医疗信息系统基础设施架构①

张林萍

(复旦大学附属金山医院 信息科，上海 200540)

摘 要：医疗信息系统持续运行依赖于高可用的信息系统基础架构，传统的双机热备技术已不能满足当前的实际需求。从共享磁盘管理角度论述故障转移型群集新技术方案，从安全角度论述局域网高可用和业务流负载均衡实施技术，操作系统安全措施及终端管理。打造一个真实高可用的信息系统基础平台，清除计划内和计划外停机。保障医疗机构的日常业务，应用软件系统可持续运行，维持正常的就医秩序和医疗服务。

关键词：持续可用性；群集；动态卷；负载均衡；终端准入

Medical Information System Infrastructure

ZHANG Lin-Ping

(Shanghai Jinshan hospital, Shanghai 200540, China)

Abstract: Medical information system has been running rely on highly available information system infrastructure construction, the traditional hot standby technology can not meet the current actual demand. This article discusses the management point of view from the shared disk type cluster failover technology program, discusses the LAN from a safety point of view and business flow load balancing, high availability implementation of technology, operating system security and terminal management. Build a real high available information systems infrastructure platform, clear the planned and unplanned outages. Medical institutions in the ordinary course of business protection, application software systems for sustainable operation, to maintain the normal order and medical services for treatment.

Keywords: continuous availability; clustering; dynamic volume; load balancing; terminal access

随着计算机科学与技术的发展，医院信息系统 HIS 从单机系统到网络共享，再到数据字化医院。医院日常医疗服务业务对 HIS 应用软件系统依赖越来越强，一旦系统宕机，若不能及时修复，将影响日常的医疗服务，引发医患矛盾，造成医疗事故。HIS 持续稳定的运行依赖于安全稳定性的高可用信息系统基础架构。本文论述建设安全稳定性的高可用信息系统基础架构的三个方面技术：第一，杜绝服务器和存储群集系统各硬件部件和软件组件单点故障，弥补传统的群集技术方案不足；第二，规划部署医疗信息网络高可用性和业务流负载均衡，结合实际应用实施配套安全策略；第三，设置操作系统安全措施，访问权限控制，终端准入及资产管理。

1 弥补传统的群集技术方案不足

传统的故障转移型群集有带共享存储的群集和不带共享存储的群集两种类型，它们各存在一些不足之处。如图 1 所示：

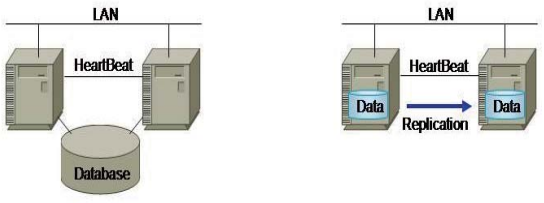

图 1 带共享存储的双机热备(左图)和不带共享存储的复制型群集(右图)

① 收稿时间:2010-04-27;收到修改稿时间:2010-05-29

1.1 典型的带共享存储的双机热备存在不足

它只能避免某台服务器硬件或软件故障所引起的系统宕机。一般情况，数据库设备文件是存储在共享存储中的，一旦共享存储发生意外故障，比如：RAID 组损坏，存储控制器微码 bug，存储背板损坏等等，都将导致群集服务组中的数据库资源无法联机；更为糟糕的情况还有，例如，MSCS 仲裁盘也在此共享存储上，整个群集都将崩溃。由此可见，传统的故障转移型群集存在共享存储单点故障，此故障发生，系统将中止运行，而且存储也不便于计划内停机维护。

1.2 不带共享存储复制型群集存在不足

不带共享存储复制型群集的原理是将主节点需要保护的分区或文件夹，通过网络复制到备节点，主备节点安装相同数据库服务或应用服务，备节点日常是不工作的，通常为主备模式 Active/Standby。当主点发生故障，群集服务组中的所有资源切换到备节点运行，主节点故障排除后，若再切回主节点，需要等待保护区从备节点到主节点反相完全同步完成，不仅等待时间较长，而且同步时争用正常数据 I/O 访问，业务应用软件系统变的迟缓。

