双机高可用系统设计与性能分析

韩德志[1,2,4] 傅 丰[3]

(广东外语外贸大学信息学院 广州 510420)[1] (中山大学广东省信息安全重点实验室 广州 510275)[2]

(黄淮学院计算机系 驻马店 463000)[3] (中山大学电子与通信工程系 广州 510275)[4]

摘 要 针对融合iSCSI,NAS,SAN的海量网络存储系统的特点,设计了一种双机高可用元数据服务器系统,该系统不仅减少了元数据服务器瓶颈,而且可充分保证存储网络系统元数据的高可用性。通过建立连续时间马尔可夫链性能分析模型,分析结果显示双机高可用系统的可用度远优于单机单路径系统。

关键词 高可用,存储网络,元数据服务器,马尔可夫模型

Design and Performance Analysis of Dual-machine High Availability System

HAN De-zhi[1,2,4] FU Feng[3]

(School of Information, Guangdong University of Foreign Studies, Guangzhou 510420, China)[1]

(Guangdong Key Laboratory of Information Security Technology, Sun Yat-sen University, Guangzhou, 510275, China)[2]

(Dept. of Computer Science, Huanghuai University, Zhumadian 463000, China)[3]

(Dept. of Electronics and Communication Engineering, Sun Yat-sen University, Guangzhou 510275, China)[4]

Abstract According to the characteristic of mass network storage system for merging iSCSI, NAS, SAN, the paper designed a dual-machine high availability system for metadata server, which not only reduces the metadata server bottleneck, but also does guarantee high availability for storage network metadata. And as shown in the performance analysis on the basis of the markovian chain model, the availability of the above dual-machine high availability system is much higher than that of a single machine system with a single path.

Keywords High availability, Storage network, Metadata server, Markovian chain model

1 引言

目前,流行的网络存储系统主要有两种:附网存储(NAS)和存储区域网(SAN)。NAS和SAN有很多优点,但也存在一些不足[1,2]。如NAS存在扩展性、高可用性,以及多个NAS管理性方面的缺陷;SAN存在构建和管理成本高、不同厂家的设备很难互操作等缺陷。针对NAS和SAN存在的缺陷,我们提出并实现了一种在IP协议下融合iSCSI,NAS,SAN的统一存储网络(简称USN)[3]。在USN中,NAS设备、iSCSI设备和SAN设备并存,用户可以以块I/O的方式访问USN中的iSCSI设备和SAN存储设备,也可以以文件I/O方式访问USN中的NAS存储设备和SAN存储设备,整个USN是一个统一的存储池。并且,USN能同时提供服务器通道和附网高速通道,向客户机提供数据,减少服务器瓶颈,提高系统的I/O速度。USN既有NAS的优点(低成本、开放性、文件共享),又有SAN的优点(高性能、高扩展性)。保证USN系统高可用的关键是保证元数据服务器的高可用性。本文针对USN系统的特点,通过设计双机高可用系统保证USN系统元数据信息的高可用性,其结构包括相互冗余的双服务器和双网络。双服务器同时工作,并相互监控对方的工作状态。当一个服务器失效时,另一服务器接替其全部业务,在减少元数据服务器瓶颈的同时,可充分保证整个系统工作的连续性。通过连续时间马尔可夫链模型对双机高可用系统的性能进行了分析,结果显示该系统在可用性方面明显优于单机系统。

2 双机高可用系统的实现方案

USN高可用系统采用"双侦测网-双端口磁盘阵列"的双机双工高可用方案。虽然双端口磁盘阵列的方案比"完全磁盘镜像"和"分布式磁盘镜像"成本要高,但是双端口磁盘阵列提供了方便的数据共享机制,可以直接支持单机应用程序的高可用性。同时,我们采用双网"心跳"侦测,避免了单网"心跳"中网络一会儿中断一会儿又恢复而引起结点误判现象。

2.1 系统的总体结构

USN双机高可用系统采用"双机-双端口磁盘阵列"的实现方案,即有两个HA节点各自运行各自的应用程序。两个HA节点的应用程序及其数据都放在共享的磁盘阵列中,同时两个HA节点之间通过Heartbeat线交换控制消息。客户

