

# 阿里云

## E-HPC 混合云搭建手册

（半导体行业）

产品版本：1.0.0

文档版本：20220101

## 法律声明

阿里云提醒您在阅读或使用本文档之前仔细阅读、充分理解本法律声明各条款的内容。如果您阅读或使用本文档，您的阅读或使用行为将被视为对本声明全部内容的认可。

1. 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档，且仅能用于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息，您应当严格遵守保密义务；未经阿里云事先书面同意，您不得向任何第三方披露本手册内容或提供给任何第三方使用。
2. 未经阿里云事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。
3. 由于产品版本升级、调整或其他原因，本文档内容有可能变更。阿里云保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在阿里云授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠道下载、获取最新版的用户文档。
4. 本文档仅作为用户使用阿里云产品及服务的参考性指引，阿里云以产品及服务的“现状”、“有缺陷”和“当前功能”的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的介绍及操作指引，但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的，阿里云不承担任何法律责任。在任何情况下，阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害，包括用户使用或信赖本文档而遭受的利润损失，承担责任（即使阿里云已被告知该等损失的可能性）。
5. 阿里云文档中所有内容，包括但不限于图片、架构设计、页面布局、文字描述，均由阿里云和/或其关联公司依法拥有其知识产权，包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意，任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外，未经阿里云事先书面同意，任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称（包括但不限于单独为或以组合形式包含“阿里云”、“Aliyun”、“万网”等阿里云和/或其关联公司品牌，上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司）。
6. 如若发现本文档存在任何错误，请与阿里云取得直接联系

版 本	日 期	修 订 内 容
20221111	2022. 11. 11	初版

# 目录

法律声明 .....	I
目录 .....	III
1. 局点环境确认 .....	1
1.1. 云上云下互通测试 .....	1
1.2. 客户侧挂载点确认 .....	1
1.3. 客户侧 dns 状态确认 .....	1
1.4. 客户侧域管理状态确认 .....	3
1.5. 客户侧调度器状态确认 .....	4
2. 混合云创建 API .....	6
2.1. 通过 API 完成 E-HPC 集群创建 .....	6
2.2. 线下挂载点接入 .....	8
2.3. 云上弹性节点在线下的预设 .....	8
2.4. 客户侧 windows AD server 接入 .....	10
2.5. 客户侧 Openldap 接入 .....	10
2.6. 客户侧 Isf 调度器接入--调度器插件部分 .....	10
3. 镜像制作 .....	17
3.1. DNS 配置的录入 .....	17
3.2. 挂载点配置 .....	17
3.3. AD server 场景下 sudo 组对接 .....	17
4. 集群测试 .....	18

## 1. 局点环境确认

局点环境是指客户在进行混合云创建前，对其自身环境的确认。由于线下环境的差异，E-HPC 混合云的尝试以最大的兼容性来满足不同的线下集群构建混合云需求，若下文中的技术细节依然有不清晰，请联系阿里云同学进行支撑。

混合云构建的局点环境确认，硬件层面需要确认打通线上线下网络，确认共享存储在云上云下可同时挂载；集群服务层面，需要确认 DNS 服务同时可以对云上云下节点进行域名解析，需要确认域控组件可对云上云下节点进行注册，客户需要根据其网段规划在其 dns 中录入预设节点的正/反向域名解析，同时需要客户将预置节点信息(hostname)写入调度器基础配置文档。具体环节如下：

### 1.1.云上云下互通测试

假设云上节点为 ecs01，云下节点为 offline01：

1) 确认节点间可以相互 ping 通；

2) 确认两节点间 ssh 可以相互访问，即 ssh 使用的端口（通常 tcp 协议的 22 端口）可以相互访问，可以使用 nc 或 telnet 进行确认；调度器多基于 ssh 服务来完成集群维护。