主机本身所能安装的磁盘是有限的几个，不便于容量和 I/O 性能的扩展，随着数据量的逐渐增长，存在空间限制和 I/O 性能瓶颈。

1.3 高可用群集案例

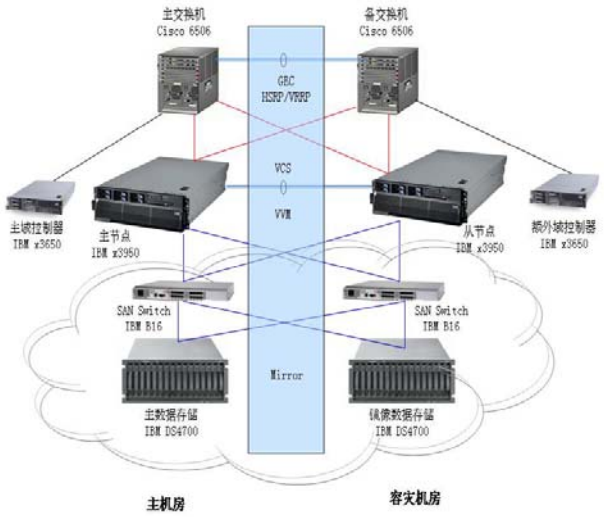

图 2　高可用的群集方案

作者在多年的医疗行业系统集成经验中，学习和研究了一种可行的高可用群集方案，而且已经在多家医院成功实施和应用。用此案例来探讨如何弥补传统的群集技术方案不足之处，以及群集新技术应用，先看方案图，如图 2 所示。

此方案图从下往上看，最底层为两台数据存储，其上一层为两台存储交换机，在一起构成一个简单而健壮的存储区域网 SAN(云图部分)；SAN 的上一层是群集主机节点，以及其他功能主机；最上层为局域网络核心交换机；整体上看分成左右两边，左边为主机房，右边为容灾机房，每层设备相互对应及热备。

此案例不同于传统群集技术的焦点在共享存储磁盘管理上，主要目的在于解决共享存储单点故障。实现关键步骤和技术如下：

第一，最低层的两台共享数据存储同样创建磁盘冗余阵列 RAID 组，由于 HIS 应用软件近似于 OLTP，随机性数据访问平凡，常选用 RAID 1+0 来提高 I/O 存取速度。例如，两台存储分别配置了 30 块，单盘容量为 146GB，速度为 4Gbps，转速为 15K，接口为 FC 的硬盘，创建 RAID 1+0 后可用容量为 146×(30÷2)≈2TB。IP SAN 不适合于数据库存储，因其使用的是 iSCSI 协议，IP SAN 比 FC SAN 稳定性较差。

第二，共享存储化分逻辑盘，Lun 的个数与大小跟据业务实际需求而定。比如，某医院应用业务数据库有 HIS、LIS、RIS、PACS，而且 LIS、RIS、PACS 业务数据库需要读写 HIS 数据库中的数据（如病人基本信息，确费操作等），所以还需要数据分布式处理 DTC 应用服务，或者是 Weblogic 等中间件服务。例如可用总容量 2TB， DTC 只是一个服务，仅一些日志数据存放在逻辑盘上，划分 10GB 空间即可；HIS 数据库相比较而言增长较快，大约划分整个容量三分之一 700GB；LIS、RIS、PACS 数据库增长不是太快，各划分 300GB；还剩余近 400GB 便于以后用于其他业务。为了让主机区分两个存储中的 Lun，便于今后维护，可将主机房内的存储每个 Lun 多化出 1GB。PACS 影像不能与数据库共用存储，防止存储大的图像时造成数据访问时延抖动，图像可以存放在性价比较高的 SATA 接口的 NAS 里。

第三，配置存储时，还要设置 Partition Mapping 群集类型和操作系统（Microsoft Cluster Service、Veritas Cluster Service 即 Veritas Storage Foundation and High

Availability Solutions、IBM AIX High Availability Cluster Multi-Processing、HP-UX ServiceGuard、Sun Solaris Cluster、Oracle Real Application Clusters），以及其所对应的 HBA 卡 WWN 号及主机名或 IP。