到稿日期:2008-02-22 本文受国家高技术研究发展计划(863)(批准号:2007AA01Z449)资助的课题和国家自然科学基金项目(60673191),广东省信息安全重点实验室项目,河南省科技计划资助项目(072100451230),广东外语外贸大学科研创新团队项目(GW2006-AT-005)资助。

韩德志(1966-),教授,博士后,研究方向为高可用、高扩展的网络存储系统,E-mail:han_dezhi88@tom.com;**傅 丰**(1969-),女,副教授,学士,主要研究方向为网络存储技术。

可通过双网络路径访问服务器 A 或服务器 B，即双交换机互为冗余。

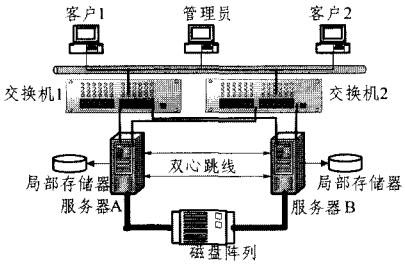

图 1　双机高可用系统的硬件结构

HA 节点间相互协作保证应用程序运行的连续性和可靠性，而应用程序数据的存储可靠性是通过磁盘阵列保证的，双端口 RAID 支持 RAID1，RAID3 和 RAID5。

双机高可用系统的硬件结构如图 1 所示，服务器 A 和 B 分别通过 RS232 和 Ethernet 连接，构成了双 Heartbeat 网络。双 SCSI 口磁盘阵列作为共享磁盘，分别和服务器 A 和 B 相连，用于保存共享的元数据。控制台通过本地以太网连接到服务器上，完成相应的管理工作。这种设计的优点是：管理方便、系统标准化和配置灵活。

双机高可用性软件分为 3 大模块，如图 2 所示。从上到下依次是 HA_ADMIN 模块、HA_KERNEL 模块和 HA_AGENT 模块。

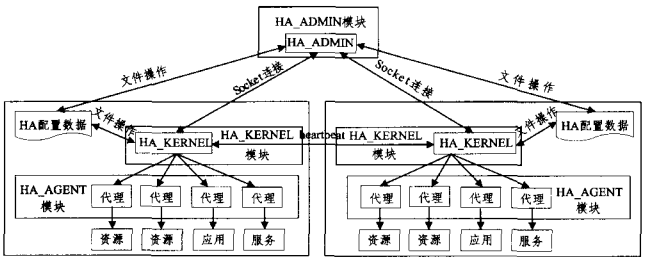

图 2　双机双工高可用服务器系统的软件结构

HA_ADMIN 模块是系统提供给外部的管理控制程序，该模块设定全局的配置信息（各种系统参数），包括组，应用，资源以及它们之间的相互依赖关系。该模块提供图形界面与用户交互。把用户输入的实时命令传递给 HA_KERNEL 模块并且需要读写配置文件。配置文件中保存了整个系统的配置信息，如系统监测的资源信息、应用信息、程序中的代理模块的信息等。

HA_KERNEL 模块实现了 HA 系统的核心逻辑，它包括 3 个子模块：HA 模块、HA_COMM 模块、HA_MONITOR 模块。其主要的功能有：分析本地、异地指令并执行、根据配置文件的设定启动各服务组。运行时根据 HA_AGENT 模块返回的服务组的状态对服务组进行 RERUN，STOP，TAKE_OVER 等操作。维护并管理 3 个进程：HA 进程、HA_COMM 进程及 HA_MONITOR 进程。

(1) HA 进程分时轮询分析本地、异地指令并执行，分时轮询与 HA_COMM，HA_MONITOR 之间的共享内存，实现信息交互以及对这两个进程的维护。

(2) HA_COMM 进程分时轮询分析本地、异地指令的执行情况，检查 HA 进程的状态，心跳间隔监测对机发来的心跳信息以及发送本机的心跳信息到对机。

(3) HA_MONITOR 进程分时轮询与 HA 之间的共享内存来实现信息交互以及对 HA 进程的维护，载入 agent 动态链接库(DLL)，并调用该动态链接库中相应的 agent 来实现对所监测的应用/服务/资源侦测。

