

中间件技术研究

张莉萍¹, 邵雄凯²

(1. 武汉科技大学中南分校信息工程学院, 武汉 430223; 2. 湖北工业大学计算机学院, 武汉 430068)

摘 要: 中间件技术屏蔽了底层分布式环境的复杂性和异构性, 简化了分布式应用程序的开发, 使分布式应用程序的健壮性、可扩展性、可用性更好。本文介绍了中间件的定义、分类、主流中间件技术平台, 能够隐藏底层网络环境的复杂性, 处理平台之间的异构性, 是分布式计算和系统集成的关键组件。本文介绍了中间件的基本概念, 阐述了一种基于系统层次的分类方法, 并详细讨论了现存各类中间件技术的特点、相关标准及发展应用情况。

关键词: 中间件; 应用集成; 分布式计算

1. 引 言

中间件 (Middleware) 是处于操作系统和应用程序之间的软件, 也有人认为它应该属于操作系统中的一部分。人们在使用中间件时, 往往是一组中间件集成在一起, 构成一个平台 (包括开发平台和运行平台), 但在这组中间件中必需要有一个通信中间件, 即中间件=平台+通信, 这个定义也限定了只有用于分布式系统中才能称为中间件, 同时还可以把它与支撑软件和实用软件区分开来^[1], 如图 1 所示。它隐藏了计算机体系结构、操作系统、编程语言和网络技术等方面的异构性, 将复杂的协议处理、分割的内存空间、数据复本、网络故障、并行操作等问题与应用程序隔离开来, 为上层应用软件提供运行与开发的环境, 帮助用户灵活、高效地开发和集成复杂的应用软件。中间件已成为构建现代分布式应用、集成系统的一种不可或缺的成分, 数据访问、企业资源规划、多媒体、无线/移动计算、电子商务、网格等都是它的典型应用领域。

2. 中间件分类

中间件可分为六类。

(1) 终端仿真/屏幕转换: 用以实现客户机图形用户接口与已有的字符接口方式的服务器应用程序

之间的互操作。

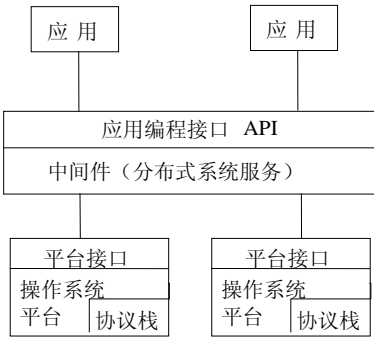


图 1 中间件的位置

(2) 数据访问中间件: 是为了建立数据应用资源互操作的模式, 对异构环境下的数据库实现联接或文件系统实现联接的中间件; 在分布式系统中, 重要的数据都集中存放在数据服务器中, 它们可以是关系型的、复合文档型、具有各种存放格式的多媒体型, 或者是经过加密或压缩存放的, 这类中间件将为在网络上虚拟缓冲存取、格式转换、解压等带来方便。

(3) 远程过程调用中间件: 通过这种远程过程调用机制, 程序员编写客户方的应用, 需要时可以调用位于远端服务器上的过程。

(4) 消息中间件: 用来屏蔽掉各种平台及协议之间的特性, 实现在不同平台之间通信, 实现分布

【作者简介】

张莉萍 (1980-), 女, 硕士, 讲师; 研究方向: 嵌入式系统、数据库。
邵雄凯 (1976-), 男, 博士, 教授; 研究方向: 数据库、软件理论。

式系统中可靠的、高效的、实时的跨平台数据传输，实现应用程序之间的协同。这是中间件中唯一不可缺少的，是销售额最大的中间件产品，主要产品有国内东方通科技公司的 TongLINK、BEA 公司的 BEA eLink、IBM 公司的 MQSeries 等，目前在 Windows 2000 操作系统中已包含了其部分功能。

