分布式工作流系统的可靠性研究*

陶 冶　范玉顺　罗海滨
清华大学自动化系，北京 100084

摘　要　工作流管理系统为企业中不同部门之间的协作以及不同应用之间的集成提供了一个功能强大的计算环境。在当前工作流技术的研究领域中，分布式工作流系统逐渐成为研究的热点，而系统的可靠性是分布式工作流系统成功实施的关键因素之一。本文介绍了相关的领域概念、系统结构、过程实例的执行以及全局控制机制，提出了一种基于企业功能领域划分的、具有很高可靠性的分布式工作流系统的设计和实施方案。

关键字　可靠性；工作流；分布式工作流系统；CORBA

一、　引言

工作流技术作为现代企业实现过程管理与过程控制的一项关键技术，它为企业的经营过程提供了一个从模型分析、建立、管理、仿真到运行的完整框架。同时，工作流管理系统通过一套集成化、可互操作的软件工具为这个框架提供了全过程的支持。经过十几年的发展，工作流技术已逐渐走向成熟，至今已被应用于银行、电信、医疗保健、生产制造等诸多领域。

但是，也必须看到，工作流技术无论在理论上还是在技术本身还不够成熟，应用范围还不够广泛。研究人员则从工作流应用系统的角度指出了目前所存在的性能方面的不足[1]，例如各工作流系统间彼此几乎都不兼容，无法胜任大规模业务，负荷能力有限，系统可靠性、可扩展性和灵活性差，而且不能提供强有力的安全保证，这主要是因为单一数据库、集中式结构、有限的通讯能力以及缺乏前瞻性的设计等原因所致。

提高工作流管理系统整体性能的一个重要策略是工作流机的分布，各分布工作流机的协作将使系统从结构上变得更加合理，它是提高工作流管理系统可靠性和可扩展性的关键之一。本文将介绍一种基于企业功能领域划分的分布式工作流机的协作模式和实施方法，提出一系列增强系统可靠性的策略，以努力解决以往集中式工作流机所固有的一些问题。

二、　企业的功能领域

在企业CIMS应用工程的实施过程中，需要首先根据企业自身特点，利用通用功能视图模型的基本构件生成对企业需求分析阶段的功能模型。在CIMOSA体系结构中[4]，功能视图反映了企业的功能结构，从上到下，依次为领域（Domain）、领域过程（Domain Process）、经营过程（Business Process）、企业活动（Enterprise Activity）和原子级的功能操作（Function Operation），其中的领域是一个很重要的概念。根据企业的经营目标，企业可以划分为若干个互相正交的领域，每个领域都有其目标、功能及其领域过程，领域的功能由领域过程实现。功能视图中领域的概念正是“面向客户”、“面向市场”的新的经营观念的体现，它使得企业的划分不再是传统的“面向职能”的陈旧模式。领域的划分和确定由企业经营目标和所受到的限制条件的集合决定。不同的领域将实现不同的企业目标，彼此之间通过事件、消息相互联系与协调，从而为企业建立新型的运行机制奠定了基础。

根据过程视图与功能视图之间的交互特性以及企业功能领域划分的思想，我们将把工作流管理系统中的核心组件工作流机与企业的建模过程有机结合，使运行时的系统能够与企业模型紧密集成。为此，需要从两个方面进行新的设置：首先，在建立每一个可执行的过程模型时，针对于最基本的活动单元，将它们与适当的功能领域相关联，这可以通过在建模工具

中添加功能领域的属性来实现，有关功能领域的信息则由企业的功能模型提供；其次，对于每一个运行时的工作流机，也同样为它们绑定不同的功能领域，这可以在注册新的工作流机时完成设置。不同领域的工作流机将负责执行不同的活动，包括激活相关的应用、生成任务表、对后继活动进行导航等。通过这种针对功能领域的绑定，企业中的每一个任务都将与某一个分布的工作流机发生联系，同时，每一个工作流机的职责也随之变得清晰和明确。

