网络考试系统容错设计与实现①

肖 毅[1]，朱幸辉[1,2]

[1](湖南农业大学 信息科学技术学院，长沙 410128)
[2](湖南农业大学 图书馆，长沙 410128)

摘 要：考试具有很强的公正性和时效性，正常的考试过程是不容许有差错和过多中断的。而在考试系统的实际应用中，常因为其存在功能和安全方面的问题导致考试数据的错误和考试过程的中断，因此如何有效地保障考试系统的正常运行成为当前一个重要研究课题。本文以基于客户机/服务器（C/S）结构的考试系统为研究背景，从安全性方面入手，采用服务器主备双机容错技术、双服务器推拉式通讯技术、客户端单线程管理技术以及客户端失效检测技术，实现了对考试系统的容错管理，提高了考试系统的可靠性和可用性。

关键词：考试系统；安全；容错；可靠性

Design and Realization of Fault Tolerate of Network Examination System

XIAO Yi[1]，ZHU Xing-Hui[1,2]

[1](Information Science and technology College, Hunan Agricultural University, Changsha, 410128, China)
[2](Library, Hunan Agricultural University, Changsha, 410128, China)

Abstract：Neither error nor interrupt was allowed in natural examination, to insure the exam be fair-minded and timely. While in the practical application of exam system, there were fault of data and interrupt of process, due to several problems on function or security, so how to ensure the normal circulation of exam system effectively becomes an important research topic recently. In this paper, based on client / server (C / S) test system, started with security, through the technologies of fault tolerate in main and spare servers、double servers push-and-pull communication、mongline manage and invalid check in client, the management of fault tolerate in exam system was designed and realized, which can enhance the reliability and usability of examination system.

Key words：examination system; security; fault tolerate; reliability

随着信息技术的发展，考试系统作为一种现代考核手段愈来愈受到重视，网络考试系统具有经济、灵活、方便、快捷等优点，但现今应用的网络考试系统或多或少存在一些运行稳定性及数据安全性方面的问题。考试的特性主要是公正性和时效性，因此必须使考试系统在保证公平公正的前提下，尽量减少甚至不发生错误，这就要求服务器在考试期间要保证系统运行的稳定及安全。

完全杜绝差错是不可能的，如何在出现问题的情况下，尽量减少损失才是重中之重。容错就是当由于种种原因在系统中出现了数据、文件损坏或丢失时，系统能够自动将这些损坏或丢失的文件和数据恢复到发生事故以前的状态，使系统能够连续正常运行的一种技术。加入容错机制可以很好的避免考试服务器因软硬件系统故障或网络通讯故障导致的考试中断和信息丢失[1-3]。

1 服务器容错

1.1 服务器容错概述

服务器容错就是采取一定方法使服务器对于错误有一定的抵抗能力。服务器错误通常可能是以下原因产生的：① 进程缺乏资源——硬件、软件和 OS 错误。

① 收稿时间:2011-08-15;收到修改稿时间:2011-08-29

© 中国科学院软件研究所 http://www.c-s-a.org.cn

② 由于程序错误引起的进程终止。③ 应用的暂时错误或者持久错误——错误的逻辑、不完整的异常处理、客户服务的版本不匹配、死锁等[4,5]。针对这些错误，服务器容错通常有以下两种方式：

(1) 应用服务器冗余容错

该方式基于软件容错，其基本原理是在同一台物理服务器上启动多个应用服务器对象，即冗余服务对象，多个服务对象同时接受并处理用户请求。在这种情况下，若其中一个服务对象的效率下降或出现错误，由于其他同类服务对象还可以正常工作，则整个系统除了性能有一点下降外，不会表现更多异常。

(2) 双机容错

该方式基于硬件容错，考试系统同时配备两台物理服务器，每台均启动应用服务程序，一台作为主服务器，另一台作为备用服务器，主备服务器同时工作，唯一区别是备用服务器不具备控制功能。当主服务器发生故障时，备用服务器自动切换成主服务器，同时启动控制功能模块。

