

Internet、Internet 软环境及其应用

Internet, its Software Environment and Application

杨 森¹ 徐泽同²

(北京理工大学计算机科学与工程系 北京100081)¹ (中国科学院数学所 北京100080)²

Abstract In this paper, internet, its software environment and application are introduced. Their current situation and research trend are discussed.

Keywords Internet, Internet software environment, Internet application, Grid

1. 引言

20世纪90年代, Internet 将地球连接成一个“信息村”。可以说, 操作系统、TCP/IP 协议族的结合构成了 Internet。处在 Internet 之上的程序设计语言、数据库、浏览器、分布式面向对象技术等等结合在一起构成了 Internet 软环境。Internet 上的应用是在 Internet 以及 Internet 软环境之上展开的。

早期的 Internet 应用主要限于科学和教育领域, 90年代进入金融、电信、CIMS、交通、能源等诸多领域。伴随世界经济全球化、信息化的步伐, 建设 Internet 之上的电子商务成为 Internet 应用的热点, 经济与 Internet 的结合成了 Internet 发展的动力。1992年后, 美国相继开展了元计算和网格(Grid)计算的研究。目前, 网格计算备受重视, 它将会极大地促进 Internet 应用的发展。它们间的依存关系如图1所示。

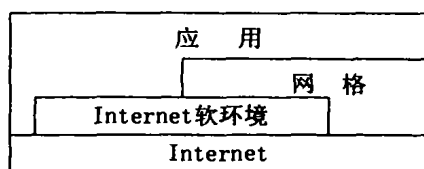


图1 Internet、Internet 软环境及应用

本文主要介绍 Internet、Internet 软环境及应用, 并探讨它们的现状和研究动向。

2. Internet

2.1 Internet 1

1969年, 美国 ARPA 建立起 ARPANET。在1977到1979年之间, 推出了 TCP/IP 体系结构和协议规范。在1980年前后, ADRPA 开始将 ARPANET 上的所有机器转向 TCP/IP 协议, 并以 ARPANET 作为主干来构建 Internet。1983年, 伯克利分校将 TCP/IP 协议族嵌入到 BSD UNIX 中, 实现了操作系统和TCP/IP的有机结合。1986年, 构建了NSFNET, 实现了与 ARPANET 的互连。随后, NSFNET 取代 ARPANET, 成了 Internet 的主干。

在整个 Internet 协议族中, TCP 和 IP 两大协议承上启下, 扮演着举足轻重的角色。由此, 我们把 Internet 网协议族叫做 TCP/IP 协议族。TCP/IP 协议屏蔽了底层物理网络的实

现细节, 向上提供一致、通用的网络数据传输手段, 构成了 Internet 网协议族的骨架和核心, 整个 Internet 技术都是围绕着它们而展开的。

2.2 Internet 2和 NGI

在1996年, 美国开始研制 Internet 2。同年, 美国又启动了 NGI(Next Generation of Internet, 下一代 Internet)计划。Internet 2的焦点是学术性研究的创新; NGI 却是建立起实验性的广域网络, 开发与测试关键性的应用。下一代互联网研究的重点包括: (1)高速主干, 着重研究 IP over DWDM 及第二代光纤网络; (2)高速接入, 如光纤、无线、微波、卫星等; (3)大容量 IP, 就是要从 IPv4 转向 IPv6, 有充足的地址支持计算机、手机、信息家电等的接入; (4)TCP/IP 多媒体协议; (5)高性能路由器及算法; (6)安全问题, 防火墙及网上信息的加密存储和传输; (7)网络管理, 管理网络的硬件, 实现网络的智能化管理; (8)QoS, 信息交流是基于 IP 的, 因此, 服务质量尤为重要。

在国内, 在国家基金委牵头领导下已经建成了 Internet 2 试验网。此外, 还开展了全光网络和 IP 网一体化的集成研究。研制光器件实现分组交换, 解决电器件分组交换与 DWDM 的高速不匹配问题, 实现由传统的光—电—光模式到光—光—光模式的转变。只有实现光器件分组交换, 网络才能真正走向高速。

3. Internet 软环境

Internet 软环境是基于分布式处理的软件, 处于操作系统之上。Internet 软环境屏蔽了操作系统和网络协议的差异, 支持分布计算, 不仅实现了应用间的互连, 而且实现了应用间的互操作和资源共享。Internet 软环境处于应用软件之下, 可以为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。