三、 分布式工作流机的体系结构

按照WfMC提出的工作流参考模型[2]，一个工作流系统包括过程定义工具、工作流机、工作流管理工具、工作流客户应用和工作流机直接调用的应用等功能模块。

同时，由于在实际应用中，工作流系统往往在不同的硬件平台、操作系统、网络协议、数据库的异构环境下运行，这对参考模型中各模块间的相互通讯和协作提出了很高的要求，CORBA 规范保证了这些要求得以实现，为系统运行提供了一个软件平台。

基于上面的考虑，现提出本系统的体系结构：

整个系统是由过程建模工具、模型仿真工具、一个总控工作流机、一个工作表管理器和多个执行工作流机组成，其中的工作表管理器实际上是一个具有专门功能的执行工作流机。

就执行机、工作表管理器和总控机的关系来说，它们构成了一个 CORBA 平台上的圆锥形模型，总控机位于锥顶，各执行机位于锥底的圆周上，工作表管理器位于锥底的圆心上。总控机能监控工作表管理器和任何一个执行机，工作表管理器和执行机向总控机报告自己的状态，工作表管理器记录和维护各执行机所执行的任务。从系统级别上说，各执行机彼此之间的地位完全对等，只要需要，它们任何两两之间都可以建立联系。它们之间的关系可以形象地用图 1 来表示：

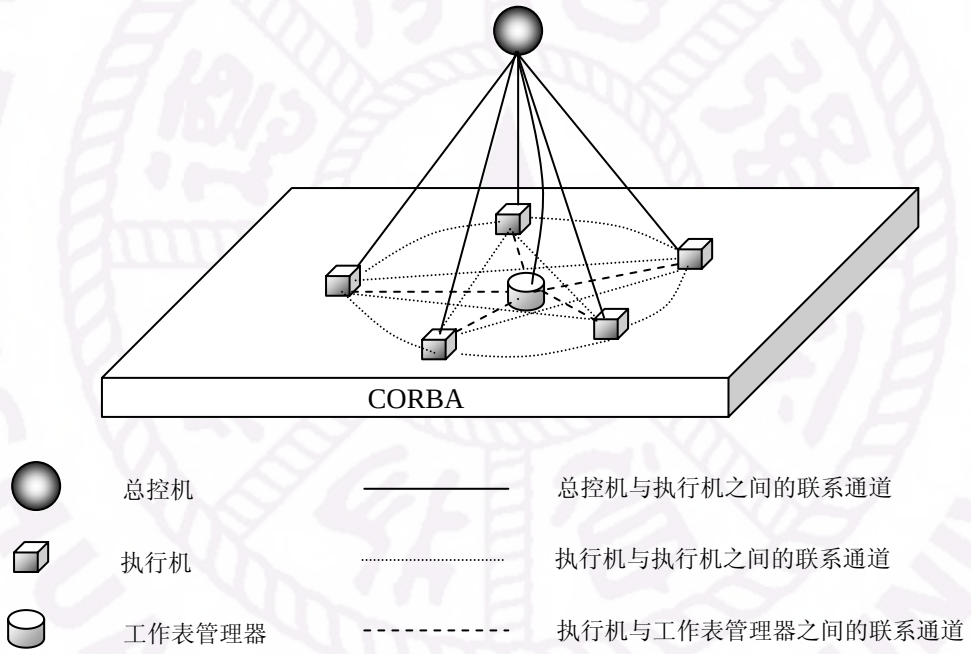

图 1　系统模型

建模工具负责创建工作流模型。在建模时，应对模型中的各个活动都绑定相应的功能领域。同时，各活动的输入输出、活动间转移的条件、非相邻活动间的数据连接都已事先定义。模型建立后，在正式投入运行之前，须经过模型仿真工具的静态仿真和动态仿真，以检验其可行性并预估其运行效果。

总控工作流机与存储过程模型的中央数据库位于同一台机器上，主要负责指导建模工具

进行模型分片、维护系统中各执行工作流机的配置信息、创建并激活过程实例、监控实例的运行状态等。执行工作流机则负责生成活动实例、激活应用、在活动之间进行导航、与工作表管理器交互等。工作表管理器则主要负责人工型活动任务项记录的生成、任务状态的更新与维护，这需要用户的参与。

