



# 新支点高可用集群软件

## **NewStart Cluster**

### 用户手册

广东中兴新支点技术有限公司

**Guangdong Zhongxing NewStart Technology Ltd.**

本手册适用于集群 5.x 版本

# 目 录

前 言 .....	3
本书约定 .....	3
正式版包装组件 .....	3
第 1 章 概 述 .....	4
第 2 章 安装前准备 .....	5
2.1 集群搭建流程预览 .....	5
2.2 相关概念 .....	5
2.3 硬件(虚拟硬件)需求列表 .....	8
2.4 操作系统支持列表 .....	8
2.5 架构规划 .....	10
2.6 配置环境 .....	12
2.7 编写服务的集群脚本 .....	14
第 3 章 安装 .....	15
3.1 产品部署 .....	15
3.2 产品注册 .....	16
3.3 产品卸载 .....	17
第 4 章 通过 cli 进行集群配置管理 .....	18
4.1 集群配置规划 .....	18
4.2 初始化配置 .....	19
4.2.1 双节点集群初始化配置 .....	19
4.2.2 多节点集群初始化配置 .....	32
4.2.3 磁盘镜像集群初始化配置 .....	36
4.3 修改配置 .....	40
4.4 集群管理 .....	53
4.4.1 节点的管理 .....	53
4.4.2 服务的管理 .....	54
4.4.3 资源的管理 .....	57
4.5 集群监控 .....	59
4.6 配置 SNMP 告警监控 (选配) .....	63
4.7 日志收集和其他 .....	68
第 5 章 通过 web 界面进行集群配置管理 .....	69
5.1 web 登陆 .....	69
5.2 双节点集群配置 (多节点集群配置类似) .....	76
5.3 磁盘镜像集群配置 .....	83
5.5 集群启动 .....	91
5.6 服务倒换测试 .....	92
第 6 章 集群的验收测试 .....	93
I. 附录 .....	1
附录 A—Clusterl 命令快速参考 .....	1
附录 B—SNMP 告警消息类型与查询说明 .....	23
附录 C—Cluster 应用脚本编写规范 .....	26

1、Cluster 操作脚本的机制 .....	26
2、手动调试脚本 .....	27
3、脚本编写方案 .....	28
附录 D—Newstart Cluster 在线注册系统使用说明 .....	30
附录 E—常见问题 .....	34
1. 如何把集群试用版本转成商用版本? .....	34
2. 运行 cli 出现 License file invalid 错误提示,如何处理? .....	34
3. 修改主机名是否影响 Cluster 配置? .....	34
4. Cluster 是否提供 Web 管理工具? .....	34
5. 如何关闭 Cluster 的 Web 管理工具? .....	35
6. 是否支持集群告警 (snmp 告警) 监控, 如何配置? .....	35
7. 操作系统重装, 是否需要重新申请 license? .....	35
8. Cluster 是否具备 LVM 锁盘功能? .....	35
9. 以下安全加固操作是否会影响 Cluster 的运行? .....	35
10. Web 登录忘记密码如何进行重置? .....	35
附录 F—典型故障排除与恢复 .....	36
■ 系统假死故障 .....	36
■ 服务 unstable 故障 .....	36
■ 集群裂脑/心跳链路 offline 故障 .....	37
■ 集群启动失败故障 .....	39
■ 数据裂脑故障 .....	40
■ web 页面无法正常显示 .....	44
附录 G—版本一键切换及授权方式变动说明 .....	45
附录 H—版本升级操作步骤 .....	46
1. 注意事项 .....	46
2. 升级前准备 .....	46
3. 升级操作 .....	49
技术支持联系 .....	57

# 前言






## 本书约定

### 1. 组织和印刷约定

格式	意义
输入的命令	<b>等宽粗体</b> （新宋体小四）
命令输出	等宽字体（新宋体小四）
命令的注释	<b><i>等宽粗斜体</i></b> （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
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平台下的高可用集群产品，目前本产品支持配置共享、磁盘镜像和多节点集群。

本手册是为新支点高可用集群系统的用户编写的，将引导您完成环境的搭建，然后通过简单操作命令行界面，方便快捷地进行配置和管理。本手册将以图文相辅、通俗易懂的形式，详细地介绍集群的安装使用过程。

光盘（或者 iso 文件）内容

文件/目录	内容说明
doc	文档目录，包括用户手册和版本发布说明
conf	配置文件
kernels	镜像驱动包
packages	集群程序包
tool	安装过程程序包
web	web 管理工具程序包
install	安装脚本
uninstall	卸载脚本
upgrade	升级脚本

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：<https://www.gd-linux.com/html/xiazai/ha/>

## 第 2 章 安装前准备

### 2.1 集群搭建流程预览

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。
2. 按照设计好的集群架构进行软硬件配置。
3. 安装集群软件并获取 License 文件。
4. 配置集群软件，该操作可在命令行或web界面中完成。
5. 安装业务软件（比如数据库系统），确保所有服务器的业务软件环境一致。
6. 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。
7. 启动集群，确认集群运行正常并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。

### 2.2 相关概念

- 节点（node）：指运行高可用集群软件的计算机。
- 主节点（也称主用节点）：通常指运行业务/服务的节点。
- 备份节点（也称备用节点）：用于监控主节点状态，一旦主节点出现故障，备份节点将接管主节点的业务/服务，从而备份节点成为主节点，而原先的主节点也将变为备份节点。
- 心跳链路（heartbeat link）：维持集群内部互联，传送心跳信息的链路。
- 服务（service）：与业务系统相关的一组资源的集合，包括：管理业务的管控脚本(application)，网络资源（NIC 和 IP），存储资源；服务可以是一种或多种资源的组合。
- 主备方式（active-standby）：指集群中服务在主节点上运行，备用节点服务处理备用状态。
- 互备方式（active-active）：指集群节点上都有服务在运行，这个时候每个节点既是主节点，同时也是备份节点，互相备份。
- 优先节点（preferred node）：指集群运行中，服务优先在该节点运行，当该节点故障恢复时，服务会回倒到优先节点。

- 服务故障迁移（服务倒换，failover）：是指当服务在主节点上运行出现故障，它将停止服务；然后在备用节点启动服务。
- 服务迁移的时间：是指服务不可用的时间，对于集群中节点 node1，node2 来说，如果服务在 node1 上运行，目前出现故障的时刻为 T1，那么当 node2 检测到 node1 出现故障时，开始在 node2 上启动该服务直到该服务可用的时刻为 T2，那么服务迁移的时间为(T2-T1)。一般服务迁移的时间(近似等于) $\approx$ 检测到服务在 node1 上出现故障的时间+在 node1 上停止服务的时间+在 node2 上启动服务的时间。服务停止和启动的时间由具体的应用来决定，而检测到服务出现错误的时间可以通过对 Cluster 的配置来调整。
- 服务检测到错误的时间：（服务检测间隔+服务检测所需时间）\*连续检测到错误的次数。如果检测到服务出现错误，那么 Cluster 会继续进行检测，直到对服务的检测错误次数达到指定的次数，Cluster 认为服务确实出错，才开始进行服务的迁移。
- 裂脑(Split-Brain)：裂脑指的是集群中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。为了降低裂脑现象发生的可能性，我们建议用户至少配置两条心跳链路。
- 看门狗(Watchdog)：Linux 看门狗是用于系统“挂死”(hanged up)或者崩溃时使服务器自动重启的技术，Linux 支持两种类型的看门狗：自带的 softdog 模块（俗称“软狗”）和硬件看门狗（俗称“硬狗”）。
- 内核恐慌(Kernel panic)功能：Linux 内核恐慌功能是集群重启服务器的一种技术，用于集群异常或者服务停止失败时自动重启服务器来尝试恢复集群/服务。
- LVM 锁：LVM 是逻辑盘管理(LogicalVolumeManager)的简称，它是 Linux 环境下对磁盘分区进行管理的一种机制,它是建立在硬盘和分区之上的一个逻辑层，用来提高磁盘分区管理的灵活性，LVM 锁可以确保 LVM 设备仅能在一个节点上被激活和挂载。
- 仲裁盘(lockdisk)，是用于在心跳全部中断时（裂脑）保证数据安全的一种解决方案。
- 服务依赖：指支持不同服务之间配置依赖关系，解决服务之间的启动、停止和倒换顺序问题；比如配置服务 A 依赖于服务 B，则服务 A 会在服务 B 启动后才启动，停止顺序会相反。**使用此功能时建议把仲裁盘也配置上。**
- 资源依赖：指支持服务的资源之间的配置依赖关系，比如应用资源 A 依赖于应用资源 B，则资源 B 先启动，然后资源 A 再启动。
- snmp 告警：当集群或服务（包含相关资源）出现异常时向网络管理端自动发送警告信息的功能。
- 持久性资源自动恢复：指持久性资源（网卡、存储等物理资源）发生异常且恢复后，服务能够自动启动并恢复运行。

- **数据裂脑**：两个节点的镜像分区数据不一致，每个节点都有对方节点没有的数据，此时需要进行数据裂脑修复，自主选择同步源节点并恢复数据一致性。
- **镜像链路（mirror link）**：节点间链路，用于镜像分区的数据同步。
- **FENCE 技术**：其核心在于解决高可用集群在出现极端问题情况下的运行保障问题，在高可用集群的运行过程中，有时候会检测到某个节点功能不正常，比如在两台高可用服务器间的心跳线突然出现故障，这时一般高可用集群技术将由于连接故障而导致系统错判服务器宕机从而导致资源的抢夺，为解决这一问题就必须通过集群主动判断及检测发现问题并将其从集群中删除以保证集群的稳定运行，Fence 技术的应用可以有效的实现这一功能。Newstart Cluster FENCE 技术目前包括 IPMI 和 SCSI3 PR 。
- **IPMI**：智能平台管理接口 (IPMI) 是一种开放标准的硬件管理接口规格，定义了嵌入式管理子系统进行通信的特定方法。IPMI 信息通过基板管理控制器 (BMC)（位于 IPMI 规格的硬件组件上）进行交流。高可用集群使用标准的 IPMI 管理接口，对服务器电源进行管理，能有效的解决系统假死问题。对故障节点进行隔离。
- **SCSI3 PR**：SCSI3 Persistent Reservation。使用 PR key 对共享磁阵进行加锁，通常一台主机会有唯一的 PR key，只有注册了 PR key 的主机才有对共享磁阵的读写权限。在高可用集群裂脑的情况下，为了防止对磁阵资源的争夺造成数据双挂，节点启动业务前可以删除对方节点的 PR key，这样就保证了对方节点与存储的连接中断，保障了用户的数据安全性。
- **Data-security**：数据安全最大化，数据同步复制过程中的处理策略,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环境。
- **Service-availability**：服务可用性最大化，数据同步复制过程中的处理策略，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环境。



## 2.3 硬件(虚拟硬件)需求列表

硬件（虚拟硬件）需求列表（节点硬件配置相同）		
服务器节点	数量	用途
硬盘	$\geq 1$	可用于： 1.安装 Linux 操作系统 2.作为镜像盘的底层存储
网卡	$\geq 2$	1. 业务链路（千兆以上网卡） 2.镜像链路（千兆以上网卡） 3.心跳链路（两条、可复用其它链路）
交换机	$\geq 1$	用于连接网卡
磁盘阵列	1	存放共享数据(选配)

## 2.4 操作系统支持列表

已确认支持的操作系统（如需支持其他Linux OS发行版/内核版本，请与我们联系）：


支持操作系统：

NewStart Carrier Grade Server Linux(简称CGSL)	CGSL V4.x/V5.x/V6.x
SUSE Linux Enterprise Server (简称SLES)	SLES 11（sp1/sp3）/12
Red Hat Enterprise Linux(简称RHEL)	RHEL 6.x/7.x/8.x
Centos	Centos 6.x/7.x/8.x

支持OS内核：

CGSLv4.x/Redhat6.x/Centos6.x	2.6.32-220.el6.x86_64
	2.6.32-358.el6.x86_64
	2.6.32-431.el6.x86_64
	2.6.32-504.el6.x86_64
	2.6.32-573.el6.x86_64
	2.6.32-642.el6.x86_64
	2.6.32-696.el6.x86_64
	2.6.32-754.el6.x86_64
CGSLv5.x/Redhat7.x/Centos7.x	3.10.0-123.el7.x86_64
	3.10.0-229.el7.x86_64
	3.10.0-327.el7.x86_64
	3.10.0-514.el7.x86_64
	3.10.0-693.el7.x86_64
	3.10.0-862.el7.x86_64
	3.10.0-957.el7.x86_64

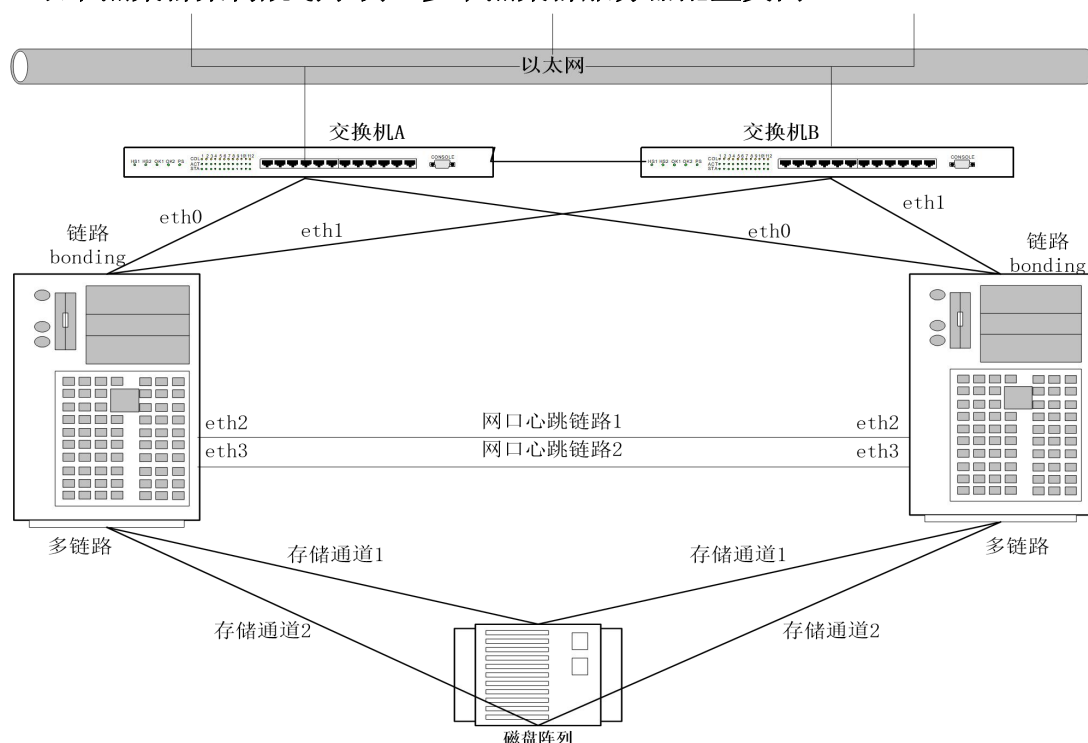
	3.10.0-1062.el7.x86_64
	3.10.0-1127.el7.x86_64
	3.10.0-1160.el7.x86_64
Sles11 sp1/sp3	3.0.101-0.47.105-default
	3.0.101-63-default
	3.0.76-0.11-default
	3.12.49-11-default
Sles12.x	4.4.138-94.39-default
	4.4.143-94.47-default
	4.4.73-5-default
	4.12.14-120-default
CGSLv6.x/Redhat8.x/Centos8.x	4.18.0-80.el8.x86_64
	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

 说明：小版本类型不同同样支持，如内核是2.6.32-220.254.el6.x86\_64，只要内核前四位是2.6.32-220，无论后面是什么小版本型号，Newstart Cluster软件都支持该内核版本，其它内核版本类似。

**Newstart Cluster**作为第三方商业软件，不依赖特定厂商的产品，支持多种 **Linux** 平台、数据库系统和其他软件，有疑问请联系我们。

## 2.5 架构规划

- 双节点集群架构规划示例（多节点集群服务器配置类同）：



双节点集群系统拓扑图（标准配置）

如上图所示，双节点集群中两台服务器A和B，假设每台服务器有四张网卡eth0-eth3。规划如下：

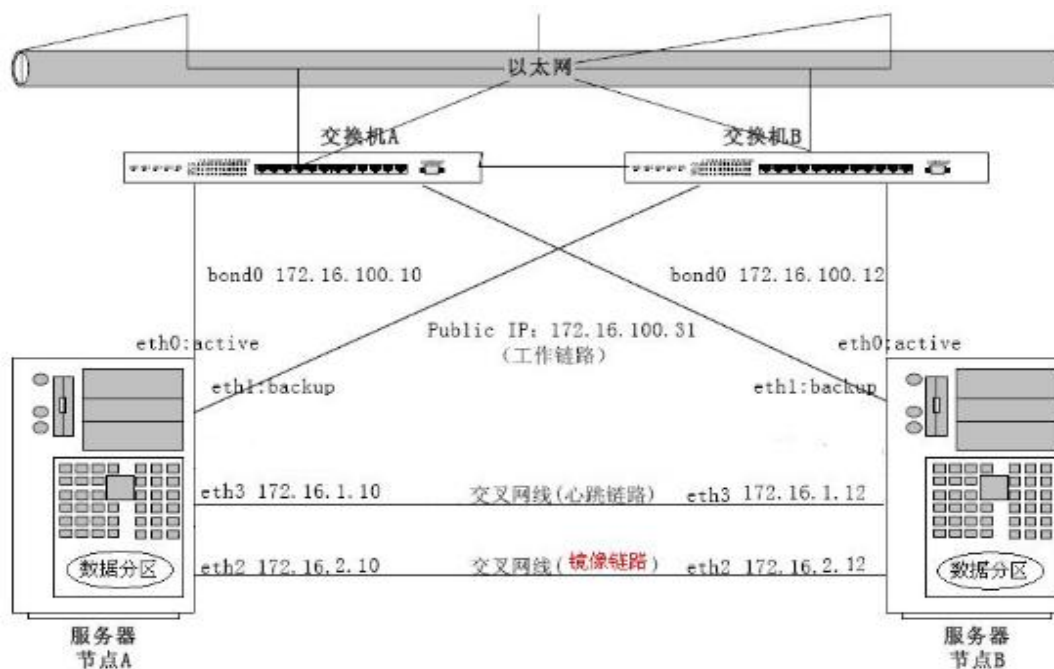
- 1、eth0和eth1设置bonding，作为业务通信链路，采用热备模式，其中eth0作为Active，连接到交换机A，eth1作为Backup，连接到交换机B。通过在Cluster中设置业务对外的业务IP地址。
- 2、eth2/eth3和另一节点对应的网卡eth2/eth3通过交叉网线直连，这些连接将作为心跳链路，负责传输心跳信号，同样支持心跳链路都走交换机。
- 3、划分vlan，规划IP网段：业务链路bond0为172.16.100.0/24网段，心跳链路eth3为172.16.1.0/24网段，eth2为172.16.2.0/24网段；这样可以防止交换机网络异常（比如网络风暴）干扰心跳链路。



注意：

- 1、业务链路、心跳链路分属不同vlan非常重要，调试集群时方便定位链路故障，运行阶段避免链路互相干扰。
- 2、如果是多节点集群，心跳链路需走交换机。

● 磁盘镜像集群架构规划示例：



磁盘镜像集群系统拓扑图（标准配置）

如上图所示，有两台节点服务器A和B搭建磁盘镜像集群，假设每台服务器有四张网卡eth0-eth3。规划如下：

- 1、eth0和eth1设置bonding，作为工作链路，采用热备模式，其中eth0作为Active，连接到交换机A，eth1作为Backup，连接到交换机B。通过在Cluster中设置业务对外的业务IP地址。
- 2、eth3和另一节点对应的网卡eth3通过交叉网线相连，负责传输心跳信号，心跳链路也支持走交换机，一般推荐两条或以上。
- 3、镜像链路用于在主、备服务器之间同步数据，镜像链路支持兼做心跳链路
- 4、划分vlan，规划IP网段：工作链路bond0为172.16.100.0/24网段，心跳链路eth3为172.16.1.0/24网段，镜像链路eth2为172.16.2.0/24网段；这样可以防止交换机网络异常（比如网络风暴）干扰心跳链路和镜像链路。

**注意：**

- 1、工作链路、心跳链路和镜像链路分属不同vlan非常重要，调试集群时方便定位链路故障，运行阶段避免链路互相干扰。
- 2、如果服务器有足够多网卡，建议镜像链路配置成bonding。

## 2.6 配置环境

### 2.6.1 仲裁盘配置（仅适用于已配置共享磁盘阵列的环境）

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。

此解决方案的优点：对于配置共享磁阵的集群不需要增加额外硬件，能大大增强系统的数据安全。

划分分区请在操作系统安装完成后手动进行分区，具体操作参考操作系统手册。



有关数据安全的注意：

- 1、服务器禁止ping IP时必须配置仲裁盘；
  - 2、如果服务器与磁盘阵列的连接方式是iscsi，不建议配置此选项；
  - 3、磁盘阵列无法划分独立的分区，不能配置此选项；
- 

### 2.6.2 划分独立磁盘分区（适用于使用本地分区进行数据同步的环境）

块设备通过网络镜像，保证了两节点数据的一致性。因此用户数据不能存放在系统分区，必须存放在独立的磁盘分区。所以首先要对磁盘进行分区，用于准备镜像的底层设备。

- 1、对磁盘进行分区

使用 fdisk 命令对磁盘进行分区，具体可查询 fdisk 使用方法，分区完成后，使用 fdisk -l 查看磁盘分区。

```
linux175:~ # fdisk -l

Disk /dev/sda: 80.0 GB, 80000000000 bytes
255 heads, 63 sectors/track, 9726 cylinders
units = cylinders of 16065 * 512 = 8225280 bytes

   Device Boot      Start         End      Blocks    Id  System
/dev/sda1             1           261      2096451    82  Linux swap / solaris
/dev/sda2            262          1566     10482412+    83  Linux
/dev/sda3            1567          2871     10482412+    83  Linux
/dev/sda4             *      2872          9726     55062787+    f  W95 Ext'd (LBA)
/dev/sda5            2872          3915      8385898+    83  Linux
/dev/sda6            3916          4889      7823623+    83  Linux
/dev/sda7            4890          5863      7823623+    83  Linux
/dev/sda8            5864          6837      7823623+    83  Linux
/dev/sda9            6838          6960       987966    83  Linux
/dev/sda10           6961          7083       987966    8e  Linux LVM
```

fdisk 分区界面

2、分区完成后，重启操作系统。

3、按照上面的步骤，对另外一台服务器进行同样的操作。

可以用分区后的普通块设备作为镜像底层设备，也可以划分逻辑卷，RAID 等作为镜像底层设备。

如果使用逻辑卷作为镜像底层设备，支持数据可用性保护功能。即备份节点数据同步前自动创建快照，同步成功删除快照。如果同步过程中主节点宕机，备份节点能把数据回滚到同步前的状态，再启动业务。保障了用户数据的连续性和可用性。请在卷组中保留一定的空间用来创建快照，空闲空间的大小安装如下公式近似计算：预计最长的故障恢复时间（小时）×每小时可能产生的数据量+镜像设备总大小×5%。