以上两种服务器容错方式各有缺点，冗余服务对象容错的优点是节约成本，不需要硬件设备的投资，但由于多个应用服务器同时处理数据，必须引入负载平衡机制，该方式的致命缺点是不能处理硬件容错和操作系统容错，一旦硬件出现故障，或者操作系统发生故障，整个考试系统将处于瘫痪状态，即使临时换用一台服务器，也会由于考生数据无法恢复导致重考。基于此，本文考试系统的服务器容错采用双机容错，以保证考试的可靠性[6,7]。

1.2 双机容错系统设计原理

综合参考众多容错技术的优缺点并结合大多数考试系统的应用环境，笔者认为在考试系统中应采用的容错技术是对传统双机容错中双机热备份方案的改进方案：两台服务器，一主一从，互相备份，共同执行同一服务，但中间不架设 RAID（Redundant Array of Inexpensive Disk）标准的磁盘阵列。当一台服务器出现故障时，可以由另一台服务器承担服务任务，从而在不需要人工干预的情况下，自动保证系统能持续提供服务[8]。

在此种容错方案中，操作系统和应用程序安装在两台服务器的本地系统盘上，整个网络系统数据是通过磁盘阵列集中管理和备份的（注意：两台服务器的配置要求完全一样）。容错系统采用“心跳”方法保证主系统与备用系统的联系。所谓“心跳”，指的是主备系统之间相互按照一定的时间间隔发送通信信号，表明各自系统当前的运行状态。一旦“心跳”信号表明主机系统发生故障，或者备用系统无法收到主机系统的“心跳”信号，则系统将认为主机系统发生故障，使主机停止工作，并将系统资源转移到备用系统上，备用系统将替代主机发挥作用，以保证网络服务运行不间断。双服务器之间采用完备的“推拉式”系统监控，使容错系统既经济实用，又能很好地保证考试系统的安全[9,10]。容错机制组成如图 1 所示。

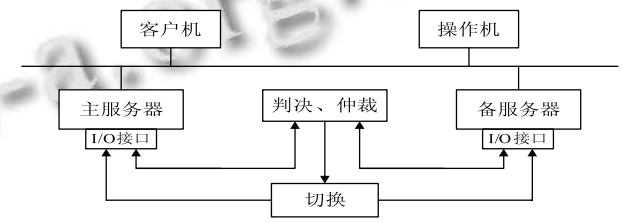

图 1　容错机制组成框图

主服务器和从服务器之间按一定时间间隔向对方发送消息，采用“推拉”模式来确认对方状态，一旦从服务器确定主服务器发生故障，能准确确定故障原因及发生位置，在发出警告的同时，接管主服务器功能，使考试不受影响。

2 考试系统容错设计与实现

2.1 主备服务器间监控及切换模块

软件系统安装完毕正式启动后，两台 PC 服务器担任不同的角色，一台为主用系统（Primary System），另一台为备用系统（Standby System），在这两台机器上都运行有一个进程叫“SMON”，它们通过以太网和串行口相互监视着对方的工作状态。在正常情况下，只有主用系统通过网络接受客户机的访问，当主用系统出现故障时，备用机就接管客户机对系统的访问，接管内容包括原主用系统的网络地址及数据库等，并重起原主用系统的各类应用程序。

系统运行时，相关运行模块通知切换模块（进行必要的注册），由切换模块通知相关模块进行模块检测，当相关模块异常时通知切换模块进行切换，切换模块主要进行统一的维护和通知进行模块检测，当相关模块检测到异常时，通知切换模块进行相应的工作。

正常情况下，主备服务器中的初始化模块在完成

© 中国科学院软件研究所 http://www.c-s-a.org.cn

双机初始化工作后将启动连接更新模块、包收发模块和心跳模块。连接更新模块将本机需要更新的相关信息通过包收发模块发送给另一主机进行处理，包收发模块在传输数据之前向心跳模块发出传输请求，确认对方服务器当前状态。心跳模块根据心跳信息检测主备机状态，若双方状态正常则向包收发模块返回“状态正常，可以传输”的响应信息。同时，心跳模块在系统运行过程中不断解析心跳报文，如果是模拟连接报文，通知模拟连接模块发起模拟连接。如果是同步信息交给同步模块进行同步处理，同步模块将同步信息通过心跳模块传给另一主机进行同步操作。

