

AIXCHINA

HACMP for AIX 学习笔记

www.aixchina.com

AIX 中国论坛发表的所有文章版权均属相关权利人所有，受《中华人民共和国著作权法》及其它相关法律的保护。

如出于商业目的使用本资料或有牵涉版权的问题请速与论坛管理员联系。管理员电子邮件：aixchina@21cn.com

Aix HACMP

IBM Certification Study Guide Test 167

Chapter 9 和 SP 有关.....	4
9.1 HACWS	4
9.1.1 硬件.....	4
9.1.2 软件.....	4
9.1.3 配置备份 cws.....	4
9.1.4 安装软件.....	4
9.1.5 HACWS 配置.....	4
9.1.6 设置和测试 HACWS.....	5
9.2 Kerberos 安全	5
9.2.1 在 HA4.3 上配置 kerberos	5
9.3 VSDs-RVSDs	6
9.3.1 Virtual Shared Disk	6
9.3.2 Recoverable Virtual Shared Disk	6
9.4 SP switch 做 HACMP 网络	6
9.4.1 switch 基本内容.....	6
9.4.2 Eprimary 管理.....	7
9.4.3 Switch 失败.....	7
Chapter 10 HA Classic vs HA/ES vs HANFS.....	8
10.1 HACMP for AIX classic.....	8
10.2 HACMP for AIX /Enhanced Scalability.....	8
10.3 HA for Network File System.....	8
10.4 相同和不同	8
10.5 如何选择.....	8

Chapter 9 和 SP 有关

9.1 HACWS

重要作用：壹台 CWS 故障 SP 仍可工作；
CWS 做维护，SP 不受影响，方便备份。

9.1.1 硬件

- (1)cws spdata fs 要在 shared disk 上，两台 cws 都可访问到。
- (2)SP 有多个 frame，应给 cws 配 8-port 卡，已保证串口够用。
- (3)cws 到 frame 的 supervisor 卡是串行直连线，主备 cws 连同一 frame 的 tty 口应相同，如：都用 tty3 来和 frame A。(节点 supervisor 卡连到菊花链串行线的哪个接口决定其是 No1 还是 No3，而该菊花链只连到电源的 J8(奇数节点)或 J7(偶节点和 switch)，用菊花链上标有 PDU 的接口。菊花链不连 cws 和 frame 的 supervisor)
frameA 上要有 Y-cable (串行口)
- (4)cws 只要一块网卡/每 cws

9.1.2 软件

AIX、PSSP、HA 都要用相同版本查手册，如 AIX4.3.2 PSSP3.1HA4.3

9.1.3 配置备份 cws

主 CWS 象无 HACWS 时一样配置。
备 CWS 的 kerberos 配置 如主 cws 是 authentication server 则备 cws 配成 secondary authentication server。如主 cws 是 authentication client of some other server，则备 cws 也是 client。

9.1.4 安装软件

除 HA 外，还有 ssp.hacws

9.1.5 HACWS 配置

- (1)用/usr/sbin/hacws/spw-apps -d 停下 spmgr、splagd、hardmon、sysctld、supfilesrv、sp_sonfigd 等守护进程

(2)配心跳线，网卡要用 boot 地址启动，/spdata fs 共享盘
(3)定义 cluster ID，NAME，node，卡应有 boot 和 service 地址，同步
(4)ssp.hacws 带了一些 script，可做为 HA 的应用服务器
建议资源组定义如下：
资源组名：hacws_group1
Node Relationship：rotating
participating Node Name：[主节点名，备节点名]
Service IP label：至少主 cws 主机名
Filesystem：/spdata
VG：包含/spdata 的 vg
Appserver：

9.1.6 设置和测试 HACWS

先让主备 CWS 可按其 hostname 寻址，再在主 CWS 上运行：
/usr/sbin/hacws/install_hacws -p 主 cws 名 -b 备 cws 名 -s 设 cws 节点名
定义事件脚本：/usr/sbin/hacws/spcw_addevents
校验：/usr/sbin/hacws/hacws_verify
检查线缆：/usr/sbin/hacws/spcw_verify_cabling
依次重启主备 cws，依次启动 HA，smit clstart
检查 hacmp.out
grep “SPCW_APPS COMPLETE” /tmp/hacmp.out

9.2 Kerberos 安全

基于 DES 算法加密
Identification 实体告诉别人自己是谁
Authentication 认证某个实体
Authorization 检查一个认证的实体是否有授权
Client 和 Server 的通信分三步
Get a ticket-granting ticket(有生命期) 用到 pwd，避免以后再用 pwd
Get a service ticket
Get the work done on the service provider

9.2.1 在 HA4.3 上配置 kerberos

HA4.3 可用 cl_setup_kerberos 定义 kerberos principals 给群集中所有的 IP labels，这样远程的 kerberized 命令可以工作。

