



银河麒麟高可用集群软件 V10¹

Booth 配置说明

麒麟软件有限公司

2024 年 4 月

¹本手册适用于银河麒麟高可用集群软件 V10 SP3 2403

目录

目录	I
1 产品概念及术语介绍	2
2 环境准备	2
3 Booth 配置	2
3.1 配置/etc/hosts	2
3.2 创建 booth.conf	3
3.3 创建票证(ticket)	4
3.4 在仲裁节点启动并启用 booth 服务	4
3.5 为每个集群创建浮动 IP(虚拟 IP)	4
3.6 初始启动的票据授权	5
3.7 资源增加票证约束	5
3.8 启动资源测试	6
3.9 自动故障切换测试	6
3.10 手动故障切换测试	7

1 产品概念及术语介绍

在企业中，针对一些企业的关键系统应用，常采用多站点多节点的集群容灾架构。生产站点和容灾站点常常分布于不同的地方，跨越距离较远。

当集群跨越多个站点时，站点之间的网络连接问题可能会导致脑裂情况。当连接断开时，一个站点上的节点无法确定另一个站点上的节点是否发生故障。此外，在两个相距太远而无法保持同步的站点之间提供高可用性服务可能会出现问題。为了解决这些问题，Pacemaker 通过使用 Booth 集群票证管理器，全面支持配置跨多个站点的高可用性集群的能力。Booth 通过票据授予情况使得在同一时间内资源可以特定运行在某一站点上，也实现跨常规集群之间的资源高可用。

2 环境准备

每个集群需要准备一个额外的 IP，booth 服务将基于这一个 IP 进行跨站点以及和仲裁节点之间进行通信。

普通集群节点上需要安装 booth、booth-core 及 booth-site 软件包，仲裁节点需要安装 booth、booth-core 及 booth-arbitrator 包。

3 Booth 配置

在跨站点高可用集群管理中，双站点+单仲裁节点的场景是最常见的一个业务场景。下面以这一场景为例，介绍 Booth 的配置方案。两个站点下集群基础配置约定如下：

集群一（hacluster）包括 host1 和 host2 节点，为集群一配置的虚拟 IP 是 172.17.127.208；

集群二（hacluster2）包括 ha1 和 ha2 节点，为集群二配置的虚拟 IP 是 172.17.127.209；

仲裁节点的 ip 是 172.17.127.203；

 注记：不同站点下的集群名称需要设置为不同的集群名称。

3.1 配置/etc/hosts

两个集群中每个节点以及仲裁节点的/etc/hosts 中写入其他集群节点的 ip，以保证后续认证通过。示例如下：

```
[host1 ~] # cat /etc/hosts
127.0.0.1      localhost      localhost.localdomain  localhost4
               localhost4.localdomain4
::1           localhost      localhost.localdomain  localhost6
               localhost6.localdomain6
172.17.127.228 host1
172.17.127.229 host2
172.17.127.231 ha1
172.17.127.232 ha2
172.17.127.203 ha203
```

3.2 创建 booth.conf

booth 的配置文件会保存在/etc/booth/路径下，通过下面的命令创建 booth 配置文件并指定集群站点 IP 和仲裁节点的 IP 信息，然后将 booth 配置信息同步到当前站点中集群其他节点上。

集群一：

```
[host1 ~] # pcs booth setup sites 172.17.127.208 172.17.127.209
arbitrators 172.17.127.203
[host1 ~] # pcs booth sync
```

 注记：将 Booth 配置文件(/etc/booth/文件夹下文件)同步到当前集群中的所有节点（单节点集群由于只有一个节点不需要执行这一步骤，以下均如此）。

仲裁节点：仲裁节点同步集群一上的 booth 配置信息。

```
[仲裁节点 ~ha203] # pcs host auth host1
[仲裁节点 ~ ha203] # pcs booth pull host1
```

集群二：集群二同步集群一上的 booth 配置信息。

```
[ha1 ~] # pcs host auth host1
[ha1~] # pcs booth pull host1
[ha1 ~] # pcs booth sync
```

3.3 创建票证(ticket)

为 Booth 配置创建票证(superticket 为自定义票证名称, 可按需修改), booth 票证用于定义资源约束, 该限制将允许资源仅在此票证被授予的集群上运行。

```
[host1 ~] # pcs booth ticket add superticket
```

3.4 在仲裁节点启动并启用 booth 服务

```
[ha203 ~]# pcs booth start  
[ha203 ~]# pcs booth enable
```