注意：

- 1、如无特殊情况请保证两台服务器的磁盘分区结果相同，用途也相同；
  - 2、如果需要配置两台服务器的镜像底层设备数据分区大小不一致，请差别不要太大。实际可用大小以最小分区的容量为准；
  - 3、镜像底层设备分区要足够大，避免出现短期内磁盘分区耗尽的情况；
  - 4、建议划分分区时保留部份未分配空间。
- 

### 2.6.3 启用 LVM 锁功能（选配）

当服务器的共享存储分区采用逻辑卷方式进行管理时，Cluster 支持启动 lvm 锁的方式来确保 LVM 设备仅能在一个节点上被激活和挂载，可有效防止业务在备机被人为拉起，降低人为因素带来的风险，启用方式如下：

- 确保环境中已正确配置 LVM 设备，命令 `lvscan` 能正确显示设备信息

```
[root@Node_A ~]# lvscan
inactive          '/dev/vg01/lv01' [200.00 MB] inherit
inactive          '/dev/vg01/lv02' [400.00 MB] inherit
inactive          '/dev/vg01/lv03' [800.00 MB] inherit
```

- 所有节点修改 `/etc/lvm/lvm.conf` 文件，设置 `locking_type` 为 2 生效；禁用 LVM 锁就把 `locking_type` 改为 1 即可。

## 2.6.4 其它系统设置及通讯端口

(1) 软件安装需要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发行版
#vi /etc/sysconfig/network  # CGSL、红帽等发行版
#hostnamectl set-hostname xxx #CGSLv5/Redhat7/Centos7/Sles12等
```

(3) 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	Web管理，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，禁止外部访问
tcp, 10018	与告警服务相关，仅限本地进程通讯，可加入防火墙，禁止外部访问
tcp, 7788~7795	镜像数据同步缺省同步端口范围
udp, 161	snmp告警服务查询端口（支持自定义），选配
udp, 162	snmp告警服务发送端口（支持自定义），选配

## 2.7 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C-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：

- web应用相关可以参考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

### 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/nsha-xxxx.iso /mnt
```

执行安装脚本开始安装（下面以双节点集群为例，其它集群安装过程类似）：

```
# /mnt/install

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

Please input the product SN: 65L948-Q93DN1-32CR1M-10035A 2
You's SN is NewStart Mirror HA product SN!
Checking cluster status ... NOT running
Installing, please wait a few minutes...
Installing the /mnt/packages/x86_64/redhat6/newstartha-5.0.10-20220801.x86_64.rpm ...
Preparing... ##### [100%]
  1:newstartha ##### [100%]
Start tserver OK.

[OK]

License files backedup in /etc/ha.d/lic/.
The cluster configuration files backedup in /etc/ha.d/.

Make /etc/ha.d/lic/newstartha.key succeeded.
Preparing... ##### [100%]
  1:drbd-utils ##### [ 33%]
  2:drbd-proxy ##### [ 67%]
  3:drbd-km-2.6.32_642.el6.##### [100%]

[OK]

Do you want to install cluster Web-based Administrative Tool [y,n](y)3
Install webadmin...
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 start web management tool!  
The component(s) is installed completely.



说明:

1. 技术支持邮箱，疑问或问题都可以发邮件到此邮箱。
2. 输入产品许可号,此处输入的是双节点集群试用序列号
3. 询问是否需要安装 web 管理工具，该处按需求选择，若是不需要 web 管理工具，填写“n”。



注意:

- 1、iso文件上传服务器时必须二进制(bin)方式上传。
- 2、所有服务器都要进行集群软件单独安装。
- 3、安装过程中会写目录/etc/init.d，/usr/bin，所以安装前不能加固以上目录。

## 3.2 产品注册

1. **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。若是前期使用试用文件，后续需要替换为商用文件，操作步骤参考[附录 E-1.如何将试用文件转换为商用文件](#)。
2. **申请 license 文件**：web 在线申请；参考“[附录 D—Newstart Cluster 在线注册系统使用说明](#)”。
3. **验证授权有效性**：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，验证有效会显示正确序列号，无效提示“License files invalid”。



提示:

- 1、如需要申请试用license，只能通过人工方式申请（请联系我们），另外服务器时间与北京时间不能相差24小时，否则license可能显示无效。集群不同产品有相应的试用SN号，可复用，按需选用：  
双节点集群试用 SN 号： ECZF6E-MB1EY8-20BN0D-50075E  
多节点集群试用SN号： 3DZ350-HE1DU2-20BT0C-50075E  
磁盘镜像集群试用 SN 号： 65L948-Q93DN1-32CR1M-10035A

### 3.3 产品卸载

挂载光盘或都 iso 文件到/mnt 子目录:

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

或者

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

执行卸载脚本

```
# /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-5.0.10-20220801.x86_64, Are you sure? [y/n](n)y
Checking cluster status ...                NOT running

uninstalling the newstartha-5.0.10-20220801.x86_64 ...                [OK]

Uninstall DRBD...                [OK]

Web Administrative Tool is running, stop it.                [OK]

Uninstall webadmin...                [OK]

[root@cgs1v405f15-1 home]#
```

## 第 4 章 通过 cli 进行集群配置管理

NewStart Cluster 的管理配置工具 cli 是命令行界面，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 集群配置规划

为保证配置的顺利进行，用户应事先规划好以下各项：

1. 配置每台服务器的主机名，主机名不要相同。
2. 配置每条链路的网卡接口名称，集群中各节点的网卡名称要求相同，同时要求配置好网卡的固定 IP 地址，各个网卡 IP 所在网段不要相同。
3. 确定业务的浮动 IP 地址。
4. 确认存储已规划好（选配），如：
  - 划分仲裁分区，详见 [2.6.1 划分仲裁分区](#)
  - 划分数据分区，建好节点的 mount 目录

## 4.2 初始化配置

集群的配置初始化一般包含集群初始化和服务初始化，如果存储使用的是镜像存储，需要在服务初始前进行镜像盘初始化。

### 4.2.1 双节点集群初始化配置

集群安装完成并申请license（确认每台服务器都已授权）后，命令行下键入cli并回车，出现以下界面：

```
# cli
NewStart High Available System. Version 5.0.10, release 20220808.
Copyright 2008-2022 Guangdong NewStart Technology Service Ltd.

You are using a trial license, and 81 days remain. Please register.1
Product: Type = Two-node cluster, SN = ECZF6E-MB1EY8-20BN0D-50075E 2

Notice: <b,c,q> -- <back, cancel, quit>    <*> -- Option which you must configure

Can not find cluster configuration file.3
Please run cluster-init to initialize Cluster Configuration.
Can not find services configuration file.
Please run service-init to initialize Services Configuration.
cli:~>
```



说明：

1. 授权使用的是试用版本许可，还可使用 81 天，正式版本没有此行提示。
2. 产品类型及许可号（SN 号），以颜色字体显示。
3. 当前环境中的集群仍未配置完成。

初始化命令摘要：

序号	命令	功能描述
1	<a href="#">cluster-init</a>	集群的初始化配置，格式： <code>cluster-init</code> 。
2	<a href="#">service-init</a>	服务的初始化配置，格式： <code>service-init</code> 。

## cluster-init

该命令用于集群的初始化配置，初始化配置完成后，用户不应该再去执行该命令，否则原有的配置将丢失。

该指令运行示例如下，分别有如下配置项及注释：(直接回车为选择默认选项，即选择（）里的内容，[]里面值表示可选选项，该配置文件在 Linux 上为 `/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
- fence configuration
- Lock disk configuration

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

```
Input the 1th node name and IP separated by spaces: [q] node1 192.168.11.3 2
```

```
Input the 2th node name and IP separated by spaces: [q] node2 192.168.11.4
```

```
How long between heartbeats(in seconds): [q] (1) 3
```

```
How long to declare heartbeat has broken(in seconds): [q] (60) 4
```

```
Do you want to enable watchdog ? [y/n] (n) 5
```

```
Add heartbeat links:
```

- 0) Multicast heartheat
- 1) Unicast heartheat
- 2) Serial heartheat

```
Select a type of heartheat: [0-2,q] 1 6
```

```

IP type:
    0) IPv4
    1) IPv6
Select a type: [0-1,q] (0) 7
Detected NICs on cluster: eth0 eth1 eth2 8
Input a network device for unicast heartbeat: [q] (eth0) eth1 9
Input node1 IPv4 address [q] (192.168.11.3)
Input node2 IPv4 address [q] (192.168.11.4)
Add another ucast heartbeat link? [y/n] (y)
Input a network device for unicast heartbeat: [c,q] (eth0) eth2
Input node1 IPv4 address [q] (192.168.12.3)
Input node2 IPv4 address [q] (192.168.12.4)
Add another ucast heartbeat link? [y/n] (y) n 10
Do you want to add other heartbeat link? [y/n] (n) 11
Do you want to enable worklink_hb ? [y/n] (y) 12
Do you want to enable kernel_panic ? [y/n] (n) 13
Do you want to add a lock disk(recommend) ? [y/n] (n) y 14
Please input the partition name(/dev/sdc): [q] /dev/sdb1 15
Warning:All data in /dev/sdb1 will be destroyed, sure to format it? [y/n] (n) y 15
Cluster information verification:

    clustername: cluster
    node: node1 192.168.11.3
    node: node2 192.168.11.4
    logfile: /var/log/ha-log
    keepalive: 1
    deadtime: 60
    watchdog: off
    ucast: eth1 192.168.11.3 192.168.11.4
    ucast: eth2 192.168.12.3 192.168.12.4
    worklink_hb: on 226.0.121.132
    lockdisk: /dev/sdb1
    kernel_panic: off
Is this information correct? [y/n,q] (y) 17
Do you want to add fence ? [y/n] (n) 18
Please run service-init to initialize you services.
cli:~>

```

双节点集群初始化示意图



说明:

1. 自定义集群名称;
2. 输入节点主机名和 IP 地址(支持配置 ipv4 或 ipv6 地址, 建议配置心跳 IP); 节点名称必须与 `uname -n` (即机器真实名称) 相同, IP 地址必须是网卡的固定 IP 地址 (比如心跳网卡 **eth2/eth3** 的 IP), 建议每个节点配置多个 IP 地址, 以空格分隔(此 IP 用于双机配置文件同步或指令传输, 一般为心跳网卡 IP);
3. 心跳间隔: 一般选择默认值 1 秒一次心跳, 注意直接回车即选择小括号里面的缺省值;

4. 判断对方服务器死的时间间隔（也称 **deadtime**），也就是多长时间没有收到对方服务器任何心跳包就认为对方已经死了，默认 60 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. 是否在集群中使用 Watchdog 看门狗技术，详细说明见后面；
6. 增加心跳链路：双节点集群支持组播、单播和串口心跳链路，推荐配置单播心跳，按规划这里将添加两条网口心跳，另注意集群中所有节点的网卡名需保持一致，否则在启动集群时会有问题；  
**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. 选择单播心跳的 ip 类型（支持 ipv4 与 ipv6 地址），这里缺省选 ipv4
8. 获取当前可用的网卡设备，这里是 eth0、eth1、eth2
9. 选择网卡设置进行心跳链路配置，这里规划两条心跳 eth1 和 eth2；
10. 已配置了二条心跳链路，这里选 n
11. 允许同时配置不同类型的心跳链路类型，一般心跳链路总数达到两条或以上即可，这里选 n；
12. 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；建议启用；
13. 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，调试完成后再启用此功能，用命令 cluster-modify 修改；**
14. 是否增加仲裁盘，推荐使用默认值；
15. 输入仲裁盘的分区名称，仲裁分区在集群节点的名称必须相同；
16. 仲裁盘分区的数据会被格式化，请确认；
17. 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”。
18. 是否配置 fence 功能，目前双节点集群支持 ipmi 和 SCSI3 方式，其配置方法后面会涉及。



#### 注意

- 1、心跳网卡不推荐使用 bonding，建议拆成 2 条心跳链路使用。
  - 2、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。
  - 3、上述配置参数不合理地方可以使用指令 cluster-modify 修改。
-

## 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] oracle 1
Is it enabled? [y/n,q] (y) 2
Please select nodes to run the service: 3
    0) node1
    1) node2
Select one or more nodes separated by spaces: [0-1,q] 0 1 4
Do you want to configure preferred node? [y/n,q] (n) y 5
Please select preferred node:
    0) node1
    1) node2
Select a node: [0-1,q] 0
Input start time out: [q] (60) 6
Input stop time out: [q] (120) 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: oracle
    multinstance: no
    enabled: yes
    starttimeout: 60
    stoptimeout: 120
    checkinterval: 30
```



```
checktimeout: 60
maxerrcount: 1
maxlockcount: 0
restartaftercheckfail: no
startanyway: no
odelist: node1 node2
autostartlist: node1 node2
preferrednode: node1
```

Is this information correct? [y/n,q] (y) <sup>13</sup>

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；
3. 列出可运行该服务的节点；
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,这里选 node1 和 node2 两节点；
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，只适用于共享/多节点版本。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:<sup>14</sup>

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] 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 22
Discovered NICs on node2: eth3 eth2 bond0 23
Input a network device: [q] (bond0) 24
Are you using the same configuration in the remaining nodes? [y/n] (y)25
Do you want to add network host IP? [y/n] (y) 26
Input IP address: [c,q] 192.168.1.194 27
Do you want add another network host IP? [y/n] (y)
Input IP address: [c,q] 192.168.1.249
Do you want add another network host IP? [y/n] (y)
Input IP address: [c,q] 192.168.1.250
Do you want add another network host IP? [y/n] (y) n 28
NIC resource information verification:29

    name: oracle_nic_0
    critical: yes
    enable: yes
    networkhostip: 192.168.1.194 192.168.1.249 192.168.1.250
    type: single
    nodename: node2    device: bond0
    nodename: node1    device: bond0

Is this information correct? [y/n,q] (y) 30
Add another NIC resource? [y/n] (n) 31
Do you want to add other resource? [y/n] (y)32

```

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；
22. 这里只有一条业务链路 bond0,选择 single；
23. 列出可供选择的网卡设备名称；
24. 选择设备 bond0；
25. 是否把配置同步到其它节点，这里选 y；
26. 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，建议配置三个或三个以上此类 IP；  
**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**
27. 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；
28. 是否再增加网络中存在的主机 IP,这里已配置三个，选择 no；
29. 列出网卡资源配置信息列表；
30. 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；
31. 是否再增加一个网卡资源，可以支持配置多这个，这里选 no；
32. 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) RawDisk
- 4) DiskMount
- b) Back to previous menu

Please select a type of resource: [0-4,b,q] 1 <sup>33</sup>

To configure IP resource the following information is required: <sup>34</sup>

- 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) 35
Input IP name: [q] (oracle_ip_0) 36
Is resource critical? [y/n,q] (y) 37
Is resource enable? [y/n,q] (y) 38
Please choose an IP type:
    0) IPv4
    1) IPv6
Select a type: [0-1,q] (0) 0 39
Discovered NICs on node2: eth3 eth2 bond0 40
Input a network device: [q] (bond0) 41
Input IP address: [q] 192.168.1.13 42
Input netmask: [q] (255.255.255.0) 43
Are you using the same configuration in the remaining nodes? [y/n] (y) 44
IP resource information verification:

    name: oracle_ip_0
    critical: yes
    enable: yes
    nodename: node2
    PubNICResName: bond0
    ipaddr: 192.168.1.13
    netmask: 255.255.255.0
    nodename: node1
    PubNICResName: bond0
    ipaddr: 192.168.1.13
    netmask: 255.255.255.0

Is this information correct? [y/n,q] (y) 45
Add another IP resource? [y/n] (n) 46
Do you want to add other resource? [y/n] (y) 47

```

服务初始化示意图3

33. 选择增加 IP 资源，主要是配置业务 IP；
34. 列出 IP 资源需要配置的属性/信息；
35. 询问是否开始配置，选择 yes 往下执行；
36. IP 资源名称，内部识别，直接回车即可；
37. 是否设定为关键资源，直接回车选择 y；
38. 是否被集群管理，直接回车选择 y；
39. 选择配置的浮动 IP 的类型，支持 IPv4 和 IPv6，这里选择 IPv4；
40. 列出可供选择的网卡设备；
41. 选择业务链路网卡，这里选择 bond0；
42. 输入业务 IP，前面规划为 192.168.1.13；
43. 配置业务 IP 的子网掩码；
44. 同步配置到其它节点；
45. 确认配置的 IP 资源信息是否正确，正确选择 y；
46. 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；
47. 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) RawDisk
    4) DiskMount
    b) Back to previous menu
Please select a type of resource: [0-4,b,q] 2 48

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) 49
Input application name: [q] (oracle_app_0) 50
Input script of application: [q] (/etc/ha.d/resource.d/oracle) /home/script/oracle.ps 51
Input the count of fault on check time out: [q] (0) 52
Is resource critical? [y/n,q] (y) 53
Is resource enable? [y/n,q] (y) 54
Application resource information verification:

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

Is this information correct? [y/n,q] (y) 55
Add another application resource? [y/n] (n) 56
Do you want to add other resource? [y/n] (y) 57

```

服务初始化示意图4

- 48. 选择增加应用资源，主要是配置业务应用，这里是 oracle；
- 49. 询问是否开始配置，选择 y 往下执行；
- 50. 应用资源名称，内部识别，直接回车即可；
- 51. 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；
- 52. 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 选项，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
- 53. 是否设定为关键资源，直接回车选择 y；

54. 是否被集群管理，直接回车选择 y ；
55. 确认配置的应用资源是否正确，正确选择 y；
56. 是否再增加一个应用资源，可以支持配置多这个；
57. 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) RawDisk
- 4) DiskMount
- b) Back to previous menu

Please select a type of resource: [0-4,b,q] **4** <sup>58</sup>

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) <sup>59</sup>

Input diskmount name: [q] (oracle\_dm\_0) <sup>60</sup>

Is resource critical? [y/n,q] (y) <sup>61</sup>

Is resource enable? [y/n,q] (y) <sup>62</sup>

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

please choose a disk type: [0-2,q] **0** <sup>63</sup>

Input block disk device: [q] (/dev/hda1) **/dev/sdb2** <sup>64</sup>

Is a network storage? [y/n,q] **n** <sup>65</sup>

Input mountpoint: [q] **/db** <sup>66</sup>

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) 67
Do you want to stop service when the disk is readonly? [y/n,q] (y) 68
When the file system mount failure, use fsck to repair it? [y/n,q] (n) 69
Diskmount resource information verification:

    name: oracle_dm_0
    critical: yes
    enable: yes
    type: disk
    device: /dev/sdb2
    mountpoint: /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) 70
Add another diskmount resource? [y/n] (n) 71
Do you want to add other resource? [y/n] (y) n 72
Do you want to add another service? [y/n] (n) 73
Please run cluster-start to start the HA system,
or run cluster-restart to restart the HA system.
cli:~>

```

服务初始化示意图5

58. 选择增加 DiskMount 存储资源，这里规划是/dev/sdb2 设备；
59. 询问是否开始配置，选择 y 往下执行；
60. 共享存储资源名称，内部识别，直接回车即可；
61. 是否设定为关键资源，直接回车选择 y；
62. 是否被集群管理，直接回车选择 y ；
63. 选择共享存储类型，这里是普通块设备,选择 0；
64. 输入使用的块设备名称： /dev/sdb2；
65. 询问是否使用网络存储（iscsi 磁阵）；如是，前面 IP 资源配置的网络存在的 IP 一定要配 iscsi 服务的 IP，这样才可以对此类磁阵进行有效的检测。
66. 输入挂载点相关信息：
  - 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；
67. 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。
  68. 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；
  69. 当设备挂载失败时是否尝试用 fsck 去修复它；
  70. 确认配置的 DiskMount 资源是否正确；
  71. 是否再增加一个 DiskMount 资源，可以支持配置多这个；
  72. 询问是否增加服务的其它资源，这里选 n；
  73. 询问是否增加另一个服务。

有些场景中双节点集群可能需要用到 NFS 服务器和 Fence 功能，这里也简单说一下配置方法。

#### NFS设备配置如下：

如果存储设备选择nfs设备，需要输入nfs服务器的IP地址和导出目录(nfs path)，其他配置项与普通的块设备一样。

```
Input nfs server ip address: 192.168.1.156
Input nfs path: /public/nfs
Input mount point:
```

#### Fence功能配置如下：

目前双节点集群支持 ipmi 及 scsi3 方式的 fence 功能，仅支持双节点集群。

```
cli:~>fence-add
fence types:
    0) ipmilan
    1) scsi
    2) other
    b) Back to previous menu
Please select a type of fence: [0-2,b,q] 0
The following is configured on the node kylin10-3
Input fence device IP address: [q] 192.168.33.25
Input login name: [q] admin
Input password: [q] abc123
...
Are you sure to use these configuration [y/n] (n) y
fence ipmi information verification:

    nodename: kylin10-3
    ipaddr: 192.168.33.25
    login: admin
    passwd: abc123

Is this information correct? [y/n,q] (y)
Are you using the same configuration in the remaining nodes? [y/n] (y) n
```



```
The following is configured on the node kylin10-4
Input fence device IP address: [q] 192.168.33.26
Input login name: [q] admin
Input password: [q] abc123
```

...

```
Are you sure to use these configuration [y/n] (n) y
fence ipmi information verification:
```

```
nodename: kylin10-4
ipaddr: 192.168.33.26
login: admin
passwd: abc123
```

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

```
Please restart the HA system to take your modification into effect.
```

## 4.2.2 多节点集群初始化配置

初始化命令摘要：

序号	命令	功能描述
1	<a href="#">cluster-init</a>	集群的初始化配置，格式： <code>cluster-init</code> 。
2	<a href="#">service-init</a>	服务的初始化配置，格式： <code>service-init</code> 。

多节点初始化配置同样是分集群初始化(`cluster-init`)与服务初始化(`service-init`)，初始化操作过程与双节点集群基本相同，主要区别在于集群初始化时支持增加两个以上节点，以及服务初始化时进行实例配置，主要过程参考如下（下面以 3+1 实例配置为例）：

```
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
- Lock disk configuration

WARNING: cluster-init will break the origin configuration file of the cluster.

Are you sure to re-create the configuration file? [y/n] (y)

Please input cluster name: [q] cluster

Input the 1th node name and IP separated by spaces: [q] linux47 192.168.8.47

Input the 2th node name and IP separated by spaces: [q] linux48 192.168.8.48

**Do you want to continue to add node? [y/n] (y) <sup>1</sup>**

Input the 3th node name and IP separated by spaces: [c,q] linux49 192.168.8.49

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

Input the 4th node name and IP separated by spaces: [c,q] linux50 192.168.8.50

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

How long between heartbeats(in seconds): [q] (1)

How long to declare heartbeat has broken(in seconds): [q] (60)

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

Add heartbeat links:

Detected NICs on linux47: eth0 eth1 eth2

Input a network device for multicast heartbeat: [q] (eth0) **eth2 <sup>2</sup>**

Add another multicast heartbeat link? [y/n] (y) n

Do you want to enable worklink\_hb ? [y/n] (y) n

Do you want to add a lock disk(recommend) ? [y/n] (n)

Do you want to enable kernel\_panic ? [y/n] (n)

Cluster information verification:

