



NewStart RDS 5.0

用户手册

广东中兴新支点技术有限公司

Guangdong Zhongxing NewStart Technology Ltd.

本手册适用于 5. x 系列版本

目 录

前 言	1
本书约定	1
正式版包装组件	1
第 1 章 概 述	2
第 2 章 安装前准备	3
2.1 集群搭建流程快速预览	3
2.2 硬件(虚拟化硬件)需求列表	3
2.3 软件需求	4
2.4 其它系统设置及通讯端口	5
2.5 编写服务的集群脚本	5
第 3 章 安装与注册	6
3.1 产品部署	6
3.2 产品注册	7
3.3 产品卸载	8
第 4 章 通过 cli 进行配置	9
4.1 远程容灾配置(RDS 2+1)	10
4.2 远程容灾配置(RDS 1+1)	27
4.3 远程容灾配置($n*(1+1)$)	30
4.4 集群启动	52
第 5 章 通过 web 进行配置	53
5.1 Web 登陆	53
5.2 RDS 2+1 初始化配置	60
5.3 RDS 1+1 初始化配置	75
5.4 集群启动	76
第 6 章 集群的验收测试	78
第 7 章 集群交互式命令	79
7.1 配置修改	80
7.2 节点管理	91
7.3 服务管理	93
7.4 资源管理	96
7.5 集群监控	99
7.6 SNMP 告警监控 (选配)	102
7.7 其它命令	104
第 8 章 集群非交互式命令	106
第 9 章 集群告警与查询	118
9.1 集群告警 oid 及说明	118
9.2 告警查询	119
附录	121
附录 A—业务管控脚本编写规范说明	121
1 脚本编写前须知	121
2 脚本编写参考	121

3 手动调试脚本	123
4 脚本编写方案	124
附录 B—在线注册系统使用说明	127
附录 C—典型故障排除与恢复	131
■ 系统假死故障	131
■ 服务 unstable 故障	131
■ 集群裂脑/心跳链路 offline 故障	132
■ 集群启动失败故障	134
■ 数据裂脑故障	135
■ web 页面无法正常显示	136
附录 D—技术支持	137

前言






本书约定

1. 组织和印刷约定

格式	意义
输入的命令	等宽粗体 （新宋体小四）
命令输出	等宽字体（新宋体小四）
命令的注释	<i>等宽粗斜体</i> （新宋体小四）

2. 安全符号约定

在本手册中可能出现下列安全符号，它们所代表的含义如下：

安全符号	意义
 危险	表示若忽视安全告诫，就有可能发生损坏设备或导致人身伤害的严重事故
 警告	表示若忽视安全告诫，就有可能发生损坏设备或导致人身伤害的事故
 注意	表示若忽视安全告诫，就有可能导致设备损坏、数据丢失、设备性能降低以及其它不可预知的后果
 提示	表示该内容是特别提示的信息
 说明	表示该内容是正文的附加信息

正式版包装组件

组件	数量
安装光盘	1
产品简介	1
安装配置手册	1
许可协议卡	1

第 1 章 概 述

本产品是广东中兴新支点技术有限公司自主研发的linux平台下的容灾产品，支持RDS1+1/RDS N+M模式。

本手册将引导您完成环境的搭建，通过图文相辅、通俗易懂的形式详细地介绍RDS的安装及使用过程。

光盘（或者 iso文件）内容

文件/目录	内容说明
conf/	集群和 web 的通用脚本函数
doc/	文档目录, 包括用户手册和简易安装手册, 版本发布说明
install	安装脚本
kernels	镜像内核包
packages/	Cluster 软件包所在目录
tools/	一些工具脚本
uninstall	卸载脚本
upgrade	集群组件升级脚本
web/	web 管理模块

如果程序光盘中的软件版本不是最新商用版本, 建议用户上我司网站下载, 网址:
<http://www.gd-linux.com>

第 2 章 安装前准备

2.1 集群搭建流程快速预览

1. 熟悉集群软件的相关概念和安装环境，确定架构和配置方案，详见[2.2](#)至[2.4](#)节说明。
2. 安装业务软件（比如oracle数据库程序），确保服务器的业务软件环境一致。
3. 安装集群软件，获取 License文件，详见“[第3章安装](#)”。
4. 配置集群软件，该操作可在命令行或web界面中完成。
5. 编写服务的集群脚本，通过在单机上调试，确保集群脚本能正常启动、停止和检测服务，方法详见“[2.5节编写服务的集群脚本](#)”。
6. 启动集群软件，查看集群的运行情况，进行服务的手动倒换和故障模拟测试，确认服务能自动故障迁移，方法详见“[第6章](#)”。

2.2 硬件(虚拟化硬件)需求列表

硬件（虚拟化硬件）需求列表（节点硬件配置相同）		
服务器节点	数量	用途
硬盘	≥ 1	可用于： 1. 安装 Linux 操作系统 2. 作为镜像盘的底层存储
网卡	≥ 2	1. 业务链路（千兆以上网卡） 2. 镜像链路（千兆以上网卡，镜像版本使用，可以兼做心跳链路） 3. 心跳链路（两条、百兆以上网卡）
交换机	≥ 1	用于网卡链路连接
磁盘阵列	1	存放共享数据(选配)

2.3 软件需求

操作系统支持列表(表中的“x”表示系列小版本):

CGSLv4.x/Redhat6.x/Centos6.x	2.6.32-220.x.el6.x86_64
	2.6.32-358.x.el6.x86_64
	2.6.32-431.x.el6.x86_64
	2.6.32-504.x.el6.x86_64
	2.6.32-573.x.el6.x86_64
	2.6.32-642.x.el6.x86_64
	2.6.32-696.x.el6.x86_64
	2.6.32-754.x.el6.x86_64
CGSLv5.x/Redhat7.x/Centos7.x	3.10.0-123.x.el7.x86_64
	3.10.0-229.x.el7.x86_64
	3.10.0-327.x.el7.x86_64
	3.10.0-514.x.el7.x86_64
	3.10.0-693.x.el7.x86_64
	3.10.0-862.x.el7.x86_64
	3.10.0-957.x.el7.x86_64
	3.10.0-1062.x.el7.x86_64
	3.10.0-1127.el7.x86_64
	3.10.0-1160.el7.x86_64
CGSLv6.x/Redhat8.x/Centos8.x	4.18.0-80.el8.x86_64
	4.18.0-147.el8.x86_64
	4.18.0-193.el8.x86_64
	4.18.0-240.el8.x86_64
	4.18.0-348.el8.x86_64
	4.18.0-372.el8.x86_64
	5.10.134-12.zncgsl6.x86_64
	5.10.134-13.zncgsl6.x86_64
	5.10.134-12.zncgsl6.aarch64
	5.10.134-13.zncgsl6.aarch64
Sles11 sp1/sp3	3.0.101-0.47.105-default
	3.0.101-63-default
	3.0.76-0.11-default
Sles12.x	3.12.49-11-default
	4.4.138-94.39-default
	4.4.143-94.47-default
	4.4.73-5-default

如需要对其它操作系统或内核提供支持, 需联系我们。

2.4 其它系统设置及通讯端口

- (1) 所有安装需要root用户权限，并且root用户使用的shell必须为bash，如果不是，请修改/etc/passwd中root帐号来设置。
- (2) 根据系统规划设置好操作系统的主机名，下面给出命令行参考步骤：
#vi /etc/HOSTNAME # SuSE早期发行版
#vi /etc/sysconfig/network # CGSL、红帽等早期发行版
#hostnamectl set-hostname xxx #CGSLv5/Redhat7/Centos7/Sles12等
- (3) 确认操作系统没做过目录或文件类的安全加固操作（已加固请先解开），否则可能会导致安装失败。
- (4) Cluster 配置和运行时需要开启如下端口：

协议和端口	用途说明
tcp, 20008	配置文件同步等
udp, 694	心跳网卡通讯，此端口不能被占用，否则Cluster无法启动
tcp, 10003	连接Web管理工具，如果不用Web，此端口可关闭
tcp, 10001、10004	与web管理组件相关，仅限本地进程通讯，可加入防火墙，禁止外部防问
tcp, 7788~7795	镜像数据同步缺省同步端口范围
tcp, 10018	与告警服务相关，仅限本地进程通讯，可加入防火墙，禁止外部防问
udp, 161	snmp告警服务查询端口，选配
udp, 162	snmp告警服务发送端口，选配

2.5 编写服务的集群脚本

服务的集群脚本规定了如何启动、停止和检测业务程序，其与/etc/init.d目录下的服务脚本的实现是类似的，具体标准可参考“[附录A-集群脚本编写规范](#)”，实际编写时可参考/etc/ha.d/resource.d目录下提供的脚本模版/例子快速编写服务双机脚本，如：

- java相关应用（比如weblogic脚本，tomcat脚本）可以参考tomcat_example.sh来编写；
- 其他应用（采用查进程方式检测）可以参考oracle_example.ps编写，数据库与应用结合的脚本参考sybase_app.ps编写；

第 3 章 安装与注册

3.1 产品部署

将安装光盘放入服务器的光驱中，挂载光盘到/mnt 子目录：

```
# mount -o ro /dev/cdrom /mnt
```

如果是 iso 文件，把 iso 文件以二进制(bin)传输方式上传到服务器/home/ 目录，挂载 iso 文件到/mnt 子目录：

```
# mount -o loop /home/xxxx.iso /mnt
```

执行安装脚本开始安装：

```
# /mnt/install
```

```

Welcome to use Guangdong NewStart Software
The NewStart Cluster Universal Installation/Uninstallation
Wizard Note: <y,n,c> -- <yes,no,cancel>
Support email: ha@gd-linux.com 1

Please input the product SN: 44CF61-BA1JL8-22BD1R-50075E 2
You's SN is NewStart Remote Disaster Recovery Software product
SN! Checking cluster status ... NOT running
Installing, please wait a few minutes...
Installing the
/mnt/packages/x86_64/redhat7/newstartha-5.0.1-20180118.x86_64.rpm ...
Preparing... ##### [100%]
Updating / installing...
1:newstartha-5.0.1-20180118##### [100%]
Start tserver OK.

[OK]

Make /etc/ha.d/lic/newstartha.key succeeded. 3
ha_kernel_dir:
/mnt/kernels/3.10.0-693.5.2.el7.x86_64
Preparing... ##### [100%]
1:drbd-utils-8.9.6-1 ##### [ 33%]
2:drbd-proxy-2.0.0-1 ##### [ 67%]
3:drbd-km-3.10.0_693.5.2.el7.x86_64##### [100%]

[OK]

Do you want to install cluster Web-based Administrative Tool [y,n] (y)

[OK]

web-based administrative tool install, deploying, please wait...
```

```
ha_httpserver installed OK !  
ha_nginx installed OK !  
ha_web installed OK !  
Setting web-based administrative tool automatically as a system service...  
[OK]  
Please remember to change the default web password after manual start web  
management tool!  
The component(s) is installed completely, you must configure it before run it.
```



说明:

1. 技术支持邮箱，使用过程中如有疑问或问题可以发邮件到此邮箱
2. 输入序列号，后面集群产品注册时会用到
3. key文件已经生成



注意:

- 1、iso文件上传服务器时必须是二进制(bin)方式上传，文本(ASCII)模式上传会导致软件安装失败。
- 2、所有服务器都要进行集群软件单独安装，否则无法配置集群。

3.2 产品注册

3.2.1 常规流程：获取试用/商用序列号-->生成 key 文件-->使用key文件去申请 license文件-->上传license文件到相应的服务器-->命令行运行“cli”或登陆web工具确认注册是否生效。

生成 key 文件：产品部署过程中会提示输入产品许可号（其可在产品包装盒里面的《新支点产品许可协议》卡片上找到），部署完成后生成的 key 文件（newstartha.key）可在/etc/ha.d/lic/目录下找到，用二进制(bin)传输方式下载到本地，并改名以方便识别，如 newstartha_ip132.key，集群中各个节点都要执行上述操作步骤。

申请 license 文件：web 在线申请, 参考“[附录一在线注册系统使用说明](#)”。

验证 license 有效性：获取 license 文件后改名为 newstartha.lic，每个节点 newstartha.lic 内容不同，key 文件与 license 文件要一一对应，然后用二进制(bin)方式上传到相应服务器的/etc/ha.d/lic/目录下，最后 shell 下运行 cli 检查 license 是否有效，验证通过会显示正确序列号，无效提示“License files invalid”。

3.2.2 产品试用:

RDS试用: nshacluster -product-switch RDS



提示:

- *) 只有CGSL操作系统才可以使用该命令, 试用期只有30天
- *) 集群没有初始化情况下需每个节点执行该命令, 试用 30 天,每个集群环境只能试用一次, 如需继续试用, 需联系我们。
- *) RDS 试用序列号: 44CF61-BA1JL8-22BD1R-50075E

至此, 软件安装完成, 要想使用集群, 必须进行配置。

3.3 产品卸载

```
# mount -o ro /dev/cdrom /mnt 挂载光盘
```

或者

```
# mount -o loop /home/xxxx.iso /mnt 挂载 iso
```

执行卸载脚本

```
# /mnt/uninstall
```

```
# /mnt/uninstall
```

```
                Welcome to use Guangdong NewStart Software
```

```
            The NewStart Cluster Universal Installation/Uninstallation
```

```
                Wizard Note: <y,n,c> -- <yes,no,cancel>
```

```
                Support email:  ha@gd-linux.com
```

```
Uninstall  newstartha-5.0.1-20180124.x86_64,  Are  you  sure?
```

```
[y/n] (n)y
```

```
    Checking  cluster status ...                NOT running
```

```
Uninstalling the  newstartha-5.0.1-20180124.x86_64 ...
```

```
[OK]
```

```
Uninstall DRBD...
```

```
[OK]
```

```
Uninstall webadmin...
```

```
[OK]
```

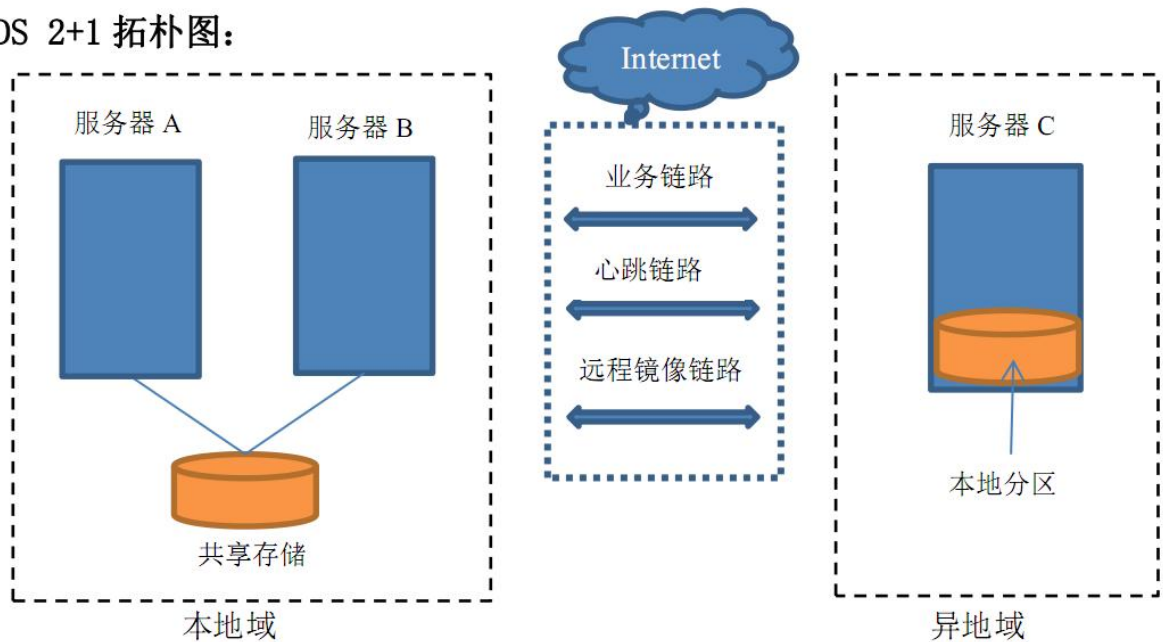
第 4 章 通过 cli 进行配置

Cli 是 NewStart 的命令行配置管理工具，以下是使用技巧：

- 1、双击 Tab键或者运行help显示集群所有命令。
- 2、运行help <集群命令>能查询该命令的功能。
- 3、利用Tab键可以像bash shell一样自动补全命令行输入。
- 4、按向上箭头按钮出现上次使用过的命令。
- 5、输入一个属性值时，如果该属性已有旧值或默认值，直接按回车键可以选择保持旧值不变或者使用默认值。
- 6、cli对输入的属性值都进行合法性校验（本地节点或者远程节点），如果校验错误，则输出错误信息，提示重新输入。
- 7、运行cluster-stat命令并按“up”或“down”选中服务后回车可查看服务所有资源运行情况，按“a”返回上一层。
- 8、运行cluster-stat命令时按Ctrl+C或“q”退回到cli，cli模式下按Ctrl+C会退出cli。

4.1 远程容灾配置(RDS 2+1)

RDS 2+1 拓扑图：



文中 RDS 2+1 配置规划参考：

容灾模式	RDS 2+1
节点	本地域节点：linux76、linux77 异地域节点：linux78
链路网卡	工作链路：ens192 (192.168.8.x) 心跳链路：ens224(192.168.11.x)、ens256(192.168.12.x) 镜像链路：ens224(192.168.11.x)
工作链路 IP	192.168.8.76、192.168.8.77、192.168.8.78
心跳链路 IP	单播心跳：192.168.11.76、192.168.11.77、192.168.11.78 单播心跳：192.168.12.76、192.168.12.77、192.168.12.78
镜像盘（镜像资源）	r0
镜像盘底层盘	本地域节点：共享存储 (/dev/sdc2) 异地域节点：本地分区 (/dev/sdb)
仲裁盘（选配）	本地域：共享存储 (/dev/sdc1)
内置镜像服务	用来管理镜像资源, 业务服务添加存储资源时自动生成，不用额外添加
业务服务	服务名字：ora（应用：Oracle）
业务浮动 IP(子网掩码)	192.168.8.62(255.255.255.0)
挂载点	/home/db
业务管控脚本	/home/oracle.sh

初始化配置

初始化命令摘要（必须按以下序列号顺序执行）

序号	命令	功能描述
1	cluster-init	集群的初始化配置，格式： <code>cluster-init</code> 。
2	mirror-init	镜像资源的初始化配置，格式： <code>mirror-init</code> 。
3	service-init	服务的初始化配置，格式： <code>service-init</code> 。



提示：`cli` 中的初始化过程是交互式操作过程，提示中中括号 “[] ” 里面的值表示可以输入的值，小括号 “() ” 里面的值表示推荐输入值，直接回车表示输入小括号里面的值。

cluster-init

执行此命令会生成新的集群配置文件，原有的集群配置会丢失，初始化过程如下：
（配置文件位置：`/etc/ha.d/ha.cf`）

```
cli:~>cluster-init
```

```
=====
Cluster Initialization Utility
=====
```

```
This utility sets up the initialization information of cluster. It
prompts you for the following information:
```

```
-Hostname
-Information about the heartbeat channels
-How long between heartbeat
-How long to declare heartbeat fails
-Watchdog configuration
-Worklink_hb configuration
-Kernel_panic configuration
-data and service policies
```

```
Please input cluster name: [q] rds_cluster 1
```

```
Input the 1th node name and IP separated by spaces: [q] linux76
192.168.11.76 2
```

```
Input the 2th node name and IP separated by spaces: [q] linux77
192.168.11.77
```

```
Do you want to continue to add node? [y/n] (y)
```

```
Input the 3th node name and IP separated by spaces: [c,q] linux78
192.168.11.78
```

```
Do you want to continue to add node? [y/n] (y) n
```

```
How long between heartbeats(in seconds): [q] (3) 3
```

```
How long to declare heartbeat has broken(in seconds): [q] (100) 4
```

```

Do you want to enable watchdog ? [y/n] (n) 5
Add heartbeat links: IP type:
    0) IPv4
    1) IPv6
Select a type: [0-1,q] (0) 6
Detected NICs on cluster: ens192 ens224 ens256
Input a network device for unicast heartbeat: [q] (ens192) ens224 7
Input linux76 IPv4 address [q] (192.168.11.76)
Input linux77 IPv4 address [q] (192.168.11.77)
Input linux78 IPv4 address [q] (192.168.11.78)
Add another ucast heartbeat link? [y/n] (y) 8
Input a network device for unicast heartbeat: [c,q] (ens192) ens256
Input linux76 IPv4 address [q] (192.168.12.76)
Input linux77 IPv4 address [q] (192.168.12.77)
Input linux78 IPv4 address [q] (192.168.12.78)
Add another ucast heartbeat link? [y/n] (y) n
Do you want to enable worklink_hb ? [y/n] (y) n 9
Do you want to enable auto_failover ? [y/n] (n) 10
Please configure data and service policies:
    0) the maximization of data security
    1) the maximization of service
    availability
Please select: [0-1,q] 0 11
Do you want to enable kernel_panic ? [y/n] (n) 12
Do you want to add domain-node ? [y/n] y 13
Please select nodes for the 1th domain-node:
    0) linux76
    1) linux77
    2) linux78
Select one or more nodes separated by spaces: [0-3,q] 0 1
Do you want to continue to add domain-node? [y/n] (y)
Please select nodes for the 2th domain-node:
    0) linux76
    1) linux77
    2) linux78
Select one or more nodes separated by spaces: [0-3,q] 2
Do you want to continue to add domain-node? [y/n] (n)
Do you want to add a lock disk(recommend) ? [y/n] (n) y 14
    0) add lockdisk to domain
    1) add lockdisk to cluster
    b) Back to previous menu
Please select an operation: [0-1,b,q] 0
Please input the partition name(/dev/sdc) for domain linux76: [c,q] /dev/sdc1

Warning:All data in /dev/sdc1 of linux76 will be destroyed, sure to format it? [y/n]
(n) y

```

Please input the partition name(/dev/sdc) for domain linux78: [c,q] c

Cluster information verification:

```
clustername: rds_cluster
node: linux76 192.168.11.76
node: linux77 192.168.11.77
node: linux78 192.168.11.78
logfile: /var/log/ha-log
keepalive: 3
deadtime: 100
watchdog: off
ucast: eth1 192.168.11.76 192.168.11.77 192.168.11.78
ucast: eth2 192.168.12.76 192.168.12.77 192.168.12.78
worklink_hb: on 226.0.60.192 ff18::e43d:3cd2
auto_failover: off
auto_run: on
maximization: data-security
kernel_panic: off
domain_node: linux76 linux77 lockdisk /dev/sdc1
domain_node: linux78
```

Is this information correct? [y/n,q] (y)¹⁵

Please run mirror-init to initialize your mirror resource.

cli:~>



说明:

- 1、自定义集群名称。
- 2、输入节点名称和 IP 地址。节点名称必须与 `uname -n`（即机器真实名称）相同，IP 地址必须是网卡的物理 IP 地址(推荐心跳网卡 IP)，建议每个节点配置多个 IP 地址，以空格分隔。
- 3、集群心跳通讯间隔；取值范围：1~10，单位为秒。
- 4、判断其它节点 `dead` 的时间间隔（也称 `deadtime`），也就是多长时间没有收到某节点任何心跳包就认为该节点已经死了，取值范围：15~300，单位为秒。
在虚拟化场景应用时需要注意：集群配置的 `deadtime` 时间要少于虚拟机的OS 启动时间（如两者相差 5 秒），防止 OS 重启时间太短带来的影响。
- 5、是否开启集群自带的 `Watchdog`（看门狗）功能，详细说明见下面注意事项，缺省不开启。
- 6、配置心跳链路,这里的心跳类型为单播心跳（心跳类型由集群类型决定，用户根据交互式向导操作即可，操作上类似），支持 `ipv4` 或 `ipv6`；注意集群中所有节点的网卡名需保持一致，否则会有问题。
在虚拟化场景应用时需要注意：用作两条心跳的两张虚拟网卡要对应不同的物理网卡，避免一张物理网卡故障，导致集群所有的心跳都断开的情景。
- 7、输入心跳网卡名称，后面直接回车，完成第一条心跳链路配置。
- 8、是否再增加一条心跳链路，这里选 `y`，重复上述 7 中操作
- 9、是否启用工作网卡链路作为备份心跳链路：当启用时，如果正常的心跳链路数目小于两条，此功能生效；如果故障的心跳链路恢复并且正常心跳链路数目大于等于两条，此功能暂时失

效；这里开启的心跳类型为组播心跳，集群所在的网络环境如对组播包的传输有限制，不建议开启，否则建议开启。

10、服务切换策略，n 表示业务故障手动切换，y 表示业务故障自动切换。

11、选择数据同步复制过程中的两种处理策略：数据安全最大化与服务持续最大化；

前者强调数据安全，优先保障用户数据一致性，适用于对数据一致性要求比较高的应用，后者优先保障用户服务的可用性，适用于用户对服务可用性要求比较高，而对丢失部分数据不敏感的应用。

12、是否启用 Linux kernel panic（内核恐慌）功能；为了保证数据安全，在Cluster异常，比如运行过程中子进程崩溃）或者服务停止失败时触发 kernel panic 使服务器重启，用户业务切换到备机，由于服务停止与集群脚本密切相关，建议集群调试阶段禁用此功能，调试完成后启用此功能，用命令cluster-modify 修改。

13、配置节点域，这里是 2+1（本地+异地）的配置，需配置两个节点域，每个域配置两个节点，不能重复配置。

14、是否增加仲裁盘，选配，这里填 y。

仲裁盘是用于在集群心跳全部中断时（裂脑）保证数据安全的一种解决方案，它是在共享磁盘阵列上划分一个很小的分区（大于 30M 即可），不创建文件系统，集群节点同时读写此分区，交换当前的成员信息，防止裂脑发生，保证数据安全，有条件都建议配置。

15、确认填写的配置信息是否正确，如果不正确，输入“n”进行重新初始化，如正确，输入“y”。



注意：

WatchDog机制基于Linux系统自带的softdog模块（俗称“软狗”），在系统“挂死”（hanged up）或者“崩溃”（crashed）时自动重启系统的。如果服务器配置了硬件看门狗（即“硬狗”），我们建议禁用软狗。Watchdog 重启系统的条件：

- 1）配置了仲裁盘，Cluster软件被‘挂起’超过（60s）无法写仲裁盘；
- 2）没有配置仲裁盘，Cluster软件被‘挂起’超过(60s)无法发送心跳包。

