命令行部署 OceanBase 三副本集群

使用命令行部署 OceanBase 数据库生产环境

官网文档：

<https://www.oceanbase.com/docs/common-oceanbase-database-cn-1000000000508275>

**实验环境介绍：**

虚拟软件：Vbox

四台虚拟机如下配置（单台）：

系统：Red Hat Enterprise Linux Server release 7.7 (Maipo)

内存：8G

(这里配置8G，后面创建resource可能不够用，后面会进行扩容到10G)

Cpu：4

网络：仅主机（host-only）模式

**部署模式：**

本文采用三副本部署模式，使用四台虚拟机，四台机器的使用情况如下：

|  |  |  |
| --- | --- | --- |
| **角色** | **机器** | **备注** |
| OBD | 192.168.56.56 | 安装在中控机上的自动化部署软件 |
| OBServer 节点 | 192.168.56.57 | OceanBase 数据库 Zone1 |
| OBServer 节点 | 192.168.56.58 | OceanBase 数据库 Zone2 |
| OBServer 节点 | 192.168.56.59 | OceanBase 数据库 Zone3 |
| OBAgent | 192.168.56.57、192.168.56.58、192.168.56.59 | OceanBase 数据库监控采集框架 |
| ODP | 192.168.56.56 | OceanBase 数据库专用的反向代理软件 |
| OCP Express | 192.168.56.56 | 基于 Web 的 OceanBase 数据库 4.x 管理工具 |