(5) 交易中间件：是在分布、异构环境下提供保证交易完整性和数据完整性的一种环境平台。在分布式事务处理系统中要处理大量事务，常常在系统中要同时做上万笔事务。在联机事务处理系统 (OLTP) 中，每笔事务常常要多台服务器上的程序顺序地协调完成，一旦中间发生某种故障时，不但要完成恢复工作，而且要自动切换系统，达到系统永不停机，实现高可靠性运行；同时要使大量事务在多台应用服务器能实时并发运行，并进行负载平衡地调度，实现昂贵的可靠性机和大型计算机系统同等的功能，为了实现这个目标，要求系统具有监视和调度整个系统的功能。根据 X/OPEN 的 DTP (Distributed Transaction Processing) 模型规定，一个分布式交易处理系统应由事务处理、通信处理以及资源管理三部分组成。BEA 公司的 TUXEDO 便是最著名的一个交易中间件产品，东方通科技公司的 TongLINK 和 TongEASY 实现了 DTP 参考模型规定，另外还有 IBM 公司的 TXSeries 是应用广泛的一个交易中间件产品。

(6) 对象中间件：在分布、异构的网络计算环境中，可以将各种分布对象有机地结合在一起，完成系统的快速集成，实现对象重用，在这个方面遵循的标准是 CORBA。对象中间件将是未来的主流，目前产品如东方通的 TONG BROKER，INPRICE 公司的 Borland Application Server。

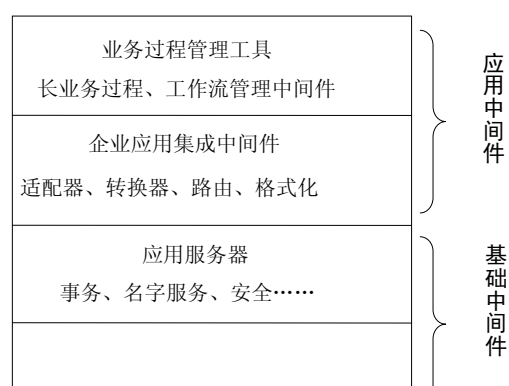
在客户/服务器模式向多层结构迁移的过程中，中间件技术得到了迅速发展和广泛的应用。传统中间件通常提供通信和事务处理功能，包括数据访问中间件、远程过程调用、面向对象、组件的中间件，基于消息的中间件等多种类型，相关技术规范也日趋成熟。而目前，随着电子商务的兴起和企业信息化建设的深入，中间件已不再仅仅实现客户机服务器的互连，还实现应用之间的互操作，也就是数据和应用的集成。

从系统层次的角度，可以把当前的中间件软件归纳为四类^[2]，从下至上依次为传统通信中间件、应用服务器、企业应用集成中间件和业务过程管理工具，如图 2 所示。其中前两种又合称为基础中间件，因为它们通常支撑单个应用系统或解决单一类问题，与平台软件，如操作系统、数据库管理系统等，结合紧密；后两种属于应用中间件的范畴，更多用于系统整合，与行业应用系统联系紧密。应用中间件通常会与多个应用系统打交道，位于系统中较高的层次，并大多基于底层中间件运行。

传统中间件发展时间较长，作为基础平台软件，具有一定的通用性，实现技术也较为复杂。该类别中囊括了多种实现机制，相关标准也比较完备。由于是高层中间件运行的基础，因而仍然占据主体地位。应用服务器在企业信息化建设中得到了广泛应用，是设计、开发、部署、运行和管理 Internet 电子商务的平台。企业应用集成中间件帮助企业集成各种新老应用程序，使它们能够跨平台互通，在企业资源规划中扮演着重要角色。正是应用服务器和企业应用集成中间件的出现改变了中间件的作用，使其从一个联结分布式应用各组成部分的通信工具，变成了一个集成工具，从而也提升了中间件的地位和重要性^[3]。据预测，未来几年增长率最高的中间件将集中在应用服务器和应用集成中间件上。业务过程管理工具帮助企业管理业务活动，为过程的产出增值，这一类工具中最突出的是 workflow 管理中间件。

3. 通信中间件

传统通信中间件根据它们提供通信的双方又可以分为两类^[4]：一类是在主动的应用程序和被动的



资源(如数据库)之间提供交互的中间件;另一类则是为主动的应用程序相互之间提供交互的中间件。

3.1 应用到服务器

最常见且用得最多的是数据库中间件,它的主要功能是为访问本地或远程数据库定义并实现一个应用编程接口 API,这样从开发者的角度来看无论是操作系统还是数据库所在位置、运行平台都被该 API 屏蔽掉了,数据格式也在该 API 处进行必要的转换。通常这类中间件由各数据库厂商自己提供,不同的数据库产品有各自独立的接口。由于针对特定数据库,这种本地中间件往往具有最佳的访问性能,但其移植性较差,一旦更换数据库就要求程序作较大的改动。ODBC 和 JDBC 是为解决这一问题制定的数据库接口标准,它们本身并不是真正的数据库中间件,只是为各类数据库定义了一个统一的访问接口。用 ODBC 或 JDBC 调用数据库时,还需要后台数据库特定的驱动程序。ODBC 与 JDBC 功能相仿,前者受到 Windows 平台的限制,后者局限于 Java 语言,两者可以桥接互通。