HA_AGENT 模块实现对应用、服务、资源的状态的监测，实时地将监测对象的状态传递给 HA_KERNEL 模块。同时从 HA_KERNEL 中取得系统对应用、服务、资源的操作命令，执行相应的各种操作。对应用、服务、资源的操作主要包括：对应用/服务包括 RUN，RERUN，STOP 等操作，对资源包括 GET，RELEASE 等操作。HA_AGENT 模块的实现形式将由多个针对不同监测对象的动态链接库组成。针对每一个应该保证其可用性的服务和资源对象都将有一个与之相应的代理程序。

2.2　管理模块的设计

HA_ADMIN 是管理模块，它通过 Socket 与 HA_COMM 模块交互、远程或本地管理系统。HA_ADMIN 模块实现对 HA 系统的配置与管理，设定全局的配置信息（各种系统参数），包括组、应用、资源以及它们之间的相互依赖关系，同时提供图形界面与用户交互、支持管理员实时命令的处理。高可用软件的管理模块是一个相对独立的模块，可以运行在本机或网络上的客户机上，它是一个独立的应用程序。管理模块主要是通过 Socket 与被管理的服务器通讯。考虑到桌面操作系统中 Windows 应用得较多，所以程序主要在 Windows 平台下开发，使用的是 Microsoft Visual C++工具来开发。模块结构如图 3 所示。

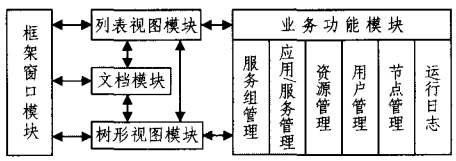

图 3　HA_ADMIN 模块结构

各个模块的功能如下。

框架窗口模块：框架窗口即程序的主窗口。框架窗口模块完成对程序主窗口、菜单栏、工具栏的管理，同时也负责列表视图窗口的创建、树形窗口的创建。

列表视图模块：列表视图是程序的主视图，负责各个业务功能的用户界面交互。列表视图模块完成对列表视图的管理，同时也负责与树形视图模块及文档模块之间的交互。最终，一部分用户的输入指令也由此模块调用业务功能执行模块来实现。

树形视图模块：树形视图是程序的导航视图，提供给用户所有程序功能的一个树形层次结构视图。树形视图模块完成对树形视图的管理。

文档模块：采用基于 Visual C++的文档——视图结构。文档模块为列表视图模块及树形视图模块提供数据，同时程序启动时一些初始化工作也由此模块完成。

业务功能执行模块：此模块负责所有业务功能的实现。所有业务功能以统一结构实现，提供统一调用接口给调用者。

2.3　核心模块的设计

HA_KERNEL 模块是系统最核心的部分，图 4 为该模块的结构图。该模块由 3 个子模块组成：HA 模块、HA_COMM 模块、HA_MONITOR 模块。3 个模块之间共享内存

通信。HA_COMM 模块负责网络通讯，包括与对机的通讯和与 HA_ADMIN 模块的通讯。HA 模块为控制模块，它从 HA_MONITRO 模块中得到资源、应用和服务的状态信息，根据这些状态信息决定对应的操作并把这些信息通过 HA_COMM 模块发送到对方计算机。根据从配置文件和对方计算机中得到的信息产生对于资源和服务的处理命令，写命令到共享内存。HA_MONITOR 模块会从共享内存中得到这些命令，而通过相应的代理执行这些命令操作。HA_MONITOR 模块的主要功能是通过各种代理监视各种资源、应用和服务的工作状态，并将获得的状态信息写入共享内存中，同时从共享内存中取 HA 模块写入的相应命令，并通过相应的代理来执行这些命令。