至此，cluster-init 配置完成，若需要修改配置，可通过命令 cluster-modify 进行修改

mirror-init

执行此命令会生成新的镜像资源配置文件，原先镜像配置会丢失，初始化过程如下：（配置文件位置：/etc/drbd.d/xx.res）

```
cli:~>mirror-init

=====
Mirror Initialization Utility
=====

It prompts you for the following information:
    - Mirror resource name
    - Mirror disk device name
    - Mirror IP address
    - Filesystem type
Input resource name: [q] (r0) 1
Input port number: [1-65535,q] (7788) 2
Please select nodes to run mirror r0:
    0) linux76
    1) linux77
    2) linux78
Select one or more nodes separated by spaces: [0-2,q] 0 1 2
Input mirror IP address for linux76: [q] 192.168.11.76 3
Input a disk used to store data on linux76: [q] /dev/sdc2 4
Input mirror IP address for linux77: [q] 192.168.11.77
Input a disk used to store data on linux77: [q] /dev/sdc2
Input mirror IP address for linux78: [q] 192.168.11.78
Input a disk used to store data on linux78: [q] /dev/sdb
Input memlimit for proxy r0 [500M/G]: [q] (500M) 5
select data compression policy for proxy
    0) low compression
    1) high compression
    2) no compression
Please select an item: [0-2,q] 2 6
DRBD resource information verification:

    name: r0
    device: /dev/drbd0
    port: 7788
    node: linux76
        disk: /dev/sdc2
        meta-disk: internal
        IP address: 192.168.11.76
    node: linux77
        disk: /dev/sdc2
        meta-disk: internal
```

```

        IP address: 192.168.11.77
node: linux78
        disk: /dev/sdb
        meta-disk: internal
        IP address: 192.168.11.78
proxy memlimit: 500M
proxy plugin: none
Is this information correct? [y/n,q] (y) 7
Do you want to add another mirror resource? [y/n] (n) 8

All the data of /dev/sdc2 on node linux76 will be destroyed. Are you sure
to continue? [y/n] (y) 9

All the data of /dev/sdb on node linux78 will be destroyed. Are you sure
to continue? [y/n] (y)

Creating meta-data for resource r0...

Create meta-data for resource r0 success

Waiting for mirror proxy starting...
start proxy success.

Waiting for mirror starting...
start mirror success.
    0) linux76
    1) linux78
Please choose the sync source for resource r0: [0-1,q] 0 10

Which filesystem type that you want to format /dev/drbd0(mirror
resource: r0) to:
    0) ext2
    1) ext3
    2) ext4
    3) xfs
    4) none
Select a filesystem type [0, 4]:2 11

mke2fs 1.42.9 (28-Dec-2013)
Filesystem label=
OS type: Linux
Block size=4096 (log=2)
Fragment size=4096 (log=2)
Stride=0 blocks, Stripe width=0 blocks
131072 inodes, 524263 blocks
26213 blocks (5.00%) reserved for the super user
First data block=0

```

```
Maximum filesystem blocks=536870912
16 block groups
32768 blocks per group, 32768 fragments per group
8192 inodes per group
Superblock backups stored on blocks:
    32768, 98304, 163840, 229376, 294912

Allocating group tables: done
Writing inode tables: done
Creating journal (8192 blocks): done
Writing superblocks and filesystem accounting information: done

.

Format disk /dev/drbd0 to ext4 success
Waiting for r0 synchronization...
r0 Resynced percentage: 100.0 %
set secondary all successful

Mirror init success!
Please run service-init to initialize you services.
```

NOTE

说明：

1. 自定义镜像资源名称，用于集群内部识别。
2. 自定义镜像链路端口。
3. 输入该节点可用于镜像链路通讯的物理网卡 IP，可配多个，使用空格分开。
4. 输入业务数据存放的存储分区名称
5. **Mirror** 代理的缓存大小，缺省 200M，取值范围无限制，可根据高峰时间内的业务数据量来确定合适值，简单估算：代理缓存值=（业务数据量速率 M/s-宽带速率 M/s）* 高峰时间/秒
6. 选择传输数据的压缩率，分为低压缩率、高压缩率和不压缩，低压缩率是比较均衡，适用大部份场景推荐配置，高压缩率适用网络延时较大的场景
7. 确认配置是否正确
8. 是否增加第二个镜像资源，这里要选 n。
9. 业务数据存放的存储分区首次使用需要格式化，这里选择y，回车即可
10. 选择数据同步源，这里同步的是业务数据，需选择存在业务数据的节点为同步源，假如没有业务数据，则同步源自行先按需选择。
11. 对镜像分区/dev/drbd0进行格式化。

至此，mirror-init 配置完成，若需要修改配置，只能先删除 mirror 才能重新配置 mirror，可通过命令 mirror-adm 进行增删

service-init

一般情况下，该命令只用于服务初始化配置中；执行此命令，原有的配置将丢失。
该指令运行示例如下：（该配置文件在 Linux 上为 /etc/ha.d/haservices.xml）
注意：服务配置 IP 资源前必须先配置网卡资源。

```
cli:~>service-init

=====
      Service Initialization Utility
=====

This utility sets up the initialization information of the service in
the HA system.
It prompts you for the following information:
    - Service information
    - Application resource information
    - Public net work interface information
    - Floating IP address information.
    - Disk Mount information
    - Raw Disk information

Input service name: [q] ora 1
Is multi instance service? [y/n,q] (n) 2
Is it enabled? [y/n,q] (y) 3
Please select nodes to run the service:
    0) linux76
    1) linux77
    2) linux78
Select one or more nodes separated by spaces: [0-2,q] 0 1 24
Do you want to configure preferred node(s)? [y/n,q] (n) 5
Input start time out: [q] (240) 6
Input stop time out: [q] (360) 7
Input check interval: [q] (30) 8
Input check time out: [q] (60) 9
Input max error count: [q] (1) 10
Restart after check result is failed? [y/n,q] (n) 11
Start service anyway when float IP exist? [y/n,q] (n) 12
Service information verification:

    name: ora
    multinstance: no
    enabled: yes
    starttimeout: 240
    stoptimeout: 360
    checkinterval: 30
    checktimeout: 60
```

```
maxerrcount: 1
maxlockcount: 0
restartaftercheckfail: no
startanyway: no
nodelist: linux76 linux77 linux78
autostartlist: linux76 linux77 linux78
Is this information correct? [y/n,q] (y)13
```

服务初始化示意图1



说明:

1. 自定义服务名称，用于集群内部识别；
所谓“服务”是集群软件管理的一个资源组，可以包括业务程序、浮动 IP 地址、网络接口、存储块设备、文件系统挂载点等资源。服务名称由用户自己定义，一般和业务名称相关。比如 oracle 的服务名可以取“oracle”，tomcat 服务名取“tomcat”；
2. 是否配置多实例服务，多实例服务可以简单理解为同一业务应用同时在多个节点上运行，这要应用本身支持。
3. 是否被集群管理，默认 yes，直接回车；如果选择 no，集群启动时不会启动服务。
4. 选择运行服务的节点，简单的说就是选择运行服务的主备机；
5. 优先节点，是指集群系统自动部署服务时，将优先考虑在该节点运行，如果优先节点发生故障或者重启后重新加入集群，在非优先节点运行的服务会倒回优先节点，优先节点与服务启动优先级的区别在于前者有服务回倒功能；
6. 启动服务的最长时间，如果应用启动时间长可适当调大，比如 120 秒；
7. 停止服务的最长时间，如果应用停止时间长可适当调大，比如 180 秒；
8. 检测服务的时间间隔，也就是多长时间检测服务状态一次，默认 30 秒；
注意：如果多个服务存在依赖关系，那么这些服务的检测间隔必须一致
9. 检测服务的超时时间，也就是执行脚本多长时间没返回算检测失败，默认 60 秒；
10. 输入最大错误数，检测出错多少次开始切换；
11. Cluster 在定时检测服务失败时是否执行本地重起；如果 yes，先对服务的资源进行停止操作，然后在本机启动服务，启动失败再切换到备份节点；如果 no，先对服务的资源进行停止操作，然后在备份节点启动服务；
12. 是否设置服务的 startanyway 选项，默认为 no；选 yes，启动服务时不 ping 浮动 IP，可加快服务的启动速度，仅用于测试；
13. 确认以上信息是否正确，选择 yes(直接回车)往下执行；

Add resources for the service:¹⁴

Resource types:

- 0) NIC
- 1) IP
- 2) Application
- 3) DiskMount
- 4) RawDisk

```

b) Back to previous menu
Please select a type of resource: [0-4,b,q] 0 15
To configure NIC resource the following information is required:16

Name of resource
Critical
Enable
Type of NIC, including single and multiple
Device
One or more network host IP(thirdparty IP)

Do you want to continue? [y/n] (y) 17
Input PubNIC name: [q] (oracle_nic_0) 18
Is resource critical? [y/n,q] (y) 19
Is resource enable? [y/n,q] (y) 20
Please choose a type of NIC:21
    0) single
    1) multiple
Select a type: [0-1,q] (0) 0
Detected NICs on linux76: ens192 ens224 ens256 22
Input a network device: [q] (ens192) 23
Are you using the same configuration in the remaining nodes? [y/n] (y)24
Do you want to add network host IP? [y/n] (y) 25
Input IP address: [c,q] 192.168.8.1 26
Do you want add another network host IP? [y/n] (y) n 27
NIC resource information verification:28

    name: oracle_nic_0
    critical: yes
    enable: yes
    networkhostip: 192.168.8.1
    type: single
    nodename: linux76 device: ens192
    nodename: linux77 device: ens192
    nodename: linux78 device: ens192
Is this information correct? [y/n,q] (y)29
Add another NIC resource? [y/n] (n) 30
Do you want to add other resource? [y/n] (y)31

```

服务初始化示意图2

14. 开始配置服务的相关资源，包括服务需要使用的业务网卡、业务 IP、业务应用、共享存储资源等；
15. 选择增加网卡资源；
16. 列出网卡资源需要配置的属性/信息；
17. 询问是否开始配置，选择 y 往下执行，选择 n 回退上一步，这里选择 y；
18. 网卡资源名称，内部识别，直接回车即可；
19. 是否设定网卡资源为关键资源，所谓关键资源就是该资源异常时会引起整

- 个服务倒换，直接回车即可；
20. 是否被集群管理，默认 yes，直接回车；
 21. 可以理解为用于服务的业务链路的数量，如果只有一条链路, 选择 single，如果有多条，选择 multiple；这里只有一条业务链路 ens192, 选择 single；
 22. 列出可供选择的网卡设备名称；
 23. 选择设备 ens192；
 24. 是否把配置同步到其它节点，这里选 y；
 25. 咨询是否增加网络中存在的主机 IP，建议配置，选择 yes；
此配置的作用是检测节点本身的网络是否正常，原理是只要任一个配置的 IP 能 ping 通就说明节点本身网络正常，如都不能 ping 通，说明本身网络存在异常；
PS：如果服务需要用于网络存储（iscsi 磁阵），这里的 IP 要填写 iscsi 服务器的 IP，建议服务上的 IP 全部添加上
 26. 输入网络中存在的主机 IP（支持 ipv4 或 ipv6 地址），这个会做检验，输入不存的 IP 时会报错；可配置多个，空格隔开；
 27. 是否再增加网络中存在的主机 IP, 这里已配置一个，选择 no；
 28. 列出网卡资源配置信息列表；
 29. 确认配置信息是否正确，正确选 y，否则选 n 重新配置该资源；
 30. 是否再增加一个网卡资源，可以支持配置多这个，这里选 no；
 31. 询问是否增加服务的其它资源，这里选 yes；

Resource types:

- 0) NIC
- 1) IP
- 2) Application
- 3) DiskMount
- 4) RawDisk
- b) Back to previous menu

Please select a type of resource: [0-4,b,q] 1 ³²

To configure IP resource the following information is required: ³³

Name of resource
Critical
Enable
A NIC which floating IP attached
An IP address
Netmask or prefix

Do you want to continue? [y/n] (y) ³⁴

Input IP name: [q] (oracle_ip_0) ³⁵

Is resource critical? [y/n,q] (y) ³⁶

Is resource enable? [y/n,q] (y) ³⁷

Please choose an IP type:

- 0) IPv4
- 1) IPv6

Select a type: [0-1,q] (0) 0 ³⁸


```

Discovered NICs on linux76: ens192 ens224 ens256 39
Input a network device: [q] (ens192) 40
Input IP address: [q] 192.168.8.62 41
Input netmask: [q] (255.255.255.0) 42
Are you using the same configuration in the remaining nodes? [y/n] (y) 43
IP resource information verification:

    name: oracle_ip_0
    critical: yes
    enable: yes
    nodename: linux76
    PubNICResName: ens192
    ipaddr: 192.168.8.62
    netmask: 255.255.255.0
    nodename: linux77
    PubNICResName: ens192
    ipaddr: 192.168.8.62
    netmask: 255.255.255.0
    nodename: linux78
    PubNICResName: ens192
    ipaddr: 192.168.8.62
    netmask: 255.255.255.0
Is this information correct? [y/n,q] (y) 44
Add another IP resource? [y/n] (n) 45
Do you want to add other resource? [y/n] (y) 46

```

服务初始化示意图3

32. 选择增加 IP 资源，主要是配置业务 IP；
33. 列出 IP 资源需要配置的属性/信息；
34. 询问是否开始配置，选择 yes 往下执行；
35. IP 资源名称，内部识别，直接回车即可；
36. 是否设定为关键资源，直接回车选择 y；
37. 是否被集群管理，直接回车选择 y；
38. 选择配置的浮动 IP 的类型，支持 IPv4 和 IPv6，这里选择 IPv4；
39. 列出可供选择的网卡设备；
40. 选择业务链路网卡，这里选择 ens19；
41. 输入业务 IP，前面规划为 192. 168. 8. 62；
42. 配置业务 IP 的子网掩码；
43. 同步配置到其它节点；
44. 确认配置的 IP 资源信息是否正确，正确选择 y；
45. 是否再增加一个业务 IP 资源，可以支持配置多这个，这里选 n；
46. 询问是否增加服务的其它资源，这里选 y；

```

Resource types:
    0) NIC
    1) IP
    2) Application

```

```

3)DiskMount
4) RawDisk
b) Back to previous menu
Please select a type of resource: [0-4,b,q] 2 47

To configure application resource the following information is required:

Name of resource
Critical
Enable
A script of application
Count of fault on check time out

Do you want to continue? [y/n] (y) 48
Input application name: [q] (oracle_app_0) 49
Input script of application: [q] (/etc/ha.d/resource.d/oracle) /home/oracle.sh 50
Input the count of fault on check time out: [q] (0) 51
Is resource critical? [y/n,q] (y) 52
Is resource enable? [y/n,q] (y) 53
Application resource information verification:

name: oracle_app_0
critical: yes
enable: yes
script: /home/oracle.sh
faultonchecktimeouts: 0

Is this information correct? [y/n,q] (y) 54
Add another application resource? [y/n] (n) 55
Do you want to add other resource? [y/n] (y) 56

```

服务初始化示意图4

- 47. 选择增加应用资源，主要是配置业务应用，这里是 oracle；
- 48. 询问是否开始配置，选择 y 往下执行；
- 49. 应用资源名称，内部识别，直接回车即可；
- 50. 填写集群脚本的完整路径。所谓集群脚本，就是集群用于管控业务的脚本；
- 51. 设定业务资源检测超时出现多少次后作为错误处理，会触发业务的倒换，0 表示不作为错误处理，缺省为 0。当设定检测超时次数时，建议配合 kernel_panic 来使用，即打开 kernel_panic，参考操作：cluster-modify；
PS:当设定检测超时次数大于 1 时，前面配置的 input max error count 的值要保持缺省值 1；即两参数值不能同时修改成大于 1
- 52. 是否设定为关键资源，直接回车选择 y；
- 53. 是否被集群管理，直接回车选择 y ；
- 54. 确认配置的应用资源是否正确，正确选择 y；
- 55. 是否再增加一个应用资源，可以支持配置多这个；
- 56. 询问是否增加服务的其它资源，这里选 y；

Resource types:

- 0) NIC
- 1) IP
- 2) Application
- 3) DiskMount
- 4) RawDisk
- b) Back to previous menu

Please select a type of resource: [0-4,b,q] 4 ⁵⁷

To configure diskmount resource the following information is required:

Name of resource

Critical

Enable

Type of diskmount, including disk, NFS and LVM

A device

A mountpoint

NetStorage that used to determine whether the network storage

Fstype(file system type)

User

Group

Mode

Options

Quota of the device

stopWhenRO that used to determine whether stop service, when the file system is readonly

FsckOpt that used to determine whether use fsck to repair file system

Do you want to continue? [y/n] (y) ⁵⁸

Input diskmount name: [q] (oracle_dm_0) ⁵⁹

Is resource critical? [y/n,q] (y) ⁶⁰

Is resource enable? [y/n,q] (y) ⁶¹

- 0) disk —— 普通块设备
- 1) nfs —— NFS共享存储
- 2) lvm —— 逻辑卷
- 3) drbd —— 镜像设备

please choose a disk type: [0-2,q] 3 ⁶²

please choose a mirror disk:

- 0) r0 /dev/drbd0

select a mirror disk: [0-0,q] 0 ⁶³

Input mountpoint: [q] /home/db ⁶⁴

Input type of file system: [q] (ext3)

Input user: [q] (root) oracle

Input group: [q] (root) oinstall

Input mode: [q] (755)

Input options: [q] (rw)

Input the quota of the device: [q] (90) ⁶⁵

```
Do you want to stop service when the disk is readonly? [y/n,q] (y)66
When the file system mount failure, use fsck to repair it? [y/n,q] (n) 67
Diskmount resource information verification:
```

```
name: oracle_dm_0
critical: yes
enable: yes
type: drbd
device: /dev/drbd0
mountpoint: /home/db
fstype: ext3
user: oracle
group: oinstall
mode: 755
options: rw
limitation: 90
stopWhenRO: yes
FsckOpt: no
NetStorage: no
```

```
Is this information correct? [y/n,q] (y) 68
Add another diskmount resource? [y/n] (n) 69
Do you want to add other resource? [y/n] (y) n 70
Do you want to add another service? [y/n] (n) 71
Please run cluster-start to start the HA system,or run cluster-restart to restart the HA
system.72
cli:~>
```

服务初始化示意图5

- 57. 选择增加 DiskMount 存储资源，这里规划是/dev/sdb2 设备；
- 58. 询问是否开始配置，选择 y 往下执行；
- 59. 共享存储资源名称，内部识别，直接回车即可；
- 60. 是否设定为关键资源，直接回车选择 y；
- 61. 是否被集群管理，直接回车选择 y ；
- 62. 选择存储类型，这里是镜像,选择 drbd，其它存储类型简单说明如下：
 - 0) disk ——普通块设备（如/dev/sda）
 - 1) nfs ——nfs 存储（如 192.168.8.x:/nfsshare）
 - 2) lvm ——逻辑卷组（如/dev/vg01/lv01）
 - 3) drbd ——镜像设备（如/dev/drbd0）



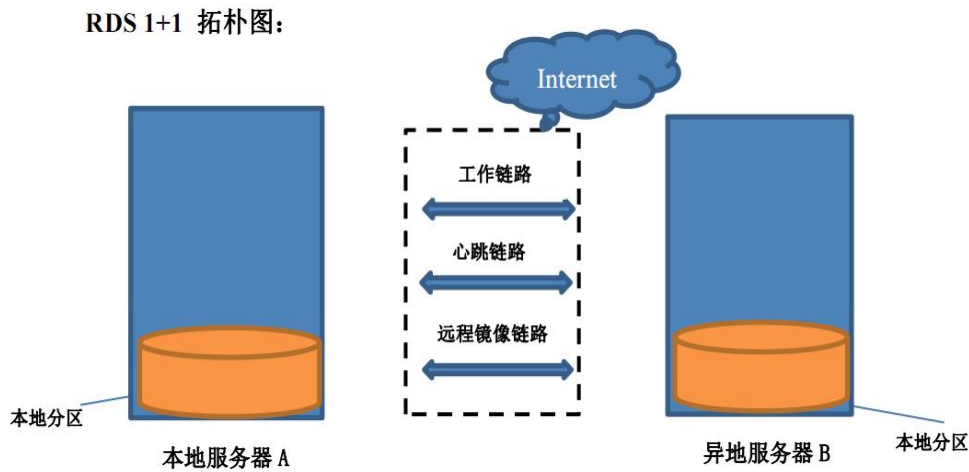
提示：针对 rds 版本，drbd 类型存储资源添加时会自动添加一个镜像服务（缺省名：MRsrvx），并且当前配置的服务会依赖该镜像服务，资源删除时同样会自动删除，此过程对使用者来说是透明的，不用关注其实现过程。

- 63. 选择要配置的镜像设备 drbd0。
- 64. 输入挂载目录及其相关信息：
mountpoint 设备挂载点（即 Linux 目录），两节点必须具有相同的挂载点；

- fstype 文件系统类型；
user 挂载目录的操作用户；
group 操作用户的群组；
mode 挂载目录的权限；
options 挂载目录对共享资源的操作权限，默认是读写；
65. 设置设备空间的使用百分比，如果使用百分比超过此值，日志输出告警信息。
 66. 当数据分区变成只读（readonly）时是否停止服务；这里补充一下，网络存储（NetStorage）不支持此属性；
 67. 当设备挂载失败时是否尝试用 fsck 去修复它；
 68. 确认配置的 DiskMount 资源是否正确；
 69. 是否再增加一个 DiskMount 资源，可以支持配置多个这个；
 70. 询问是否增加服务的其它资源，这里选 n；
 71. 询问是否增加另一个服务。
 72. 服务配置完成，启动集群生效
-

至此，软件配置完成。

4.2 远程容灾配置(RDS 1+1)



文中 RDS 1+1 的规划参考:

容灾模式	RDS 1+1
节点	本地节点: Linux76 异地节点: linux77
链路网卡	工作链路: ens192 (192.168.8.x) 心跳链路: ens224(192.168.11.x)、ens256(192.168.12.x) 镜像链路: ens224(192.168.11.x)
工作链路 IP	192.168.8.76、192.168.8.77
心跳链路 IP	单播心跳: 192.168.11.76、192.168.11.77 单播心跳: 192.168.12.76、192.168.12.77
镜像盘 (镜像资源)	r0
镜像盘底层盘	本地分区: /dev/sdb
业务服务	服务名字: ora (应用: Oracle)
业务浮动 IP(子网掩码)	192.168.8.62(255.255.255.0)
挂载点	/home/db
业务管控脚本	/home/oracle.sh

初始化配置

HA 不同类型产品的初始化流程是类似的，这有利于使用者尽快熟悉相应产品而。下面的初始化配置不再对初始化过程进行过多重复解释，仅对差异的地方进行补充说明。

初始化命令摘要（必须按以下序列号顺序执行）

序号	命令	功能描述
1	cluster-init	集群的初始化配置，格式： <code>cluster-init</code> 。
2	mirror-init	镜像资源的初始化配置，格式： <code>mirror-init</code> 。
3	service-init	服务的初始化配置，格式： <code>service-init</code> 。

cluster-init

```
cli:~>cluster-init
=====
Cluster Initialization Utility
=====

This utility sets up the initialization information of cluster. It prompts you for
the following information:
    - Hostname
    - Information about the heartbeat channels
    - How long between heartbeat
    - How long to declare heartbeat fails
    - Watchdog configuration
    - Worklink_hb configuration
    - Kernel_panic configuration
    - data and service policies

Please input cluster name: [q] RDS_cluster
Input the 1th node name and IP separated by spaces: [q] linux76
192.168.11.76
Input the 2th node name and IP separated by spaces: [q] linux77
192.168.11.77
Do you want to continue to add node? [y/n] (y)n
How long between heartbeats(in seconds): [q] (3)
How long to declare heartbeat has broken(in seconds): [q] (100)
Do you want to enable watchdog ? [y/n] (n)
Add heartbeat links: IP   type:
    0 ) IPv4
    1 ) IPv6
```

```

Select a type: [0-1,q] (0) 7
Detected NICs on cluster: ens192 ens224 ens256
Input a network device for unicast heartbeat: [q] (ens192) ens224
Input linux76 IPv4 address [q] (192.168.11.76)
Input linux77 IPv4 address [q] (192.168.11.77)
Add another ucast heartbeat link? [y/n] (y)
Input a network device for unicast heartbeat: [c,q] (ens192) ens256
Input linux76 IPv4 address [q] (192.168.12.76)
Input linux77 IPv4 address [q] (192.168.12.77)
Add another ucast heartbeat link? [y/n] (y) n
Do you want to enable worklink_hb ? [y/n] (y) n
Do you want to enable auto_failover ? [y/n] (n)
Please configure data and service policies:
    0) the maximization of data security
    1) the maximization of service availability
Please select: [0-1,q] 0
Do you want to enable kernel_panic ? [y/n] (n)
Do you want to add a lock disk(recommend) ? [y/n] (n)
Cluster information verification:
clustername: rds_cluster
    node: linux76 192.168.11.76
    node: linux77 192.168.11.77
    logfile: /var/log/ha-log
    keepalive: 3
    deadtime: 100
    watchdog: off
    ucast: ens224 192.168.11.76 192.168.11.77
    ucast: ens256 192.168.12.76 192.168.12.77
    worklink_hb: off
    auto_failover: off
    auto_run: on
    maximization: data-security
    kernel_panic: off
Is this information correct? [y/n,q] (y)
Please run mirror-init to initialize your mirror resource.

