
LightDB 高可用和逻辑复制更换 IP 操作流程

目录

一、 更换操作流程.....	1
1. 登录到主数据库查询集群状态.....	1
2. 将集群 ltcluster 置为 pause 状态.....	1
3. 主库、备库、witness 库均停止 ltclusterd.....	1
4. 更改主库、备库、witness 库 ltcluster 配置文件 ip.....	1
5. 更改备库 lightdb.auto.conf 中 IP 地址.....	1
6. 所有节点更新环境变量和 lt_hba.conf*.....	2
7. 主库、备库、witness 停止数据库.....	2
8. 更改主、备、witness 主机物理 ip.....	2
9. 服务器重启后启动数据库.....	2
10. 确认当前主从复制情况.....	2
11. 订阅服务器订阅端更改订阅信息（*如未配置逻辑复制，该步骤可跳过）.....	3
12. 登录主库再次确认复制状态（*如未配置逻辑复制，该步骤可跳过）.....	4
13. 每个节点重新注册 ltcluster.....	4
14. 验证集群状态.....	5
15. 停掉主库、备库的 keepalived.....	5
16. 编辑 keepalived.conf 文件（主备都执行）.....	5
17. root 重新启动 keepalived 进程.....	5

修订记录

日期	修订版本	描述	作者
2023/08/17	V1.0	LightDB 高可用和逻辑复制更换 IP 操作流程	姚崇

审阅

姓名	职位

分发

姓名	单位

一、更换操作流程

1. 登录到主数据库查询集群状态

首选确保集群状态正常，均为 running 状态

```
# su - lightdb
$ ltcluster -f $LTHOME/etc/ltcluster/ltcluster.conf service status
ID | Name | Role | Status | Upstream | ltclusterd | PID | Paused? | Upstream last seen
-----+-----+-----+-----+-----+-----+-----+-----+-----
1 | lightdbCluster101010165501 | primary | * running | | running | 26176 | no | n/a
2 | lightdbCluster101010265501 | standby | running | lightdbCluster101010165501 | running | 17335 | no | 1 second(s) ago
4 | lightdbCluster101010465501 | witness | * running | lightdbCluster101010165501 | running | 4914 | no | 0 second(s) ago
```

2. 将集群 ltcluster 置为 pause 状态

```
$ ltcluster -f $LTHOME/etc/ltcluster/ltcluster.conf service pause
NOTICE: node 1 (lightdbCluster101010165501) paused
NOTICE: node 2 (lightdbCluster101010265501) paused
NOTICE: node 4 (lightdbCluster101010465501) paused
[lightdb@hs-10-20-148-122 ltcluster]$ ltcluster -f $LTHOME/etc/ltcluster/ltcluster.conf service status
ID | Name | Role | Status | Upstream | ltclusterd | PID | Paused? | Upstream last seen
-----+-----+-----+-----+-----+-----+-----+-----+-----
1 | lightdbCluster101010165501 | primary | * running | | running | 26176 | yes | n/a
2 | lightdbCluster101010265501 | standby | running | lightdbCluster101010165501 | running | 17335 | yes | 0 second(s) ago
4 | lightdbCluster101010465501 | witness | * running | lightdbCluster101010165501 | running | 4914 | yes | 0 second(s) ago
```

3. 主库、备库、witness 库均停止 ltclusterd

登录到三个节点，分别查询 ltclusterd 进行，并 kill 掉

```
# su - lightdb
$ ps -ef | grep ltclusterd
lightdb 17335 1 0 16:06 ? 00:00:01 /data2/lightdb/lightdb_22.2//lightdb-x/13.3-22.2/bin/ltclusterd -d -f
/data2/lightdb/lightdb_22.2//lightdb-x/13.3-22.2/etc/ltcluster/ltcluster.conf -p /data2/lightdb/lightdb_22.2//lightdb-
x/13.3-22.2/etc/ltcluster/ltclusterd.pid
kill -9 17335
```

4. 更改主库、备库、witness 库 ltcluster 配置文件 ip

三个节点均执行，优先更改备库，然后主库和 witness

```
# su - lightdb
$ cd $LTHOME/etc/ltcluster
$ pwd
/data2/lightdb/lightdb_22.2/lightdb-x/13.3-22.2/etc/ltcluster
[lightdb@hs-10-20-148-126 ltcluster]$ vim ltcluster.conf
node_id=2
node_name='lightdbCluster101010265501'
conninfo='host=10.10.10.2 port=65501 user=ltcluster dbname=ltcluster connect_timeout=2'
# 第三行 conninfo 对应 ip 地址改为新 ip 地址
```