```
clustername: oracle
node: linux47 192.168.8.47
node: linux48 192.168.8.48
node: linux49 192.168.8.49
node: linux50 192.168.8.50
logfile: /var/log/ha-log
keepalive: 1
deadtime: 60
watchdog: off
mcast: eth2 226.0.228.131 694 1 0
worklink_hb: off
kernel_panic: off
```

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

Please run service-init to initialize you services.



说明:

1. 选择添加两个以上节点，这里选择 y 进行操作，共添加 4 个节点（3+1）
2. 选择心跳网卡，多节点仅支持组播心跳，与双节点配置的心跳类型有区别

```
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] mult\_service

**Is multi instance service? [y/n,q] (n) y<sup>1</sup>**

Is it enabled? [y/n,q] (y)

...此处打印过程省略<sup>2</sup>

Add instances for the service mult\_service:

Please input the instance ID: [0-99,q] (0)<sup>3</sup>

Input instance name: [q] (mult\_service\_0)

This service 'mult\_service' has the following resources:

- 0) mult\_service\_nic\_0
- 1) mult\_service\_ip\_0
- 2) mult\_service\_ip\_1
- 3) mult\_service\_ip\_2
- 4) mult\_service\_app\_0

Select resources separated by spaces: [0-4,q] 0 1 4<sup>4</sup>

Resources for instance ID 0 information verification:

```
id: 0
name: mult_service_0
member: mult_service_nic_0
member: mult_service_ip_0
member: mult_service_app_0
```

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

Do you want to add another instance? [y/n] (y)

Please input the instance ID: [0-99,q] (1)

This service 'mult\_service' has the following resources:

- 0) mult\_service\_nic\_0
- 1) mult\_service\_ip\_0
- 2) mult\_service\_ip\_1
- 3) mult\_service\_ip\_2
- 4) mult\_service\_app\_0

Select resources separated by spaces: [0-4,c,q] 0 2 4

Resources for instance ID 1 information verification:

```
id: 1
name: mult_service_1
member: mult_service_nic_0
member: mult_service_ip_1
member: mult_service_app_0
```

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

Do you want to add another instance? [y/n] (y) <sup>5</sup>

Please input the instance ID: [0-99,q] (2)

This service 'mult\_service' has the following resources:

```
0) mult_service_nic_0
1) mult_service_ip_0
2) mult_service_ip_1
3) mult_service_ip_2
4) mult_service_app_0
```

Select resources separated by spaces: [0-4,c,q] 0 3 4

Resources for instance ID 2 information verification:

```
id: 1
name: mult_service_1
member: mult_service_nic_0
member: mult_service_ip_2
member: mult_service_app_0
```

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

Do you want to add another instance? [y/n] (y) n

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. 是否配置多实例服务，这里选 y
2. 此次配置过程与前面双节点集群 service-init 类似，按向导操作，这里不重复
3. 添加实例 ID；每个实例分配一个可识别的 ID，目前只支持数字
4. 配置该实例包含的所有资源
5. 添加下一个实例，操作类似，范例共添加 3 个实例，组成 3+1



#### 注意

- 1、多个实例启动机制是集群先分配好所有实例运行的节点，然后同时启动。
- 2、多实例版本同样支持双节点集群的服务配置，运行节点选择其中的两个节点即可，适合服务器资源复用，可达到资源最大化利用目的。

## 4.2.3 磁盘镜像集群初始化配置

初始化命令摘要：

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

### cluster-init

磁盘镜像集群初始化包括集群、镜像资源和服务初始化，并且需要严格遵守该操作顺序,集群初始化（`cluster-init`）过程参考如下：

```
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] cluster
Input the 1th node name and IP separated by spaces: [q] aa40 192.168.8.40
Input the 2th node name and IP separated by spaces: [q] aa41 192.168.8.41
How long between heartbeats(in seconds): [q] (1)
How long to declare heartbeat has broken(in seconds): [q] (60)
Do you want to enable watchdog ? [y/n] (n)
Add heartbeat links:
    0) Multicast heartheat
    1) Unicast heartheat
    2) Serial heartheat
Select a type of heartheat: [0-2,q] 1
IP type:
    0) IPv4
    1) IPv6
```

```
Select a type: [0-1,q] (0)
Detected NICs on linux_61: eth2 bond0 trunk0
Input a network device for multicast heartbeat: [q] (eth2)
Add another multicast heartbeat link? [y/n] (y) n
Do you want to add other heartbeat link? [y/n] (n)
Do you want to enable worklink_hb ? [y/n] (y)
Please configure data and service policies:
    0. the maximization of data security
    1. the maximization of service availability
Please select: [0-1,q] 0 1
Do you want to enable kernel_panic ? [y/n] (n)
Cluster information verification:

    clustername: aa
    node: aa40 192.168.8.40
    node: aa41 192.168.8.41
    logfile: /var/log/ha-log
    keepalive: 1
    deadtime: 60
    watchdog: off
    mcast: eth2 226.0.117.247 694 1 0
    worklink_hb: on 226.0.154.244
    maximization: data-security
    kernel_panic: off

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



#### 说明:

1. 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。

## mirror-init

下面是镜像资源的初始化（mirror-init）过程：

```
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

mirror resource configuration file exists, Do you want to recreate it and re-initialize mirror? [y/n] (y)

Input resource name: [q] (r0) <sup>1</sup>

Input port number: [1-65535,q] (7788) <sup>2</sup>

Input an unused disk on aa41: [q] /dev/sdb <sup>3</sup>

Input mirror IP address for aa41: [q] 192.168.12.41 <sup>4</sup>

Input an unused disk on aa40: [q] (/dev/sdb) <sup>5</sup>

Input mirror IP address for aa40: [q] 192.168.12.40 <sup>6</sup>

DRBD resource information verification:

```
name: r0
device: /dev/drbd0
local node: aa41
    IP address: 192.168.12.41
    disk: /dev/sdb
remote node: aa40
    IP address: 192.168.12.40
    disk: /dev/sdb
```

Is this information correct? [y/n,q] (y) <sup>7</sup>

Do you want to add another mirror resource? [y/n] (n) <sup>8</sup>

All the data of /dev/sdb on node aa41 will be destroyed. Are you sure to continue?  
[y/n] (y) <sup>9</sup>

All the data of /dev/sdb on node aa40 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 starting...

```

start mirror success.
Waiting for r0 synchronization...
r0 Resynced percentage: 100.0 % 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, 3]:1 11
mke2fs 1.39 (29-May-2006)
Filesystem label=
OS type: Linux
Block size=4096 (log=2)
Fragment size=4096 (log=2)
393216 inodes, 786399 blocks
39319 blocks (5.00%) reserved for the super user
First data block=0
Maximum filesystem blocks=805306368
24 block groups
32768 blocks per group, 32768 fragments per group
16384 inodes per group
Superblock backups stored on blocks:
    32768, 98304, 163840, 229376, 294912
Writing inode tables: done
Creating journal (16384 blocks): done
Writing superblocks and filesystem accounting information: done

Format disk /dev/drbd0 to ext3 success
set secondary all successful
Mirror init success! 12

```

磁盘镜像初始化示意图



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、输入节点1的本地磁盘分区名称，注意两个节点的磁盘分区名称要相同；
- 4、输入节点1镜像链路地址
- 5、输入节点2的本地磁盘分区名称
- 6、输入节点2镜像链路地址
- 7、确认配置镜像资源是否正确
- 8、是否增加另一下镜像资源
- 9、准备对配置的分区进行格式化
- 10、 两节点分区进行数据同步，100%表示同步完成
- 11、 对镜像分区进行格式化
- 12、 镜像资源格式化完成。



## 4.3 修改配置

集群运行过程中，可能会遇到现网环境调整情况，此时集群也要做出相应调整，下面列出了常用配置修改的操作指令，按需参考。

命令摘要：

序号	命令	功能描述
1	<a href="#">heartbeat-add</a>	增加心跳链路
2	<a href="#">heartbeat-delete</a>	删除心跳链路
3	<a href="#">cluster-modify</a>	修改集群的初始化配置。
4	<a href="#">lockdisk-adm</a>	仲裁盘管理
5	<a href="#">cluster-show</a>	显示集群的配置信息
6	<a href="#">service-add</a>	向配置文件添加服务
7	<a href="#">service-show</a>	显示指定服务的配置信息
8	<a href="#">service-modify</a>	修改一个服务的配置信息
9	<a href="#">resource-add</a>	增加服务的相关资源
10	<a href="#">resource-delete</a>	删除服务的相关资源
11	<a href="#">resource-modify</a>	修改服务的资源设定
12	<a href="#">service-delete</a>	从配置文件里删除服务
13	<a href="#">resource-depend</a>	设定服务的资源依赖
14	<a href="#">service-depend</a>	设定服务之间的依赖关系
15	<a href="#">check-config</a>	检查服务配置是否正确，配置不正确集群禁止启动
16	<a href="#">mirror-adm</a>	磁盘镜像管理，包括增加与删除
17	<a href="#">fence-add</a>	增加fencing功能
18	<a href="#">fence-delete</a>	删除fencing功能
19	<a href="#">instance-config</a>	增加、删除和显示已配置实例，多节点集群特有
20	<a href="#">ipvs-config</a>	添加、删除、修改、显示负载均衡相关配置
21	<a href="#">modify-config</a>	把旧双机版本的配置文件转换成新版本的配置文件
22	<a href="#">product-switch</a>	版本类型转换操作

### heartbeat-add

增加心跳链路heartbeat-add。该指令运行示例如下：

```
cli:~>heartbeat-add
Add heartbeat links:
    0) Multicast heartheat
    1) Unicast heartheat
    2) Serial heartheat
Select a type of heartheat: [0-2,q] 0
Discovered NICs on node1: eth3 eth2 bond0
Input a network device for multicast heartbeat: [q] (bond0)
Add another multicast heartbeat link? [y/n] (y) n
Do you want to add other heartbeat link? [y/n] (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。该指令运行示例如下：

```
cli:~>heartbeat-delete
    0) mcast: eth3 226.0.91.53 694 1 0
    1) mcast: bond0 226.0.73.179 694 1 0
    2) mcast: eth2 226.0.62.75 694 1 0
    3) serial: /dev/ttyS0
    b) Back to previous menu
Select a heartbeat link channel to delete: [0-3,b,q] 1
Do you want to save the configuration? [y/n] (y)
Please restart the HA system to take your modification into effect.
```

## cluster-modify

修改集群配置命令。该指令运行示例如下：

```
cli:~>cluster-modify
Configuration of HA system:
    0) clustername: cluster
    1) node: node1 192.168.1.3
    2) node: node2 192.168.1.4
    3) logfile: /var/log/ha-log
    4) keepalive: 1
    5) deadtime: 60
    6) watchdog: off
    7) mcast: eth3 226.0.91.53 694 1 0
    8) mcast: eth2 226.0.62.75 694 1 0
    9) serial: /dev/ttyS0
   10) worklink_hb: on 226.0.121.132
   11) lockdisk: /dev/sdb1
   12) kernel_panic: off
    b) Back to previous menu
Select an item to modify: [0-12,b,q]
```

## lockdisk-adm

仲裁分区管理命令。该指令运行示例如下：

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

## cluster-show

显示当前集群的配置情况，确认配置是否正确。

```
cli:~>cluster-show

configuration of the cluster:

    clustername: cluster
    node: node1 192.168.1.3
    node: node2 192.168.1.4
    logfile: /var/log/ha-log
    keepalive: 1
    deadtime: 60
    watchdog: off
    mcast: eth3 226.0.91.53 694 1 0
    mcast: eth2 226.0.62.75 694 1 0
    serial: /dev/ttyS0
    worklink_hb: on 226.0.121.132
    lockdisk: /dev/sdb1
    kernel_panic: off
```

## service-add

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

## service-show

查看服务配置信息命令 service-show。

```
cli:~>service-show
Current service:
    0) oracle
    b) Back to previous menu
Select a(n) service: [0-0,b,q] 0
```

```
service
-----
name: oracle
enabled: yes
starttimeout: 60
stoptimeout: 120
checkinterval: 30
checktimeout: 60
maxerrcount: 1
restartaftercheckfail: no
startanyway: no
nodelist: node1 node2
autostartlist: node1 node2
    preferrednode: node1
... 过程省略
Show other service? [y/n] (n)
```

## service-modify

修改服务配置信息命令 `service-modify`。该指令运行示例如下：

```
cli:~>service-modify
Services of the cluster:
    0) oracle
    b) Back to previous menu
Please select a service to modify: [0-0,b,q] 0
Property information:
    0) enabled: yes
    1) starttimeout: 60
    2) stoptimeout: 120
    3) checkinterval: 30
    4) checktimeout: 60
    5) maxerrcount: 1
    6) restartaftercheckfail: no
    7) startanyway: no
    8) nodelist: node1 node2
    9) autostartlist: node1 node2
    10) preferrednode: node1
    b) Back to previous menu
Please select an item to modify: [0-10,b,q]
```

## resource-add

修改已存在的服务，增加相关资源，如按需增加一个业务 ip，增加磁阵等。该指令运行示例如下：

```
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

修改已存在的服务，删除相关资源。该指令运行示例如下：

```
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。示例如下：

```
cli:~>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**

## service-delete

删除服务命令service-delete。该指令运行示例如下：

cli:~>**service-delete**

- 0) oracle
- b) Back to previous menu

Please select a service to delete: [0-0,b,q]

## resource-depend

设定服务中资源的依赖关系（原则：父资源依赖于子资源）。例如下面设定 oracle 服务的 IP 资源为父资源，网卡资源为子资源，则 IP（父资源）必须在网卡（子资源）启动后才启动。

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)

## service-depend

设定服务之间的依赖关系（原则：父服务依赖于子服务）；此功能建议与仲裁盘配合使用，即把仲裁盘也配置好，仲裁盘配置方法见前面的 [lockdisk-adm](#)。

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) oracle
- 1) monitor

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

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

- 0) forced      表示具有依赖关系的任一服务异常，所有服务倒换。
- 1) fixed      表示父服务异常，不影响子服务，反之所有服务倒换。

Please select a type of dependency: [0-1,q] 1

- 0) local      表示依赖的服务都在本地运行
- 1) remote    表示依赖的服务都有远端运行

```

    2) global  表示依赖的服务允许在本地或远端运行
Please select a kind of relationship: [0-2,q] 2
Do you want to configure another service dependency? [y/n] (n)
Service dependency information verification:

    Parent:monitor          Child:oracle          Relationship:global
Dependency-type:fixed

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

```

## check-config

检查本节点服务配置是否正确,该指令运行如下:

```

cli:~>check-config 检测正常
Configuration file checked ok.

cli:~>check-config 检测错误
/etc/ha.d/haservices.xml:27: parser error : Premature end of data in tag haservices line
2

Can not parse document /etc/ha.d/haservices.xml successfully.
Configuration file checked error.

```

## mirror-adm

镜像资源的增删管理指令，增加操作与[mirror-init](#)过程类似，下面仅展示删除操作过程。

```

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

Warning: mirror resource r0 has been used in service.
Are you sure to delete it? [y/n] (n) y

```



## fence-add

如环境需要增加 fencing 功能，可直接运行 fence-add 按向导操作。

```
cli:~>fence-add
fence types:
    0) ipmilan
    1) scsi
    2) other
    b) Back to previous menu
Please select a type of fence: [0-2,b,q] 0
The following is configured on the node aa40
Input fence device IP address: [q]
```

## fence-delete

如不需配置，直接删掉。

```
cli:~>fence-delete
...
```

## instance-config

用于增加、删除和显示已配置实例，cli中运行instance-config，按向导操作，下面以实例增加为例，其它操作类似：

```
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)

...

## ipvs-config

用于负载均衡配置添加和删除，真实服务器信息的增加和删除，配置修改与查看指令,这里主要看后两者操作。

**cli:~>ipvs-config**

- 0) IPVS add
- 1) IPVS delete
- 2) Real IP add
- 3) Real IP delete
- 4) IPVS modify
- 5) Display IPVS configuration
- b) Back to previous menu

```

Please select an operation: [0-5,b,q] 4
    0) modify virtual IP
    1) modify real IP
    b) Back to previous menu
Please select an operation: [0-1,b,q] 0
virtual IP:
    0) 192.168.8.151
    b) Back to previous menu
Please select a virtual IP to modify: [0-2,b,q] 0
    0) ip: 192.168.8.151
    1) netmask: 255.255.255.0
    3) nic: eth0
    4) service: http
    5) port: 80
    6) protocol: tcp
    7) scheduler: wlc 1
    8) checktype: negotiate 2
    9) nodelist: linux57 linux58
    b) Back to previous menu
Please select an item to modify: [0-8,b,q] 7
Load balancing scheduling algorithm:
    0) rr - Round Robin
    1) wrr - Weighted Round Robin
    2) lc - Least Connections
    3) wlc - Weighted Least Connections
    4) lbic - Locality-Based Least Connections
    5) lbicr - Locality-Based Least Connections with Replication
    6) dh - Destination Hashing
    7) sh - Source Hashing
    8) sed - Shortest Expected Delay
    9) nq - Never Queue
Please select a scheduling algorithm: [0-9,c,q] (3) 03
Modify other property of VIP? [y/n] (n)
Please restart the HA system to take your modification into effect.
cli:~>

```



说明:

1、修改负载均衡调度算法,相关算法说明如下:

- \*) rr - 轮询方式, 把请求按顺序分配给业务服务器
- \*) wrr - 带权重的轮询方式, 权重越大, 分配的请求越多
- \*) lc - 按最少连接数分配请求
- \*) wlc - 带权重的最少连接数分配方式, 权重越大, 分配的请求越多
- \*) lbic - 用于cache服务器集群, 优先分配集群中访问过的服务器, 然后才适用“最少链接”的分配原则
- \*) lbicr - 用于cache服务器集群, 与lbic算法区别在于多了服务器组的维护, 优先分配cache服务器集群中访问过的服务器, 然后才适用“最少链接”原则把服务器加入服务器组

- \*) dh - 根据目标地址从分配的散列表找出对应的服务器进行请求分配
  - \*) sh - 根据请求地址从分配的散列表找出对应的服务器进行请求分配
  - \*) sed - 使用sed算法进行请求分配
  - \*) nq - 分配请求无需队列, 如果服务器连接数为0就直接分配
- 2、修改负载均衡服务器监控真实服务器(真实业务应用所在机器)的方式, 缺省有以下几种, 说明如下:
- \*) negotiate: 连接到真实服务器, 请求返回指定的应答字符串, 根据返回情况判断。
  - \*) connect: 连接到真实服务器, 根据TCO/IP连接情况判断。
  - \*) ping: 通过ping真实服务器上配置的浮动IP来判断
  - \*) A number :尝试指定数量的连接真实服务器测试
  - \*) off: 不监控真实服务器
- 3、修改调度算法为rr(Round Robin),配置修改都需要重启集群才能生效。

```
cli:~>ipvs-config
0) IPVS add
1) IPVS delete
2) Real IP add
3) Real IP delete
4) IPVS modify
5) Display IPVS configuration      ————查看配置
b) Back to previous menu
Please select an operation: [0-5,b,q] 5
IPVS configuration information:
virtual ip: 192.168.8.151
netmask: 255.255.255.0
nic: eth0
service: http
protocol: tcp
port: 80
scheduler: rr
checktype: negotiate
nodelist: linux57 linux58
real ip: 192.168.8.100
workmode: g
weight: 1
real ip: 192.168.8.101
workmode: g
weight: 1
real ip: 192.168.8.102
workmode: g
weight: 1

virtual ip: 192.168.8.251
netmask: 255.255.255.0
nic: eth0
service: none
protocol: tcp
```

```
port: 10024
scheduler: wlc
checktype: negotiate
nodelist: linux57 linux58
real ip: 192.168.8.66
workmode: g
weight: 1
real ip: 192.168.8.77
workmode: g
weight: 1
```

## modify-config

用于把旧双机版本的配置文件转换成新版本的配置文件

```
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.
```

## product-switch

用于 HA 版本切换后，版本之间的配置转换。

```
cli:~>product-switch
Switch product success.
```



说明：

版本切换操作流程详细请参考[附录G—版本一键切换及授权方式变动说明](#)。

## 4.4 集群管理

集群管理包括对节点和服务的管理，对部分资源的管理,对snmp告警的管理等。

### 4.4.1 节点的管理

命令摘要：

序号	命令	功能描述
1	<a href="#">cluster-start</a>	启动集群软件，格式： <code>cluster-start [nodename]</code>
2	<a href="#">cluster-stop</a>	关闭集群软件，格式： <code>cluster-stop [nodename]</code> 。
3	<a href="#">cluster-restart</a>	重启集群软件，格式： <code>cluster-restart [nodename]</code> 。
4	<a href="#">cluster-restart-only</a>	用于只重启HA集群系统，不重启业务

#### Cluster-start

启动集群软件，如果服务脱离了集群管理，cluster-start不会带起服务。

#### Cluster-stop

停止集群软件，如果服务脱离了集群管理，cluster-stop只停止集群软件。

#### 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。

#### Cluster-restart-only

用于只重启HA集群系统，不重启业务。



提示：

1. [ ]为可选参数：为空表示在所有节点上都执行；如果为某节点名，则只在该节点上执行。
2. 修改集群配置、心跳配置或服务配置信息后，需要重启双机集群软件才能生效。

## 4.4.2 服务的管理

命令摘要：

序号	命令	功能描述
1	<a href="#">service-start</a>	手动启动服务，格式： service-start [service name] [hostname]
2	<a href="#">service-stop</a>	手动停止服务，格式： service-stop [service name]
3	<a href="#">service-enable</a>	使服务加入集群管理，格式： service-enable [service name]
4	<a href="#">service-disable</a>	使服务脱离集群管理，格式： service-disable [service name]
5	<a href="#">service-migrate</a>	手动切换服务，格式：service-migrate [service name] [destination hostname]
6	<a href="#">clear-fault</a>	清除服务的软件错误标识

### service-start

手动启动服务，如果不带参数则以交互方式执行命令。

```
Cli:~>service-start
Current service:
    0) oracle
    b) Back to previous menu
Please select a service: [0-0,b,q] 0
Current node:
    0) node1
    1) node2
    b) Back to previous menu
Please select a node: [0-1,b,q] 0
Send message to HA system to start service oracle on node1.
```

### Service-stop

手动停止服务，如果不带参数则以交互方式执行命令。

```
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-enable 和 service-disable

一般情况下，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。

**Service-enable**使服务加入集群管理，运行此命令的前提：

- 1、集群正在运行，节点状态正常(UP)，运行[show-status](#)查看即知
- 2、运行[show-status](#)查看服务状态，服务enable属性为no。