3) 确认 DNS 服务可以使用，通常为 53 端口（注意 DNS 需要 UDP 协议），可以使用 nc 或 telnet 进行确认；

4) 其他服务可互通确认，如 windows AD server, Openldap, nis 等需要的端口。

### 1.2.客户侧挂载点确认

假设待使用挂载点为：03cd949fd9-pis41.cn-hangzhou.nas.aliyuncs.com

在挂载点上建立文件夹/test1，使用 mount 命令，在云上云下节点建立挂载点，测试挂载后能读到挂载点上内容。

### 1.3.客户侧 dns 状态确认

我们以 centos 上 bind 组件实现的 dns server 为例（windows 平台 dns server 搭建后，测试方式一致）：

云上 DNS 配置：

a) 集群所在 zone 确定，常规配置在 /etc/named.rfc1912.zones，也可在 /etc/named.conf 中，因其引用 named.rfc1912.zones。具体配置如下图，其中文件 zheng 负责正向解析，即 hostname 到 ip 解析，文件 fan 为反向解析，负责 ip 到 hostname 解析。

```
zone "ehpc" IN {
    type master;
    file "zheng";
    allow-update { none; };
};

zone "200.16.172.in-addr.arpa" IN {
    type master;
    file "fan";
    allow-update { none; };
};
```

b) dns 服务端口及访问限制，如下图所示，设定为 53 端口，任何主机都可以使用该 dns server。

```
options {
    listen-on port 53 { 172.16.200.199; };
    listen-on-v6 port 53 { ::1; };
    directory "/var/named";
    dump-file "/var/named/data/cache_dump.db";
    /* statistics-file "/var/named/data/named_stats.txt";
    memstatistics-file "/var/named/data/named_mem_stats.txt";
    recursing-file "/var/named/data/named.recursing";
    secroots-file "/var/named/data/named.secroots";
    */
    allow-query { any; };
}
```

提供dns服务的主机ip, 并使用53端口

支持任何主机使用

c) 正反向解析中录入集群节点信息，我们以 vcenter 和 testDNS1 两个节点为例，配置文件位置位于/var/named/目录下，具体配置如下，这里我们将 vcenter.ehpc 的域名解析映射到 172.16.200.199，将 testDNS1.ehpc 域名映射到 172.16.200.198。

```
>_ 3. root@dnsServerTest:/var/named x
$TTL 1D
@      IN SOA  @ rname.invalid. (
                                0      ; serial
                                1D      ; refresh
                                1H      ; retry
                                1W      ; expire
                                3H )    ; minimum

      NS      @
      A       172.16.200.199
      AAAA    ::1
vcenter A    172.16.200.199
testDNS1  A   172.16.200.198
~
```

d) 反向解析配置

```
> 3. root@dnsServerTest:/var/named ✕  
$TTL 1D  
@ IN SOA @ rname.invalid. ( 0 ; serial  
1D ; refresh  
1H ; retry  
1W ; expire  
3H ) ; minimum  
NS @  
A 172.16.200.199  
AAAA ::1  
199 PTR vcenter  
198 PTR testDNS1
```