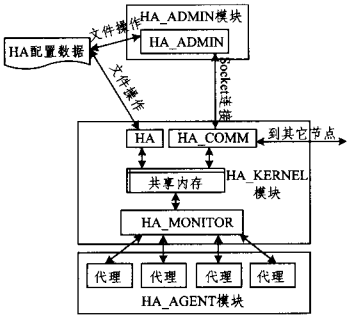

图 4　HA_KERNEL 模块的结构

由于篇幅所限，HA_AGENT 的设计不做介绍。

3　双机高可用系统性能分析

为了便于对比，分两种情况进行性能分析：第一种情况是单服务器单网络路径（交换机）系统的可用性分析；第二种情况是双机高可用性系统（双服务器双交换机）的可用性分析。

3.1　单机单路径系统的可用性

图 5 和图 6 分别为单机系统和相互冗余的双机系统的状态转换图。单机系统只有正常工作和不正常工作两种状态，如图 5 所示：1 表示正常工作状态，2 表示不正常工作状态。双机系统有 3 种工作状态：两个机器都能正常工作、一个机器能正常工作和两个机器都不能正常工作，如图 6 所示：1 表示两个机器都能正常工作，2 表示只有一个机器能正常工作，3 表示两个机器都不能正常工作。其中常数 λ 是故障率，常数 μ 是故障修复率。

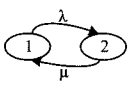

图 5　单机系统的状态转换图

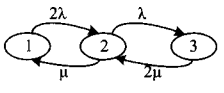

图 6　双机系统的状态转换图

由文献 [4]可知，节点冗余的可用度公式为

$$A_n=1-P_n=1-\left(\frac{\lambda}{\mu+\lambda}\right)^n$$

其中 n 是互为冗余的节点数，可得图 5 单机系统和图 6 双机系统的可用度公式分别为

$$A_1=1-P_1=1-\frac{\lambda}{\lambda+\mu} \tag{1}$$

$$A_2=1-P_2=1-\left(\frac{\lambda}{\mu+\lambda}\right)^2=\left(1-\left(\frac{\lambda}{\lambda+\mu}\right)\right)\left(1+\left(\frac{\lambda}{\lambda+\mu}\right)\right)=\left(1+\frac{\lambda}{\lambda+\mu}\right)A_1 \tag{2}$$

其中 A_1 为单机系统的可用度，A_2 为双机系统的可用度。因为 $\left(1+\frac{\lambda}{\lambda+\mu}\right)>1$，所以 $A_2>A_1$，由此可知双机系统的可用度比单机系统的可用度要高。由文献[4]可知单机单路径的可用度 A_{11} 为

$$A_{11}=\pi_1=\frac{1}{1+\frac{\lambda_1}{\mu_1}+\frac{\lambda_2}{\mu_2}} \tag{3}$$

其中常数 λ_1 是服务器的故障率，常数 μ_1 是服务器的故障修复率，常数 λ_2 是路径（交换机）的故障率，常数 μ_2 是路径（交换机）的故障修复率。

3.2　双机高可用系统的可用性能分析

从图 1 可知，双机高可用性系统包括互相冗余的双机及双网络路径。当一个主机出现故障时，正常主机接替失效主机的业务；当一个网络路径出现故障，所有 I/O 请求都经过正常路径传输。为了使用连续时间马尔可夫链建立双机高可用系统的性能分析模型，我们进行如下假设：(1)假设服务器（主机）、网络交换机等部件的平均无故障时间分别为 $MTTF_1$ 和 $MTTF_2$，它们的平均故障修复时间分别为 $MTTR_1$ 和 $MTTR_2$。并且，$MTTF_1$，$MTTF_2$，$MTTR_1$ 和 $MTTR_2$ 都服从负指数分布，主机和网络交换机的失效率分别为 λ_1，λ_2（$\lambda_1=1/MTTF_1$，$\lambda_2=1/MTTF_2$），其故障修复率分别为 μ_1，μ_2（$\mu_1=1/MTTR_1$，$\mu_2=1/MTTR_2$）；(2)系统中的各个部件的失效和修复都相互独立；(3)系统中各个失效部件都是可修复的；(4)当系统中的两主机出现失效时，整个系统不能工作，双交换机不再出现故障。同理，当两交换机失效时，整个系统停止工作，主机不再出现故障；(5)假设双机高可用系统中的双端口 RAID 的性能是可靠的，其失效率相对服务器和交换机的失效率可以忽略，RAID 与某一服务器相连的端口失效表示该服务器失效；(6)如果一交换机失效，一主机与正常交换机相连的网卡或连线失效时该主机被认为失效。同理，一主机失效，另一主机正常，并且两交换机都能正常工作，但交换机与正常主机相连的网卡或连线失效时，该交换机被认为失效。