5. 更改备库 lightdb.auto.conf 中 IP 地址

host 对应此处地址更改为新的主库的 IP 地址

```
$ cd $LTDATA
$ vim lightdb.auto.conf
# Do not edit this file manually!
# It will be overwritten by the ALTER SYSTEM command.
```

```
primary_conninfo = 'host=10.20.148.122 port=65501 user=ltcluster application_name=lightdbCluster101010265501
connect_timeout=2'
primary_slot_name = 'ltcluster_slot_2'
restore_command = 'cp /data2/lightdb/lightdb_22.2//lightdb-x/13.3-22.2/archive/%f %p'
```

6. 所有节点更新环境变量和 lt_hba.conf*

环境变量更新为最新的 IP 地址（配置文件在 lightdb 用户下的 ~/.bashrc，修改该文件中的 LTHOST 为新 IP），**暂时无需使用 source ~/.bashrc 生效**

同时 lt_hba.conf 确认 ltcluster 登录为免密，防止后续重新注册 ltcluster 失败

```
$ vim ~/.bashrc
export LTPORT=65501
export LTUSER=lightdb
export LTHOST=10.20.148.122
export LTHOME=/data2/lightdb/lightdb_22.2//lightdb-x/13.3-22.2
export LTDATA=/data2/lightdb/lightdb_22.2//lightdb-x/13.3-22.2/data/defaultCluster/
export PATH=${LTHOME}/bin:${LTHOME}/tools/bin:${LTHOME}/tools/sbin:${PATH}
export LD_LIBRARY_PATH=${LTHOME}/lib:${LTHOME}/lib/lttext:${LTHOME}/tools/lib64:${LD_LIBRARY_PATH}

$ vim $LTDATA
$ lt_hba.conf 上方配置，红色为三个新的 IP 地址
host all lightdb 10.20.148.122/32 trust
host ltcluster ltcluster 10.20.148.122/32 trust
host replication ltcluster 10.20.148.122/32 trust
host all lightdb 10.20.148.126/32 trust
host ltcluster ltcluster 10.20.148.126/32 trust
host replication ltcluster 10.20.148.126/32 trust
host all lightdb 10.20.148.128/32 trust
host ltcluster ltcluster 10.20.148.128/32 trust
host replication ltcluster 10.20.148.128/32 trust
```

7. 主库、备库、witness 停止数据库

分别登录到 3 台服务器，停止数据库

```
$lt_ctl -D $LTDATA stop
```

8. 更改主、备、witness 主机物理 ip

此处有主机运维人员完成，**更改 IP，重启服务器**

9. 服务器重启后启动数据库

分别登录到三台服务器执行

```
$ lt_ctl -D $LTDATA start
2023-08-17 16:47:58.022138T @ postmaster 00000[2023-08-17 16:47:52 CST] 0 [17012] HINT: Future log output will appear
in
...
server started
```

10. 确认当前主从复制情况

默认逻辑复制会发生中断，所以 pg_stat_replication 中只记录一条物理流复制信息，由于主库 IP 地址发生改变所以逻辑复制中断

```
# 次查询返回一行
lightdb@postgres=# select * from pg_stat_replication ;
-[ RECORD 1 ]-----+-----
pid          | 29015
usesysid     | 20802
username     | ltcluster
application_name | lightdbCluster101010265501
```

```

client_addr      | 10.20.148.126
client_hostname  |
client_port     | 38280
backend_start   | 2023-08-17 16:51:34.410962+08
backend_xmin    |
state           | streaming
sent_lsn        | 0/80BAF530
write_lsn       | 0/80BAF530
flush_lsn       | 0/80BAF530
replay_lsn      | 0/80BAF530
write_lag       | 00:00:00.000078
flush_lag       | 00:00:00.00018
replay_lag      | 00:00:00.00022
sync_priority   | 1
sync_state      | sync
reply_time      | 2023-08-17 16:53:04.596151+08

# 此查询返回 2 行 (*如未配置逻辑复制, 此查询返回一行, 即下面的第一行)
lightdb@postgres=# select * from pg_replication_slots ;
-[ RECORD 1 ]-----+-----
slot_name          | ltcluster_slot_2
plugin             |
slot_type          | physical
datoid             |
database           |
temporary          | f
active             | t
active_pid         | 29015
xmin              |
catalog_xmin       |
restart_lsn        | 0/80BC3428
confirmed_flush_lsn |
wal_status         | reserved
safe_wal_size      | 8040729560
-[ RECORD 2 ]-----+-----
slot_name          | subname
plugin             | ltoutput
slot_type          | logical
datoid             | 20870
database           | testdb
temporary          | f
active             | f
active_pid         |
xmin              |
catalog_xmin       | 17971
restart_lsn        | 0/80B8DB40
confirmed_flush_lsn | 0/80B8E868
wal_status         | reserved
safe_wal_size      | 8040729560

