短事务、强实时双机容错系统的研究

李宏亮 金士尧 胡华平 王志英
(国防科学技术大学并行与分布处理国家重点实验室 长沙 410073)

摘 要 在军事、工业控制以及电子商务系统中存在着大量的高可用、短事务、强实时应用.在这些应用中,采用双机系统具有较高的性能价格比,如何保证双机系统的强实时性、高可用度和服务"不断流",是其中的关键技术难题.文中着重论述了系统可用度、故障检测、结果判别和状态切换中的关键问题.在理论的指导下,给出了实现策略和实际测试数据.测试数据表明本方案完全满足系统的要求,并且在具体工程实践中得到了应用,取得了明显的效果.

关键词 短事务;强实时;容错;双机系统;状态切换
中图法分类号: TP302

A Study of Short Transaction, Hard Real Time Dual Systems

LI Hong-Liang JIN Shi-Yao HU Hua-Ping WANG Zhi-Ying
(*National Laboratory of Parallel and Distributed Processing, National University of Defense Technology, Changsha* 410073)

Abstract There are many short-transaction applications, with requirements of high-availability and hard real time, in military, industry and electronic commerce. The dual systems are used in these applications with high performance-cost ratio. How to ensure the hard real time, the high availability and the "non stop" service are key problems in dual systems. A full solution, which can meet the requirements of the dual systems, is proposed in the paper. The system availability, fault detection, result comparison and state transition are emphasized in the paper. Firstly, the architecture of dual systems and the concepts such as physical processor, logic processor are defined. The system availability of dual systems with cut in success ratio is proposed also. Secondly, the fault classification is defined and the algorithms of "non fixed interval" fault detection, the results comparison, the system cut-in are proposed in the paper. These techniques can ensure the high efficiency, the synchronization between dual systems, and the report of faults rapidly. These means which can be applied to reduce the cut-in time are proposed with analysis. The solution has been applied in an important national project. At last, the test data fetched from that project are displayed in the paper, which also shows that the achievement of the solution in the project is outstanding.

Keywords short-transaction; hard real time; fault tolerance; dual systems; state transition

1 引 言

分布实时系统是计算机的一个重要应用方向,而其中的短事务实时应用也成为一个突出的应用分支,广泛应用于军事、工业控制以及日益兴起的银行、电信和与 Internet 相关的电子商务.在这些分布短事务实时系统中,事务的到达非常频繁,并且不允

收稿日期:2000-06-27;修改稿收到日期:2002-06-28.**李宏亮**,男,1975 年生,博士,主要研究方向为计算机系统结构、容错技术、复杂系统仿真等. E-mail: hlli@nudt.edu.cn. **金士尧**,男,1935 年生,教授,博士生导师,主要研究方向为计算机仿真、容错技术等. **胡华平**,男,1964 年生,博士后,副研究员,主要研究方向为计算机性能评价、网络安全等. **王志英**,男,1956 年生,博士,教授,博士生导师,主要研究方向为计算机体系结构、并行计算.

许丢失任何事务，任何微小的差错都可能导致重大的物质经济损失，因此如何保证高的可用度和高的系统处理能力是他们共同面临的问题.

这些短事务、强实时系统的特征有：

(1) 实时性要求高. 事务的处理时间短，一般为 0.1s 量级；响应时间有限制，一般为秒级；事务的到达率高，一般为 50～200 次/s；

(2) 需要不间断地进行处理，不允许出现事务的丢失或者“断流”；

(3) 可用度要求高，通常要求全系统可用度在 99.99%以上.

在以前大多数的容错研究和实践中，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相对会带来高的代价，甚至显得有些多余. 因此，在许多实时应用中，越来越多地采用双系统的结构. 如在计算机容错领域中处于领先地位的 Compaq 公司的 Trucluster 系统①～③和喜马拉雅(Himalaya，原天腾公司 Tandem Co.)系列④，HP 公司的双机双控容错系统方案⑤以及美国 LEGATO(VINCA)公司服务器群集(CLUSTERING)容错系统 Co-Standby Server⑥等，其关键结构均采用双系统或者双进程/处理器/IO 等的热备份方式，在通过高速网络连接和磁盘阵列进行互联的基础上，实现了通用的基于磁盘阵列的软硬件容错系统.