目前, Internet 软环境技术主要包括: OMG 组织推出的 CORBA 体系、Sun 公司推出的 J2EE(Java 2 Enterprise Edition)体系和微软公司推出的 Microsoft .NET 体系。它们的成分大致分为5类: 对象环境、数据存取环境、消息环境、交易环境和应用服务器等。它们均可以很好地实现 Internet 之上的程序设计语言、数据库、浏览器、分布式面向对象技术等的结合, 进而为 Internet 上的分布式应用开发提供强有力的支持。这样, 就可以提高应用软件系统的开发效率, 增强系统稳定

杨 森 博士生, 研究方向为分布式面向对象技术、虚拟现实等; 徐泽同 多种 Internet 应用。

研究员, 研究方向为程序设计语言、分布式数据库、Internet 软环境及

性,便于系统维护管理,具有良好的伸缩性和可扩展性,降低系统的投资风险。

Internet 软环境已成为分布式应用的关键性软件,广泛适用于金融、电信、交通、军事等领域。Internet 软环境产品不仅仅是一种软件产品,更重要的是一种计算模型和标准。在软环境所提供的网络分布式开发平台之上,能够开发出大量适合于各种应用领域的软件,从而带动软件业的发展,形成规模效应。

3.1 OMG 的 CORBA

CORBA 是 OMG 组织提出的一个基于 OMA (Object Management Architecture) 体系结构的开放的、分布式对象计算结构,提供了异构计算环境互操作的标准,是面向对象分布式系统中的典型代表。CORBA 在分布式环境中为解决异构环境的应用提供了基础平台,实现了应用程序之间的相互通信。在图2中,给出了 CORBA 的核心构件 ORB 的结构。

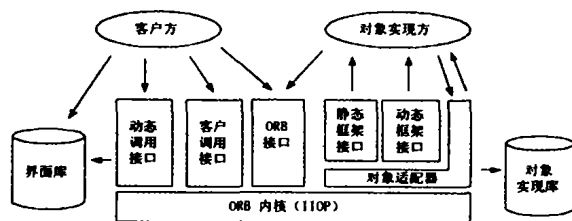


图2 CORBA ORB 的结构

CORBA 规范有以下特点:

- CORBA 提供了软件总线机制,使用接口定义语言 IDL 编译实现与 C++、Java 等语言的映射,便于应用系统的集成。实现了客户方和实现方的完全分离,改变了面向过程调用机制的 C/S 模式。
- CORBA 规范使用代理来完成客户的抽象服务请求的映射;自动发现和搜寻服务器;自动设定路由,完成服务器方任务的实现。
- CORBA 规范是依据 OMA 制订的工业标准,采用逐层设计和实现的方法。

在 CORBA 3.0 版本中,增加了新的组成部分:①推出了组件模型 CCM (CORBA Component Model),便于脚本语言将组件合成一个应用系统,最大限度地推动软件的重用,便于灵活地实现动态配置;②引入了异步消息处理机制,提高了实际应用系统的并行度;③为了增强基于不同 ORB 构建的分布式系统之间的互操作性,CORBA 引入了可移植对象适配器 POA (Portable Object Adapter),POA 还具有线程管理、对象生命周期策略等策略管理功能;④针对嵌入式系统,CORBA 在新规范当中加入了 Minimum CORBA,此外还引入了实时 CORBA 和容错 CORBA 解决实时性、可靠性、高可用性的要求;⑤可互操作命名服务 INS (Interoperable Naming Service) 实现了远程调用的对象引用,EJB 与 CORBA 的互连接就需要借助于 INS;⑥为了提高分布式系统的安全性,加入了防火墙规范,提出了传输层、应用层上的 TCP Firewall、SOCKS Proxy、GIOP Proxy 三种安全解决方案。

20 世纪 90 年后, CORBA 就已成功地应用到电信、金融、CIMS 等应用领域。在国内,电信、金融等领域也已经开始引入 CORBA 技术。CORBA 的实现产品有: IONA 的 Orbix、Borland 的 VisiBroker、BEA 的 ObjectBroker、IBM 的 ComponentBroker,而目前 Orbix 市场占有率较高。

• 96 •

3.2 Oracle 的 OAS

近年来, Oracle 实现了 OAS (Oracle Application Server),是数据库技术、Web 技术、CORBA 体系的很好结合。OAS 提供了 PL/SQL、JWeb、LiveHTML、JCORBA、C、Perl、ODBC、EJB 等多种插件,实现了软总线、软插件、软件重用的软件工程体系,先进而实用。