运行service-enable命令后再查看服务状态，确认服务enable属性变为yes；表明服务加入了集群软件管理，并确认服务状态正常。

此命令修改服务的enabled属性为yes，见[service-init](#)说明。

```
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)
注意：如果服务间存在依赖关系时，子服务enable，父服务同样会被enable，但父服务disable时，子服务不受影响，服务间不存在依赖关系不受影响。
Send message to the HA system to enable service oracle.
```

**Service-disable**使服务脱离集群管理，运行此命令的前提：

- 1、集群正在运行，节点状态正常(UP)，运行[show-status](#)查看即知
  - 2、运行[show-status](#)查看服务状态，服务enable属性为yes。
- 运行service-disable命令后再查看服务状态，确认服务enable属性变为no，表明服务脱离了集群管理。

此命令修改服务的enabled属性为no，见[service-init](#)说明。

```
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)
注意：服务间存在依赖关系时会有上述提示，处理原则与enable类似。
Send message to the HA system to disable service oracle.
```



注意

运行service-enable一定要在集群的节点状态正常(UP)时，否则可能导致服务倒换或者停止。



## Service-migrate

手动切换服务，执行命令后请运行[show-status](#)查看服务状态，确认服务能正常切换。

有多种执行方式：

### I. 交互方式切换服务，比较简单，适合手动操作

```
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) node2
    b) Back to previous menu
Please select a node: [0-0,b,q] 0
Send message to migrate service oracle from node1 to node2.
```

### I. 非交互方式切换单个服务

```
cli:~>service-migrate Node_B Node_A oracle
Send message to migrate service oracle from Node_B to Node_A.
```

### I. 非交互方式切换所有服务

```
cli:~>service-migrate Node_B Node_A
Send message to migrate all services from Node_B to Node_A.
```

### I. 此命令不会切换脱离了集群管理的服务。

## 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。

```
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。

### 4.4.3 资源的管理

资源属于服务的一部分，一般情况下，当服务启动时资源自动启动，服务停止时资源停止，这是Cluster对资源的自动管理；对资源的手动管理主要用于业务升级等场合。

命令摘要：

序号	命令	功能描述
1	<a href="#">resource-start</a>	手动启动服务的某个资源
2	<a href="#">resource-stop</a>	手动停止服务的某个资源
3	<a href="#">resource-enable</a>	管控服务中的特定资源
4	<a href="#">resource-disable</a>	不管控服务中的特定资源
5	<a href="#">resource-critical</a>	修改服务中的某一资源为关键资源。
6	<a href="#">resource-noncritical</a>	修改服务中的某一资源为非关键资源

#### resource-start

启动服务中的某个资源，如业务 ip,网卡，应用等，示例如下：

```
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，操作界面与 resource-start 类似，这里不再列举。说明一下，如果 Cluster 重新启动，此应用会随服务自动启动。

## Resource-enable

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。

```
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
```

## 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；操作界面与 resource-enable 类似，这里不再列举。

## Resource-critical

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。

```
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.
```

## Resource- noncritical

设定服务中的某个资源为非关键资源，当其运行状态异常时不会引起整个服务倒换，但有会相关日志信息输出。因操作界面与 resource-critical 类似，这里不再列举。

## 4.5 集群监控

集群监控主要包括节点、心跳链路、服务状态和服务相关资源状态的监控。

命令摘要：

序号	命令	功能描述
1	<a href="#">cluster-stat</a>	持续显示集群状态，每隔 1 秒自动刷新，格式为： cluster-stat
2	<a href="#">show-status</a>	显示一次集群当前状态，格式为：showstatus

### cluster-stat

每隔 1秒自动刷新一下状态，包括集群各节点工作状态、心跳链路状态、仲裁链路状态和各服务运行状态。该指令运行示例如下：

cli:~> <b>cluster-stat</b>					
Press Ctrl-C or 'Q' to exit					
Note: Ldisk – LockDisk, <b>ethx</b> – Abnormal link					
Select a service to view the status of resources: <b>oracle</b>					
Use the arrow keys to select: Up-First, Down-Last, Left-Previous, Right-Next					
NodeName	Status	HeartbeatLink	ServiceName	ServiceStatus	Enable
node1	UP	eth3,eth2	oracle	stopped	YES
		Ldisk(ONLINE)			
node2	UP	eth3,eth2	oracle	running	YES
		Ldisk(ONLINE)			

cluster-stat示意图

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
<b>NodeName</b>	集群中节点名称
<b>Status</b>	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
<b>HeartbeatLink</b>	心跳链路，会显示链路名称，如链路异常，会显示“-” 如配置了仲裁盘（Ldisk）也会这里显示
<b>ServiceName</b>	用户自定义的服务名称，见 <a href="#">service_init</a>
<b>ServiceStatus</b>	服务几种状态： running: 服务正在运行 stopped: 服务没运行

	<p>stopping: 服务正在停止  starting: 服务正在启动  unstable: 服务不稳定  unknown: 服务状态未知</p> <p><b>显示服务在什么状态下异常:</b>  str: Cluster启动服务失败  stu: 服务启动成功后, 之后某个时刻Cluster检测到服务状态不正常  stp: Cluster停止服务失败</p> <p><b>显示引起服务异常的原因:</b>  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>
<b>Enable</b>	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 (oracle), 按up或down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 oracle's resource status:					
Name	Type	Enable	Critl	NodeName	
ResourceStatus					
oracle_nic_0	nic	yes	yes	node1	running
				node2	running
oracle_ip_0	ip	yes	yes	node1	stopped
				node2	running
oracle_dm_0	disk	yes	yes	node1	stopped
				node2	running
oracle_app_0	app	yes	yes	node1	stopped
				node2	running

## 资源状态示意图

输出说明：

项	含义
<b>Name</b>	服务中的资源名称，用于内部识别
<b>type</b>	资源类型：如ip、nic(网卡)、disk(共享存储)和app(业务应用)
<b>NodeName</b>	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
<b>Enable</b>	资源可管理与否。Yes表示可管理
<b>Critical</b>	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

当手动停止node2节点的oracle服务时，服务自动切换到node1节点，cluster-stat显示如下，注意node2节点的oracle服务状态为**stopped:stu (A)**：

```
cli:~>cluster-stat
          Press Ctrl-C or 'Q' to exit
          Note: Ldisk - LockDisk

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

NodeName  Status  HeartbeatLink  ServiceName  ServiceStatus  Enable
node1     UP      eth3,eth2      oracle        running         YES
          Ldisk(ONLINE)

node2     UP      eth3,eth2      oracle        stopped:stu(A)  YES
          Ldisk(ONLINE)

str start fail, stp stop fail, stu status fail, op operation fail
A Application, I IP, S Storage, W Work link, B Split Brain
P Ping, T Time out, L Lock or unlock fail, M Mount
```



**警告：**Cluster依赖网卡运行，因此在Cluster运行过程中，禁止对网卡进行重启或者停止操作，否则可能导致集群状态异常甚至数据损坏。下列操作会重启网

卡：

- 1、 Suse Linux用yast配置网卡并确认。
- 2、 CGSL和类Redhat Linux发行版运行setup进行网络配置。
- 3、 在终端运行/etc/init.d/network restart等。

如果要重启网卡，请先停止集群。

---

## Show-status

打印当前状态，包括Cluster各节点工作状态、各服务运行状态、心跳链路状态。  
运行示例如下（如以下的多节点集群状态图）：

```
cli:~>show-status
```

NodeName	Status	HeartbeatLink	ServiceName	ServiceStatus	Enable
-----	-----	-----	-----	-----	-----
node1	UP	eth3,eth2 Ldisk(ONLINE)	oracle	running	YES
node2	UP	eth3,eth2 Ldisk(ONLINE)	oracle	stopped:stu(A)	YES
x64-4	UP	eth3,eth2 Ldisk(ONLINE)			

str – start fail   stp – stop fail   stu – status fail   op – operation error  
A – Application   I – IP   S – Storage   M – Mount  
W – Work link   T – Time out   B – Split brain  
P – Ping   L – Lock or unlock fail

## 4.6 配置 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。snmp 告警服务配置有两种配置方式，包括 SNMP v2c 和 SNMP v3，按需选择配置。

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

若配置 SNMP v2c，配置告警源 IP 及告警 IP（网管端 IP）如下：

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) Initialize SNMP  
1) Add SNMP source IP  
2) Add SNMP trap IP  
3) Add SNMPv3 user  
b) Back to previous menu

Please select an item to add: [0-1,b,q] 0 <sup>1</sup>

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

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

- 0) SNMP v2c
- 1) SNMP v3

Please select SNMP version to configure: [0-1,q] (0) <sup>2</sup>

Input trap IP address: [q] 192.168.1.194 <sup>3</sup>

Input trap port number between 0-65535: [q] (162)

Add another trap IP? [y/n] (n)

Input read-write community: [q] (public) <sup>4</sup>

Input SNMP source port number between 0-65535: [q] (161)

Add SNMP source IP? [y/n] (y)

Input IP address: [q] 192.168.111.2 <sup>5</sup>



```
Input netmask: [q] (255.255.255.0)
Select a NIC to start the IP.
Detected NICs on linux21: eth2 eth0 eth1
Input a network device: [q] (eth0)
SNMP information verification:
```

```
SourcePort: 161
NIC: eth0
SourceIP: 192.168.111.2
Netmask: 255.255.255.0
Rwcommunity: public
```

```
trapip: 192.168.1.194
trapport: 162
```

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

```
*      0) Initialize SNMP
        1) Add SNMP source IP
        2) Add SNMP trap IP
        3) Add SNMPv3 user
        b) Back to previous menu
```

```
Please select an item to add: [0-3,b,q] b
```

```
cli:~>
```

 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、选择告警服务类型, 这里配置 SNMP v2c。
- 3、配置告警 IP 和端口, 即配置告警发往哪个主机。
- 4、配置 snmp 读写团体串, 缺省为公用的 public,可按需修改成私有串体团。
- 5、配置告警源 IP。

若配置 SNMP v3, 配置告警源 IP 及告警 IP (网管端 IP) 如下:

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) Initialize SNMP
        1) Add SNMP source IP
        2) Add SNMP trap IP
        3) Add SNMPv3 user
        b) Back to previous menu
```

```
Please select an item to add: [0-1,b,q] 0
```

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

A unique rwcommunity

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

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

0) SNMP v2c

1) SNMP v3

Please select SNMP version to configure: [0-1,q] (0) 1 <sup>1</sup>

The following configuration SNMP V3 trap information:

Input trap IP address: [q] **192.168.1.194** <sup>2</sup>

Input trap port number between 0-65535: [q] (162)

Input security engine ID: [q] (0x0102030a0b0c) <sup>3</sup>

The following configuration SNMP V3 trap user information:

Input SNMPv3 username: [q] (Trap\_v3user) <sup>4</sup>

Please choose a security level:

0) noAuthNoPriv

1) authNoPriv

2) authPriv

Select a level: [0-2,q] (1) 2 <sup>5</sup>

Please choose an authentication protocol:

0) MD5

1) SHA

Select an authentication protocol: [0-1,q] (1) <sup>6</sup>

Input authentication protocol passphrase: [q] (Auth\_1234) <sup>7</sup>

Please choose a privacy protocol:

0) DES

1) AES

Select a privacy protocol: [0-1,q] (1) <sup>8</sup>

Input privacy protocol passphrase: [q] (Priv\_1234) <sup>9</sup>

Add another trap IP? [y/n] (n)

Input SNMP source port number between 0-65535: [q] (161)

Add SNMP source IP? [y/n] (y) <sup>10</sup>

Input IP address: [q] **192.168.111.2**

Input netmask: [q] (255.255.255.0)

Select a NIC to start the IP.

Detected NICs on linux21: eth2 eth0 eth1

Input a network device: [q] (eth0)

The following configuration SNMP V3 user information:

Input SNMPv3 username: [q] (SNMPv3\_user)

Please choose a security level:

0) noAuthNoPriv

1) authNoPriv

```

    2) authPriv
Select a level: [0-2,q] (1)2
Please choose an authentication protocol:
    0) MD5
    1) SHA
Select an authentication protocol: [0-1,q] (1)
Input authentication protocol passphrase: [q] (Auth_1234)
Please choose a privacy protocol:
    0) DES
    1) AES
Select a privacy protocol: [0-1,q] (1)
Input privacy protocol passphrase: [q] (Priv_1234)
Do you want to continue adding another SNMP V3 user? [y/n] (n)
SNMP information verification:

    SourcePort: 161
    NIC: eth0
    SourceIP: 192.168.111.2
    Netmask: 255.255.255.0

    user: SNMPv3_user
    security_level: authPriv
    auth: SHA
    auth_passwd: Auth_1234
    priv: AES
    priv_passwd: Priv_1234

    trapip: 192.168.1.194
    trapport: 162
    engineID: 0x0102030a0b0c
    trap_user: Trap_v3user
    security_level: authPriv
    auth: SHA
    auth_passwd: Auth_1234
    priv: AES
    priv_passwd: Priv_1234

Is this information correct? [y/n,q] (y)
*    0) Initialize SNMP
    1) Add SNMP source IP
    2) Add SNMP trap IP
    3) Add SNMPv3 user
    b) Back to previous menu
Please select an item to add: [0-3,b,q] b
cli:~>

```



说明:

- 1、选择配置告警服务 SNMPv3 功能。
- 2、配置告警目的 IP，即配置告警发往哪个主机
- 3、引擎 ID，主要用于验证使用，缺省即可。
- 4、添加目的告警用户。
- 5、选择验证加密方式，有三种验证加密方式，noAuthNoPriv 不开启验证，authNoPriv 开启验证不加密，authPriv 开启验证加密。
- 6、有两种验证方式，包括 MD5 和 SHA 验证。
- 7、添加验证密码，缺省 Auth\_1234。
- 8、有两种加密方式，包括 DES 和 AES 加密方式。
- 9、添加验证密码，缺省 Auth\_1234。
- 10、配置告警源，包括 IP、网卡、验证加密方式。

## snmp-delete

交互式指令，删除已配置的告警源 IP 及告警 IP，示例如下：

```
cli:~>snmp-delete
    0) Delete SNMP configuration
    1) Delete SNMP source IP
    2) Delete SNMP trap IP
    3) Delete SNMPv3 user
    b) Back to previous menu
Please select an item to delete: [0-3,b,q] 0
Do you want to save the changes? [y/n] (y)
Please restart the HA system to take your modification into effect.
```

## 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

```
cli:~>snmp-modify
    0) SourcePort: 161
    1) NIC: eth0
    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]
```

## 4.7 日志收集和其他

有问题请先阅读用户手册排查解决，无法解决请运行命令 `hareport` 收集信息，把集群节点的信息文件和 `cli->show-status` 或者 `cluster-stat` 输出发送到 [ha@gd-linux.com](mailto:ha@gd-linux.com)，我们有专门的工程师负责解答和解决。

```
node2:~ # hareport
Gathering NewstartHA configuration and log files now, Please wait a moment.....
Gather NewstartHA configuration and log files is ok,
the file is /var/log/haconfig-node2.tar.gz
```

### 查看某一命令的帮助信息

```
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 5 章 通过 web 界面进行集群配置管理

### 5.1 web 登陆

在某一节点上输入 **webadminserver start**，启动 web

```
[root@linux26 lic]# webadminserver start
```

```
Starting NewstartHA webadmin server.....
```

输入 **webadminserver status**，running 说明 web 启动成功

```
[root@linux26 lic]# webadminserver status
```

```
NewstartHA webadmin server is running...
```

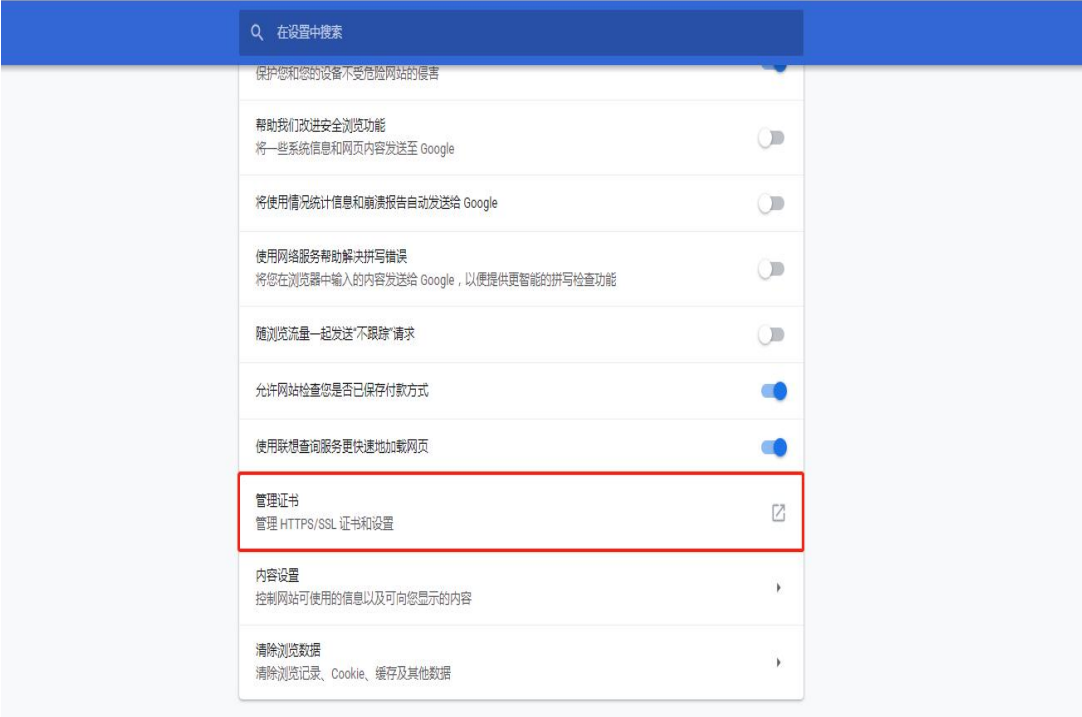


说明：

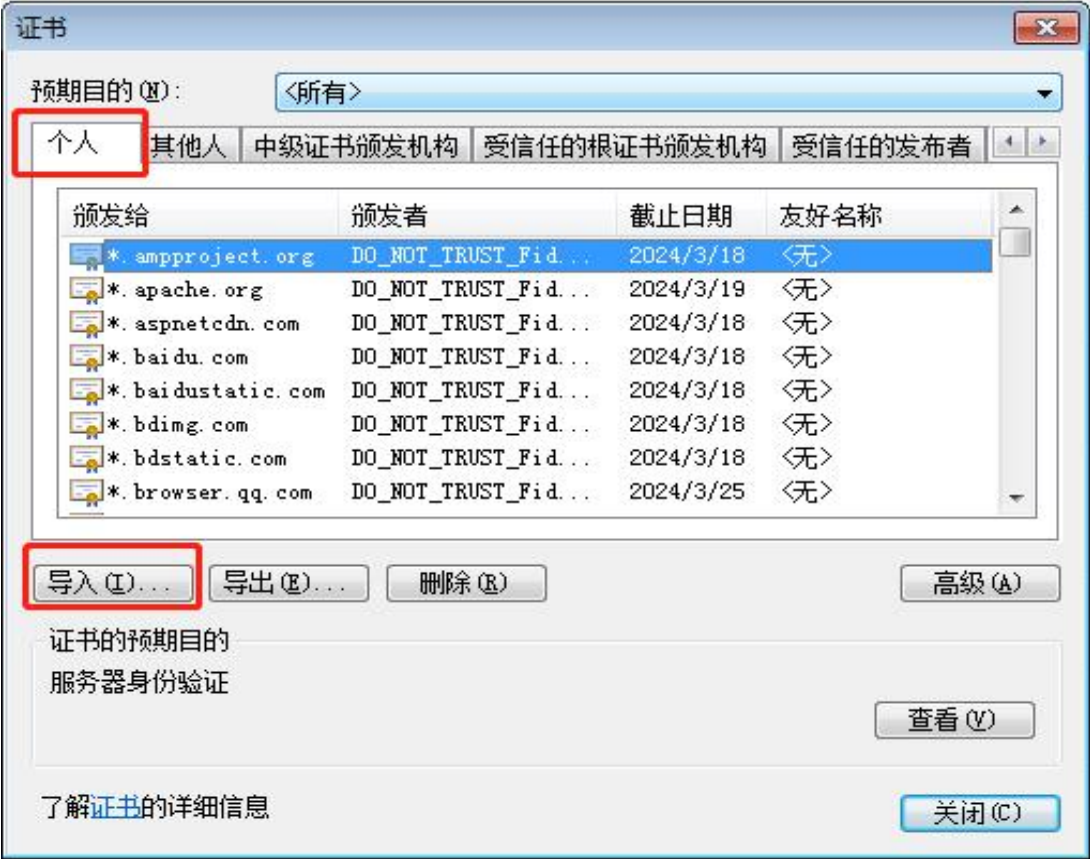
- 登陆 web 前需要下载证书，并将证书导入浏览器，推荐浏览器：Chrome、火狐、IE10，证书路径 /opt/NewStartHA/web/nginx/conf/mykey.p12。（密码：HA@NewStart123\_）

导入证书操作步骤如下：

5.1.1 以谷歌浏览器为例：找到“设置——高级——管理证书”



5.1.2 选择“个人”，点击“导入”