OCP Express节点需要配置java要求1.8版本。

# 一、准备工作

## 1.1、四个节点关闭并禁用selinux

# setenforce 0

# sed -i "s/SELINUX=enforcing/SELINUX=disabled/g" /etc/selinux/config

## 1.2、四个节点关闭并禁止防火墙自启动

# systemctl stop firewalld

# systemctl disable firewalld

## 1.3、四个节点关闭NetworkManager

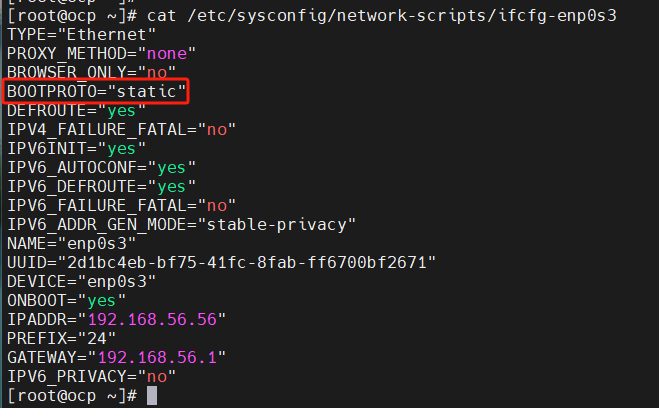
# systemctl stop NetworkManager

# systemctl disable NetworkManager

这里建议关闭NetworkManager，使用手工网络配置，否则后面初始化ob的时候会网络冲突卡住。

## 1.4、四个节点手工配置网络

这里主要是使用静态IP



## 1.5、四个节点创建admin用户

# useradd admin

# passwd admin

## 1.6、四个节点做时间同步

这里以OBD中控机（192.168.56.56）作为ntp服务器，其他三台同步该机器时钟：

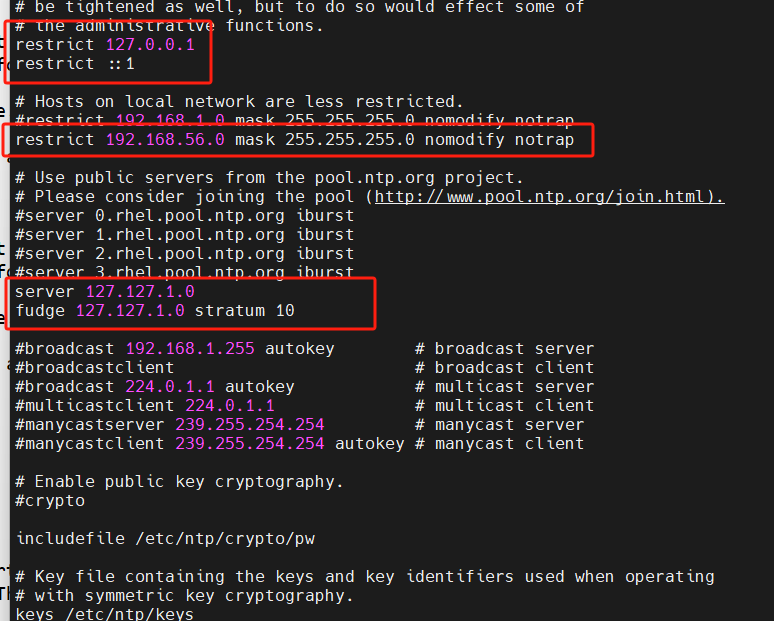
安装ntp

配置本地yum源，安装ntp

# yum install -y ntp

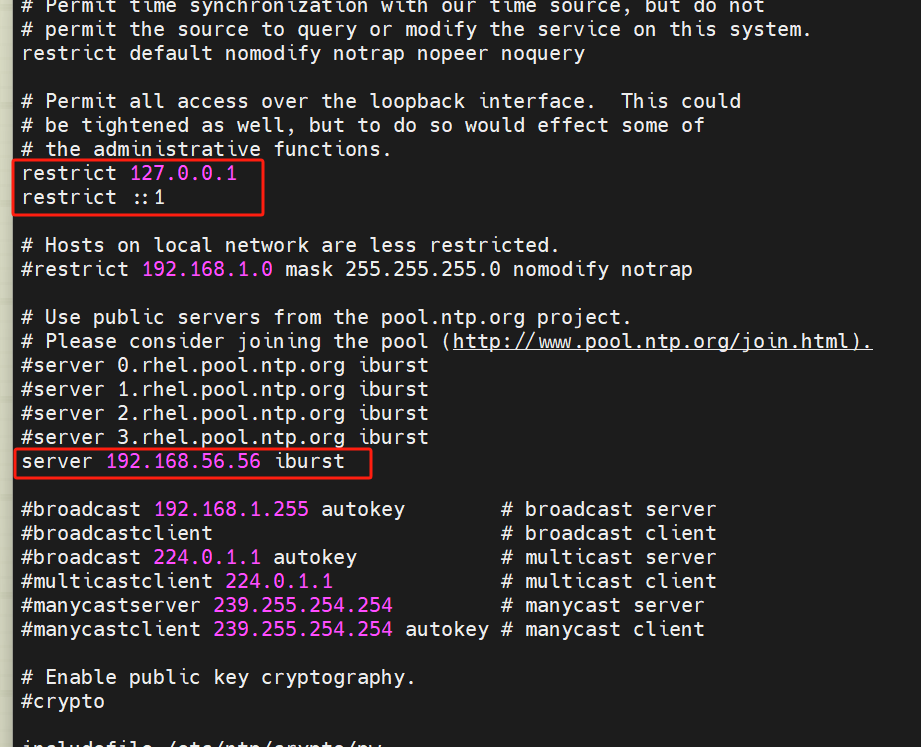
192.168.56.56：

# vi /etc/ntp.conf



192.168.56.57、192.168.56.58、192.168.56.59三台机器配置相同即可：

# vi /etc/ntp.conf



修改好配置文件后，192.168.56.56首先启动ntpd进程，生效配置文件，其他三台停止ntpd进程，手工从192.168.56.56同步时间之后再启动ntpd：

192.168.56.56：

# systemctl start ntpd

192.168.56.57、192.168.56.58、192.168.56.59：

# systemctl stop ntpd

# ntpdate 192.168.56.56

# systemctl start ntpd

## 1.7、四个节点调整limits.conf

# vi /etc/security/limits.conf

添加如下内容

root soft nofile 655350

root hard nofile 655350

\* soft nofile 655350

\* hard nofile 655350

\* soft stack unlimited

\* hard stack unlimited

\* soft nproc 655360

\* hard nproc 655360

\* soft core unlimited

\* hard core unlimited