但是这些系统多数采用热备份甚至冷备份的双系统工作方式，工作机组的故障检测时间比较长(根据应用和故障类型的不同，时间有所不同)，通过实际应用和测试，当系统出现故障时，系统的双机/主备机切换时间都在 10 秒级，甚至数分钟，导致系统处理的“断流”. 因此，热备份或者冷备份的系统结构在可用度、保证系统的连续处理等方面，难于满足短事务强实时系统的需求. 双系统容错技术的研究仍有着重要的现实意义.

2　短事务容错双系统的体系结构

根据以上的分析，采用了双工的工作方式来满足短事务、强实时系统的需求. 根据系统的特征，下面给出了短事务、强实时系统的一个通用的模型. 在这个模型中，根据实际系统的情况，将一个事务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，通常情况下，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，在这里，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.

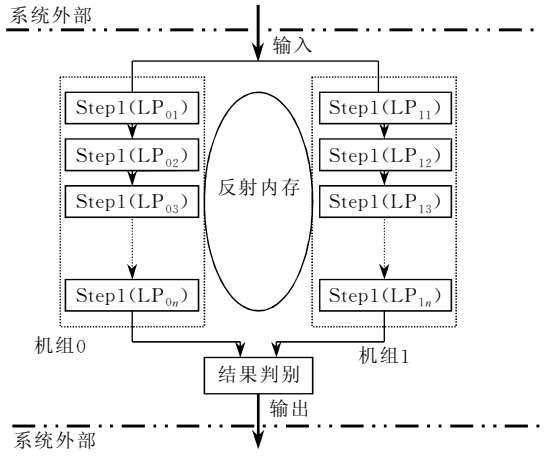


图 1　短事务、强实时系统逻辑示意图

实际上，多个逻辑处理器可以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，以保证基本的容错需要. 各个不同的物理处理器之间可以用反射存储器(reflective memory)来实现紧耦合. 通常情况下，要求容错双系统的结构对外是透明的，这样，系统对外只能有单个输入和单个输出. 因此，两个机组采用的是同源输入，系统的输出结果采用一定的策略从两个机组中选取. 实际上，由于并没有一个独立的处理结点来进行结果判别，因此两个机组都需要进行结果判别，为了保证只有一个机组输出，需要使用值班机(in duty)标志.

从上面的双系统结构图可以直观地看出，要满足系统的实时性、可用度和不断流的要求，除了提高单个逻辑处理器的软硬件本身的可靠性和可用度以外，还严重地依赖下列因素：系统发现故障的能力，包括监测时间和成功率；系统状态的转换能力，包括转换时间和成功率. 根据文献[1]的分析，在考虑了双机的切换问题(包括切入成功率，与此相关的切入时间和再次切入的时间以及故障判别问题)后，完整的短事务、双系统的稳态可用度为[1]

$$A_{\text{double}}^{(\infty)} = \left(1+\frac{2\lambda}{\mu}\right)\Big/\left(1+2\,\frac{\lambda(1-D)}{\alpha'}+2\,\frac{\lambda}{\alpha}+2\,\frac{\lambda}{\mu}+2\,\frac{\lambda^2}{\mu^2}+\frac{2(1-C)\lambda}{\beta}\right) \tag{1}$$

① http://www.tru64unix.compaq.com/cluster/.
② Compaq Co. TruCluster 5.1 Technical Overview, 2001.
③ Compaq Co. TruCluster 5.1 Highly Available Applications.
④ http://himalaya.compaq.com.
⑤ http://www.hp.com.
⑥ http://www.legato.vinca.com/.

其中,λ 为平均失效率为平均无故障时间($MTBF$)的倒数,β 为故障诊断率,为平均诊断时间的倒数,μ 为平均维修率,为平均维修时间($MTTR$)的倒数,α 为切入失效率,为平均切入时间的倒数,C 为故障判别率,α' 为再次切入失效率为再次切入时间的倒数(重启双工时间的倒数),D 为切入成功率.