基于上述假设，我们建立双机高可用系统的连续时间马尔可夫链模型（Continuous Time Markov Chain Model，简称 CTMCM），CTMCM 如图 7 所示。在图 7 中，当系统处于状态 1,2,3,4 时是可用的，当系统处于状态 3,6,7,8 时不可用。

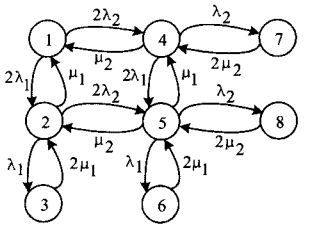

图 7　双机高可用系统对应的 CTMCM

系统处于各个状态时，系统各部件所处的状态如下：状态 1，两主机两交换机都正常，系统能正常工作；状态 2，一主机正常，一主机失效，两交换机正常，系统能工作；状态 3，两主机失效，两交换机正常，系统不能工作；状态 4，两主机正常，一交换机失效，一交换机正常，系统能工作；状态 5，一主机正常，一主机失效，一交换机正常，一交换机失效，系统能工作；状态 6，两主机失效，一交换机失效，一交换机正常，系统不能工作；状态 7，两主机正常，两交换机失效，系统不能工作；状

态 8，一主机正常，一主机失效，两交换机失效，系统不能工作。

为了通过 CTMCM 对双机高可用系统进行性能分析，我们让 π_i 表示系统处在第 i 个状态时的平衡状态概率。根据连续时间的马尔可夫链的状态方程我们有转移概率矩阵 $P(t)$ 为：

$P(t)=[p_{ij}(t)]$，$P(0)=I$

$Q=[q_{ij}]$ 为转移概率矩阵 $P(t)$ 的无限小生成元，q_{ij} 是从状态 i 到状态 j 的转换速率。由于双机高可用系统中的各个状态构成的连续时间马尔可夫链的转移概率 $P_{ij}(t)$ 只与状态 i,j 有关而与时间 t 无关，因此双机高可用系统中的各个状态构成的连续时间马尔可夫链是齐次的。设双机高可用系统中的各个状态构成的连续时间马尔可夫链的状态空间为 S，从图 7 可知：$S=\{1,2,3,4,5,6,7,8\}$，并且除 S 本身之外不存在其它任何闭集，所以由 CTMCM 构成的连续时间马尔可夫链是不可约的。

根据齐次不可约连续时间马尔可夫链的性质可得线性方程组：

$$\begin{cases}\pi Q=0\\ \sum\limits_{i\in S}\pi_i=1\end{cases} \tag{4}$$

其中：$\pi=\{\pi_1,\pi_2,\pi_3,\pi_4,\pi_5,\pi_6,\pi_7,\pi_8\}$，并且双机高可用性系统的转移概率矩阵 $P(t)$ 的无限小生成元为

$$Q=\begin{bmatrix}
-(2\lambda_1+2\lambda_2) & 2\lambda_1 & 0 & 2\lambda_2 & 0 & 0 & 0 & 0\\
\mu_1 & -(\mu_1+\lambda_1+2\lambda_2) & \lambda_1 & 0 & 2\lambda_2 & 0 & 0 & 0\\
0 & 2\mu_1 & -2\mu_1 & 0 & 0 & 0 & 0 & 0\\
\mu_2 & 0 & 0 & -(\mu_2+2\lambda_1+\lambda_2) & 2\lambda_1 & 0 & \lambda_2 & 0\\
0 & \mu_2 & 0 & \mu_1 & -(\lambda_1+\lambda_2+\mu_1+\mu_2) & \lambda_1 & 0 & \lambda_2\\
0 & 0 & 0 & 0 & 2\mu_1 & -2\mu_1 & 0 & 0\\
0 & 0 & 0 & 2\mu_2 & 0 & 0 & -2\mu_2 & 0\\
0 & 0 & 0 & 0 & 2\mu_2 & 0 & 0 & -2\mu_2
\end{bmatrix}$$

