL，MHA MANAGER

我这里一主两从，模拟故障时我仅仅是手动关闭了 146 节点的 MYSQL 服务，并没有关机 146 主机，所以你会发现 SSH 可用。

==========

Tue Nov 27 10:19:45 2018 - [warning] Got error on MySQL select ping: 2006 (MySQL server has gone away)＃MHA 管理端检测到有一个 MYSQL 节点离线了
Tue Nov 27 10:19:45 2018 - [info] Executing SSH check script: exit 0 ＃马上发起健康检查，系统会发起三次检查，如下：
Tue Nov 27 10:19:45 2018 - [info] HealthCheck: SSH to 10.0.0.146 is reachable.
Tue Nov 27 10:19:47 2018 - [warning] Got error on MySQL connect: 2003 (Can't connect to MySQL server on '10.0.0.146' (111))
Tue Nov 27 10:19:47 2018 - [warning] Connection failed 2 time(s)..
Tue Nov 27 10:19:49 2018 - [warning] Got error on MySQL connect: 2003 (Can't connect to MySQL server on '10.0.0.146' (111))
Tue Nov 27 10:19:49 2018 - [warning] Connection failed 3 time(s)..
Tue Nov 27 10:19:51 2018 - [warning] Got error on MySQL connect: 2003 (Can't connect to MySQL server on '10.0.0.146' (111))
Tue Nov 27 10:19:51 2018 - [warning] Connection failed 4 time(s)..
Tue Nov 27 10:19:51 2018 - [warning] Master is not reachable from health checker!＃经过上面三次检查，无法获取到健康状态
Tue Nov 27 10:19:51 2018 - [warning] Master 10.0.0.146(10.0.0.146:3306) is not reachable!＃这里已经下结论了，说 146 这个主节点不可用了。
Tue Nov 27 10:19:51 2018 - [warning] SSH is reachable. ＃SSH 是可用的
Tue Nov 27 10:19:51 2018 - [info] Connecting to a master server failed. Reading configuration file /etc/masterha_default.cnf and /etc/mha/app1.cnf again, and trying to connect to all servers to check server status.. ＃连接到主节点服务器失败，重新读取 MHA 配置文件（默认有两个地方：/etc/masterha_default.cnf and /etc/mha/app1.cnf），并连接配置文件中所有节点和检查状态。
Tue Nov 27 10:19:51 2018 - [warning] Global configuration file /etc/masterha_default.cnf not found. Skipping. ＃MHA 发现/etc/masterha_default.cnf 不存在。

