基于等级保护的网络容灾系统模型*)

蔡皖东
(西北工业大学计算机学院　西安710072)

摘　要　等级保护是信息安全技术的重要原则之一，作为信息安全重要组成部分的网络容灾系统同样要遵循这一原则。本文详细论述了容灾系统等级划分、4种不同等级的网络容灾系统模型以及灾难管理等内容。

关键词　等级保护，网络容灾，数据复制，灾难管理

Grading Protection-Based Network Disaster-Tolerant System Models

CAI Wan-Dong
(College of Computer Science, Northwestern Polytechnical University, Xi'an 710072)

Abstract　Grading protection is an important tenet in information security technology. This tenet is also applicable to the network disaster-tolerant system. In this paper, the disaster-tolerant system grading, four network disaster-tolerant system models and disaster management are discussed in detail.

Keywords　Grading protection, Network disaster-tolerant, Data replication, Disaster management

1　引言

随着网络技术的快速发展和广泛应用，在金融、证券、税务以及民航等行业中，开始采用“数据大集中”模式来构架网络信息系统，即各个基层单位子系统所产生的业务数据不再由本地计算机处理和保存，而是集中在数据中心进行统一的处理和存储。这种模式反映了网络计算模型逐步从传统的“分布式计算，分布式存储”模型向先进的“分布式计算、集中式存储”模型转变，更加有利于数据管理和安全。

由于网络系统的开放性、复杂性和各种天灾人祸因素的客观存在，使网络系统面临着自然灾害、恐怖事件、物理故障、黑客攻击、病毒破坏等来自于自然界和人为的各种威胁。一旦系统发生灾难性事件，不仅导致系统服务中断，而且还会破坏系统中所存储的大量数据和信息，从而造成不可挽回的灾难性后果。例如，美国“9.11”事件给很多企业带来了灭顶之灾。因此，采用“数据大集中”模式的网络信息系统，通常都要采用网络容灾技术来保护系统和数据，提高网络系统的抗毁能力。

网络容灾技术的基本思路是在异地建立和维护一个备份系统，利用资源冗余性和地理分散性来保证网络系统对灾难事件的抵御能力。网络容灾系统核心技术主要有两个方面：数据复制和应用切换。数据复制是指在异地建立一个数据备份系统，实时或定时地备份本地关键性数据。应用切换是指在数据复制的基础上，在异地建立一个应用系统的远程镜像，在发生灾难情况下，快速地切换到远程应用系统，恢复系统运行。

近年来，网络容灾技术受到国内外的高度重视，开发出一些网络容灾技术和产品。如何选择网络容灾技术和产品来构架网络容灾系统成为人们关注的焦点，这里涉及到诸多因素，如系统恢复时间、灾难发生时丢失的数据量、系统成本和运行费用等，从而产生不同的容灾等级。从信息安全的角度，等级保护是信息安全技术的重要原则之一，容灾系统作为信息安全的重要组成部分，同样也要遵循这一原则。

本文首先讨论了容灾系统等级划分，然后提出和论述4种不同等级的网络容灾系统模型，最后探讨了灾难管理问题。

2　容灾系统等级划分

美国SHARE用户组和IBM公司于1992年共同提出了灾难恢复等级概要SHARE(Summary of Disaster Recovery Tiers)78标准，根据备份数据量大小、本地系统和远程备份系统之间的距离和网络连接方法、灾难发生时丢失的数据量、系统恢复时间、系统成本和运行费用等因素，将容灾系统分为7个等级，参见表1，分别适用于不同的系统规模和应用场合。

在SHARE 78标准中，符合现代网络容灾要求的容灾系统是第4～6级，它们具有类似的系统框架，其差别在于数据复制软件的功能差异和备份站点的硬件配置不同。第4级容灾系统只需配置远程系统备份软件即可工作，第5级容灾系统依赖于数据库系统的两步提交法来保证数据的同步，第6级容灾系统则需要配置更可靠的数据管理软件和专用的硬件设备，以保证灾难事件发生时零数据丢失和备份站点的即时切换。