一旦心跳模块检测到状态异常，则驱动故障检测模块进行故障检测、诊断和处理，故障检测模块发现故障后，将诊断结果返回给心跳模块，心跳模块将通知切换模块进行切换操作，切换模块将切换消息通过心跳模块传给另一主机。

2.2 主备服务器一致性

双机系统发生切换存在两种状态：主机正常运行，备机重新加入整个考试系统和双机在运行中某一服务器故障退出系统。这两种状态发生时，既要保证切换不影响考试运行，又要保证双机能同步工作，就要求设计时必须考虑双机的一致性，其中包括：状态一致性和数据一致性[11,12]。

(1) 状态一致性

状态一致性是主备服务器切换时最重要的部分。系统中，如果有 TCP 连接存在，主服务器就要把来自 TCP 客户端的 TCP 报文通过双机系统专用的传输通道转发给备份服务器，而一旦传输过程中有报文丢失，主备服务器之间状态就不能同步。因此在备机进入的时候，需要让主服务器暂停对 TCP 报文的处理，先将 TCP 报文进行缓冲，当主备服务器状态达到一致后，再将缓冲的 TCP 报文转发给备份服务器，使双机能够同时处理 TCP 报文。

还有一种特殊情况是当双机系统中单机出现永久故障，需要将故障机从系统中脱离。系统从双机正常状态跃变为单机正常状态。由于单机正常状态系统的可用度降低，因此应该尽量避免，并且应尽快恢复成双机正常状态，以保证整个系统的可靠运行。

(2) 数据一致性

数据一致性是指主机和备机的考生数据、配置数据和统计数据保持一致。当主服务器组卷时，将主机中所有考生生成的试卷目录及相关用户等信息复制到备机上。而在考试中，为节省网络开销，考生客户端仅和主服务器联系，不和备服务器发生联系，考生考试情况由主机及时发送给备机。

当考生客户端和主机失去联系时，在反复重试多次无效的情况下，及时向备机报告。备机在收到客户端报告后，会向其他客户端确认，同时向主机发出询问信息，若确认所有客户端均与主机失去联系，且未接到主机反馈信号，则认为主机发生故障，将本身转换为主机，接管所有业务，同时将从失去联系开始，将所有数据填入标记，一旦故障恢复，将所有带标记的数据重新传送给主机，以保持数据一致性[4,5]。本文重点讨论数据复制过程中数据的一致性问题。

在数据的复制过程中，我们必须保证数据的传输顺序，确保备机的执行顺序与主机一致，而且确保到达备机的数据无丢失，也无重复。这就要求数据传输做到不漏传，不重传，也不能颠倒顺序传。下图说明了我们如何达到上述目的。

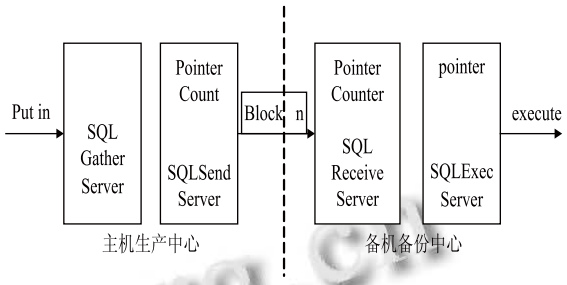

图 2 双机数据复制传输模型

如图 2 所示，在 GatherServer 与 SendServer 之间和 ReceiveServer 与 ExecServer 之间不存在数据传输，他们之间采用数据文件的重命名机制，数据传输主要发生在 SendServer 与 ReceiveServer 之间。为保证数据传输的可靠性，在对象调用时采用同步机制，包括生产中心加载服务器对 GatherServer 的调用、SendServer 对 ReceiveServer 的调用以及 ExecServer 对加载服务器的调用。

GatherServer 将接收到的数据往数据文件后面追加，因此文件中数据的先后顺序与生产中心加载服务器提交的顺序是一致的，但需要在数据文件中预留一片存储区供后续的服务器记录文件指针用。