Tue Nov 27 10:19:51 2018 - [info] Reading application default configuration from /etc/mha/app1.cnf.. ＃MHA 再次读取/etc/mha/app1.cnf 配置文件，由此可见 MHA 是优先读取/etc/masterha_default.cnf 配置文件。
Tue Nov 27 10:19:51 2018 - [info] Reading server configuration from /etc/mha/app1.cnf.. ＃正在从/etc/mha/app1.cnf 文件中读取配置
Tue Nov 27 10:19:52 2018 - [info] GTID failover mode = 1
Tue Nov 27 10:19:52 2018 - [info] Dead Servers:＃系统判定这是一台死去的节点服务器
Tue Nov 27 10:19:52 2018 - [info] 10.0.0.146(10.0.0.146:3306)
Tue Nov 27 10:19:52 2018 - [info] Alive Servers:＃系统判定这是活动的节点服务器
Tue Nov 27 10:19:52 2018 - [info] 10.0.0.145(10.0.0.145:3306)
Tue Nov 27 10:19:52 2018 - [info] 10.0.0.147(10.0.0.147:3306)
Tue Nov 27 10:19:52 2018 - [info] Alive Slaves:
Tue Nov 27 10:19:52 2018 - [info] 10.0.0.145(10.0.0.145:3306) Version=5.6.38-log (oldest major version between slaves) log-bin:enabled ＃系统判定这是在节点中最古老的从服务器，换句话说这台可以优先作为新的主服务器，并且检测到已开启了 bin-log
Tue Nov 27 10:19:52 2018 - [info] GTID ON
Tue Nov 27 10:19:52 2018 - [info] Replicating from 10.0.0.146(10.0.0.146:3306)＃开始从原 MYSQL 主节点中复制
Tue Nov 27 10:19:52 2018 - [info] 10.0.0.147(10.0.0.147:3306) Version=5.6.38-log (oldest major version between slaves) log-bin:enabled ＃系统判定这是在节点中最古老的从服务器，换句话说这台可以优先作为新的主服务器，并且检测到已开启了 bin-log
Tue Nov 27 10:19:52 2018 - [info] GTID ON
Tue Nov 27 10:19:52 2018 - [info] Replicating from 10.0.0.146(10.0.0.146:3306)＃开始从原 MYSQL 主节点中复制
Tue Nov 27 10:19:52 2018 - [info] Checking slave configurations..＃检查从服务器的配置
Tue Nov 27 10:19:52 2018 - [info] read_only=1 is not set on slave 10.0.0.145(10.0.0.145:3306).
Tue Nov 27 10:19:52 2018 - [info] read_only=1 is not set on slave 10.0.0.147(10.0.0.147:3306).
Tue Nov 27 10:19:52 2018 - [info] Checking replication filtering settings..#检查复制过滤设置
Tue Nov 27 10:19:52 2018 - [info] Replication filtering check ok.#检查复制过滤设置完成
Tue Nov 27 10:19:52 2018 - [info] Master is down!＃主节点服务器已停机
Tue Nov 27 10:19:52 2018 - [info] Terminating monitoring script. ＃停止监控
Tue Nov 27 10:19:52 2018 - [info] Got exit code 20 (Master dead). ＃退出，原因是主节点死了
Tue Nov 27 10:19:52 2018 - [info] MHA::MasterFailover version 0.57.
Tue Nov 27 10:19:52 2018 - [info] Starting master failover. ＃开始故障转移
Tue Nov 27 10:19:52 2018 - [info]
Tue Nov 27 10:19:52 2018 - [info] * Phase 1: Configuration Check Phase.. ＃第一阶段：配置检查
Tue Nov 27 10:19:52 2018 - [info]
Tue Nov 27 10:19:53 2018 - [info] GTID failover mode = 1
Tue Nov 27 10:19:53 2018 - [info] Dead Servers:＃死掉的节点服务器
Tue Nov 27 10:19:53 2018 - [info] 10.0.0.146(10.0.0.146:3306)
Tue Nov 27 10:19:53 2018 - [info] Checking master reachability via MySQL(double check)... ＃通过 MYSQL 检查主机的可用性（双重校验）
Tue Nov 27 10:19:53 2018 - [info] ok.
Tue Nov 27 10:19:53 2018 - [info] Alive Servers:＃活动的服务器
Tue Nov 27 10:19:53 2018 - [info] 10.0.0.145(10.0.0.145:3306)
Tue Nov 27 10:19:53 2018 - [info] 10.0.0.147(10.0.0.147:3306)
Tue Nov 27 10:19:53 2018 - [info] Alive Slaves:

Tue Nov 27 10:19:53 2018 - [info] 10.0.0.145(10.0.0.145:3306) Version=5.6.38-log (oldest major version between slaves) log-bin:enabled ＃系统判定这是在节点中最古老的从服务器，换句话说这台可以优先作为新的主服务器，并且检测到已开启了 bin-log
Tue Nov 27 10:19:53 2018 - [info] GTID ON
Tue Nov 27 10:19:53 2018 - [info] Replicating from 10.0.0.146(10.0.0.146:3306) ＃开始从原 MYSQL 主节点中复制
Tue Nov 27 10:19:53 2018 - [info] 10.0.0.147(10.0.0.147:3306) Version=5.6.38-log (oldest major version between slaves) log-bin:enabled ＃系统判定这是在节点中最古老的从服务器，换句话说这台可以优先作为新的主服务器，并且检测到已开启了 bin-log
Tue Nov 27 10:19:53 2018 - [info] GTID ON
Tue Nov 27 10:19:53 2018 - [info] Replicating from 10.0.0.146(10.0.0.146:3306) ＃开始从原 MYSQL 主节点中复制
Tue Nov 27 10:19:53 2018 - [info] Starting GTID based failover. ＃启动基于 GTID 的故障转移
Tue Nov 27 10:19:53 2018 - [info]
Tue Nov 27 10:19:53 2018 - [info] ** Phase 1: Configuration Check Phase completed. ＃第一阶段：检查配置文件完成
Tue Nov 27 10:19:53 2018 - [info]
Tue Nov 27 10:19:53 2018 - [info] * Phase 2: Dead Master Shutdown Phase.. ＃第二阶段：停用坏的主节点阶段
Tue Nov 27 10:19:53 2018 - [info]
Tue Nov 27 10:19:53 2018 - [info] Forcing shutdown so that applications never connect to the current master.. ＃强制关闭使应用程序不会再连接到当前主节点服务器
Tue Nov 27 10:19:53 2018 - [info] Executing master IP deactivation script: ＃执行主 IP 失活脚本
Tue Nov 27 10:19:53 2018 - [info] /usr/local/bin/master_ip_failover --orig_master_host=10.0.0.146 --orig_master_ip=10.0.0.146 --orig_master_port=3306 --command=stopssh --ssh_user=root ＃执行 VIP 切换脚本