#vi /etc/security/limits.d/nproc.conf

添加如下内容

\* soft nproc 655360

\* hard nproc 655360

## 1.8、四个节点调整sysctl.conf

# vi /etc/sysctl.conf

添加如下内容

# for oceanbase

## 修改内核异步 I/O 限制

fs.aio-max-nr=1048576

## 网络优化

net.core.somaxconn = 2048

net.core.netdev\_max\_backlog = 10000

net.core.rmem\_default = 16777216

net.core.wmem\_default = 16777216

net.core.rmem\_max = 16777216

net.core.wmem\_max = 16777216

net.ipv4.ip\_local\_port\_range = 3500 65535

net.ipv4.ip\_forward = 0

net.ipv4.conf.default.rp\_filter = 1

net.ipv4.conf.default.accept\_source\_route = 0

net.ipv4.tcp\_syncookies = 1

net.ipv4.tcp\_rmem = 4096 87380 16777216

net.ipv4.tcp\_wmem = 4096 65536 16777216

net.ipv4.tcp\_max\_syn\_backlog = 16384

net.ipv4.tcp\_fin\_timeout = 15

net.ipv4.tcp\_max\_syn\_backlog = 16384

net.ipv4.tcp\_tw\_reuse = 1

net.ipv4.tcp\_tw\_recycle = 1

net.ipv4.tcp\_slow\_start\_after\_idle=0

vm.swappiness = 0

vm.min\_free\_kbytes = 2097152

fs.file-max = 6573688

# 修改进程可以拥有的虚拟内存区域数量

vm.max\_map\_count = 655360

# 此处为 OceanBase 数据库的 data 目录

kernel.core\_pattern = /data

注意：max\_map\_count 配置不合理的情况下，可能会导致严重的内存泄露。

# 二、下载安装配置JAVA

部署 OCP Express的机器需要安装配置java，要求1.8版本

我这里在192.168.56.56上进行配置

<https://github.com/dragonwell-project/dragonwell8/releases/download/dragonwell-extended-8.14.15_jdk8u362-ga/Alibaba_Dragonwell_Extended_8.14.15_x64_linux.tar.gz>