整个系统采用基于 CORBA 规范的 Orbix Daemon 和 Orbix for Web 软件总线及其服务为系统支撑环境，系统提供的所有服务都经过 IDL 的定义与封装，经编译后以 CORBA 对象的形式出现。这些对象遵循 IIOP 协议，能够接受嵌于 Web 的客户端的调用请求，正因如此，系统中所有的用户界面、管理员界面都是基于 Web 的。图 2 给出了系统的体系结构，并演示了一个简单的过程实例的运行流程。

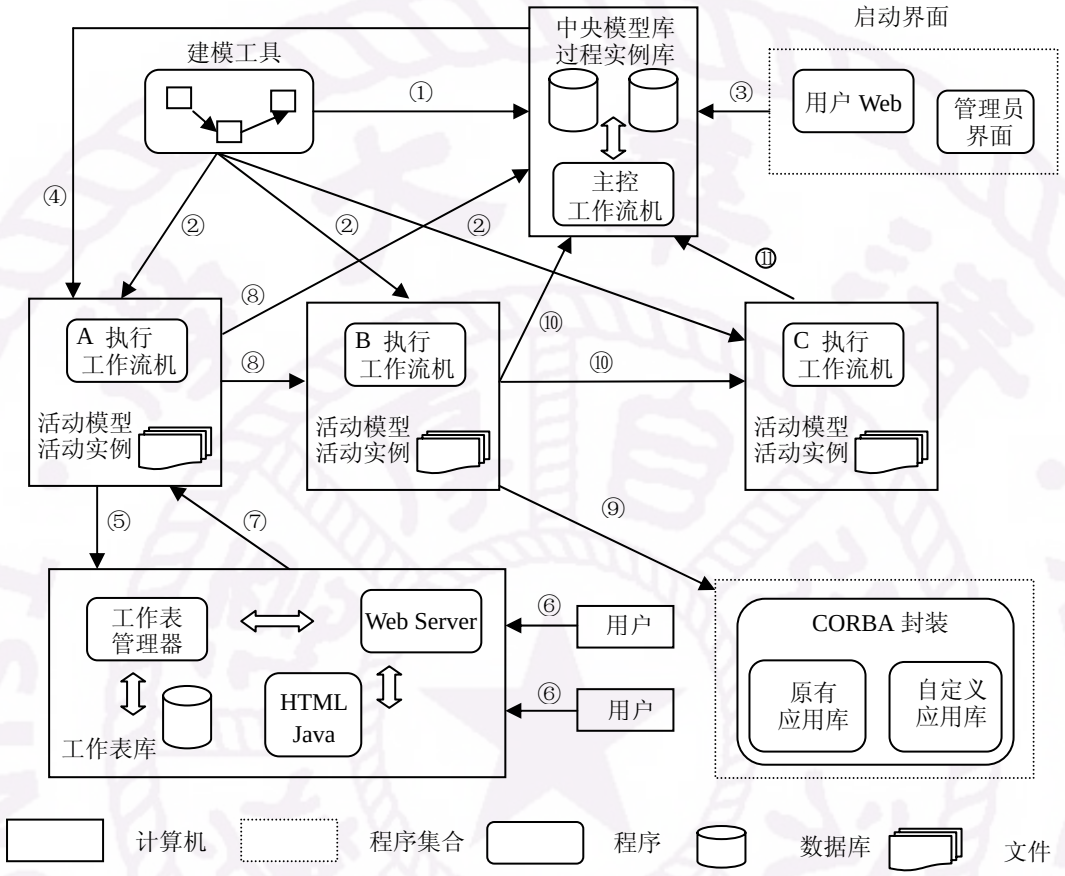

图 2　系统体系结构和运行实例

具体的运行步骤解释如下：

（1）通过建模工具，一个由三个活动组成的过程模型被建立，并被保存到中央数据库中；

（2）过程模型被分片传递到三个领域的工作流机上；

（3）该过程被某一启动界面启动，首先由总控工作流机生成一个过程实例，然后通过查找数据库中的模型信息及有关工作流机的配置信息，获得负责执行第一个活动的工作流机的主机地址；