＃开始 VIP 切换
IN SCRIPT TEST====/sbin/ifconfig eth0:0 down==/sbin/ifconfig eth0:0 10.0.0.148/24===

Disabling the VIP on old master: 10.0.0.146 ＃将 146 节点主的 VIP 禁用
Tue Nov 27 10:19:53 2018 - [info] done. ＃完成
Tue Nov 27 10:19:53 2018 - [warning] shutdown_script is not set. Skipping explicit shutting down of the dead master. 脚本中没有设置跳过坏的主节点的显式关闭
Tue Nov 27 10:19:53 2018 - [info] * Phase 2: Dead Master Shutdown Phase completed. ＃第二阶段：关闭坏的主节点完成
Tue Nov 27 10:19:53 2018 - [info]
Tue Nov 27 10:19:53 2018 - [info] * Phase 3: Master Recovery Phase.. ＃第三阶段：恢复主节点阶段
Tue Nov 27 10:19:53 2018 - [info]
Tue Nov 27 10:19:53 2018 - [info] * Phase 3.1: Getting Latest Slaves Phase.. ＃获取最新的从状态
Tue Nov 27 10:19:53 2018 - [info]
Tue Nov 27 10:19:53 2018 - [info] The latest binary log file/position on all slaves is mysql-bin.000006:231 ＃所有从节点中最新的 BIN－LOG 和 POSITION 是：mysql-bin.000006:231
Tue Nov 27 10:19:53 2018 - [info] Retrieved Gtid Set: 37e290de-f146-11e8-82d5-000c29c80a99:1-2 ＃检索 GTID 设置
Tue Nov 27 10:19:53 2018 - [info] Latest slaves (Slaves that received relay log files to the latest):＃最新的从节点（收到的最新中继日志文件）

Tue Nov 27 10:19:53 2018 - [info] 10.0.0.145(10.0.0.145:3306) Version=5.6.38-log (oldest major version between slaves) log-bin:enabled ＃145
Tue Nov 27 10:19:53 2018 - [info] GTID ON
Tue Nov 27 10:19:53 2018 - [info] Replicating from 10.0.0.146(10.0.0.146:3306) ＃从 146 开始复制
Tue Nov 27 10:19:53 2018 - [info] 10.0.0.147(10.0.0.147:3306) Version=5.6.38-log (oldest major version between slaves) log-bin:enabled ＃146
Tue Nov 27 10:19:53 2018 - [info] GTID ON
Tue Nov 27 10:19:53 2018 - [info] Replicating from 10.0.0.146(10.0.0.146:3306) ＃从 146 开始复制
Tue Nov 27 10:19:53 2018 - [info] The oldest binary log file/position on all slaves is mysql-bin.000006:231 ＃所有从节点中最新的 BIN－LOG 和 POSITION 是：mysql-bin.000006:231
Tue Nov 27 10:19:53 2018 - [info] Retrieved Gtid Set: 37e290de-f146-11e8-82d5-000c29c80a99:1-2 ＃检索 GTID 设置
Tue Nov 27 10:19:53 2018 - [info] Oldest slaves: 最新的从节点（或者说运行最久的，数据存的最多的）
Tue Nov 27 10:19:53 2018 - [info] 10.0.0.145(10.0.0.145:3306) Version=5.6.38-log (oldest major version between slaves) log-bin:enabled
Tue Nov 27 10:19:53 2018 - [info] GTID ON
Tue Nov 27 10:19:53 2018 - [info] Replicating from 10.0.0.146(10.0.0.146:3306) ＃从 146 开始复制
Tue Nov 27 10:19:53 2018 - [info] 10.0.0.147(10.0.0.147:3306) Version=5.6.38-log (oldest major version between slaves) log-bin:enabled
Tue Nov 27 10:19:53 2018 - [info] GTID ON
Tue Nov 27 10:19:53 2018 - [info] Replicating from 10.0.0.146(10.0.0.146:3306)＃从 146 开始复制
Tue Nov 27 10:19:53 2018 - [info]
Tue Nov 27 10:19:53 2018 - [info] * Phase 3.3: Determining New Master Phase.. ＃第三阶段：确定新的 MYSQL 主节点
Tue Nov 27 10:19:53 2018 - [info]
Tue Nov 27 10:19:53 2018 - [info] Searching new master from slaves..＃从 从节点中扫描新的主节点
Tue Nov 27 10:19:53 2018 - [info] Candidate masters from the configuration file:＃配置文件中指定的候选主节点
Tue Nov 27 10:19:53 2018 - [info] Non-candidate masters:＃未定义
Tue Nov 27 10:19:53 2018 - [info] New master is 10.0.0.145(10.0.0.145:3306)新的 MYSQL 主节点是 145
Tue Nov 27 10:19:53 2018 - [info] Starting master failover.. ＃开始主节点故障转移
Tue Nov 27 10:19:53 2018 - [info]
From:
10.0.0.146(10.0.0.146:3306) (current master) ＃当前的主节点，下
 +--10.0.0.145(10.0.0.145:3306)
 +--10.0.0.147(10.0.0.147:3306)