可使用迅雷下载，很快，网页下载可能会慢

上传到虚拟机，解压安装

# 解压资源包

cd /usr/local

tar -zxvf Alibaba\_Dragonwell\_Extended\_8\*.tar.gz

# 调整目录名

mv dragonwell\* java

# 进入java目录

cd java

# 创建软连

ln -s `pwd`/bin/java /usr/bin/java

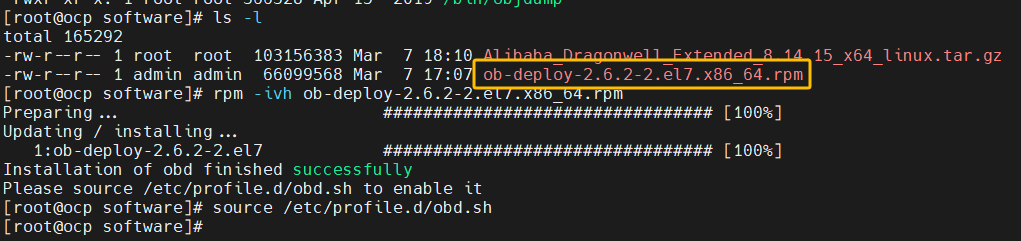
# 三、下载安装配置OBD

官网下载：

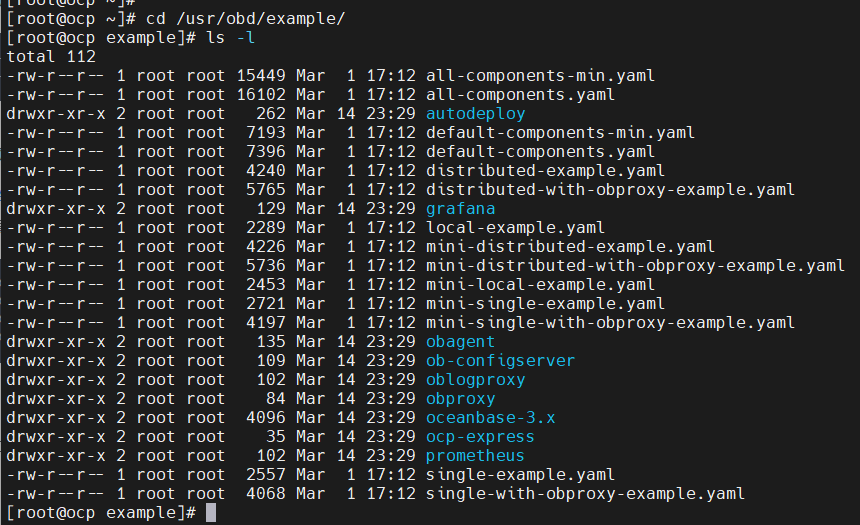
<https://www.oceanbase.com/softwarecenter>



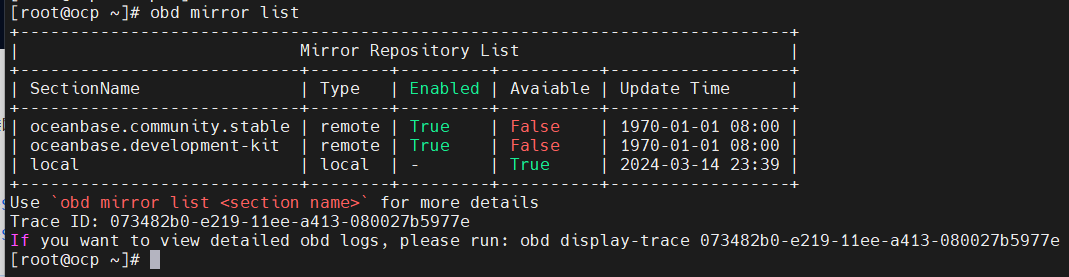
上传到OBD中控机（192.168.56.56）上，安装：



同时/usr/obd/example下面会出现ob集群配置文件模板，用来配置ob集群：

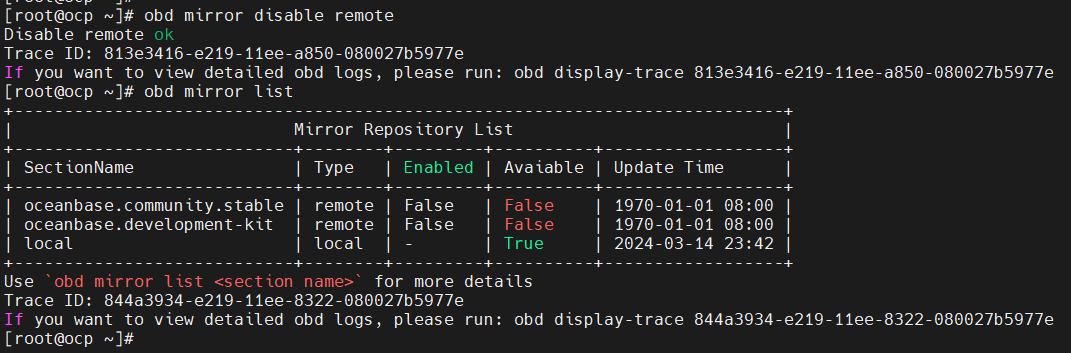


作为一个部署工具，OBD 自带仓库信息，当机器可以连接网络时，无须配置，执行 obd mirror list 命令即可查看 OBD 仓库中的软件。因为我是仅主机模式，离线的，所以看不到网络上的远程仓库信息：



