



OceanBase 0.4.2

配置指南

文档版本：Beta 02

发布日期：2013.11.30

支付宝（中国）网络技术有限公司·OceanBase 团队

前言

概述

本文档主要介绍搭建OceanBase 0.4.2 主备RootServer、主备UpdateServer和主备集群的方法。

读者对象

本文档主要适用于：

- 安装工程师。
- 数据库管理工程师。

通用约定

在本文档中可能出现下列各式，它们所代表的含义如下。

| 格式 | 说明 |
|-----------------|--|
| 警告 | 表示可能导致设备损坏、数据丢失或不可预知的结果。 |
| 注意 | 表示可能导致设备性能降低、服务不可用。 |
| 小窍门 | 可以帮助您解决某个问题或节省您的时间。 |
| 说明 | 表示正文的附加信息，是对正文的强调和补充。 |
| 宋体 | 表示正文。 |
| 粗体 | 表示命令行中的关键字（命令中保持不变、必须照输的部分）或者正文中强调的内容。 |
| 斜体 | 用于变量输入。 |
| { a b ... } | 表示从两个或多个选项中选取一个。 |
| [] | 表示用“[]”括起来的部分在命令配置时是可选的。 |

修订记录

修改记录累积了每次文档更新的说明。最新版本的文档包含以前所有文档版本。

| 版本和发布日期 | 说明 |
|---------------------|---------------------------------|
| Beta 02（2013-11-30） | 第一次发布Beta版本，适用于OceanBase 0.4.2。 |
| 01（2013-10-30） | 第一次正式发布，适用于OceanBase 0.4.1。 |

联系我们

如果您有任何疑问或是想了解 OceanBase 的最新开源动态消息，请联系我们：

支付宝（中国）网络技术有限公司·OceanBase 团队

地址：杭州市万塘路 18 号黄龙时代广场 B 座；邮编：310099

北京市朝阳区东三环中路 1 号环球金融中心西塔 14 层；邮编：100020

邮箱：alipay-oceanbase-support@list.alibaba-inc.com

新浪微博：<http://weibo.com/u/2356115944>

技术交流群（阿里旺旺）：853923637

目 录

| | |
|----------------------------------|--------|
| 1 配置主备 RootServer..... | - 1 - |
| 1.1 基本信息 | - 1 - |
| 1.2 配置免登录 | - 1 - |
| 1.3 安装 Pacemaker 及其依赖包 | - 2 - |
| 1.4 配置并启动 corosync..... | - 4 - |
| 1.4.1 修改“corosync.conf”文件 | - 4 - |
| 1.4.2 新增“pcmk”和“authkey”文件 | - 6 - |
| 1.4.3 启动 corosync | - 6 - |
| 1.5 配置资源 | - 6 - |
| 1.5.1 拷贝监控脚本 | - 6 - |
| 1.5.2 配置资源文件 | - 7 - |
| 1.5.3 启动程序 | - 9 - |
| 1.6 启动和验证 OceanBase..... | - 12 - |
| 1.7 常用操作 | - 16 - |
| 2 配置主备 UpdateServer | - 18 - |
| 3 配置主备集群 | - 19 - |
| 3.1 基本信息 | - 19 - |
| 3.2 采用 RPM 包安装时的集群部署方法..... | - 19 - |
| 3.3 采用源码安装时的集群部署方法..... | - 22 - |

1 配置主备 RootServer

配置主备 RootServer 双机前，请先参考《OceanBase 0.4.2 安装指南》的“1 安装前须知”至“4 采用源码安装”，在各服务器上正确安装 OceanBase 软件。

1.1 基本信息

假设主备 RootServer 的服务器信息如[表 1-1](#)所示。

表 1-1 RootServer 信息

| 类型 | IP | Virtual IP | 主机名 |
|--------------|------------|--------------|--------------------|
| 主 RootServer | 10.10.10.2 | 10.10.10.254 | obtest2.alipay.net |
| 备 RootServer | 10.10.10.3 | 10.10.10.254 | obtest3.alipay.net |

注意：RootServer 的 VIP（Virtual IP）地址和本机 IP 地址必须在同一个网段，可根据“1.2.1 配置 VIP”进行配置。

1.2 配置免登录

主备 RootServer 之间打通 SSH 免登录，需要注意以下几点：

- 主备 RootServer 双向进行免登录。
- RootServer 本机对本机可以进行免登录。

主备 RootServer 之间打通 SSH 免登录操作步骤如下：

1. 以 **admin** 用户登录主 RootServer 服务器（10.10.10.2）。
2. 执行以下命令，并按“Enter”键，直至生成公钥。
ssh-keygen -t rsa
3. 执行以下命令，并根据提示输入登录密码，配置免登录。
ssh-copy-id admin@10.10.10.2
ssh-copy-id admin@10.10.10.3
4. 参考“步骤 1”至“步骤 3”，配置备 RootServer 服务器免登录。