### 5.1.3 点击“下一步”

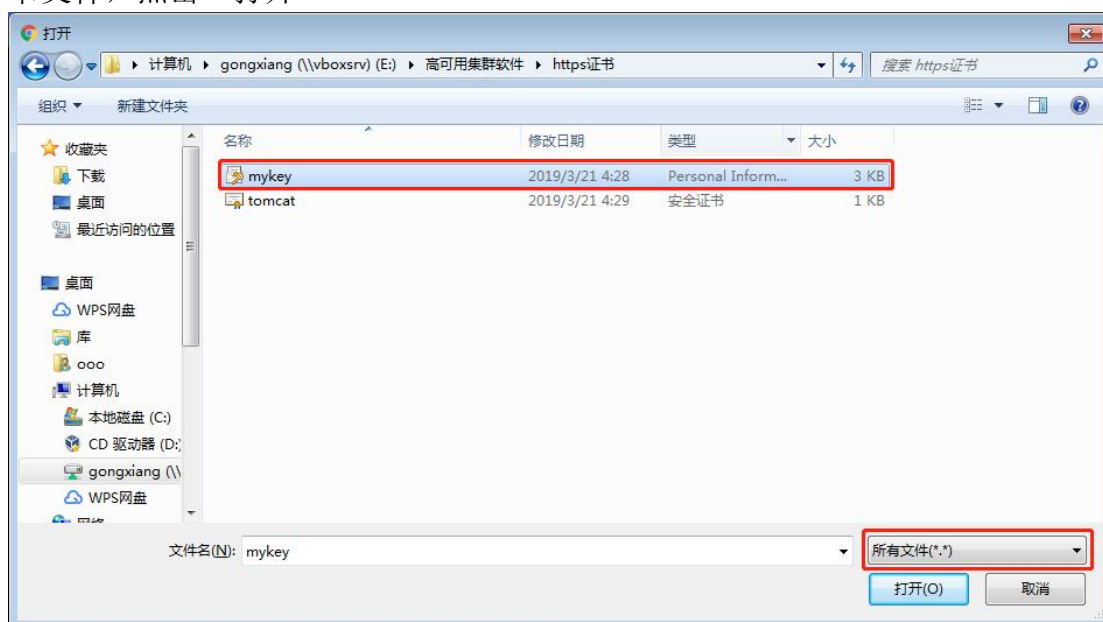


### 5.1.4 点击“浏览”





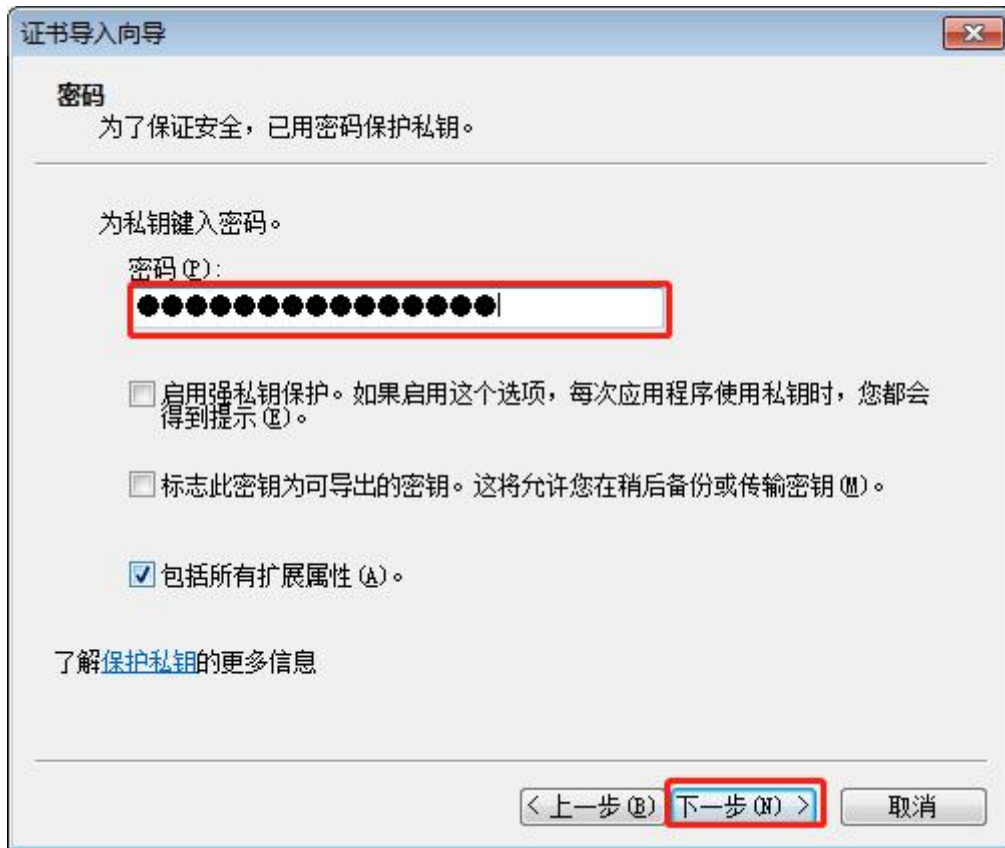
5.1.5 找到下载好的证书文件路径，下拉框选择“所有文件”，找到“mykey”证书文件，点击“打开”



5.1.6 点击“下一步”



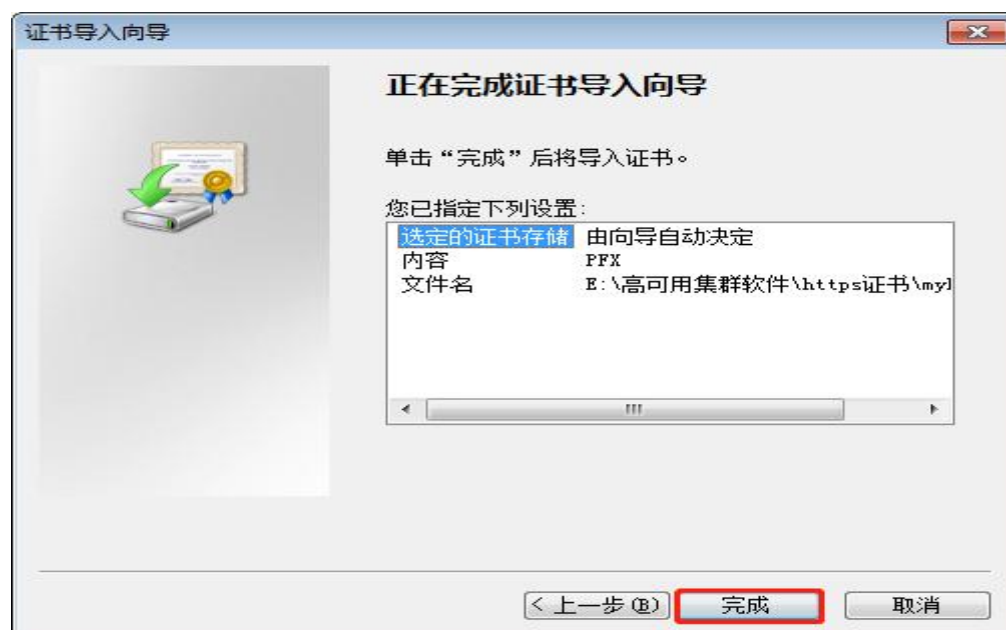
5.1.7 输入证书密码“HA@NewStart123\_”，点击“下一步”



#### 5.1.8 选择“根据证书类型，自动选择证书存储”，点击“下一步”



### 5.1.9 点击“完成”



### 5.1.10 导入成功



5.1.11 关闭浏览器，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”



### 5.1.12 弹出提示后选择证书，点击确认



### 5.1.13 进入登陆界面，如下:



© 2004 - 2019 广东中兴新支点技术有限公司，版权所有。

**缺省 web 超级管理员/密码：“SuperAdmin / HA@newstart123\_”**

Note:

- 1、出现访问异常时，检查防火墙的影响和 10003 端口是否被占用。
- 2、web 默认存在三个用户，分别是超级管理员 SuperAdmin、管理员 Admin、普通用户 User，密码均为“HA@newstart123\_”：

**超级管理员：**能够增加、修改、查看和删除配置文件，能够进行启停集群的操作及查看集群状态，能够添加、删除管理员和普通用户及修改用户的 web 登录密码；

**管理员：**能够增加、修改、查看和删除配置文件，能够进行启停集群的操作及查看集群状态；

**普通用户：**能够查看配置文件及集群状态；

3、所有用户进行 web 首次登录或者距离上次修改密码超过 1 个月，需要修改密码后才能进入 web 界面。

## 5.2 双节点集群配置（多节点集群配置类似）

### 5.2.1 集群参数配置

Web 登录后需进行集群的参数配置（页面中一些参数的缺省值保存即可），截图如下：

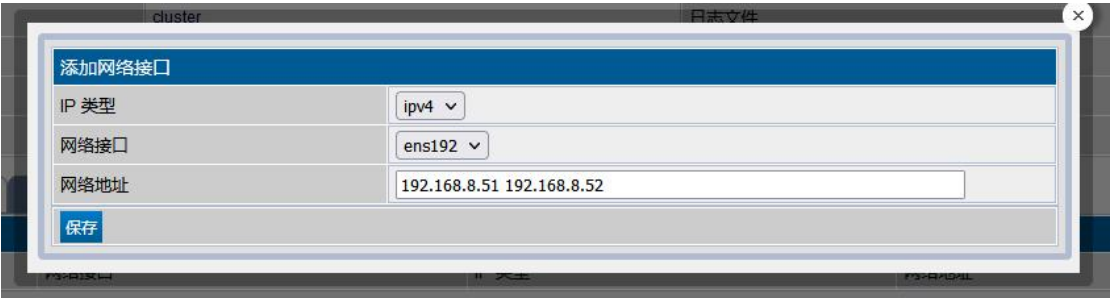
#### Note:

集群名：集群内部识别，自行定义。

本地主机 IP：IP 地址必须是网卡的固定 IP 地址，建议每个节点配置多个 IP 地址，以空格分隔(此 IP 用于双机配置文件同步或指令传输，一般为心跳网卡 IP)。

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。功能了解请查看 [2.6.1](#) 中的相关说明。

心跳链路配置



Note:

- 单播心跳添加步骤：1、基本参数配置--2、单播心跳--3、点击“添加”按钮。
- Cluster 支持多播、单播和串口心跳类型，推荐使用单播心跳，另外注意心跳链路总数不少于两条。
- IP 类型：支持 IPv4 和 IPv6 两种类型。
- 网络接口：填写对应的网卡名字。

## 5.2.2 服务参数配置

添加服务



添加服务	
服务名	db
服务类型	<input checked="" type="checkbox"/> 普通
镜像资源	<input type="checkbox"/> r0
域节点	cgs1v504f14 cgs1v504f16 (多个节点请使用空格分开)
域节点启动顺序	cgs1v504f14 cgs1v504f16 (多个节点请使用空格分开)
优先节点	cgs1v504f14 ▾
启动服务超时时间 (20~1200秒)	60
停止服务超时时间 (20~1200秒)	120
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
集群管理服务	<input checked="" type="checkbox"/> 管理
服务检测失败后	<input type="checkbox"/> 重启
<button>保存</button>	

**Note:**

- 服务名：集群内部识别，自行定义。
- 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。



配置网卡资源

NewStart HA  
5.0.10-20220902.x86\_64  
Product: Type = Two-node cluster

实时监控

服务监控

基本配置信息

集群服务配置

db

服务依赖配置

资源依赖配置

日志服务管理

其他服务管理

服务参数配置

服务属性参数

服务名	db	集群管理服务	<input checked="" type="checkbox"/> 管理
优先节点	cgslv504f14	域节点	cgslv504f14 cgslv504f16
域节点启动顺序	cgslv504f14 cgslv504f16		
启动服务超时时间 (20~1200秒)	60	停止服务超时时间 (20~1200秒)	120
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
---	-------	----

添加NIC资源

服务名

db

公共网卡资源名

db\_nic\_0

第三方主机IP

0.0.0.0

可管理资源

☒是

关键资源

☒是

主机名

cgslv504f14

设备名

ens192  
ens224  
ens256

主机名

cgslv504f16

设备名

ens192  
ens224  
ens256

Note:

- 1、 第三方主机 IP; 建议填写与业务链路同一网段的网络中已存在的 IP, 用于检测链路是否正常。
- 2、 设备名; 指业务链路物理网卡。

NewStart HA  
5.0.10-20220902.x86\_64  
Product: Type = Two-node cluster

实时监控

服务监控

基本配置信息

集群服务配置

db

服务依赖配置

资源依赖配置

日志服务管理

其他服务管理

服务参数配置

服务属性参数

服务名	db	集群管理服务	<input checked="" type="checkbox"/> 管理
优先节点	cgslv504f14	域节点	cgslv504f14 cgslv504f16
域节点启动顺序	cgslv504f14 cgslv504f16		
启动服务超时时间 (20~1200秒)	60	停止服务超时时间 (20~1200秒)	120
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
1	db_nic_0	<div><div></div><div></div><div></div></div>

公共网卡资源配置

公共网卡资源名	db_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	cgslv504f14	设备名	ens256
主机名	cgslv504f16	设备名	ens256

79



配置浮动 IP 资源

服务参数配置

服务属性参数

服务名	db	集群管理服务	<input checked="" type="checkbox"/> 管理
优先节点	cgs1v504f14	域节点	cgs1v504f14 cgs1v504f16
域节点启动顺序	cgs1v504f14 cgs1v504f16		
启动服务超时时间 (20~1200秒)	60	停止服务超时时间 (20~1200秒)	120
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

IP 资源

#	资源名标识	操作
---	-------	----

添加IP资源

服务名	db		
IP资源名	db_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
cgs1v504f14	ens192	192.168.55.15	255.255.255.0
cgs1v504f16	ens192	192.168.55.15	255.255.255.0
保存			

Note:

- a、公共网卡资源名：指的是运行业务的物理网卡。
- b、IP 类型：支持配置 ipv4/ipv6 的浮动 IP。

服务参数配置

服务属性参数

服务名	db	集群管理服务	<input checked="" type="checkbox"/> 管理
优先节点	cgs1v504f14	域节点	cgs1v504f14 cgs1v504f16
域节点启动顺序	cgs1v504f14 cgs1v504f16		
启动服务超时时间 (20~1200秒)	60	停止服务超时时间 (20~1200秒)	120
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

IP 资源

#	资源名标识	操作
1	db_ip_0	

IP资源配置

IP资源名	db_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
cgs1v504f14	ens192	192.168.55.15	255.255.255.0
cgs1v504f16	ens192	192.168.55.15	255.255.255.0

配置挂载资源

服务参数配置

服务属性参数

服务名	db	集群管理服务	<input checked="" type="checkbox"/> 管理
优先节点	cgslv504f14	域节点	cgslv504f14 cgslv504f16
域节点启动顺序	cgslv504f14 cgslv504f16		
启动服务超时时间 (20~1200秒)	60	停止服务超时时间 (20~1200秒)	120
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

DiskMount资源

#

资源名标识

操作

添加DiskMount资源

服务名	db
DiskMount资源名	db_dm_0
磁盘文件格式	disk
磁盘设备名	/dev/sdd6
挂载点	/home/hamount
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:

- a、磁盘文件格式；支持 NFS、LVM 和 DISK 三种类型的共享磁存方式。
- b、文件系统；主要手动能挂载成功的文件系统都可支持。
- c、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。
- d、网络存储；ISCSI 和 NFS 等都是网络存储，需选择 yes。
- e、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。

服务参数配置

服务属性参数

服务名	db	集群管理服务	<input checked="" type="checkbox"/> 管理
优先节点	cgsiv504f14	域节点	cgsiv504f14 cgsiv504f16
域节点启动顺序	cgsiv504f14 cgsiv504f16		
启动服务超时时间 (20~1200秒)	60	停止服务超时时间 (20~1200秒)	120
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

DiskMount资源

#	资源名标识	操作
1	db_dm_0	<div><div></div><div></div></div>

DiskMount资源

DiskMount资源名	db_dm_0		
磁盘文件格式	disk	磁盘设备名	/dev/sdd6
挂载点	/home/hamount	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

配置应用资源

服务参数配置

服务属性参数

服务名	db	集群管理服务	<input checked="" type="checkbox"/> 管理
优先节点	cgsiv504f14	域节点	cgsiv504f14 cgsiv504f16
域节点启动顺序	cgsiv504f14 cgsiv504f16		
启动服务超时时间 (20~1200秒)	60	停止服务超时时间 (20~1200秒)	120
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

App资源

#	资源名标识	操作
---	-------	----

添加应用资源

服务名	db
应用资源标识	db_app_0
应用脚本路径	/etc/ha.d/resource.d/myfaked.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>	

服务参数配置

服务属性参数

服务名	db	集群管理服务	<input checked="" type="checkbox"/> 管理
优先节点	cgslv504f14	域节点	cgslv504f14 cgslv504f16
域节点启动顺序	cgslv504f14 cgslv504f16		
启动服务超时时间 (20~1200秒)	60	停止服务超时时间 (20~1200秒)	120
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

App资源

#	资源名标识	操作
1	db_app_0	 

应用资源配置

应用资源标识	db_app_0
应用脚本路径	/etc/ha.d/resource.d/myfaked.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

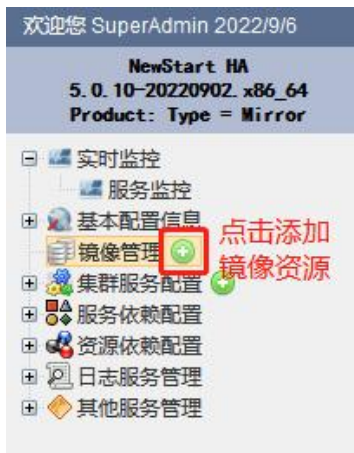
## 5.3 磁盘镜像集群配置

### 5.3.1 集群参数配置

配置方法与双节点集群基本相同，这里不再重复，具体请参考 [5.1.1](#) 小节。

### 5.3.2 磁盘镜像配置

添加镜像资源



配置镜像资源

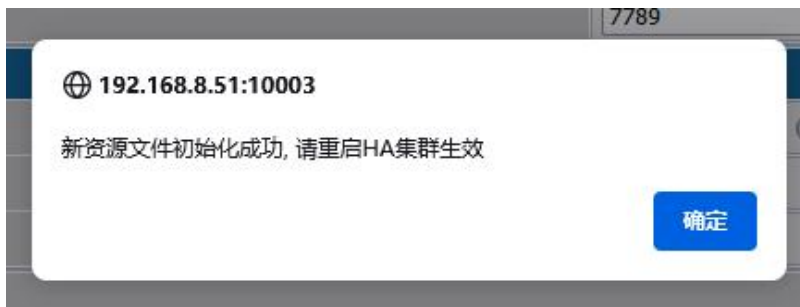
镜像参数配置增加

资源名	r0
cgslv504f14	
磁盘设备	/dev/sdb
IP地址(多个IP请用空格分开)	192.168.8.51
端口号(0~65535)	7789
cgslv504f16	
磁盘设备	/dev/sdb
IP地址(多个IP请用空格分开)	192.168.8.52
端口号(0~65535)	7789
Common	
同步源(主节点)	<input checked="" type="radio"/> cgslv504f14 <input type="radio"/> cgslv504f16
文件系统	ext3
同步方式	立刻同步
保存 取消	

## 数据同步中



镜像初始化完成，首次初始化不需要进行重启 HA 操作



### Note:

- 磁盘设备：底层磁盘，用于生成镜像资源，主备磁盘大小尽量保持一致
- IP：数据链路
- port：端口，默认即可
- 主节点：选择同步源
- 文件系统：按需选择，包括 xfs/ext2/ext3/ext4
- 同步方式：分为立即同步和暂不同步，选择暂不同步时直接添加配置文件，启动集群后会自动进行同步化
- 镜像资源配置完成并保存后会自动进行数据同步，同步时间与存储容量及网络有关。

## 5.3.3 服务参数配置

配置方法与双节点集群基本相同，这里不再重复，具体请参考 [5.1.2](#) 小节

## 5.4 配置告警信息（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，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ens192IP 类型ipv4

源IP192.168.56.45子网掩码255.255.255.0

源端口号(0~65535)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。  
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。

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public1

源IP参数

网卡资源名ens192IP 类型ipv4

源IP192.168.56.45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	192.168.8.51	162	

增加

Note:  
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，如下：

SNMP参数配置

版本类型

V2

V3

源IP参数

网卡资源名

ens192

IP 类型

ipv4

源IP(浮动IP)

子网掩码

源端口号

161

保存

取消

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。

SNMP参数配置

源IP参数

网卡资源名

eth1

IP 类型

ipv4

源IP(浮动IP)

12.12.12.21

子网掩码

255.255.255.0

源端口号

161

修改

源引擎ID (选配)

增加

v3用户

增加

目的IP参数

增加

源 IP 参数保存后，可配置源引擎 ID（选配）、v3 用户、目的 IP 参数三项。

点击增加，配置源引擎 ID（选配），若是不需要配置，可忽略

SNMP参数配置

源IP参数

网卡资源名

eth1

IP 类型

ipv4

源IP(浮动IP)

12.12.12.21

子网掩码

255.255.255.0

源端口号

161

修改

源引擎ID (选配)

增加



SNMP参数配置 

源IP参数

网卡资源名	eth1	IP 类型	ipv4 
源IP(浮动IP)	12.12.12.21	子网掩码	255.255.255.0
源端口号	161		

修改

源引擎ID (选配)

添加源引擎ID

源引擎ID (选配)	0x0102030a0b0c
------------	----------------

保存 取消

源引擎 ID（选配）：查询告警验证，配置输入字符时必须为十六进制；

## 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	eth1	IP 类型	ipv4 
源IP(浮动IP)	12.12.12.21	子网掩码	255.255.255.0
源端口号	161		

修改

源引擎ID (选配)

0x0102030a0b0c  

引擎ID	0x0102030a0b0c
------	----------------

v3用户

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

保存 取消

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。

点击增加，配置目的 IP 参数配置

目的IP参数

增加

SNMP参数配置

源IP参数

网卡资源名	eth1	IP 类型	ipv4
源IP(浮动IP)	12.12.12.21	子网掩码	255.255.255.0
源端口号	161		

修改

源引擎ID (选配)

0x0102030a0b0c

引擎ID	0x0102030a0b0c
------	----------------

v3用户

SNMPv3\_user

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

增加

目的IP参数

添加目的IP与端口号

目的IP	192.168.8.23	目的端口号	162
引擎ID	0x0102030a0b0c		
用户名称	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

配置告警目的 IP 和端口：即配置告警信息接收端。  
引擎 ID：主要用于接收告警验证使用，缺省即可，配置输入字符时必须为十六进制。

SNMPV3 配置完成

SNMP参数配置 

源IP参数

网卡资源名	eth1	IP 类型	ipv4 
源IP(浮动IP)	12.12.12.21	子网掩码	255.255.255.0
源端口号	161		

修改

源引擎ID (选配)

0x0102030a0b0c  

引擎ID	0x0102030a0b0c
------	----------------

v3用户

SNMPv3\_user 

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

增加

目的IP参数

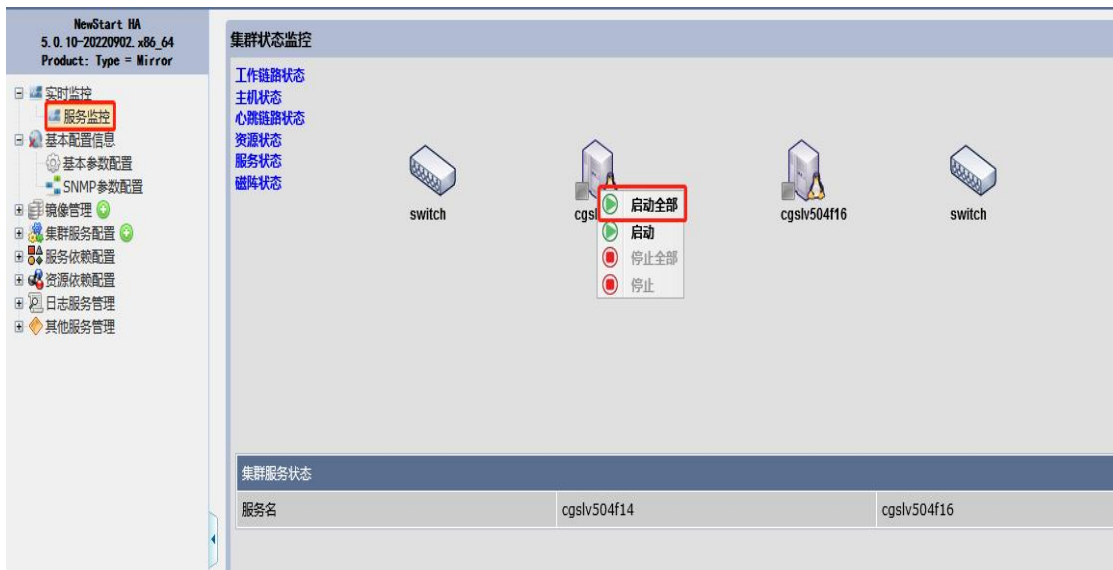
192.168.8.23  

目的IP	192.168.8.23	目的端口号	162
引擎ID	0102030a0b0c		
用户名称	Trap_v3user	安全级别	authPriv 
验证协议	SHA 	验证密码	••••••
加密协议	AES 	加密密码	••••••