（4）调用工作流机 A 的接口函数，由 A 生成新的活动实例；

（5）该活动是一个人工型活动，因此，A 将激活工作表管理器在数据库中生成相应的工作项，工作项的分配方式是采取用户自主的 PULL 类型；

（6）用户访问自己的工作表，并执行任务；任务完成后，用户通过同样的界面进行提交，系统更新工作表；

（7）工作表管理器通知执行工作流机 A，该活动实例已经完成；
（8）工作流机 A 通过本地的模型分片文件及配置信息，通知工作流机 B 执行过程实例中的第二个活动，同时通知主控工作流机它已经完成第一个活动；
（9）执行工作流机 B 调用其他应用程序来完成第二个活动，这是一个 CORBA 封装的自动应用；
（10）类似地，工作流机 B 通知工作流机 C 来执行第三个活动，同时通知主控工作流机它已经完成第二个活动；
（11）C 通知主控工作流机，该过程实例已经执行完毕。

在系统能够提供工作流服务以前，必须按照一定的顺序进行启动以建立工作流运行环境，通常是先启动总控工作流机，然后启动工作表管理器和各个执行工作流机。系统运行时各执行工作流机均处于待命状态。企业可以根据自身的情况，按照实际要求来配置执行工作流机，比如每个领域或者每个部门配置一个。

图 2 中 CORBA 封装的应用包括系统自带的原有应用库和用户自定义的应用库，这些应用都是一些诸如打印文档、发送邮件、定时等功能单一且固定，从而可重用度较大的应用。这给用户提供了一个开放的接口，用户可以按系统提供的规范接口编写或设定他们用得较多的一些特殊应用。由于这些应用以软件组件的形式出现，在以后可以对之进行扩充和修改，从而可以提高系统的柔性和可重用性，便于企业信息系统的不断改进和重组。

四、 系统的可靠性研究

在实际运行时，系统可能会面临各种各样的错误或异常，这些错误或异常可能源于工作流管理系统本身，也可能源于一些外部因素(如：人工错误、网络或硬件异常等)。一个好的工作流管理系统应该确保存在一个事先定义好的框架或机制来处理错误或从错误中恢复。因此，可靠性成为分布式工作流系统性能的一个重要指标。

可靠性是指在规定的运行环境下、在规定的时间周期内，系统正常运行并正确执行所要求的功能的概率。系统可靠性强意味着外部环境或内部状态出现变化或异常时，系统仍能够完成其预先设定的功能，而不至于使系统停止运行甚至崩溃。

提高系统的可靠性，也就是要提高系统处理异常的能力。异常通常可以分为系统异常和逻辑异常。系统异常是指由于系统结构设计不合理，或者由于不可预料的意外情况而导致工作流不能正常推进；逻辑异常是指由于程序中的逻辑缺陷而导致一定条件下工作流出现死锁或非正常终止等异常情况。

五、 提高系统可靠性的策略

本系统针对预防错误和出错恢复两方面，在设计时采取了预防性和补偿性两方面的措施来处理异常，预防性措施包括避错措施和查错措施，补偿性措施包括容错措施和纠错措施。避错和容错主要针对系统异常，查错和纠错主要针对逻辑异常。图 3 为本系统的可靠性设计框架：

1. 避免使总控机成为系统的核心

总控机就是对整个分布式工作流系统的配置和运行情况进行总体监控的服务器，它负责指导工作流模型的分片、过程模型的实例化、过程入口的导航、系统数据库的更新和维护等重要工作。在系统设计时，如果将所有对保证系统正常运行来说非常重要的工作都由总控机来完成，则一旦总控机出现故障而不能启动或运行，则整个系统必然会陷入瘫痪状态 。因此，本系统在设计时坚持的一个重要思想就是弱化总控机对工作流运行的导航和干预作用，避免它成为整个系统不可或缺的核心。

总控机在系统开始运行之前需要根据各执行机的功能领域将模型的各个活动分配至相应的执行机，创建过程实例并导航至过程实例的入口活动所在的执行机。模型的分片通过分片文件来实现，该文件保存了相应活动的前驱和后继活动中与导航相关的信息。一旦系统进