经典的双系统可用度计算,就是在 $D=1,\alpha=\alpha'=\infty,C=1,\beta=\infty$ 的特殊条件下的可用度.这是双系统可用度的极限:

$$A_{\text{double}}^{(\infty)} \mid_{\substack{C=D=1\\ \alpha=\infty,\alpha'=\infty}} = \left(1+\frac{2\lambda}{\mu}\right)\Big/\left(1+\frac{2\lambda}{\mu}+\frac{2\lambda^2}{\mu^2}\right) \tag{2}$$

采用同构双系统,在典型值的计算中,可以获得 99.99995%的可用度.考虑故障判别成功率 $C=0.5$,且 $\beta=1/(30\ \text{min})$,则该双工系统的可用度约为 99.997%.进一步考虑双工切换的成功率 $D=1/(30\ \text{min})$,则该双工系统的可用度约为 99.997%.进一步考虑双工切换的成功率 $D=0.5$ 且 $\alpha=1/(1\text{s})$,则它的可用度只有 99.993%.因此,在双系统工程中,必须重视实现故障判别成功率和切入成功率的有关技术设计.

故障判别成功率的高低,依赖于故障检测机制的快速性、准确度和结果判别的准确度;而切入成功与否,依赖于切入的时机、切入的时间等因素.

3 故障检测与结果判别

3.1 故障检测

故障检测是双系统中最基础的一步,对于实时系统来说,故障检测应该满足及时准确的定位故障、占用系统开销小以提高故障判别的成功率.根据实际系统的特点,系统故障可以分为[2,3]:

(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.

① 硬件故障.硬件设备故障.

② 软件故障.程序设计中出现的逻辑错误.

③ 崩溃.

在短事务、强实时系统中,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.检查点时间间隔的设置如图 2 所示.

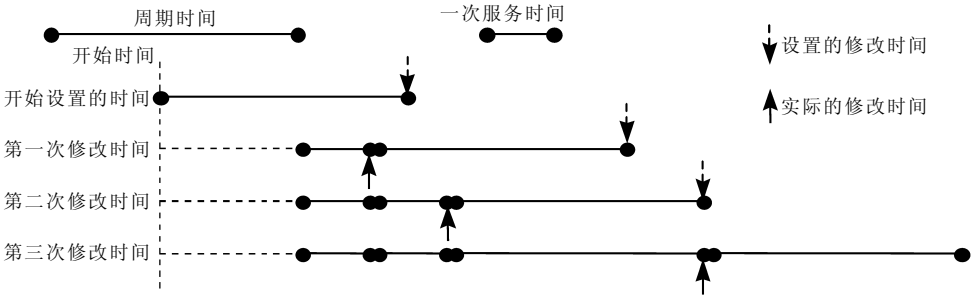


图 2 检查点时间设置示意图

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.

由于结果比较是短事务、实时双系统中任何事务处理都需要经历的步骤,同时,也是系统中事务处理的一个同步点,因此,我们把故障检测放到结果判别部分进行.

3.2 结果判别

对于短事务、实时双系统来说,结果判别的作用可以分为两个方面:

(1) 进行结果选择,保证系统的可用度;

(2) 检测机组故障.

因此,结果判别的优劣直接关系到整个系统运行结果的正确性,同时,由于结果判别是系统的一个同步点,因此,也成为系统的一个潜在的瓶颈,影响到系统运行性能的高低.因此,结果判别应该满足:

(1) 高效;

(2) 保证两个机组的同步,防止双机组的失步;

(3) 及时准确报告故障.

通常,可以把结果判别阶段分为结果配对和结果选择两个阶段.

在结果配对阶段,需要将两个机组对于同一个

事务的处理结果进行匹配，由于两个机组的环境以及系统的一些无法确定的因素，结果配对时一个机组的结果通常需要等待另一个机组的结果，对于短事务、实时双系统来说，每个事务的处理时间是一定的，不可能无限期地等待，因此，需要在每个事务流入系统时，记录时戳，进行超时判断. 当出现事务超时时，应该单工发送结果，保证事务能够按时从系统中流出，并且将单工发送的事务编号记录到单工日志中，防止超时的另外一个结果重复发送，同时，还需要将将超时机组报告监控系统，以便监控系统收集信息，做出决定.