3.3 Sun 的 J2EE

Sun 公司在 1995 年推出的 Java 是一种跨机器平台的程序设计语言,即所谓“一次编写,到处运行”。1999 年 Sun 发布 J2EE,它适合于书写门户和电子商务网站,使 Java 从一个编程语言发展成为全功能的服务器端应用开发环境。

在图3中,给出了 J2EE 的体系结构图,包括:

- 运行单元 J2EE 提供了 Applet 容器、Application Client 容器、Web 容器、EJB 容器。还提供了 Applet、Application Client 等程序的运行环境;
- 应用接口 J2EE 提供了 JMS、JAAS、JTA、JaveMail/JAF、JAXP、JDBC、Connector 等,实现与其它应用系统的集成;
- 基础平台 有 J2SE (Java 2 Standard Edition) 和 J2ME (Java 2 Micro Edition), J2SE 用于编写 Application、Applet, J2ME 用来开发各种嵌入应用, J2EE 是构建在 J2SE 之上的。

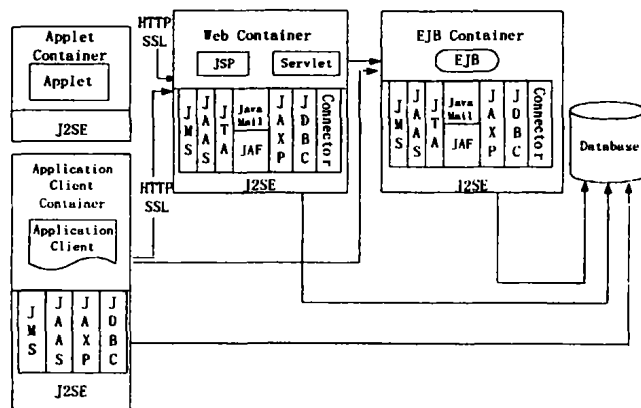


图3 J2EE 体系结构

为了提供一个跨平台的软件环境,J2EE 提供了下列主要组成部分:(1)Java Naming and Directory Interface 为 J2EE 提供了命名和目录服务技术,为 CORBA、LDAP、DNS、NIS、RMI 等多种命名和目录服务提供了统一的访问方式;(2)Java Message Service 是 J2EE 的工作流技术,用于消息服务;(3)Java Authentication and Authorization Service 用于验证用户和对访问进行控制;(4)Java Transaction API 是 J2EE 的事务管理技术,用于确保与事务应用程序、资源管理程序、事务处理监视等复杂事务资源的互操作性;(5)Java Mail 是 J2EE 的邮件技术;(6)Java API for XML Parsing 用于阅读、操作和产生 XML 文档;(7)Java Data Base Connectivity 为 J2EE 提供了访问关系数据库的接口;(8)RMI-IIOP 是在 OMG 的 IIOP 之上提供了 Java RMI API 的实现,支持开发客户和服务器的远程接口;(9)Connector 是 J2EE 与异构企业信息系统互连接的体系结构标准;(10)Java Servlet 是 Java 的服务器端的编程技术,在 J2EE 中已变为底层框架支持;(11)Java Server Pagers 是 J2EE 的页面编程技术;(12)Enterprise Java Beans 是 J2EE 的组件技术。

Java 既是 Internet 语言,又是 Internet 软环境。BEA 的

WebLogic、IBM 的 WebSphere、HP 2001年推出的 IOE(Internet Operation Environment),都是基于 J2EE 标准的 Internet 软环境。

3.4 微软的 Microsoft.NET

分布式计算已成为在分布计算资源中保持互操作性的主要手段,但在多层次、基于组件的框架上实现互操作应用仍然受到限制。

微软公司在2000年6月发布了它的 Microsoft.NET 计划。以 SOAP 为基础的 Microsoft.NET 就是试图要减少局限性,快速地实现互操作。

Microsoft.NET 的基本理念是:不再关注单个的网站与 Internet 连接的单个设备,而是让计算机群、相关设备和服务协调工作,提供更加广泛和丰富的解决方案以及服务,而不再是一座座的孤岛。