既然是离线模式，直接禁用远程仓库：

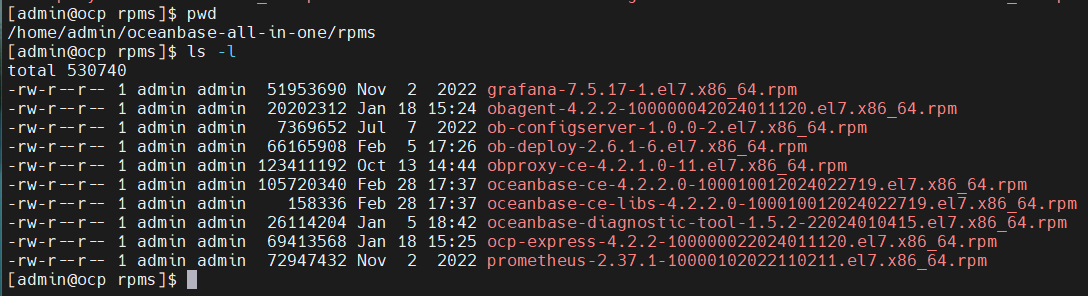
# obd mirror disable remote



禁用以后发现Enabled一列变成了False。

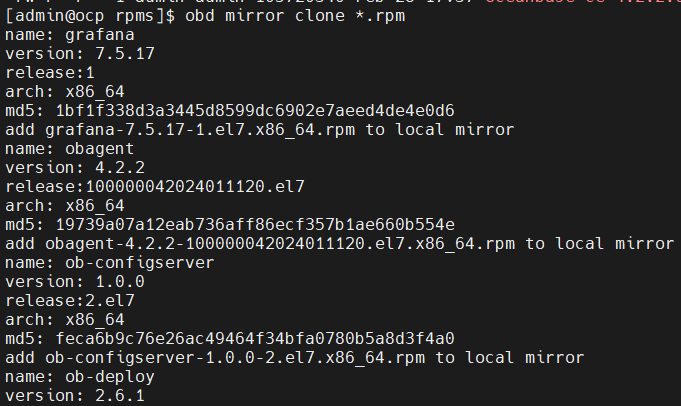
然后需要我们手工准备软件仓库，去官网下载部署所需要的各个组件的rpm安装包。

<https://www.oceanbase.com/softwarecenter>



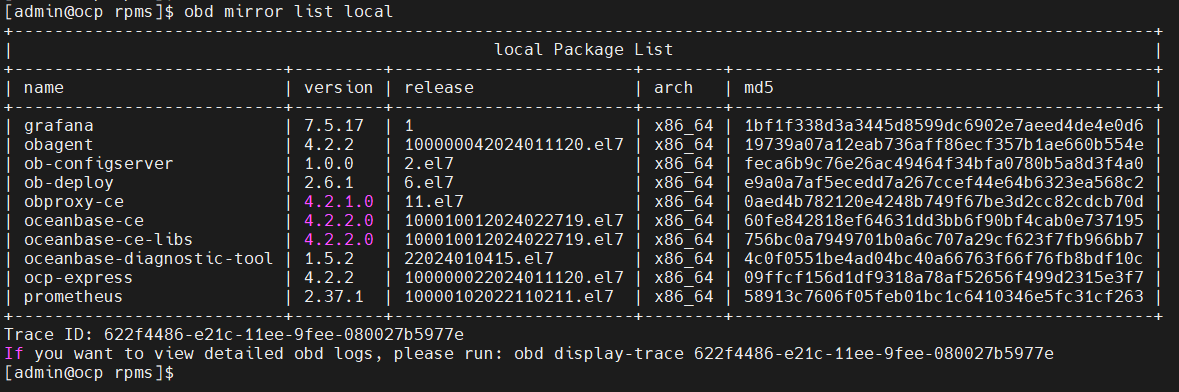
因为我之前用白屏（all-in-one）部署过，所以这里直接使用all-in-one里面现成的rpm包，进入到rpm包存放目录，使用如下命令创建本地仓库：

$ obd mirror clone \*.rpm



查看本地仓库：

$ obd mirror list local



# 四、选择并编辑集群配置文件

对于配置文件的选择，官方文档里有说明：



我这里选择的是mini-distributed-with-obproxy-example.yaml，并加入了obagent和ocp-express，这里就不需要Prometheus 和 Grafana了。

接下来对其进行编辑，主要内容如下：

注意：appname决定了集群名称

user:

username: admin

password: admin

oceanbase-ce:

servers:

- name: server1

ip: 192.168.56.57

- name: server2

ip: 192.168.56.58

- name: server3

ip: 192.168.56.59

global:

devname: enp0s3

cluster\_id: 1

memory\_limit: 6G

system\_memory: 1G

datafile\_size: 1G

datafile\_next: 2G

datafile\_maxsize: 20G

log\_disk\_size: 14G

cpu\_count: 4

production\_mode: false

enable\_syslog\_wf: false

enable\_syslog\_recycle: true

max\_syslog\_file\_count: 4

appname: my\_ob\_cls

server1:

mysql\_port: 2881

rpc\_port: 2882

home\_path: /home/admin/observer

data\_dir: /data

zone: zone1

server2:

mysql\_port: 2881

rpc\_port: 2882

home\_path: /home/admin/observer

data\_dir: /data

zone: zone2

server3:

mysql\_port: 2881

rpc\_port: 2882

home\_path: /home/admin/observer

data\_dir: /data

zone: zone3

obproxy-ce:

depends:

