国 防 科 技 大 学 学 报
第 22 卷 第 6 期 JOURNAL OF NATIONAL UNIVERSITY OF DEFENSE TECHNOLOGY Vol.22 No.6 2000

文章编号：1001-2486（2000）06-0057-03

强实时双系统中容错技术研究*

李宏亮[1]，胡国平[2]，胡华平[1]，赵龙[1]

（1．国防科技大学计算机学院，湖南 长沙 410073；2．总装备部测通所，北京）

摘 要：在关键任务部门中，实时双系统的应用越来越广泛。如何保证高的可用度和高的响应时间是它们共同面临的问题。在本文中，全面论述了强实时双系统容错技术中的关键问题，包括系统结构定义、故障检测、状态切换。实际应用的结果表明，方案可以满足应用系统高可靠、强实时的需求，取得了良好的效果。

关键词：实时；双系统；故障检测；状态切换

中图分类号：TP311 文献标识码：A

A Study of Fault-Tolerant in Hard Real Time Dual System

LI Hong-liang[1]，HU Guo-ping[2]，HU Hua-ping[1]，ZHAO Long[1]

（1．College of Computer，National Univ．of Defense Technology，Changsha 410073，China

2．The Institute of Measurement and Communication，Beijing）

Abstract：The hard real time dual systems are applied in crucial tasks widely．The common problem they face is how to guarantee the high availability and the high response time．A full discussion of the key problems in hard real time dual system is presented，including the definition of system architecture，fault-detection and state-transition．The practice shows that the solution can meet the requirement of high availability and hard real time，and the achievement is outstanding．

Key words：real time；dual system；fault detection；state transition

实时系统是计算机应用的一个重要方向，已经广泛应用于航天、军事、工业控制以及电子商务中。同时，在这些实时系统中，任何微小的差错都可能导致重大的经济损失，因此如何保证高的可用度是他们共同面临的问题[1] [5] [7]。

在以前的大多数研究和实践中，系统的容错采用的是三模或者多模容余。但是，随着计算机性能的不断提高，单模块的平均故障时间在延长，这样，使用三模或者多模容余会带来比较高的代价，甚至有些多余[2] [3]。因此，在许多的实时应用中，越来越多地采用双系统的结构。因此，双系统容错技术的研究有着重要的现实意义。

我们研究的实时系统的特点为：实时性要求高（响应时间和处理时间）；需要不间断地进行处理，可用度要求高；处理的对象主要为短事务。

根据系统的特点，下面给出了实时系统的一个通用的模型。在这个模型中，我们将一个事物的处理分成多个步骤，每个步骤用一个逻辑处理器（Logic Processor LP）来进行处理，每一个处理步骤的输出是下一个处理步骤的输入，从而构成一个宏流水结构。我们把一个完整的处理流程称为一个机组。系统的逻辑示意图如图 1 所示：

实际上，多个逻辑处理器可对应到一个物理处理器上，但是，机组和机组之间必须使用不同的物理处理器。各个不同的物理处理器之间可用高速网络或者反射存储器（reflective memory）来实现紧耦合。在通常情况下，两个机组采用同源输入，系统的输出结果采用一定的策略从两个机组中选取。

* 收稿日期：2000-06-23
基金项目：95 国家预研重点资助项目（15.2.3）
作者简介：李宏亮（1975-），男，博士研究生。

1 故障检测技术

故障检测是双系统的最基础的一步。对于实时系统来说，故障检测应该满足：及时准确地定位故障，占用系统开销小。根据实际系统的特点，我们把系统的故障分为：

(1) 偶然故障：由偶然事件引发的程序执行过程中出现的难以重复出现的故障

(2) 永久故障：系统中重复出现的故障。又可以分为

· 硬件故障：硬件设备故障

· 软件故障：程序设计中出现的逻辑错误

· 崩溃

我们采用了以下的检测方法：

(1) 系统状态区。为了及时检测各个逻辑处理器的状态，需要设立全局的状态区。其结构如图2所示。

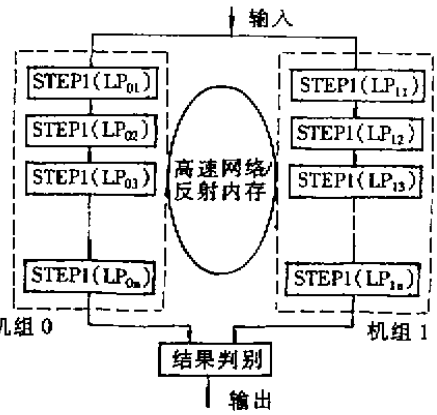

图1 应用系统逻辑示意图

Fig.1 The Logic figure of application system

(2) 检查点检测和心跳检测相结合。心跳检测指的是在一段时间间隔向外广播自身的状态（通常为“存活”状态）并且检查其它节点的“存活”状态。心跳检测的检测时间和检测的间隔时间是心跳检测的关键问题。设置的间隔时间过于频繁，将会影响到系统的正常运行，占用系统资源；而设置的间隔时间太长，则检测会比较迟钝，影响故障检测的及时性，最终影响到整个系统计算的正确。同时，由于定时器的处理占用的系统资源比较多，因此应该尽量避免定时器中断的产生。

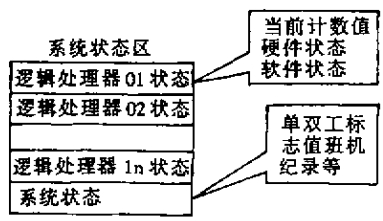

图2 全局系统状态区示意图

Fig.2 The figure of system status area

在我们的短事务实时系统中，我们采用了定时检测和动态检测相结合的方法来设置心跳检测的间隔时间，即设置的心跳检测的间隔时间为一个定时间隔 t_1 和处理一定数量事务的时间 t_2 最小值。