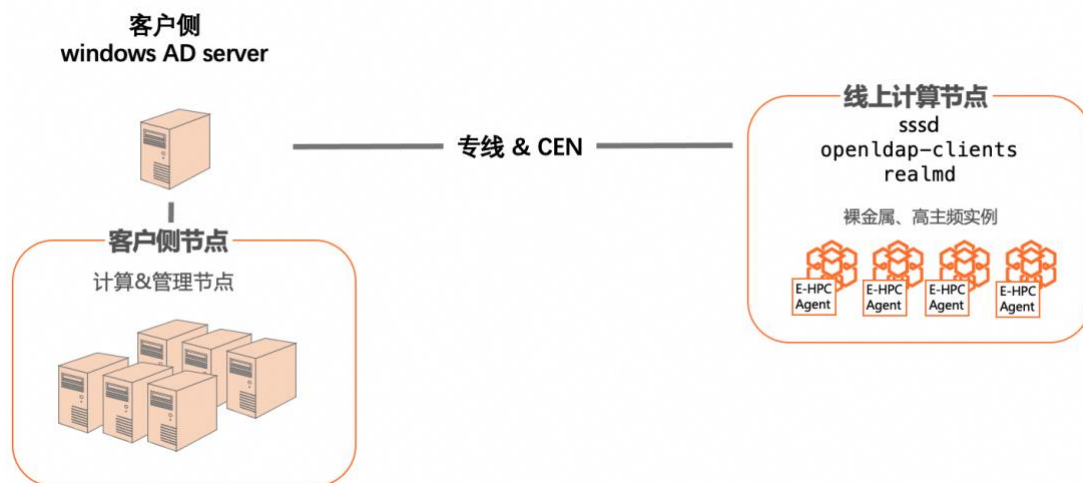
e) 重启 named 服务,测试 server 端可以提供 dns 服务,之后测试 vcenter 和 testDNS1 可以相互解析,注意,为了配合之后调度器中对节点的命名,我们在 resolv.conf 文件中加入对 ehpc 的搜索,并将 dns 解析服务器都设置成之前配置的 dns 服务器位置。

```
[root@dnsServerTest named]# cat /etc/resolv.conf  
options timeout:2 attempts:3 rotate single-request-reopen  
; generated by /usr/sbin/dhclient-script  
search ehpc  
nameserver 172.16.200.199  
nameserver 100.100.2.138  
[root@dnsServerTest named]# ssh vcenter  
root@vcenter's password:  
Last login: Thu Nov 3 10:25:26 2022 from 100.104.86.54  
Welcome to Alibaba Cloud Elastic Compute Service !  
[root@dnsServerTest ~]# ssh testDNS1  
root@testdns1's password:  
Last login: Wed Nov 2 16:25:23 2022 from 172.16.200.199  
Welcome to Alibaba Cloud Elastic Compute Service !  
[root@testDNS1 ~]# ssh vcenter  
root@vcenter's password:
```

## 1.4.客户侧域管理状态确认

### 1) windows AD server 状态确认

E-HPC 混合云方案推荐使用云上资源直连线下 AD server 的方案,如下图:



如上图案例中，线下 AD server 构建的域为 ldap-win002.lework.com，那么云上 ecs01 加入该节点通过如下命令可以完成，该命令需要输入 AD server 端进行认证的用户的密码（如 Administrator，也可以是其他）。

```
realm join --user=Administrator -v ldap-win002.lework.com
```

之后根据客户使用习惯对 sssd 文件进行修改并重启 sssd 服务，即可在 centos 主机上使用 AD server 中已有的账户，如 wutest。

```
systemctl restart sssd
```

```
su wutest
```

## 2) Openldap 状态确认

同 AD server 的接入方案的架构，我们推荐云上节点直连线下 Openldap 服务。具体命令如下，其中 ldapserver 的值为线下 ip，basedn 也是现在使用的域：

```
authconfig --enableldap --ldapserver='ldap://192.168.0.232/,ldap:/// --  
ldapbasedn='dc=ehpc-sh-MDiQgx5qaQ,dc=com' --update
```

```
systemctl restart nslcd
```

## 1.5.客户侧调度器状态确认

确认已挂载 LSF 安装目录

由于 LSF 客户端的安装需要获取 LSF 安装目录，所以需要确定挂载的共享存在是否包含 LSF 安装目录，例如安装目录为 /opt/lsf（以下内容将以该目录为 LSF 安装目录叙述）

修改 lsf.cluster.ehpc 加入节点并启动 LSF 客户端



修改/opt/lsf/conf/lsf.cluster.ehpc 文件，将客户端节点以对应格式加入到配置文件节点列表中并保存

```
执行 /opt/lsf/10.1/install/hostsetup --top="/opt/lsf" --boot="y"
```

配置 lsf.sudoers 文件

```
echo 'LSF_STARTUP_PATH=/opt/lsf/10.1/linux3.10-glibc2.17-x86_64/etc/' >> /etc/lsf.sudoers
```

```
echo 'LSF_STARTUP_USERS="lsfadmin"' >> /etc/lsf.sudoers
```

```
chmod 600 /etc/lsf.sudoers
```

配置 LSF 调度器环境变量并启动客户端服务

```
执行 echo 'source /opt/lsf/conf/profile.lsf' >> /etc/profile
```

```
执行 echo 'source /opt/lsf/conf/cshrc.lsf' >> /etc/csh.cshrc
```