```



说明:

- 1、RDS 1+1 仅需配置 2 个节点，不需要进行域节点配置。
- 2、本地节点与异地节点的 IP 可以属于不同网段，保证相互之间能通讯即可。

mirror-init

镜像资源初始化过程与前面 [RDS 2+1 的 mirror-init](#) 类似，可直接参考，这里不重复。

service-init

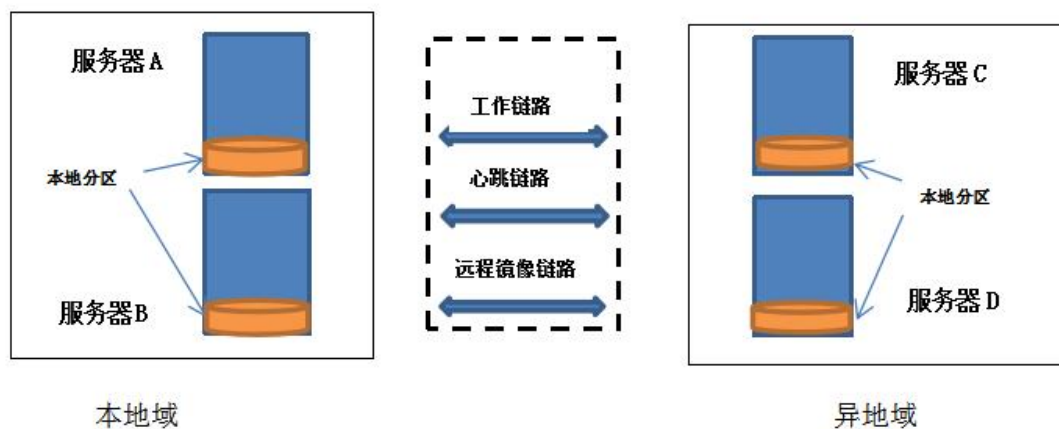
服务初始化过程与前面 [RDS 2+1 的 service-init](#) 类似，可直接参考，这里不重复。

4.3 远程容灾配置($n*(1+1)$)

此场景是远程容灾中比较特殊的场景，支持配置场景为：本地域每台服务器配置一个业务应用，并且业务间存依赖关系，当任一节点上的业务发生故障时，本地所有业务应用一并倒换到远端来保证业务系统的持续运行。

下面以 $2*(1+1)$ 场景配置为例进行说明：

异地 $2*(1+1)$ 拓扑图：



文中异地 2*（1+1）配置规划参考：

容灾模式	2*(1+1)
节点	本地域节点：Linux70、linux71 异地域节点：linux72、linux73
链路网卡	工作链路：ens192 (192.168.8.x) 心跳链路：ens224(192.168.11.x)、ens256(192.168.12.x) 镜像链路：ens224(192.168.11.x)
工作链路 IP	192.168.8.70、192.168.8.71、192.168.8.72、192.168.8.73
心跳链路 IP（单播）	192.168.11.70、192.168.11.71、192.168.11.72、192.168.11.73 192.168.12.70、192.168.12.71、192.168.12.72、192.168.12.73
镜像盘（镜像资源）	r0:linux70、linux72 （业务：oracle） r1:linux71、linux73 （业务：管理系统）
镜像盘底层盘	Linux70:/dev/sdb Linux71:/dev/sdb Linux72:/dev/sdb Linux73:/dev/sdb
业务服务	服务名字：ora（应用：Oracle）、man_sys(应用：管理系统) 其中 man_sys 依赖 ora
业务运行节点	Oracle(ora)：linux70、linux72 管理系统（man_sys）：linux71、linux73
业务浮动 IP(子网掩码)	Oracle:192.168.8.22(255.255.255.0) Man_sys:192.168.8.23(255.255.255.0)
挂载点	/home/db /home/man_sys_data
业务管控脚本	/home/oracle.sh /home/man_sys_data

初始化配置

HA 不同类型产品的初始化流程是类似的，这有利于使用者尽快熟悉相应产品。

初始化命令摘要（必须按以下序列号顺序执行）

序号	命令	功能描述
1	cluster-init	集群的初始化配置，格式：cluster-init。
2	mirror-init	镜像资源的初始化配置，格式：mirror-init。
3	service-init	服务的初始化配置，格式：service-init。

cluster-init

集群初始化过程与前面 [RDS 2+1 的 cluster-init](#) 类似,配置参数说明可以直接参考,因场景特殊,这里还是把过程重新演示一次。

```
cli:~>cluster-init
```

```
=====
Cluster Initialization Utility
=====
```

```
Built: Mon Mar 16 15:07:43 2020
```

This utility sets up the initialization information of cluster.

It prompts you for the following information:

- Hostname
- Information about the heartbeat channels
- How long between heartbeat
- How long to declare heartbeat fails
- Watchdog configuration
- Worklink_hb configuration
- Kernel_panic configuration
- data and service policies

Please input cluster name: [q] cluster

Input the 1st node name and IP separated by spaces: [q] **linux70**
192.168.11.70 192.168.12.70

Input the 2nd node name and IP separated by spaces: [q] **linux71**
192.168.11.71 192.168.12.71

Do you want to continue to add node? [y/n] (y)

Input the 3rd node name and IP separated by spaces: [c,q] **linux72**
192.168.11.72 192.168.12.72

Do you want to continue to add node? [y/n] (y) n

Input the 4th node name and IP separated by spaces: [c,q] **linux73**
192.168.11.73 192.168.12.73

Do you want to continue to add node? [y/n] (y) n

How long between heartbeats(in seconds): [q] (3)

How long to declare heartbeat has broken(in seconds): [q] (100)

Do you want to enable watchdog ? [y/n] (n)

Add heartbeat links:

IP type:

- 0) IPv4
- 1) IPv6

```

Select a type: [0-1,q] (0)
Detected NICs on cluster: ens192 ens224 ens256
Input a network device for unicast heartbeat: [c,q] (ens224)
Input linux70 IPv4 address [q] (192.168.11.70)
Input linux71 IPv4 address [q] (192.168.11.71)
Input linux72 IPv4 address [q] (192.168.11.72)
Input linux73 IPv4 address [q] (192.168.11.73)
Add another ucast heartbeat link? [y/n] (y)
Input a network device for unicast heartbeat: [c,q] (ens256)
Input linux70 IPv4 address [q] (192.168.12.70)
Input linux71 IPv4 address [q] (192.168.12.71)
Input linux72 IPv4 address [q] (192.168.12.72)
Input linux73 IPv4 address [q] (192.168.12.73)
Add another ucast heartbeat link? [y/n] (y) n
Do you want to enable worklink_hb ? [y/n] (y)
Do you want to enable auto_failover ? [y/n] (n)
Please configure data and service policies:
    0. the maximization of data security
    1. the maximization of service availability
Please select: [0-1,q] 0
Do you want to enable kernel_panic ? [y/n] (n)
Do you want to add domain-node ? [y/n] (y)
Please select nodes for the 1th domain-node:
    0) linux70
    1) linux71
    2) linux72
    3) linux73
Select one or more nodes separated by spaces: [0-3,q] 0 1
Do you want to continue to add domain-node? [y/n] (y)
Please select nodes for the 2th domain-node:
    0) linux70
    1) linux71
    2) linux72
    3) linux73
Select one or more nodes separated by spaces: [0-3,q] 2 3
Do you want to continue to add domain-node? [y/n] (n)
Do you want to add a lock disk(recommend) ? [y/n] (n)
Cluster information verification:

    clustername: cluster
    node: linux70 192.168.11.70 192.168.12.70

```

```
node: linux71 192.168.11.71 192.168.12.71
node: linux72 192.168.11.72 192.168.12.72
node: linux73 192.168.11.73 192.168.12.73
logfile: /var/log/ha-log
keepalive: 3
deadtime: 100
watchdog: off
ucast: ens192 192.168.8.70 192.168.8.71 192.168.8.72
192.168.8.73
ucast: ens224 192.168.11.70 192.168.11.71 192.168.11.72
192.168.11.73
ucast: ens256 192.168.12.70 192.168.12.71 192.168.12.72
192.168.12.73
worklink_hb: on 226.0.114.150 ff18::a434:3392
auto_failover: off
auto_run: on
maximization: data-security
kernel_panic: off
domain_node: linux70 linux71
domain_node: linux72 linux73

Is this information correct? [y/n,q] (y)
Please run mirror-init to initialize your mirror resource.
cli:~>
```



说明:

- 1、2*（1+1）同样需要配置域，按规划这里配置两个域，按本地和远端的划分。
- 2、本地节点与异地节点的 IP 可以属于不同网段，保证相互之间能通讯即可。

mirror-init

镜像初始化过程与前面 [RDS 2+1 的 mirror-init](#) 类似,配置参数说明可以直接参考,因场景特殊,这里还是把过程重新演示一次。

```
cli:~>mirror-init

=====
Mirror Initialization Utility
=====

It prompts you for the following information:
- Mirror resource name
- Mirror disk device name
- Mirror IP address
- Filesystem type
Input resource name: [q] (r0)
Input port number: [1-65535,q] (7788)
Please select nodes to run mirror r0:
    0) linux70
    1) linux71
    2) linux72
    3) linux73
Select one or more nodes separated by spaces: [0-3,q] 0 2
Input mirror IP address for linux70: [q] 192.168.11.70
Input a disk used to store data on linux70: [q] /dev/sdb
Input mirror IP address for linux72: [q] 192.168.11.72
Input a disk used to store data on linux72: [q] /dev/sdb
Input memlimit for proxy r0 [500M/G]: [q] (500M)
select data compression policy for proxy
    0) low compression
    1) high compression
    2) no compression
Please select an item: [0-2,q] (2)
DRBD resource information verification:

    name: r0
    device: /dev/drbd0
    port: 7788
    node: linux70
        disk: /dev/sdb
        meta-disk: internal
        IP address: 192.168.11.70
    node: linux72
        disk: /dev/sdb
        meta-disk: internal
```

```

        IP address: 192.168.11.72
        proxy memlimit: 500M
        proxy plugin: none

Is this information correct? [y/n,q] (y)
Do you want to add another mirror resource? [y/n] (n) y
Input resource name: [q] (r1)
Input port number: [1-65535,q] (7789)
Please select nodes to run mirror r1:
    0) linux70
    1) linux71
    2) linux72
    3) linux73
Select one or more nodes separated by spaces: [0-3,q] 1 3
Input mirror IP address for linux71: [q] 192.168.11.71
Input a disk used to store data on linux71: [q] /dev/sdb
Input mirror IP address for linux73: [q] 192.168.11.73
Input a disk used to store data on linux73: [q] /dev/sdb
Input memlimit for proxy r1 [500M/G]: [q] (500M)
select data compression policy for proxy
    0) low compression
    1) high compression
    2) no compression
Please select an item: [0-2,q] (2)
DRBD resource information verification:

    name: r1
    device: /dev/drbd1
    port: 7789
    node: linux71
        disk: /dev/sdb
        meta-disk: internal
        IP address: 192.168.11.71
    node: linux73
        disk: /dev/sdb
        meta-disk: internal
        IP address: 192.168.11.73
    proxy memlimit: 500M
    proxy plugin: none

Is this information correct? [y/n,q] (y)
Do you want to add another mirror resource? [y/n] (n)

All the data of /dev/sdb on node linux70 will be destroyed. Are you sure
to continue? [y/n] (y)

All the data of /dev/sdb on node linux72 will be destroyed. Are you sure

```

```

to continue? [y/n] (y)

Creating meta-data for resource r0...

Create meta-data for resource r0 success

All the data of /dev/sdb on node linux71 will be destroyed. Are you sure
to continue? [y/n] (y)

All the data of /dev/sdb on node linux73 will be destroyed. Are you sure
to continue? [y/n] (y)

Creating meta-data for resource r1...

Create meta-data for resource r1 success

Waiting for mirror proxy starting...
start proxy success.

Waiting for mirror starting...
start mirror success.
    0) linux70
    1) linux72
Please choose the sync source for resource r0: [0-1,q] 0
    0) linux71
    1) linux73
Please choose the sync source for resource r1: [0-1,q] 0

Which filesystem type that you want to format /dev/drbd0(mirror
resource: r0) to:
    0) ext2
    1) ext3
    2) ext4
    3) xfs
    4) none
Select a filesystem type [0, 4]:2

mke2fs 1.42.9 (28-Dec-2013)
Filesystem label=
OS type: Linux
Block size=4096 (log=2)
Fragment size=4096 (log=2)
Stride=0 blocks, Stripe width=0 blocks
131072 inodes, 524263 blocks
26213 blocks (5.00%) reserved for the super user
First data block=0
Maximum filesystem blocks=536870912

```



```

16 block groups
32768 blocks per group, 32768 fragments per group
8192 inodes per group
Superblock backups stored on blocks:
    32768, 98304, 163840, 229376, 294912

Allocating group tables: done
Writing inode tables: done
Creating journal (8192 blocks): done
Writing superblocks and filesystem accounting information: done


Format disk /dev/drbd0 to ext4 success

Which filesystem type that you want to format /dev/drbd1(mirror
resource: r1) to:
    0) ext2
    1) ext3
    2) ext4
    3) xfs
    4) none
Select a filesystem type [0, 4]:2

mke2fs 1.42.9 (28-Dec-2013)
Filesystem label=
OS type: Linux
Block size=4096 (log=2)
Fragment size=4096 (log=2)
Stride=0 blocks, Stripe width=0 blocks
131072 inodes, 524263 blocks
26213 blocks (5.00%) reserved for the super user
First data block=0
Maximum filesystem blocks=536870912
16 block groups
32768 blocks per group, 32768 fragments per group
8192 inodes per group
Superblock backups stored on blocks:
    32768, 98304, 163840, 229376, 294912

Allocating group tables: done
Writing inode tables: done
Creating journal (8192 blocks): done
Writing superblocks and filesystem accounting information: done

```

```

Format disk /dev/drbd1 to ext4 success

Waiting for r0 synchronization...
r0 Resynced percentage: 100.0 %

Waiting for r1 synchronization...
r1 Resynced percentage: 100.0 %
set secondary all successful

Mirror init success!
Please run service-init to initialize you services.
cli:~>

```

NOTE

说明:

- 1、按规划，这里添加两个镜像盘，其中 r0 属于 oracle (linux70、linux72)，r1 属于管理系统 (linux71、linux73)
- 2、如果盘同步时间过长，可以暂时先跳过同步环节，具体操作方式联系我们。

service-init

服务初始化过程与前面 [RDS 2+1 的 service-init](#) 类似,配置参数说明可以直接参考，因场景特殊，这里还是把过程重新演示一次。

```

cli:~>service-init

=====
      Service Initialization Utility
=====

This utility sets up the initialization information of the service in
the HA system.
It prompts you for the following information:
    - Service information
    - Application resource information
    - Public net work interface information
    - Floating IP address information.
    - Disk Mount information
    - Raw Disk information

Input service name: [q] ora
Is multi instance service? [y/n,q] (n)
Is it enabled? [y/n,q] (y)

```

Please select nodes to run the service:

- 0) linux70**
- 1) linux71
- 2) linux72**
- 3) linux73

Select one or more nodes separated by spaces: [0-3,q] **0 2**

Do you want to configure preferred node(s)? [y/n,q] (n)

Input start time out: [q] (240)

Input stop time out: [q] (360)

Input check interval: [q] (30)

Input check time out: [q] (60)

Input max error count: [q] (1)

Restart after check result is failed? [y/n,q] (n)

Start service anyway when float IP exist? [y/n,q] (n)

Service information verification:

name: ora

multinstance: no

enabled: yes

starttimeout: 240

stoptimeout: 360

checkinterval: 30

checktimeout: 60

maxerrcount: 1

maxlockcount: 0

restartaftercheckfail: no

startanyway: no

odelist: linux70 linux72

autostartlist: linux70 linux72

Is this information correct? [y/n,q] (y)

Add resources for the service:

Resource types:

- 0) NIC
- 1) IP
- 2) Application
- 3) DiskMount
- 4) RawDisk
- b) Back to previous menu

Please select a type of resource: [0-4,b,q] 0

To configure NIC resource the following information is required:

Name of resource

Critical

Enable

Type of NIC, including single and multiple

```

Device
One or more network host IP(thirdparty IP)

Do you want to continue? [y/n] (y)
Input PubNIC name: [q] (ora_nic_0)
Is resource critical? [y/n,q] (y)
Is resource enable? [y/n,q] (y)
Please choose a type of NIC:
    0) single
    1) multiple
Select a type: [0-1,q] (0)
Detected NICs on linux70: ens192 ens224 ens256
Input a network device: [q] ens192
Are you using the same configuration in the remaining nodes? [y/n] (y)
Do you want to add network host IP? [y/n] (n)
NIC resource information verification:

    name: ora_nic_0
    critical: yes
    enable: yes
    type: single
    nodename: linux70    device: ens192
    nodename: linux72    device: ens192

Is this information correct? [y/n,q] (y)
Add another NIC resource? [y/n] (n)
Do you want to add other resource? [y/n] (y)
Resource types:
    0) NIC
    1) IP
    2) Application
    3) DiskMount
    4) RawDisk
    b) Back to previous menu
Please select a type of resource: [0-4,b,q] 1

To configure IP resource the following information is required:

    Name of resource
    Critical
    Enable
    A NIC which floating IP attached
    An IP address
    Netmask or prefix

Do you want to continue? [y/n] (y)
Input IP name: [q] (ora_ip_0)

```

```

Is resource critical? [y/n,q] (y)
Is resource enable? [y/n,q] (y)
Please choose an IP type:
    0) IPv4
    1) IPv6
Select a type: [0-1,q] (0)
Detected NICs on linux70: ens192 ens224 ens256
Input a network device: [q] ens192
Input IP address: [q] 192.168.8.22
Input netmask: [q] (255.255.255.0)
Are you using the same configuration in the remaining nodes? [y/n] (y)
IP resource information verification:

    name: ora_ip_0
    critical: yes
    enable: yes
    nodename: linux70
    PubNICResName: ens192
    ipaddr: 192.168.8.22
    netmask: 255.255.255.0
    nodename: linux72
    PubNICResName: ens192
    ipaddr: 192.168.8.22
    netmask: 255.255.255.0

Is this information correct? [y/n,q] (y)
Add another IP resource? [y/n] (n)
Do you want to add other resource? [y/n] (y)
Resource types:
    0) NIC
    1) IP
    2) Application
    3) DiskMount
    4) RawDisk
    b) Back to previous menu
Please select a type of resource: [0-4,b,q] 2

To configure application resource the following information is required:

    Name of resource
    Critical
    Enable
    A script of application
    Count of fault on check time out

Do you want to continue? [y/n] (y)
Input application name: [q] (ora_app_0)

```

```

Input script of application: [q] (/etc/ha.d/resource.d/ora)
/home/oracle.sh
Input the count of fault on check time out: [q] (0)
Is resource critical? [y/n,q] (y)
Is resource enable? [y/n,q] (y)
Application resource information verification:

    name: ora_app_0
    critical: yes
    enable: yes
    script: /home/oracle.sh
    faultonchecktimeouts: 0

Is this information correct? [y/n,q] (y)
Add another application resource? [y/n] (n)
Do you want to add other resource? [y/n] (y)
Resource types:
    0) NIC
    1) IP
    2) Application
    3) DiskMount
    4) RawDisk
    b) Back to previous menu
Please select a type of resource: [0-4,b,q] 3

To configure diskmount resource the following information is required:

    Name of resource
    Critical
    Enable
    A device
    A mountpoint
    NetStorage that used to determine whether the network storage
    Fstype(file system type)
    User
    Group
    Mode
    Options
    Quota of the device
    stopWhenRO that used to determine whether stop service, when the
file system is readonly
    FsckOpt that used to determine whether use fsck to repair file
system

Do you want to continue? [y/n] (y)
Input diskmount name: [q] (ora_dm_0)
Is resource critical? [y/n,q] (y)

```

```

Is resource enable? [y/n,q] (y)
    0) disk
    1) nfs
    2) lvm
    3) drbd
please choose a disk type: [0-3,q] 3
please choose a mirror disk
    0) r0 /dev/drbd0
    1) r1 /dev/drbd1
select a mirror disk: [0-1,q] 0
Input mountpoint: [q] /home/db
Input type of file system: [q] (ext3) ext4
Input user: [q] (root) oracle
Input group: [q] (root) oinstall
Input mode: [q] (755)
Input options: [q] (rw)
Input the quota of the device: [q] (90)
Do you want to stop service when the disk is readonly? [y/n,q] (y)
When the file system mount failure, use fsck to repair it? [y/n,q] (n)
Diskmount resource information verification:

    name: ora_dm_0
    critical: yes
    enable: yes
    type: drbd
    device: /dev/drbd0
    mountpoint: /home/db
    fstype: ext4
    user: oracle
    group: oinstall
    mode: 755
    options: rw
    limitation: 90
    stopWhenRO: yes
    FsckOpt: no
    NetStorage: no
    notcheck: no

Is this information correct? [y/n,q] (y)
Add another diskmount resource? [y/n] (n)
Do you want to add other resource? [y/n] (y) n
Do you want to add another service? [y/n] (n) y
Input service name: [q] man_sys
Is multi instance service? [y/n,q] (n)
Is it enabled? [y/n,q] (y)
Please select nodes to run the service:
    0) linux70

```

- 1) linux71
- 2) linux72
- 3) linux73

Select one or more nodes separated by spaces: [0-3,q] 1 3

Do you want to configure preferred node(s)? [y/n,q] (n)

Input start time out: [q] (240)

Input stop time out: [q] (360)

Input check interval: [q] (30)

Input check time out: [q] (60)

Input max error count: [q] (1)

Restart after check result is failed? [y/n,q] (n)

Start service anyway when float IP exist? [y/n,q] (n)

Service information verification:

```

name: man_sys
multinstance: no
enabled: yes
starttimeout: 240
stoptimeout: 360
checkinterval: 30
checktimeout: 60
maxerrcount: 1
maxlockcount: 0
restartaftercheckfail: no
startanyway: no
nodelist: linux71 linux73
autostartlist: linux71 linux73

```

Is this information correct? [y/n,q] (y)

Add resources for the service:

Resource types:

- 0) NIC
- 1) IP
- 2) Application
- 3) DiskMount
- 4) RawDisk
- b) Back to previous menu

Please select a type of resource: [0-4,b,q] 0

To configure NIC resource the following information is required:

```

Name of resource
Critical
Enable
Type of NIC, including single and multiple
Device
One or more network host IP(thirdparty IP)

```



```

Do you want to continue? [y/n] (y)
Input PubNIC name: [q] (man_sys_nic_0)
Is resource critical? [y/n,q] (y)
Is resource enable? [y/n,q] (y)
Please choose a type of NIC:
    0) single
    1) multiple
Select a type: [0-1,q] (0)
Detected NICs on linux71: ens192 ens224 ens256
Input a network device: [q] ens192
Are you using the same configuration in the remaining nodes? [y/n] (y)
Do you want to add network host IP? [y/n] (n)
NIC resource information verification:

    name: man_sys_nic_0
    critical: yes
    enable: yes
    type: single
    nodename: linux71    device: ens192
    nodename: linux73    device: ens192

Is this information correct? [y/n,q] (y)
Add another NIC resource? [y/n] (n)
Do you want to add other resource? [y/n] (y)
Resource types:
    0) NIC
    1) IP
    2) Application
    3) DiskMount
    4) RawDisk
    b) Back to previous menu
Please select a type of resource: [0-4,b,q] 1

To configure IP resource the following information is required:

    Name of resource
    Critical
    Enable
    A NIC which floating IP attached
    An IP address
    Netmask or prefix

Do you want to continue? [y/n] (y)
Input IP name: [q] (man_sys_ip_0)
Is resource critical? [y/n,q] (y)
Is resource enable? [y/n,q] (y)

```

```

Please choose an IP type:
    0) IPv4
    1) IPv6
Select a type: [0-1,q] (0)
Detected NICs on linux71: ens192 ens224 ens256
Input a network device: [q] ens192
Input IP address: [q] 192.168.8.23
Input netmask: [q] (255.255.255.0)
Are you using the same configuration in the remaining nodes? [y/n] (y)
IP resource information verification:

    name: man_sys_ip_0
    critical: yes
    enable: yes
    nodename: linux71
    PubNICResName: ens192
    ipaddr: 192.168.8.23
    netmask: 255.255.255.0
    nodename: linux73
    PubNICResName: ens192
    ipaddr: 192.168.8.23
    netmask: 255.255.255.0

Is this information correct? [y/n,q] (y)
Add another IP resource? [y/n] (n)
Do you want to add other resource? [y/n] (y)
Resource types:
    0) NIC
    1) IP
    2) Application
    3) DiskMount
    4) RawDisk
    b) Back to previous menu
Please select a type of resource: [0-4,b,q] 2

To configure application resource the following information is required:

    Name of resource
    Critical
    Enable
    A script of application
    Count of fault on check time out

Do you want to continue? [y/n] (y)
Input application name: [q] (man_sys_app_0)
Input script of application: [q] (/etc/ha.d/resource.d/man_sys)
/home/man_sys.sh

```

Input the count of fault on check time out: [q] (0)

Is resource critical? [y/n,q] (y)

Is resource enable? [y/n,q] (y)

Application resource information verification:

```
name: man_sys_app_0
critical: yes
enable: yes
script: /home/man_sys.sh
faultonchecktimeouts: 0
```

Is this information correct? [y/n,q] (y)

Add another application resource? [y/n] (n)

Do you want to add other resource? [y/n] (y)

Resource types:

- 0) NIC
- 1) IP
- 2) Application
- 3) DiskMount
- 4) RawDisk
- b) Back to previous menu

Please select a type of resource: [0-4,b,q] 2

To configure application resource the following information is required:

```
Name of resource
Critical
Enable
A script of application
Count of fault on check time out
```

Do you want to continue? [y/n] (y)

Add another application resource? [y/n] (n)

Do you want to add other resource? [y/n] (y)

Resource types:

- 0) NIC
- 1) IP
- 2) Application
- 3) DiskMount
- 4) RawDisk
- b) Back to previous menu

Please select a type of resource: [0-4,b,q] 3

To configure diskmount resource the following information is required:

```
Name of resource
Critical
```

```

    Enable
    A device
    A mountpoint
    NetStorage that used to determine whether the network storage
    Fstype(file system type)
    User
    Group
    Mode
    Options
    Quota of the device
    stopWhenRO that used to determine whether stop service, when the
file system is readonly
    FsckOpt that used to determine whether use fsck to repair file
system

Do you want to continue? [y/n] (y)
Input diskmount name: [q] (man_sys_dm_0)
Is resource critical? [y/n,q] (y)
Is resource enable? [y/n,q] (y)
    0) disk
    1) nfs
    2) lvm
    3) drbd
please choose a disk type: [0-3,q] 3
please choose a mirror disk
    0) r1 /dev/drbd1
select a mirror disk: [0-0,q] 0
Input mountpoint: [q] /home/man_sys_data/
Input type of file system: [q] (ext3)
Input user: [q] (root)
Input group: [q] (root)
Input mode: [q] (755)
Input options: [q] (rw)
Input the quota of the device: [q] (90)
Do you want to stop service when the disk is readonly? [y/n,q] (y)
When the file system mount failure, use fsck to repair it? [y/n,q] (n)
Diskmount resource information verification:

    name: man_sys_dm_0
    critical: yes
    enable: yes
    type: drbd
    device: /dev/drbd1
    mountpoint: /home/man_sys_data/
    fstype: ext3
    user: root
    group: root

```

```

mode: 755
options: rw
limitation: 90
stopWhenR0: yes
FsckOpt: no
NetStorage: no
notcheck: no

Is this information correct? [y/n,q] (y)
Add another diskmount resource? [y/n] (n)
Do you want to add other resource? [y/n] (y)
Resource types:
    0) NIC
    1) IP
    2) Application
    3) DiskMount
    4) RawDisk
    b) Back to previous menu
Please select a type of resource: [0-4,b,q] b
Do you want to add another service? [y/n] (n)
Please run cluster-start to start the HA system,
or run cluster-restart to restart the HA system.
cli:~>

```



说明:

- 1、这里添加两个应用，分别为 ora 和 man_sys。
- 2、重点关注规划中指定业务的运行节点及其对应的存储，如：
 - 1) ora 使用镜像盘是 r0，业务运行节点为 linux70 和 linux72
 - 2) man_sys 使用镜像盘是 r1，业务运行节点为 linux71 和 linux72

最后，再按规划进行业务依赖关系的配置，依赖原则是父依赖子，模式须选择：**remote-force**

```

cli:~>service-depend
    0) Link (add service dependency)
    1) Unlink (delete service dependency)
    2) Display (display service dependency)
    b) Back to previous menu
Please select an operation: [0-2,b,q] 0

To configure service dependency the following information is required:

    Parent service
    Child service
    Type of dependency, including fixed and forced
    Relationship, including local, remote and global

```

Notice: The parent depends on the child being running before going running itself.