SP 上用 setup_authent 做 SP 相关的 kerberos 设置，IP labels 来自于 SDR，因 SDR 不允许一个接口多个 Iplabels，对有 IPAT 的情况，HA 的 kerberos 设置要做 2 次，或用 setup_authent 直接设，或用 setup-server 间接调用 setup_authent。

手工也可设，用 kadmin。

必须的 principal 有 rcmd , godm2 个 , 格式 godm.had1_stby@IBM.com。

将 principals 加入 kerberos DB 后 , 还要将它们从 DB 中解出 , 加到各节点的 /etc/krb_srvtab 文件中。现在可以在 root 的。Klogin 和/etc/krb.realms 中用新的 principals。

9.3 VSDs-RVSDs

9.3.1 Virtual Shared Disk

在 pssp 的 ssp.csd.vsd 文件集中。

VSD 只支持裸 lv , 不支持 fs。VSD 没有上锁机制 , 要通过 Oracle Parallel Server 来完成。一个节点可以自身没有 HD , 没有 VSD 进程 , 但仍可成为其它节点的 VSD 的 client。VSD 可以使用各种 IP 网络 , 出于性能考虑 , 建议用 Switch 网络。VSD 对访问本地 HD 的影响很小。

VSD 结构图

三种数据来源

VSD 状态图

buddy buffer

略

9.3.2 Recoverable Virtual Shared Disk

物理上提供对 disk 的 2 个以上节点的连接。

RVSD 功能图

RVSD 相关守护进程功能

略

9.4 SP switch 做 HACMP 网络

早期 HPS 或 HiPS 又称为 TB2switch HACMP/ES V4.3 PSSP3.1 以后不再支持。有缺点 , 一个节点的变化影响整个网络其它节点。

SP switch 又称为 TB3 switch , 克服上述缺点。

9.4.1 switch 基本内容

switch 网络是点到点网络 , 配成 private 网络。要支持 IPAT , 就要用 ARP。用 IPAT , 还要用 ifconfig alias 将 service 和 boot 地址赋于 CSS0 卡。SDR 中有 switch 卡的 base IP 地址 , 不要用于 HA 网络 , 会引起 PSSP 混乱。所有 SP switch 网络地址的掩码和 CSS0 base IP 地址的掩码一致。

9.4.2 Eprimary 管理

Eprimary 节点管理 SPS 的交通路由事件，有 backup 节点。

HACMP/ES4.3 以前(不含 4.3)有 primary 节点管理功能，如果升级到 V4.3，switch 也升为 spswitch，这些管理已经内置，HA 不再管理。

用 `odmget -q 'name=EPRIMARY' HACMPSP2` 看是否有管理 Eprimary。

升级前用 `/usr/es/sbin/cluster/events/utlis/cl_HPS_Eprimary unmanage` 将其置为不管理。

9.4.3 Switch 失败

Switch 卡是单点失败，应认为节点失败。参考 2.6.2.1 的定义，关掉节点，让接管发生。如该节点是 Eprimary 节点，且是 Spswitch，SP 软件会选一个新的 Eprimary 节点，不须要 HA 软件的介入。

Chapter 10 HA Classic vs HA/ES vs HANFS

10.1 HACMP for AIX classic

群集管理通过 NIM(网络接口模块)管理节点卡和网络状态。

10.2 HACMP for AIX /Enhanced Scalability

HACMP/ES 不再通过 NIM，而是通过一个原来为 SP 开发的技术，现在从 PSSP 中分离出来的 RSCT 技术来实现。该包原来叫 ssp.ha 文件集，现在集成在 HACMP/ES 中，RISC System Cluster Techonlogy。

RSCT 包含以下内容（为什么用 HACMP/ES 的原因？）：

- (1)Event Manager；
- (2)Group Service；
- (3)Topology Service；

(4)安全性上，可选择标准的，用/.rhosts 来满足同步要求；增强的，用 kerberos 来认证同步时的远程命令。

10.3 HA for Network File System

只支持 cascading 资源组，HANFS 只支持 2 个节点，基于 Classic HA 结构，在 NFS 功能上能处理双重请求和保证单机失败接管时锁状态的正确性。HANFS 和 HACMP Classic 不能混用于一个 cluster 中，但可在同一物理网络上。

10.4 相同和不同

结构上是相同的，都需要 Cluster Manager 控制节点，追踪状态，激发事件，不同之处在于追踪状态的技术不同。HACMP/ES V4.3 可支持 SP 和非 SP 机器，但如用在 SP 上，要求 PSSP3.1，而 PSSP3.1 又要求 Spswitch，当然 SP 也可用传统的 HACMP，这样对 PSSP 就没有要求。

10.5 如何选择

如果是 HPS Switch，只能用不高于 PSSP2.4，因此只能用 HACMP classic 或 HACMP/ES V4.2.2 及以下。无 switch 或新的 spswitch 可选功能更强的 HA。用于 NFS，则用 HANFS。