配置完成后，分别在主备 RootServer 上执行以下命令，均可以免登录：

- **ssh admin@10.10.10.2**

- **ssh admin@10.10.10.3**

1.3 安装 Pacemaker 及其依赖包

阿里内部可直接使用“**sudo yum install pacemaker.x86_64**”命令，进行安装。

手动安装 Pacemaker 及其依赖包操作步骤如下：

1. 以 **root** 用户登录主 RootServer 服务器（10.10.10.2）。
2. 执行以下命令，获取 Pacemaker 及其依赖包。
wget
https://github.com/alibaba/oceanbase/blob/oceanbase_0.4/PacemakerAndDepLib/PacemakerAndDepLib.tar?raw=true
3. 执行以下命令，解压缩“PacemakerAndDepLib.tar”。
tar -xvf PacemakerAndDepLib.tar
4. 执行以下命令，进入“PacemakerAndDepLib”目录。
cd ~/PacemakerAndDepLib
5. 根据表 1-2，按序号进行安装 Pacemaker 及其依赖包。如果在安装过程中还需其他依赖，请根据提示信息自行获取安装。

表 1-2 安装 Pacemaker 以及其依赖包

| 序号 | 安装 RPM 包 |
|----|--|
| 1 | rpm -ivh libtalloc-2.0.1-1.1.el6.x86_64.rpm |
| 2 | rpm -ivh keyutils-1.4-3.el6.x86_64.rpm |
| 3 | rpm -ivh cifs-utils-4.8.1-5.el6.x86_64.rpm |
| 4 | rpm -ivh libtool-ltdl-2.2.6-15.5.el6.x86_64.rpm |
| 5 | rpm -ivh libtool-ltdl-devel-2.2.6-15.5.el6.x86_64.rpm |
| 6 | rpm -ivh cluster-glue-libs-1.0.5-2.el6.x86_64.rpm |
| 7 | rpm -ivh libtasn1-2.3-3.el6.x86_64.rpm |
| 8 | rpm -ivh gnutls-2.8.5-4.el6.x86_64.rpm |
| 9 | rpm -ivh libxslt-1.1.26-2.el6.x86_64.rpm |

| 序号 | 安装 RPM 包 |
|----|---|
| 10 | <code>rpm -ivh pacemaker-libs-1.1.6-3.el6.x86_64.rpm</code> |
| 11 | <code>rpm -ivh libibverbs-1.1.5-3.el6.x86_64.rpm</code> |
| 12 | <code>rpm -ivh librdmacm-1.0.14.1-3.el6.x86_64.rpm</code> |
| 13 | <code>rpm -ivh</code> <code>perl-Module-Pluggable-3.90-119.el6_1.1.x86_64.rpm</code> <code>perl-Pod-Escapes-1.04-119.el6_1.1.x86_64.rpm</code> <code>perl-Pod-Simple-3.13-119.el6_1.1.x86_64.rpm</code> <code>perl-version-0.77-119.el6_1.1.x86_64.rpm</code> |
| 14 | <code>rpm -ivh lm_sensors-libs-3.1.1-10.el6.x86_64.rpm</code> |
| 15 | <code>rpm -ivh lm_sensors-3.1.1-10.el6.x86_64.rpm</code> |
| 16 | <code>rpm -ivh net-snmp-libs-5.5-37.el6.x86_64.rpm</code> |
| 17 | <code>rpm -ivh net-snmp-5.5-37.el6.x86_64.rpm</code> |
| 18 | <code>rpm -ivh corosync-1.4.1-4.el6.x86_64.rpm</code> <code>corosynclib-1.4.1-4.el6.x86_64.rpm</code> |
| 19 | <code>rpm -ivh perl-TimeDate-1.16-11.1.el6.noarch.rpm</code> |
| 20 | <code>rpm -ivh clusterlib-3.0.12.1-23.el6.x86_64.rpm</code> |
| 21 | <code>rpm -ivh pacemaker-cluster-libs-1.1.6-3.el6.x86_64.rpm</code> |
| 22 | <code>rpm -ivh pacemaker-cli-1.1.6-3.el6.x86_64.rpm</code> |
| 23 | <code>rpm -ivh cluster-glue-1.0.5-2.el6.x86_64.rpm</code> |
| 24 | <code>rpm -ivh tcp_wrappers-7.6-57.el6.x86_64.rpm</code> |
| 25 | <code>rpm -ivh quota-3.17-16.el6.x86_64.rpm</code> |

| 序号 | 安装 RPM 包 |
|----|--|
| 26 | rpm -ivh libgssglue-0.1-11.el6.x86_64.rpm |
| 27 | rpm -ivh libtirpc-0.2.1-5.el6.x86_64.rpm |
| 28 | rpm -ivh rpcbind-0.2.0-8.el6.x86_64.rpm |
| 29 | rpm -ivh libevent-1.4.13-1.el6.x86_64.rpm |
| 30 | rpm -ivh nfs-utils-lib-1.1.5-4.el6.x86_64.rpm nfs-utils-1.2.3-15.el6.x86_64.rpm |
| 31 | rpm -ivh resource-agents-3.9.2-7.el6.x86_64.rpm |
| 32 | rpm -ivh pacemaker-1.1.6-3.el6.x86_64.rpm |