由式 (4)可得下列线性方程组：

$$\begin{cases}
\pi_1+\pi_2+\pi_3+\pi_4+\pi_5+\pi_6+\pi_7+\pi_8=1\\
-(2\lambda_1+2\lambda_2)\pi_1+\mu_1\pi_2+\mu_2\pi_4=0\\
2\lambda_1\pi_1-(\mu_1+\lambda_2+2\lambda_2)\pi_2+2\mu_1\pi_3+\mu_2\pi_5=0\\
\lambda_1\pi_2-2\mu_1\pi_3=0\\
2\lambda_2\pi_1-(\mu_2+2\lambda_1+\lambda_2)\pi_4+\mu_1\pi_5+2\mu_2\pi_7=0\\
2\lambda_2\pi_2+2\lambda_1\pi_4-(\lambda_1+\lambda_2+\mu_1+\mu_2)\pi_5+2\mu_1\pi_6+2\mu_2\pi_8=0\\
\lambda_1\pi_5-2\mu_1\pi_6=0\\
\lambda_2\pi_4-2\mu_2\pi_7=0\\
\lambda_2\pi_5-2\mu_2\pi_8=0
\end{cases} \tag{5}$$

通过双机高可用系统的 CTMCM 可知，其可用度由状态 1、状态 2、状态 4 和状态 5 的稳定概率决定。因为状态 1、状态 2、状态 4 和状态 5 是 4 种可用的状态，而状态 3、状态 6、状态 7 和状态 8 是 4 种不可用的状态，所以双机高可用系统的可用度应为

$A_{22}=\pi_1+\pi_2+\pi_4+\pi_5$

解方程组(5)可得双机高可用系统的可用度：

$$A_{22}=\frac{\mu_1\mu_2^2+\mu_1^2\mu_2+4\lambda_1\mu_1\mu_2+4\lambda_2\mu_1\mu_2+2\lambda_1\mu_2^2+4\lambda_1^2\mu_2+8\lambda_1\lambda_2\mu_1+8\lambda_1\lambda_2\mu_2+2\lambda_2\mu_1^2+4\lambda_2^2\mu_1+8\lambda_1^2\lambda_2+8\lambda_1\lambda_2^2}{b+c+ad} \tag{6}$$

a,b,c,d 为：

$$\begin{cases}
a=4\lambda_1\lambda_2\mu_1+2\lambda_2\mu_1^2+4\lambda_2^2\mu_1+2\lambda_2\mu_1\mu_2\\
b=\dfrac{(\mu_2+2\lambda_1+\mu_1+2\lambda_2)(\mu_1^2\mu_2-2\lambda_2\mu_1\mu_2-\lambda_1\lambda_2\mu_2-\lambda_2^2\mu_1)}{\mu_1}\\
c=\dfrac{4\lambda_1\mu_1\mu_2^2+8\lambda_1^2\mu_1\mu_2+8\lambda_1\lambda_2\mu_2+4\lambda_1\mu_1^2\mu_2+2\lambda_1^2\mu_2^2+4\lambda_1^3\mu_2+4\lambda_1^2\lambda_2\mu_2+2\lambda_1^2\mu_1\mu_2}{2\mu_1}\\
d=(1+\dfrac{\lambda_1}{2\mu_1}+\dfrac{\lambda_2}{2\mu_2})\cdot\dfrac{\mu_2+2\lambda_1}{\mu_1}+1+\dfrac{\lambda_2}{2\mu_2}
\end{cases}$$