由于SHARE标准提出得比较早，容灾系统等级划分和实现模型依赖于当时的技术水平，存在一定的局限性。近年来，随着宽带网络技术、集群计算技术和网络存储技术的发展，“分布式计算、集中式存储”网络计算模型占据主导地位。在这种网络计算模型中，网络服务系统与数据存储系统相分离，数据集中存储在网络存储器上，通过高速的存储域网络(SAN)链路与网络服务器连接，提供网络集中存储服务，为构建高级别的网络容灾系统提供了重要的技术支撑，对容灾系统等级划分、容灾质量和实现技术产生了巨大的影响。

*)本文受航空基础科学基金项目(项目号：03F53031)和西安市工业攻关项目(项目号：GG200312)的资助。蔡皖东　教授，主要从事计算机网络、网络信息安全等方面研究。

表1 SHARE 78标准等级划分

等级	备份方式	恢复时间	恢复质量	系统成本	说明
0	本地备份		不能恢复数据	最低廉	不提供任何灾难恢复能力
1	汽车运送	1个星期左右	不能完全恢复数据	比较低	基于冷备份站点的异地数据备份
2	汽车运送	1天左右	不能完全恢复数据	有所增加	基于热备份站点的异地数据备份
3	网络传送	一天之内	不能完全恢复数据	有所增加	基于热备份站点的异地数据备份
4	网络传送	小时级	最近一次数据备份以后的数据丢失	有所增加	基于热备份站点的异地数据备份，并具有应用切换功能
5	网络传送	分钟级	处于传送中尚未完成提交的数据丢失	大大增加	基于数据镜像的异地数据备份，采用两步提交法来保证数据同步，并具有应用切换功能
6	网络传送	立即恢复	零数据丢失	最昂贵	基于数据镜像的异地数据备份，保证数据同步和零数据丢失，同时具有应用切换功能

3 网络容灾系统模型

一个完整的容灾系统应当包括应用容灾和数据容灾两个方面，应用容灾是指通过应用切换机制提供面向应用系统的容灾功能，避免因灾难事件中断应用系统运行。数据容灾是指通过数据复制机制提供面向数据存储的容灾功能，避免因灾难事件毁坏关键性数据，这也是容灾系统中最重要的部分。根据对网络容灾的支撑能力、灾难恢复质量以及系统构建成本等因素，我们提出4种不同等级的网络容灾系统模型。

3.1 单机容灾系统模型

在单机容灾系统中，在本地建立一个主机系统和网络存储器，在异地建立一个网络存储器镜像，两个网络存储器之间通过SAN实现网络连接，提供数据复制通道，参见图1。应用程序运行在主机系统上，通过前端网络提供网络服务，数据存储在网络存储器上，通过数据复制功能将数据备份到远程网络存储器上。

单机容灾系统是一种廉价的网络容灾系统，主要为了解决数据容灾问题，但没有应用容灾功能，在主机系统发生硬件和软件故障时会中断应用系统运行。

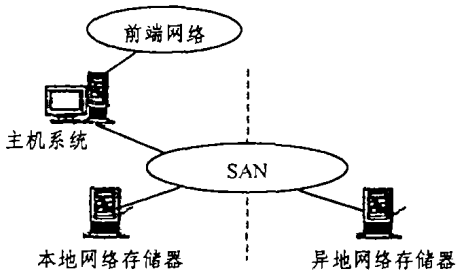

图1 单机容灾系统模型

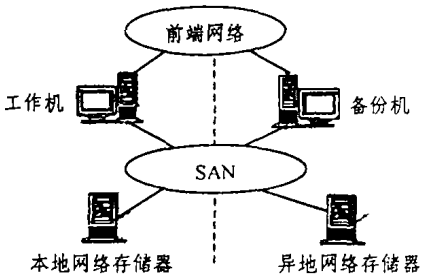

图2 双机容灾系统模型

3.2 双机容灾系统模型

在双机容灾系统中，本地和异地各自建立一套主机系统和网络存储器，主机系统之间通过前端网络来连接，网络存储器之间通过SAN来连接，参见图2。主机系统之间通过“心搏(Heartbeat)”检测机制来检测彼此的工作状态，如果发现工作机出现异常，备份机立即切换成工作机，维持应用程序的正常运行。无论哪个主机成为工作机，本地数据都会复制到远程网络存储器上，以实现数据容灾。