6. 参考“步骤 1”至“步骤 5”，在备 RootServer 服务器上安装 Pacemaker。

1.4 配置并启动 corosync

主要介绍配置和启动 corosync 的方法，包括：修改“corosync.conf”文件、新增“pcmk”和“authkey”文件和启动 corosync。

1.4.1 修改“corosync.conf”文件

“corosync.conf”的配置除 IP 外，其余参数配置请与步骤中的举例保持一致。

1. 以 **root** 用户登录主 RootServer 服务器（10.10.10.2）。
2. 执行以下命令，进入“/etc/corosync”目录。
cd /etc/corosync
3. 执行以下命令，复制“corosync.conf.example”文件。
cp corosync.conf.example corosync.conf
4. 使用 **vi** 编辑器，按照以下举例修改“corosync.conf”文件。其中两个“member”分别为主备 RootServer 的 IP，“bindnetaddr”为本机 IP。

```
# Please read the corosync.conf.5 manual page
compatibility: whitetank

totem {
    version: 2
```



```

    secauth: off
    threads: 0
    netmtu: 1300
    token: 3000
    interface {
        ringnumber: 0
        member {
            memberaddr: 10.10.10.2
        }
        member {
            memberaddr: 10.10.10.3
        }
        bindnetaddr: 10.10.10.2
        mcastport: 5543
        ttl: 1
    }
    transport: udpu
}

logging {
    fileline: off
    to_stderr: no
    to_logfile: yes
    to_syslog: yes
    logfile: /var/log/cluster/corosync.log
    debug: on
    timestamp: on
    logger_subsys {
        subsys: AMF
        debug: off
    }
}

amf {
    mode: disabled
}

aisexec {
    user: root
    group: root
}

```

5. 参考“步骤 1”至“步骤 4”，配置备 RootServer 服务器的“corosync.conf”文件。

1.4.2 新增“pcmk”和“authkey”文件

“pcmk”和“authkey”文件在主备 RootServer 中必须相同，因此我们可以在主 RootServer 服务器中创建“pcmk”和生成“authkey”文件，然后再拷贝到备 RootServer 服务器。

1. 以 **root** 用户登录主 RootServer 服务器（10.10.10.2）。
2. 使用 **vi** 编辑器，新建“/etc/corosync/service.d/pcmk”文件。

```
service {  
    ver: 0  
    name: pacemaker  
}
```

3. 将“/etc/corosync/service.d/pcmk”文件拷贝到备 RootServer 服务器。
scp /etc/corosync/service.d/pcmk root@10.10.10.3:/etc/corosync/service.d
4. 执行以下命令，生成“/etc/corosync/authkey”文件。
/usr/sbin/corosync-keygen
5. 将“/etc/corosync/authkey”文件拷贝到备 RootServer 服务器。
scp /etc/corosync/authkey root@10.10.10.3:/etc/corosync

1.4.3 启动 corosync

在主备 RootServer 服务器上分别启动 corosync。

1. 以 **root** 用户登录主 RootServer 服务器（10.10.10.2）。
2. 执行以下命令，启动 corosync。
service corosync start
3. 执行以下命令，查看 corosync 是否启动成功。
ps aux | grep corosync
4. 参考“步骤 1”至“步骤 3”，启动备 RootServer 服务器的 corosync。

1.5 配置资源

配置资源包括：拷贝监控脚本、配置资源文件和启动程序。

1.5.1 拷贝监控脚本

将 RootServer 的监控脚本拷贝到“/usr/lib/ocf/resource.d/heartbeat”目录中。监控脚本为“rootserver”文件，可从 OceanBase 源代码中的“/script/ha”目录下获取。

1. 以 **root** 用户登录主 RootServer 服务器（10.10.10.2）。
2. 执行以下命令，拷贝 rootserver 的监控脚本到
“/usr/lib/ocf/resource.d/heartbeat”目录。
**cp /home/admin/oceanbase_install/script/ha/RootServer
/usr/lib/ocf/resource.d/heartbeat/**
3. 以 **root** 用户登录备 RootServer 服务器。
4. 执行以下命令，拷贝 rootserver 的监控脚本到
“/usr/lib/ocf/resource.d/heartbeat”目录。
**cp /home/admin/oceanbase_install/script/ha/RootServer
/usr/lib/ocf/resource.d/heartbeat/**