执行 source /etc/profile 之后，执行 service lsf status 进行服务状态确认

#### 5) 检查客户端节点是否加入

在线下调度器节点上执行 yes | lsadmin reconfig 重新配置 lsf 集群，reconfig 之后检查 qhost，查看是否存在该客户端节点

客户端节点执行 yes | lsadmin limstartup , yes | lsadmin resstartup, yes | lsadmin hstartup 来启动 LSF 守护进程服务

## 2. 混合云创建 API

### 2.1. 通过 API 完成 E-HPC 集群创建

由于线下集群的差异性，建议使用 E-HPC 的 api 完成集群创建，具体使用参数如下：

```
const Core = require('@alicloud/pop-core');

var client = new Core({

  accessKeyId: 'xxxxxxxxxx', //客户 keyId

  accessKeySecret: 'xxxxxxxxxx', //客户 key 密钥

  endpoint: 'https://ehpc.cn-hangzhou.aliyuncs.com', //混合云云上集群创建服务点

  apiVersion: '2018-04-12'

});

var params = {

  "RegionId": "cn-hangzhou", //混合云云上集群创建区域

  "Name": "armOnline6-zailong", //混合云云上集群名称

  "OsTag": "CentOS_7.6_64", //集群所需操作系统标签

  "ImageId": "m-bp136x83f14rf27u3z9c",

  "VpcId": "vpc-bp1q4vwldug6bojdg6lcd", //集群所需 VPC

  "VSwitchId": "vsw-bp1fjyh1hiqbuiyhxznid", //集群所需交换机

  "EcsOrder.Compute.InstanceType": "ecs.c7.large", //集群创建所需实例类型

  "ZoneId": "cn-hangzhou-k", //实例所在区域

  "Nodes.1.HostName": "manager", //线下集群管理节点命名

  "Nodes.1.IpAddress": "172.16.100.78", //线下集群 ip 地址

  "Nodes.1.Role": "ResourceManager", //线下集群该节点角色描述

  "Nodes.1.AccountType": "custom", //线下节点使用的域控管理，如客户自持则为 custom
}
```

```
"Nodes.1.SchedulerType": "custom", //线下调度器, 如果是客户自持都选择 custom
"Nodes.2.HostName": "proxymgr", //云上代理节点名称
"Nodes.2.IpAddress": "172.16.200.252", //云上代理节点 IP
"Nodes.2.Role": "AccountManager", //不用修改
"Nodes.2.AccountType": "custom", //不用修改
"Nodes.2.SchedulerType": "lsf", //云下调度器名称
"Location": "ProxyOnline", //集群位置, eda 混合云都位 ProxyOnline
"VolumeType": "local",
"VolumeMountpoint": "03cd949fd9-pis41.cn-hangzhou.nas.aliyuncs.com", //挂载点位置
"Password": "AliHPC123", //集群密码
"WinAdPar.AdUser": "Administrator", //使用 windows AD 时, AD 侧的用户名
"WinAdPar.AdUserPasswd": "Alihpc123", // AD 侧的用户的密码
"WinAdPar.AdDc": "ad-hybrid001.ehpcad.com", //AD 侧给云上主机使用规划的用户域
"WinAdPar.AdIp": "172.16.200.33", //线下 windows AD server 的 IP 地址
// "OpenldapPar.LdapServerIp": "172.16.200.23", //线下 Openldap 的 IP 地址
// "OpenldapPar.BaseDn": "ehpctest", //线下 Openldap 管控账户所使用的 Dn
"ClientVersion": "1.0.1" //集群版本
"Plugin": ' {"pluginMod": "image", "pluginLocalPath": "/opt/plugin" }' //插件模式及路径
}

var requestOption = {
    method: 'GET',
};
```

```
client.request('CreateHybridCluster', params, requestOptions).then((result) =>
{
    console.log(JSON.stringify(result)); //混合云创建接口
}, (ex) => {
    console.log(ex);
})
```

## 2.2. 线下挂载点接入

当前，混合云接入的方式有两种，可以灵活的支持线上或线下的挂载点。

若“VolumeType”:“local”, 则 ehpc 混合云线上集群会按照镜像（下章详述）中已挂载的位置，这样方面线上节点直接对接完成客户线下的应用（如/opt 下的应用），同时直接对接完成/home 目录下的用户文件。推荐使用此种方式。