结果判别是双系统中关键的部分. 系统运行是否正确，容错性能的好坏都与他密切相关. 结果判别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，双工系统的结果判别成功率取决于结果比较的错误判断依据. 在我们的系统中，采用了过程控制质量的原理[1]. 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. 在结果判别阶段，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. 整个结果判别算法如下所示：

```
输入：本机组事务处理结果缓冲区(Local Buffer)，它机组事务处理结果缓冲区(Remote Buffer)
输出：发送缓冲区(Send out Buffer)
start：
  while(TRUE)
    取系统时间 TIMEnow；
    i=0，j=0；
    while(i<Length of Local Buffer)
      if (Local Buffer 中的第 i 个事务 Ti 有效)
        提取事务 Ti 进入系统的时间 TIMEi，事务编号 Ni，处理结果 Ri 和质量报告 QRi；
        while(j<Length of Remote Buffer)
          if (Remote Buffer 中的第 j 个事务 Tj 有效 and 事务编号=Ni)
            提取事务 Tj 处理结果 Rj 和质量报告 QRj；
            //Result Comparing
            if (Ri=Rj and Induty)
              发送到 Send out Buffer
            else
              QRbetter=max(QRi，QRj)；
              QRbetter degrade
              发送到 Send out Buffer
            end if
          end if
        end while
        if (TIMEnow-TIMEi>THESHOLD)
          if (Ni<单工发送日志)
            cancel；
          else
            发送到 Send out Buffer
            记录到单工发送日志
          end if
        end if
      end if
    end while
  end while
```

4　状态切换

双系统的另一个关键问题是机组(或者结点)的切换，状态切换主要是指双工系统变成单工系统或单工系统切换成双工系统. 前者称为切出，后者称为切入. 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，并且保证切入后系统的同步运行.

4.1　切　出

当一个结点出现永久故障时，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，故障结点恢复时的状态从正常结点获得，因此切出操作不需要备份状态，相对比较简单.

切出时机：

(1) 某一个处理结点崩溃(“心跳”停止)，或者

(2) 某一个机组出现永久故障，或者

(3) 接收到控制台发出的切出命令.

4.2　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内部的状态，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. 实现方案如图 3 所示.