- oceanbase-ce

servers:

- 192.168.56.56

global:

listen\_port: 2883

prometheus\_listen\_port: 2884

home\_path: /home/admin/obproxy

enable\_cluster\_checkout: false

cluster\_name: my\_ob\_cls

skip\_proxy\_sys\_private\_check: true

enable\_strict\_kernel\_release: false

obagent:

depends:

- oceanbase-ce

servers:

- name: server1

ip: 192.168.56.57

- name: server2

ip: 192.168.56.58

- name: server3

ip: 192.168.56.59

global:

home\_path: /home/admin/obagent

ocp-express:

depends:

- oceanbase-ce

- obproxy-ce

- obagent

servers:

- name: server1

ip: 192.168.56.56

global:

home\_path: /home/admin/ocp-server

memory\_size: 1G

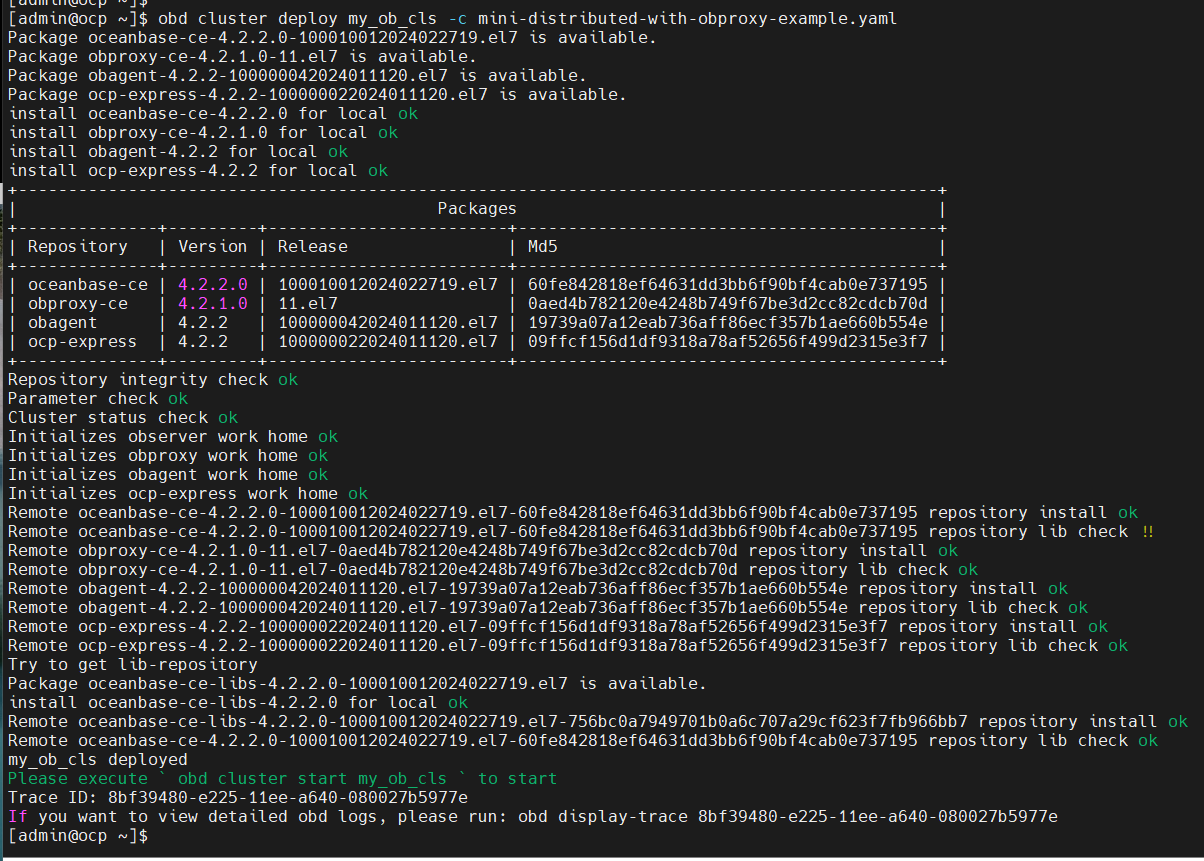
admin\_passwd: OCEANbase@#123

logging\_file\_total\_size\_cap: 10GB

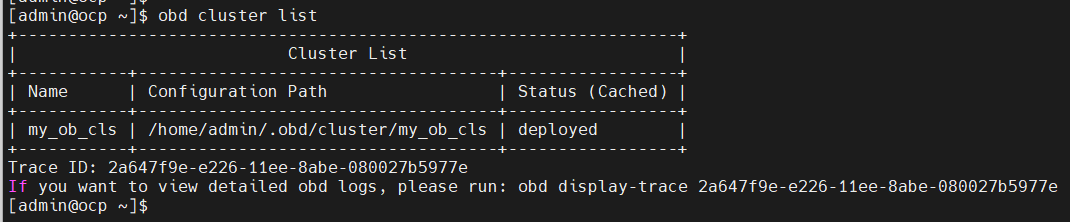
# 五、部署OceanBase集群

部署OceanBase集群

$ obd cluster deploy my\_ob\_cls -c mini-distributed-with-obproxy-example.yaml

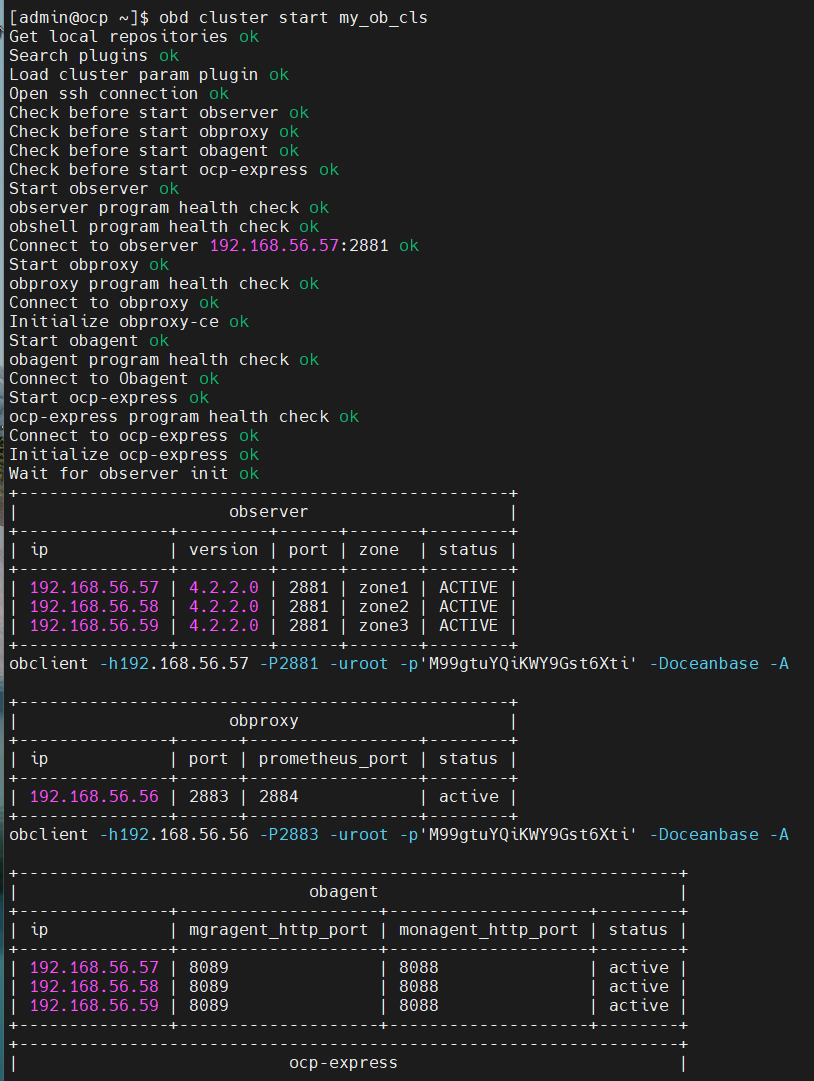


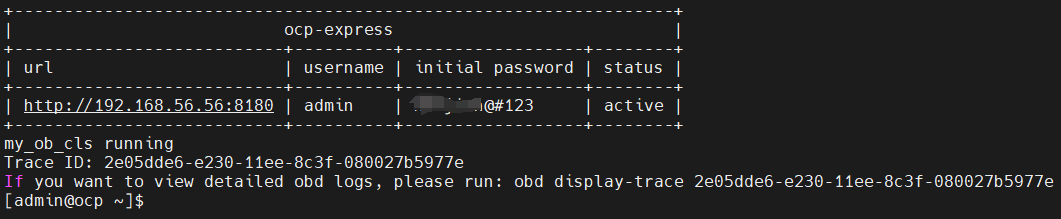