1.5.2 配置资源文件

资源文件为“rs.xml”文件，可从 OceanBase 源代码中的“/script/ha”目录下获取。

1. 以 **root** 用户登录主 RootServer 服务器（10.10.10.2）。
2. 使用 **vi** 编辑器，修改“/home/admin/oceanbase_install/script/ha/rs.xml”文件，如黑体部分所示，配置说明请参见[表 1-3](#)，其余参数请保持缺省值。

```
<resources>
  <group id="rootserver-group">
    <primitive class="ocf" provider="heartbeat" type="IPAddr2" id="ip-alias-rs">
      <instance_attributes id="ip-alias-rs-instance_attributes">
        <nvpair name="ip" id="ip-alias-rs-instance_attributes-ip"
value="10.10.10.254"/>
        <nvpair name="nic" id="ip-alias-rs-instance_attributes-nic"
value="eth0:0"/>
      </instance_attributes>
      <operations>
        <op id="ip-alias-rs-monitor-2s" interval="2s" name="monitor"/>
      </operations>
      <meta_attributes id="ip-alias-rs-meta_attributes">
        <nvpair id="ip-alias-rs-meta_attributes-target-role" name="target-role"
value="Started"/>
      </meta_attributes>
    </primitive>
    <primitive class="ocf" id="ob-rootserver" provider="heartbeat"
type="RootServer">
      <instance_attributes id="ob-rootserver-instance_attributes">
        <nvpair id="ob-rootserver-instance_attributes-basedir" name="basedir"
value="/home/admin/oceanbase/"/>
        <nvpair id="ob-rootserver-instance_attributes-pidfile" name="pidfile"
```

```

value="/home/admin/oceanbase/run/rootserver.pid"/>
    <nvpair id="ob-rootserver-instance_attributes-rsip" name="rsip"
value="localhost"/>
    <nvpair id="ob-rootserver-instance_attributes-rsport" name="rsport"
value="2500"/>
    <nvpair id="ob-rootserver-instance_attributes-montool"
name="montool" value="/home/admin/oceanbase/bin/ob_ping"/>
    <nvpair id="ob-rootserver-instance_attributes-user" name="user"
value="admin"/>
  </instance_attributes>
  <operations>
    <op id="ob-rootserver-monitor-2s" interval="2s" name="monitor"/>
    <op id="ob-rootserver-start" interval="0s" name="start"
timeout="600s"/>
    <op id="ob-rootserver-stop" interval="0s" name="stop" timeout="30s"/>
  </operations>
  <meta_attributes id="ob-rootserver-meta_attributes">
    <nvpair id="ob-rootserver-meta_attributes-target-role"
name="target-role" value="Started"/>
    <nvpair id="ob-rootserver-meta_attributes-resource-stickiness"
name="resource-stickiness" value="INFINITY"/>
    <nvpair id="ob-rootserver-meta_attributes-resource-failure-stickiness"
name="resource-failure-stickiness" value="-INFINITY"/>
  </meta_attributes>
</primitive>
</group>
</resources>

```

表 1-3 参数说明

| 参数 | 说明 |
|---|------------------------------|
| <pre> <nvpair name="ip" id="ip-alias-rs-instance_attributes-ip" value="10.10.10.254"/> </pre> | RootServer 的 VIP。 |
| <pre> <nvpair name="nic" id="ip-alias-rs-instance_attributes-nic" value="eth0:0"/> </pre> | 绑定 RootServer 的 VIP 的虚拟网卡名称。 |

| 参数 | 说明 |
|--|-----------------------|
| <pre><nvpair id="ob-rootserver-instance_attributes-basedir" name="basedir" value="/home/admin/oceanbase/" /></pre> | OceanBase 的安装目录。 |
| <pre><nvpair id="ob-rootserver-instance_attributes-pidfile" name="pidfile" value="/home/admin/oceanbase/run/rootserver.pid" /></pre> | “rootserver.pid”文件路径。 |
| <pre><nvpair id="ob-rootserver-instance_attributes-rsreport" name="rsreport" value="2500" /></pre> | RootServer 的服务端口。 |
| <pre><nvpair id="ob-rootserver-instance_attributes-montool" name="montool" value="/home/admin/oceanbase/bin/ob_ping" /></pre> | “ob_ping”文件路径。 |

3. 使用 **vi** 编辑器，在“/home/admin”目录下创建“constraints.xml”文件，内容如下所示。

```
<constraints>
  </constraints>
```

1.5.3 启动程序

启动程序，实现主备 RootServer 服务器的 VIP 漂移。

1. 以 **root** 用户分别登录主备 RootServer 服务器(10.10.10.2 和 10.10.10.3)。
2. 根据[表 1-4](#)，按序号启动程序。

表 1-4 启动程序

| 序号 | 方式 | 命令 |
|----|----------------------------------|--|
| 1 | 在主 RootServer 服务器中，右侧列的命令分别执行两次。 | <pre>sudo crm_attribute --type crm_config --attr-name symmetric-cluster --attr-value true sudo crm_attribute --type crm_config --attr-name stonith-enabled --attr-value false sudo crm_attribute --type rsc_defaults --name resource-stickiness --update 100</pre> |
| 2 | 在备 RootServer 服务器中，右侧列的命令分别执行两次。 | <pre>sudo crm_attribute --type crm_config --attr-name symmetric-cluster --attr-value true sudo crm_attribute --type crm_config --attr-name stonith-enabled --attr-value false sudo crm_attribute --type rsc_defaults --name resource-stickiness --update 100</pre> |

| 序号 | 方式 | 命令 |
|----|---|--|
| 3 | <p>在主 RootServer 服务器中，右侧列的命令执行一次。</p> <p><i>说明： 请注意修改命令中“rs.xml”和“constraints.xml”文件所在路径。</i></p> | <pre> set no-quorum-policy = ignore sudo crm_attribute --type crm_config --attr-name no-quorum-policy --attr-value ignore sudo cibadmin --query --obj_type=resources sudo cibadmin --replace --obj_type=resources --xml-file /home/admin/oceanbase_install/script/ha/rs.xml sudo cibadmin --query --scope=constraints sudo cibadmin --replace --obj_type=constraints --xml-file /home/admin/constraints.xml sudo cibadmin --query --scope=constraints </pre> |