事务处理监控器 TPM,又称为交易中间件,也是当前应用最广泛的中间件之一。它能支持数以万计的客户进程对服务器的并发访问,使系统具有极强的扩展性,因此适于电信、金融、证券等拥有大量客户的领域,在对效率、可靠性要求严格的关键任务系统中具有明显优势。TPM 一般支持负载均衡和服务组件的复本管理,支持分布式两阶段提交,保证事务完整性和数据完整性,并具有安全认证和故障恢复等功能。TPM 有同步调用、异步调用、管道通信、会话、广播、通知、队列、发布订阅等通信方式,能很好地满足应用开发的要求。著名的 TPM 如 BEA 的 Tuxedo, IBM 的 CICS。

3.2 应用到应用

为主动应用模块提供通信的中间件按照其实现机制,主要有三种:远程过程调用 RPC,对象请求代理 ORB 和面向消息的中间件 MOM。三种机制都已成熟,其中 ORB 和 MOM 仍然具有广泛的用途。

RPC 是最早出现的中间件形式,它能够隐藏平台的异构性,支持透明调用远程过程,为通信双方作数据传递和格式转换。但 RPC 是一种同步调用机制,要求双方同时在线,对延迟大、故障率高的网

络环境(如 Internet)不适合,而且编程模型级别较低。现在流行的基于 xml 的简单对象访问协议 SOAP 就使用了 RPC 机制。

ORB 为通信双方传递调用请求,负责编、解码参数并激活服务对象。它的好处在于提供面向对象的编程模型,具有位置透明性和平台无关性,并可为分布式应用环境提供多种基本设施,如名录服务、事件服务、生命周期服务、安全服务和事务服务等。它的缺点和 RPC 类似,即本质上仍是一种同步通信机制,系统容错能力低,扩展性不强。CORBA 是一个得到 800 多家厂商支持的、成熟的 ORB 标准,它除了平台无关性外还提供语言无关性,即基于 CORBA 中间件的分布式应用各模块可以分别用多种编程语言实现。国内外都有很多符合 CORBA 规范的中间件软件,最著名的是 IONA 的 Orbix 系列产品。不同厂商的 CORBA 产品可以通过 IIOP 协议互通。除 CORBA 外,微软的 COM+/.NET 技术和 SUN 公司的 J2EE 在 ORB 中间件中也有了一定的影响力。J2EE 提供了面向组件的模型,而且主要服务采用了 CORBA 体系结构;COM+/.NET 因为受限于 Windows 平台的特点而使其应用受到了一定的限制。

MOM 是基于报文传递的网络通信机制的自然延伸,它为不同平台、甚至使用不同网络协议的通信双方提供消息的存储和传递机制,使它们不需直接交互。MOM 是一种异步通信机制,因为没有同步建立过程,也不需要调用参数编、解码,所以效率较高。而且它是一种松耦合的组织形式,有更强的扩展性和灵活性,更适合建立企业级或跨企业的大规模分布式系统,特别是事件驱动的应用。它的缺点是不适合有实时要求的应用,另外从编程的角度看,抽象级别较低,容易出错,不易调试。MOM 可看作是实际需求和抽象等级间的一种折衷^[1]。现有大量的 MOM 软件,如微软的 MSMQ、IBM 的消息排队系统 MQ Series 等,它们一般都支持多种协议和操作系统。但是 MOM 的标准化程度不高,各厂家产品一般不兼容,这容易造成对特定厂商的依赖,影响系统的可移植性、交互性和灵活性。

4. 应用服务器

随着企业应用环境逐渐向 Internet 迁移,传统中间件面临的网络分布计算环境发生了很大变化,

受到诸如用户数量不可估计、大规模并发访问、分布式长事务、网络不提供 QoS 保证等挑战。应用服务器就是在这种面向 Internet 的 Web 计算环境下出现的新型中间件。