增加

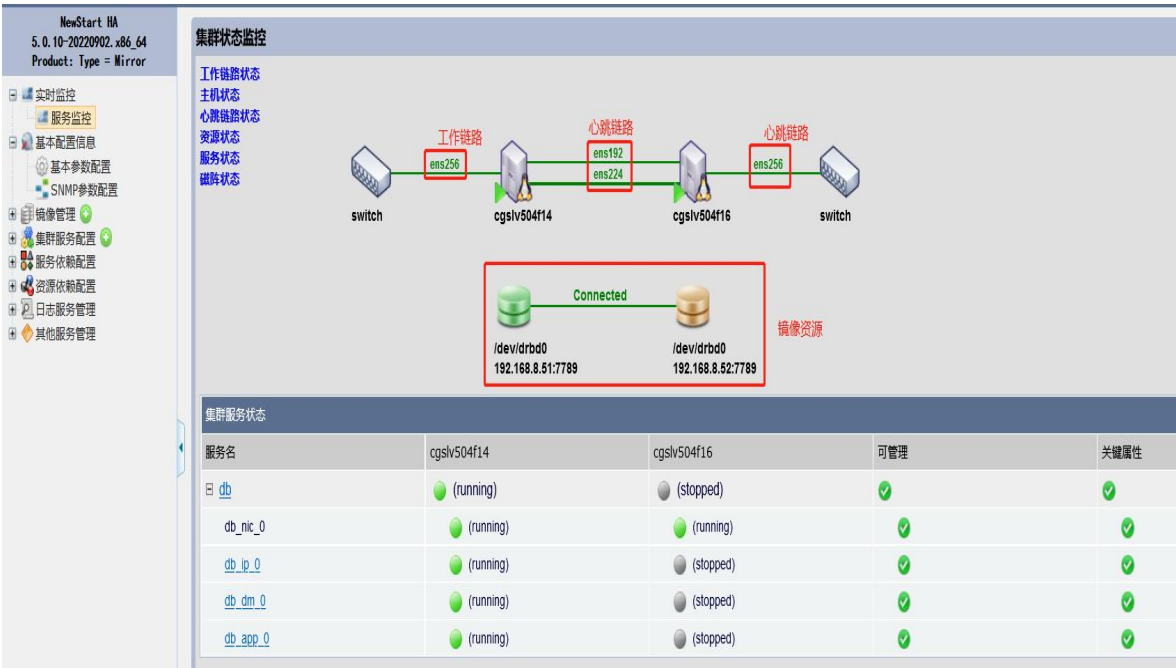
至此，web 下集群的配置完成。

## 5.5 集群启动



**Note:**

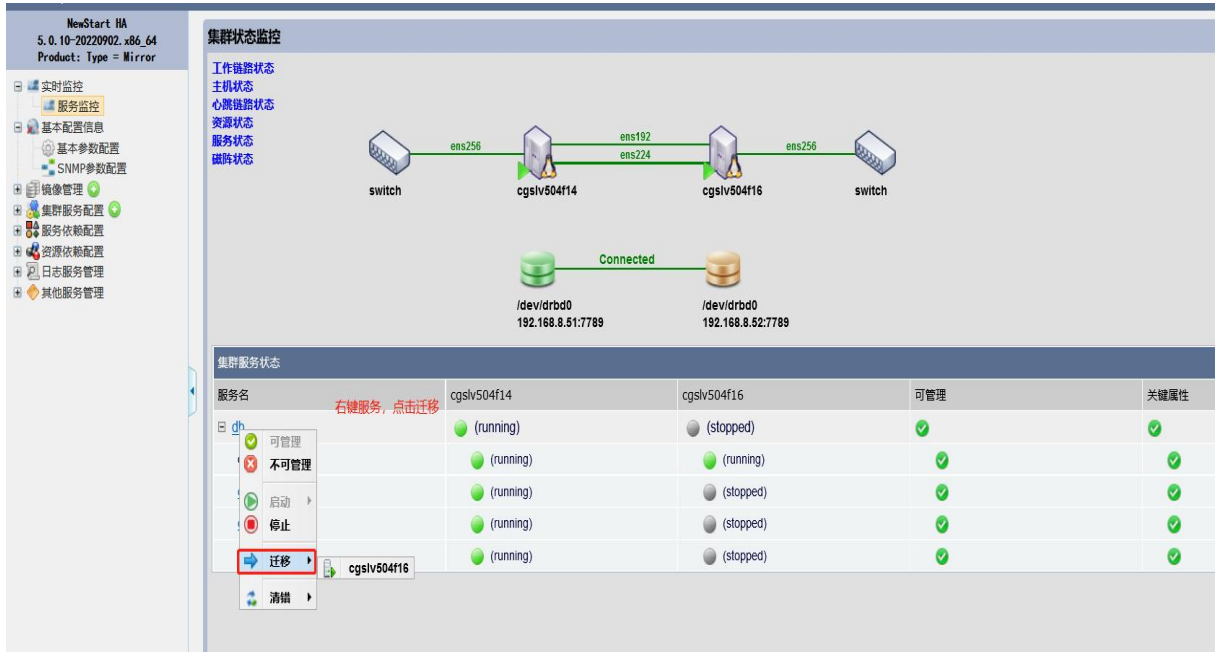
点击监控菜单，可在右边管理集群的启动和停止；查看集群和服务运行状态；上面截图操作执行后，正常情况下，稍等片刻就会显示如下。



**Note:**

- a. 集群状态监控：可用于查看各链路的运行状态
- b. 服务状态：可查看服务及其各资源的具体运行情况。

# 5.6 服务倒换测试



**Note:**  
集群配置完成后需要进行服务倒换测试，需要保证来回倒换（迁移）都正常。

## 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。
- I. 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。

# I. 附录

## 附录 A—Clusterl 命令快速参考

交互式命令：

命令	功能描述
<a href="#">cluster-init</a>	集群的初始化配置，详细用法请点击右边命令跳转查看，下同
<a href="#">mirror-init</a>	镜像资源初始化，磁盘镜像版本特有
<a href="#">service-init</a>	服务的初始化配置
<a href="#">heartbeat-add</a>	增加心跳链路
<a href="#">heartbeat-delete</a>	删除心跳链路
<a href="#">cluster-modify</a>	修改集群的初始化配置
<a href="#">lockdisk-adm</a>	仲裁盘管理
<a href="#">cluster-show</a>	显示集群的配置信息
<a href="#">service-add</a>	向配置文件添加服务
<a href="#">service-show</a>	显示指定服务的配置信息
<a href="#">service-modify</a>	修改一个服务的配置信息
<a href="#">resource-add</a>	增加服务的相关资源
<a href="#">resource-delete</a>	删除服务的相关资源
<a href="#">resource-modify</a>	修改服务的资源设定
<a href="#">service-delete</a>	从配置文件里删除服务
<a href="#">resource-depend</a>	设定服务的资源依赖
<a href="#">service-depend</a>	设定服务之间的依赖关系
<a href="#">check-config</a>	检查服务配置是否正确，配置不正确Cluster禁止启动
<a href="#">cluster-start</a>	启动集群软件
<a href="#">cluster-stop</a>	关闭集群软件
<a href="#">cluster-restart</a>	重启集群软件
<a href="#">cluster-restart-only</a>	用于只重启HA集群系统，不重启业务
<a href="#">service-start</a>	手动启动服务
<a href="#">service-stop</a>	手动停止服务
<a href="#">service-enable</a>	使服务加入集群管理
<a href="#">service-disable</a>	使服务脱离集群管理
<a href="#">service-migrate</a>	手动切换服务
<a href="#">clear-fault</a>	清除服务的软件错误标识
<a href="#">resource-start</a>	手动启动服务的某个资源
<a href="#">resource-stop</a>	手动停止服务的某个资源
<a href="#">resource-enable</a>	管控服务中的特定资源
<a href="#">resource-disable</a>	不管控服务中的特定资源
<a href="#">resource-critical</a>	修改服务中的某一资源为关键资源
<a href="#">resource-noncritical</a>	修改服务中的某一资源为非关键资源

<a href="#"><u>cluster-stat</u></a>	持续显示集群状态，每隔 1 秒自动刷新
<a href="#"><u>show-status</u></a>	显示一次集群当前状态
<a href="#"><u>snmp-add</u></a>	配置告警源IP及告警IP（网管端IP）
<a href="#"><u>snmp-delete</u></a>	删除告警源IP及告警IP
<a href="#"><u>snmp-modify</u></a>	修改告警源IP及端口设定
<a href="#"><u>mirror-adm</u></a>	镜像资源管理，包括增、删镜像，建议使用非交互命令nshamirror来操作。
<a href="#"><u>fence-add</u></a>	增加fencing功能，支持ipmi和scsi3
<a href="#"><u>fence-delete</u></a>	删除fencing功能
<a href="#"><u>instance-config</u></a>	用于增加、删除和显示已配置的实例
<a href="#"><u>ipvs-config</u></a>	用于负载均衡配置添加和删除
<a href="#"><u>modify-config</u></a>	把旧双机版本的配置文件转换成新版本的配置文件
<a href="#"><u>product-switch</u></a>	切换版本相关配置
recover-config-files	恢复配置文件
recover-remote-license	恢复对端的license文件
domain-config	域节点配置，仅适用于RCS/RDS
mirror-mode-switch	镜像同步模式切换，仅适用于RCS/RDS
quit	退出cli
exit	退出cli
help	展示交互式命令具体用法信息



## 非交互式命令：

命令介绍：

命令	描述
<a href="#">nshacluster</a>	配置 Cluster 集群基本信息
<a href="#">nshamirror</a>	配置和管理镜像模块
<a href="#">nshasrv</a>	配置和管理服务
<a href="#">nshares</a>	配置和管理资源
<a href="#">nshasnmpp</a>	配置和管理snmp告警
<a href="#">nshastart</a>	启动集群
<a href="#">nshastop</a>	停止集群
<a href="#">nshastat</a>	获取集群状态
<a href="#">nshaget</a>	获取集群版本
<a href="#">nshahide</a>	用于特殊功能的开启
<a href="#">nshaopsrv</a>	用于未启动集群情况下启停服务wei
<a href="#">nshacli</a>	用于非root用户执行非交互式命令

命令格式约定

符号	描述
{ }	大括号中的参数为必选项
[ ]	方括号中的参数为可选项
< >	尖括号的内容用实际值代替
	用于分隔若干选项，表示二选一或多选一
{x   y   z}	表示必须选择x，y，z 中的一个
[x {y   z}]	方括号的内容是可选，但若选择了方括号中的内容，则必须选择大括号中y，z中的一个

命令返回值：

返回值	描述
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

备注：

- 1) 命令的参数可以缩写。如 -add 参数， 可以缩写为-a 或者-ad;
- 2) 设置某个属性值为空，用双引号。如：  
nshasrv -modify srv1 preferrednode "" ;

## 集群配置

名称：**nshacluster**

功能：该命令用来配置Cluster集群基本信息，配置保存在文件/etc/ha.d/ha.cf中。

格式：

# **nshacluster -help vv**

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

-add {<cluster-name> | -node <node-name> -ip <ipaddr1 [ipaddr2] [ipaddr3]> |  
-hb {-mcast | -mcast6 | -ucast | -ucast6 | -serial} <nic-name>}  
[ipaddrs] | -lockdisk <disk-name> -format {yes | no}}  
-delete {-node <node-name> | -hb {-mcast | -mcast6 | -ucast | -ucast6 | -serial}  
<nic-name>} | -lockdisk <disk-name>  
-valid-lockdisk  
-format-lockdisk <disk>  
-add-domain <nodes> [-lockdisk <disk-name> -format {yes | no}]  
-del-domain <nodes>  
-modify <key> <value>  
-product-switch [<product-type>]  
-upgrade  
-chmod [<mode>] [<file>]  
-recover-config-files [<destination node>]  
-recover-lic <destination node>  
-get {product | version}  
-sn <sn>  
-display [code | product-type | <key>]  
-help

参数解释

参数	说明
-add <cluster-name>	添加集群名称,保存到集群配置文件中，并产生一些默认的集群信息，可用display参数查看
-add -node <node-name> -ip <ipaddr1 [ipaddr2] [ipaddr3]>	添加集群节点和IP
-add -hb {-mcast   -mcast6   -ucast   -ucast6   -serial} <nic-name> [ipaddrs]	添加心跳链路，可以添加的心跳链路有多播、单播和串口
-add -lockdisk <disk-name> -format {yes   no}	添加仲裁盘， yes表示格式化该仲裁盘
-delete -node <node-name>	删除集群节点
-delete -hb {-mcast   -mcast6   -ucast	删除心跳链路

-ucast6   -serial} <nic-name>}	
-delete -lockdisk <disk-name>	删除仲裁盘
-valid-lockdisk	重置仲裁盘，适用于已配置服务，但没有启动过集群
-format-lockdisk <disk>	初始化仲裁盘，适用于仲裁盘首次添加。
-add-domain <nodes> [-lockdisk <disk-name> -format {yes   no}]	配置节点域，适用于RCS/RDS集群
-del-domain <nodes>	删除节点域，适用于RCS/RDS集群
-modify <key> <value>	修改集群配置信息，key值可用display参数查看到和修改的结果
-upgrade	更新配置文件
-chmod	更新集群文件权限； [<mode>]:文件权限，不带默认644权限； [<file>]: 文件名称，不带文件，则默认全部集群文件都要更新权限
-recover-config-files [<destination node>]	恢复集群配置文件至其它节点，默认对像为其它所有节点
-recover-lic <remote node>	remote node:远程节点名称
-get {product   version}	product:产品类型；version:版本
-sn <sn>	sn:序列号，检查序列号是否有效
-display [code]	不带code参数显示集群配置信息，带code参数显示命令执行结果返回码的意义
-help	帮助参数，显示命令的用法

#### 举例：

```
linux176:~ # nshacluster -dis
configuration of the cluster:

    clustername: gdlc
    node: linux175 192.168.2.175 192.168.10.175
    node: linux176 192.168.2.176 192.168.10.176
    logfile: /var/log/ha-log
    keepalive: 1
    deadtime: 60
    watchdog: on
    mcast: eth2 226.0.222.76 694 1 0
    worklink_hb: off
    kernel_panic: off

linux176:~ # nshacluster -modify watchdog off
Please restart the HA system to take your modification into effect.
linux176:~ # nshacluster -dis
configuration of the cluster:
```

```
clustername: gdlc
node: linux175 192.168.2.175 192.168.10.175
node: linux176 192.168.2.176 192.168.10.176
logfile: /var/log/ha-log
keepalive: 1
deadtime: 60
watchdog: off //已修改
mcast: eth2 226.0.222.76 694 1 0
worklink_hb: off
kernel_panic: off
```

## 镜像配置

名称: **nshamirror**

功能: 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。

格式:

# **nshamirror -help vv**

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>] [-force]
-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>  
 -detach <res|dev>  
 -adjust <res|dev> [-node <node>]  
 -proxy-stop [-node <node>]  
 -proxy-start [-node <node>]  
 -connect <res|dev> [-node <node>]  
 -disconnect <res|dev> [-node <node>]  
 -convert-md <res|dev> [-node <node>] [-force]  
 -selinux [-node <node>]  
 -down <res|dev>  
 -up <res|dev>  
 -display [ code | <resname> | status-xml | online-xml | dump-xml [<resname>] |  
 meta-data <resname>]  
 -help

#### 参数解释

参数	说明
-add -res <resname> -node <nodename> -ip <ipaddr> [-port <portnum>] -disk <disk> [-metadisk <mdisk>]	添加镜像资源以及相关属性，resname代表自定义资源名，portnum代表端口号
-delete <resname>	删除镜像资源
-modify -res <resname> <key> <value>	修改镜像资源
-valid <resname> [-node <node>] [-force]	使已配置的镜像资源生效，支持强制模式
-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>	修改镜像配置，配合display <resname>来使用
-sync-rate <rate> [<resname>]	临时修改镜像同步速率
-start [-res <resname>] [-node <node>]	启动镜像资源
-stop [-res <resname>] [-node <node>]	停止镜像资源
-set <res dev> {-primary   -secondary} [-force]	设置镜像资源为主/备
-full-sync <resname>	全同步镜像资源
-splitbrain-recovery <res dev> -node <sync source>	数据裂脑恢复
-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，适用于底层盘扩容后的环境下执行

-add-link <res dev> -node <node> -ip <ipaddr> [-port <portnum>]	配置多条镜像链路
-switch-link <res dev> [-ip <ipaddr>] [-force]	镜像链路切换，仅适用于配置多条镜像链路场景
-switch-mode <mode>	切换集群同步模式，仅适用于RCS/RDS集群
-open-node <res dev> [-node <node>]	镜像资源运行用到，用户使用过程中一般不涉及
-create-script <res dev>	内部服务或内部资源添加时用到，用户使用过程中一般不涉及
-detect-mirror-mode <ipaddr res dev>	检测当前集群适用的集群模式:同城或远程，仅适用于RCS/RDS集群
-status [<res dev>] [detail]	查询当前镜像盘运行状态，detail可选参数表示可查看详细信息
-attach <res dev>	重新连接镜像盘的底层设备
-detach <res dev>	临时下线当前节点镜像盘
-adjust [-node <node>]	镜像参数即时生效
-proxy-stop [-node <node>]	停止同步代理模块
-proxy-start [-node <node>]	启动同步代理模块
-connect <res dev> [-node <node>]	触发节点间镜像盘连接尝试
-disconnect <res dev> [-node <node>]	断开节点间镜像盘连接
-convert-md <res dev> [-node <node>] [-force]	镜像资源元数据转换
-selinux [-node <node>]	刷新镜像模块的selinux策略
-down <res dev>	停止镜像资源
-up <res dev>	启动镜像资源
-display [ code   <resname>   status-xml   online-xml   dump-xml [<resname>]   meta-data <resname>]	显示镜像资源相关参数
-help	帮助参数，显示命令的用法

## 举例

```
[root@linux64 ~]# nshamirror -start -res r0 -node linux64
```

```
Waiting for mirror starting...
start mirror success.
```

## 服务配置

名称: **nshasrv**

功能: 该命令用来配置和管理服务, 配置保存在文件/etc/ha.d/haservices.xml中。

格式:

# **nshasrv -help vv**

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,配置保存到服务配置文件中,并产生一些默认的服务配置参数信息,可用display参数查看
-delete <srv>	删除服务
-modify <srv> <attr> <value>	修改服务属性,属性可用display参数查看
-link <parent srv> <child srv> <relationship> <depend type>	配置服务依赖。relationship包含: local、remote、global; depend type包含: fixed和forced。
-unlink <parent srv> <child srv>	删除服务依赖
-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>	使能服务,让Cluster管控服务

-disable <srv>	去使能服务，让服务脱离Cluster的管控
-migrate <srv> [<instance id>] <destination node>	迁移某个[实例]服务到某个节点运行
-clearfault <srv> [<node>]	清楚服务上错误标识，可指定某个节点
-checkcfg [<srv>]	检查服务是否配置完整
-addinstance <srv> <instance id> [<instance name>] <nodes> <res>	增加多实例服务及实例编号，instance name:实例名称，instance id:实例编号，指行实例运行节点及关联资源
-delinstance <srv> <instance id> [<res>]	删除多实例服务及实例编号，instance id:实例编号
-addmrrsrv <srv> {[<mirror resname>]   [<nic> <ipaddr> <netmask> <instance id>]}	添加MrrRcs资源
-delmrrsrv <srv> <resname>	删除MrrRcs资源
-display [-name <srv>   code   link   -instance <srv>]	不带任何参数，display只显示整个集群中的所有服务名称；-name 显示某个服务的具体配置信息；code 参数显示命令执行结果返回码；link参数显示服务依赖配置信息；-instance显示实例服务信息
-help	帮助参数，显示命令的用法

## 举例

```
linux176:~ # nshasrv -disp -name myfaked
Information of service 'myfaked':
Property information:
    enabled: yes
    preferrednode: linux175
    starttimeout: 60
    stoptimeout: 120
    checkinterval: 30
    checktimeout: 60
    maxerrcount: 1
    restartaftercheckfail: no
    startanyway: no

Resource information:
    myfaked_app_001
    dm
    myfaked_ip_01
    myfaked_nic_01

linux176:~ # nshasrv -modify myfaked preferrednode ""
linux176:~ # nshasrv -disp -name myfaked
Information of service 'myfaked':
```



```
Property information:
    enabled: yes
    preferrednode: //已修改
    starttimeout: 60
    stoptimeout: 120
    checkinterval: 30
    checktimeout: 60
    maxerrcount: 1
    restartaftercheckfail: no
    startanyway: no

Resource information:
    myfaked_app_001
    dm
    myfaked_ip_01
    myfaked_nic_01
```

## 资源配置

名称: **nshares**

功能: 该命令用来配置和管理资源，配置保存在文件/etc/ha.d/haservices.xml中。

格式:

**# nshares --help vv**

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>	为某个服务添加资源， type为资源类型，可用-display type 查看到Cluster的资源类型
-delete <res>	删除某个资源
-modify <res> <attr> <value>	修改资源属性，属性可用display参数查看
-link <parent res> <child res>	配置资源依赖
-unlink <parent res> <child res>	删除资源依赖
-start <res>	启动某个资源
-stop <res>	停止某个资源
-enable <res>	使能资源
-disable <res>	去使能资源
-critical <res>	将资源设为关键资源
-noncritical <res>	将资源设为非关键资源
-checkcfg <res>	检查某个资源是否配置完整
-display [-name <res>   type   code   link]	不带任何参数，display所有资源名称；-name 显示某个资源的具体配置信息；type参数显示资源类型；code 参数显示命令执行结果返回码；link参数显示资源依赖配置信息
-help	帮助参数，显示命令的用法

## 举例

```
linux176:~ # nshares -add my_ip IP myfaked
linux176:~ # nshares -display
Resources of the service 'myfaked':
    myfaked_app_001
    dm
    myfaked_ip_01
    myfaked_nic_01
    my_ip

linux176:~ # nshares -display -name my_ip
IP resource 'my_ip' property information:
    ipaddr:
    netmask: 255.255.255.0
    PubNICResName:
    enable: yes
    critical: yes

linux176:~ # nshares -modif my_ip ipaddr 192.168.50.100
linux176:~ # nshares -modif my_ip PubNICResName eth1
linux176:~ # nshares -display -name my_ip
```

IP resource 'my\_ip' property information:

ipaddr: 192.168.50.100  
netmask: 255.255.255.0  
PubNICResName: eth1  
enable: yes  
critical: yes

