**Percona XtraDB Cluster Documentation**



**General introduction**

1. The Cluster consists of Nodes. Recommended configuration is to have at least 3 nodes, but you can make it running with 2 nodes as well.

2. Each Node is regular MySQL / Percona Server setup. The point is that you can convert your existing MySQL / Percona Server into Node and roll Cluster using it as a base. Or otherwise – you can detach Node from Cluster and use it as just a regular server.

3. Each Node contains the full copy of data. That defines XtraDB Cluster behavior in many ways. And obviously there are benefits and drawbacks.

Benefits of such approach:

• When you execute a query, it is executed locally on the node. All data is available locally, no need for

remote access.

• No central management. You can loose any node at any point of time, and the cluster will continue to

function.

• Good solution for scaling a read workload. You can put read queries to any of the nodes.

Drawbacks:

• Overhead of joining new node. The new node has to copy full dataset from one of existing nodes. If it is

100GB, it copies 100GB.

• This can’t be used as an effective write scaling solution. There might be some improvements in write

throughput when you run write traffic to 2 nodes vs all traffic to 1 node, but you can’t expect a lot. All

writes still have to go on all nodes.

• You have several duplicates of data, for 3 nodes – 3 duplicates.

**Prerequisites**

At least 3 nodes with:

* CentOS 6.4 Updated
* EPEL 6-8

To install EPEL 6-8 execute following command

http://ftp.jaist.ac.jp/pub/Linux/Fedora/epel/6/i386/epel-release-6-8.noarch.rpm

* ip(6)tables and SELinux's disabled
* ntpd ON
* Each node should mount (usually in /var/lib/mysql) a disk or partition of adequate size (100 or better 200GB)

Note: if the nodes are virtual, the virtual instances **must** run on different hardware.

**Installing Percona XtraDB Cluster Using Percona Software Repositories**

rpm -Uhv <http://repo.percona.com/testing/centos/6/os/noarch/percona-testing-0.0-1.noarch.rpm>

rpm -Uhv <http://www.percona.com/downloads/percona-release/percona-release-0.0-1.x86_64.rpm>

yum install Percona-XtraDB-Cluster-server-55 Percona-XtraDB-Cluster-client-55 Percona-XtraDB-Cluster-galera-2

### Configure Percona

On all three nodes execute the following commands (substitute IPLIST's value with the actual comma-separated IPs of your cluster):

export IPLIST="192.168.122.51,192.168.122.52,192.168.122.53"

export MYIP=`ip -f inet addr show eth0 | grep inet | awk '{print $2}'|awk -F"/" '{print $1}'`

export CLUSTER\_NAME="mysql\_cluster"

cat << EOF > /etc/my.cnf

[mysqld]

#set-variable = max\_connect\_errors=999999999

datadir=/var/lib/mysql

user=mysql

innodb\_log\_file\_size=64M

innodb\_file\_per\_table=1

innodb\_locks\_unsafe\_for\_binlog=1

# Path to Galera library

wsrep\_provider=/usr/lib64/libgalera\_smm.so

# Cluster connection URL contains the IPs of node#1, node#2 and node#3

wsrep\_cluster\_address=gcomm://$IPLIST

# In order for Galera to work correctly binlog format should be ROW

binlog\_format=ROW

# MyISAM storage engine has only experimental support

default\_storage\_engine=InnoDB

# This is a recommended tuning variable for performance

innodb\_locks\_unsafe\_for\_binlog=1

# This changes how InnoDB autoincrement locks are managed and is a requirement for Galera

innodb\_autoinc\_lock\_mode=2

# Node #1 address

wsrep\_node\_address=$MYIP

# SST method

wsrep\_sst\_method=xtrabackup

# Cluster name

wsrep\_cluster\_name=$CLUSTER\_NAME

# Authentication for SST method

wsrep\_sst\_auth="sstuser:SST\_PWD"

EOF

yum -y install xinetd

echo "mysqlchk 9200/tcp" >> /etc/services

chkconfig xinetd on

sed -i s'+/usr/bin/clustercheck+/usr/bin/clustercheck\n\tserver\_args\t= clustercheckuser CLUSTERCHECK\_PWD+' /etc/xinetd.d/mysqlchk

sed -i 's+log\_type.\*+log\_type = FILE /var/log/xinetd.log+' /etc/xinetd.conf

sed -i 's+log\_on\_success.\*+log\_on\_success =+' /etc/xinetd.conf

service xinetd start

### Bootstrap the Percona cluster

Temporary choose a node as primary. In the primary node execute the commands:

/etc/init.d/mysql start --wsrep-cluster-address="gcomm://"

chkconfig mysql on

mysql -u root -e "CREATE USER 'sstuser'@'localhost' IDENTIFIED BY 'SST\_PWD'"

mysql -u root -e "GRANT RELOAD, LOCK TABLES, REPLICATION CLIENT ON \*.\* TO 'sstuser'@'localhost';"

mysql -u root -e "GRANT PROCESS ON \*.\* TO 'clustercheckuser'@'localhost' IDENTIFIED BY 'CLUSTERCHECK\_PWD'; FLUSH PRIVILEGES;"

On the other nodes execute:

/etc/init.d/mysql start

chkconfig mysql on

Now the MySQL-Percona cluster should be up & running (you can check with ”ps -ef|grep mysql” that the processes are running on all nodes). On all nodes of the cluster execute this commands (substitute 192.168.122 with your net address):

mysql -u root -e "GRANT ALL PRIVILEGES ON \*.\* TO 'root'@'192.168.122.%' IDENTIFIED BY '' WITH GRANT OPTION;"

### Basic checks

On any node check the correct working of the script /usr/bin/clustercheck:

[root@percona1 ~]# /usr/bin/clustercheck clustercheckuser CLUSTERCHECK\_PWD

HTTP/1.1 200 OK

Content-Type: text/plain

Connection: close

Content-Length: 40

Percona XtraDB Cluster Node is synced.

Clustercheck, via xinetd, is able to communicate the sync status also to external machines; on your destkop try the following command:

dorigoa@lxadorigo 15:06:08 ~>telnet percona1.pd.infn.it 9200

Trying 192.168.122.51...

Connected to percona1.pd.infn.it.

Escape character is '^]'.

HTTP/1.1 200 OK

Content-Type: text/plain

Connection: close

Content-Length: 40

Percona XtraDB Cluster Node is synced.

Connection closed by foreign host.

Note that both commands above, are saying that the cluster is synced. If any problem occurred during installation/configuration, “not synced” is printed instead.

### Notes about starting/stopping/restarting the Percona Cluster

Above, it has been showed how to bootstrap the cluster: the primary node must be started with the option **--wsrep-cluster-address="gcomm://** and the other nodes without it (just: service mysql start).

When the cluster is running, each node can be simply restarted (without the **--wsrep-cluster-address=gcomm://** option).

**But be careful**: not using the option is allowed only when at least one node is up. If the cluster is starting from a situation in which all nodes are OFF, one of them must be started with the option (the choice of which node is based on the assumption that that node has the last database's data update).