有人将应用服务器看作是 Web 下的当代事务处理监控器,但实际上除了 TPM 的可伸缩、可靠、高效等特性外,应用服务器还具有 TPM 没有的执行和管理应用逻辑、定位和实例化对象以及控制对象生命周期的能力^[5],它侧重有状态的服务,而且比 TPM 支持更标准、更开放的协议和 API,支持更大规模的用户。作为三层结构的中间一层,应用服务器被看作是设计、开发、部署、运行和管理 Internet 电子商务应用的平台,它的主要功能包括:提供事务性 Web 应用的跨平台执行环境,提供事务、安全、应用集成和 QoS 在内的一系列基本服务,以及一系列开发、部署、管理 Web 应用的工具。

应用服务器是当前发展最快的中间件类型之一,国内外都有相应产品,这些产品从实现机制上可以分为脚本、面向对象和对象三种模式^[5]。脚本模式技术简单,但扩展性、重用性低,而且缺乏集成遗留系统和事务处理能力;面向对象模式的特点是用面向对象语言编写脚本,可重用性较好,但缺乏相应规范,使用范围和可移植性受到限制;对象模式是最复杂的一种,其模块化程度高,支持组件重用,而且有相应的标准和规范。其中最突出的规范是 J2EE,它为应用服务器的实现提供了一个完整的底层框架,使得不同 J2EE 应用服务器上的应用组件可以互操作,移植性好,得到了广泛认可。当前多数应用服务器产品遵循的是 J2EE 规范。J2EE 应用服务器还有 2 个评测基准: J2EE 兼容性测试 CTS 和 J2EE 应用服务器可伸缩性和性能测试基准 ECperf。主流 J2EE 应用服务器,如 BEA 的 Weblogic, IBM 的 Websphere 都已通过了 CTS 测试。

5. 其他中间件技术

5.1 企业应用集成中间件

传统的企业应用程序通常专注于相对具体和复杂的业务工作,很少考虑互相之间的接口,而随着企业的发展和合并,为保护已有投资,迫切要求这些系统能够共享信息、协同工作。但这些信息系统一般都是历经一段较长时间、逐步发展起来的,

它们在组织结构、数据格式、通信协议等方面可能完全异构,要实现集成非常困难,即使利用企业资源规划(ERP)也不能完全解决问题^[3]。

在许多专门的集成方案因为集成规模等原因相继被淘汰之后,出现了企业应用集成(EAI)的概念。EAI 就是利用通用中间件(如 MOM、ORB)融合多种应用的集成方案^[6],基于中间件的方案可以通过封装应用软件、共享信息、协调工作流程等降低集成复杂度。企业应用集成平台是为实现 EAI 提供基本功能组件、服务和开发工具的基础设施,它扩展了中间件,减少了集成工作量,因此我们把它视为一种更高层次的中间件(企业应用集成中间件)。企业应用集成中间件将许多集成步骤自动化,例如数据路由、映射和转换^[3],使用应用集成中间件已经成为实现 EAI 的必要手段。

可用的企业应用集成中间件很多,最突出的如 Neon 的 MQIntegrator、Vitria 的 BusinessWare 等,它们在中间件市场上占据着越来越重要的地位,尤其被类似于银行、电信等拥有众多网点(子系统)的单位所需要。虽然 OMG 组织有一些与 EAI 有关的规范文档,但是无论是 EAI 自身还是应用集成中间件都没有统一标准,它们使用的技术与具体的应用对象有关。市场上的企业应用集成中间件产品也存在很大的区别,没有哪一种产品能单独提供完整的 EAI 解决方案。

典型的企业应用集成中间件一般完成数据集成和业务逻辑集成两种集成功能,数据集成使异构平台上的数据源可以互相访问和交换数据,而业务逻辑集成使不同的应用程序可以共享业务逻辑,从而使企业的业务活动能够跨平台执行。

不同的企业应用集成中间件通常在通信模式、集成方法、底层中间件和服务四个方面存在差异^[6]。通信模式、集成方法和中间件是 EAI 方案的核心。通信模式分为同步通信和异步通信,两者都广泛地应用于多种场合:同步通信意味着高度耦合,要求有可靠的网络设施,适合于发起方需要接收方发回应答结果和确认接收的场合,如访问数据库;异步通信不需要协同操作,耦合程度比同步通信低,当通信的目的只是为了转移信息时,适合选用异步通信,例如数据更新、事件通知等。集成方法分为消息传递和接口定义,两种都需要适配器来作为应用程序的接口:在消息传