这里部署完毕，查询一下：



根据部署提示启动集群：

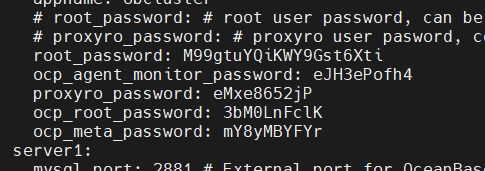
$ obd cluster start my\_ob\_cls





通过上面可以看到自动生成的root密码，使用如下命令可以看到集群配置以及自动生成的各项密码：

$ obd cluster edit-config my\_ob\_cls



# 六、连接数据库进行操作

## 5.1、安装obclient并连接数据库

官网下载rpm包上传到服务器安装：

<https://www.oceanbase.com/softwarecenter>

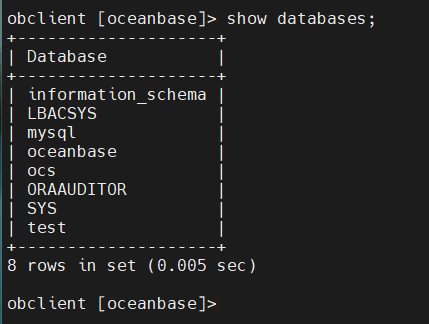


# rpm -ivh obclient-2.2.3-1.el7.x86\_64.rpm

以下两种连接方式：

$ obclient -h192.168.56.57 -P2881 -uroot -p'M99gtuYQiKWY9Gst6Xti' -Doceanbase -A

$ obclient -h192.168.56.56 -P2883 -uroot -p'M99gtuYQiKWY9Gst6Xti' -Doceanbase -A



## 5.2、创建resource unit

obclient [oceanbase]> CREATE RESOURCE UNIT u\_s MAX\_CPU 1, MEMORY\_SIZE '2G', MAX\_IOPS 1280,LOG\_DISK\_SIZE '2G', MIN\_IOPS=1024, IOPS\_WEIGHT=1;

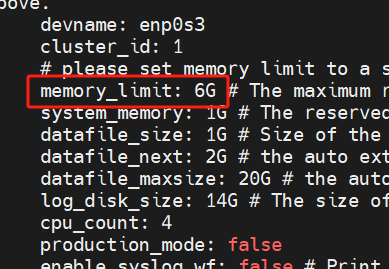
删除resource unit命令：

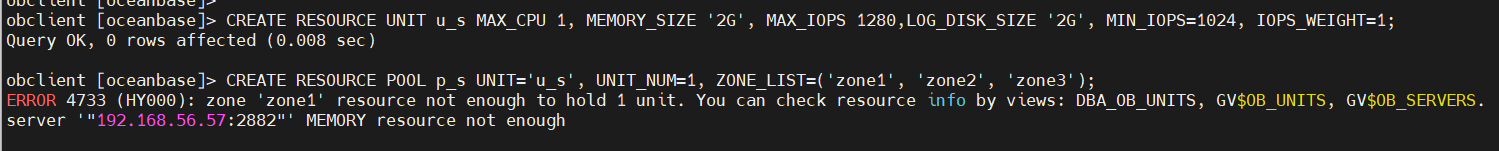
drop resource unit u\_s;

## 5.3、创建resource\_pool

obclient [oceanbase]> CREATE RESOURCE POOL p\_s UNIT='u\_s', UNIT\_NUM=1, ZONE\_LIST=('zone1', 'zone2', 'zone3');