Conversely, parent stops before child does so.

Do you want to continue? [y/n] (y)

0) ora

1) man_sys

Please select a parent service: [0-1,c,q] 1

Please select a child service: [0-1,c,q] 0

Relationship:

0) global

1) local

2) remote

Please select a kind of relationship: [0-2,q] 2

Dependency type:

0) forced

Please select a type of dependency: [0-0,q] 0

Do you want to configure another service dependency? [y/n] (n)

Service dependency information verification:

Parent:man_sys Child:ora Relationship:remote

Dependency-type:forced

Is this information correct? [y/n,q] (y)

0) Link (add service dependency)

1) Unlink (delete service dependency)

2) Display (display service dependency)

b) Back to previous menu

Please select an operation: [0-2,b,q] q

Are you sure you want to quit? [y/n] (y)

cli:~>

至此，产品配置完成，其命令及使用习惯与其它产品基本一致，这里不再细说。

4.4 集群启动

集群初始化配置完成后就可以进行集群启动，这里有几点需要注意：

- *) 集群启动，任一节点上执行命令即可
- *) 集群管控的服务（资源组）会自动随集群启动
- *) **RDS 2+1** 的服务间启动的方式是串行（即逐个服务启），缺省情况下按服务配置先后顺序，**RDS 1+1** 缺省无此限制

```
cli:~>cluster-start
Configuration file checked ok.
[linux76]Starting High-Availability services:
Done.
.
[linux77]Starting High-Availability services:
Configuration file checked ok.
Done.
.
```

第 5 章 通过 web 进行配置

前期准备：

- 1) 确认 web 组件后台服务运行正常：

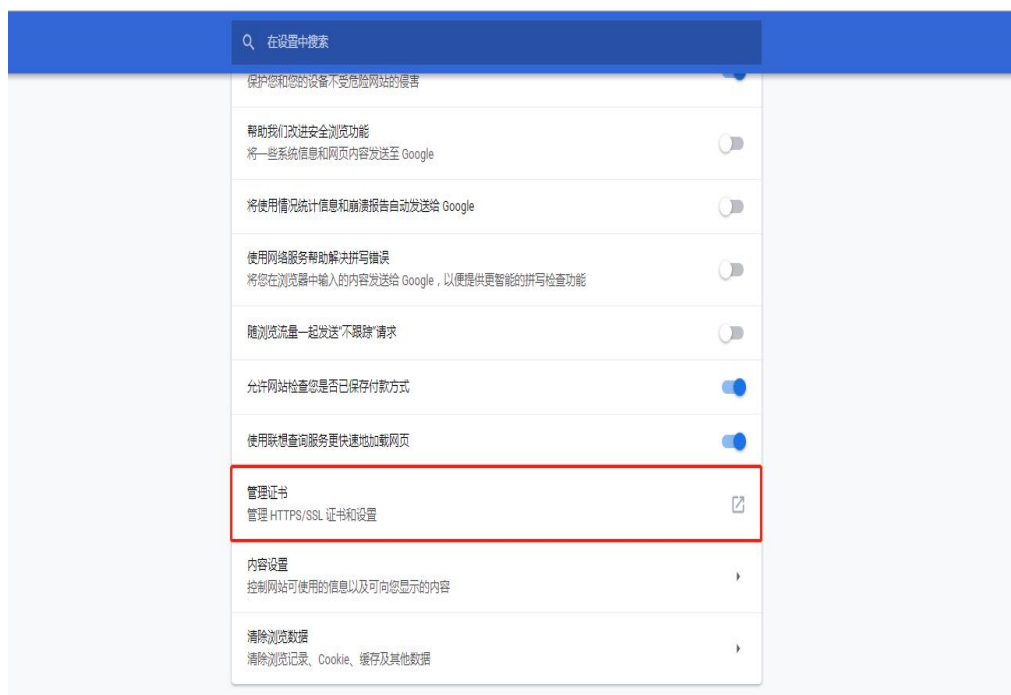
```
# webadminserver start
Starting NewstartHA webadmin server.....
# webadminserver status
NewstartHA webadmin server is running...
```

- 2) 访问 web 管理页面前需要导入安全证书（HA 采用双向加密通讯方式），具体操作下文会有说明，安全证书位置：/opt/NewStartHA/web/nginx/conf/mykey.p12
- 3) 推荐浏览器：Chrome、火狐、IE10。

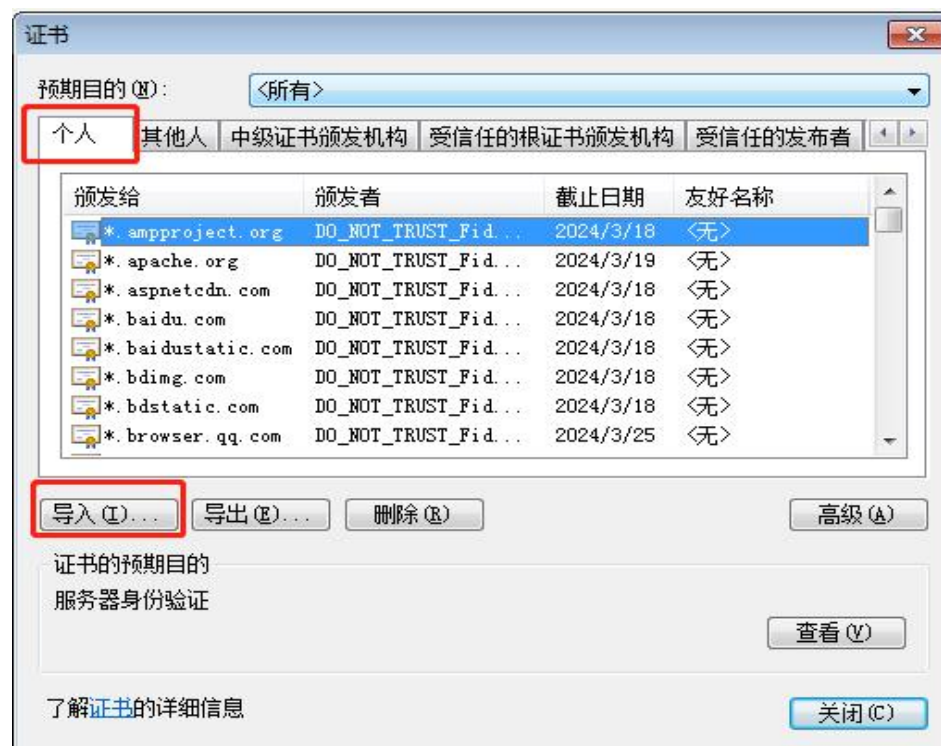
5.1 Web 登陆

5.1.1 浏览器加载证书

以谷歌浏览器为例：找到“设置——高级——管理证书”



选择“个人”，点击“导入”



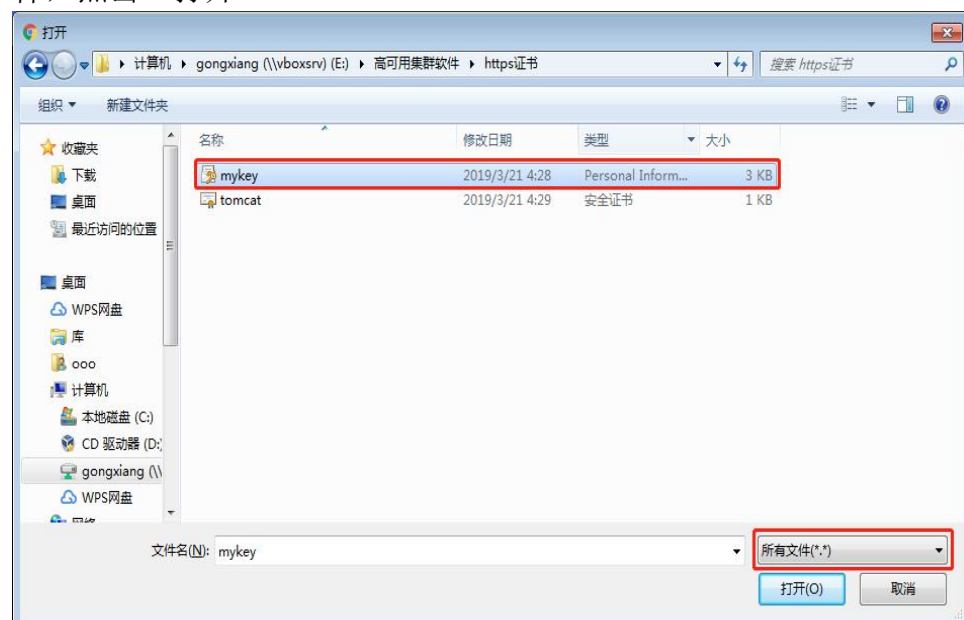
点击“下一步”



点击“浏览”



找到下载好的证书文件路径，下拉框选择“所有文件”，找到“mykey”证书文件，点击“打开”



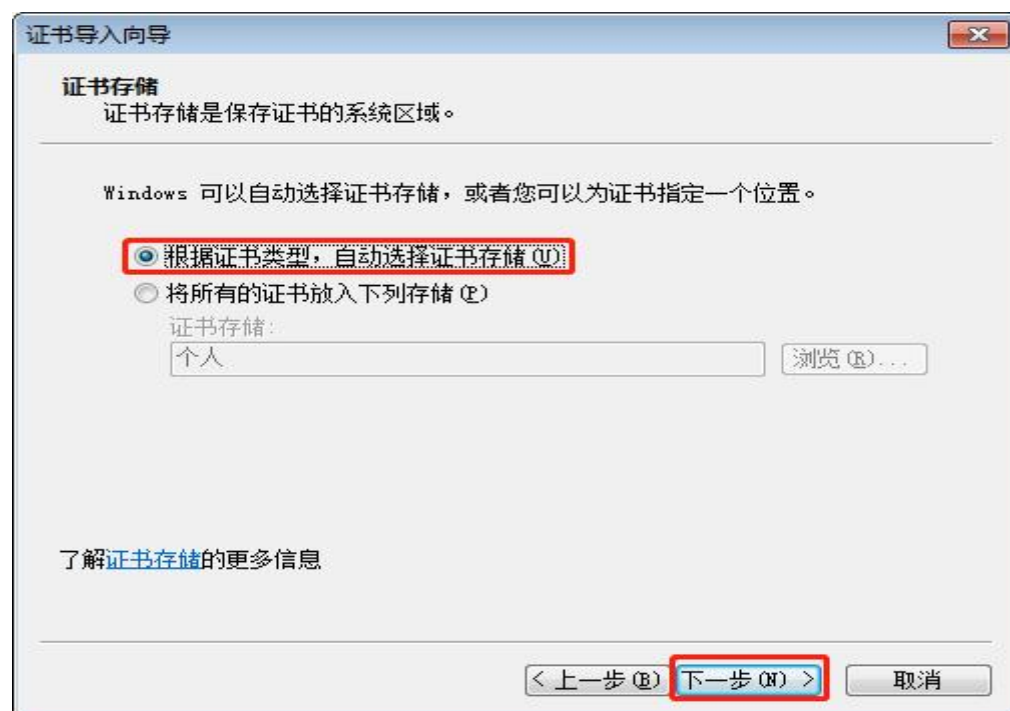
点击“下一步”



输入证书密码“HA@NewStart123_”，点击“下一步”



选择“根据证书类型，自动选择证书存储”，点击“下一步”



点击“完成”



导入成功



5.1.2 访问 web 管理页面

重新打开浏览器，按以下要求输入地址，登陆 web 界面：

https://节点物理 IP:10003（如：**https://192.168.1.19:10003**），弹出提示后点开“隐藏详情”，点击“继续前往”



弹出提示后选择证书，点击确认



进入登陆界面，如下：



The image shows the login interface for NewStart HA. It features a blue header with the NewStart logo and the text "登录 NewStart HA" and "请输入您的用户名和密码登录管理界面。". Below the header, there are two input fields: "用户名" (Username) with the value "SuperAdmin" and "密码" (Password) with masked characters. There are also radio buttons for "中文" (Chinese) and "English", and a "登录" (Login) button. At the bottom, there is a copyright notice: "© 2004 - 2019 广东中兴新支点技术有限公司, 版权所有。"

Note:

- 1) 缺省 web 用户名/密码: “SuperAdmin / HA@newstart123_”
- 2) 出现访问异常时，检查防火墙的影响和端口是否被占用。
- 3) web 时效性: 设定 10 分钟无命令调取或者无操作后，页面会失效，需要刷新浏览器，页面会自动返回登录页面，重新登录即可。

5.2 RDS 2+1 初始化配置

5.2.1 集群初始化

集群基本配置：

NewStart HA
5.0.6-20190625.x86_64
Product: Type = Remote disaster recovery

基本配置信息

集群基本配置

集群参数配置

集群名	rds_cluster		
心跳间隔时间(1-10秒)	3	工作链路心跳	<input checked="" type="checkbox"/>
宕机间隔时间(15-300秒)	100	内核恐慌	<input type="checkbox"/>
最大化策略	数据安全最大化	watchdog检测配置	<input type="checkbox"/>
故障切换	手动	仲裁盘 (选配)	

集群主机

主机名	主机IP	删除
linux76	192.168.8.76	<input type="button" value="0"/>
linux77	192.168.8.77	<input type="button" value="0"/>
linux78	192.168.8.78	<input type="button" value="0"/>

节点域

节点域	仲裁盘	删除
linux76 linux77	/dev/sdc1	<input type="button" value="0"/>
linux78		<input type="button" value="0"/>

保存

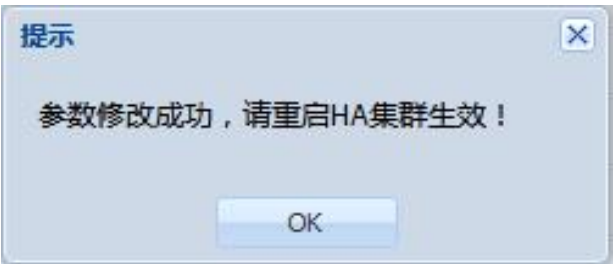
Note:

集群名：内部识别，自行定义

主机名：若类型为 2+2，则在集群主机和节点域加 1 个节点；

仲裁盘：用于在心跳全部中断时保证数据安全的一种解决方案，推荐使用，节点域下的仲裁盘类似，也是选配；

配置完成后，点击保存，提示修改成功，首次配置不用重启集群，以下配置类似



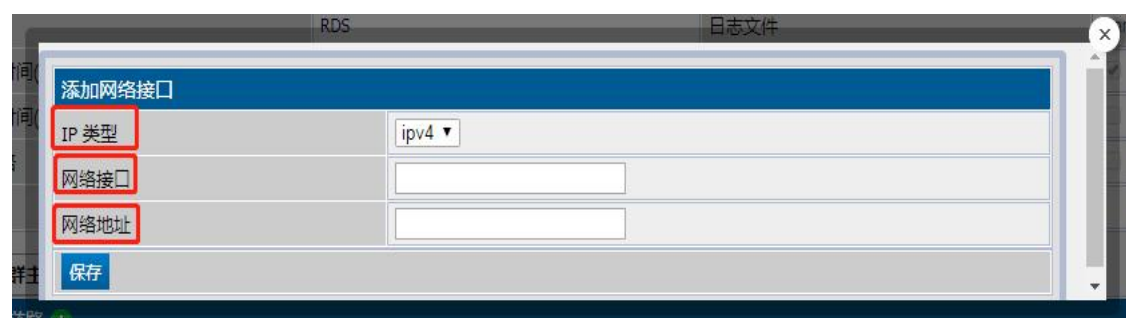
集群参数配置成功后，会自动跳转到监控页面



Note:

基本参数配置：可以进行修改集群参数配置

心跳链路配置



Note:

IP 类型：支持 IPv4 和 IPv6

网卡心跳链路：推荐配置单播心跳，并且建议配置两条及以上的心跳网络接口

网络地址：心跳链路地址，有多少个节点必须配置多少个心跳 IP 地址，如添加网卡 eth1 后，由于是 2+1 类型，有三个节点机器，其网络地址必须添加三个心跳链路地址。

成功添加心跳链路

集群主机	多播心跳	单播心跳	串口心跳	节点域
网卡心跳链路 +				
#	网络接口	IP 类型	网络地址	删除
1	eth1	ipv4	192.168.11.21 192.168.11.22 192.168.11.29	0
2	eth2	ipv4	192.168.12.21 192.168.12.22 192.168.12.29	0

节点域增删

集群主机	多播心跳	单播心跳	串口心跳	节点域
节点域 +				
#	节点名	仲裁盘	删除	
domain1	linux76 linux77	/dev/sdc1		
domain2	linux78			

Note:
节点域： 可进行节点域的增加和删除操作。

5.2.2 镜像初始化

将鼠标移至“镜像管理”会出现增加按钮，点击增加会弹出镜像配置页面



资源名	r0
linux76	
磁盘设备	/dev/sdc2
IP地址(多个IP请用空格分开)	192.168.11.76
端口号(0~65535)	7788
linux77	
磁盘设备	/dev/sdc2
IP地址(多个IP请用空格分开)	192.168.11.77
端口号(0~65535)	7788
linux78	
磁盘设备	/dev/sdb
IP地址(多个IP请用空格分开)	192.168.11.78
端口号(0~65535)	7789
Common	
同步源(主节点)	<input checked="" type="radio"/> linux76 <input type="radio"/> linux77 <input type="radio"/> linux78
文件系统	ext4 ▼
同步方式	立刻同步 ▼
proxy	
缓存限制	500 M ▼
压缩率	无压缩 ▼
保存 取消	

Note:

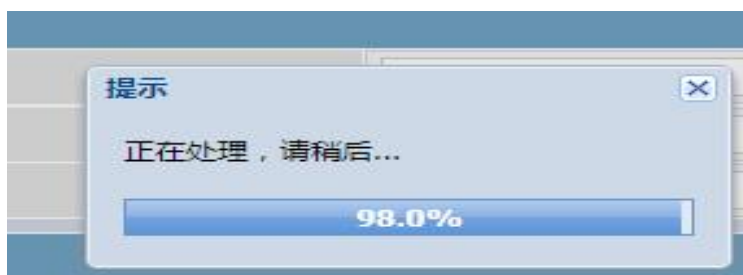
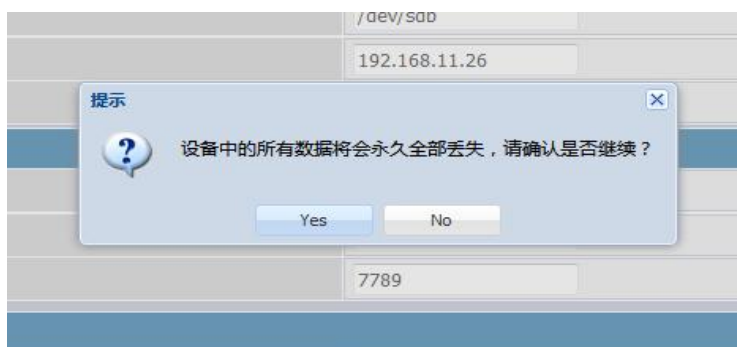
资源名：可以自定义，缺省 **rx** (x:0,1,2,...)

磁盘设备：磁盘设备是指存放业务数据的分区(也称数据分区)，如果环境中的数据分区有数据,则文件系统不用填写,否则会清除所有业务数据。

同步源：同步的是业务数据，需选择存在业务数据的节点为同步源，假如没有业务数据，则同步源自行先按需选择。

同步方式：立即同步和暂不同步；若选择“暂不同步”，则在配置完成后，启动集群时会自动进行格式化。

选择“立即同步”，会提示数据分区首次使用需要格式化，这里 yes 即可





Note:

镜像盘正在同步中，同步完成后会提示初始化成功。

选择“暂不同步”



Note:

1. 暂不同步不会出现同步进度条，稍等一会后会弹出提示框“新资源文件初始化成功”。
2. 集群所有配置完成后，启动集群时会自动进行格式化。

成功添加镜像资源

镜像参数配置	
linux76	
磁盘设备	/dev/sdc2
IP地址(多个IP请用空格分开)	192.168.11.76
端口号(0~65535)	7788
linux77	
磁盘设备	/dev/sdc2
IP地址(多个IP请用空格分开)	192.168.11.77
端口号(0~65535)	7788
linux78	
磁盘设备	/dev/sdb
IP地址(多个IP请用空格分开)	192.168.11.78
端口号(0~65535)	7788
proxy	
缓存限制	500 M ▼
压缩率	无压缩 ▼

Note:

镜像初始化成功后，可进行修改和删除操作

5.2.3 服务初始化

将鼠标移至“集群服务配置”会出现增加按钮，点击增加弹出服务配置页面

The image shows a 'Add Service' (添加服务) configuration window. It contains the following fields and options:

- 服务名** (Service Name): ora
- 服务类型** (Service Type): ☒ 普通 (Normal) ☐ mirror
- 镜像资源** (Image Resource):
- 域节点** (Domain Node): linux76 linux77 linux78 (多个节点请使用空格分开)
- 域节点启动顺序** (Domain Node Start Order): linux76 linux77 linux78 (多个节点请使用空格分开)
- 优先节点** (Priority Node): linux76 ▼
- 启动服务超时时间 (20~1200秒)** (Start Service Timeout): 240
- 停止服务超时时间 (20~1200秒)** (Stop Service Timeout): 360
- 服务检测间隔时间 (15~120秒)** (Service Check Interval): 30
- 服务检测超时时间 (20~1200秒)** (Service Check Timeout): 60
- 最大出错次数** (Maximum Error Count): 1
- 集群管理服务** (Cluster Management Service): ☒ 管理
- 服务检测失败后** (After Service Check Failure): ☐ 重启

At the bottom, there is a '保存' (Save) button.

Note:

服务名：集群内部识别，自行定义。

服务类型：普通类型是指业务服务(如 oracle、tomcat 等)。

提示：针对 rds 版本，drbd 类型存储资源添加时会自动添加一个镜像服务（缺省名：MRsrvx），并且当前配置的服务会依赖该镜像服务，资源删除时同样会自动删除。镜像服务会在资源 DiskMount 配置完后自动生成，此过程对使用者来说是透明的，不用关注其实现过程。

域节点：集群中节点名称,即主机名。

域节点启动顺序：集群中节点启动顺序。

优先节点：主要体现为当故障恢复时,服务会回倒该优先节点。

最大错误次数：检测出错多少次开始切换。

添加服务参数配置成功

服务参数配置

服务属性参数  

服务名	ora	集群管理服务	<input checked="" type="checkbox"/> 管理
优先节点	linux76	域节点	linux76 linux77 linux78
域节点启动顺序	linux76 linux77 linux78		
启动服务超时时间 (20~1200秒)	240	停止服务超时时间 (20~1200秒)	360
服务检测间隔时间 (15~120秒)	30	服务检测超时时间 (20~1200秒)	60
最大出错次数	1	服务检测失败后	<input type="checkbox"/> 重启

PubNIC IP DISKMOUNT APP

公共网卡资源 

#	资源名标识	操作
---	-------	----

Note:

服务属性参数：可进行修改和删除服务


服务资源：PubNIC、IP、DiskMount、APP

服务资源配置，资源根据需求可以添加多个

NIC 配置

选择 PubNIC，点击增加按钮

PubNIC IP DISKMOUNT APP

公共网卡资源 

#	资源名标识
---	-------

弹出配置页面

服务名	ora		
公共网卡资源名	ora_nic_0		
第三方主机IP	0.0.0.0		
可管理资源	<input checked="" type="checkbox"/> 是		
关键资源	<input checked="" type="checkbox"/> 是		
主机名	linux76	设备名	eth0 ▲ eth1 eth2 ▼
主机名	linux77	设备名	eth0 ▲ eth1 eth2 ▼
主机名	linux78	设备名	eth0 ▲ eth1 eth2 ▼
			

Note:

公共网卡资源名：自定义，缺省为服务名_nic_x (x:0,1,2,...)