递集成方法中, 通信消息既包括控制信息也包括数据, 它的耦合度较低, 应用灵活、容易扩展, 不足之处在于应用软件对消息的处理是不可见的, 除非有清楚的文档和服务说明; 在接口定义方式中, 接口规定了应用软件能调用的操作, 要处理的数据都通过接口传输, 这种方式要求有严格定义的接口, 易于排错和维护, 但可能难于改变和扩展。通信模式和集成方法都与采用的底层中间件有关, EAI 采用最多的基础中间件是 MOM 和 ORB, RPC、数据库中间件和 TPM 则相对使用的较少, ORB 因为其面向对象的特点, 非常适合于基于组件的集成。通常一个好的 EAI 结构会综合应用 MOM 和 ORB 技术以最大限度地解决问题。服务是基本通信和中间件功能的扩展, 重要的服务包括目录、生存周期、安全、数据转换、连续存储、事件通知等, 一个好的应用集成中间件应该支持大多数或所有这些服务。

5.2 工作流管理中间件

工作流管理中间件是一个人机结合的系统, 它通过定义、创建工作流, 以及协调、监督、管理工作流的运行(例如进度检查、工作量平衡等), 实现了企业业务处理中各项活动的合作、控制和通信的自动化。工作流管理中间件一般基于客户/服务模式, 提供图形化的工作流定义工具, 并越来越多地基于 Web 界面操作, 其使用集中在保险、银行、法律等人员密集型的办公环境。

工作流的标准化程度很高, 工作流管理联盟 WfMC 从定义、术语到客户端 API 和审计数据为工作流定义了一系列规范^[7], 使异构的工作流产品能互操作, 并与其他的信息服务(如电子邮件)集成。这些规范已被 130 多个成员接受, 包括工作流中间件厂商、用户和研究人員。

5.3 Web 服务

值得一提的还有 Web 服务, 这是新近流行起来的一种面向服务的分布式计算范式, 虽然还有待进一步完善, 但它的简单性、灵活性和开放性已经对现有中间件技术形成了很大冲击。

Web 服务的技术核心是基于 xml 和 HTTP 的三个标准: Web 服务描述语言 WSDL, 用于服务发布

的统一描述、发现和集成规范 UDDI, 以及简单对象访问协议 SOAP。通过自描述性的 xml 和广泛使用的 HTTP 标准, Web 服务提供了一种简便的方法, 将运行在 Intranet、Internet 上的应用集成在一起。相比于典型的 EAI 解决方案, Web 服务更便于开发、维护和使用, 松散耦合的集成方式使它更灵活, 开放的标准也使它更便宜。问世不久, Web 服务就得到了广泛的支持, 以 WebSphere、.NET 为代表的多种软件平台都开始提供 Web 服务开发工具。

此外, Web 服务还被称作“中间件的中间件”^[8], 这是因为传统的中间件技术虽然能够在一定条件下隐藏某些系统异构性, 但彼此之间却是异构的, 基于不同中间件的信息系统很难互联互通。而基于开放的 Internet 标准这一特点使 Web 服务有能力成为这些异构中间件之间沟通的桥梁。如图 3 所示, 各种技术, 包括异构的中间件系统, 都可以通过 SOAP 协议互通。

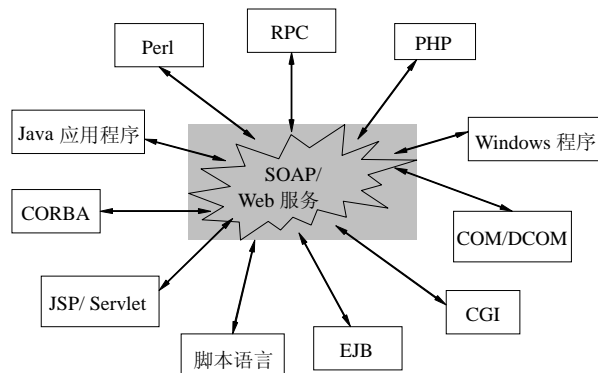


图3 各种技术通过 SOAP 互通

6. 结 论

中间件技术是构造分布式应用和系统集成的关键。从系统层次的角度, 可以把中间件划分为传统通信中间件、应用服务器、企业应用集成中间件和业务过程管理工具四类, 前两种合称为基础中间件, 后两种合称应用中间件。有一些中间件技术已

(下转第 49 页)