

# 基于开源系统的高可用性集群应用研究

吴鑫

(南京审计学院 江苏南京 211815)

**【摘要】:** (本文介绍了集群的定义及内涵) 分析了集群的特点高可用性集群系统的实现原理, 并以提供高可用 Web 服务为例, 给出了开源系统下高可用集群的配置实例。

**【关键词】** 开源系统; 集群; 高可用性

随着硬件价格的逐步下降, PC 服务器已经不是什么高端设备了。而近些年虚拟化的发展, 架设一台服务器已经是很容易的事情。通过组建集群来对关键服务提供高可用性 (High-availability, 简称 HA), 已经是很多园区网中的一种选择。本文在基于对高可用性集群的原理进行分析和研究的基础上, 具体给出配置和实现高可用性集群的实例。

## 1. 集群的定义及内涵

集群是一种计算机系统, 一般通过将两台或多台计算机组合起来完成工作。逻辑上可以把集群看作一台计算机。集群之间一般通过局域网或者其他网络连接起来。一般情况下, 集群的性价比要比单个计算机要高很多。集群的类型一般分为三类。一为高可用性集群, 即当某一节点产生故障时可以在很短的时间内由其他节点接管服务, 保证服务的持续性, 多用于关键服务等场合, 备用节点正常情况下只监听主节点状态, 并不对外提供服务。二为高性能集群, 由多台廉价的普通 PC/服务器构成集群, 实现同等投资下更高的性能产出, 多用于科学计算等领域。三为负载均衡集群, 即将服务的负载分摊到集群中的各个节点, 可以提供比单台服务器更强的服务能力。相对于 HA 集群, 负载均衡集群需要有一或者多台均衡器, 负责处理并分发用户的请求到后端的真实服务器上。

图 1 是一个简单的 HA 集群的示意图

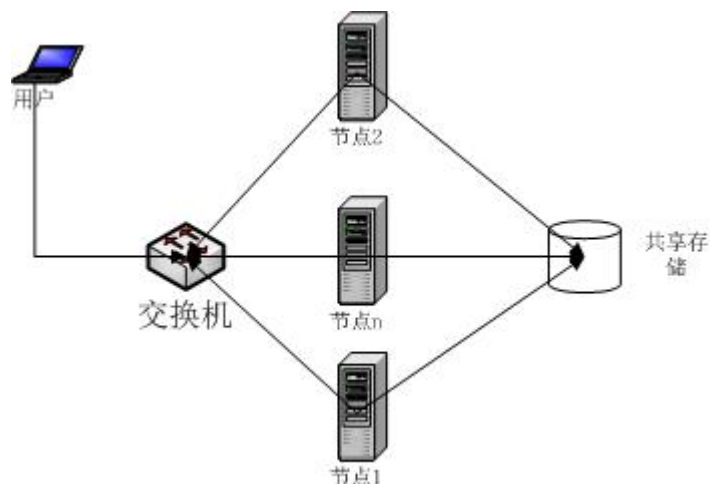


图 1 HA 集群的示意图

从图 1 中我们可以看到，整个系统主要部分为  $n$  台服务器节点 ( $n \geq 2$ )，共享存储以及可供所有节点互联的网络。当集群开始工作时，根据优先级设置，存在一个主节点，此主节点拥有整个集群的 IP 地址，并对外提供服务。集群中其他的节点作为备用节点，通过特定的协议检测主节点的操作系统、服务软件状态，当发现主节点无法提供服务时，备用节点自动接管集群的 IP 地址并对外提供服务。