 注记: 不要在集群的任何节点上手动启动或启用 Booth, 因为 Booth 在这些集群中作为 Pacemaker 资源运行。

3.5 为每个集群创建浮动 IP(虚拟 IP)

booth 服务之间需要配置一个单独的 ip 进行通信, 设置命令如下:

集群一:

```
[host1 ~] # pcs booth create ip 172.17.127.208
```

集群二:

```
[ha1 ~] # pcs booth create ip 172.17.127.209
```

 注记: 这里的 ip 是前面为 booth 准备的一个业务 ip。这一命令也会创建一个名字为 booth-booth-group 的组资源。

启动 booth-booth-group 组资源:

集群一:

```
[host1 ~] # pcs resource enable booth-booth-group
```

集群二:

```
[ha1 ~] # pcs resource enable booth-booth-group
```

3.6 初始启动的票据授权

初次启动 booth 以后，需要为集群授予票据，命令如下：

```
[ha1 ~] # pcs booth ticket grant superticket
```

 注记：在选定的一个集群的一个节点上执行即可，必须在资源启动以后才可以执行。

```
[root@ha1 ~]# pcs booth status
TICKETS:
ticket: superticket, leader: 172.17.127.209, expires: 2023-02-20 09:51:33

PEERS:
site          172.17.127.208, last recv: 2023-02-20 09:51:20
              Sent pkts:36 error:0 resends:0
              Recv pkts:36 error:0 authfail:0 invalid:0

site          172.17.127.209, last recv: 2023-02-20 09:51:33
              Sent pkts:4 error:0 resends:0
              Recv pkts:4 error:0 authfail:0 invalid:0
```

可以看到此时票据被授权到 hacluster2 上。

3.7 资源增加票证约束

Booth 提供了票据的核心概念，只有在集群被授予票据以后集群上和票据绑定的资源才可以运行。以 Dummy 资源为例，进行票据和资源的约束绑定（在实际场景中，将 dummy 这一资源 id 换成对应的资源 id 即可）：

集群一：

```
[host1 ~] # pcs constraint ticket add superticket dummy
```

集群二：

```
[ha1 ~] # pcs constraint ticket add superticket dummy
```

3.8 启动资源测试

集群一：

```
[host1 ~] # pcs resource enable dummy
```

集群二：

```
[ha1 ~] # pcs resource enable dummy
```

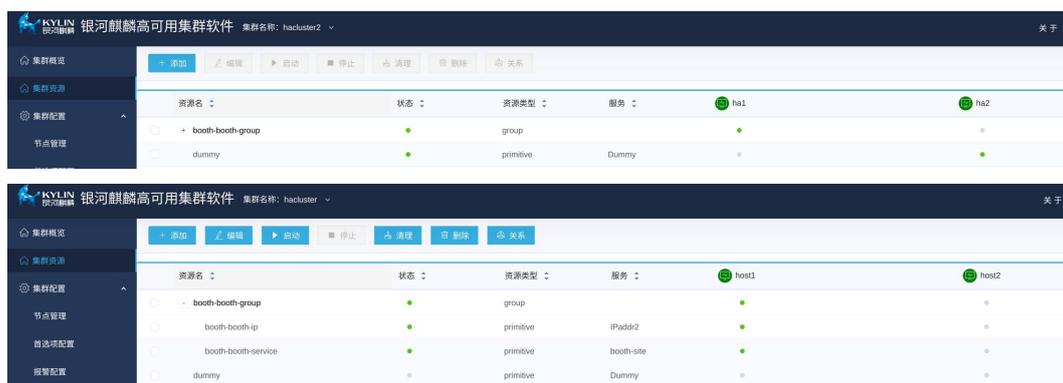


图 3-1 启动资源测试

可以看到 dummy 资源在 hacluster2 上可以启动，但是 dummy 在 hacluster 集群上由于没有被授予票据而无法启动。

3.9 自动故障切换测试

假设票据原来在 hacluster2 上面，关闭 hacluster2（模拟集群二宕机的场景），然后在过了原来票据的 expire 时间以后，可以看到票据从 hacluster2 切换到 hacluster 上，说明 booth 自动进行了故障转移。

```
[root@ha1 booth] pcs cluster stop --all
[root@host1 booth]# pcs booth status
TICKETS:
ticket: superticket, leader: 172.17.127.208, expires: 2023-02-20 10:41:21
PEERS:
site          172.17.127.208, last recv: 2023-02-20 10:38:52
              Sent pkts:44 error:0 resends:0
              Recv pkts:44 error:0 authfail:0 invalid:0
```

```

site          172.17.127.209, last recv: 2023-02-20 10:36:05
Sent pkts:53 error:0 resends:0
Recv pkts:53 error:0 authfail:0 invalid:0