## SNMP 告警配置

名称: nshasnmpp

功能: 该命令用来配置snmp，配置信息保存在文件/etc/ha.d/haservices.xml和NSHA-agent.conf中。

格式:

**# nshasnmpp -help vv**

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

-add [-version {v2c|v3}]

-add -user <username> -sec\_level {noAuthNoPriv|authNoPriv|authPriv}  
[-authentication {MD5|SHA} -auth\_passwd <passphrase>] [-privacy  
{DES|AES} -priv\_passwd <passphrase>]

-add -sourceip <ipaddr>

-add -trapip <ipaddr> [-port <trapport>] [-community <community>] [{-engineID  
<id> -user <username>}  
{-sec\_level {noAuthNoPriv|authNoPriv|authPriv}} [-authentication  
{MD5|SHA} -auth\_passwd <passphrase>]  
[-privacy {DES|AES}} -priv\_passwd <passphrase>]]

-add -engineID <id>

-delete [-sourceip | -trapip <ipaddr> | -user <username> | -engineID <id>]

-modify <attr> <value>

-encrypt

-display [type|code|xml]

-help

参数解释

参数	说明
-add [-version {v2c v3}]	选择增加snmp版本类型，分为v2和v3
-add -user <username> -sec_level {noAuthNoPriv authNoPriv authPriv}	snmpv3版本增加用户和用户的相关属性; -user: 用户名称，自定义; -sec_level: 安全级别，分为三种noAuthNoPriv: 非加密非验证 authNoPriv: 验证非加密

[-authentication {MD5 SHA} -auth_passwd <passphrase>] [-privacy {DES AES} -priv_passwd <passphrase>]	authPriv: 既验证既加密; -authentication: 验证协议, 分为两种MD5和SHA; -auth_passwd: 验证密码, 可自定义; -privacy: 加密协议, 分为两种DES和AES; -priv_passwd: 加密密码, 可自定义;
-add -sourceip <ipaddr>	增加告警源IP地址
-add -trapip <ipaddr> [-port <trapport>] [-community <community>] [{-engineID <id> -user <username>} {-sec_level {noAuthNoPriv authNoPriv auth Priv}}] [-authentication {MD5 SHA} -auth_passwd <passphrase>] [-privacy {DES AES}} -priv_passwd <passphrase>]]	-add -trapip <ipaddr>: 增加目的IP地址; -port <trapport>: 增加目的端口; -community <community>: 团体属性配置 -engineID <id> -user <username>: 增加目的引擎ID和用户, 引擎ID必须是16进制字符; -sec_level: 安全级别, 分为三种noAuthNoPriv: 非加密非验证 authNoPriv: 验证非加密 authPriv: 既验证既加密; -authentication: 验证协议, 分为两种MD5和SHA; -auth_passwd: 验证密码, 可自定义; -privacy: 加密协议, 分为两种DES和AES; -priv_passwd: 加密密码, 可自定义;
-add -engineID <id>	增加告警源引擎ID, 必须是16进制字符
-delete [-sourceip   -trapip <ipaddr>  -user <username>   -engineID <id>]	-delete不带参数, 删除所有snmp配置信息; 若带上参数, 只删除对应的条目信息
-modify <attr> <value>	修改SNMP属性, 属性可用display参数查看
-encrypt	对snmpv3配置文件的用户密码进行加密
-display [type   code   xml]	不带任何参数, display显示SNMP配置信息; type参数显示SNMP可修改的属性; code 参数显示命令执行结果返回码; xml参数查看snmp的配置信息;
-help	帮助参数, 显示命令的用法

### 举例

```
linux176:~ # nshasnmpp -display
SourcePort: 161
NIC: eth1 //修改告警网卡

user: SNMPv3_user
security_level: authNoPriv
auth: SHA
auth_passwd: ODg4OEtaSEhMX0Vabm9zRCopKC8=

trapip: 192.168.8.24
trapport: 162
```

```

engineID: 0102030a0b0c
trap_user: Trap_v3user
security_level: authNoPriv
auth: SHA
auth_passwd: ODg4OEtaSEhMX0Vabm9zRCopKC8=
linux176:~ # nshasnmpp -modify NIC eth2
linux176:~ # nshasnmpp -display
SourcePort: 161
NIC: eth2 //修改完成

user: SNMPv3_user
security_level: authNoPriv
auth: SHA
auth_passwd: ODg4OEtaSEhMX0Vabm9zRCopKC8=

trapip: 192.168.8.24
trapport: 162
engineID: 0102030a0b0c
trap_user: Trap_v3user
security_level: authNoPriv
auth: SHA
auth_passwd: ODg4OEtaSEhMX0Vabm9zRCopKC8=

```

## CLUSTER 启动

名称：nshastart

功能：该命令用来启动集群。

格式：

**# nshastart -help**

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

[node name]

**-display** [code]

**-help**

参数解释

参数	说明
[node name]	可选参数，如果带节点启动，则是指定在该节点上启动Cluster，如果不带节点启动，则

	Cluster自动选择一节点启动
-display [code]	不带任何参数，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

## CLUSTER 停止

名称：nshastop

功能：该命令用来停止集群。

格式：

# nshastop -help

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

[node name]

**-display** [code]

**-help**

参数解释

参数	说明
[node name]	可选参数，如果带节点停止，则是指定在该节点上停止Cluster，如果不带节点停止，则所有节点都停止Cluster
-display [code]	不带任何参数，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

## CLUSTER 状态获取

名称: nshastat

功能: 该命令用来获取集群的状态, 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。

格式:

# **nshastat -help**

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

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

参数解释

参数	说明
-service <srv> <node>	获取某个服务在某节点上的状态, 状态码表示的意义可通过 <b>-display code</b> 查看
-resource <res> <node>	获取某个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> <node>	获取心跳链路状态, 支持获取指定节点的心跳链路状态
-status [detail]	显示整个集群状态, <b>detail</b> 支持显示详细信息
-abnormal <srv> [<node>]	获取节点服务状态类型
-display {code   hbtype}	<b>code</b> 参数显示所获取状态码的意义, <b>hbtype</b> 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

## CLUSTER 版本及类型获取

名称: nshaget

功能: 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。

格式:

# **nshaget -help**

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)
-help	帮助参数, 显示命令的用法

## CLUSTER 特定场景支持

名称: nshahide

功能: 此命令用于开启集群特定场景支持功能, 添加后, 配置信息保存在 /etc/ha.d/hide.cf。

格式:

# **nshahide**

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

**-add** <key> <value>  
**-delete** <key>  
**-display** [code|key]  
**-help**

展示特殊功能的具体信息及用法, value下前面的值就是开启特殊功能所需要的数值

# **nshahide -display vv**



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
network_upgrade	Time to wait for network upgrade. Do not stop services during network upgrade(all NICs down). Default 0 seconds means no waiting
worklink_maxerrcnt	Maximum number of working link detection errors. default 0
arping	1 Arping instead of Ping. default 0
ignore_consist	1 Ignore the consistency of configuration files. default 0
fuser_kill	0 Auto, 1 fuser -km device, 2 fuser -km mountpoint. default 0
run_same_domain	1 All services run on the same domain-node. default 0

#### 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>	并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

## 举例

如，开启单网卡启动集群功能

```
linux176:~ # nshahide -add singleNIC 1
linux176:~ # cat /etc/ha.d/hide.cf
singleNIC 1
```

## 服务启停操作

名称：nshaopsrv

功能：此命令用于在未启动集群的情况下启停服务。

格式：

**# nshaopsrv -help**

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

**-start** [<srv>]

**-stop** [<srv>]

**-help**

## 非 root 用户执行非交互式命令操作

名称：nshacli

功能：此命令用于非root用户可以执行非交互式命令。

格式：

**# nshacli**

Description:

Tool nshacli for non-root user management cluster.

Usage:

nshacli <command> [<option>] [<args>] ...

Commands:

nshacluster	Configuration and management of cluster
nshasrv	Configuration and management of services
nshares	Configuration and management of resources
nshamirror	Configuration and management of mirror

nshasnmpp	Configure SNMP
nshastart	Start cluster
nshastop	Stop cluster
nshastat	Display cluster status
nshahide	Configuration of cluster special functions
nshaget	Get version information for the cluster
nshaopsrv	Start or stop services without starting the cluster
hareport	Collect log information

### 举例

如，ha用户收集集群日志

```
[ha@linux176]:~ # nshacli hareport
Gathering NewstartHA configuration and log files now, Please wait a moment.....
Gather NewstartHA configuration and log files is ok,
the file is /var/log/haconfig-linux176.tar.gz
```

## 使用非交互式命令配置集群和服务范例：

#配置集群：

```
nshacluster -add cluster
```

```
nshacluster -add -node cgslv606-1 -ip 192.168.11.45 192.168.12.45
```

```
nshacluster -add -node cgslv606-2 -ip 192.168.11.46 192.168.12.46
```

```
nshacluster -add -hb -mcast ens224
```

```
nshacluster -add -hb -mcast ens226
```

```
nshacluster -modify worklink_hb on
```

#开始配置服务

```
nshasrv -add OA
```

```
nshasrv -modify OA multinstance no
```

```
nshasrv -modify OA nodelist 'cgslv606-1 cgslv606-2'
```

```
nshasrv -modify OA autostartlist 'cgslv606-1 cgslv606-2'
```

#配置业务链路网卡资源

```
nshares -add OA_nic_0 NIC OA
```

```
nshares -modify OA_nic_0 'cgslv606-1 device' 'ens192'
```

```
nshares -modify OA_nic_0 'cgslv606-2 device' 'ens192'
```

#配置业务IP资源

```
nshares -add OA_ip_0 IP OA
```

```
nshares -modify OA_ip_0 'cgslv606-1 PubNICResName' ens192
```

```
nshares -modify OA_ip_0 'cgslv606-1 ipaddr' 66.66.66.66
```

```
nshares -modify OA_ip_0 'cgslv606-2 PubNICResName' ens192
```

```
nshares -modify OA_ip_0 'cgslv606-2 ipaddr' 66.66.66.66
```

#配置存储资源

```
nshares -add OA_dm_0 DiskMount OA
```

```
nshares -modify OA_dm_0 type disk
```

```
nshares -modify OA_dm_0 device /dev/sdd9
```

```
nshares -modify OA_dm_0 mountpoint /home/hamount
```

#配置业务应用资源

```
nshares -add OA_app_0 Application OA
```

```
nshares -modify OA_app_0 script /etc/ha.d/resource.d/myfaked.sh
```

## 附录 B—SNMP 告警消息类型与查询说明

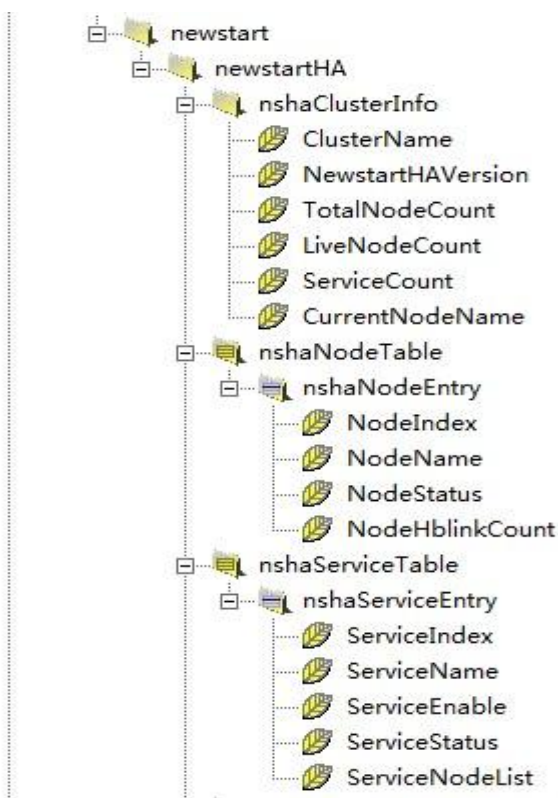
1) SNMP 告警的消息类型如下，可按需查阅：

Newstart Cluster SNMP Trap 接口				
告警类型	Trap Oid	严重级别	Trap Lable Name	Trap Description
资源告警	1.3.6.1.4.1.36895.1.90 0.1.0.2	Warning	resourceMonitorTimeoutTrap	资源监视已超时
	1.3.6.1.4.1.36895.1.90 0.1.0.3	Error	ResourceNotGoingOfflineTrap	Newstart Cluster 无法使资源脱机
	1.3.6.1.4.1.36895.1.90 0.1.0.5	Warning	ResourceWentOnlineByItselfTrap	资源自动联机
	1.3.6.1.4.1.36895.1.90 0.1.0.6	Error	ResourceFaultedTrap	资源在某个节点出现故障
服务告警	1.3.6.1.4.1.36895.1.90 0.2.0.1	Information	serviceOnlineTrap	服务启动成功
	1.3.6.1.4.1.36895.1.90 0.2.0.2	Information	serviceOfflineTrap	服务停止成功
	1.3.6.1.4.1.36895.1.90 0.2.0.4	Error	serviceFaultedTrap	服务故障
	1.3.6.1.4.1.36895.1.90 0.2.0.5	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。
	1.3.6.1.4.1.36895.1.90 0.2.0.6	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
	1.3.6.1.4.1.36895.1.90 0.2.0.7	Information	serviceInitiatingForSwitchingTrap	服务准备发生切换
	1.3.6.1.4.1.36895.1.90 0.2.0.8	SevereError	serviceConcurrencyViolationTrap	服务在集群中的多个节点处于联机状态
	1.3.6.1.4.1.36895.1.90 0.2.0.9	Information	serviceRestInRspnToPerstResGoOnlineTrap	服务组正在重新启动，因为已从故障恢复了持久性资源。
	1.3.6.1.4.1.36895.1.90 0.2.0.12	Error	serviceOPTimeoutTrap	对端服务操作超时

	1.3.6.1.4.1.36895.1.90 0.2.0.3	Information	serviceDisabledTrap	服务的 enable 属性 变为 no
节点告警	1.3.6.1.4.1.36895.1.90 0.3.0.2	Warning	clusterSystemRestartingman uallyTrap	用户手动重启集群 (用户通过 web 或命 令行发出了指令， 集群停止动作将要 开始执行)
	1.3.6.1.4.1.36895.1.90 0.3.0.3	Error	clusterSystemInDangerTrap	运行 Newstart Cluster 的节点进入 危险状态（心跳全 段）
	1.3.6.1.4.1.36895.1.90 0.3.0.4	SevereError	clusterSystemFaultedTrap	Newstart Cluster 因 故障而关闭
	1.3.6.1.4.1.36895.1.90 0.3.0.5	Information	clusterSystemJoinedClusterT rap	运行 Newstart Cluster 的节点加入 集群
	1.3.6.1.4.1.36895.1.90 0.3.0.6	Information	clusterSystemExitedManuall yTrap	Newstart Cluster 已 从先前正在运行的 节点中正常退出。
心跳告警	1.3.6.1.4.1.36895.1.90 0.4.0.1	Information	heartbeatOnlineTrap	心跳恢复
	1.3.6.1.4.1.36895.1.90 0.4.0.2	Warning	heartbeatOfflineTrap	心跳断开

## 2) SNMP 查询说明

目前 Cluster 仅支持对集群、节点和服务的相关信息查询，具体以下面列表说明为准。



查询说明：

	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 C—Cluster 应用脚本编写规范

- 集群应用脚本需包含 4 个参数选项，分别为 start, stop, forcedstop, status。

start 启动服务

stop 停止服务

forcedstop 强迫停止 (只在stop失败时被调用)

status 服务当前状态检测

- 脚本路径

脚本可以放置在用户指定的目录，如果不能确定目录权限，可放置在 /etc/ha.d/resource.d/ 目录；但是不要放在磁盘阵列上。

- 脚本文件是unix格式(区别于Windows格式)，必须具有可读和可执行权限，修改命令如下：

```
#chmod a+rx 脚本名
```

```
#dos2unix 脚本名
```

### 1、Cluster 操作脚本的机制

目前 Cluster 以 root 用户执行脚本。

- **启动应用脚本：**服务启动的最后步骤是启动应用脚本（即执行 start），对启动操作的返回值不做检测；2s 后检测服务的所有资源（即执行 status），如果都返回 0 Cluster 认为服务正常运行。
- **检测应用脚本：**在服务的缺省配置下(服务可管理)，集群启动后，主备机都会按设定时间间隔（缺省 30s）执行 status，如果返回 0 Cluster 认为服务正常运行，返回非 0 认为服务停止。这里分两种情况：
  1. 在服务运行节点（即主机）执行“script status” 返回非 0，然后开始停止服务。
  2. 在备机执行“script status” 返回 0，然后停止脚本；停止失败就会在集群范围内报告服务为 unstable，在 cli->cluster-stat 可以看到。所以有时候我们看到服务 unstable，但是实际业务正在运行的“奇怪现象”。
- **停止服务：**对服务的所有资源执行“script stop”，对停止操作的返回值不做检测；2s 后检测服务的所有资源（即执行 status），如果都返回非 0 Cluster 认为服务停止。



## 2、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]: 2022/06/28_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]: 2022/06/28_16:53:21 debug: start_resource:/bin/bash --login
/home/script/oracle.sh start
grm[14591]: 2022/06/28_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]: 2022/06/28_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 为例：/etc/ha.d/resource.d/ora\_script.ps)

#### 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）；

### 3、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 tnslnr | 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]: 2021/12/18_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]: 2021/12/18_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 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-lm (ESTABLISHED)
--------	------	--------	-----	------	-------	--	-----	---

tnslnr	10134	oracle	8u	IPv4	22260		TCP	*:ncube-lm (LISTEN)
--------	-------	--------	----	------	-------	--	-----	---------------------

tnslnr	10134	oracle	11u	IPv4	22314		TCP	e222:ncube-lm->e222:59034 (ESTABLISHED)
--------	-------	--------	-----	------	-------	--	-----	---

## 附录 D—Newstart Cluster 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

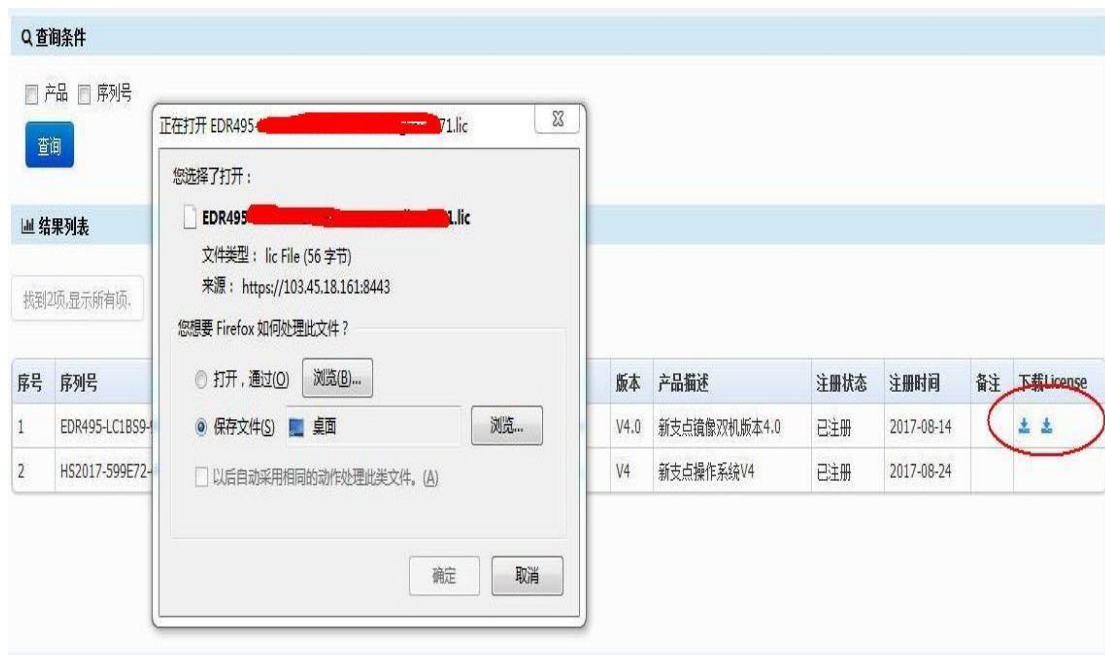
申请人	<input type="text" value="客户A"/>
项目信息	<input type="text" value="REDACTED项目"/>
手机号码	<input type="text" value="139xxxxxxxx"/>
邮箱	<input type="text" value="test@zte.com.cn"/>
公司或部门	<input type="text" value="REDACTED"/>

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 E—常见问题

### 1. 如何把集群试用版本转成商用版本？

**A:** 需要在两台服务器重新生成新的 key 文件，执行命令：

`/usr/lib64/newstartha/makekey` 或 `/usr/lib/newstartha/makekey`，输入商用序列号，若无报错，`newstartha.key` 文件则会生成在 `/etc/ha.d/lic/` 目录下。

**B:** 申请 license 参考 [3.2 节产品注册](#)，替换试用文件不会中断业务。

### 2. 运行 cli 出现 License file invalid 错误提示,如何处理？

**A:** 常见于授权申请操作有误或服务器硬件更换，此情况需要重新进行授权，把集群中所有节点的旧 key 文件及问题节点重新生成的 key 文件（key 生成工具：`/usr/lib(64)/newstartha/makekey`）一并发到 `ha@gd-linux.com`,并说明当前问题。

### 3. 修改主机名是否影响 Cluster 配置？

**A:** 修改主机名会对 Cluster 造成一定影响：

系统主机名相关文件

SuSE: `/etc/HOSTNAME`

RedHat: `/etc/sysconfig/network`

主机名相关文件：

`/etc/ha.d/ha.cf`

`/etc/ha.d/haservice.xml`

建议把 2 台机器下列文件删除，然后重起 Cluster

`/var/lib/newstartha/hb_generation`

`/var/lib/newstartha/hb_uuid`

### 4. Cluster 是否提供 Web 管理工具？

**A:** 提供 cli（命令行）和 web 工具进行集群、服务的初始化配置和日常管理，web 工具的详细操作可参考第五章节。



## 5. 如何关闭 Cluster 的 Web 管理工具 ？

A: 停止 Web:

```
# webadminserver stop
# webadminserver status  ——检查 Web 工具当前状态
```

系统启动时不启动 Web

```
# chkconfig webadminserver off
```

## 6. 是否支持集群告警（snmp 告警）监控，如何配置？

A: 支持，详细配置方法请参考 4.6 节。

## 7. 操作系统重装，是否需要重新申请 license？

A: 不需要，把原先的 key 及 license 文件覆盖拷贝回去即可。

## 8. Cluster 是否具备 LVM 锁盘功能？

A: 支持，请参考 [2.6.3 节](#)。

## 9. 以下安全加固操作是否会影响 Cluster 的运行？

- 修改 SSH 的版本，将 V1 版本修改为 Swan
- 修改 root 用户密码
- 禁止 roo 用户直接 SSH
- 关闭操作系统中的 rlogin、telnet、echo、discard、daytime、chargen、dtspc、exec、rexec、ntalk、finger、uucp 等服务；