To:
10.0.0.145(10.0.0.145:3306) (new master)#新的主节点
 +--10.0.0.147(10.0.0.147:3306)
Tue Nov 27 10:19:53 2018 - [info]

Tue Nov 27 10:19:53 2018 - [info] * Phase 3.3: New Master Recovery Phase..＃第 3.3 阶段：新的主节点恢复阶段
Tue Nov 27 10:19:53 2018 - [info]
Tue Nov 27 10:19:53 2018 - [info] Waiting all logs to be applied.. ＃等待所有日志应用完成
Tue Nov 27 10:19:53 2018 - [info] done.＃完成
Tue Nov 27 10:19:53 2018 - [info] Getting new master's binlog name and position.. ＃获取新的主节点 BIN－LOG 和 POSITION
Tue Nov 27 10:19:53 2018 - [info] mysql-bin.000004:231 ＃这是获取结果
Tue Nov 27 10:19:53 2018 - [info] All other slaves should start replication from here. Statement should be: CHANGE MASTER TO MASTER_HOST='10.0.0.145', MASTER_PORT=3306, MASTER_AUTO_POSITION=1, MASTER_USER='repl', MASTER_PASSWORD='xxx'; ＃所有的从节点应该是从这里复制，语句是：CHANGE MASTER TO MASTER_HOST='10.0.0.145', MASTER_PORT=3306, MASTER_AUTO_POSITION=1, MASTER_USER='repl', MASTER_PASSWORD='xxx';

主节点恢复完成，文件和 POS 以及获取到的 GTID 是 mysql-bin.000004, 231, 37e290de-f146-11e8-82d5-000c29c80a99:1-2,
e9b9d3d1-f12f-11e8-8244-000c29b152c9:1-4

Tue Nov 27 10:19:53 2018 - [info] Master Recovery succeeded. File:Pos:Exec_Gtid_Set: mysql-bin.000004, 231, 37e290de-f146-11e8-82d5-000c29c80a99:1-2,
e9b9d3d1-f12f-11e8-8244-000c29b152c9:1-4
Tue Nov 27 10:19:53 2018 - [info] Executing master IP activate script:＃执行主 IP 激活脚本
Tue Nov 27 10:19:53 2018 - [info] /usr/local/bin/master_ip_failover --command=start --ssh_user=root --orig_master_host=10.0.0.146 --orig_master_ip=10.0.0.146 --orig_master_port=3306 --new_master_host=10.0.0.145 --new_master_ip=10.0.0.145 --new_master_port=3306 --new_master_user='mha' --new_master_password=xxx
这里两个未知选项，可能是软件版本不支持的原因。
Unknown option: new_master_user
Unknown option: new_master_password

IN SCRIPT TEST====/sbin/ifconfig eth0:0 down==/sbin/ifconfig eth0:0 10.0.0.148/24===

Enabling the VIP - 10.0.0.148/24 on the new master - 10.0.0.145 ＃在 145 节点上启动 VIP
Tue Nov 27 10:19:53 2018 - [info] OK. ＃完成
Tue Nov 27 10:19:53 2018 - [info] ** Finished master recovery successfully. ＃完成主节点恢复
Tue Nov 27 10:19:53 2018 - [info] * Phase 3: Master Recovery Phase completed. 第三阶段：主节点恢复阶段完成
Tue Nov 27 10:19:53 2018 - [info]
Tue Nov 27 10:19:53 2018 - [info] * Phase 4: Slaves Recovery Phase.. ＃第四阶段：从节点恢复阶段
Tue Nov 27 10:19:53 2018 - [info]
Tue Nov 27 10:19:53 2018 - [info]
Tue Nov 27 10:19:53 2018 - [info] * Phase 4.1: Starting Slaves in parallel.. 第 4.1 阶段：并行启动从节点
Tue Nov 27 10:19:53 2018 - [info]
Tue Nov 27 10:19:53 2018 - [info] -- Slave recovery on host 10.0.0.147(10.0.0.147:3306) started, pid: 4949. Check tmp log /var/log/mha/app1/10.0.0.147_3306_20181127101952.log if it takes