还可以在线上线下共享存储上创建 ehpc 所需要的文件路径，并进行挂载。如：“VolumeType”赋值为空，同时对“VolumeMountpoint”赋值（如上节 api 示例），E-HPC 会默认在该挂载点下创建集群文件夹，例如 ehpc.xxxxx，该文件夹内包含/opt, /home, /ehpcdata 三个文件。如需挂载其他文件夹，可在镜像中进行约定。线下集群挂载线上节点的场景下，可使用此种方式。

## 2.3. 云上弹性节点在线下的预设

在客户侧 dns 中，将预设节点的 hostname 和 ip 节点的正向解析进行录入。我们约定云上节点的命名规则为，计算节点名称前缀+数字索引，单个虚拟交换机下支持一种计算节点名前缀，例如：计算节点名为 queue，则进行池内弹性伸缩的节点列表为：queue001, queue002, 最大为 queue250。索引值为该节点的 ip 地址最后一位，例如：云上节点规划使用的交换机网段为：172.16.26.0/24，则 queue001，对应的 ip 地址为 172.16.26.1，queue002 对应 172.16.26.2，queue250 对应 172.16.26.250。这些预设节点都需要写入到 dns 的正向解析中（反向选填）。同时将这些预设节点名称，写入到 lsf 调度器对应配置文件内。具体如下图。

DNS 服务器上正向解析的预设节点配置

```
$TTL 1D
@      IN SOA  @ rname.invalid. (
                                0      ; serial
                                1D     ; refresh
                                1H     ; retry
                                1W     ; expire
                                3H )   ; minimum

      NS      @
      A       172.16.25.198
      AAAA    ::1
vcenter A     172.16.25.198
manager A     172.16.25.188
AdDns2 A      172.16.25.197
compute000 A   172.16.25.190
proxymgr A     172.16.26.251
queue001 A     172.16.26.1
queue002 A     172.16.26.2
queue003 A     172.16.26.3
queue004 A     172.16.26.4
queue005 A     172.16.26.5
queue006 A     172.16.26.6
queue007 A     172.16.26.7
```

lsf.cluster 文件中的预设节点配置

```
Begin Host
HOSTNAME model type server RESOURCES #Keywords
#apple Sparc55 SUNSOL 1 (sparc bsd) #Example
#peach DEC3100 DigitalUNIX 1 (alpha osf1)
#banana HP9K778 HPPA 1 (hp68k hpux)
#mango HP735 HPPA 1 (hpux cs)
#grape SGI4D35 SGI5 1 (irix)
#lemon PC200 LINUX 1 (linux)
#pear IBM350 IBMAIX4 1 (aix cs)
#plum PENT_100 NTX86 1 (nt)
#berry DEC3100 ! 1 (ultrix fs bsd mips dec)
#orange ! SUNSOL 1 (sparc bsd) #Example
#prune ! ! 1 (convex)
manager ! ! 1 (mg)
proxymgr ! ! 1 ()
compute000 ! ! 1 ()
queue001 ! ! 1 ()
queue002 ! ! 1 ()
queue003 ! ! 1 ()
queue004 ! ! 1 ()
queue005 ! ! 1 ()
queue006 ! ! 1 ()
End Host
```

lsb.hosts 文件中预设节点的配置

```

Begin Host
HOST_NAME MXJ    r1m    pg    ls    tmp    DISPATCH_WINDOW    AFFINITY # Keywords
#hostA    ()    3.5/4.5    15/    12/15    0    ()    (Y) # Example
#hostB    !    3.5    15/18    12/    0/    (5:19:00-1:8:30    20:00-8:30) (Y)
#hostC    1    3.5/5    18    15    ()    ()    (Y) # Example
#hostD    !    ()    ()    ()    ()    ()    (Y) # Example
#hostE    4    ()    ()    ()    ()    ()    (Y) # Example
default    !    ()    ()    ()    ()    ()    (Y) # Example
proxymgr    !    ()    ()    ()    ()    ()    (Y)
compute000    !    ()    ()    ()    ()    ()    (Y)
queue001    !    ()    ()    ()    ()    ()    (Y)
queue002    !    ()    ()    ()    ()    ()    (Y)
queue003    !    ()    ()    ()    ()    ()    (Y)
queue004    !    ()    ()    ()    ()    ()    (Y)
queue005    !    ()    ()    ()    ()    ()    (Y)
queue006    !    ()    ()    ()    ()    ()    (Y)
End Host