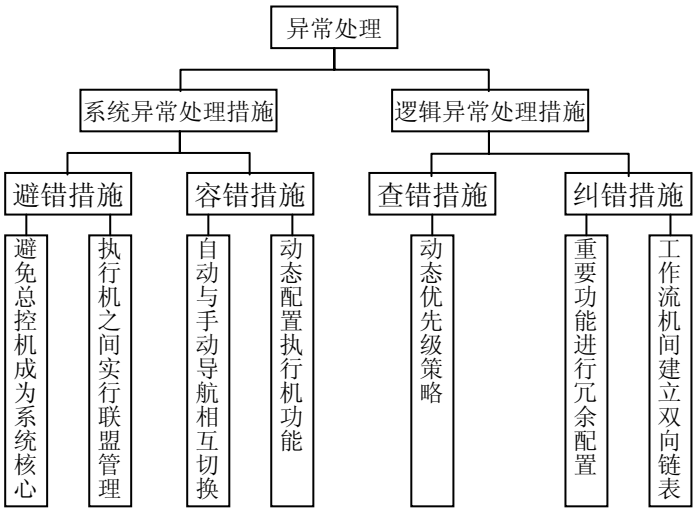

图 3　系统可靠性设计框架

入正常运行之后，则执行机和总控机之间只存在单向的信息传递通道，即执行机向总控机报告本身状态的改变，而总控机并不向执行机下达任何指令。在这一点上，执行机是完全自治的，它在活动间的导航完全通过保存在执行机本地的实例文件来进行。因此，一旦系统进入正常运行状态之后，即使总控机出现故障甚至停止运行，也不会影响工作流的推进。正是分片文件使执行机具备了自主导航的能力，从而在保证系统整体性能的前提下避免了总控机成为系统的核心，提高了系统的可靠性。

2.　系统运行时在执行机之间实行联盟管理

由于在模型中为每一个活动和执行机都指定了所属的领域，因此一旦工作流模型建立后，则此工作流将会涉及到哪些执行机在系统运行之前就已经确定了。从系统的角度来说，这些所涉及到的执行机就临时结成了一个联盟，当工作流全部运行结束之后，这个联盟也就随之解散了。在这个联盟中，系统将根据一定的规则，比如以工作负荷或工作性质为标准，自动推选出一个盟主，该盟主自己本身既是一个执行机，同时又负责对本联盟内部事务的管理。可以说总控机是系统级的管理层，盟主是功能级的管理层。

首先，盟主将保存本联盟在运行期间的所有全局数据。在工作流运行期间，各执行机之间的信息联系方式包括传递和共享，相当于 Email 方式和 BBS 方式。全局数据很多时候包括大量的文档，将全局数据保存在盟主所在的执行机上时，避免了数据在不必要环节上的传递，不但降低了网络的负担，而且给数据管理带来了很大的方便，如数据一致性的维护、数据访问和操作权限的管理等在传递方式下难以处理的问题在共享的方式下变得很容易。

其次，盟主在运行期间将协调各执行机之间的冲突。在工作流运行时，各执行机之间有时难免会出现时序互锁或资源共享上的冲突，若不进行协调，则系统可能会陷入停滞状态，此时，盟主根据一定的协调机制，在发生冲突的执行机之间采取某种措施，比如强令某一执行机先挂起或将任务转移到另一冲突之外的执行机去执行，来保证工作流的顺利推进。