```

可以看到此时 hacluster 集群上的 dummy 资源在得到了 superticket 以后成功启动了。



图 3-2 资源成功启动

3.10 手动故障切换测试

Booth 不仅支持对票据的自动管理，同样支持对票据的手动管理。进行手动管理需要将 /etc/booth/booth.conf 的配置文件中 ticket 段下 mode 模式改为“MANUAL”：

```

[root@host1 booth]# cat booth.conf
authfile = /etc/booth/booth.key
site = 172.17.127.208
site = 172.17.127.209
arbitrator = 172.17.127.203
ticket="superticket"
mode = MANUAL

```

将配置修改同步到其他节点上：

```

[root@host1 booth]# pcs booth sync
[root@ha1 booth]# pcs booth pull host1
[root@ha1 booth]# pcs booth sync
[root@ha203 booth]# pcs booth pull host1

```

booth 状态检查：

```

[root@ha1 ~]# pcs booth status
TICKETS:

```

```
ticket: superticket, leader: 172.17.127.208, expires: INF [manual mode]

PEERS:
site          172.17.127.208, last rcv: 2023-02-20 09:50:11
              Sent pkts:3 error:0 resends:0
              Recv pkts:3 error:0 authfail:0 invalid:0

arbitrator    172.17.127.203, last rcv: 2023-02-20 09:48:55
              Sent pkts:1 error:0 resends:0
              Recv pkts:1 error:0 authfail:0 invalid:0
```

可以看到此时 ticket 是被授予在 hacluster 这个集群上。

 注记：这个命令需要在各个集群上 booth-booth-group 运行的节点以及仲裁节点上才能正常运行。需要注意如果 arbitrator 上 last rcv 为 1970 年说明当前集群和仲裁节点之间的通信服务存在问题。

在集群二 ha1 节点上手动给 hacluster2 集群授权票据：

```
[root@ha1 ~]# pcs booth ticket grant superticket
[root@ha1 ~]# pcs booth status
TICKETS:
ticket: superticket, leader: 172.17.127.209, expires: 2023-02-20 09:51:33

PEERS:
site          172.17.127.208, last rcv: 2023-02-20 09:51:20
              Sent pkts:36 error:0 resends:0
              Recv pkts:36 error:0 authfail:0 invalid:0

site          172.17.127.209, last rcv: 2023-02-20 09:51:33
              Sent pkts:4 error:0 resends:0
              Recv pkts:4 error:0 authfail:0 invalid:0
```

可以看到此时票据被授权到 hacluster2 上。

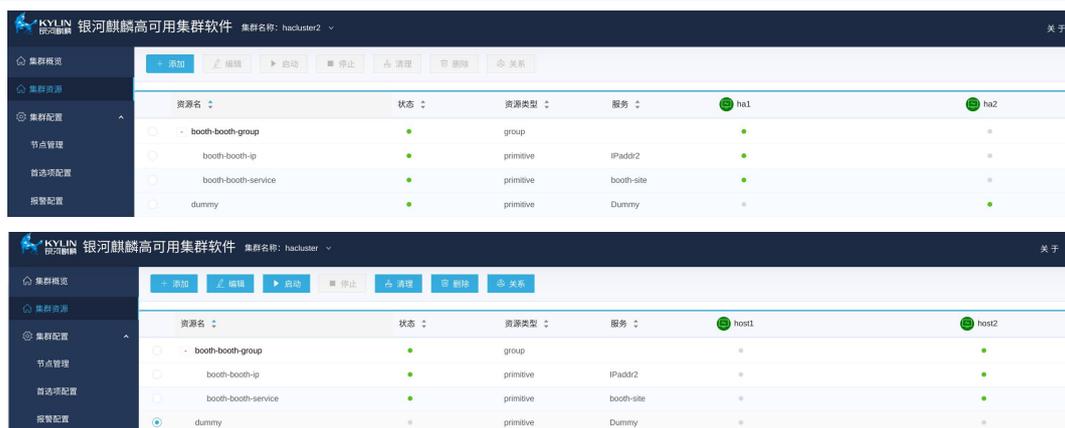


图 3-3 启动资源测试

可以看到 dummy 资源在 hacluster2 上可以启动，但是 dummy 在 hacluster 集群上由于失去票据没办法启动。