```

将云上计算节点的前缀名写入到 proxymgr 节点上的 /usr/local/cloudnode\_prefix 文件中，以便能够正确识别云上计算节点负载

## 2.4.客户侧 windows AD server 接入

"WinAdPar.AdUser": "Administrator", //使用 windows AD 时，AD 侧的用户名

"WinAdPar.AdUserPasswd": "Alihpc123", // AD 侧的用户的密码

"WinAdPar.AdDc.l": "ad-hybrid001.ehpcad.com", //AD 侧给云上主机使用规划的用户域

"WinAdPar.AdIp": "172.16.200.33", //线下 windows AD server 的 IP 地址

如上所示，为线下 AD server 接入时需要输入的配置参数。AD server 中需要负责对用户管理进行配置。推荐在 AD server 中创建至少两个组（group）对用户进行管理，以支持区分 sudo 权限用户和普通用户。

## 2.5.客户侧 Openldap 接入

"OpenldapPar.LdapServerIp": "172.16.200.23", //线下 Openldap 的 IP 地址

"OpenldapPar.BaseDn": "ehpctest", //线下 Openldap 管控账户所使用的 Dn

如上所示，线上扩容节点会以线下 openldap 为 server，并获取其上对用户、用户组和权限的配置。

## 2.6.客户侧 lsf 调度器接入--调度器插件部分

将插件解压到 manager 共享路径（联系团队获取插件）

```
mkdir /opt/plugin
```

```
mv /root/lsf_hybrid_dns_plugin_v2.zip /opt/plugin/
```

```
cd /opt/plugin/
```

```
unzip lsf_hybrid_dns_plugin_v2.zip
```