3. 执行 **crm status** 命令，查看是否成功启动程序，并记录 VIP 所在节点。系统显示如下，则说明启动成功。

```

=====
Last updated: Fri Oct 18 10:49:29 2013
Last change: Fri Oct 18 10:48:31 2013 via cibadmin on obtest2.alipay.net
Stack: openais
Current DC: obtest3.alipay.net - partition with quorum
Version: 1.1.6-3.el6-a02c0f19a00c1eb2527ad38f146ebc0834814558
2 Nodes configured, 2 expected votes
2 Resources configured.
=====

Online: [ obtest2.alipay.net obtest3.alipay.net ]

Resource Group: rootserver-group
    ip-alias-rs      (ocf::heartbeat:IPAddr2):      Started obtest2.alipay.net
    ob-rootserver    (ocf::heartbeat:RootServer):    Stopped

```

```
Failed actions:
  ob-rootserver_start_0 (node=obtest2.alipay.net, call=6, rc=1, status=complete):
unknown error
  ob-rootserver_start_0 (node=obtest3.alipay.net, call=7, rc=1, status=complete):
unknown error
```

1.6 启动和验证 OceanBase

启动 OceanBase，并验证主 RootServer 故障时，可切换到备 RootServer。

* 启动 OceanBase

启动 OceanBase 并设置“10.10.10.2”服务器为主 RootServer。

注意：启动 ChunkServer 前请先启动 RootServer，否则 ChunkServer 在一段时间后会自动结束进程。

1. 以 **admin** 用户分别登录各 OceanBase 服务器。
2. 执行以下命令，进入 OceanBase 安装目录。
cd /home/admin/oceanbase
3. 启动 RootServer、UpdateServer、ChunkServer 和 MergeServer。参数说明如[表 1-5](#)所示。

注意：启动 ChunkServer 前请先启动 RootServer，否则 ChunkServer 在一段时间后会自动结束进程。

- “10.10.10.2”中，启动主 RootServer/UpdateServer/Listener
bin/rootserver -r 10.10.10.254:2500 -R 10.10.10.254:2500 -i eth0 -C 1
bin/updateserver -r 10.10.10.254:2500 -p 2700 -m 2701 -i eth0
bin/mergeserver -r 10.10.10.254:2500 -p 2800 -z 2828 -i eth0 -t lms
- “10.10.10.3”中，启动备 RootServer
bin/rootserver -r 10.10.10.254:2500 -R 10.10.10.254:2500 -i eth0 -C 1
- 依次在“10.10.10.4”和“10.10.10.5”中，启动
ChunkServer/MergeServer
bin/chunkserver -r 10.10.10.254:2500 -p 2600 -n obtest -i eth0
bin/mergeserver -r 10.10.10.254:2500 -p 2800 -z 2880 -i eth0

表 1-5 常用操作

| 服务器 | 参数 | 说明 |
|--------------|----|---|
| RootServer | -r | 需要启动的 RootServer 的 IP 地址和服务端口。主备 RootServer 时，为需要启动的 RootServer 的 VIP 地址和服务端口。 格式：-r [IP]:[Port] |
| | -R | 主集群的 RootServer 的 IP 地址和端口。主备 RootServer 时，为主集群的 RootServer 的 VIP 地址和服务端口。 格式：-R [IP]:[Port] |
| | -i | 设置绑定的网卡。 格式：-i [NIC Name] |
| | -C | 设置集群 ID，必须为数字。 格式：-C [Cluster ID] |
| UpdateServer | -r | 所在集群的 RootServer 的 IP 地址和端口。主备 RootServer 时，为所在集群的 RootServer 的 VIP 地址和服务端口。 格式：-r [IP]:[Port] |
| | -p | 设置当前 UpdateServer 的服务端口。 格式：-p [Port] |
| | -m | 每日合并操作时，ChunkServer 请求合并数据所用的端口。 格式：-m [Port] |

| 服务器 | 参数 | 说明 |
|-------------|----|---|
| | -i | 设置绑定的网卡。 格式: -i [NIC Name] |
| ChunkServer | -r | 所在集群的 RootServer 的 IP 地址和端口。主备 RootServer 时, 为所在集群的 RootServer 的 VIP 地址和服务端口。 格式: -r [IP]:[Port] |
| | -p | 设置当前 ChunkServer 的服务端口。 格式: -p [Port] |
| | -n | APP 名称。与“4.6 创建各 Server 所需目录”中 sstable 的父目录名称保持一致。 格式: -n [APP Name] |
| | -i | 设置绑定的网卡。 格式: -i [NIC Name] |
| MergeServer | -r | 所在集群的 RootServer 的 IP 地址和端口。主备 RootServer 时, 为所在集群的 RootServer 的 VIP 地址和服务端口。 格式: -r [IP]:[Port] |
| | -p | 设置当前 MergeServer 的服务端口。 格式: -p [Port] |
| | -z | 设置 MergeServer 的 MySQL 的协议端口。 格式: -z [Port] |

| 服务器 | 参数 | 说明 |
|----------|----|---|
| | -i | 设置绑定的网卡。 格式: -i [NIC Name] |
| Listener | -r | 所在集群的 RootServer 的 IP 地址和端口。主备 RootServer 时, 为所在集群的 RootServer 的 VIP 地址和服务端口。 格式: -r [IP]:[Port] |
| | -p | 设置 Listener 的服务端口。 格式: -p [Port] |
| | -z | 设置 Listener 的 MySQL 的协议端口。缺省值为 2828, 不建议修改。 格式: -z [Port] |
| | -i | 设置绑定的网卡。 格式: -i [NIC Name] |
| | -t | 将该 MergeServer 进程指定为 Listener。 格式: -t lms |