```

11. 订阅服务器订阅端更改订阅信息 (*如未配置逻辑复制, 该步骤可跳过)

订阅数据库登录到业务库进行查询

```

$ltsql -p 65502 -d testdb -h 127.0.0.1
ltsql (13.3-22.2)
Type "help" for help.

lightdb@postgres=# select * from pg_subscription;
  oid | subdbid | subname | subowner | subenabled |          subconninfo          | subslotname | subsynccommit | subpublications
-----+-----+-----+-----+-----+-----+-----+-----+-----
20842 | 20807 | subname | 10 | t | hostaddr= 10.10.10.1 port=65501 password=lightdb123 user=lightdb dbname=testdb | subname | off | {pubname}
(1 row)

```

```
lightdb@testdb=# ALTER SUBSCRIPTION subname CONNECTION 'hostaddr=10.20.148.122 port=65501 password=lightdb123 user=lightdb
dbname=testdb';
ALTER SUBSCRIPTION
lightdb@testdb=# select * from pg_subscription;
 oid | subdbid | subname | subowner | subenabled |          subconninfo          | subslotname | subsynccommit | subpublications
-----+-----+-----+-----+-----+-----+-----+-----+-----
 20842 | 20807 | subname | 10 | t          | hostaddr=10.20.148.122 port=65501 password=lightdb123 user=lightdb dbname=testdb | subname     | off           | {pubname}
(1 row)
```

12. 登录主库再次确认复制状态 (*如未配置逻辑复制, 该步骤可跳过)

```
lightdb@postgres=# select * from pg_stat_replication ;
 pid | usesysid | username | application_name | client_addr | state | write_lsn | write_lag | flush_lag | replay_lag | sync_state | reply_time
-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----
 29015 | 20802 | ltcluster | lightdbCluster101010265501 | 10.20.148.126 | streaming | 0/80D50338 | 00:00:00.00011 | 00:00:00.000199 | 00:00:00.000241 | sync | 2023-08-17 17:01:51.159759+08
 2880 | 10 | lightdb | subname | 10.20.148.128 | streaming | 0/80D50338 | | | | async | 2023-08-17 17:01:16.382978+08
(2 rows)
```