数据在 SendServer 与 ReceiveServer 之间传输时涉及到传输顺序控制问题。我们的解决办法是在两个服

© 中国科学院软件研究所 http://www.c-s-a.org.cn

务器端各设一个计数器（count 和 counter），SendServer 每发一个数据块都给 count 增加一定的步长值，并将该计数值随数据块传给 ReceiveServer，ReceiveServer 将数据块成功写入文件后将 count 的值赋给 counter，并将其写入文件预留的存储区。ReceiveServer 每次收到一个数据块后都要先将传过来的计数值与上一次收到的计数值作比较，如果顺序是对的就处理该数据块，如果顺序不正确（是前面已处理完的或跳过了一些没处理的块）则返回一个所需要的数据块的标识值，丢弃当前数据块，SendServer 将根据返回的数据块的标识值计算出正确的数据块位置重新发送。当 ReceiveServer 在一个文件还没传完时就 down 掉，重启时将首先从数据文件预留存储区读取计数值以控制下次收到正确的数据块。

SendServer 中的 pointer 变量记录了发送文件当前的指针位置，每次数据块发送成功后就将该值写入文件中预留的指针存储区。这样，当 SendServer 在数据文件还没发送完就 down 掉时，服务器重启时将首先读取文件预留存储区中的指针值，以实现数据的断点续传。在 ExecServer 也设置了一个文件指针变量 pointer，其作用与 SendServer 中的完全类似。一是用于控制每次数据的读取位置，保证数据按正确的位置读取；二是起到断点续传的作用。

由上所述，我们选择了恰当的数据备份策略，采取了一系列的控制措施，基于主机文件系统实现了数据库系统中关键数据在 TCP/IP 网络上的实时同步远程备份，并且保证了数据的一致性传输。

2.3 客户端失效容错

在考试系统的实际使用过程中，除了服务器会出现各种故障外，客服端也经常会出现问题，引起客户端和服务器之间通讯的中断，有些情况下因故障不能被及时发现导致数据的丢失。因此，笔者考虑在服务器容错的基础上引入客户端失效容错，主要包括客户端失效检测和失效处理两部分[13,14]。

(1) 客户端失效检测

客户端失效监测必须在“心跳”机制下保证客户端能和服务器之间始终保持有效的联系，其基本模式主要分为以下几个部分：

传输“心跳”。由考试系统客户端定期地向服务器发送心跳信息，通告它们依然存活，如果超过某一期限没有收到“心跳”，服务器则认为其失效，相当于客户端向服务器“推”（push）失效事件，我们称之为“推”模式。该模式下服务器和客户端之间发送 one-way 消息，所以该方法比较高效。该模式下服务器和客户端之间的消息交换如图 3 所示。

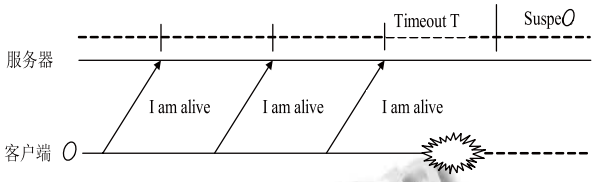

图 3　传输“心跳”中的监控消息

轮询检测。服务器定期轮循检查被监控实体，询问它是否存活，如果客户端作出应答，则意味着它没有失效。相当于服务器从实体“拉”（pull）失效事件，我们称之为“拉”模式。该模式下服务器和客户端之间是 two-way 的消息，因此比“推”模式低效，但由于客户端是被动的，所以该模式更容易被应用开发者使用，它们无需任何有关失效监控所需的信息。该模式下服务器和客户端之间的消息交换如图 4 所示。

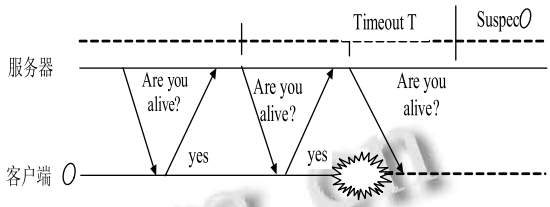

图 4　轮询检测中的监控消息

根据考试系统服务器双机容错的特点，系统采用传输“心跳”方式进行客户端失效检测。

(2) 失效处理

一旦客户端出现失效，服务器将立即提示管理人员该客户端链接断开，并停止对该考生计时，客户端发生失效检测的原因一般有以下几种：① 网络连接中断：针对这种情况，只需由管理员将网络接通即可，一旦网络接通，服务器收到“心跳”，失效解除，考生计时继续。② 客户端对象出错，无法向服务器发送心跳：中止客户端考试进程，重新启动，进行第二次登录，为了确保考试系统的安全性，系统设定第二次登录必须由管理员完成，需要二次登录密码。③ 操作系统出错或断电：重新启动客户端，系统自动连接。④ 客户端主机硬件故障：可在服务器端指定可以更换机器。