双机容灾系统同时提供了数据容灾和应用容灾功能。由于主机系统之间通过远程链路检测心搏信号，存在一定的网络延迟，对应用切换的质量产生一定的影响。

3.3 单集群容灾系统模型

在单集群容灾系统中，在本地建立一个网络集群系统，应用程序运行在集群系统上。集群系统的数据集中存储在网络存储器上，本地和异地网络存储器之间通过SAN来连接，实现数据复制，参见图3。

在集群系统中，各个主机节点可以分布在本地或异地不同的地理位置上，通过故障管理机制来维护各个节点的可用性，从而实现应用容灾功能。同时，通过适当的负载分配算法将网络负载均衡地分配到可用节点上，从而改善了应用系统的性能和服务质量。

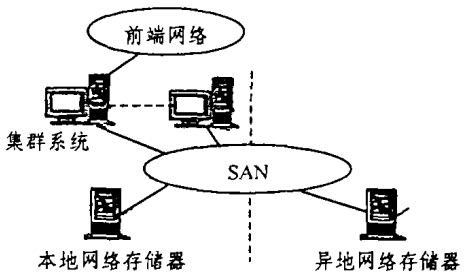

图3 单集群容灾系统模型

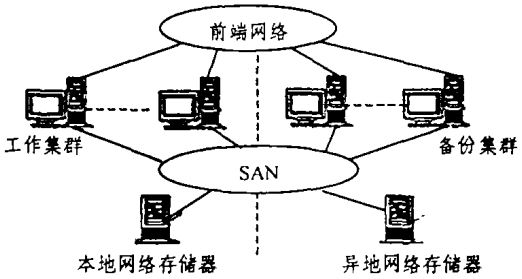

图4 双集群容灾系统模型

3.4 双集群容灾系统模型

在双集群容灾系统中，本地和异地各自建立一套集群系统和网络存储器，集群系统之间通过前端网络来连接，通过心搏检测机制来检测彼此的工作状态，如果发现工作集群出现

异常，备份集群立即切换成工作集群，维持应用系统的正常运行。网络存储器之间通过SAN来连接，实现数据复制功能，参见图4。

在双集群系统中，应用容灾性能和服务质量都是最好的，但系统成本较高。由于灾难事件是小概率事件，异地的备份系统可能长期处于待机状态，可以适当地降低系统配置。一种优化的方案是在异地建立一个单机系统，作为应用容灾的备份机，而本地仍采用集群系统。

在这四种网络容灾系统模型中，每个网络存储器还可以选择配置本地数据备份系统，如磁盘、磁带机等，以近程方式对数据进行实时和定时备份。

3.5 数据复制模式

上述的网络容灾系统模型都提供了数据容灾功能，而数据容灾质量取决于数据复制模式。数据复制模式主要有如下3种：

(1) 同步复制模式：在主存储系统和备份存储系统之间保持完全数据镜像状态。当主机发出一个磁盘I/O操作(主要是写操作)请求时，该请求被同时发送到两个存储系统中，并且必须等到两个存储系统都完成处理后，才向主机返回确认信号，表示本次磁盘I/O操作的完成，主机可以继续下一次磁盘I/O操作。

这种模式的优点是能够保持两个存储系统中数据完全一致。缺点是当两个存储系统之间通信链路速率较低时，将产生比较明显的I/O延迟，对主机的性能产生一定的影响。这种模式要求在两个存储系统之间的通信链路必须使用高速专线，因此系统造价较高。