## 2. HA 集群系统所需要关注以及解决的问题。

- (1) 节点之间服务状态的互相检测（一般称为心跳, Heartbeat），以及安全认证。
  - a) 节点首先应可以互相访问。我们可以通过将所有节点连接到标准 IP 网络，或者通过直连网线、串口线等互联两个节点来实现这一功能。如果通过网络线进行互联，那么服务器必须有两块网卡，一块对外提供服务，一块供心跳检测。。
  - b) 节点运行状态的检测。节点应当具有检测集群中其他节点状态的能力，包括操作系统状态，系统服务状态等，以便当主节点遇到故障无法提供服务时自动切换到备用节点。最简单的检测是系统级别的检测，即备用节点在配置的 Keepalive 时间内通过 TCP/IP 协议或者私有协议 ping 主节点，当发现主节点在超时时间内无法 ping 通时（可能是主节点死机，网络故障等）则认为该节点已无法提供服务，备节点立刻接管集群的虚拟 IP，并对外提供服务。该方式的优点是配置简单，缺点在于，当操作系统或者网络并未崩溃而仅仅是服务软件停止服务时，该种检测机制就无能为力了。因此，更深层次的是对系统服务进行监控。备节点会检测主节点服务的状态，当发现主节点无法提供服务时，启动切换机制，将服务切换到自身，即使此时主节点的网络依旧是畅通的。该种方式的优点是可以达到更高的可用性——这也是 HA 集群的目的，缺点就是配置起来稍显复杂，需要根据提供的不同服务进行不同的配置，另外也不是所有的服务都可以支持这种 HA 方式。同时由于此时可能主节点的操作系统层依旧在正常工作，因此主节点要考虑到集群 IP 的释放问题，这应当由 HA 软件来实现。
  - c) 最后，考虑到安全性，节点之间应当拥有加密、认证等机制。加密是为了保证集群的节点之间通信的安全，防止机密数据被窃取。认证则是避免未经授权的节点加入到集群中，或者未经授权的节点状态信息影响到节点的切换，从而影响到集群的正常运行。最基础的认证方式是在物理层通过串口线或者直连网线连接两台系统，同时在配置文件中关闭其他认证方式。但是对于不具备该种物理环境的情况，我们就

可以通过 IP 网络加上 HA 软件内置的认证机制，如 MD5, SHA1 等来实现。

- (2) 节点状态、IP 和 MAC 地址的管理、切换。
  - a) 每个节点应当有自身的 IP 地址，供状态检测，设备、资源管理等用途，同时整个集群也应当有一个公用 IP 发布给外界提供服务。正常状态下，主节点拥有集群的 IP 并提供服务，在主节点出现故障时，备用节点自动发现故障，并接管虚拟 IP 来提供服务。
  - b) 在数据链路层，集群 IP 可以对应虚拟或者真实的 MAC 地址。虚拟 MAC 即整个集群虚拟出一个 MAC 地址对应集群的 IP 地址，正常状态下主节点负责对外宣告自身拥有该 IP 和 MAC 地址，类似于路由器中常用到的 VRRP（虚拟路由冗余协议），在节点切换过程中，MAC 地址不发生变化，只是由另一台设备宣告拥有该 MAC 地址。另一种方式则是每个节点依旧拥有自身的 MAC 地址，由当前工作的节点负责宣告自身为该 IP 的拥有者，即工作节点上有两个 IP 对应同一 MAC，节点切换时集群 IP 的对应虚拟 MAC 的所有者由主节点变为备节点。
- (3) 节点文件资源的同步。
  - a) 无论什么服务，都必须有相应的文件资源。我们必须保证每个节点都能访问到同样的文件资源，即文件资源的同步。我们可以采取独立存储的方式，即分别拷贝 n 份相同文件到集群的 n 个节点上，但这种方式问题在于当我们需要修改服务内容（如修改网站的 html 文件）时，带来的工作量也是原来的 n 倍。对于数据库实时性要求很高的等应用，通过这种方式是完全不可能实现的。
  - b) 因此我们可以采取共享存储的方式来实现文件的同步。可以是 NFS 网络存储，也可以是直连磁盘阵列，或者 iSCSI, FC SAN 等存储网络。多台节点都可以访问一个集中的存储设备，如需修改文件，只需通过一个节点修改即可。我们只需要配置好每个节点对于共享存储的访问权限即可。