A: 不会有影响

## 10. Web 登录忘记密码如何进行重置？

A: 若是忘记登录 web 的密码，需要发送邮件到 ha@gd-linux.com 邮箱进行账号密码重置。

## 附录 F—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动备份服务器。

### ■ 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	node1	node2	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](mailto: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
node1	UP	eth3,eth2,-- LDisk(ONLINE)	oracle	running	YES
node2	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 <对方节点心跳链路网卡的 IP> 通

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，可能原因如下请一一排查：

### 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
[node2]Starting High-Availability services:
Configuration file checked ok.
  NewStartHA failure [rc=3]. ..failed
License file invalid.
```

```
Configuration file checked ok.
[node1]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 &  
或者/usr/lib64/newstartha/tsever &

## ■ 数据裂脑故障

数据裂脑跟上述的集群裂脑情况不一样，数据裂脑一般出现在网络中断（也不排除有其它异常情况导致），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：

两个节点的 cs 都显示 StandAlone 状态

```
[root@linux21 home]# nshamirror -status
version: 8.4.11-1 (api:1/proto:86-101)
GIT-hash: 66145a308421e9c124ec391a7848ac20203bb03c build by root@centos120, 2019-10-28 07:22:40
0: cs:StandAlone ro:Primary/Unknown ds:UpToDate/Outdated   r-----
   ns:0 nr:0 dw:663888 dr:72979 al:18 bm:0 lo:0 pe:0 ua:0 ap:0 ep:1 wo:f oos:328
```

```
[root@linux22 home]# nshamirror -status
version: 8.4.11-1 (api:1/proto:86-101)
GIT-hash: 66145a308421e9c124ec391a7848ac20203bb03c build by root@centos120, 2019-10-28 07:22:40
0: cs:StandAlone ro:Secondary/Unknown ds:UpToDate/Outdated   r-----
   ns:0 nr:0 dw:663480 dr:4455 al:3 bm:0 lo:0 pe:0 ua:0 ap:0 ep:1 wo:f oos:0
```

或节点 1：cs"StandAlone"，节点 2：cs"WfConnection"状态

### 2、排查环境

- 先检查下镜像链路是否正常；
- 镜像同步端口 7788~7795 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、若两边 cs 状态均为"StandAlone"，则需手动干预修复。

**场景 1：**如果以当前服务运行节点（primary 节点）数据为同步源，同步到对端，则需要以下操作：

先查看镜像资源名 nshamirror -display

```
[root@linux19 ~]# nshamirror -display
Mirror resources:
   r0
```

在**服务运行节点**执行修复命令 `nshamirror -splitbrain-recovery r0 -node 'hostname'`，如下：

```
[root@linux21 ~]# nshamirror -splitbrain-recovery r0 -node 'hostname'
StandAlone

Manually split-brain recover successful
```

# **nshamirror -status** 持续观察状态

若数据在同步，直到变为 **Connected** 状态即恢复正常，如下：

```
[root@linux22~]# nshamirror -status
version: 8.4.11-1 (api:1/proto:86-101)
GIT-hash: 66145a308421e9c124ec391a7848ac20203bb03c build by root@centos120, 2019-10-28 07:22:40
0: cs:SyncSource ro:Primary/Secondary ds:UpToDate/Inconsistent A r-----
   ns:0 nr:0 dw:784060 dr:33204 al:21 bm:0 lo:0 pe:0 ua:0 ap:0 ep:1 wo:f oos:45828
     [>.....] sync'ed: 8.4% (45828/45828)K
     finish: 0:01:54 speed: 0 (0) K/sec
[root@linux22 ~]# nshamirror -status
version: 8.4.11-1 (api:1/proto:86-101)
GIT-hash: 66145a308421e9c124ec391a7848ac20203bb03c build by root@centos120, 2019-10-28 07:22:40
0: cs:SyncSource ro:Primary/Secondary ds:UpToDate/Inconsistent A r-----
   ns:11672 nr:0 dw:784088 dr:45872 al:21 bm:0 lo:11 pe:1 ua:12 ap:0 ep:1 wo:f oos:34184
     [=====>.....] sync'ed: 33.4% (34184/45828)K
     finish: 0:00:02 speed: 11,644 (11,644) K/sec
[root@linux22 ~]# nshamirror -status
version: 8.4.11-1 (api:1/proto:86-101)
GIT-hash: 66145a308421e9c124ec391a7848ac20203bb03c build by root@centos120, 2019-10-28 07:22:40
0: cs:SyncSource ro:Primary/Secondary ds:UpToDate/Inconsistent A r-----
   ns:31128 nr:0 dw:784088 dr:64304 al:21 bm:0 lo:0 pe:0 ua:0 ap:0 ep:1 wo:f oos:14728
     [=====>.....] sync'ed: 75.0% (14728/45828)K
     finish: 0:00:00 speed: 15,548 (15,548) K/sec
[root@linux22 ~]# nshamirror -status
version: 8.4.11-1 (api:1/proto:86-101)
GIT-hash: 66145a308421e9c124ec391a7848ac20203bb03c build by root@centos120, 2019-10-28 07:22:40
0: cs:SyncSource ro:Primary/Secondary ds:UpToDate/Inconsistent A r-----
   ns:45856 nr:0 dw:784088 dr:79032 al:21 bm:0 lo:0 pe:0 ua:0 ap:0 ep:1 wo:f oos:0
     [=====>.....] sync'ed: 100.0% (0/45828)K
     finish: 0:00:00 speed: 15,276 (15,276) K/sec
[root@linux22~]# nshamirror -status
version: 8.4.11-1 (api:1/proto:86-101)
GIT-hash: 66145a308421e9c124ec391a7848ac20203bb03c build by root@centos120, 2019-10-28 07:22:40
0: cs:Connected ro:Primary/Secondary ds:UpToDate/UpToDate A r-----
   ns:45856 nr:0 dw:784088 dr:79032 al:21 bm:0 lo:0 pe:0 ua:0 ap:0 ep:1 wo:f oos:0
```

**场景 2:** 如果不是以**当前服务运行节点**为数据同步源, 是以**对端节点**为最新数据作为同步源, 则需要进行以下操作:

步骤 1、进入 cli, 执行命令: `service-stop` , 然后选中对应的服务停止

```
cli:~>service-stop
Current service:
    0) NshaService
    1) myfaked
    b) Back to previous menu
Please select a service: [0-1,b,q] 1
Send message to HA system to stop service myfaked.
cli:~>show-status

Date: Wed Apr  1 17:43:13 2020
Note: MI - Mirror, dl - Diskless
Note: LDisk - LockDisk
Note: show-status detail - Display detailed status
```

NodeName	Status	HeartbeatLink	ServiceName	ServiceStatus	Enable
linux21	UP	eth1,eth2	NshaService	running	YES
		MI(r0,--)	myfaked	stopped	YES
linux22	UP	eth1,eth2	NshaService	stopped	YES
		MI(r0,--)	myfaked	stopped	YES

步骤 2、到对端节点设置为 **primary**, 执行 `nshamirror -set r0 -primary`, 操作如下:

```
[root@linux22 ~]# nshamirror -set r0 -primary

resource r0 set primary on node linux20 successful
```

步骤 3、在该节点运行修复命令 `nshamirror -splitbrain-recovery r0 -node 'hostname'`, 操作如下:

```
[root@linux22 ~]# nshamirror -splitbrain-recovery r0 -node 'hostname'
StandAlone

Manually split-brain recover successful
```

若数据在同步, 直到变为 **Connected** 状态即恢复正常, 如下:



```
[root@linux22 home]# nshamirror -status
version: 8.4.11-1 (api:1/proto:86-101)
GIT-hash: 66145a308421e9c124ec391a7848ac20203bb03c build by root@centos120, 2019-10-28 07:22:40
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
```

步骤 4、进入 cli，执行命令：service-start，然后选中对应的服务启动

```
cli:~>service-start
Current service:
    0) myfaked
    b) Back to previous menu
Please select a service: [0-0,b,q] 0
Current node:
    0) linux21
    1) linux22
    2) auto
    b) Back to previous menu
Please select a node: [0-2,b,q] 1
Send message to HA system to stop service myfaked.
cli:~>show-status

Date: Wed Apr  1 17:43:13 2020
Note: MI - Mirror, dl - Diskless
Note: LDisk - LockDisk
Note: show-status detail - Display detailed status


```

NodeName	Status	HeartbeatLink	ServiceName	ServiceStatus	Enable
linux21	UP	eth1,eth2 MI(r0,eth1)	NshaService myfaked	running stopped	YES YES
linux22	UP	eth1,eth2 MI(r0,eth1)	NshaService myfaked	stopped <b>running</b>	YES YES

若此步骤无法修复，见 3.2；

3.2、若 3.1 无法修复，考虑重新初始化备机镜像盘。



**禁止所有节点都进行此操作，否则会把镜像盘所有数据擦掉，数据不可恢复，如有需要帮助请先联系 HA 厂家再进行下面的操作。**

在备节点运行：

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

```
# nshamirror -valid r0
```

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

```
# nshamirror -status 观察状态，若出现数据同步，再看到 Connected 状态即恢复正常
```

```
[root@linux22 home]# nshamirror -status
```

```
version: 8.4.11-1 (api:1/proto:86-101)
```

```
GIT-hash: 66145a308421e9c124ec391a7848ac20203bb03c build by root@centos120, 2019-10-28 07:22:40
```

```
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 wo:f oos:0
```

## ■ web 页面无法正常显示

表现如下：

页面过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；

```
[root@linux19 ~]# webadminserver status
```

```
NewstartHA webadmin server is NOT running...
```

## 附录 G—版本一键切换及授权方式变动说明

版本支持集群多个版本（双节点、镜像）一键切换及节点软、硬件变动一键恢复授权功能。

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（双节点、镜像）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）
    - \*）添加镜像资源：cli->mirror-init/mirror-adm
    - \*）修改 disk 资源，修改 device 及 type 参数
  - 镜像切双机）
    - \*）修改 disk 资源，修改 device 及 type 参数
  - \*）disk 资源增删操作指令：cli-> 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”

>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

节点软、硬件变动一键恢复授权功能：（注：该版本只支持 cgs1 操作系统。）  
支持备机系统重装、硬件部件更换、集群中任一服务器故障更换、虚拟机克隆等。  
操作流程：

- 1) 问题节点故障恢复后，需要把操作系统配置修改成与原环境一致，包括节点名、IP、网卡名、集群配置文件（/etc/ha.d/ha.cf、haservices.xml）等
- 2) 问题节点运行 cli 查看授权情况
- 3) 如授权无效，则正常节点进入 cli，运行授权恢复命令：  
recover-remote-license，输入需要恢复的节点名，等待执行完成
- 4) 问题节点再次运行 cli 查看授权恢复情况

PS：如果正常节点执行 recover-remote-license 失败，优先检查问题节点 IP 是否改回原来配置？网络是否通？

## 附录 H—版本升级操作步骤

### 1. 注意事项

#### 1.1 是否需要重新申请 license 文件

\*) 小版本升级不需要重新申请 license，两台服务器/etc/ha.d/lic/目录下的 newstartha.key 及 newstartha.lic 文件升级前备份，升级后覆盖恢复即可。

\*) 大版本升级(如 4.0 升级到 5.0)涉及到商务及序列号变更，需要重新申请 license。

#### 1.2 若需要远程协助升级，需至少提前一天发邮件申请，并包含以下信息

\*) 几套双机；

\*) 升级前后版本信息；

\*) 项目名称：哪个省（海外写国家）的什么项目；

\*) 具体操作时间（北京时间）；

\*) 现场联系人员；

\*) 集群配置文件及环境信息，可直接在命令行下执行 hareport 命令收集，两个节点都要。

### 2. 升级前准备

#### 升级前检查表

序号	检查项目	检查标准（操作指令）	检查结果
1	查看当前 HA 版本	HA 版本升级前的版本（cli）	
2	查看服务运行状态	所有服务处于 running 状态（cluster-stat）	
3	获取所需的升级软件	获取所需的配套升级软件（我司官网）	
4	备份 HA 文件	HA 文件完毕（见 附录 H-2.4）	
5	记录 SN 号	执行 cli 查看	
6	备份 key 和 lic 文件	需要备份两台服务器目录下/etc/ha.d/lic 下的 newstartha.key 和 newstartha.lic 文件	

## 2.1 查看当前运行版本

使用命令 `cli`，进入 HA 操作命令行，查看 HA 版本及 SN 号。

```
# cli

NewStartHA

NewStart High Available System. Version Cluster_4.0.2.06, release 20160617.
Copyright 2008-2016 Guangdong NewStart Technology Service Ltd.

You are using a trial license, and 89 days remain. Please register.

Product: Type = Mirror, SN = 65L948-Q93DN1-32CR0M-10035A

Note: <b,c,q> -- <back,cancel,quit> <*> -- Option which you must configure
```

## 2.2 查看业务运行状态

运行 `cli->cluster-stat` 观察服务 1 分钟，确保所有服务状态 `running`，工作链路都是 `up` 状态，心跳链路 `Heartbeat link` 显示无闪烁。如果某服务状态为 `starting`，`stopping`，说明有动作未完成，请等待确保动作完成。下图是集群正常显示：

```
Note: MI - Mirror, dl - Diskless
Note: LDisk - LockDisk, ethx - Abnormal link

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

NodeName	Status	HeartbeatLink	ServiceName	ServiceStatus	Enable
linux19	UP	eth1,eth2 MI(r0,eth1)	eeee	running	YES
linux20	UP	eth1,eth2 MI(r0,eth1)	eeee	stopped	YES

## 2.3 获取所需软件版本

访问我司官网 [www.gd-linux.com](http://www.gd-linux.com)，依次选中【下载】---【高可用集群软件】获取需要的集群版本最新的 iso 包，为确保软件版本的完整性，建议下载后通过 sha 码来对比进行校验。



## 2.4 备份 HA 文件

小版本的配置文件虽然在升级过程不会覆盖，但是以防万一，需要将授权文件和配置文件进行备份，两台服务器的 HA 版本升级完毕后将所有文件重新恢复至原目录下。

1) 两台服务器的授权文件备份:

**/etc/ha.d/lic/newstartha.key;**

**/etc/ha.d/lic/newstartha.lic;**

2) 任一服务器的配置文件备份:

**/etc/ha.d/ha.cf;**

**/etc/ha.d/haservice.xml;**

**/etc/drbd.d/\*\*\*.res (若是镜像版本需要备份);**

**相关业务脚本;**

查询业务脚本位置：进入 cli--执行 service-show--服务名--查找 Application 下的 script，如下：

```
cli:~>service-show
Current service:
    0) eeee
    b) Back to previous menu
Select a(n) service: [0-0,b,q] 0
.....
Application
-----
name: eeee_app_0
script: /etc/ha.d/resource.d/myfaked.sh
enable: yes
```



说明：若是有个服务或者多个 app 资源，需要找出所有的业务脚本，并进行备份。

### 3. 升级操作

升级操作分两种情况进行，一是小版本升级操作，二是大版本升级操作，无论是两种情况，都需要执行上述的 **1.注意事项** 和 **2.升级前准备** 的步骤。

#### 3.1 小版本升级操作

小版本升级不需要重新申请 license。

##### 3.1.1 业务脱离 HA 的管理

1) 使用 cli 命令进入 HA 操作命令行

```
# cli
```

2) 使用 service-disable，将所有服务置为 disable 状态，若是有多多个服务，继续执行 service-disable 命令，将每一个服务都确保处于 disable 状态。

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

3) 使用 `cluster-stat` 查看服务状态，确保所有服务均为 `disable` 状态，服务的 `Enable` 属性均为 `NO`，如下图：

Note: MI - Mirror, dl - Diskless

Note: LDisk - LockDisk, ethx - Abnormal link

Select a service to view the status of resources: eeee

Use the arrow keys to select: Up-First, Down-Last, Left-Previous, Right-Next

NodeName	Status	HeartbeatLink	ServiceName	ServiceStatus	Enable
linux19	UP	eth1,eth2 MI(r0,eth1)	eeee	unknown	no
linux20	UP	eth1,eth2 MI(r0,eth1)	eeee	unknown	no

### 3.1.2 停止集群

在 `cli` 中执行 `cluster-stop` 停止集群(只在一个节点执行即可)，成功后退出 `cli`。

```
cli:~>cluster-stop
[linux21]Stopping High-Availability services:
Done.
.
[linux22]Stopping High-Availability services:
Done.
.
cli:~>
```

### 3.1.3 挂载 iso 包

将最新下载的的 `iso` 包上传至 `/home` 下，然后挂载到 `/mnt` 目录

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



### 3.1.4 升级

执行升级脚本: **/mnt/upgrade**

```
#/mnt/upgrade

Welcome to use NewStart Cluster Upgrade Wizard
Note: <y,n,c> -- <yes,no,cancel>

1) NewStart HA Server Program and CLI Administrative Tool
2) Web-based Administrative Tool
3) Mirror component

Please select the options? [1,2,3,c](1) 1

NewStart HA is installed. Do you want to upgrade it ?[y,n](y)
Uninstall...

Uninstalling the newstartha-Cluster ...
[OK]
Install...
Installing, please wait a few minutes...
Installing the /mnt/nsha/x86_64/redhat6/newstartha-Cluster_ ...
Preparing... #####
[100%]
1:newstartha
##### [100%]
Start tserver OK.
License files backuped in /etc/ha.d/lic/.
The cluster configuration files backuped in /etc/ha.d/.

Upgrade HA success.
#/mnt/upgrade

Welcome to use NewStart Cluster Upgrade Wizard
Note: <y,n,c> -- <yes,no,cancel>

1) NewStart HA Server Program and CLI Administrative Tool
2) Web-based Administrative Tool
3) Mirror component
```

Please select the options? [1,2,3,c](1)2 <sup>2</sup>

Webtool is stoped.

Uninstall webadmin...

[OK]

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

ha\_nginx installed OK !

tomcat installed ok!

web-based administrative tool installed ok!

Attention!! HA provides secure encrypted communication,please load certificates to browsers:

/opt/NewStartHA/web/nginx/conf/mykey.p12 (Default SN:HA@NewStart123\_)

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!



说明:

- 1)是升级 HA 组件；2)是升级 web 组件；3)是镜像模块。
2. 建议将 1)和 2)组件同时升级了，若是操作系统内核模块没有变化，不需要升级 3)组件，同时升级 1)和 2)组件时，升级完一个组件，重复上述步骤，再进行升级另外一个组件。

### 3.1.5 恢复授权及配置文件

操作建议请参考上述 [附录 H-2.4](#)。

### 3.1.6 备机升级

在另外一台机器上同样执行[附录 H-3.1.3](#)至[附录 H-3.1.5](#)中的操作。

### 3.1.7 启动集群并恢复业务监管

- 1) 使用 cli 命令进入 HA 操作命令行，查看 HA 版本，此时版本应该为升级之后的版本号

```
# cli
```

- 2) 在 HA 命令行界面执行 cluster-start

```
cli:~>cluster-start
```

3) 执行 `service-enable`，监管业务，如果有多个服务，待前一个服务状态成为 `running` 之后，再次执行 `service-enable` 来恢复 HA 对下一个业务程序的监控管理。

```
cli:~>service-enable
```

4) 使用 `cluster-stat` 查看服务状态，所有服务均显示为 `running`，此时说明 HA 升级成功

```
cli:~>cluster-stat
```

Note: MI - Mirror, dl - Diskless

Note: LDisk - LockDisk, ethx - Abnormal link

Select a service to view the status of resources: eeee

Use the arrow keys to select: Up-First, Down-Last, Left-Previous, Right-Next

NodeName	Status	HeartbeatLink	ServiceName	ServiceStatus
Enable				
linux19	UP	eth1,eth2 MI(r0,eth1)	eeee	running YES
linux20	UP	eth1,eth2 MI(r0,eth1)	eeee	stopped YES

## 3.2 大版本升级操作

大版本升级(如 4.0 升级到 5.0)由于序列号不通用，涉及到商务及序列号变更，需要重新申请 `license`，要确保已经购买了新版本的商用序列号，再进行升级操作。

### 3.2.1 停止集群及业务

由于大版本升级需要停止业务，因此需要提前准备在合适的时间进行升级操作

进入 cli，执行命令停止集群及业务，并执行命令检查是否停止成功：

```
cli:~>cluster-stop
[linux21]Stopping High-Availability services:
Done.

.
[linux22]Stopping High-Availability services:
Done.

.
cli:~>show-status
The HA system is not running now.
```

### 3.2.2 卸载旧版本

挂载旧版本的 iso 文件到/mnt 子目录：

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

或者将装有旧版本的光盘挂载到/mnt 子目录：

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

卸载旧版本，执行卸载脚本：

```
# /mnt/uninstall

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

Uninstall newstartha-Cluster_3.0.3.1-20160918.x86_64, Are you sure? [y/n](n)y
Checking ...                NOT running

Uninstalling the newstartha-Cluster_3.0.3.1-20160918.x86_64 ... [OK]

Uninstall DRBD... [OK]

Web Administrative Tool is running, uninstall will try stop it. [OK]

Uninstall webadmin... [OK]
```

### 3.2.3 安装新版本

参考 [第 3 章的 3.1 产品部署](#) 步骤，进行安装新版本。

### 3.2.4 恢复授权及配置文件

1. 授权文件需要重新申请 license，操作步骤参考 [第 3 章的 3.2 产品注册](#)。

2. 恢复配置文件，操作步骤参考 [附录 H-2.4](#)。

3.2.5 备机升级

在另外一台机器上同样执行[附录 H-3.2.2](#)至[附录 H-3.2.4](#)中的操作。

3.2.6 启动集群及业务

1) 使用 cli 命令进入 HA 操作命令行，查看 HA 版本，此时版本应该为升级之后的版本号

```
# cli
```

2) 在 HA 命令行界面执行 cluster-start

```
cli:~>cluster-start
```

3) 使用 cluster-stat 查看服务状态，待所有服务均显示为 running，此时说明 HA 升级成功

```
cli:~>cluster-stat

Date: Wed Apr 17 16:32:18 2017
Note: MI - Mirror, dl - Diskless
Note: LDisk - LockDisk, ethx - Abnormal link

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

NodeName      Status  HeartbeatLink  ServiceName  ServiceStatus  Enable
linux19       UP      eth1,eth2      eeee         running        YES
              MI(r0,eth1)
linux20       UP      eth1,eth2      eeee         stopped        YES
              MI(r0,eth1)
```

3.3 内核升级后的 HA 版本升级操作

若是操作系统内核升级后，HA 现有版本不支持，需要升级新的版本进行适配，HA 版本升级操作参考 [附录 H-3.2 大版本升级操作](#)，区别在于是否重新申请

license:

1. 若是小版本升级不需要重新申请 license。
2. 若是大版本升级(如 4.0 升级到 5.0)需要购买商用序列号重新申请 license。

## 技术支持联系

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) 将在每个服务器命令下运行 `hareport` 生成的压缩包文件，发送过来
- 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

■ 项目信息（国外项目请注明国家或地区）

■ 联系人、联系方式

■ 公司名称或工作部门

### 2、 联系方式

技术服务邮箱： [ha@gd-linux.com](mailto: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：51066