- 以 **admin** 用户登录主 RootServer 所在的 OceanBase 服务器 (10.10.10.2)。

说明: 以下操作仅第一次启动 OceanBase 各 Server 后需要执行。

- 执行以下命令, 进入 OceanBase 安装目录。
cd /home/admin/oceanbase
- 依次执行以下命令, 初始化 OceanBase。
bin/rs_admin -r 10.10.10.254 -p 2500 set_obi_role -o OBI_MASTER
bin/rs_admin -r 10.10.10.254 -p 2500 -t 60000000 boot_strap
- 执行以下命令, 清空“constraints.xml”。
sudo cibadmin --replace --obj_type=constraints --xml-file /home/admin/constraints.xml

- 依次执行以下命令，设置“10.10.10.2”服务器为主 RootServer。
sudo crm resource failcount ob-rootserver set obtest2.alipay.net 0
sudo crm resource failcount ob-rootserver set obtest3.alipay.net 0
sudo crm resource failcount ob-rootserver set obtest2.alipay.net 1

*** 验证**

验证主备 RootServer 是否可以正常切换方法如下：

- 执行 **crm status** 命令，查看 VIP 在哪个服务器上，则该服务器为主 RootServer。
- 使用 **kill -9** 命令，结束主 RootServer 服务进程。
- 执行 **crm status** 命令，查看 VIP 所在节点是否迁移到备 RootServer 服务器。
- 重新启动原主 RootServer 服务，则该服务器为新备 RootServer，原备 RootServer 服务器为新主 RootServer。
- 使用 **kill -9** 命令，结束新主 RootServer 服务进程。
- 执行 **crm status** 命令，查看 VIP 所在节点是否迁移到新备 RootServer 服务器。

1.7 常用操作

其他常用操作如[表 1-6](#)所示。

表 1-6 常用操作

| 目的 | 说明 |
|--------|---|
| 查看主备状态 | 执行 sudo crm status 命令，查看主备状态。 |
| 查看节点状态 | 执行 sudo crm resource failcount ob-rootserver show obtest2.alipay.net 命令，查看“obtest2.alipay.net”服务器状态。 |

| 目的 | 说明 |
|----------------|--|
| 主备迁移 | <p>依次执行以下命令，设置 OceanBase146001.cm4 服务为主 RootServer。</p> <pre>sudo cibadmin --replace --obj_type=constraints --xml-file /home/admin/constraints.xml</pre> <pre>sudo crm resource failcount ob-rootserver set obtest2.alipay.net 0</pre> <pre>sudo crm resource failcount ob-rootserver set obtest3.alipay.net 0</pre> <pre>sudo crm resource failcount ob-rootserver set obtest3.alipay.net 1</pre> |
| 查看 corosync 日志 | <p>执行 sudo cat /var/log/cluster/corosync.log 命令，查看 corosync 日志。</p> |
| 查看 crm 命令帮助 | <ol style="list-style-type: none"> 1. 执行 sudo crm 命令，进入 crm 的管理台。 2. 执行 help 命令，查看帮助。 |

2 配置主备 UpdateServer

UpdateServer 主备由 RootServer 内部算法进行选举决定，因此无需进行特定设置。请直接根据《OceanBase 0.4.2 安装指南》进行安装、启动。

注意：在配置免登录时，需要配置主备 UpdateServer 间的双向免登录。

3 配置主备集群

配置 OceanBase 主备集群前，请先参考《OceanBase 0.4.2 安装指南》的“1 安装前须知”至“4 采用源码安装”，在各服务器上正确安装 OceanBase 软件。

3.1 基本信息

假设主备 OceanBase 主备集群各服务器的 IP，如[表 3-1](#)所示。

表 3-1 集群信息

| 集群 | RootServer/UpdateServer | ChunkServer/MergeServer |
|-----|-------------------------|--------------------------|
| 主集群 | 10.10.10.2 | 10.10.10.4, 10.10.10.5 |
| 备集群 | 10.10.10.12 | 10.10.10.14, 10.10.10.15 |

3.2 采用 RPM 包安装时的集群部署方法

如果您采用源码安装，则跳过本小节。

采用 RPM 包安装部署 OceanBase 主备集群时，配置一键脚本“deploy.conf”及初始化方法如下：

1. 以 **admin** 用户登录 OceanBase 服务器（10.10.10.2）。
2. 执行以下命令，进入一键脚本目录。
cd ~/oceanbase_install/Script
3. 执行以下命令，复制配置文件。
cp oceanbase.conf.template deploy.conf
4. 使用 **vi** 编辑器，修改配置文件“deploy.conf”，如黑体部分所示。参数说明见注释部分。