另外，盟主在系统运行出错时将提供出错恢复所需的数据。

3.　实行动态优先级策略

在实际生产和业务流程中，各执行任务都有一定的执行优先级，但在整个流程中各优先级不是固定不变的，执行环境、执行时间和资源状况的变化，或意外情况的出现等因素通常

会导致优先级提升和下降。若工作流系统中各活动的优先级固定不变，则可能会由于应该立即执行的活动没立即执行而导致工作流推进出现异常，至少是延缓了工作流推进的进度。

因此，本系统采取了动态优先级的策略。以活动的执行期限为例，有些活动有一个 Deadline 属性，表示该活动必须在某一时间点之前完成。在工作流运行期间，系统会检测各活动的 Deadline，当某个尚未执行的活动快要达到它的 Deadline 时，系统会提升它的优先级，并将这一信息通知它的前驱活动，前驱活动的优先级也因之而得到提升，由此上溯，直至找到选择分支出现的地方，并提升分支入口处活动的优先级。当工作流推进到该选择分支时，刚才这一分支会优先执行。对于其它如环境、资源等变化，系统同样根据一定的原则来改变活动的优先级。通过这种方式，系统可在一定程度上预测出可能发生的异常，并采取相应的措施来尽量避免这种异常的出现。

4. 动态配置执行机的功能

在系统中，每个执行机都配置在一台计算机上，也就是配置了一个可向建模工具、总控机、工作表管理器和用户提供服务的 CORBA 对象。同时，系统中的原有应用或用户自定义的应用，比如打印服务、传真服务、定时器服务、计数器服务、电子邮件服务等等，都通过 CORBA 封装成一个个独立的服务实体。这些实体通过 Orbix Server Manager 进行注册，配置到执行机所在的计算机里。从本质上看，虽然执行机和上面这些应用都是系统提供的服务对象，但可以认为这些应用是外挂在执行机下的，因为在系统运行中导航到某台执行机后，用户就能获得这台执行机所能提供的所有服务的一个清单。这里实际上已将执行机的概念虚化了，因为它本身并不提供某种具体的面向用户或者面向业务的服务，它起到的作用就是对系统所提供服务的检索和管理，同时保证系统的正确导航。

有了这种功能的动态配置作保障，若在工作流运行过程中某一执行机因为故障不能执行某任务，这时可很方便地将该任务移植到另一执行机上，系统仍可以正确导航，这保证了工作流的正常推进，实际上就提高了整个系统的可靠性。

5. 系统提供自动导航与手动导航相互切换的功能

在前面的讨论中，所有的工作流都是自动导航的。自动导航有其本身的优点，但若在运行中因为某种不确定性因素而导致系统不能正确导航，则工作流就会因之而中断。出于这点考虑，系统提供了自动导航与手动导航相互切换的功能。它并不消除自动导航所需的信息，只是将系统中各执行机的配置信息提供给用户，由用户按需要选取其中的某一个或几个执行机，使系统改变自动导航的轨迹，而将任务转移至用户指定的执行机上执行。在必要时，用户还可将工作流切换回自动导航，这种切换可以增强系统的可靠性。

6. 运行中建立工作流机之间的双向链接，以提供可靠的出错恢复

系统可靠性的一个很重要的方面体现在系统的纠错能力上。在实际生产和事务处理中，有时会因后面执行的任务出错而需要返工。本系统为了支持这种返工而在执行机之间建立双向链接，以备在系统出错时提供可靠的过程回滚，可回滚的最大步数由系统管理员设定。

执行机之间的双向链接将在系统运行时建立，也就是每个执行机完成自己的任务后导航到后继执行机，这时它同时要将自己的配置信息通知其所有后继执行机，每个后继执行机将这些信息保存在其前驱执行机列表中。同时，每个执行机在执行任务之前都会缺省地将当前条件和一些重要数据进行备份，以备出错恢复，当然用户也可选择不备份。对保存在盟主所在执行机上的全局数据和文档而言，需要建立数据和文档的备份队列，队列长度等于可回滚的最大步数。当系统需要回滚时，可以从队列中可靠地恢复历史数据和文档。

7. 重要功能在执行机上进行冗余配置

在一个工作流中一般存在一些重要的或最常用的功能，这些功能应保证其在需要时能可靠地执行，这一目标可以通过将它们在不同执行机上进行冗余配置来达到。

在工作流推进过程中，当执行机导航至其后继执行机时，它将设定一个检测故障的定时

器。若在定时期间内它一直没收到后继执行机的反馈，则它认为后继执行机出现了故障，而将工作流导航到冗余功能所在的那台执行机上去。从概率论的角度来说，若不考虑各执行机发生故障的相关性，则多台执行机同时发生故障的概率远远小于一台发生故障的概率，因此，通过功能冗余一般可以保证任务的正常完成。