当事物频繁到达时，$t_1 > t_2$，这时定时器不会激活。而当事务比较少时，定时器的激活不会影响到系统的性能。实践表明，这种方法很好地解决了这个问题。

(3) 自检。系统运行一段时间以后，可以在系统负载比较轻的情况下，采用一些标准子程序来进行测试。

(4) 结果比较。结果比较是双系统中重要和关键的部分。系统运行是否正确，容错性能的好坏都与它密切相关。结果比较是判断双系统中是否存在故障的主要措施。但是，由于当双系统的结果比较不一样时，难以确认故障的位置，因此，结果比较是双系统中的关键部分，同时也是争议比较大的地方。在我们的系统中，我们采用下面的方法来判断故障的位置：

· 复执：当结果比较不相同时，两个机组都将该事务重新提交到各自第一个逻辑处理器的输入端，重新进行处理，这样，同一个事物具有多个处理结果，从而可以通过少数服从多数来确定正确的结果和故障节点，通过复执，可以检测出所有的偶然故障和部分永久故障。

· 质量报告：使用过程控制质量的原理，在每个逻辑处理器处理之后，附加上一个质量报告，对本阶段的处理质量进行评估。从而，可以对每个事务给出一个综合的质量报告。在结果比较阶段，当结果不相同时，可以根据质量报告等级和软硬件状态来决定取舍[5]。

2 故障恢复技术

对于分布式系统来说，常用的故障恢复技术可以分为前向恢复和后向恢复。

(1) 前向恢复。前向恢复技术指的是系统从故障中恢复时，从出错时刻以后的某一时刻点开始恢复。采用前向恢复时，事务的执行时间要比后向恢复短，任务的完成时间可预测性强。但是，选择恢复的时机通常比较困难，状态恢复困难。在实时短事务双系统中，由于处理的对象是短事务，事务和事物之间通常是独立的，我们采用将故障节点的状态恢复成工作节点的“切入”方法，达到前向恢复。

(2) 后向恢复。后向恢复技术指的是系统从故障恢复时，退回到以前的某一个状态，重新开始处理。后向恢复需要设置检查点，在每个检查点处保存节点状态，以便进行恢复时是用于恢复状态，检查点设置的时间间隔和保存的状态是后向恢复中的一个重点。同时，后向恢复需要避免出现多米诺效应。对于实时短事务双系统来说，我们将检查点设置为每个（或者若干个）事务处理结束处；当结果比较不相同时，我们将两个机组恢复到前一次保存的状态，并将事务重新构造出来，达到状态恢复的目的。由于事务的处理时间短，事物之间是独立的，而且两个机组之间的通讯仅仅限于结果比较，因此，可以避免多米诺效应。实践说明这种方法是有效的。

3 切换技术

双系统的一个关键问题是机组或者节点的切换。切换应该保证事务的连续处理，不丢失、不阻塞事务，事务处理不超时。

(1) 切出。当一个节点出现永久故障时，我们需要将该机组从系统中脱离。从而系统由双工变成单工。由于单工时系统的可用度降低，因此应该尽量避免切出，并且应该尽快恢复成双工，以保证系统的可靠运行。使用切出和切入操作时，采用的是前向恢复的方法，故障节点恢复时的状态从正常节点获得，因此切出操作不需要备份状态，相对比较简单。

切出时机选择：某一个处理节点崩溃（“心跳”停止），或者某一个机组出现永久故障，或者接收到控制台发出的切出命令。

(2) 切入。当一个机组修复完成后，加入到系统中，并且将系统状态恢复到双工的过程称为切入。切入过程中最重要的是保证两个机组状态的一致。由于系统中事务的到达是连续的，而每个事务的处理都可能改变机组的状态，因此，在切入操作时，需要让工作机组暂停处理，当两个机组状态一致之后，再同步处理事务。因此，我们采用了双缓冲队列的方式，取得了明显的效果。具体工作流程略。

采用这种方案，可以保证服务不会丢失，并且一些静态数据，或者可以从数据库中得到的数据，可以在“切入就绪”之前完成，而真正开始切入后，再恢复动态的关键的数据，从而可以使得切入操作对系统的影响最小。

4 结束语

强实时双系统广泛地应用于许多关键部门和关键任务。本文论述了关键技术，能够满足系统的强实时性、高可用、服务不断流的要求，较好地解决了双系统中通常会遇到的结果判别的问题。这个解决方案已经成功地应用到了重大工程实践中，取得到良好的效果。

参考文献：

[1] 金士尧，胡华平，李宏亮. 具有容错结构的高可用计算机双系统研究 [J]. 中国工程科学，1999，1 (3).

[2] D I Heimann et al，Dependability Modeling for Computer System [C]，1991 Proc. of RAM，1991.

[3] L Tomek et al，Reliability modeling of Life-Critical，Real-Time Systems [C]，Proceedings of the IEEE，1994，82 (1)：108 ~ 121.

[4] 李宏亮 等. 高可用实时系统中故障检测及故障恢复技术的研究 [J]. 计算机工程与科学，1999，21 (6).

[5] 胡华平，金士尧，李宏亮. 高可用、强实时分布式系统的切换技术研究 [J]. 国防科技大学学报，1999，21 (6).

[6] Deconinck G. User-Triggered Checkpointing and Rollback in Massively Parallel Systems，1996.

[7] 金士尧等 . 强实时高可靠的群机系统设计与论证 [J]，计算机工程与科学，1997，19 (A1)：1 ~ 5.