注意：所有以“#@”开始的行有特殊含义，不允许当注释删除。

```
#@begin_global [settings]
# rs_admin 工具的位置，请勿修改。
rs_admin=./bin/rs_admin
# OceanBase 的安装目录。
ob_home=/home/admin/oceanbase
[public]
# APP 名称。
```

```

appname=obtest
# 主集群 ID, 与集群名称对应, 即以#@begin_cluster_x 和#@end_cluster_x 开头的行。
# OceanBase 内部使用纯数字 ID, 即该配置中的数字部分为 ob 内部使用的集群 ID 号。
# 如果不指定集群 ID, 则默认使用数字最小的集群为主集群。
master_cluster_id=cluster_1
# 网络接口名称, 默认是 bond0。放到不同的 section 下可以单独为那个 section 中的 server 进行配置。
devname=eth0
[rootserver]
# RootServer 的服务端口。
port=2500
# RootServer 存放 commitlog 的目录。
# 执行脚本后, 会在“/home/admin/oceanbase/data”下创建“rs_commitlog”目录, 并软链接到“/data/log/rs_commitlog”。
commitlog_dir=/data/log/rs_commitlog
[chunkserver]
# ChunkServer 的端口。
port=2600
# ChunkServer 使用的磁盘数。
# 需要已经建立/data/{1..max_disk_num}的目录。
max_disk_num=8
[mergeserver]
# MergeServer 的服务端口。
port=2800
# MergeServer 的 MySQL 端口。
sql_port=2880
# 部署在 RootServer 上的 Listener 端口。
# 不建议修改!
lms_port=2828
[updateserver]
# UpdateServer 的服务端口。
port=2700
# UpdateServer 用于每日合并的端口。
inner_port=2701
# UpdateServer 转储用的磁盘的数目。
# 需要已经建立/data/{1..max_disk_num}的目录。
max_disk_num=8
# UpdateServer 存放 commitlog 的目录。
# 执行脚本后, 会在“/home/admin/oceanbase/data”下创建“ups_commitlog”目录, 并软链接到“/data/log/ups_commitlog”。
commitlog_dir=/data/log/ups_commitlog
#@end_global

#@begin_init_config

```



```

# 各 Server 启动时使用的配置项。
[rootserver]

[chunkserver]

[mergeserver]

[updateserver]
log_sync_type=1
#@end_init_config

#@begin_cluster_1
[public]
[rootserver]
# vip 为主备 RootServer 的虚拟 IP。
# 当 RootServer 单机时，请输入 RootServer 的实际 IP。
vip=10.10.10.2
# 主备 RootServer 的 IP 地址。
10.10.10.2
[updateserver]
# 主备 UpdateServer 的 IP 地址。
10.10.10.2
[chunkserver]
# 所有 ChunkServer 的 IP 地址。
10.10.10.4
10.10.10.5
[mergeserver]
# 所有 MergeServer 的 IP 地址。
10.10.10.4
10.10.10.5
#@end_cluster_1

#@begin_cluster_2
[public]
[rootserver]
# vip 为主备 RootServer 的虚拟 IP。
# 当 RootServer 单机时，请输入 RootServer 的实际 IP。
vip=10.10.10.12
# 主备 RootServer 的 IP 地址。
10.10.10.12
[updateserver]
# 主备 UpdateServer 的 IP 地址。
10.10.10.12
[chunkserver]

```

```
# 所有 ChunkServer 的 IP 地址。
10.10.10.14
10.10.10.15
[mergeserver]
# 所有 MergeServer 的 IP 地址。
10.10.10.14
10.10.10.15
#@end_cluster_2
```

5. 执行以下命令，一键启动及初始化。参数说明如[表 3-2](#)所示。
./oceanbase.pl init --force deploy.conf

表 3-2 参数说明

| 参数 | 说明 |
|--------------|---|
| oceanbase.pl | 运行脚本名称。 |
| init | 操作类型。初始化环境、启动并初始化集群。在首次安装时使用。 |
| --force | 强制执行，可省略。 |
| -c | “-c 1”表示只对 cluster_1 进行初始化。如果不指定集群 ID，则初始化配置文件中的所有集群。 |
| deploy.conf | 配置文件名称。 |

3.3 采用源码安装时的集群部署方法

如果您采用 RPM 安装，则跳过本小节。

OceanBase 主备集群部署时，各 Server 启动及初始化的操作步骤如下：

1. 以 **admin** 用户分别登录各 OceanBase 服务器。
2. 执行以下命令，进入 OceanBase 安装目录。
cd /home/admin/oceanbase
3. 启动 RootServer、UpdateServer、ChunkServer 和 MergeServer。参数说明如[表 3-3](#)所示。

注意：启动 ChunkServer 前请先启动 RootServer，否则 ChunkServer 在一段时间后会自动结束进程。

- “10.10.10.2”中，启动主集群的 RootServer/UpdateServer/Listener
bin/rootserver -r 10.10.10.2:2500 -R 10.10.10.2:2500 -i eth0 -C

1

bin/updateserver -r 10.10.10.2:2500 -p 2700 -m 2701 -i eth0
bin/mergeserver -r 10.10.10.2:2500 -p 2800 -z 2828 -i eth0 -t
lms