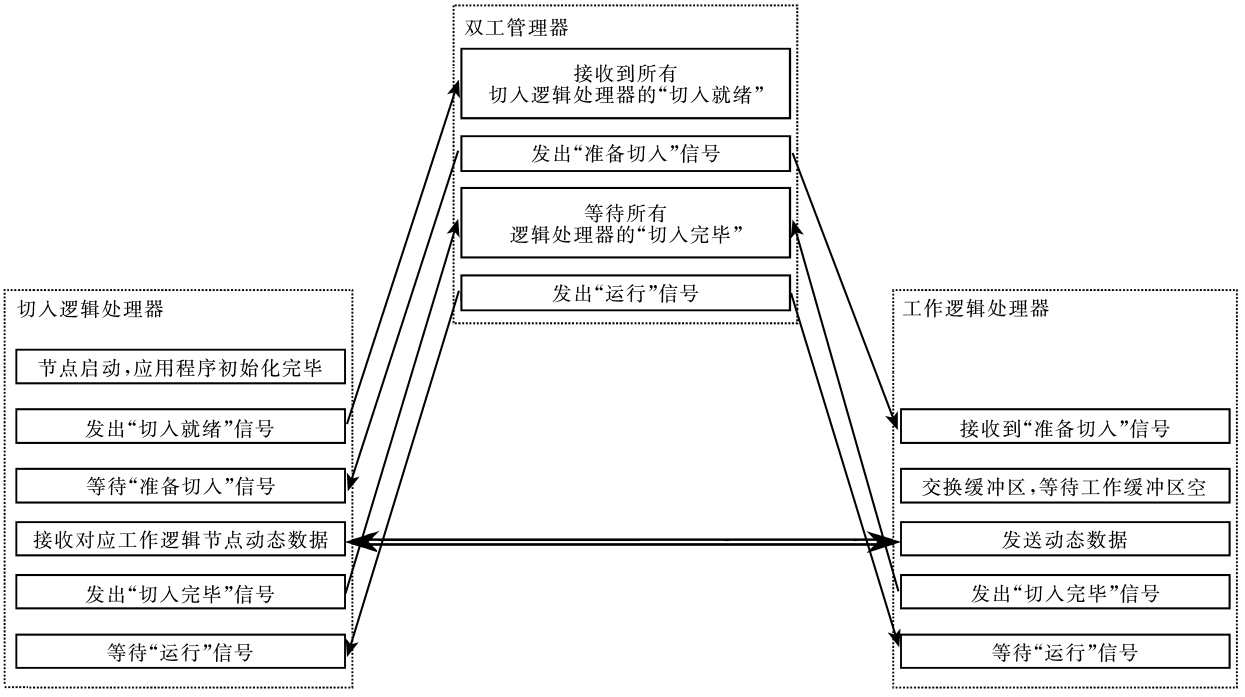


图 3 双系统切入流程示意图

采用这种方案,可以保证服务不会丢失,并且静态数据可以从数据库或其它途径中获得,在“切入就绪”之前完成,而在开始切入后,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.

切入时机

假设:

切入时刻 t_0

最佳切入时刻 t_1 $\}t_1-t_0=t_a$,选择最佳切入时延;

迁移现场时刻 t_2 $\}t_2-t_1=t_b$, 迁移现场时延.

其中,选择最佳切入时延 t_a 与切入时刻以前进入系统而尚未处理的请求服务数量 n 以及平均服务时间 $\overline{t_s}$ 有关,取平均值有

$$\overline{t_a}=\bar{n}\cdot\overline{t_s} \tag{3}$$

迁移现场时延 t_b 与迁移的现场与大小以及迁移现场的单位传输时间 t_c 有关,通常现场中包含两部分,一部分现场是时间函数 f_t,另一部分是非时间函数 $f_{\bar{t}}$,与时间无关的现场可以在切入时刻之前先期准备完毕.所以,

$$\overline{t_b}=f\cdot\overline{t_c}=(f_t+f_{\bar{t}})\cdot\overline{t_c}\xrightarrow{f_{\bar{t}}=0}f_t\cdot\overline{t_c} \tag{4}$$

因此,切入所需的时间为

$$t=\overline{t_a}+\overline{t_b}=\bar{n}\cdot\overline{t_s}+f_t\cdot\overline{t_c} \tag{5}$$

当系统进行切入时,工作机组和切入机组都会暂停工作,因此,切入时间的长短,直接关系到切入操作是否会造成事务处理的超时或者丢失.要减小切入操作的时间,可以采取以下措施:

(1) 减少 f 的大小.通常现场中包含两部分,一部分现场是时间函数 f_t,另一部分是非时间函数 $f_{\bar{t}}$.与时间无关的现场可以在切入时刻之前先期准备完毕.

(2) 选择 f_t 较小的时刻.当处理宏流水线中没有事务时,系统处于相对稳定的阶段,需要移动的现场数据比较小.

(3) 减小 t_c 的时间,即增加传输速度.在使用反射存储器来通信时,在通常的通信方式中,发送方需要传送完所有的数据时,接收方才可以开始接收,造成比较大的传输延迟.因此,引入了滑动窗口的方式,对数据分片同时传送,隐藏通信延迟.对于比较大的数据,这种传输方式的引入,取得了明显的效果.

(4) 减小 $\overline{t_s}$.可以通过增加处理器的数量或者处理能力来减小事务的平均处理时间.

(5) 选择 n 和 K 比较小的时候.由于系统的负载是有一定波动的,因此,也可以选择负载比较轻的时候来切入,以避免事务的丢失.当 n 和 K 比较大时,切入操作会停止执行.

5 性能分析

5.1 实际应用介绍

本方案已经应用到了国家重点工程的信息处理中心,并已连续运行两年多,在遇到突发事件,甚至

修改应用程序时，都能够通过双系统的切换来保证系统不停机、不丢失事务. 在实际运行过程中获得了许多测量数据.