3.3 双机高可用系统的性能指标估算

下面以单机单路径和双机双路径为例来估算双机高可用系统的可用性。我们做如下假设：

(1) USN 元数据服务器系统的平均故障时间间隔为

$MTTF_1\geqslant 1000$h

其失效率为

$\lambda_1=1/MTTF_1=1\times10^{-3}$/h

(2) 与元数据服务器相连的网络交换机的平均故障时间间隔为

$MTTF_2\geqslant 2000$h

其失效率为：

$\lambda_2=1/MTTF_2=5\times10^{-4}$/h

(3) USN 元数据服务器的平均故障修复时间为 2h，则其修复率为

$\mu_1=0.5$/h

(4) 交换机的平均故障修复时间为 1h，则其修复率为

$\mu_2=1.0$/h

基于上述假设，根据式(3)和式(6)可得单机单路径系统及双机高可用系统的可用度分别为

$$A_{11}=99.75062\% \tag{7}$$

$$A_{22}=99.99912\% \tag{8}$$

从式(7)和(8)可知单机单路径系统一年停机时间为21.9h，而双机双路径高可用系统一年停机时间只有 4.62min。

如果令 $\lambda=\lambda_1$，$\mu=\mu_1$，由式(1)和式(2)可得

$$A_1=1-\frac{\lambda}{\lambda+\mu}=1-\frac{\lambda_1}{\lambda_1+\mu_1}=0.99800=99.800\% \tag{9}$$

$$A_2=1-\left(\frac{\lambda}{\mu+\lambda}\right)^2=1-\left(\frac{\lambda_1}{\mu_1+\lambda_1}\right)^2=99.99960\% \tag{10}$$

从式(7),(8),(9)和(10)可知,双机互相冗余系统和双机双路径高可用系统的可用度都可达到"5 个 9",即 0.99999,而单机系统和单机单路径系统的可用度只有"2 个 9",即 0.99。

通过以上性能指标估算,我们可得出结论:在服务器级和网络级(交换机)通过增加各级的冗余度可提高各级部件的可用度,从而提高整个系统的可用性。

结束语 对于融合 NAS,iSCSI 和 SAN 的复杂海量网络存储系统,保证元数据服务器系统的高可用是整个存储网络系统实现高可用的关键。采用共享双端口 RAID 的双机高可用系统可充分保证系统元数据服务器系统的高可用性。在 USN 元数据服务器双机高可用系统中,整个软件结构分成 3 大部分:可视化的控制台模块、高可用核心模块和高可用代理模块,双服务器通过 3 个模块协同工作保证整个元数据服务器系统的高可用性。建立连续时间的马尔可夫链模型对双机高可用系统的可用性进行分析结果显示:双机高可用系统的可用度远优于单机单路径系统,这从理论上进一步证明了双机高可用系统能为系统提供高的可用性。

参 考 文 献

[1] Barraza O. Achieving 99.9998% Storage Uptime and Availability. Dot Hill Systems Corp, 2002. http://www.dothill.com/products/whitepapers/5-9s_wp.pdf

[2] Rueda A, Pawlak A. Pioneers of the reliability theories of the past 50 years // 2004 Annual Symposium-RAMS, 2004. NY USA: ACM Press, 2004: 102-109

[3] 韩德志,余顺争,谢长生. 融合 NAS 和 SAN 的统一存储网络系统的设计与实现. 电子学报,2006(11)

[4] 韩德志. 统一存储网络高可用关键技术研究[D]. 武汉:华中科技大学图书馆,2005

(上接第 54 页)

表 4 变迁保证函数与健康调节

变迁	保证函数 G	健康调节 K
t_1	web1. h⊥f1≤C(s_0)⊥f_1	C(s_0)−(0,0,10,5)⊥ζ
t_2	web2. h⊥f2≤C(s_0)⊥f_2	null
t_3	ips. h⊥f3≤C(s_0)⊥f_3	C(s_0)+(0,0,20,10)⊥ζ
t_4	web1. h⊥f1≤C(s_3)⊥f_1	C(s_3)−(0,0,10,5)⊥ζ
t_5	web2. h⊥f2≤C(s_3)⊥f_2	null
t_6	vid. h⊥f1≤C(s_1)⊥f_1	null
t_7	vid. h⊥f2≤C(s_2)⊥f_2	null