根据实际 LSF 软件环境调整插件内容

在制作 LDAP 自定义镜像的同时来配置计算节点 LSF 客户端启动服务（需要根据具体 LSF 安装位置调整）

3.1 创建 script 目录：`mkdir -p /usr/local/lsf/script`

3.2 创建 log 目录：`mkdir -p /usr/local/lsf/log`

3.3 创建启动进程脚本：`/usr/local/lsf/script/lsfinit.sh`

```
#!/usr/bin/env bash
```

```
#
```

```
# EHPC lsf, (c)Copyright 2022
```

```
#
```

```
# Install the LSF master, compute in ALIYUN EHPC platform
```

```
LSF_DIR=/opt/lsf
```

```
LOG_FILE=/usr/local/lsf/log/lsfinit.log
```

```
EHPC_CONF=/root/ehpc.conf
```

```
LSF_ADMIN=lsfadmin
```

```
IS_MANAGER=
```

```
IS_LOGIN=
```

```
IS_COMPUTE=
```

```
LSF_CLUSTER_CONF=${LSF_DIR}/conf/lsf.cluster.ehpc
```

```
LSF_HOST_CONF=${LSF_DIR}/conf/lsbatch/ehpc/configdir/lsb.hosts
```

```
LSF_QUEUE_CONF=${LSF_DIR}/conf/lsbatch/ehpc/configdir/lsb.queues
```

```
LSF_PARAM_CONF=${LSF_DIR}/conf/lsbatch/ehpc/configdir/lsb.params
```

```
LSF_PROFILE_ENV="source ${LSF_DIR}/conf/profile.lsf"
```

```
LSF_CSHRC_ENV="source ${LSF_DIR}/conf/cshrc.lsf"
```

```
echo -e "" | tee -a $LOG_FILE

echo -e "-----" | tee -a
$LOG_FILE

mkdir -p /usr/local/lsf/log

touch ${LOG_FILE}

# 0. Check lsf service

source /etc/profile && service lsf status > /dev/null 2>&1

if [ "$?" -eq "0" ]; then

    echo -e "INFO: lsf service is startup, skip follow steps" | tee -a $LOG_FILE

    exit 0

else

    source /etc/profile && service lsf start > /dev/null 2>&1

    source /etc/profile && service lsf status > /dev/null 2>&1

    if [ "$?" -eq "0" ]; then

        echo -e "INFO: lsf daemons is restartup, skip follow steps" | tee -a
$LOG_FILE

        exit 0

    else

        echo -e "INFO: lsf daemons is , skip follow steps" | tee -a $LOG_FILE

    fi

fi

# 1. Check EHPC config

if [[ ! -f /root/ehpc.conf ]] ; then

    echo -e "ERROR: E-HPC service is not installed." | tee -a $LOG_FILE

    exit 1
```



```
fi

# 2. Retry check user lsfadmin

count=0

while [ 0 -eq 0 ]

do

    id -u $LSF_ADMIN > /dev/null 2>&1

    if [ "$?" -eq "0" ]; then

        echo -e "INFO: check lsf_admin success." | tee -a $LOG_FILE

        break;

    else

        let count=${count}+1

        if [ ${count} -eq 60 ]; then

            echo -e "ERROR: check lsf_admin exceed 5 min, exit" | tee -a
$LOG_FILE

            exit;

        fi

        echo -e "WARN: check lsf_admin ${count} times failed." | tee -a
$LOG_FILE

        sleep 5

    fi

done

# 3. Get current node type by ehpc.conf Role

NODE_ROLE=`cat ${EHPC_CONF} | grep Role | awk -F "=" '{print $2}'`

# 4. Loop the Roles

for ROLE in `echo $NODE_ROLE | tr "," " "`
```

```
do

    # Scheduler

    if [[ "$ROLE" == "ResourceManager" ]] ; then

        IS_MANAGER=1

    fi

    # Login

    if [[ "$ROLE" == "LoginNode" ]] ; then

        IS_LOGIN=1

    fi

    # Compute

    if [[ "$ROLE" == "ComputeNode" ]] ; then

        IS_COMPUTE=1

    fi

    # ProxyMgr

    if [[ "$ROLE" == "ProxyManager" ]] ; then

        IS_PROXYMGR=1

    fi

done

# 5. Install LSF client

if [[ "X$IS_COMPUTE" == "X1" ]] || [[ "X$IS_PROXYMGR" == "X1" ]] ; then

    COMPUTE_HOST=`hostname -s`

    echo -e "INFO: Start to install LSF ${COMPUTE_HOST} ..." | tee -a $LOG_FILE

    ${LSF_DIR}/10.1/install/hostsetup --top="${LSF_DIR}" --boot="y" | tee -
a $LOG_FILE

    echo "${LSF_PROFILE_ENV}" >> /etc/profile
```

```
echo "${LSF_CSHRC_ENV}" >> /etc/csh.cshrc
```

```
source /etc/profile && service lsf start | tee -a $LOG_FILE
```

```
source /etc/profile && service lsf status | tee -a $LOG_FILE
```

```
echo -e "INFO: Finish to install LSF ${COMPUTE_HOST} ..." | tee -a $LOG_FILE
```

```
fi
```

```
echo -e "INFO: Init successfully." | tee -a $LOG_FILE
```

```
exit 0
```

### 3.4 创建启动服务： lsf-init.service

```
[Unit]
```

```
Description=Install AIP by E-HPC node roles
```

```
After=local-fs.target    network-online.target    cloud-final.service    ehpc-  
config.service
```

```
Requires=network-online.target cloud-final.service ehpc-config.service
```

```
Wants=local-fs.target
```

```
[Service]
```

```
Type=oneshot
```

```
ExecStart=/bin/bash /usr/local/lsf/script/lsfinit.sh
```

```
RemainAfterExit=yes
```

```
Restart=no
```

```
[Install]
```

```
WantedBy=multi-user.target
```

3.6 拷贝启动服务到系统目录: `cp lsf-init.service /lib/systemd/system`

3.7 load 启动服务目录: `systemctl daemon-reload`

3.8 enable 启动服务目录: `systemctl enable lsf-init`

使用上一步的自定义镜像通过 API 创建带有插件的混合云集群

## 3. 镜像制作

上一章节中，我们讲到了 E-HPC 混合云场景下在调度器及域控组件对接上的必要输入及基本原理。本章节将围绕如何制作适合该场景下的镜像展开，镜像是云计算提供的能力，按需制作镜像并在集群扩容时选择所需镜像，可以保证计算节点配置的一致性，并大大节约运维与操作成本。