(2) 半同步复制模式：在这种模式中，主机发出的写磁盘操作请求首先发送给主存储系统，主存储系统完成处理后立即向主机返回确认信号，然后再将数据复制到备份存储系统中。在两个存储系统未完成数据同步之前，主机只能执行读磁盘操作，不能执行写磁盘操作。只有备份存储系统执行完数据更新后，主机才能继续下一次写磁盘操作。

这种模式基于如下事实：大多数应用程序在写数据之前先要执行读数据操作。利用这一时间差能够明显地改善主机的响应性能。

(3) 异步复制模式：在这种模式中，主机对主存储系统的写操作与数据复制操作无关，只要主存储系统返回确认信号，主机就可以连续地执行写操作，而不考虑向备份存储系统的数据复制是否完成。通常，数据复制操作是以后台方式定时执行的，或者采用手工方式来完成。

在这种模式下，两个存储系统之间的数据复制操作不会影响主机的性能，但存在着两个存储系统中的数据不一致问题，不能保证数据的完全恢复。

通常，异地备份中心可以建在同一城市或者不同城市，取决于容灾质量要求和资金投入。同步复制和半同步复制适合于同城复制，异步复制模式则适合于异城复制。

4 灾难管理

为了应对灾难事件，除了在网络容灾系统模型和数据复制等技术层面上必须满足容灾要求之外，灾难管理也是必须的。灾难管理是一个全过程的动态管理，包含事先、事中和事后等阶段。

1. **事先管理** 在这个阶段，主要的管理工作是评估潜在的灾难风险及其影响，制定灾难应急计划和相关管理制度。

(1) 灾难风险评估：分析和评估潜在的灾难事件源、发生灾难事件的可能性、灾难事件可能造成的社会影响和经济损失、灾难应对措施等，评估结果将为设计和构造网络容灾系统提供重要的参考依据。

(2) 灾难应急响应计划：制定灾难应急响应计划和灾难恢复计划，包括数据备份和恢复计划、应用系统切换和恢复程序、灾难事件取证方法、灾难事件影响和损失评估方法以及其它管理制度等。

2. **事中管理** 在这个阶段，主要的管理工作是监管容灾系统工作状态，维护系统正常运行，定期进行灾难恢复演练。

(1) 容灾系统监管和维护：对容灾系统运行状态进行有效监管，包括通信链路、工作/备份主机、工作/备份存储器等设备的工作状态和安全状况，在发生异常时及时维护和处理，保证容灾系统始终处于正常运行状态。

(2) 灾难恢复定期演练：灾难事件是一种突发性的小概率事件，定期地进行系统恢复演练是必要的。一方面能够测试和检验容灾系统的灾难恢复能力，另一方面还能够训练有关人员应对和处理灾难事件的能力，并且不断地完善灾难应急响应计划。

3. **事后管理** 这个阶段也是灾难应急响应阶段，主要的工作是按照灾难应急响应计划对已发生的灾难事件进行调查取证和损失评估，并快速和最大化地恢复系统运行。

结束语 "9.11"事件给很多企业敲响了警钟，人们开始重视网络容灾问题。"数据大集中"业务处理模式又给网络容灾技术的应用和发展提供了很大的空间。由于网络容灾系统的构建成本和运行费用都很高，给用户带来很大的负担，这相当于为系统购买保险，尽管灾难的发生是小概率事件。因此，在充分评估灾难事件可能造成经济损失和社会影响的情况下，采用基于等级保护的思想来构建网络容灾系统是必要的，使得用户投资与被保护资产的价值相平衡。

另外，只有容灾技术措施是不够的，还要重视灾难管理工作，这样才能有效地应对灾难事件。

参 考 文 献

1 蔡皖东，等. 一种网络容灾系统的设计与实现[J]. 计算机工程，2004(4)

2 蔡皖东，等. 基于SAN的快速灾难恢复技术[J]. 计算机科学，2002(2)

3 蔡皖东. 基于SAN的高可用性网络存储解决方案[J]. 小型微型计算机系统，2001(3)

4 IBM公司. Disaster Recovery Strategies with Tivoli Storage Management [R]. ibm.com/redbooks，2000