第三方主机 IP：指网络上存在的有效 IP，用于检测自身网络是否正常，若不能保证该 IP 一直 ping 通，建议不进行配置，缺省即可。

设备名：指业务网卡

添加 NIC 资源成功

PubNIC

IP

DISKMOUNT

APP

公共网卡资源

#	资源名标识	操作
1	ora_nic_0	<div><div></div><div></div></div>

公共网卡资源配置

公共网卡资源名	ora_nic_0	第三方主机IP	0.0.0.0
可管理资源	<input checked="" type="checkbox"/> 是	关键资源	<input checked="" type="checkbox"/> 是
主机名	linux76	设备名	eth0
主机名	linux77	设备名	eth0
主机名	linux78	设备名	eth0

Note:
公共网卡资源：可添加多个 NIC 资源
操作：可修改和删除 NIC 资源配置

IP 配置

选择 IP，点击增加按钮

PubNIC		IP	DISKMOUNT	APP
IP 资源				
#	资源名标识			

弹出配置页面

添加IP资源

服务名

ora

IP资源名

ora_ip_0

IP 类型

ipv4

可管理资源

☒ 是

关键资源

☒ 是

主机名

公共网卡资源名

浮动 IP 地址

子网掩码

linux76

eth0

192.168.8.64

255.255.255.0

linux77

eth0

192.168.8.64

255.255.255.0

linux78

eth0

192.168.8.64

255.255.255.0

保存

Note:
IP 类型：支持 IPv4 和 IPv6
公共网卡资源名：要与配置的资源 NIC 的网卡一致
浮动 IP 地址：业务 IP

添加资源 IP 成功

PubNIC		IP	DISKMOUNT	APP
IP 资源 				
#	资源名标识		操作	
1	ora_ip_0		 	
IP资源配置				
IP资源名	ora_ip_0		IP 类型	ipv4
可管理资源	<input checked="" type="checkbox"/> 是		关键资源	<input checked="" type="checkbox"/> 是
主机名	公共网卡资源名		浮动 IP 地址	子网掩码
linux76	eth0		192.168.8.64	255.255.255.0
linux77	eth0		192.168.8.64	255.255.255.0
linux78	eth0		192.168.8.64	255.255.255.0

Note:

IP 资源：可添加多个 IP 资源

操作：可修改和删除 IP 资源配置

DiskMount 配置

选择 DISKMOUNT，点击增加按钮

PubNIC	IP	DISKMOUNT	APP
DiskMount资源 			
#	资源名标识		

弹出配置页面

添加DiskMount资源	
服务名	ora
DiskMount资源名	ora_dm_0
磁盘文件格式	drbd
磁盘设备名	/dev/drbd0
挂载点	/home/db
文件系统	ext3
挂载的用户	root
挂载的用户组	root
挂载的权限	755
挂载参数	rw
文件系统修复	no
网络储存	no
容量阈值百分比	95 %
磁盘只读时停止	<input checked="" type="checkbox"/> 是
可管理资源	<input checked="" type="checkbox"/> 是
关键资源	<input checked="" type="checkbox"/> 是
保存	

Note:

磁盘文件格式：选择 drbd 格式

磁盘设备名：文件格式若选择 drbd 后，磁盘会自动选择 drbd0。

文件系统：主要手动能挂载成功的文件系统都可支持。

容量阈值百分比：意思是当共享存储的容量超过设定的百分比时就会提示。

磁盘只读时停止：当共享存储的文件系统变成只读时,服务会停止

添加 DiskMount 成功

PubNIC	IP	DISKMOUNT	APP
DiskMount资源 			
#	资源名标识		操作
1	ora_dm_0		 
DiskMount资源			
DiskMount资源名	ora_dm_0		
磁盘设备名	/dev/drbd0		
挂载点	/home/db	文件系统	ext3
挂载的用户	root	挂载的用户组	root
挂载的权限	755	挂载参数	rw
文件系统修复	no	网络储存	no
容量阈值百分比	95	磁盘只读时停止	<input checked="" type="checkbox"/> 是
可管理资源	<input checked="" type="checkbox"/> 是	关键资源	<input checked="" type="checkbox"/> 是

Note:

DiskMount 资源：可添加多个 DiskMount 资源

操作：可修改和删除 DiskMount 资源配置

APP 资源配置

选择 APP，点击增加按钮

PubNIC	IP	DISKMOUNT	APP
App资源 +			
#	资源名标识		

弹出配置页面

添加应用资源

服务名	ora
应用资源标识	ora_app_0
应用脚本路径	/home/oracle.sh
检测超时次数当错误处理	0
可管理资源	<input checked="" type="checkbox"/> 是
关键资源	<input checked="" type="checkbox"/> 是
<div>保存</div>	

Note:

应用脚本路径：业务脚本

检测超时次数当错误处理：超时次数达到后会以错误处理，为资源异常。

添加 APP 资源成功

PubNIC

IP

DISKMOUNT

APP

App资源

#	资源名标识	操作
1	ora_app_0	

应用资源配置

应用资源标识	ora_app_0		
应用脚本路径	/home/oracle.sh		
检测超时次数当错误处理	0		
可管理资源	<input checked="" type="checkbox"/> 是	关键资源	<input checked="" type="checkbox"/> 是

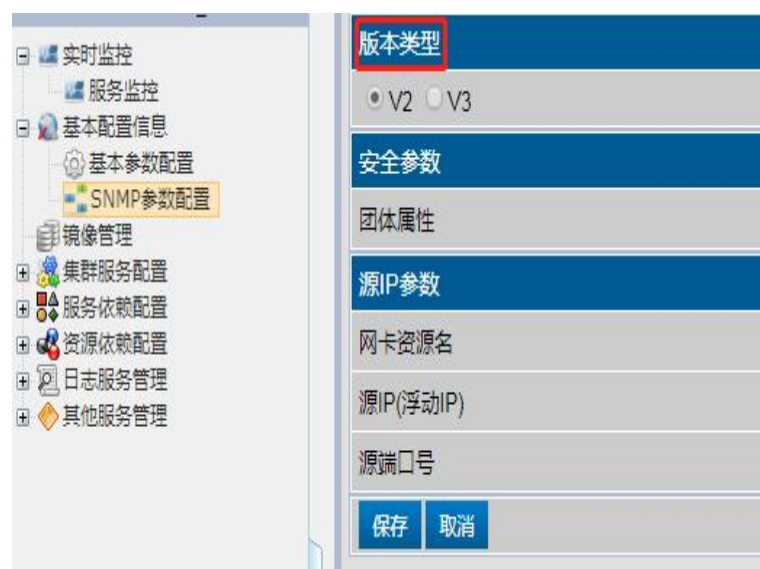
Note:

APP 资源：可添加多个 APP 资源

操作：可修改和删除 APP 资源配置

5.2.4 告警服务配置（选配）

点击“基本配置信息” -- “SNMP 参数配置” -- “+”号



Note:

版本类型：包括 SNMPV2 和 SNMPV3 版本，区别在于 SNMPV3 版本根据需求可添加验证加密方式。

若选择配置 SNMPV2 版本告警服务，配置如下：

SNMP参数配置

版本类型

V2

V3

安全参数

团体属性

public

源IP参数

网卡资源名

IP 类型

ipv4

源IP(浮动IP)

子网掩码

源端口号

161

保存

取消

团队属性：缺省为公用的 public，可按需修改成私有串体团。

源 IP 参数：配置告警源的网卡、IP 地址、掩码和端口，即配置告警信息发起端。

配置完参数后，点击保存

点击“保存”后，进行配置目的 IP 参数

SNMP参数配置

安全参数

团体属性

public

源IP参数

网卡资源名

eth0

IP 类型

ipv4

源IP

192.168.8.250

子网掩码

255.255.255.0

源端口号(0~65535)

161

修改

目的IP参数

#	目的IP	目的端口号(0~65535)	删除
1	192.168.8.19	162	<div></div>

增加

Note:

目的 IP 参数：配置告警目的 IP 和端口，即配置告警信息接收端

若选择配置 SNMPV3 版本告警服务，配置如下：

SNMP参数配置

版本类型

☐ V2 ☒ V3

源IP参数

网卡资源名

IP 类型

ipv4 ▼

源IP(浮动IP)

子网掩码

源端口号

161

保存

取消

源 IP 参数：配置告警源的网卡、IP 地址、掩码和端口，即配置告警信息发起端。
配置完参数后，点击保存

点击增加，配置 V3 用户信息

SNMP参数配置 ⓘ

源IP参数

网卡资源名

eth0

IP 类型

ipv4 ▼

源IP(浮动IP)

12.12.12.121

子网掩码

255.255.255.0

源端口号

161

修改

v3用户

增加

目的IP参数

增加

SNMP参数配置 ⓘ

源IP参数

网卡资源名

eth0

IP 类型

ipv4 ▼

源IP(浮动IP)

12.12.12.121

子网掩码

255.255.255.0

源端口号

161

修改

v3用户

用户名称

SNMPv3_user

安全级别

authPriv ▼

验证协议

SHA ▼

验证密码

Auth_1234

加密协议

AES ▼

加密密码

Priv_1234

保存

取消

目的IP参数

增加

Note:

安全级别：分为三种，包括 authNoPriv 验证不加密，noAuthNoPriv 不验证不加密，authPriv 验证加密。

验证协议：分为 SHA 和 MD5 验证，按需选择。

加密协议：分为 AES 和 DES 加密，按需选择。

配置完 V3 用户信息后，点击保存

点击增加，配置目的 IP 参数配置

SNMP参数配置

源IP参数

网卡资源名	eth0	IP 类型	ipv4
源IP(浮动IP)	12.12.12.121	子网掩码	255.255.255.0
源端口号	161		

修改

v3用户

SNMPv3_user

用户名	SNMPv3_user	安全级别	authPriv
验证协议	SHA	验证密码	Auth_1234
加密协议	AES	加密密码	Priv_1234

增加

目的IP参数

增加

目的IP参数

添加目的IP与端口号

目的IP		目的端口号	162
引擎ID	0102030a0b0c		
用户名	Trap_v3user	安全级别	authPriv
验证协议	SHA	验证密码	Auth_1234
加密协议	AES	加密密码	Priv_1234

保存 取消

目的 IP 参数：配置告警目的 IP 和端口，即配置告警信息接收端。
引擎 ID：主要用于验证使用，缺省即可。

SNMPV3 配置完成

SNMP参数配置

源IP参数

网卡资源名	eth0	IP 类型	ipv4
源IP(浮动IP)	12.12.12.121	子网掩码	255.255.255.0
源端口号	161		

修改

v3用户

SNMPv3_user

用户名	SNMPv3_user	安全级别	authPriv
验证协议	SHA	验证密码	Auth_1234
加密协议	AES	加密密码	Priv_1234

增加

目的IP参数

192.168.8.23

目的IP	192.168.8.23	目的端口号	162
引擎ID	0102030a0b0c		
用户名	Trap_v3user	安全级别	authNoPriv
验证协议	SHA	验证密码	Auth_1234

增加

至此，web 上 RDS 2+1 的配置完成。

5.3 RDS 1+1 初始化配置

5.3.1 集群初始化

集群参数化配置

NewStart HA
5.0.6-20190625.x86_64
Product: Type = Remote cluster

基本配置信息

集群基本配置

集群参数配置

集群名	rcs_cluster		
心跳间隔时间(1-10秒)	3	工作链路心跳	<input checked="" type="checkbox"/>
宕机间隔时间(15-300秒)	100	内核恐慌	<input type="checkbox"/>
最大化策略	数据安全最大化	watchdog检测配置	<input type="checkbox"/>
故障切换	手动		

集群主机

主机名	主机IP	删除
linux76	192.168.8.76	<input type="button" value="0"/>
linux77	192.168.8.77	<input type="button" value="0"/>

节点域

节点域	仲裁盘	删除
-----	-----	----

保存

Note:

- 1.节点域：RDS 1+1 不需要添加节点域即可点击保存，RDS 支持多个 1+1 配置。
- 2.本地节点与异地节点的 IP 可以属于不同网段，保证相互之间能通讯即可。
- 3.其它配置与 [5.2.1 节](#) 配置基本类似，区别在于添加 RDS 一个 1+1 不用配置节点域，若是添加了两个及以上的 1+1，可点击“+”增加节点域。

5.3.2 镜像初始化

配置过程与 [5.2.2 节](#) 类似，不进行重复说明。

5.3.3 服务初始化

配置过程与 [5.2.3 节](#) 类似，不进行重复说明。

5.3.4 告警服务配置（选配）

配置过程与 [5.2.4 节](#) 类似，不进行重复说明。

至此，web 上 RDS 1+1 的配置完成。

5.4 集群启动

集群初始化配置完成后就可以进行集群启动，这里有几点需要注意：

- *) 集群启动，任一节点上执行命令即可
- *) 集群管控的服务（资源组）会自动随集群启动
- *) RDS 2+1 的服务间启动的方式是串行（即逐个服务启），缺省情况下按服务配置先后顺序，RDS 1+1 缺省无此限制

点击“服务监控”——右键“任一节点名”——选择“启动全部集群”



Note:

可以单个节点启动，也可以全部节点启动

执行后，正常情况下，RDS 2+1 监控页面就会显示如下：

集群状态监控					
服务状态					
主机名	主机状态	心跳链接状态	服务名	服务状态	可管理
 linux76	 (up)	 eth1   (ONLINE)	 NshaService	 (running)	
			 MRsrv0:0	 (running)	
			 ora	 (running)	
 linux77	 (up)	 eth1   (ONLINE)	 NshaService	 (stopped)	
			 MRsrv0	 (stopped)	
			 ora	 (stopped)	
 linux78	 (up)	 eth1 	 NshaService	 (stopped)	
			 MRsrv0:1	 (running)	
			 ora	 (stopped)	
本地存储状态					
资源名称	linux76	linux77	linux78		
 n0-dev/drbd0	 online		 online		

Note:

服务状态：查看集群和服务运行状态；

告警服务 NshaService、镜像服务 MRsrv、服务 ora：服务状态运行正常，处于可管理状态

本地储存状态：运行正常。

执行后，正常情况下，RDS 1+1 监控页面就会显示如下：

集群状态监控

服务状态					
主机名	主机状态	心跳链路状态	服务名	服务状态	可管理
 linux76	 (up)	 eth1  eth2	 NshaService	 (running)	
			 ora	 (running)	
 linux77	 (up)	 eth1  eth2	 NshaService	 (stopped)	
			 ora	 (stopped)	
本地存储状态					
资源名称		linux76		linux77	
 r0-dev/drbd0		 online		 online	

Note:

服务状态：查看集群和服务运行状态；

告警服务 NshaService、服务 ora：服务状态运行正常，处于可管理状态

本地储存状态：运行正常。

第 6 章 集群的验收测试

集群软件出厂之前已经过各种严格测试，但是其与系统、硬件及业务脚本存在一定关联，需要在工程现场进行以下两项测试，以确保主机故障时备机能自动接管业务，实现持续性运行。

- I. 启动集群软件，查看集群的运行情况，进行服务的倒换测试，确认服务能倒换成功（包括主机到备机、备机到主机两种倒换）；进行业务测试，确认服务运行状态与实际业务状态一致。
- II. 故障测试：节点集群程序重启能否自动启动并加入集群；主机重启或者关机服务能否倒换到备机；其他故障测试一般可以不做。

第 7 章 集群交互式命令

交互式命令（cli）比较简单，只要按交互式操作向导都可快速完成所需操作，按用途可分以下几类。

分类	命令	功能描述
配置修改	cluster-init	集群配置初始化，生成集群配置文件
	cluster-modify	修改集群初始化后配置文件
	cluster-show	显示集群配置文件信息
	heartbeat-add	增加心跳链路
	heartbeat-delete	删除心跳链路
	service-init	服务配置初始化，生成服务配置文件
	service-add	向配置文件添加一个服务（资源组）
	service-delete	从配置文件里删除一个服务（资源组）
	service-modify	修改一个服务的配置信息
	service-show	显示指定服务的配置信息
	service-depend	不同服务间建立依赖关系
	resource-add	添加服务（资源组）的相关资源
	resource-delete	删除服务（资源组）的相关资源
	resource-modify	修改服务的资源设定
	resource-depend	不同资源间建立依赖关系
	mirror-init	镜像配置初始化，生成镜像配置文件
	mirror-adm	镜像资源添加或删除
	mirror-mode-switch	同城或异地集群类型切换
	modify-config	把旧版本的配置文件转换成新版本的配置文件
	instance-config	实例增、删和显示
	product-switch	切换版本相关配置
	domain-config	节点域配置
	recover-config-files	恢复所有配置文件到其它节点
	lockdisk-adm	仲裁盘管理
节点管理	cluster-start	启动集群
	cluster-stop	停止集群
	cluster-restart	集群重启
服务管理	service-start	手动启动服务，前提是集群已启动
	service-stop	手动停止服务，前提是集群已启动
	service-migrate	手动切换服务
	service-enable	集群管控服务，缺省配置下是管控状态
	service-disable	服务脱离集群管控，常用于业务在线升级

	clear-fault	清除服务的软件错误标识
资源管理	resource-start	资源手动启动
	resource-stop	资源手动停止
	resource-enable	集群管控资源
	resource-disable	资源脱离集群管控
	resource-critical	设定资源为关键资源，故障会倒换，缺省都是关键资源
	resource-noncritical	设定资源为非关键资源，故障不会倒换
状态监控	cluster-stat	持续显示当前集群运行状态
	show-status	显示当前集群状态
SNMP告警	snmp-add	配置告警源IP及告警IP（网管端IP）
	snmp-delete	删除告警源IP及告警IP
	snmp-modify	修改告警源IP及端口设定
其它命令	check-config	检查集群所有配置文件配置是否正常
	recover-remote-license	恢复授权文件到指定节点，用于节点故障授权恢复
	exit、quit	退出cli
	help	查用cli中的命令用法

注意： 其它命令参数可参考手册“NewStart Cluster 5.0 用户手册.pdf”

7.1 配置修改

对集群所涉及的配置文件进行操作，包括集群配置文件（/etc/ha.d/ha.cf）、镜像资源配置文件（/etc/drbd.d/xxx.res）和服务配置文件（/etc/ha.d/haservices.xml）。其中命令 cluster-init、mirror-init 和 service-init 前面初始化配置已解说过，这里不重复。

cluster-modify

修改集群配置文件参数。

格式：cli->cluster-modify

```
cli:~>cluster-modify
Configuration of HA system:
    0) clustername: rds_cluster
    1) node: linux76 192.168.8.76
    2) node: linux77 192.168.8.77
    3) node: linux78 192.168.8.78
```

```

4) node: linux79 192.168.8.79
5) logfile: /var/log/ha-log
6) keepalive: 3
7) deadtime: 60
8) watchdog: off
9) ucast: ens224 192.168.11.76 192.168.11.77 192.168.11.78
192.168.11.79
10) ucast: ens256 192.168.12.76 192.168.12.77 192.168.12.78
192.168.12.79
11) worklink_hb: off
12) auto_failover: off
13) auto_run: on
14) maximization: data-security
15) kernel_panic: on
16) domain_node: linux76 linux77
17) domain_node: linux78 linux79
b) Back to previous menu
Select an item to modify: [0-17,b,q] 7
Input new deadtime: [c,q] (90) 30
Modify another item? [y/n] (n)
Do you want to save the changes? [y/n] (y)
Please restart the HA system to take your modification into effect.

```

cluster-show

查看当前集群配置文件。

格式: cli->cluster-show

```

cli:~>cluster-show
configuration of the cluster:
  clustername: rds_cluster
  node: linux76 192.168.8.76
  node: linux77 192.168.8.77
  node: linux78 192.168.8.78
  node: linux79 192.168.8.79
  logfile: /var/log/ha-log
  keepalive: 3
  deadtime: 30
  watchdog: off
  ucast:  ens224   192.168.11.76   192.168.11.77   192.168.11.78
192.168.11.79
  ucast:  ens256   192.168.12.76   192.168.12.77   192.168.12.78
192.168.12.79

```

```
worklink_hb: off
auto_failover: off
auto_run: on
maximization: data-security
kernel_panic: on
domain_node: linux76 linux77
domain_node: linux78 linux79
```

heartbeat-add

增加一条心跳链路。

格式: cli->heartbeat-add

```
cli:~>heartbeat-add
Add heartbeat links:
IP type:
    0) IPv4
    1) IPv6
Select a type: [0-1,q] (0)
Detected NICs on cluster: ens192 ens224 ens256
Input a network device for unicast heartbeat: [q] (ens192)
Input linux76 IPv4 address [q] (192.168.8.76)
Input linux77 IPv4 address [q] (192.168.8.77)
Input linux78 IPv4 address [q] (192.16.8.8.78)
Input linux79 IPv4 address [q] (192.168.8.79)
Add another ucast heartbeat link? [y/n] (y) n
Do you want to save the configuration? [y/n] (y)
Please restart the HA system to take your modification into effect.
```

heartbeat-delete

删除心跳链路，要求至少保留一条心跳链路。

格式: cli->heart-delete

```
cli:~>heartbeat-delete
0) ucast: ens224 192.168.11.76 192.168.11.77 192.168.11.78 192.168.11.79
1) ucast: ens256 192.168.12.76 192.168.12.77 192.168.12.78 192.168.12.79
2) ucast: ens192 192.168.8.76 192.168.8.77 192.168.8.78 192.168.8.79
b) Back to previous menu
Select a heartbeat link channel to delete: [0-2,b,q] 2
Do you want to save the configuration? [y/n] (y)
Please restart the HA system to take your modification into effect.
```

service-add

添加服务（资源组），运行service-add，服务配置过程与service-init类似，这里不再介绍。

格式：cli->service-add

service-delete

删除服务配置文件中已配置的服务（资源组）。

格式：cli->service-delete

```
cli:~>service-delete
    0) oracle
    b) Back to previous menu
Please select a service to delete: [0-0,b,q] 0
```

service-modify

修改服务配置，如修改ora_tom服务停止超时时间如240秒。

格式：cli->service-modify

```
cli:~>service-modify
Services of the cluster:
    0) ora_tom
    1) MRsrv0
    b) Back to previous menu
Please select a service to modify: [0-1,b,q] 0
Property information of 'ora_tom':
    0) enabled: yes
    1) multinstance: no
    2) starttimeout: 240
    3) stoptimeout: 360
    4) checkinterval: 30
    5) checktimeout: 60
    6) maxerrcount: 1
    7) maxlockcount: 0
    8) restartaftercheckfail: no
    9) startanyway: no
   10) nodelist: linux76 linux77 linux78 linux79
   11) autostartlist: linux76 linux77 linux78 linux79
```

```

    12) preferrednode:
    b) Back to previous menu
Please select an item to modify: [0-12,b,q] 3
Input new stoptimeout: [c,q] (360) 240
Modify other property of service? [y/n] (n)
Modify other service? [y/n] (n)
Do you want to save the changes? [y/n] (y)
Please restart the HA system to take your modification into effect.

```

service-depend

设定不同服务之间的依赖关系（规则：父服务依赖于子服务，启动服务时子服务先启，停止服务时反之），目前支持两种依赖关系：

- 1) local-forced: *依赖的服务都在本地运行*
 - 2) global-fixed: *依赖的服务允许在本地或远端运行*
- 格式: cli->service-depend

```

cli:~>service-depend
    0) Link (add service dependency)
    1) Unlink (delete service dependency)
    2) Display (display service dependency)
    b) Back to previous menu
Please select an operation: [0-2,b,q] 0

To configure service dependency the following information is required:

    Parent service
    Child service
    Type of dependency, including fixed and forced
    Relationship, including local, remote and global

    Notice: The parent depends on the child being running before going
    running itself.
    Conversely, parent stops before child does so.

Do you want to continue? [y/n] (y)
    0) ora_tom
    1) MRsrv0
Please select a parent service: [0-1,c,q] 0
Please select a child service: [0-1,c,q] 1
Relationship:
    0) global
    1) local

```

```

Please select a kind of relationship: [0-1,q] 1
Dependency type:
    0) forced
Please select a type of dependency: [0-0,q] 0
Do you want to configure another service dependency? [y/n] (n)
Service dependency information verification:

    Parent:ora_tom      Child:MRsrv0      Relationship:local
Dependency-type:forced

Is this information correct? [y/n,q] (y)

```

resource-add

对配置的服务进行修改，增加相关资源，如按需增加一个业务 ip，增加磁阵等。
格式：cli->resource-add

```

cli:~>resource-add
Services of the cluster:
    0) oracle
    b) Back to previous menu
Please select a service to add resource: [0-0,b,q] 0
Resource types:
    0) NIC
    1) IP
    2) Application
    3) RawDisk
    4) DiskMount
    b) Back to previous menu
Please select a type of resource: [0-4,b,q] 1

```

resource-delete

删除已存在服务的相关资源。
格式：cli->resource-delete

```

cli:~>resource-delete
Services of the cluster:
    0) oracle
    b) Back to previous menu

```



```
Please select a service to delete resource: [0-0,b,q] 0
Resources of service 'oracle':
    0) oracle_nic_0
    1) oracle_ip_0
    2) oracle_app_0
    3) oracle_dm_0
    b) Back to previous menu
Please select a resource to delete: [0-3,b,q]
```

resource-modify

修改已存在的服务的资源，如更改业务 ip 等。

格式: resource-modify

```
cli:~>resource-modify
Services of the cluster:
    0) oracle
    b) Back to previous menu
Please select a service to modify resource: [0-0,b,q] 0
Resources of service 'oracle':
    0) oracle_nic_0
    1) oracle_ip_0
    2) oracle_app_0
    3) oracle_dm_0
    b) Back to previous menu
Please select a resource to modify: [0-3,b,q] 1
```

resource-depend

设定同一服务中的不同资源之间的依赖关系（原则：父资源依赖于子资源, 资源启动时子资源先启，停止时反之）。

格式: cli:~>resource-depend

```
cli:~>resource-depend
    0) Link (add resource dependency)
    1) Unlink (delete resource dependency)
    2) Display (display resource dependency)
    b) Back to previous menu
Please select an operation: [0-2,b,q] 0
To configure resource dependency the following information is required:
```

```

    Parent resource
    Child resource

    Notice: The parent depends on the child being running before going
    running itself.
    Conversely, parent stops before child does so.
    Do you want to continue? [y/n] (y)
    Services of the cluster:
        0) oracle
        1) monitor
        b) Back to previous menu
    Please select a service to configure resource dependency: [0-1,b,q] 0
        0) oracle_nic_0
        1) oracle_ip_0
        2) oracle_app_0
        3) oracle_dm_0
    Please select a parent resource: [0-3,c,q] 1
    Please select a child resource: [0-3,c,q] 0
    Do you want to configure another resource dependency? [y/n] (n)
    Resource dependency information verification:

        Parent: oracle_ip_0    Child: oracle_nic_0
    Is this information correct? [y/n,q] (y)

```

mirror-adm

用于增加或删除一个镜像（mirror）资源，镜像资源的增加过程与 mirror-init 类似，这里不重复。

格式：cli:~>mirror-adm

```

cli:~>mirror-adm
    0) add mirror resource
    1) delete mirror resource
    b) Back to previous menu
    Please select an operation: [0-1,b,q] 1
        0) r0
        b) Back to previous menu
    which do you want to delete: [0-0,b,q] 0