5.2 双系统切换性能

对系统的切换时间进行了测试，切入逻辑处理器需要从数据库和硬盘中读取约 3.2G 的数据，需要恢复的动态数据约为 2MB，事务的到达率为 100 次/s. 此时的测试结果列于表 1.

表 1　系统切换性能

测试项目	切入时间 (ms)	切出时间 (ms)	值班位交换 (ms)	切入准备时间 (min)
测试 1	212	<1	<1	20
测试 2	213	<1	<1	21
测试 3	246	<1	<1	19
测试 4	268	<1	<1	20
测试 5	256	<1	<1	20

5.3 单双工处理能力比较

由于单工时，结果判别部分可以认为不工作，即基本不会影响到系统的性能. 当系统处于双工状态时，需要进行结果判别. 单双工处理能力的比较实际上是对双工结果判别对系统性能影响的测试. 当系统出现事务丢失时，认为系统达到了处理的极限，测试结果列于表 2.

表 2　单双工处理能力比较

测试项目	双工处理极限(次/s)	单工处理极限(次/s)
测试 1	230	240
测试 2	225	230
测试 3	220	230
测试 4	230	235

从表 2 可以看出，双工处理对系统性能影响增量小于 4%.

5.4 切入操作对系统处理能力的影响

在实际系统中，平均事务处理时间 t_s 为 40ms，需要动态传输的数据 f_t 约为 1MB，反射内存的单位传输时间为 $t_c=0.025$s/MB，每个事务的最大处理时间 t_{max} 为 360ms，系统共有 8 个逻辑处理器，这时(根据实际应用的特点，测试时产生的事务服从均匀分布)：

(1)当事务的到达率为 100 次/s 时，切入操作不会出现事务的丢失情况，输入和输出的事务数量相等.

(2)当事务的到达率为 150 次/s 时，切入操作会引起事务处理超时.

从测试数据可以看出，切入操作对系统的性能有一定的影响，在切入时，需要选择负载比较轻的时刻进行.

6 结束语

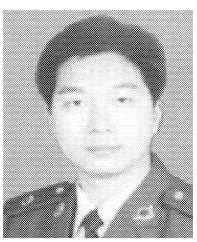
短事务、强实时系统将会广泛地应用到社会的许多关键部门和关键任务中去，本文针对这种系统，提出了一个完整的解决方案，方案能够满足系统的强实时性、高可用、服务不断流的要求，较好地解决了双工系统中通常会遇到的结果判别的问题，而且在进行状态切换时，能够保证服务“不断流”，双工处理占用的系统开销小，基本不会影响到系统的处理能力. 测试结果表明，方案完全能够满足此类系统的需求. 并且这个解决方案已经成功地应用到了重大工程实践中，取得了良好的效果.

参 考 文 献

1 Jin Shi-Yao, Hu Hua-Ping, Li Hong-Liang. A study of a high-availability dual computer system with fault tolerant architecture. Engineering Science, 1999, 1(3):46～50(in Chinese)
(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. 中国工程科学, 1999,1(3): 46～50)

2 Li Hong-Liang, Wen Mei, Zhang Chun-Yuan *et al*. Research of the fault-detection and fault-recovery technique in high availability real-time systems. Computer Engineering and Science, 1999, 21(6): 57～59(in Chinese)
(李宏亮，文　梅，张春元等. 高可用实时系统中故障检测及故障恢复技术的研究. 计算机工程与科学，1999，21(6)：57～59)

3 Deconinck G. User-triggered checkpointing and rollback in massively parallel systems[Ph D dissertation]. Katholieke University Leuven, Belgium, 1996



LI Hong-Liang, male, born in 1975, Ph. D.. His research interests include computer architecture, fault tolerance.

JIN Shi-Yao, male, born in 1935, professor, Ph. D. supervisor. His research interests include computer architecture, fault tolerance technology, complex systems simulation.

HU Hua-Ping, male, born in 1964, Ph. D.. His research interests include computer performance evaluation, network security.

WANG Zhi-Ying, male, born in 1956, Ph. D., professor. His research interests include computer architecture, parallel computing.