### 3.1.DNS 配置的录入

若客户对阿里云云助手或挂载阿里云上共享存储有需求，则线上扩容节点的操作系统中需要在 dns 的寻址配置文件/etc/resolv.conf 文件中保留阿里云默认地址中的一个，另一个地址可配置用户自己的 dns。

阿里云 ecs 上的 dns 配置修改，请参考如下链接，注意/etc/cloud/cloud.cfg 的配置。

[https://help.aliyun.com/knowledge\\_detail/188420.html?spm=a2c4g.11186623.0.0.5c001c217pUglM](https://help.aliyun.com/knowledge_detail/188420.html?spm=a2c4g.11186623.0.0.5c001c217pUglM)

```
options timeout:2 attempts:3 rotate single-request-reopen
; generated by /usr/sbin/dhclient-script
search ehpcad ← 客户hostname解析所在域
nameserver 172.16.25.198 ← 客户自身dns地址
nameserver 100.100.2.138 ← 阿里云dns地址
```

### 3.2.挂载点配置

挂载点的使用配合 local 模式下的 VolumeType。注意制作带挂载点的镜像时，除了需要进行 mount 操作外，还需要在/etc/fstab 内进行修改。

### 3.3.AD server 场景下 sudo 组对接

若客户需要接入 AD 中的 sudo 权限组和普通权限组，则需要在镜像的 sudo 配置文件中进行修改，例如 ad server 中我们规定 adminSudo 组内的用户为 sudo 权限用户，则需要 visudo 后，在最优一行加入授权操作，如下。

```
%adminSudo    ALL=(ALL)        ALL
```

```
[root@queue001 ~]# id lsfadmin
```

```
uid=272801108(lsfadmin) gid=272800513(domain users) 组 =272800513(domain users),272801114(adminsudo)
```

## 4. 集群测试

### 集群创建测试



线下集群中用户同步至线上节点

```
[root@queue001 ~]# id lsfadmin
```

```
uid=272801108(lsfadmin) gid=272800513(domain users) 组 =272800513(domain users), 272801114(adminsudo)
```

线上节点挂载线下存储

使用 `df -h` 命令查看。

线上线下节点间互 ping 及互 ssh

此处目的为测试 dns 的能力。

用户在线上线下节点间的免密环境

使用 root 以外的用户，在节点间 ssh，可以免密跳转。

预设节点配置完成后调度器重启后对预设节点的发现

```
[root@manager ~]# lsrfrestart
Restarting all LIMs ...

Checking configuration files ...
No errors found.

Restart only the master candidate hosts? [y/n] n
Do you really want to restart LIMs on all hosts? [y/n] y
Restart LIM on <manager> ..... done
Restart LIM on <proxymgr.ehpcad> ..... done
Restart LIM on <compute000> ..... done
Restart LIM on <queue001.ehpcad> ..... done
Restart LIM on <queue002.ehpcad> ..... done
Restart LIM on <queue003.ehpcad> ..... done
Restart LIM on <queue004.ehpcad> ..... done
Restart LIM on <queue005.ehpcad> ..... done

Trying unavailable hosts :
Restart LIM on <queue006.ehpcad> ..... █
```

作业提交后，可以在线上线下一起运行

```
[lsfadmin@manager ~]$ bjobs -a
JOBID  USER  STAT  QUEUE      FROM_HOST  EXEC_HOST  JOB_NAME  SUBMIT_TIME
105    lsfadmi  EXIT  normal    manager    queue003.eh test1.lsf  Nov  8 15:57
106    lsfadmi  EXIT  normal    manager    queue003.eh test1.lsf  Nov  8 15:59
107    lsfadmi  DONE  normal    manager    queue003.eh test1      Nov  8 16:00
109    lsfadmi  EXIT  normal    manager    -          test1      Nov  8 16:22
110    lsfadmi  DONE  normal    manager    2*queue001. test1      Nov  8 16:24
                                2*compute000
```

作业提交和完成后，集群的自动伸缩