第四，在存储性能优化设置方面，尽量做到条带大小除磁盘个数等于文件系统分配单元大小，等于数据库分配单元（页）大小。

第五，存储交换机需要化分 Zone，目的在于隔离每个存储控制器到存储交换机 FC 信号，防止相互干扰。此案例中只要将每台存储交换机连接存储的两个接口分别与两台主机的 HBA 卡接口化分在一个 Zone 中即可。

第六，为了让主机 HBA 卡更好的工作，可指定拓扑结构为 Point to Point，指定传输速率为 4Gbps。

第七，主机上存储多路径软件可以使用存储厂商提供的，例如：IBM 的 RDAC/MPIO，HP 的 Secure Path，HDS 的 HDML，EMC 的 Power Path。如果是不同类型存储可以使用第三方的多路径软件，例如 Veritas 的 DMP。

第八，主机端磁盘管理尤为重要，为了使两台存储中对应化分的业务逻辑盘实现同时读写数据，即 RAID 1 镜像功能，先需要创建动态磁盘组。本案例使用 Veritas Volume Management，当然根据实际需要也可以采用其他的卷管理工具，如比 LVM、ASM 中的镜像。假设主机识别主机房存储 DTC、HIS、LIS、RIS 对应业务 Lun1、Lun2、Lun3、Lun4 分别为 Harddisk1、Harddisk2、Harddisk3、Harddisk4，主机识别容灾机房存储 DTC、HIS、LIS、RIS 对应业务 Lun5、Lun6、Lun7、Lun8 分别为 Harddisk5、Harddisk6、Harddisk7、Harddisk8，可将 Harddisk1 和 Harddisk5 划到一个 Dynamic Disk Group，作为 DTC_VMDG，可将 Harddisk2 和 Harddisk6 划到一个 Dynamic Disk Group，作为 HIS_VMDG，可将 Harddisk3 和 Harddisk7 划到一个 Dynamic Disk Group，作为 LIS_VMDG，可将 Harddisk4 和 Harddisk8 划到一个 Dynamic Disk Group，作为 PACS_VMDG。接下来再创建动群集镜像卷，本案例分别为 DTC_LV、HIS_LV、LIS_LV、RIS_LV、PACS_LV。

本案例中为什么不把存储的镜像功能交给 SAN 本身的 Mirror 来完成呢？存储透明的提供给主机镜像卷不是更好吗？原因是，如果使用不同厂家存储镜像几乎不可能实现；即使为相同品牌，但不同型号或批次的存储，微码版本很可能不一致，实现镜像也很难。另外，通过主机来实现镜像容易管理维护，可操作性较强，其对主机的资源开销较小，对存储本身没有太多要求，比较灵活，也便于今后在线升级扩展。

第九，群集软件最好选用不带仲裁机制软件，否则仲裁盘或其数据、日志损坏可能导致所有群集服务组中的资源脱机。

HA 不能完全理想化的防止停机，其故障转移切换要中断片刻，面向数据连接的 C/S 应用程序需要重新建立连接。HA 软件本身也会发生错误，比如 HA 程序 Bug 等，可能导致整个群集瘫痪。但像 VCS 这类群集软件，群集服务即使不启动，也可以强行启动数据库服务。因使用 VVM 管理动态磁盘组，共享盘不会发生“Split Brain 脑裂”现象。

第十，最后将相应的业务动态磁盘组、群集镜像卷、虚拟 IP、虚拟服务器名、数据库实例等群集资源配置到其对应的服务组中，例如：DTC_SG、HIS_SG、LIS_SG、RIS_SG、PACS_SG。每个服务组所有资源可以自由切换到不同的主机节点上运行，例如：主节点运 DTC 服务和 HIS 数据库实例，从节点运行 LIS、RIS、PACS 数据库实例，以致达到主从式双机互备，业务负载均衡，不同于传统的主备式双机热备。