```

modify-config

转换集群所有配置文件，使之与当前集群的版本相匹配。

格式: cli:~>modify-config

```
cli:~>modify-config
```

```
Modify /etc/ha.d/ha.cf successful. Please manual copy to other node.
```

```
Modify /etc/ha.d/haservices.xml successful. Please manual copy to other node.
```

instance-config

用于增加、删除和显示已配置实例，实例属于多实例服务中的概念，每个实例也是一组资源组，多实例服务应用于同一业务应用在不同节点同时运行。

格式: cli:~>instance-config

```
cli:~>instance-config
```

```
0) Instance add      ——增加实例
```

```
1) Instance delete  ——删除实例
```

```
2) Display instance ——显示已配置实例
```

```
b) Back to previous menu
```

```
Please select an operation: [0-2,b,q] 0
```

```
To configure multi-instance the following information is required:
```

```
Instance ID
```

```
Select resource for the instance
```

```
Do you want to continue? [y/n] (y)
```

```
Services of the cluster:
```

```
0) mult_instance
```

```
b) Back to previous menu
```

```
Please select a service to add instance: [0-0,b,q] 0
```

```
Please input the instance ID: [0-99,q] (0)
```

```
This service 'mult_instance' has the following resources:
```

```
0) mult_instance_nic_0
```

```
1) mult_instance_ip_0
```

```
2) mult_instance_ip_1
```

```
3) mult_instance_ip_2
```

```
4) mult_instance_app_0
```

```
Select resources separated by spaces: [0-4,q] 0 1 4
```

```
Resources for instance ID 0 information verification:
```

```
mult_instance_nic_0
```

```
mult_instance_ip_0
```

```
mult_instance_app_0

Is this information correct? [y/n,q] (y)
Do you want to add another instance? [y/n] (y)
Do you want to add another instance? [y/n] (y)
Please input the instance ID: [0-99,q] (1)
This service 'mult_instance' has the following resources:
    0) mult_instance_nic_0
    1) mult_instance_ip_0
    2) mult_instance_ip_1
    3) mult_instance_ip_2
    4) mult_instance_app_0
Select resources separated by spaces: [0-4,c,q] 0 2 4
Resources for instance ID 1 information verification:

    mult_instance_nic_0
    mult_instance_ip_1
    mult_instance_app_0

Is this information correct? [y/n,q] (y)
Do you want to add another instance? [y/n] (y)
...
```

product-switch

用于集群不同版本（双机、镜像、异地、多节点、rds）配置转换或产品试用。

```
cli:~>product-switch ----转换配置
Switch product success.

cli:~>product-switch rds -----产品试用30天
Switch product success.
```

 说明:

版本一键切换流程:

- 1) 操作前，先停止集群: cli->cluster-stop
- 2) 使用商用 SN 重新生成 key 文件:/usr/lib64/makekey + SN
- 3) 进入 cli，运行切换指令:product-switch
- 4) 此步骤分几种情况，对应以下几种操作:
 - 双机切镜像/异地)
 - *) 添加 drbd 资源
 - *) 修改 disk 资源，修改 device 及 type 参数
 - 镜像与异地互切)

- *) 无需额外操作, 跳过
镜像/异地切双机)
 - *) 修改 disk 资源, 修改 device 及 type 参数
 - *) disk 资源增删操作指令: resource-add/ resource-delete
- 5) 集群重启生效: cli->cluster-restart
- PS:
- 1、disk 资源修改流程: cli->resource-modify >选择 dm 资源>选择 “2) type”>最后修改 “0) device” 。
- 2、如果使用 web 操作, 执行上述步骤 3 后需重启 web 工具: webadminserver restart

domain-config

用于节点域的添加和删除。

格式: cli:~>domain-config

```
cli:~>domain-config
    0) Add domain-node
    1) Delete domain-node
    b) Back to previous menu
Please select an operation: [0-1,b,q] 1
domain-node:
    0) domain_node linux76 linux77
    1) domain_node linux78 linux79
    b) Back to previous menu
Please select a domain-node to delete: [0-1,b,q] 0
cli:~>domain-config
    0) Add domain-node
    1) Delete domain-node
    b) Back to previous menu
Please select an operation: [0-1,b,q] 0
Please select nodes for domain-node:
    0) linux76
    1) linux77
    2) linux78
    3) linux79
Select one or more nodes separated by spaces: [0-3,b,q] 0 1
```

recover-config-files

把当前集群所有配置文件恢复到其它节点, 常用于集群中节点硬件更换或软件重装后的集群恢复。

格式: cli:~>recover-config-files

```
cli:~>recover-config-files
Restore configuration files to remote nodes successfully!
```

lockdisk-adm

仲裁盘管理命令，可进行增，删操作。

格式: cli:~>lockdisk-adm

```
cli:~>lockdisk-adm
    0) add lock disk
    1) delete lock disk
    b) Back to previous menu
Please select an operation: [0-1,b,q]
```

mirror-mode-switch

用于同城或远程集群类型一键切换，其中 sync 表现同城模式，async 表示远程模式：

格式: cli:~>mirror-mode-switch

```
cli:~>mirror-mode-switch
There are two types of mirror-mode:
    0) sync
    1) async
    b) Back to previous menu
Please select a type: [0-1,b,q]
```

7.2 节点管理

Cluster-start

启动集群（支持所有节点的集群启动或指定节点的集群启动），服务会随着启动，如果服务脱离了集群管理，cluster-start不会带起服务。

格式: cli:~>cluster-start [nodename]

```
cli:~>cluster-start ----启动整个集群
Configuration file checked ok.
[linux76]Starting High-Availability services:
Done.

.
[linux77]Starting High-Availability services:
Configuration file checked ok.
Done.

.
[linux78]Starting High-Availability services:
Configuration file checked ok.
Done.

.
[linux79]Starting High-Availability services:
Configuration file checked ok.
Done.

.

cli:~> cluster-start linux76 ----启动指定节点集群
[linux76]Starting High-Availability services:
Configuration file checked ok.
Done.
```

Cluster-stop

停止集群（支持所有节点的集群停止或指定节点的集群停止），如果服务脱离了集群管理，cluster-stop只停止集群软件。

格式：格式：cli:~>cluster-stop [nodename]

```
cli:~> cluster-stop ----停止整个集群
[linux77]Stopping High-Availability services:
Done.

.
[linux78]Stopping High-Availability services:
Done.
```

```
.
[linux79]Stopping High-Availability services:
Done.

.
[linux76]Stopping High-Availability services:
Done.

.
cli:~> cluster-stop linux79 ----停止指定节点集群
[linux79]Stopping High-Availability services:
Done.

.
```

Cluster-restart

重启集群软件，相当于先运行cluster-stop停止集群，再运行cluster-start启动集群，这里不重复。

7.3 服务管理

service-start

手动启动服务。

格式: Cli:~>service-start

```
Cli:~>service-start
Current service:
    0) oracle
    b) Back to previous menu
Please select a service: [0-0,b,q] 0
Current node:
    0) susell-3
    1) susell-4
    b) Back to previous menu
Please select a node: [0-1,b,q] 0
Send message to HA system to start service oracle on susell-3.
```


Service-stop

手动停止服务。

格式: Cli:~>service-start

```
Cli:~>service-stop
Current service:
    0) oracle
    b) Back to previous menu
Please select a service: [0-0,b,q] 0
Send message to HA system to stop service oracle.
```

Service-migrate

手动切换服务，执行后可运行cluster-stat持续查看集群状态。

格式: cli:~>service-migrate

```
cli:~>service-migrate
Select service to migrate:
Current service:
    0) oracle
    b) Back to previous menu
Please select a service: [0-1,b,q] 0
Select the destination node:
Current node:
    0) susell-4
    b) Back to previous menu
Please select a node: [0-0,b,q] 0
Send message to migrate service oracle from susell-3 to susell-4.
```

Service-enable 和 service-disable

当集群启动时服务自动启动，集群停止时服务停止，服务运行过程中，集群软件定时检测服务状态，我们把这种关系称为“服务被集群管理”或者“服务加入集群管理”；当服务脱离集群管理时，服务与集群软件就没有任何关系了，比如：停止集群不影响服务运行；如果服务间存在依赖关系，相关依赖服务会一并enable或disable.

格式: Cli:~>service-enable

Cli:~>service-disable

Service-enable使服务加入集群管理，运行此命令的前提：集群正在运行，节点状态正常(UP)，服务状态为running/stoped，服务enable属性为no。

Service-disable使服务脱离集群管理，运行此命令的前提：集群正在运行，节点状态正常(UP)，服务状态为running/stoped，服务enable属性为yes。

```
Cli:~>service-enable
Current service:
    0) oracle
    b) Back to previous menu
Please select a service: [0-1,b,q] 0
There are services depend on service oracle, they will be enabled too.
Do you want to continue? [y/n] (y)
Send message to the HA system to enable service oracle.

Cli:~>service-disable
Current service:
    0) oracle
    b) Back to previous menu
Please select a service: [0-1,b,q] 0
There are services depend on service oracle, they will be disabled too.
Do you want to continue? [y/n] (y)
Send message to the HA system to disable service oracle.
```



说明：

- 1、服务enable成功或后，服务的enable属性为yes；服务disable成功后，服务的enable属性为no。
- 2、集群以上状态信息可通过cli->cluster-stat持续查看。

Clear-fault

清除服务的软件错误标志位。运行此命令前先运行cluster-stat命令查看服务状态，发现有错误标志位，使用该命令清除软件错误标准，再查看服务状态，确认服务状态正常。

格式：Cli:~>clear-fault

```
Cli:~>clear-fault
Current service:
    0) oracle
    b) Back to previous menu
```

```
Please select a service: [0-1,b,q] 1
Do you want to clear fault on all nodes [yes/no] (yes)
Send message to the HA system to clear fault of service monitor on all.
```



注意

1. 服务处于unstable状态时，运行clear-fault可能会导致服务停止，只有已经查明造成unstable的原因并解决后，才可以使用此命令；已商用系统请谨慎使用。
2. 执行clear-fault时如果服务在集群中处于停止状态，执行后服务保持原停止状态，如需启动服务，需要手动执行service-start。

7.4 资源管理

服务是资源的集合，可以看作是资源组，当服务启动时所有相关资源自动启动，服务停止时资源同样全部停止，这是Cluster对资源的自动管理，一般情况下对服务进行操作可满足大部份的日常需求。

resource-start

启动服务中的某个资源，如业务 ip, 网卡，应用等。

格式: cli:~>resource-start

```
cli:~>resource-start
Current service:
    0) oracle
    b) Back to previous menu
Please select a service: [0-1,b,q] 0
    0) name: oracle_app_0    script: /etc/ha.d/resource.d/myfaked.sh
    1) name: oracle_dm_0
    2) name: oracle_ip_0
    b) Back to previous menu
Please select a resource: [0-2,b,q]
```

resource-stop

停止服务中的某个资源，如业务 ip, 网卡，应用等。

格式: cli:~>resource-stop

```
cli:~>resource-stop
Current service:
    0) ora_tom
    1) MRsrv0
    2) bbbb
    b) Back to previous menu
```

```

Please select a service: [0-2,b,q] 2
    0) name: bbbb_ip_0
    1) name: bbbb_app_0    script: /home/myfaked.sh
    b) Back to previous menu
Please select a resource: [0-1,b,q] 1
Send message to HA system to stop resource bbbb_app_0 in service bbbb

```



说明:

1、如果 Cluster 重启，上述操作的资源会随服务自动启动。

resource-enable 和 resource-disable

resource-enable即管控服务中的特定资源，即集群监控并管理该资源的状态，出现异常时会有错误日志输出或触发一些事件，如是该资源为关键资源，甚至触发整个服务的倒换。缺省情况下服务的全部资源都会管控。

resource-disable 不管控服务中的特定资源，即该资源的运行状态不会影响整个服务运行。

格式: Cli:~>resource-enable

Cli:~>resource-disable

```

Cli:~>resource-enable
Current service:
    0) oracle
    b) Back to previous menu
Please select a service: [0-1,b,q] 0
    0) name: oracle_ip_0
    1) name: oracle_app_0    script: /etc/ha.d/resource.d/myfaked.sh
    2) name: oracle_dm_0
    b) Back to previous menu
Please select a resource: [0-2,b,q] 0
Please restart the HA system to take your modification into effect

cli:~>resource-disable
Current service:
    0) ora_tom
    1) MRsrv0
    2) bbbb
    b) Back to previous menu
Please select a service: [0-2,b,q] 2
    0) name: bbbb_ip_0
    1) name: bbbb_app_0    script: /home/myfaked.sh
    b) Back to previous menu

```

```
Please select a resource: [0-1,b,q] 1
Please restart the HA system to take your modification into effect.
```



说明:

1、**resource-enable** 和 **resource-disable** 操作需重启集群后才能生效, 即下次生效, 这与服务的脱管 (**service-disable**) 与管控(**service-enable**)的最大区别在此。

resource-critical 和 resource-noncritical

resource-critical的作用是设定服务中的某个资源为关键资源, 当其运行状态异常时会引起整个服务倒换, 缺省情况下都是关系资源, 关系资源故障会引起服务(资源组)倒换。

resource-noncritical的作用与**resource-critical**相反。

格式: Cli:~>resource-critical

Cli:~>resource-noncritical

```
Cli:~>resource-critical
Current service:
    0) oracle
    b) Back to previous menu
Please select a service: [0-1,b,q] 0
    0) name: oracle_ip_0
    1) name: oracle_app_0    script:
/etc/ha.d/resource.d/myfaked.sh
    2) name: oracle_dm_0
    b) Back to previous menu
Please select a resource: [0-2,b,q] 0
Please restart the HA system to take your modification into effect.
```

```
cli:~>resource-noncritical
Current service:
    0) ora_tom
    1) MRsrv0
    2) bbbb
    b) Back to previous menu
Please select a service: [0-2,b,q] 2
    0) name: bbbb_ip_0
    1) name: bbbb_app_0    script: /home/myfaked.sh
    b) Back to previous menu
Please select a resource: [0-1,b,q] 1
Please restart the HA system to take your modification into effect.
```

7.5 集群监控

cluster-stat

每隔 1秒自动刷新整个集群状态，包括集群的节点状态、心跳链路状态、仲裁链路状态和服务运行状态等。

格式：cli:~>cluster-stat

```
cli:~>cluster-stat
          Press Ctrl-C or 'Q' to exit
          Note: MI - Mirror, ---- - Abnormal link

Select a service to view the status of resources: MRsrv0
Use the arrow keys to select: Up-First, Down-Last, Left-Previous, Right-Next

NodeName   Status   HeartbeatLink   ServiceName   ServiceStatus   Enable
linux76     UP       ens224, ens256  MRsrv0:0      running         YES
           MI (r0, ens224)  ora_tom        running        YES
           MI (r1, ens256)
linux77     UP       ens224, ens256  MRsrv0        stopped         YES
           ora_tom      stopped        YES
linux78     UP       ens224, ens256  MRsrv0:1      running         YES
           MI (r0, ens224)  ora_tom        stopped        YES
           MI (r1, ens256)
linux79     UP       ens224, ens256  MRsrv0        stopped         YES
           ora_tom      stopped        YES
```

 说明：

(注意状态图中如果链路位置出现闪烁，则表示链路出现异常)

项	含义
NodeName	集群中节点名称
Status	节点状态，分以下三种： Up：节点正常 Down：节点故障 Unknown：节点状态未知
HeartbeatLink	1、心跳链路，会显示链路名称，如链路异常，会显示” - ” 2、配置镜像盘会在此列显示，格式:MI（镜像名称，链路）

	3、配置了仲裁盘（Ldisk）也会这里显示
ServiceName	用户自定义的服务名称，见 service_init
ServiceStatus	<p>服务状态说明：</p> <p>running: 服务正在运行 stopped: 服务没运行 stopping: 服务正在停止 starting: 服务正在启动 unstable: 服务不稳定，集群无法管控服务 unknown: 服务状态未知</p> <p>出现以下关键字表现服务状态异常或曾出现过异常：</p> <p>str: 服务启动失败，如str:(原因) stu: 服务正常运行中，服务周期性检测时检测到服务状态异常，如stu:(原因) stp: 服务停止失败，如: stp: (原因)</p> <p>显示引起服务异常的原因：</p> <p>op: 执行失败 N: 网络存储异常 I: IP地址 A: 应用程序 M: 挂载在文件系统的磁盘设备 S: 存储，如共享磁盘 W: 业务链路 T: 超时 B: 浮动IP在对方节点存在, 启动服务失败，双机处于裂脑状态 P: 无法ping通任何一个网络中存在的IP L: 启动服务时仲裁盘抢锁失败</p>
Enable	服务可管理与否。Yes表示可管理。可使用service-enable以及service-disable修改

上述状态图中，缺省选中并显示第一个服务，按up或down或left或right可以切换显示不同服务，回车后可以查看资源组状态，返回上一层按字母a，如下：

Press 'q' or 'Ctrl-C' to exit					
Press 'a' to return cluster status interface					
Service ora_tom's resource status:					
Name	Type	Enable	Crit1	NodeName	ResourceStatus
ora_tom_nic_0	nic	yes	yes	linux76	running

				linux77	running
				linux78	running
				linux79	running
ora_tom_ip_0	ip	yes	yes	linux76	running
				linux77	stopped
				linux78	stopped
				linux79	stopped
ora_tom_dm_0(r0)	disk	yes	yes	linux76	running
				linux77	stopped
				linux78	stopped
				linux79	stopped
ora_tom_dm_1(r1)	disk	yes	yes	linux76	running
				linux77	stopped
				linux78	stopped
				linux79	stopped
ora_tom_app_0	app	yes	yes	linux76	running
				linux77	stopped
				linux78	stopped
				linux79	stopped
ora_tom_app_1	app	yes	yes	linux76	running
				linux77	stopped
				linux78	stopped
				linux79	stopped

 说明：

项	含义
Name	服务中的资源名称，用于内部识别
type	资源类型：如ip、nic(网卡)、disk(共享存储)和app(业务应用)
NodeName	集群中的节点名
ResourceStatus	资源状态的状态有以下几种： unknown：资源状态未知，常见于资源disable情况 running：资源正在运行 stopped：资源已停止 starting：资源正在启动 stopping：资源正在停止 running pending：此状态表示应用已执行启动，但是否启动成功待确认 stopped pending：此状态表示应用已执行停止，但是

	否停止成功待确认 stopping admin: 资源异常，需要人工干预
Enable	资源可管理与否。Yes表示可管理
Critical	是否为关键资源，关键资源意味着该资源异常时会触发服务倒换

Show-status

获取当前集群状态，获取一次，获取的集群状态信息与cluster-stat输出内容相同，区别在于只获取一次，这里不重复。

格式：cli:~>show-status

7.6 SNMP 告警监控（选配）

集群支持发送 snmp 告警提示维护人员集群的变化，以便及时发现和解决问题。告警内容的详细说明及解读请参考附录：SNMP 告警消息类型与查询说明

snmp-add

配置告警源 IP 及告警 IP（网管端 IP）。

格式：cli:~>snmp-add

```
cli:~>snmp-add

To configure SNMP the following information is required:

    One source IP
    One or more trap IP

Do you want to continue? [y/n] (y)
*      0) SNMP source IP and read-write community
*      1) SNMP trap IP
        b) Back to previous menu
Please select an item to add: [0-1,b,q] 0a

To configure SNMP source IP the following information is required:

    A unique network interface card(NIC)
    A unique source IP
```

```

A unique source port number between 0-65535

Do you want to continue? [y/n,q] (y)
Please choose a network device:
    0) eth3
    1) eth2
    2) bond0
Select a network device: [0-2,q] 1
Input source IP address: [q] 192.168.111.2
Input netmask: [q] (255.255.255.0)
Input source port number between 0-65535: [q] (161) 163
Input read-write community: [q] (public) b
SNMP information verification:

    SourcePort: 163
    NIC: eth2
    SourceIP: 192.168.111.2
    Netmask: 255.255.255.0
    Rwcommunity: public

Is this information correct? [y/n,q] (y)
*    0) SNMP source IP and read-write community
*    1) SNMP trap IP
    b) Back to previous menu
Please select an item to add: [0-1,b,q] 1c

To configure SNMP trap IP the following information is required:

    One or more trap IP and its port number between 0-65535

Do you want to continue? [y/n,q] (y)
Input trap IP address: [q] 192.168.1.194
Input trap port number between 0-65535: [q] (162)
Add another trap IP? [y/n] (n)
SNMP trap IP and port information verification:

    Trap IP: 192.168.1.194    Port: 162

Is this information correct? [y/n,q] (y)
Please restart the HA system to take your modification into effect.

```



说明:

- 1、配置告警源的网卡、IP 地址、掩码和端口,即配置告警发起端信息。
- 2、配置 snmp 读写团体串,缺省为公用的 public,可按需修改成私有串团体。

3、配置告警 IP 和端口，即配置告警发往哪个主机；

snmp-delete

支持删除整个 snmp 的配置、已配置的告警源 IP、告警 IP。

格式：cli:~>snmp-delete

```
cli:~>snmp-delete
    0) Delete SNMP configuration
    1) Delete SNMP source IP
    2) Delete SNMP trap IP
    b) Back to previous menu
Please select an item to delete: [0-2,b,q]
```

snmp-modify

修改配置的告警源设定，如括网卡、IP 地址、掩码和端口等。

格式：cli:~>snmp-modify

```
cli:~>snmp-modify
    0) SourcePort: 163
    1) NIC: eth2
    2) SourceIP: 192.168.111.2
    3) Netmask: 255.255.255.0
    4) Rwcommunity: public
    b) Back to previous menu
Please select an attribute to modify: [0-4,b,q]
```

7.7 其它命令

check-config

用于检测集群的配置文件是否正常。

格式：cli:~>check-config

```
cli:~>check-config
```

```
Configuration file checked ok.
```

recover-remote-license

恢复授权文件到指定节点，常用于集群中某一节点故障恢复后的授权恢复。

格式：cli:~>recover-remote-license

```
cli:~>recover-remote-license
Input remote node name: [q] linux77
Recovery license success!
```

exit、quit

退出集群管理工具 cli。

格式：cli:~>exit

cli:~>quit

help

查看 cli 中非交互命令的用法。

格式：cli->help 命令

```
cli:~>help cluster-start

NAME
    cluster-start - start HA system and all services

SYNOPSIS
    cluster-start [hostname]

DESCRIPTION
    Start the HA system, if not specify the hostname, it will start
    the HA system
    on all nodes in the cluster. It will start the HA system on remote
    host at first,
    and start the HA system on localhost at last.

    hostname - the node name in the cluster
```

第 8 章 集群非交互式命令

集群提供以下非交互式命令：

命令	描述
nshacluster	配置和管理集群基本信息
nshamirror	配置和管理镜像资源
nshasnmpp	配置和管理snmp告警
nshastart	启动集群
nshastop	停止集群
nshaget	获取集群版本
nshares	配置和管理资源
nshasrv	配置和管理服务
nshastat	获取集群状态
nshahide	用于特殊功能开启



说明：

- 1) 命令参数中的符号说明：
{par}，大括号里面的参数为必选，至少选一项
[par]，中括号里面的参数为可选
<val>，尖括号里面的值为必填，以实际值替代
|，用于分隔若干可选选项，通常与大括号或中括号配合使用
- 2) 常规命令返回值：
255或-1，命令执行失败
0，命令执行成功
1，配置已经存在
2，配置不存在
5，命令直接退出
其他，系统错误码或特殊含义
- 3) 设置某个属性值为空或者分隔开的字符串，用双引号。如：
nshasrv -modify srv1 preferrednode ""
nshacluster -add-domain "node1 node2"
- 4) 其它命令参数可参考手册“NewStart Cluster 5.0 用户手册.pdf”

8.1 nshacluster

用来配置Cluster集群基本信息，配置信息保存在文件/etc/ha.d/ha.cf中。

Usage: nshacluster [OPTION] [ARGUMENT]...