如果是8G内存，集群edit\_config里面memory\_limit配置6G的话资源不够导致resource\_pool创建失败。





## 5.4、集群资源扩容

关闭集群：

$ obd cluster stop my\_ob\_cls

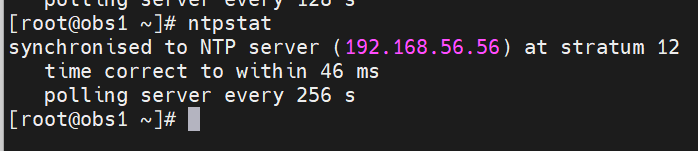
关闭虚拟机调整虚拟机内存和cpu配置

# init 0

调整完以后启动虚拟机

注意检查是否时间正常ntp同步，如果不同步会导致集群无法启动 ：

#ntpstat



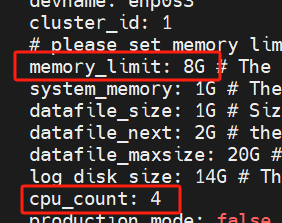
启动集群：

$ obd cluster start my\_ob\_cls

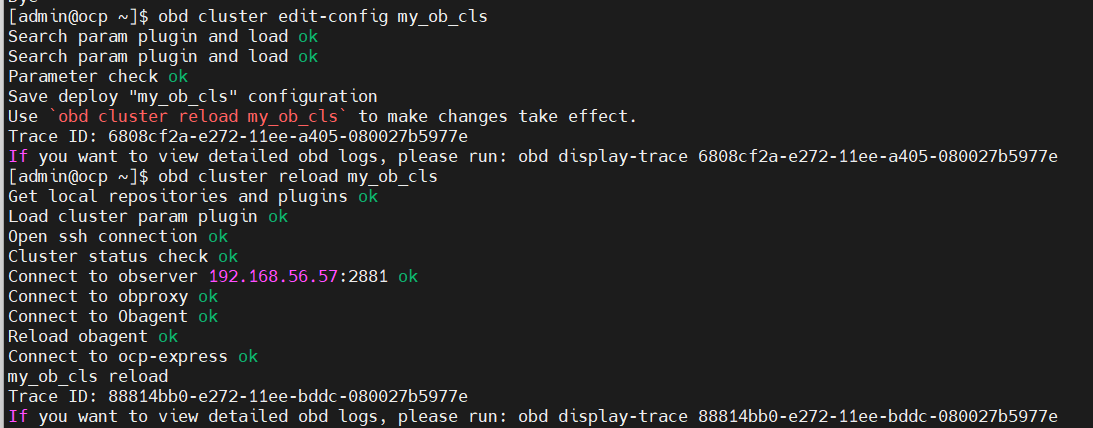
正常启动好集群以后，修改配置文件

$ obd cluster edit-config my\_ob\_cls

调整需要扩容的地方



保存退出，会提示如何生效新的配置文件，按照提示生效配置文件：



重新登录数据库

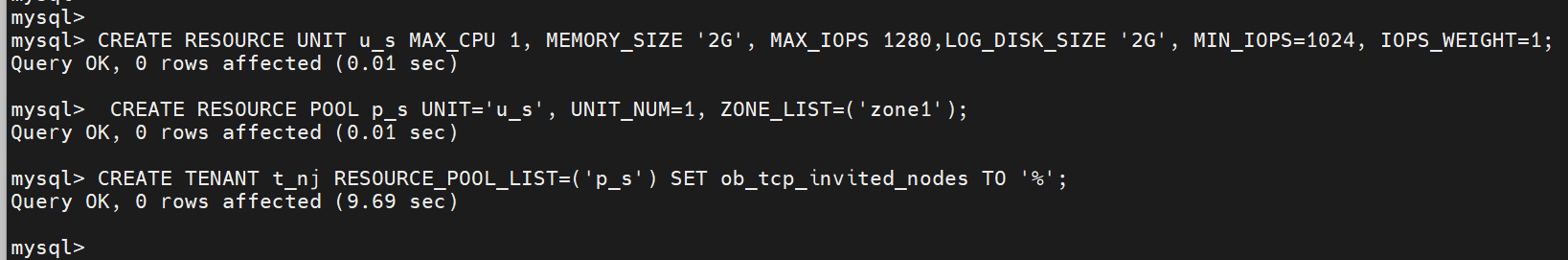
$ obclient -h192.168.56.57 -P2881 -uroot -p'M99gtuYQiKWY9Gst6Xti' -Doceanbase -A

创建resource pool:

obclient [oceanbase]> CREATE RESOURCE POOL p\_s UNIT='u\_s', UNIT\_NUM=1, ZONE\_LIST=('zone1', 'zone2', 'zone3');

## 5.5、创建租户并允许所有客户端IP连接数据库

obclient [oceanbase]> CREATE TENANT t\_nj RESOURCE\_POOL\_LIST=('p\_s') SET ob\_tcp\_invited\_nodes TO '%';



删除租户命令

drop tenant t\_nj;

## 5.6、使用新建的租户登录ob，默认密码为空

$ obclient -h192.168.56.57 -P2881 -uroot@t\_nj -A

执行建库，建表，建索引，插入数据等操作：

obclient [(none)]> create database njdb;

obclient [(none)]> create table t1(id int,name varchar(40));

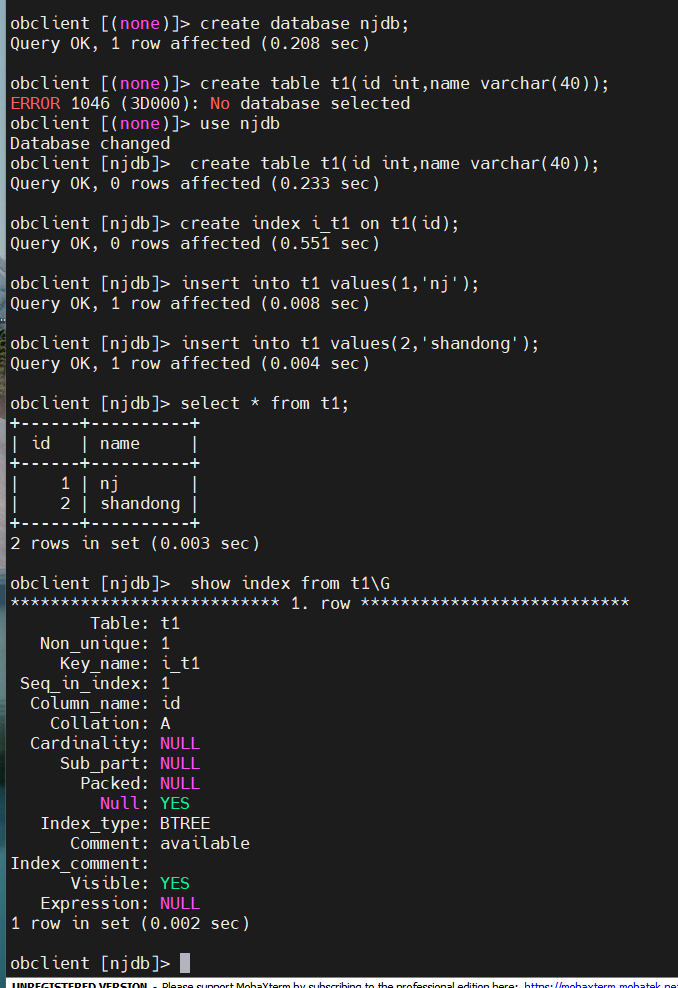
obclient [njdb]> create index i\_t1 on t1(id);

obclient [njdb]> insert into t1 values(1,'nj');

obclient [njdb]> insert into t1 values(2,'shandong');

obclient [njdb]> select \* from t1;

obclient [njdb]> show index from t1\G



# 七、总结

这里我遇到一个踩坑点，一开始没注意NetworkManager，导致网络冲突，初始化数据库的时候一直卡在这，observer/log/rootservice.log里面一直刷错误日志：

[errcode=-4038] monitor root service failed(ret=-4038, ret=“OB\_NOT\_MASTER”)

关闭NetworkManager以后就好了。