time.. ＃在 147 节点是恢复从节点完成，检查日志：
/var/log/mha/app1/10.0.0.147_3306_20181127101952.log
Tue Nov 27 10:19:55 2018 - [info]
Tue Nov 27 10:19:55 2018 - [info] Log messages from 10.0.0.147 ... ＃开始应用 147 日志
Tue Nov 27 10:19:55 2018 - [info]
Tue Nov 27 10:19:53 2018 - [info] Resetting slave 10.0.0.147(10.0.0.147:3306) and starting replication from the new master 10.0.0.145(10.0.0.145:3306).. 重置从节点 147 并从新的主节点 145 开始复制
Tue Nov 27 10:19:53 2018 - [info] Executed CHANGE MASTER. 执行更改主节点
Tue Nov 27 10:19:54 2018 - [info] Slave started. 启动 SLAVE 线程
Tue Nov 27 10:19:54 2018 - [info] gtid_wait(37e290de-f146-11e8-82d5-000c29c80a99:1-2, e9b9d3d1-f12f-11e8-8244-000c29b152c9:1-4) completed on 10.0.0.147(10.0.0.147:3306). Executed 0 events.＃执行了 0 个事务
Tue Nov 27 10:19:55 2018 - [info] End of log messages from 10.0.0.147. ＃147 日志信息应用完成
Tue Nov 27 10:19:55 2018 - [info] -- Slave on host 10.0.0.147(10.0.0.147:3306) started.＃从节点 147 启动完成
Tue Nov 27 10:19:55 2018 - [info] All new slave servers recovered successfully.＃所有 MYSQL 从节点服务器恢复完成
Tue Nov 27 10:19:55 2018 - [info]
Tue Nov 27 10:19:55 2018 - [info] * Phase 5: New master cleanup phase..＃第五阶段：新主节点清理阶段
Tue Nov 27 10:19:55 2018 - [info]
Tue Nov 27 10:19:55 2018 - [info] Resetting slave info on the new master.. ＃在新的主节点上复置 SLAVE 信息
Tue Nov 27 10:19:55 2018 - [info] 10.0.0.145: Resetting slave info succeeded.＃145 节点重置 SLAVE 信息完成
Tue Nov 27 10:19:55 2018 - [info] Master failover to 10.0.0.145(10.0.0.145:3306) completed successfully. ＃主节点 145 故障转移完成
Tue Nov 27 10:19:55 2018 - [info] Deleted server2 entry from /etc/mha/app1.cnf . ＃从/etc/mha/app1.cnf 配置文件中删除 server2 标签
Tue Nov 27 10:19:55 2018 - [info]

----- Failover Report ----- ＃故障报告

app1: MySQL Master failover 10.0.0.146(10.0.0.146:3306) to 10.0.0.145(10.0.0.145:3306) succeeded ＃MYSQL 主节点故障转（146）转到（145）完成

Master 10.0.0.146(10.0.0.146:3306) is down! ＃主节点 146 已停机

Check MHA Manager logs at db02:/var/log/mha/app1/manager for details.＃从/var/log/mha/app1/manager 检查 MHA MANAGER 详细日志

Started automated(non-interactive) failover.＃开始非交互式的自动故障转移
Invalidated master IP address on 10.0.0.146(10.0.0.146:3306) ＃无效的主 IP
Selected 10.0.0.145(10.0.0.145:3306) as a new master. ＃选择 145 作为了新的主节点
10.0.0.145(10.0.0.145:3306): OK: Applying all logs succeeded. ＃145 应用所有日志完成
10.0.0.145(10.0.0.145:3306): OK: Activated master IP address. ＃145 主 IP 活动的

10.0.0.147(10.0.0.147:3306): OK: Slave started, replicating from 10.0.0.145(10.0.0.145:3306)＃147 已启动从节点，并从 145 开始复制
10.0.0.145(10.0.0.145:3306): Resetting slave info succeeded. ＃145 重新设置从节点信息完成
Master failover to 10.0.0.145(10.0.0.145:3306) completed successfully. ＃主节点 145 故障转移完成