-add {<cluster-name> | -node <node-name> -ip <ipaddr1 [ipaddr2]

```

[ipaddr3]> | -hb {-mcast | -mcast6 | -ucast | -ucast6 | -serial}
<nic-name> | -lockdisk <disk-name> -format {yes | no}}
-delete {-node <node-name> | -hb {-mcast | -mcast6 | -ucast | -ucast6 |
    -serial} <nic-name> | -lockdisk <disk-name>}}
-valid-lockdisk
-add-domain <nodes> [-lockdisk <disk-name> -format {yes | no}]
-del-domain <nodes>
-modify <key> <value>
-product-switch [<product-type>]
-upgrade
-display [code | product-type | <key>]
-help

```

参数解释:

参数	说明
-add <cluster-name>	添加集群名称，自定义，用于内部识别
-add -node <node-name> -ip <ipaddr1 [ipaddr2] [ipaddr3]>	添加集群节点及其IP，ip允许多个，空格分隔
-add -hb {-mcast -mcast6 -ucast -ucast6 -serial} <nic-name>	添加心跳链路 *) 支持ipv4和ipv6 *) 心跳类型为单播(ucast)、多播(mcast)和串口(serial) *) nic-name为操作系统识别的网卡名
-add -lockdisk <disk-name> -format {yes no}	添加仲裁盘 *) disk-name，共享存储名称（如/dev/sdc1） *) yes或no表示是否格式化该仲裁盘，一般都要填yes
-delete -node <node-name>	删除集群节点
-delete -hb {-mcast -mcast6 -ucast -ucast6 -serial} <nic-name>	删除心跳链路 *) 支持ipv4和ipv6 *) 心跳类型为单播(ucast)、多播(mcast)和串口(serial) *) nic-name为操作系统识别的网卡名
-delete -lockdisk <disk-name>	删除仲裁盘
-valid-lockdisk	使仲裁盘生效，用于重新初始化仲裁盘
-modify <key> <value>	修改集群配置信息，key及value值可用display查看到和修改的结果
-display [code product-type <key>]	不带参数显示集群配置信息，带参数显示相应的值，code表返回值，product-type查询集群类型
-add-domain <nodes> [-lockdisk <disk-name> -format {yes no}]	增加节点域； [-lockdisk <disk-name> -format {yes no}]可选择增加域节点的仲裁盘

-del-domain <nodes>	删除节点域
-product-switch [<product-type>]	用于集群类型切换或产品试用。 *)product-type类型用于产品试用, 类型有ha、mirror、cluster、RDS和rds五种
-upgrade	集群配置更新
-help	帮助参数, 显示所有命令

8.2 nshamirror

该命令用来配置和管理镜像, 配置信息保存在/etc/drbd.d/r0.res中。

Usage: nshamirror [OPTION] [ARGUMENT]...

```

-add -res <resname> -node <nodename> -ip <ipaddrs> [-port <portnum>]
      -disk <disk> [-metadisk <mdisk>]
-delete <resname>
-modify -res <resname> <key> <value>
-valid <resname>[-node <node>]
-format <fstype> -res <resname> [-mnt <mountpoint>]
-global-modify {global | <resname>} <section1> [<section2>] <key><value>
-sync-rate <rate> [<resname>]
-start [-res <resname>] [-node <node>]
-stop [-res <resname>] [-node <node>]
-set <res|dev> {-primary | -secondary} [-force]
-full-sync <res|dev>
-splitbrain-recovery <res|dev> -node <sync source>
-state <res|dev> [cs | ds | ro] [xmlonline [<nodename>]]
-resize <res|dev> [<size>K|M|G]
-add-link <res|dev> -node <node> -ip <ipaddr> [-port <portnum>]
-switch-link <res|dev> [-ip <ipaddr>] [-force]
-switch-mode <mode>
-open-node <res|dev> [-node <node>]
-create-script <res|dev>
-detect-mirror-mode <ipaddr|res|dev>
-status [<res>|<dev>] [detail]
-attach <res|dev>
-adjust [-node <node>]
-proxy-stop [-node <node>]
-proxy-start [-node <node>]
-display [code | <resname> | status-xml | online-xml | dump-xml [resname]]
-help

```

参数解释:

参数	说明
-add -res <resname> -node <nodename> -ip <ipaddrs> [-port <portnum>] -disk <disk> [-metadisk <mdisk>]	添加镜像资源: *) resname, 资源名, 可自定义 *) nodename, 主机名 *) ipaddrs, 镜像链路ip, 可配多个 *) portnum, 端口号, 可自定义 *) disk, 数据盘使用的磁盘设备 *) mdisk, 元数据盘使用的磁盘设备
-delete <resname>	删除已配置的镜像资源
-modify -res <resname> <key> <value>	修改已配置的镜像资源文件 (如: /etc/drbd.d/r0.res)
-valid <resname>	使已配置的镜像资源生效
-format <fstype> -res <resname> [-mnt <mountpoint>]	格式化已配置的镜像资源, 支持格式化后直接挂载到指定目录
-global-modify {global <resname>} <section1> [<section2>] <key> <value>	修改镜像资源的全局配置文件 (如: /etc/drbd.d/global_common.conf)
-sync-rate <rate> [<resname>]	临时修改指定镜像资源的同步速率
-start [-res <resname>] [-node <node>]	支持镜像资源在指定节点启动
-stop [-res <resname>] [-node <node>]	支持镜像资源在指定节点停止
-set <res dev> {-primary -secondary} [-force]	设置/强制设置镜像资源的角色为 primary或secondary
-full-sync <res dev>	指定某一镜像资源执行全同步操作
-splitbrain-recovery <res dev> -node <sync source>	镜像资源数据裂脑手动恢复, 可支持指定同步源
-state <res dev> [cs ds ro] [xmlonline [<nodename>]]	查询指定镜像资源的状态 (包括连接状态、底层设备状态和读写状态), 可支持获取指定节点的以上状态
-resize <res dev> [<size>K M G]	刷新镜像盘大小, 用于磁盘扩容
-add-link <res dev> -node <node> -ip <ipaddr> [-port <portnum>]	增加数据链路, 目前最多三条
-switch-link <res dev> [-ip <ipaddr>] [-force]	切换数据链路, 通常配合force参数来使用
-switch-mode <mode>	切换集群的模式为同城(sync)或远程(async)
-open-node <res dev> [-node <node>]	使指定节点的镜像配置生效, 用于2+1或2+2集群
-create-script <res dev>	重新创建镜像服务的管控脚本, 特殊用途
-detect-mirror-mode <ipaddr res dev>	检测当前集群适用的集群模式: 同城(sync)或远程(async)
-status [<res> <dev>] [detail]	

-attach <res dev>	重新连接镜像盘的底层设备
-adjust [-node <node>]	镜像参数即时生效
-proxy-stop [-node <node>]	停止procy模块
-proxy-start [-node <node>]	启动procy模块
-display [code <resname> status-xml online-xml dump-xml [resname]]	支持显示当前镜像命令执行返回码、镜像配置、镜像当前状态信息
-help	帮助参数，显示所有命令

8.3 nshasrv

该命令用来配置和管理服务，配置信息保存在文件：`/etc/ha.d/haservices.xml`

Usage: nshasrv [OPTION] [ARGUMENT]...

```

-add <srv>
-delete <srv>
-modify <srv> <attr> <value>
-link <parent srv> <child srv> <relationship> <depend type>
-unlink <parent srv> <child srv>
-start <srv> [<node>]
-stop <srv>
-enable <srv>
-disable <srv>
-migrate <srv> [<instance id>] <destination node>
-clearfault <srv> [<node>]
-checkcfg [<srv>]
-addinstance <srv> <instance id> [<instance name>] <nodes> <res>
-delinstance <srv> <instance id> [<res>]
-addmrrsrv <srv> {[<mirror resname>] | [<nic> <ipaddr> <netmask> <instance id>]}
-delmrrsrv <srv> <resname>
-display [-name <srv> | code | link | -instance <srv>]]
-help

```

参数解释：

参数	说明
-add <srv >	添加服务, 自定义服务名, 属性值为缺省值
-delete <srv>	删除已添加的服务
-modify <srv> <attr> <value>	修改服务属性（即“attr”值），属性可用display参数查看
-link <parent srv> <child srv> <relationship> <depend type>	配置服务间依赖。 relationship: local和global

	depend type: fixed和forced
-unlink <parent srv> <child srv>	删除服务间依赖
-start <srv> [<node>]	启动某个服务，启动节点可选，如不选，Cluster自动选择节点启动。
-stop <srv>	停止正在运行的服务
-enable <srv>	管控服务，前提是服务处于脱离状态
-disable <srv>	使服务脱离，前提是服务处于管控状态
-migrate <srv> [<instance id>] <destination node>	手动迁移服务到指定节点运行； instance id: 实例服务的编号
-checkcfg [<srv>]	检查服务是否配置完整，有效
-addinstance <srv> <instance id> [<instance name>] <nodes> <res>	增加多实例服务及实例编号； <instance name>: 实例服务名称；
-delinstance <srv> <instance id> [<res>]	删除实例服务
-addmrrsrv <srv> {[<mirror resname>] [<nic> <ipaddr> <netmask> <instance id>]}	增加镜像服务； mirror resname: 镜像资源名称； nic: 添加镜像服务的网卡资源 ipaddr: 添加镜像服务的ip资源 netmask: 添加镜像服务的子网掩码 instance id: 实例服务的编号
-delmrrsrv <srv> <resname>	删除镜像服务
-display [-name <srv> code link -instance <srv>]]	1、不带任何参数，显示整个集群中的所有服务名称 2、name，显示指定服务的配置信息 3、code，显示命令执行结果返回码 4、link，显示配置的服务依赖信息 5、-instance，显示实例服务
-help	帮助参数，显示命令

8.4 nshares

该命令用来配置和管理资源，配置信息保存在文件：/etc/ha.d/haservices.xml

Usage: nshares [OPTION] [ARGUMENT]...

-add <res> <type> <srv>
-delete <res>
-modify <res> <attr> <value>
-link <parent res> <child res>
-unlink <parent res> <child res>

```

-start <res>
-stop <res>
-enable <res>
-disable <res>
-critical <res>
-noncritical <res>
-checkcfg <res>
-display [-name <res> | type | code | link]
-help

```

参数解释:

参数	说明
-add <res> <type> <srv>	为指定服务添加资源 1、res, 资源名, 自定义, 可使用display参数参考类似命名规则 2、type为资源类型, 可使用display type 查看 3、srv, 已存在的服务名
-delete <res>	删除某个资源, 资源名(res)可使用display参数查看
-modify <res> <attr> <value>	修改资源属性 1、res, 可使用display参数查看 2、attr, 可用-display [-name <res>] 参数查看 3、value, 使用实际值代替
-link <parent res> <child res>	配置同一服务(资源组)的资源间依赖, 规则是父资源依赖子资源, 资源名(res)可使用display参数查看。
-unlink <parent res> <child res>	删除资源间依赖关系, 资源名(res)可使用display参数查看。
-start <res>	启动资源, 资源名(res)可使用display参数查看。
-stop <res>	停止某个资源. 资源名(res)可使用display参数查看
-enable <res>	管控资源, 前提是资源处理脱管状态, 重启集群生效, 资源名(res)可使用display参数查看。
-disable <res>	资源单独脱管, 前提是资源处理管控状态, 重启集群生效, 资源名(res)可使用display参数查看。
-critical <res>	将资源设为关键资源, 缺省资源均为关键资源, 资源名(res)可使用display参数查看。
-noncritical <res>	将资源设为非关键资源, 前提是资源

	处理管控状态，非关键资源异常不会触发服务倒换，资源名(res)可使用display参数查看。
-checkcfg <res>	检查资源是否配置完整，资源名(res)可使用display参数查看。
-display [-name <res> type code link]	1、display，显示所有资源名称 2、-name，显示资源的具体配置信息 3、Type，显示资源类型 4、Code，显示命令执行结果返回码 5、Link，显示资源依赖配置信息
-help	帮助参数，显示命令

8.5 nshasnmpp

用来配置snmp，配置信息保存在文件/etc/ha.d/haservices.xml和/etc/ha.d/NSHA-agent.conf

Usage: nshasnmpp [OPTION] [ARGUMENT]...

-add [-sourceip <ipaddr> | -trapip <ipaddr> [-port <trapport>]]

-delete [-sourceip | -trapip <ipaddr>]

-modify <attr> <value>

-display [type|code]

-help

参数解释:

参数	说明
-add [-sourceip <ipaddr> -trapip <ipaddr> [-port <trapport>]]	-add不带其他参数，默认生成SNMP的基本配置信息，可用display查看，并用modify修改属性。 1、-sourceip <ipaddr>，添加告警源IP 2、-trapip <ipaddr> [-port <trapport>]，添加告警目的IP和端口，支持执行多次的方式添加多个告警目的IP
-delete [-trapip <ipaddr>]	-delete不带参数，删除所有snmp配置信息。 1、-trapip <ipaddr>，删除告警目的IP
-modify <attr> <value>	修改SNMP 1、属性(attr)可用display或-display type查看 2、value，使用实际值代替
-display [type code]	1、-display，显示SNMP配置信息 2、type，显示SNMP可修改的属性 3、code，显示命令执行结果返回码

-help	帮助参数，显示命令的用法
-------	--------------

8.6 nshastart

该命令用来启动集群。

Usage: nshastart [OPTION] [ARGUMENT]...

[node name]

-display [code]

-help

参数解释：

参数	说明
[node name]	可选参数，如果带节点名启动，是指在该节点上启动集群，否则所有节点集群启动
-display [code]	不带任何参数或code 参数均显示命令执行结果返回码
-help	帮助参数，显示命令的用法

8.7 nshastop

该命令用来停止集群。

Usage: nshastop [OPTION] [ARGUMENT]...

[node name]

-display [code]

-help

参数解释：

参数	说明
[node name]	可选参数，如果带节点名停止，是指在该节点上停止集群，否则所有节点集群停止
-display [code]	不带任何参数或code 参数显示命令执行结果返回码
-help	帮助参数，显示命令的用法

8.8 nshastat

该命令用来获取集群的相关状态，包括服务状态、资源状态、节点状态、业务

链路状态、心跳链路状态等。

Usage: nshastat [OPTION] [ARGUMENT]...

-service <srv> <node>
-resource <res> <node>
-node <node>
-worklink <device> <node>
-hblink <device> <node>
-status
-abnormal <srv> [<node>]
-display {code | hbtype}
-help

参数解释:

参数	说明
-service <srv> <node>	获取服务在某节点上的状态
-resource <res> <node>	获取资源在某节点上的状态
-node <node>	获取节点状态
-worklink <device> <node>	获取业务链路在某个节点的状态
-hblink <device>	获取心跳链路状态
-status	显示整个集群状态
-abnormal <srv> [<node>]	查询业务资源状态是否异常, 定制化需求
-display {code hbtype}	code 参数显示所获取状态码的意义, hbtype参数显示心跳类型的几种类型;
-help	帮助参数, 显示命令的用法

8.9 nshaget

该命令用于获取集群信息, 包括版本、集群类型, 序列号及试用授权信息

Usage: nshaget [OPTION] [ARGUMENT]...

-version
-version-detail
-product [<sn>]
-license-days
-sn
-help

参数解释:

参数	说明
-version	获取集群版本, 如v3, v4, v5...
-version-detail	获取集群版本具体信息

-product [<sn>]	获取集群类型, 如mirror
-license-days	获取授权剩余天数(试用序列号)
-sn	获取SN号
-help	帮助参数, 显示命令的用法

8.10 nshahide

此命令用于开启集群特定功能。

Usage: nshahide [OPTION] [ARGUMENT]...

-add <key> <value>

-delete <key>

-display [code|key]

-help

Key has the following values:

key	value
singleNIC	1 cluster system can use a single NIC. default 0
samehost	1 cluster system hostname can be the same. default 0
snmpdelay	SNMP delay time of send trap message. default 0 seconds
snmptrapcheck	1 SNMP trap2sink(trap IP) active send trap message, inactive do not send. default 0
worklink_hb_trap	0 close worklink_hb SNMP trap message. default 1
snmptrapsrv	If the service is running, send trap message, otherwise do not send. service name as value
mirrorwait	Time to wait for remote mirror to be detected. default 0 seconds
high-security	1 mirror data high security. default 0
mirror_remote_mount	1 mirror device can be mount during synchronization. default 0
mirror_fence	0 mirror fence off. default 1
mirror_diskless	1 Running in diskless mode. default 0
mirror_congest	1 Auto deal with congestion. default 0
mirror_link_fast_switch	1 fast switching mirror link. default 0
srv_disabled_show	1 after the service is disabled, all resources show the original state. default 0
proxy_wait_time	Time waiting for mirror-proxy to brush data. default 60 seconds

参数解释:

参数	说明
-add <key> <value>	开启功能，key及value参考上述值
-delete <key>	关掉已开启的特定功能
-display [code key]	显示命令执行返回值或存在哪些key值
-help	帮助参数，显示命令的用法

第 9 章 集群告警与查询

NewStart 集群产品支持通过配置 snmp 告警的方式上报集群运行过程中出现的异常事件，方便管理员快速响应。同时支持在告警接受端向成功配置告警的集群环境发出集群信息查询请求。

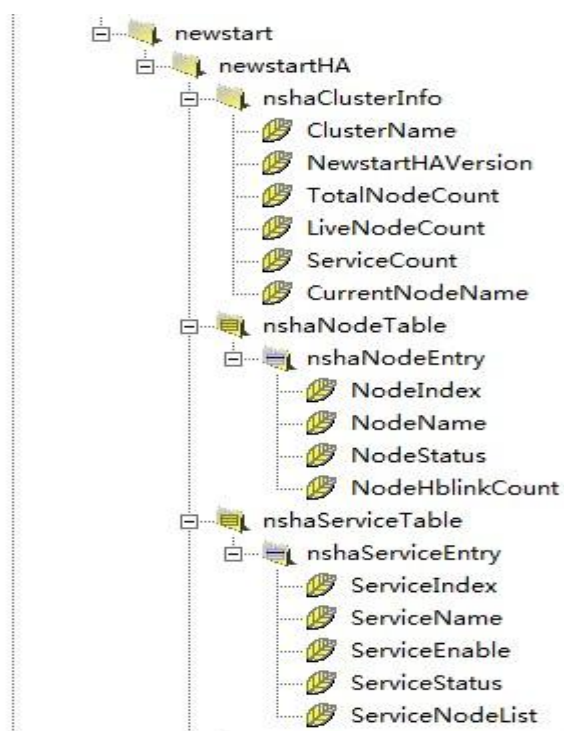
9.1 集群告警 oid 及说明（如果需要 mib 库，请联系我司）

Newstart Cluster SNMP Trap 接口				
告警类型	Trap Oid	严重级别	Trap Label Name	Trap Description
资源告警	1.3.6.1.4.1.36895.1.900.1.0.2	Warning	resourceMonitorTimeoutTrap	资源监视已超时
	1.3.6.1.4.1.36895.1.900.1.0.3	Error	ResourceNotGoingOfflineTrap	Newstart Cluster 无法使资源脱机
	1.3.6.1.4.1.36895.1.900.1.0.5	Warning	ResourceWentOnlineByItselfTrap	资源自动联机
	1.3.6.1.4.1.36895.1.900.1.0.6	Error	ResourceFaultedTrap	资源在某个节点出现故障
	1.3.6.1.4.1.36895.1.900.1.0.11	Information	resourceFaultRecoveryTrap	资源在某个节点恢复
服务告警	1.3.6.1.4.1.36895.1.900.2.0.1	Information	serviceOnlineTrap	服务启动成功
	1.3.6.1.4.1.36895.1.900.2.0.2	Information	serviceOfflineTrap	服务停止成功
	1.3.6.1.4.1.36895.1.900.2.0.3	Information	serviceDisabledTrap	服务的 enable 属性变为 no
	1.3.6.1.4.1.36895.1.900.2.0.4	Error	serviceFaultedTrap	服务故障
	1.3.6.1.4.1.36895.1.900.2.0.5	SevereError	serviceFaultedAndNowhereToFailoverTrap	指定的服务在可以使其联机的所有节点上出现故障。无法将该服务故障转移到任何节点。
	1.3.6.1.4.1.36895.1.900.2.0.6	Warning	serviceRestartingTrap	服务正在重新启动，以对持久性资源进入联机状态做出响应
	1.3.6.1.4.1.36895.1.900.2.0.7	Information	serviceInitiatingForSwitchingTrap	服务准备发生切换

	1.3.6.1.4.1.36895.1.900.2.0.9	Information	serviceRestInRspnToPers tResGoOnlineTrap	服务组正在重新启动，因为已从故障恢复了持久性资源。
	1.3.6.1.4.1.36895.1.900.2.0.12	Error	serviceOPTimeoutTrap	对端服务操作超时
	1.3.6.1.4.1.36895.1.900.2.0.13	Information	serviceDisabledTrap	服务的 enable 属性变为 yes
	1.3.6.1.4.1.36895.1.900.3.0.3	Error	clusterSystemInDangerTrap	运行 Newstart Cluster 的节点进入危险状态（心跳全段）
	1.3.6.1.4.1.36895.1.900.3.0.4	SevereError	clusterSystemFaultedTrap	Newstart Cluster 因故障而关闭
	1.3.6.1.4.1.36895.1.900.3.0.5	Information	clusterSystemJoinedClusterTrap	运行 Newstart Cluster 的节点加入集群
	1.3.6.1.4.1.36895.1.900.3.0.6	Information	clusterSystemExitedManuallyTrap	Newstart Cluster 已从先前正在运行的节点中正常退出。
心跳告警	1.3.6.1.4.1.36895.1.900.4.0.1	Information	heartbeatOnlineTrap	心跳恢复
	1.3.6.1.4.1.36895.1.900.4.0.2	Warning	heartbeatOfflineTrap	心跳断开

9.2 告警查询

目前集群仅支持“集群”、“节点”和“服务”的信息进行查询，具体以下面列表说明为准。



	查询项	描述
集群信息	ClusterName	集群名称
	NewstartHAVersion	集群版本号
	TotalNodeCount	集群中的节点数
	LiveNodeCount	集群中活动的节点数
	ServiceCount	集群中服务个数
	CurrentNodeName	反馈当前信息的节点名称
节点信息	NodeName	集群中的节点名称
	NodeStatus	集群中节点的状态
	NodeHblinkCount	心跳链路个数
服务信息	ServiceName	显示集群中服务名称
	ServiceEnable	服务的 enable 属性（可管控）
	ServiceStatus	服务的运行状态
	ServiceNodeList	服务的可运行节点

附录

附录 A—业务管控脚本编写规范说明

1 脚本编写前须知

1.1、业务管控脚本须包含 4 个参数：start（启动）、stop（停止）、forcedstop（强制停止）和 status（运行状态检测）。

1.2、脚本存放位置及权限

脚本文件原则上可以存放在操作系统任一目录下，但不建议放在磁盘阵列；另外脚本文件必须是unix格式（如修改命令：dos2unix 脚本名），必须具有可读和可执行权限（如修改命令：chmod a+rx 脚本名）。

1.3、集群操作脚本的机制（目前集群仅支持root用户执行业务管控脚本）

*) 业务管控脚本启动

服务启动的最后步骤是启动业务（即执行“脚本+ start 参数”），对启动操作的返回值不做检测；2s 后检测服务的所有资源，如果都返回 0，集群认为服务正常运行，启动成功。

*) 业务管控脚本检测

在服务缺省配置下(服务可管理)，集群启动完成后，主备机都会周期性进行服务检测（即执行“脚本+ status 参数”），如果返回值为 0，说明服务正常运行，返回非 0 认为服务停止。

*) 业务管控脚本停止

服务停止时会执行“脚本+ stop 参数”，对返回值不做检测, 2s 后会检测服务的所有资源，哪果都返回非 0，集群认为服务停止，停止成功。

2 脚本编写参考

正常情况下，业务本身的启、停都有相应的操作脚本，对于集群脚本的 start/stop，如果业务是以 root 用户起停，就直接调用业务脚本；如果以普通用户起停，格式如下：

```
su - user -c "业务脚本"
```

集群脚本的 forcedstop 一般是查找业务进程、强杀进程。
所以下面重点说明如何编写脚本的 status 部分。

ps+grep 类型脚本 (pgrep)

这是我们优先考虑的脚本编写方案。

1. 采用查进程名称的方式，只要进程名称存在就认为业务正常；对进程名的要求是具有唯一性，比如单词 java、tomcat 就不太适合此类脚本。

2. 统一规定书写格式：ps -fwu，参数含义如下：

-w Wide output.

-f does full-format listing. This option can be combined with many other UNIX-style options to add additional columns.

-u Select by effective user ID (EUID) or name. 指定用户 id 或者名称，如果是 root 用户，应该不写 -u，虽然没有错但是不符合常规。

3. grep 要求加参数 -w (word 单词匹配)，可以提高效率：

-w: Select only those lines containing matches that form whole words. The test is that the matching substring must either be at the beginning of the line, or preceded by a non-word constituent character. Similarly, it must be either at the end of the line or followed by a non-word constituent character. Word-constituent characters are letters, digits, and the underscore.

4. 不要再使用 “>/dev/null 2>&1”，这会导致 Cluster 日志无法记录脚本出错信息；但是可以使用 “>/dev/null”。

5. grep -vq grep 放在语句最后，以提高效率

6. 脚本 stop 禁止使用 kill -9 这种强杀进程的停止方式，可以使用 kill。

7. 脚本函数里不要调用 exit (直接退出)，改用 return 返回；

exit 只在主脚本里调用，具体可参考 Cluster 自带脚本/etc/ha.d/resource.d/myfaked.sh。

举例如下：

1) oracle 脚本：

```
ps -efww | grep tnslnsr | grep -v grep >/dev/null 2>&1
```

此脚本只适用于一个 oracle SID 的情况，效率低；

应改为：

```
ps -fwu oracle | grep -w ${LISTENER_NAME} | grep -vq grep
```

LISTENER_NAME 是脚本中定义的监听器名称。

2) sybase 脚本：

```
ps -fwu $SYBASE_USER | grep -w dataserver | grep -vq grep
```

如果不是这样写的，此问题可能会以其他现象的形式出现：停共享磁阵盘资源失败，强制停止才成功

```

nrm[d13073]: 2011/12/18_23:40:48 debug:
  stop_resource:/etc/ha.d/resource.d/Filesystem /dev/sda1 /ivasdata/sybase ext3 rw
sybase sybase 755 stop
umount: /ivasdata/sybase: device is busy  umount 报告 sybse 还在访问设备
然后不停地报
nrm[d13073]: 2011/12/18_23:40:53 debug: resource post-stop check,
udb_sybase_for_udb_mount_0 is NOT stopped.
原因是 Filesystem 发送 SIGTERM 停止 sybase。

```

3) zxin 脚本写法:

```
ps -efww |grep -v grep|grep -qw zxmoni
```

应改为:

```
ps -fwu zxinl0 | grep -w zxmoni | grep -vq grep
```

在一台服务器上可能有多个zxmoni进程运行,但是属于不同的用户,所以不能写成-efww。

查端口类型脚本

如tomcat, weblogic等,这类型脚本运行需要java虚拟机,如果用grep查询输出结果很长,请参考/etc/ha.d/resource.d/tomcat_example.sh脚本快速编写脚本。

举例: 查询属于 oracle 用户并且监听 1521 端口的进程

```

[root@e222 ~]# lsof -a -i :1521 -u oracle
COMMAND  PID   USER   FD   TYPE DEVICE SIZE NODE NAME
oracle   9932  oracle 18u  IPv4  22313      TCP e222:59034->e222:ncube-lm (ESTABLISHED)
tnslsnr  10134  oracle  8u  IPv4  22260      TCP *:ncube-lm (LISTEN)
tnslsnr  10134  oracle 11u  IPv4  22314      TCP e222:ncube-lm->e222:59034 (ESTABLISHED)

```

3 手动调试脚本

复杂脚本的运行往往需要连接 IP 和访问磁阵,并设置磁阵挂载点的宿主和属性,操作比较麻烦,调试脚步的快捷步骤如下:

1. 配置好后启动 Cluster, 等服务启动, 从/var/log/ha-log 提取相关信息

```
#grep "start_res" /var/log/ha-log
```

```
grm[14591]: 2013/06/28_16:53:21 debug:
```

```
start_resource:/etc/ha.d/resource.d/Filesystem /dev/sdc2 /home/db disk ext3 rw
oracle oinstall 755 90 0 0 0 start
```

```
grm[14591]: 2013/06/28_16:53:21 debug: start_resource:/bin/bash --login
```

```
/home/script/oracle.sh start
```

```
grm[14591]: 2013/06/28_16:53:21 debug:
start_resource:/usr/lib/newstartha/resource.d/LastingRes -t netcard -d bond0 -o start
-i 192.168.1.192 -i 192.168.1.193 -i 192.168.1.247
grm[14591]: 2013/06/28_16:53:23 debug:
start_resource:/etc/ha.d/resource.d/IPaddr 192.168.11.13/24/bond0 start
```

2. 停止集群，粘帖以上红色粗体字启动 IP 和挂载点，然后以 root 用户开始调试脚本了，这是模拟 Cluster 的操作，比较快捷。

3. 开始调试脚本（以 oracle 脚本调试为例）

1) 测试 oracle 启动

```
#!/etc/ha.d/resource.d/ora_script.ps start
#sleep 30; /etc/ha.d/resource.d/ora_script.ps status; echo $?
由于 oracle 启动需要一些时间，所以 sleep 30 秒后检查状态，确保启动成功$?
返回 0，否则返回非 0（通常等于 1）。
```

2) 测试 oracle 停止

```
#!/etc/ha.d/resource.d/ora_script.ps stop
#sleep 60; /etc/ha.d/resource.d/ora_script.ps status; echo $?
确保停止成功后$?返回非 0（通常等于 1），失败返回 0 说明 oracle 未完全停止
```

3) 如果手动调试通过，但是 Cluster 启动失败？

这一般与定义的用户环境变量有关。请在脚本中引用相关配置文件：

小数点+空格+配置文件，比如：

```
. /etc/profile
. /home/oracle/.bash_profile （Redhat 等发行版）
. /home/oracle/.profile （Suse 等发行版）
```

4. 调试完成后，停止手动启动的业务、挂载点、IP，启动集群然后倒换服务

5. 如果 Cluster 启动服务成功，但是切换到备机失败？

修改配置，设定服务在备机启动，是否成功？

1) 成功，说明脚本 status 部分存在问题

2) 失败，cluster-stat 显示 str (A)，说明备机的环境存在问题（或者与主机不一致：比如 oracle 的用户 id 在两个节点不同）；

4 脚本编写方案

一般业务有自己的启动和停止脚本，对于集群脚本的 start/stop，如果业务是以 root 用户起停，就直接调用业务脚本；如果以普通用户起停，格式如下：

```
su - user -c "业务脚本"
```

集群脚本的 forcedstop 一般是查找业务进程、强杀进程。

所以下面重点说明如何编写脚本的 status 部分。

ps+grep 类型脚本 (pgrep)

这是我们优先考虑的脚本编写方案。

1. 采用查进程名称的方式，只要进程名称存在就认为业务正常；对进程名的要求是具有唯一性，比如单词 java、tomcat 就不太适合此类脚本。

2. 统一规定书写格式：ps -fwu，参数含义如下：

-w Wide output.