需要说明的是，系统进行导航时会将各执行机的当前工作负荷作为一个因素考虑，以决定导航的方向，使各执行机的工作负荷尽可能平衡，以减轻工作流运行当中的瓶颈现象。

六、 结论

本文所提出的分布式工作流系统的方案，按照对企业功能模型的需求分析与设计说明，工作流管理系统将在过程定义与执行这两个层次上对企业的功能领域提供支持。在定义过程时，将最基本的活动单元与适当的功能领域相关联；在执行过程时，不同的活动将由不同领域的工作流机来执行与导航。同时，系统从避错、查错、容错和纠错四个方面采取了多种措施来提高系统的可靠性。这种分布式工作流系统的方案不仅可以解决原有集中式工作流管理系统所存在的一些瓶颈，而且能够与企业的建模过程紧密集成，从而形成一个完整的、可靠的、从建模到实施的过程体系，能较好地满足实际生产和业务流程的要求。从开发出来的原型产品来看，本系统具有很高的可靠性。

参考文献

1 范玉顺，吴澄，“工作流管理技术研究与产品现状及发展趋势”， <<计算机集成制造系统 CIMS>>，Vol.6，No.1，Jan.，2000，pp.1-7

2 Haibin Luo, Yushun Fan, “CIMFlow: A Workflow Management System Based on Integration Platform Environment”,7th IEEE International Conference on Emerging Technologies and Factory Automation,1999，pp.233-241

3 范玉顺，吴澄，王刚，高展，“集成化企业建模方法与工具系统研究“，《计算机集成制造系统-CIMS》，Vol 6，No.3，2000，pp.1-5

4 Alonso G, Agrawal D, Abbadi El A, et al. Functionality and Limitations of Current Workflow Management Systems. IEEE Expert, 1997,12(5).

5 WfMC. Workflow Management Coalition Terminology and Glossary (WfMC-TC-1011). Technical Report, Workflow Management Coalition, Brussels, 1996.

6 Nina Edelweiss, Mariano Nicolao. Workflow Modeling: Exception and Failure Handling Representation, Computer Science, SCCC '98. 1998

7 Dong-Soo Han, Jae-Yong Shim. Design and Implementation of a Distributed Transactional Workflow System, IEEE TENCON, 1999

Research on the Reliability of Distributed Workflow System

Tao Ye, Fan Yushun, Luo Haibin

Department of Automation，Tsinghua University，Beijing 100084

Abstract Workflow management system (WfMS) has been used to provide a powerful computation environment for the integration of different applications and the cooperation of different organizations in enterprises. In the current research field of workflow technologies, the distributed workflow engines have gradually caught more and more attentions, and the reliability

of the whole system is one of the critical factors to the successful implementation of the distributed workflow system. In this paper, a distributed workflow system with high reliability that runs on different enterprise function domains is introduced. Several measures to improve the reliability of the system are proposed. Meanwhile, the relevant concept of domain, the system infrastructure, the implementation of a process instance and the global control mechanism are also presented.

Key words Reliability；Workflow；Distributed Workflow System；CORBA

作者简介：

陶　冶　男，清华大学自动化系国家 CIMS 工程技术研究中心硕士研究生。98 年本科毕业于清华大学自动化系。现从事工作流及相关集成技术的研究。

范玉顺　男，教授，博士生导师。90 年博士毕业于清华大学自动化系，93～95 年获洪堡奖学金在德国柏林工业大学从事合作研究。国际自动控制联合会先进制造技术委员会委员，863/CIMS 重大关键技术攻关专家组成员，青岛海洋大学兼职教授，已完成多项 863、国家自然科学基金项目，发表论文 80 余篇，完成专著两本。现在清华大学自动化系国家 CIMS 工程技术研究中心工作。

罗海滨　男，清华大学自动化系国家 CIMS 工程技术研究中心博士研究生。97 年本科毕业于清华大学自动化系。现从事工作流及相关集成技术的研究。