© 中国科学院软件研究所 http://www.c-s-a.org.cn

3 结语

本容错系统采用的仍然是传统的双机容错原理，但是很好的结合了实际使用情况，而且能很有效的保证考试的正常运行。因为网络考试一般应用于学校的机房中，因此找两台甚至多台拥有完全一样的硬件和软件系统的计算机比较简单，这样不仅能实现设备的冗余，而且能实现操作系统、应用软件、数据等的同步和冗余，最关键的是系统相同可以在发生故障时实现无缝切换。经过多次对考试系统可靠性和稳定性，特别是容错模块的反应速度的测试，模拟现实中所发生的各种错误操作及紧急情况，观察得出考试系统在出现很多常见错误时仍能正常运行，基本可以达到 2 秒级的实时效果，真正达到无断点运行。

参考文献

1 栾好利.基于局域网的计算机考试系统研究与实现.东北大学学报, 2006,13(10):33−35.

2 梅晓勇,颜君彪,侯识忠.网络环境下的考试系统应用设计与实现.计算机工程与应用, 2003,8(26):43−46.

3 全渝娟,范荣强.基于 Web 的远距离考试系统.计算机应用与软件, 2003,15(7):92.

4 陈胜功.容错计算机技术及应用研究.北京:航空工业出版社, 2000:57−59.

5 王珍熙.可靠性、冗余及容错技术.北京:航空工业出版社, 1991:243−247.

6 郝莹.网上考试系统中的安全机制.微机发展, 2007,26(3):88−92.

7 淡战平,候义斌.面向应用级的纯软件双机热备份机制设计与实现.计算机工程与应用, 2000,56(6):104−105,140.

8 南英,陈士格,戴冠中.容错控制进展.北京:航空工业出版社, 1993:62-67.

9 Flavin Cristian. Understanding fault-tolerant distributed systems. Comm. of ACM, 2001,34(2):57-58.

10 Nam HC, Kim J, Hong SJ, Lee SG. Probabilistic Check pointing IEICE Transactions on Information and Systems, 6(17), June 2002:1093−1104.

11 姚耀文,刘希挥,王作新.双计算机容错系统的故障诊断模型研究.华南理工大学学报(自然科学版)，1999,78(7):39−44.

12 朱幸辉.大规模事务处理监测系统的研究与实现.长沙:国防科技大学, 2005.

13 简新红.Corba component model 系统监控与管理的设计与实现.长沙:国防科技大学, 2005.

14 Michael N.Nelson, Brent B. Welch, et al. Caching in the Sprite Network File System. ACM Transactions on Computer Systems, 1988,6(1):134−154.

(上接第 99 页)

Telecomunication System. Washington, DC: IEEE Computer Society, 2000: 28.

10 Almeida V, Bestavros A, Crovella M, De Oliveira A. Characterizing reference locality in the WWW. Proc. of the 1996 4th International Conference on Parallel and Distributed Information Systems. Washington, DC: IEEE Computer Society, 1996:92−103.

11 Otoo E, Olken F, Shoshani A. Disk Cache Replacement Algorithm for Storage Resource Managers in Data Grids. Proc. of the 2002 ACM/IEEE Conference on Supercomputing. Los Alamitos: IEEE Computer Society, 2002. 1−15.

12 Cao P, Irani S. Cost-Aware WWW Proxy Caching Algorithms. Proc. of the 1997 Usenix Symposium on Internet Technologies and Systems (USITS-97). Berkeley: USENIX Association, 1997:193−206.

13 O'Neil EJ, O'Neil PE, Weikum G. The LRU-K page replacement algorithm for database disk buffering. Proc. of the 1993 ACM SIGMOD International Conference on Management of Data. New York: ACM, 1993: 297−306.

© 中国科学院软件研究所 http://www.c-s-a.org.cn