这样的一种群集架构，可以防止任一单点故障。如果是一台存储需要关机维护，或者损坏，其群集服务功能不受影响，只是速度有所下降，因为可同时读操作的盘减少了，当该存储修复好之后，只需要 Reactive 或 Replace 磁盘，镜像卷会自动生成。如果是一个控制器损坏，一台存储交换机损坏，一个 SFP 模块损坏，一个 FC 链路断开，一块 HBA 卡损坏，也不影响 I/O 读写，因为有冗余路径和多路径软件支持。如果是某台主机节点发生故障宕机，其上所有群集服务组资源会转移到优先级较高的主机节点上运行。即使是整个主机房发生意外，如火灾、电源故障等，然而容灾机房群集服务器能接管所有群集服务组资源，保障业务应用软件系统可持续运行。

主机房与容灾机房之间用光缆连接，如果距离在 500m 以内，可选用 50um 波长的多模光缆；如果距离大于 500m，在几千米或几十千米之间，可选用单模光缆，相应设备或模块也要使用单模类型。

群集主机节点双网卡绑定，做负载均衡或故障转

移。分别与两台局网核心交换机相连接，下一章详细论述网络高可用性。

2　医院网络高可用性

对医院信息系统而言，只是服务器高可用还是不够的，如果局域网络出现故障，也将导致信息无法传递，业务应用软件系统因此中断。由此可见局域网高可用性也很重要。

2.1 典型的高可用医院局域网

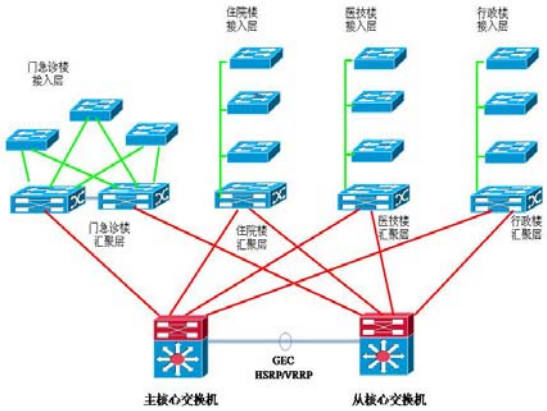

图 3　医院局域网拓扑图

首先，两个核心交换使用 VRRP 配置为主从式双机互备，启用 MSTP 多根域，并对应设置 VLAN 接口优先级和 Spanning-tree 优先级，达到业务流负载均衡。例如：主核心交换机处理转发门急诊楼 VLAN10 和住院楼 VLAN20 的业务流；从核心交换机处理转发医技楼 VLAN30 和行政楼 VLAN40 的业务流。

其次，门急诊区域汇聚层交换机冗余部署，从而保证门急诊关键窗口业务应用软件不因一台汇聚交换机损坏而中断。住院、医技、行政通常为单汇聚即可。

最后，接入层交换机用于连接 PC 工作站，应设置安全措施，比如：BPDU 保护或快速收敛，有效防止 Dos 攻击；启用 dot1x 协议，对接入 PC 安全准入认证；MAC+IP+端口绑定，即安全又能防止 ARP 病毒；针对每个接口设置增强访问控制列表 ACL，限制其出或入 IP、TCP、UDP 及端口，禁止 PC 之间或 PC 与服务器之间不必要的通信，还能有效阻断网络病毒传播路径。

另外，汇聚层和接入层交换机需要冷备一到两台，把所有交换机配置上传到 TFTP 服务器，当有交换机损坏时，可及时灌入配置并更换。

2.2 安全可靠的医院广域网

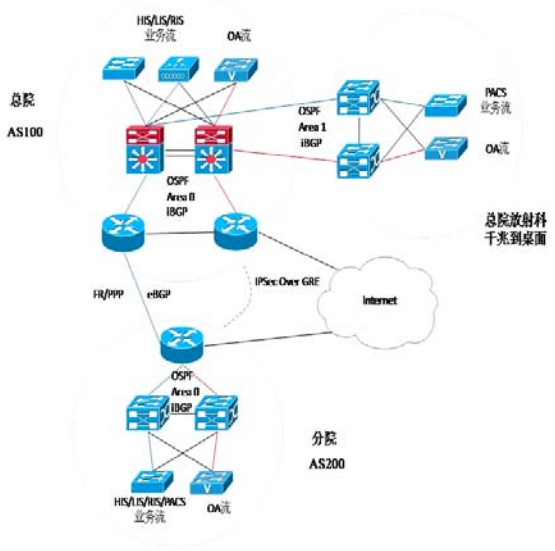

图 4　医院广域网拓扑图

总院区域的 IGP 路由协议为 OSPF，划分 HIS/LIS/RIS 数据与 PACS 图像 2 个区域 Area 0 和 Area 1，数据与图像区域之间三层核心交换使用 IBGP 路由。