-f does full-format listing. This option can be combined with many other UNIX-style options to add additional columns.

-u Select by effective user ID (EUID) or name. 指定用户 id 或者名称，如果是 root 用户，应该不写-u，虽然没有错但是不符合常规。

3. grep 要求加参数-w (word 单词匹配)，可以提高效率：

-w: Select only those lines containing matches that form whole words. The test is that the matching substring must either be at the beginning of the line, or preceded by a non-word constituent character. Similarly, it must be either at the end of the line or followed by a non-word constituent character. Word-constituent characters are letters, digits, and the underscore.

4. 不要再使用 “>/dev/null 2>&1”，这会导致 Cluster 日志无法记录脚本出错信息；但是可以使用 “>/dev/null”。

5. grep -vq grep 放在语句最后，以提高效率

6. 脚本 stop 禁止使用 kill -9 这种强杀进程的停止方式，可以使用 kill。

7. 脚本函数里不要调用 exit (直接退出)，改用 return 返回；

exit 只在主脚本里调用，具体可参考 Cluster 自带脚本/etc/ha.d/resource.d/myfaked.sh。

举例如下：

1) oracle 脚本：

```
ps -efww | grep tnslnsr | grep -v grep >/dev/null 2>&1
```

此脚本只适用于一个 oracle SID 的情况，效率低；

应改为：

```
ps -fwu oracle | grep -w ${LISTENER_NAME} | grep -vq grep
```

LISTENER_NAME 是脚本中定义的监听器名称。

2) sybase 脚本：

```
ps -fwu $SYBASE_USER | grep -w dataserver | grep -vq grep
```

如果不是这样写的，此问题可能会以其他现象的形式出现：停共享磁阵盘资源失败，强制停止才成功


```

nrmd[13073]: 2011/12/18_23:40:48 debug:
  stop_resource:/etc/ha.d/resource.d/Filesystem/dev/sda1/ivasdata/sybase ext3 rw
sybase sybase 755 stop
umount: /ivasdata/sybase: device is busy  umount 报告 sybse 还在访问设备
然后不停地报
nrmd[13073]: 2011/12/18_23:40:53 debug: resource post-stop check,
udb_sybase_for_udb_mount_0 is NOT stopped.
原因是 Filesystem 发送 SIGTERM 停止 sybase。

```

4) zxin 脚本写法:

```
ps -efww |grep -v grep|grep -qw zxmoni
```

应改为:

```
ps -fwu zxin10 | grep -w zxmoni | grep -vq grep
```

在一台服务器上可能有多个zxmoni进程运行,但是属于不同的用户,所以不能写成-efww。

查端口类型脚本

如tomcat, weblogic等,这类型脚本运行需要java虚拟机,如果用grep查询输出结果很长,请参考/etc/ha.d/resource.d/tomcat_example.sh脚本快速编写脚本。

举例: 查询属于 oracle 用户并且监听 1521 端口的进程

```

[root@e222 ~]# lsof -a -i :1521 -u oracle
COMMAND  PID   USER   FD   TYPE DEVICE SIZE NODE NAME
oracle   9932  oracle 18u   IPv4  22313      TCP e222:59034->e222:ncube-1m (ESTABLISHED)
tnslsnr  10134  oracle 8u    IPv4  22260      TCP *:ncube-1m (LISTEN)
tnslsnr  10134  oracle 11u   IPv4  22314      TCP e222:ncube-1m->e222:59034 (ESTABLISHED)

```

附录 B—在线注册系统使用说明

1、访问在线注册系统

使用主流浏览器 (IE、firefox、chrome 等) 访问中兴新支点官网 (<http://www.gd-linux.com>), 点击“帮助”----“在线注册”, 跳转在线注册系统, 参考截图如下:



2、用户注册、用户登陆和用户密码找回

用户登陆: 浏览器输入访问地址---->跳转到用户登陆界面---->输入帐户信息登陆



用户注册: 浏览器输入访问地址---->跳转到用户登陆界面---->点击“xxx 注册”---->跳转用户注册页面---->输入相关信息完成新用户注册 (注意邮箱必须是可用的)

注册信息

 点击按钮进行验证 

注册

返回

用户密码找回：浏览器输入访问地址——跳转到用户登陆界面——点击“xxx 密码找回”——跳转密码找回界面，用户密码会重置并发送到注册邮箱

找回密码

 点击按钮进行验证 

发送邮件

返回

3、序列号（SN）在线注册

3.1 登陆——输入序列号（SN）——自动校验序列号可用性，显示“产品名称”，“版本信息”等

尊敬的用户：

感谢您购买NewStart软件产品，请填写以下软件注册信息，提交成功后，从“序列号管理-序列号列表-下载license”处点击[下载license](#)文件，license会自动保存到系统中，以后可登陆下载。如有问题请致电(+86)133-1877-2180,(+86)400-033-0108. 联系人：HA技术支持，邮箱：ha@gd-linux.com。

注意：不支持注册已注册的序列号和试用序列号

注册序列号

序列号	<input type="text" value="REDACTED"/>
产品名称	<input type="text" value="NewStart Mirror"/>
产品版本	<input type="text" value="V4.0"/>
产品描述	<input type="text" value="新支点镜像双机版本4.0"/>

待授权文件1 未选择文件。

待授权文件2 未选择文件。

待授权文件是指newstartha.key文件

申请者信息

申请人	<input type="text"/>
项目信息	<input type="text"/>
手机号码	<input type="text"/>
	<input type="text"/>

3.2 然后填写申请人信息-----点击最下方“注册”按钮，成功注册后会有提示，并告之 license 的获取方式并跳转到相关下载页面。

注册序列号

序列号	<input type="text" value="REDACTED"/>
产品名称	<input type="text" value="NewStart Mirror"/>
产品版本	<input type="text" value="V4.0"/>
产品描述	<input type="text" value="新支点镜像双机版本4.0"/>
待授权文件1	<input type="button" value="浏览..."/> newstartha_node1.key
待授权文件2	<input type="button" value="浏览..."/> newstartha_node2.key

申请者信息

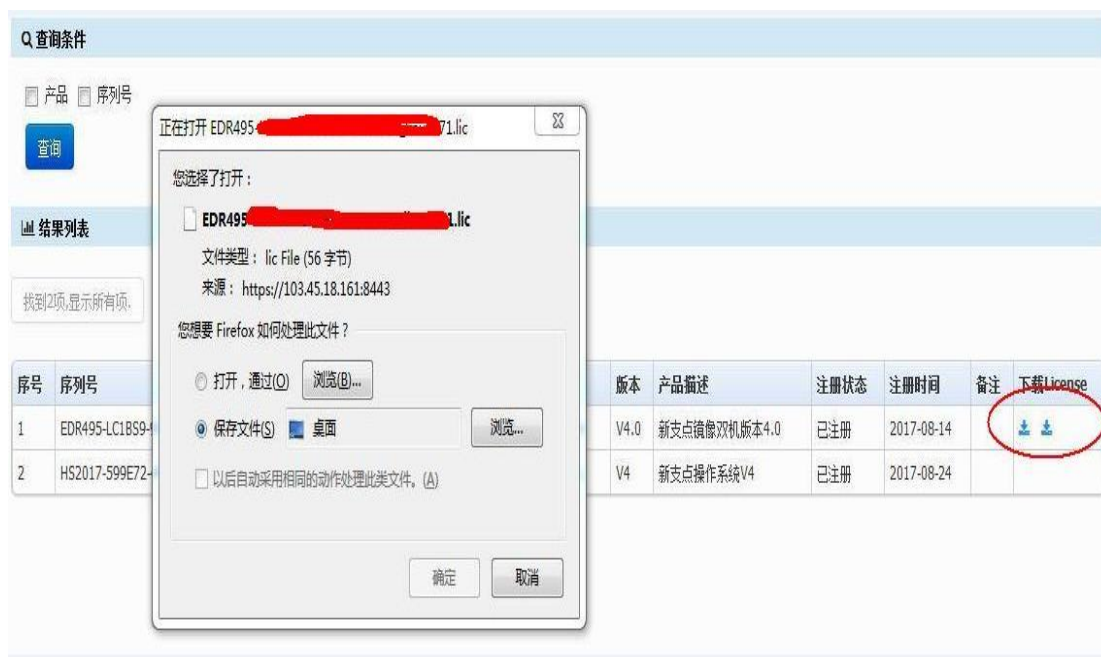
申请人	<input type="text" value="客户A"/>
项目信息	<input type="text" value="REDACTED项目"/>
手机号码	<input type="text" value="139xxxxxxxx"/>
邮箱	<input type="text" value="test@zte.com.cn"/>
公司或部门	<input type="text" value="REDACTED"/>

注册

返回



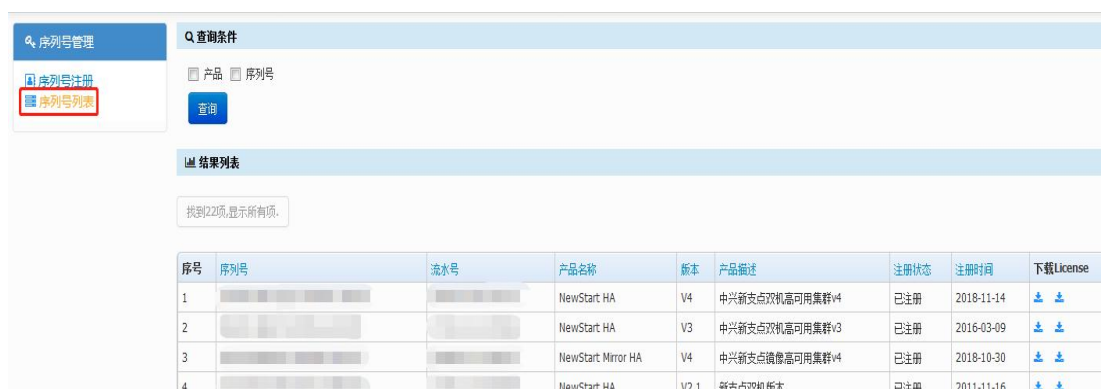
3.3 下载成功注册后的 license。



下载的 license 文件重命名为 newstartha.lic 后以二进制方式上传到对应服务器的/etc/ha.d/lic 目录下即完成注册。

4、历史注册信息查询与 license 找回

用户登陆----点击左侧“序列号列表”----右边页面就会显示已注册的序列号并支持序列号查询，也提供当时注册的 license 下载



附录 C—典型故障排除与恢复

■ 系统假死故障

系统假死是一种特殊的故障场景，至今仍是业界一大难题，当双机环境中出现可能故障时，如何判断及做出规范操作？

从 Cluster 角度出发，可以从现场的以下表现来做出判断：

故障节点无法通过 ssh 和 telnet 进行连接，无法响应其它应用请求，控制台无响应，并且系统能继续收发双机心跳包(这点可以从双机中正常节点系统上做出简单判断，该节点 ha-log 日志中没有打印对方节点 dead 的异常信息)。

推荐操作：

重启假死服务器使得备机接管服务（可考虑采用硬重启方式），不推荐手动倒换服务到备机服务或者此情况下重启备份服务器。

■ 服务 unstable 故障

在双机集群运行过程中，可能会遇到一种故障现象：服务 unstable。当出现该状况时请务必注意，因为服务出现该状态意味着其已脱离了 Cluster 的管控，即集群不管控该服务所有资源，所括业务 IP，磁阵，应用等，属高风险表现。进入 cli 查询集群状态时服务会有以下类似的打印：

Service status:

Name	susell-3	susell-4	Enable
myfaked	unstable:stp(A)	unstable:stp(A)	YES

推荐操作：

- 1、 排查引起服务 unstable 的原因，可通过 Cluster 日志定位，有需要把两节点的 ha-lo 日志压缩打包发给厂家分析，公共技术支持邮箱：ha@gd-linux.com。
- 2、 故障定位及解决后，可执行 clear-fault 清除异常状态，最后执行 service-start 来启动服务。

■ 集群裂脑/心跳链路 offline 故障

裂脑指的是集群中的节点彼此失去了联系，但集群软件都正常运行，这样将导致资源的竞争，可能会损坏数据。

裂脑故障常规表现如下：

1. 所有业务运行正常；
2. 节点各自运行正常，但显示其它节点 down 或 unknow；
3. 心跳链路显示 ‘—’，备机服务一般会出现 **stopped:str(B)** 的状态显示，下面是主机执行 `cli->cluster-stat` 的结果。

NodeName	Status	HeartbeatLink	ServiceName	ServiceStatus	Enable
susell-3	UP	eth3, eth2, — LDisk (ONLINE)	oracle	running	YES
susell-4	DOWN	—, —, — LDisk (ONLINE)	oracle	unknown	YES

推荐操作：

裂脑故障重点是解决心跳链路 offline 问题，可按以下方法排查。

1、首先确保 ping 通<对方节点心跳链路网卡的 IP>

心跳链路网卡与业务链路要属于不同的 vlan，或者配置成不同网段的 IP 地址。心跳链路能够 ping 通，`cli->cluster-stat` 却显示 offline，因为它们使用的协议不同，Ping 使用 ICMP 协议，Cluster 心跳链路使用 UDP 组播协议。

2、启动了操作系统防火墙

运行 `iptables -L` 查看防火墙规则。关闭防火墙或者打开 udp 端口：694；

3、尝试用直连网线连接两台服务器的心跳网卡，配成私有网络；看能否解决(多节点集群需走交换机，不能直连)。

4、对于刀片服务器/vmware 虚拟机环境

要安装 3.0 以上版本才能解决，因为实现了 IGMP 协议(俗称“三层组播协议”)。

5、另外一种分析方法是查看网卡配置和抓包分析：

```
# tcpdump -ni X1 port 694
```

```
# ifconfig X1
```

PS：X1 换成实际网卡名，比如 eth1, bond0, trunk1 等，如果有多张网卡，对每张网卡都需抓包分析。

例如：

```
[root@linux173 ~]# ifconfig eth1
```

```
eth1      Link encap:Ethernet  HWaddr 00:0A:EB:30:81:24
          inet      addr:192.168.10.173          Bcast:192.168.10.255
Mask:255.255.255.0
          inet6 addr: fe80::20a:ebff:fe30:8124/64 Scope:Link
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:2635153 errors:0 dropped:0 overruns:0 frame:0
RUNNING 表示网卡工作正常，没插网线时网卡没有此标志；
```

抓报举例：

```
[root@e222 ~]# tcpdump -ni eth0 port 694|grep "226.0.12.17"
listening on eth0, link-type EN10MB (Ethernet), capture size 96 bytes
15:51:19.685283 IP 192.168.1.222.48628 > 226.0.12.17.ha-cluster: UDP,
length 190
15:51:19.731718 IP 192.168.1.223.49469 > 226.0.12.17.ha-cluster: UDP,
length 191
...
```

以上说明抓到 192.168.1.222, 192.168.1.223 的心跳包（正常），因为局域网内存在多套集群，抓包时 grep "226.0.12.17" 只显示本集群的心跳包，而组播地址 226.0.12.17 可以使用 cli->cluster-show 命令查看获取，如：

```
cli:~>cluster-show
```

Configuration of the cluster:

```
watchdog_fork: off
watchdog_load: 0
node: e222 192.168.1.222
node: e223 192.168.1.223
logfile: /var/log/ha-log
keepalive: 1
deadtime: 15
mcast: eth1 226.0.12.17 694 1 0
```

如果心跳 eth1 不通也没有抓到从 eth1 发出的心跳包，这说明对方 eth1 的数据包没有到达本节点，请用户检查网络通畅情况；反之，如果心跳 eth1 不通但是有抓到从 eth1 发出的心跳包（极少见），看重启 Cluster 能否解决。

6、如果还是没有解决，需要人工分析，请把如下发给我们分析：

- * cli 命令输出
- * /etc/ha.d/ha.cf 和 haservices.xml 文件
- * 2 个节点所有心跳网卡的 ifconfig 输出和抓包结果的输出

■ 集群启动失败故障

集群命令行管理工具 cli 上启动集群 (cluster-start) 时, 可能会遇到启动失败的情况, 下面罗列一些表现及相关排查解决方法。

1. 双机启动, 均报 NewStartHA failure [rc=6]. ..failed.

推荐操作:

此报错常见于 Cluster 获取配置文件中的一些资源失败, 此异常一般需要结合 /var/log/ha-log 日志来定位, 现场人员可以筛选其中的 error 信息来快速定位, 当然, 有需要也可把相关资料发给厂家来协助定位。

2. 执行 cluster-start 时报 rc=3 错误, 如下所示:

```
cli:~>cluster-start
[susell-4]Starting High-Availability services:
Configuration file checked ok.
NewStartHA failure [rc=3]. ..failed
License file invalid.
```

```
Configuration file checked ok.
[susell-3]Starting High-Availability services:
NewStartHA failure [rc=3]. ..failed
License file invalid.
```

推荐操作:

此类报错常见于 license 文件(/etc/ha.d/lic/newstartha.lic)不存在或无效, 解决方法是重新生成 key 文件, 发到 ha@gd-linux.com 重新申请 license, 具体操作方法可以参考附录 C—常见问题及解答的第三点。

3. cli 中执行 cluster-start 时报 WARN: connect 'xxx(主机名)' tserver failed, please see ha-log for details.

推荐操作:

此提示常见于双机的两节点的 tsever 进程无法建立正常通信, 一般是由于以下两原因引起:

- 操作系统防火墙影响, 可以尝试关闭防火墙试试。
- tsever 进程没有起来 (ps -ef | grep tsever), 此时需要手动把 tsever 启起来, 参考操作: /usr/lib/newstartha/tsever &

■ 数据裂脑故障

在某种异常情况下，主备机数据出现不一致，主备无法切换服务，此时需要人工干预修复。

数据裂脑故障常规表现如下：

1、查看状态

cat /proc/drbd

主备两个节点出现以下异常状态时，则可能出现数据裂脑

```
[root@localhost ~]# cat /proc/drbd
version: 8.3.16 (api:88/proto:86-97)
GIT-hash: a798fa7e274428a357657fb52f0ecf40192c1985 build by root@localhost.localdomain, 2017-02-22 16:22:31
0: cs:StandAlone ro:Primary/Unknown ds:UpToDate/Outdated r-----
ns:0 nr:0 dw:15100 dr:135697 al:54 bm:0 lo:0 pe:0 ua:0 ap:0 ep:1 wo:f oos:13940

[root@localhost ~]# cat /proc/drbd
version: 8.3.16 (api:88/proto:86-97)
GIT-hash: a798fa7e274428a357657fb52f0ecf40192c1985 build by root@localhost.localdomain, 2017-02-22 16:22:31
0: cs:StandAlone ro:Secondary/Unknown ds:Outdated/DUnknown r-----
ns:0 nr:0 dw:0 dr:0 al:0 bm:0 lo:0 pe:0 ua:0 ap:0 ep:1 wo:f oos:868
```

或节点 1: cs"StandAlone", 节点 2: cs"WFCConnection"状态。

2、排查环境

先检查下镜像链路是否正常，端口是否被拦截，防火墙的配置等因素后再往下操作；

3、开始操作修复

3.1、若节点 1: cs"StandAlone", 节点 2: cs"WFCConnection", 此时先在 StandAlone 节点上尝试链接资源：

drbdadm connect r0

cat /proc/drbd

查看状态，若开始出现同步，则等待同步完恢复；若状态变化：节点 1: cs"StandAlone", 节点 2: cs"StandAlone", 则需手动修复，见 3.2；

3.2、若两边 cs 状态均为"StandAlone"，则需手动干预修复，在 **primary** 节点（即业务运行节点）运行：

nshamirror -splitbrain-recovery r0 -node `hostname`

cat /proc/drbd

观察状态，若数据同步，再到 Connected 状态即恢复正常；若此步骤无法修复，见 3.3；

3.3、若 3.2 无法修复，考虑重新初始化备机镜像盘。**禁止主、备都进行此操作，会把镜像盘所有数据擦掉，数据不可恢复，如有需要请联系我们再进行下面的操作。**

在备节点运行:

```
# nshamirror -stop -res r0 -node `hostname`
```

```
# nshamirror -valid r0
```

```
# nshamirror -start -res r0 -node `hostname`
```

cat /proc/drbd 观察状态, 若出现数据同步, 再到 Connected 状态即恢复正常

```
[root@linux21 home]# cat /proc/drbd
version: 8.4.11-1 (api:1/proto:86-101)
GIT-hash: 66145a308421e9c124ec391a7848ac20203bb03c build by root@linux171, 2018-11-26
14:44:50
0: cs:Connected ro:Primary/Secondary ds:UpToDate/UpToDate A r-----
   ns:67572 nr:0 dw:66548 dr:3080916 al:22 bm:0 lo:0 pe:0 ua:0 ap:0 ep:1 wo:f oos:0
```

■ web 页面无法正常显示

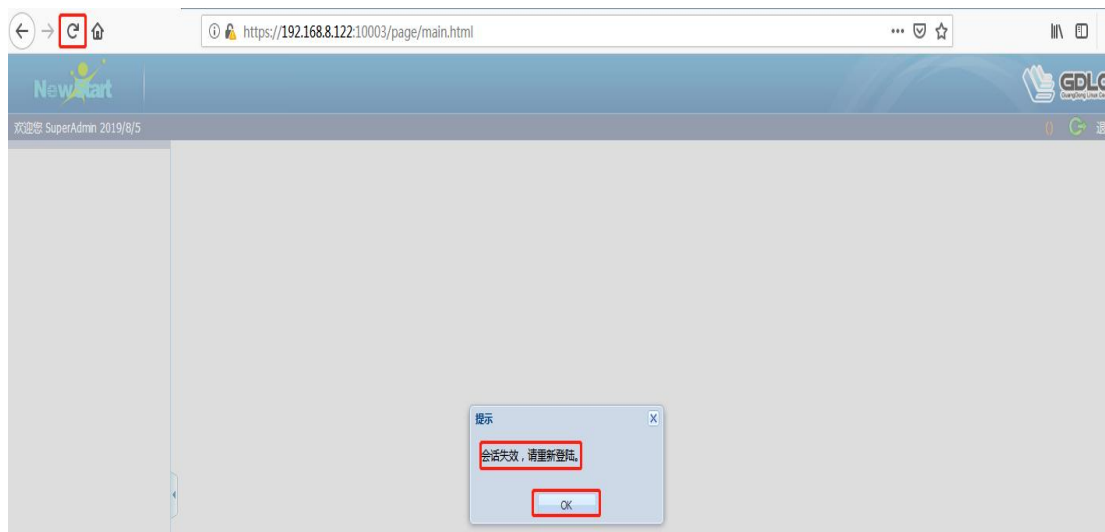
表现如下:

页面一段时间没有进行操作后, 页面显示不正常或者无响应。



推荐操作:

1、检查页面时效性, 目前 web 设定 10 分钟无命令调取或者无操作后, 页面会失效, 需要刷新浏览器, 页面会自动返回登录页面, 重新登录即可;



2、检查服务器是否将 web 程序关闭: webadminserver status;

附录 D—技术支持

广东省 Linux 公共服务技术支持中心（广东中兴新支点技术有限公司）衷心感谢您使用本公司集群产品。您的选择是我们前进的动力，您的支持是我们成功的基石。在使用我们的产品时，如发现什么问题或有什么建议，请与我们联系，我们将不胜感谢。

1、 与技术支持联系

在联系技术支持之前，请确保您的计算机符合产品文档中所列的系统要求，请提供以下信息：

■ 问题说明：

- 1) [命令 hareport](#) 生成的文件和 `cli->show-status` 或者 `cluster-stat` 输出
- 2) 描述问题首次出现的时间及当时所做的操作
- 3) 最近所做的软件配置更改和网络更改

■ 可用内存、磁盘空间

■ 网络结构

■ 路由器、网关和 IP 地址信息

2、 联系方式

技术服务邮箱： ha@gd-linux.com

技术服务电话：（86）13318772180，400-033-0108 转 2

公司电话：（86）020-87048587/87048575/87048576

公司传真：（86-20）87048317

公司名称：广东中兴新支点技术有限公司

公司网址： <http://www.gd-linux.com>

地址：广州市天河区高唐软件园基地高普路 1021 号 E 栋 601 室

邮编：510663