• 用户以 $\mathcal{R}_1(u)=\{staffer, tester\}$ 访问,ζ=(1,0,1,1),指派给用户的权限为{*web1*,*ips*,*vid*},如图 4 中所示,可能发生的变迁有 t_1,t_3,t_4,t_6。若初始状态为(50,0,70,85),可发生 t_1,但是不可发生 t_1t_6,即不可通过 *web1* 获得 *vid*,因为 t_1 发生时产生了健康衰减,导致 $C(s_1)$=(50,0,70,85)−(0,0,10,5)⊥ζ=(50,0,60,80),$G(t_6)$=(30,50,70,80)⊥(0,0,1,1)≤$C(s_1)$⊥(0,0,1,1)=false。但可发生 $t_3t_4t_6$,即通过 *ips* 获得 *web1*,再获得 *vid*,由于 t_3 发生产生了健康增长:$C(s_3)$=(50,0,70,85)+(0,0,20,10)⊥ζ=(40,0,90,95),$G(t_4)$=true;$C(s_1)$=(40,0,90,95)−(0,0,10,5)⊥ζ=(40,0,80,90),$G(t_6)$=true。

• 用户以 $\mathcal{R}_2(u)=\{staffer, visitor\}$ 访问,ζ=(1,0,1,0)。指派给用户的权限为{*web2*,*ips*,*vid*},可能发生的变迁是 t_2,t_3,t_5,t_7。若初始状态为(40,0,50,0),t_2,t_7 均不可发生,因为 $G(t_2)$=(20,30,55,70)⊥(1,0,1,0)≤(40,0,50,0)⊥(1,0,1,0)=false。但能发生 $t_3t_5t_7$,因为 t_3 发生产生了健康增长,导致 $G(t_5)$=$G(t_7)$=true。

由此可看出,在 ESRBAC 机制下,用户在不同的环境下会有不同的访问权限,而且访问权限会随着用户具体的访问行为动态地变化。ESRBAC 将用户的动态环境纳入到 RBAC 机制中,增强了动态环境下访问控制的安全性和灵活性。

结束语 本文提出了基于环境安全的角色访问控制模型 ESRBAC 模型,将 RBAC 模型中的角色与一定的环境策略相关联,通过一个能反映用户环境安全性的端点模型,将环境上下文的安全性纳入到 RBAC 权限分配的策略中来,提高了 RBAC 模型对网络动态环境的适应性。并用着色 Petri 网的方法描述分析了一个基于 ESRBAC 的应用实例模型,证明了 ESRBAC 模型的可用性和安全性。

参 考 文 献

[1] Sandhu R S, Coyne E J, Feinstein H, et al. Role-based access control models. IEEE Computer, 1996, 29(2): 38-47

[2] Joshi J B D, Bertino E, Latif U, et al. A generalized temporal role-based access control model. IEEE Transactions on Knowledge and Data Engineering, 2005, 7(1): 4-23

[3] Covington M J, Sastry M R, Manohar D J. Attribute-based Authentication Model for Dynamic Mobile Environments // Proc. of the 3rd International Conference of Security in Pervasive Computing. 2006: 227-242

[4] Shafiq B, Joshi J B D, Bertino E, et al. Secure Interoperation in a Multidomain Environment Employing RBAC Policies. IEEE Transactions on Knowledge and Data Engineering, 2005, 17(11): 1557-1577

[5] Covington M J, Long W, Srinivasan S, et al. Securing context-aware applications using environment roles // Proc. of the 6th ACM Symposium on Access Control Models and Technologies. 2001: 10-20

[6] Teo L, Ahn G J, Zheng Y L. Dynamic and Risk-aware Network Access Management // Proc. of the 7th ACM Symposium on Access Control Models and Technologies. 2003: 217-230

[7] Chakraborty S, Ray I. TrustBAC - Integrating Trust Relationships into the RBAC Model for Access Control in Open Systems // Proc. of the 11th ACM Symposium on Access Control Models and Technologies. 2006: 49-58

[8] Sandhu R S. Lattice-based access control models. IEEE Computer, 1993, 26(11): 9-19

[9] Jiang Y X, Lin C, Yin H, et al. Security Analysis of Mandatory Access Control Model // Proc. of 2004 IEEE International Conference on Systems. 2004: 5013-5018

[10] 欧阳凯,周敬利,夏涛,等. 基于 SSL VPN 接入机制的研究. 计算机科学,2005,32(5):59-63

[11] 王瑜,卿斯汉. 一种新的访问控制模型 TBPM — RBAC. 计算机科学,2005,32(2):169-172