### 3. HA 集群的安装配置实例

目前市场上的高可用集群产品除了各大厂商如 Redhat, Microsoft, Symantec 等的商业产品外，使用比较广泛的是 Linux 下的开源产品，如 Linux Heartbeat

(<http://www.linux-ha.org/>)。下面以它为例研究如何实现高可用性集群的操作，并配合 Apache Web Server，对外提供 Web 服务。

虽然并非必须，但是为了避免不必要的问题，建议服务器采用相同的硬件以及操作系统、软件版本，也可以事先在虚拟机上进行测试，通过之后再到真实环境下部署。

Linux 系统选择的是 CentOS 5.3，这是一款基于 Redhat Enterprise Linux 的源代码再编译版本。在 CentOS 下安装 Linux-HA，只需在联网状态下运行 `#yum install heartbeat` 即可，如果有缺失的组件也会自动进行安装。当然另一种方式是编译安装，这种方式的定制性更

强。

Heartbeat 的启动、停止命令为`#/etc/rc.d/init.d/heartbeat start/stop`，在首次启动之前，应当配置好配置文件。

Heartbeat 的配置文件包括 `ha.cf`，`resource` 以及 `authkeys`，分别是主配置文件，资源文件以及认证文件。文件的位置在 `/etc/ha.d/` 下。`ha.cf` 中，大部分配置默认即可，需要关注的是以下几个地方

```
keepalive 2 /检测状态的间隔时间/  
deadtime 30 /认为对方节点失去响应的时间/  
warntime 10 /警告时间/  
  
bcast eth1 /用于节点通信的网卡/  
ucast eth1 192.168.100.1 /用于节点通信的 IP 地址/  
node linux-ha-a /两个节点的 Linux 主机名/  
node linux-ha-b
```

除了通信的 IP 地址不同，两台节点需设定相同的参数。

认证文件 `authkeys`，Heartbeat 支持 3 种认证方式，安全性是递增的，如最基本的 CRC 校验只提供数据传输时错误的检测，建议只用在可靠的物理网络下。我们这里采用 MD5 认证，

```
auth 1  
1 md5 thisisakey
```

`Resource` 文件的配置，

```
linux-ha-a 10.1.1.1 httpd::start
```

相应含义，分别是当前主机名，集群的虚拟 IP 地址，对应需要集群的服务，服务的启动参数。

HA 安装好之后，我们继续安装 Apache，在 Cent OS 下运行 `#yum install httpd` 即可安装。随后通过 `#service httpd start` 即可启动。为了检测 HA 的工作状态，我们可以分别在两个节点上放置不同的 html 主文件以标识对应的节点（如“this is node a”），并分别通过 http 访问两台节点查看 http 服务是否正常启动。

随后我们启动 HA，如果一切正常，我们可以通过浏览器访问集群的 IP 看到网页，目前我们可以看到的就是主节点。

开始检测 HA 的运行状况，不断 ping 主节点 IP，接下来断开主节点的网线，或者停止服务，我们可以看到会丢失一些 ping 包，网页暂时无法打开。随后 ping 恢复正常，网页也可以打开，但是此时显示生效的服务器是节点 B。

通过以上配置，我们已经实现了一个能够提供 Web 服务的高可用集群。

#### 4. 结束语

在介绍了高性能集群的配置实例后，我们还可以扩展一下思路，比如实验中我们是两台节点运行一个高可用集群，如果要求比较高，我们甚至可以将两个 HA 集群组合成一个更高

等级的 HA 集群，这样可以达到更高的可用性。前面我们提到过负载均衡集群也可以通过将一个 HA 集群看作负载均衡集群中的一个普通节点来实现高可用+负载均衡的集群形式。总之，集群的组建方式有多种，我们应当根据实际的需求以及条件来构建。

#### 参考文献

- [1]、 利用 Heartbeat 实现高可用集群系统，胡小勇；余翔湛；张宏莉，微计算机信息, 2009
- [2]、 一种高可靠、高可用集群计算平台的设计与实现，刘天田；龚念；杨升春，舰船电子工程, 2008