- 依次在“10.10.10.4”和“10.10.10.5”中，启动主集群的
 ChunkServer/MergeServer
bin/chunkserver -r 10.10.10.2:2500 -p 2600 -n obtest -i eth0
bin/mergeserver -r 10.10.10.2:2500 -p 2800 -z 2880 -i eth0
- “10.10.10.12”中，启动备集群的
 RootServer/UpdateServer/Listener
bin/rootserver -r 10.10.10.12:2500 -R 10.10.10.2:2500 -i eth0
-C 2
bin/updateserver -r 10.10.10.12:2500 -p 2700 -m 2701 -i eth0
bin/mergeserver -r 10.10.10.12:2500 -p 2800 -z 2828 -i eth0 -t
lms
- 依次在“10.10.10.14”和“10.10.10.15”中，启动备集群的
 ChunkServer/MergeServer
bin/chunkserver -r 10.10.10.12:2500 -p 2600 -n obtest -i eth0
bin/mergeserver -r 10.10.10.12:2500 -p 2800 -z 2880 -i eth0

说明： obtest 为《OceanBase 0.4.2 安装指南》的“4.6 创建各 Server 所需目录”中创建 sstable 所在的目录。

表 3-3 参数解释

| 服务器 | 参数 | 说明 |
|------------|----|---|
| RootServer | -r | 需要启动的 RootServer 的 IP 地址和服务端口。主备 RootServer 时，为需要启动的 RootServer 的 VIP 地址和服务端口。 格式：-r [IP]:[Port] |
| | -R | 主集群的 RootServer 的 IP 地址和端口。主备 RootServer 时，为主集群的 RootServer 的 VIP 地址和服务端口。 格式：-R [IP]:[Port] |
| | -i | 设置绑定的网卡。 格式：-i [NIC Name] |

| 服务器 | 参数 | 说明 |
|--------------|----|---|
| | -C | 设置集群 ID，必须为数字。 格式：-C [Cluster ID] |
| UpdateServer | -r | 所在集群的 RootServer 的 IP 地址和端口。主备 RootServer 时，为所在集群的 RootServer 的 VIP 地址和服务端口。 格式：-r [IP]:[Port] |
| | -p | 设置当前 UpdateServer 的服务端口。 格式：-p [Port] |
| | -m | 每日合并操作时，ChunkServer 请求合并数据所用的端口。 格式：-m [Port] |
| | -i | 设置绑定的网卡。 格式：-i [NIC Name] |
| ChunkServer | -r | 所在集群的 RootServer 的 IP 地址和端口。主备 RootServer 时，为所在集群的 RootServer 的 VIP 地址和服务端口。 格式：-r [IP]:[Port] |
| | -p | 设置当前 ChunkServer 的服务端口。 格式：-p [Port] |
| | -n | APP 名称。与“4.6 创建各 Server 所需目录”中 sstable 的父目录名称保持一致。 格式：-n [APP Name] |

| 服务器 | 参数 | 说明 |
|-------------|----|---|
| | -i | 设置绑定的网卡。 格式: -i [NIC Name] |
| MergeServer | -r | 所在集群的 RootServer 的 IP 地址和端口。主备 RootServer 时, 为所在集群的 RootServer 的 VIP 地址和服务端口。 格式: -r [IP]:[Port] |
| | -p | 设置当前 MergeServer 的服务端口。 格式: -p [Port] |
| | -z | 设置 MergeServer 的 MySQL 的协议端口。 格式: -z [Port] |
| | -i | 设置绑定的网卡。 格式: -i [NIC Name] |
| Listener | -r | 所在集群的 RootServer 的 IP 地址和端口。主备 RootServer 时, 为所在集群的 RootServer 的 VIP 地址和服务端口。 格式: -r [IP]:[Port] |
| | -p | 设置 Listener 的服务端口。 格式: -p [Port] |
| | -z | 设置 Listener 的 MySQL 的协议端口。缺省值为 2828, 不建议修改。 格式: -z [Port] |

| 服务器 | 参数 | 说明 |
|-----|----|--|
| | -i | 设置绑定的网卡。 格式: -i [NIC Name] |
| | -t | 将该 MergeServer 进程指定为 Listener。 格式: -t lms |

4. 以 **admin** 用户登录主集群的主 RootServer 服务器（10.10.10.2）。
5. 执行以下命令，进入 OceanBase 安装目录。
cd /home/admin/oceanbase
6. 依次执行以下命令，初始化 OceanBase，参数说明如[表 3-4](#)所示。
bin/rs_admin -r 10.10.10.2 -p 2500 set_obi_role -o OBI_MASTER
bin/rs_admin -r 10.10.10.12 -p 2500 set_obi_role -o OBI_SLAVE
bin/rs_admin -r 10.10.10.2 -p 2500 -t 60000000 boot_strap

表 3-4 参数说明

| 参数 | 说明 |
|----|---|
| -r | RootServer 的 IP 地址。主备 RootServer 时，为 RootServer 的 VIP 地址。 格式: -r [IP] |
| -p | RootServer 的端口号。 格式: -p [Port] |
| -o | 指定主备集群 RootServer。 主集群格式: set_obi_role -o OBI_MASTER 备集群格式: set_obi_role -o OBI_SLAVE |
| -t | 命令的超时时长。 单位: 微秒。 格式: -t [Time] boot_strap |