13. 每个节点重新注册 ltcluster

主节点、从节点、witness 节点分别执行

```
##### 主库执行
$LTHOME/bin/ltcluster -f $LTHOME/etc/ltcluster/ltcluster.conf primary register -h 10.20.148.122 -p 65501 -F
WARNING: following problems with command line parameters detected:
 database connection parameters not required when executing PRIMARY REGISTER
INFO: connecting to primary database...
INFO: "ltcluster" extension is already installed
NOTICE: primary node record (ID: 1) updated
lightdb@ltcluster=# select * from ltcluster.nodes;
 node_id | upstream_node_id | active | node_name | type | location | priority | conninfo
-----+-----+-----+-----+-----+-----+-----+-----
 repluser | slot_name | config_file
-----+-----+-----
 4 | 1 | t | lightdbCluster101010465501 | witness | default | 0 | host=10.10.10.4 port=65501 user=ltcluster dbname=ltcluster
connect_timeout=2 | ltcluster | /data2/lightdb/lightdb_22.2/lightdb-x/13.3-22.2/e
tc/ltcluster/ltcluster.conf
 2 | 1 | t | lightdbCluster101010265501 | standby | default | 100 | host=10.10.10.2 port=65501 user=ltcluster dbname=ltcluster
connect_timeout=2 | ltcluster | ltcluster_slot_2 | /data2/lightdb/lightdb_22.2/lightdb-x/13.3-22.2/e
tc/ltcluster/ltcluster.conf
 1 | | t | lightdbCluster101010165501 | primary | default | 100 | host=10.20.148.122 port=65501 user=ltcluster dbname=ltcluster
connect_timeout=2 | ltcluster | ltcluster_slot_1 | /data2/lightdb/lightdb_22.2/lightdb-x/13.3-22.2/e
tc/ltcluster/ltcluster.conf
(3 rows)
$LTHOME/bin/ltclusterd -d -f $LTHOME/etc/ltcluster/ltcluster.conf -p ${LTHOME}/etc/ltcluster/ltcluster.pid

##### 备份库执行
$LTHOME/bin/ltcluster -f $LTHOME/etc/ltcluster/ltcluster.conf standby register -h10.20.148.122 -p 65501 -F
$LTHOME/bin/ltclusterd -d -f $LTHOME/etc/ltcluster/ltcluster.conf -p ${LTHOME}/etc/ltcluster/ltcluster.pid

##### witness 库执行
$LTHOME/bin/ltcluster -f $LTHOME/etc/ltcluster/ltcluster.conf witness register -h 10.20.148.122 -p 65501 -F
$LTHOME/bin/ltclusterd -d -f $LTHOME/etc/ltcluster/ltcluster.conf -p ${LTHOME}/etc/ltcluster/ltcluster.pid
```

14. 验证集群状态

```
$ ltcluster -f $LTHOME/etc/ltcluster/ltcluster.conf service status
```

ID	Name	Role	Status	Upstream	ltclusterd	PID	Paused?	Upstream last seen
1	lightdbCluster101010165501	primary	* running		running	1381	yes	n/a
2	lightdbCluster101010265501	standby	running	lightdbCluster101010165501	running	28972	yes	0 second(s) ago
4	lightdbCluster101010465501	witness	* running	lightdbCluster101010165501	running	96567	yes	1 second(s) ago

```
$ ltcluster -f $LTHOME/etc/ltcluster/ltcluster.conf service unpause
```

```
NOTICE: node 1 (lightdbCluster101010165501) unpaused
NOTICE: node 2 (lightdbCluster101010265501) unpaused
NOTICE: node 4 (lightdbCluster101010465501) unpaused
[lightdb@hs-10-20-148-122 defaultCluster]$
```

```
$ ltcluster -f $LTHOME/etc/ltcluster/ltcluster.conf service status
```

ID	Name	Role	Status	Upstream	ltclusterd	PID	Paused?	Upstream last seen
1	lightdbCluster101010165501	primary	* running		running	1381	no	n/a
2	lightdbCluster101010265501	standby	running	lightdbCluster101010165501	running	28972	no	0 second(s) ago
4	lightdbCluster101010465501	witness	* running	lightdbCluster101010165501	running	96567	no	0 second(s) ago

15. 停掉主库、备库的 keepalived

```
[root@standby1 ~]# kill -9 `cat /var/run/keepalived.pid`
[root@standby1 ~]# ps aux | grep keepalived
root      14675  0.0  0.0 112808  964 pts/1    R+   01:05   0:00 grep --color=auto keepalived
```

16. 编辑 keepalived.conf 文件（主备都执行）

```
$ vim $LTHOME/etc/keepalived/keepalived.conf
#更改下面 virtual_ipaddress 对应地址为新的 vip 地址
vrrp_instance lightdb {
    state MASTER
    interface ens160
    virtual_router_id 5
    priority 100
    advert_int 2
    authentication {
        auth_type PASS
        auth_pass lightdb123
    }
    virtual_ipaddress {
        10.20.148.11
    }
}
```

17. root 重新启动 keepalived 进程

```
# root 执行
# cd $LTHOME/../../keepalived-1.3.5-x86/sbin
/home/lightdb/lightdb/lightdb-x/13.3-22.2
./keepalived -f $LTHOME/etc/keepalived/keepalived.conf

./keepalived -f /home/lightdb/lightdb/lightdb-x/13.3-22.1/etc/keepalived/keepalived.conf
```