分院数据流根据策略路由、路由优先级、前缀列表等进行选路，业务流使用专线，OA 流使用 VPN (IPSec Over GRE)，两者互为备份。

总院与分院骨干区域 Area 0 用 MP 进行捆绑和 CHAP 验证。

总院与分院之间通过 EBGP 来进行路由更新，在路由通告时，进行 BGP 报文安全验证。

为了确保总院和分院网络安全，业务禁止访问 Internet，OA 允许访问 Internet。业务与 OA 工作站也不能相互访问。

另外，为了更好的利用网络资源，建创未来的数字化医院，可将无线、语音和视频等统一纳入到此 IP 网络中。

3　操作系统安全措施及终端管理

3.1 操作系统安全措施

群集主机节点操作系统本身也需要定期备份，防

止操作系统更新补丁或病毒破坏系统文件等引发意外宕机。

加强用户密码管理：密码必须符合复杂性要求，最短密码长度建议为 6 位以上，密码最长使用期限建议为 60 天。防止他人猜测操作系统管理员密码。

帐户安全管理：锁定阈值建议设置登录 3 次失败锁定该帐户，帐户锁定时间建议设为 30 分钟，在此后复位帐户锁定计数器建议设为 30 分钟。防止他人利用黑客软件暴力破解操作系统管理员密码。

共享应用程序或文件权限控制：防止越权使用应用程序，比如管理工具等。

禁止不必要的服务和共享资源：防止安全隐患。客户机与服务器管理员不要设置相同密码：防止友好验证。

3.2 工作站安全管理

终端准入 EAD：所有经过身份认证的 MAC 地址并且安装了客户端软件的机器可以接入网络，防止一些非法的或者外来的新机器接入到内网，只有在经过管理人员的许可之后，才可以访问内网资源，降低外来非法机器导致的内网安全风险。

USB 存储和网络设备管理：防止内网工作站使用外界程序，传入病毒，破坏内网数据。

资产管理：记录工作站计算机硬件和软件配置，防止故意更换硬盘（外带程序），安装非法软件等。

进程管理：查看和管理各客户端当前和曾经运行的进程，终止非法进程，对指定的程序外的程序进行禁止，能够防止客户端通过修改文件名和目录的方式运行非法程序。

事件报警：管理员可以设置报警的规则，比如工作站运行某个特定程序时报警。

审计管理：管理员可以统计查看到终端 PC 用户登录事件、关机事件、打开窗口事件、操作文档事件等一系列的操作事件日志，通过这些日志，管理员可以详细了解用户的操作历史。

为了方便日常工作，桌面管理系统还应用包括：远程协助、软件分发（打补丁）等功能。

4 结论

为了保障医院信息系统 365*24 持续运行，需要一个高可用的 IT 基础平台，要求此基础平台主机、存储、网络、操作系统等，由于计划内停机维护、升级扩展，或者由于意外故障都不影响 OLTP 业务持续运行。

(1) 本文论证了医疗信息系统服务器和存储高可用群集新技术方案，此 HA 方案，不仅能解决主机本身故障，还能解决共享存储故障。

(2) 本文还研究了医疗信息系统网络高可用性，以及业务流负载均衡的规划和部署，并施行网络安全策略，保护通信连续性。

(3) 本文最后阐述操作系统和终端安全管理措施，作为以上二点的补充，三个方面一起构成了一个持续可用性的医疗信息系统基础设施架构。

参考文献

1 何萍,索仲良.论如何构建医院信息系统的安全体系.计算机应用与软件,2007,24(10):202－204.