SOAP(Simple Object Access Protocol)是一种使用 XML 作为编码格式的协议,XML 是 Internet 信息交换标准,是 Microsoft.NET 的核心技术。XML 类似于 CORBA 的 IIOP,DCOM 的 ORPC 或 Java 远程方法调用的 Java 远程方法协议(JRMP)。

Microsoft.NET 核心技术内容就是基于 XML 标准的符号信息框架、自然用户界面、整体数字媒体支持、支持个人信息管理与控制的隐私权技术,以及为了安全和无缝安装、更新、漫游和离线操作而开发的新型动态传递系统。在图4中,给出了 Microsoft.NET 的开发框架。

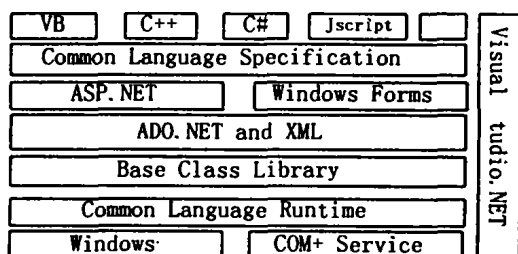


图4 Microsoft.NET 的开发框架

在 Microsoft.NET 中访问数据库的技术是 ADO.NET,它提供了一组用来连接数据库、运行命令、返回记录集的类库。ASP.NET 是 Microsoft.NET 中的网络编程结构,用来建造、运行和发布网络应用。

Microsoft.NET 技术体系不仅仅支持网络应用的开发、部署和运行,而且计划把过去个人计算软件加以“网络化”,支持个人用户通过网上预定,获得诸如我们熟悉的 Office 等个人计算软件的使用许可。

Microsoft.NET 平台是专门为所有网站服务和应用的快速开发、集成而设计的综合解决方案。

虽然 OMG 的 CORBA、ORACLE 的 OAS、SUN 的 J2EE、微软的 Microsoft.NET 等解决方案是以各自不同的概念、模型等形式出现的,但其中有相当部分是借鉴、吸收其它分布式计算结构的设计思想而发展起来的。随着这些技术的不断发展和完善,它们势必会逐步融合。

4. Internet 应用和网格研究

从技术上看,Internet 应用可划为三代:70~80年代主要是 E-mail 应用;90年代主要是 Web 信息浏览和电子商务应用;第三代将会是信息服务网格(Information Service Grid)。从行业角度而言,Internet 早期的应用主要限于科学研究和

教育领域,进而会与经济结合,应用领域包括金融、电信、CIMS、交通等领域。随着经济活动全球化、信息化的发展,电子商务会成了研究和应用的热点。网络和经济相结合是网络发展的必由之路。

在1992年,美国就着手元计算到网格计算的研究。网格计算(grid computing),就是通过网络连接地域上分布的各类计算机(包括机群)、数据库、各类设备和存储设备,形成对用户相对透明的虚拟高性能计算环境,其应用包括分布式计算、高吞吐量计算、协同工程和数据查询等诸多功能。

Internet 的作用是将各种计算机连接在一起,网格将各种信息资源连接在一起。网格计算模式已经发展成为连接和统一各类远程资源的一种基础结构。网格计算是一个广域范围的“无缝的集成和协同计算环境”。

在 Internet/Web 上,数据和信息资源零散地分布在各个网络站点上。而在信息网格中,资源被统一管理,用户可以通过网格门户(portal)透明地使用整个网络资源。用户看到的是一个逻辑门户上若干与自己相关的频道,而不必在成千上万的网站中搜索自己想要的信息。

在国外,最著名的网格计算研究是美国的 Globus 项目。该项目的主要研究目标有两个:其一是网格技术的研究;其二是相应软件的开发和标准的制定。此外,Globus 项目还涉及到网格应用的开发及试验床的建立。最近,Globus 项目提出了网格的体系结构模型。在图5中,给出了 Globus 的体系结构。

Globus 结构的底层是元计算实验床和实验系统,它们是网络连接的网格计算节点。Globus 工具包构件是 Globus 的核心服务。

Globus 工具包构件是 Globus 的核心服务,主要有:资源管理 GRAM,提供资源分配和进程管理;通信管理 Nesus,提供单点和多点通信;安全管理 GSI,提供认证和相关的安全服务;信息管理 MDS,提供分布式访问结构和状态信息;健康和状态管理 NGM,监控系统构件的健壮情况和状态;远程数据访问 GASS,通过串行和并行接口来远程访问数据;执行管理 GEM,构建、缓存和定位各执行。

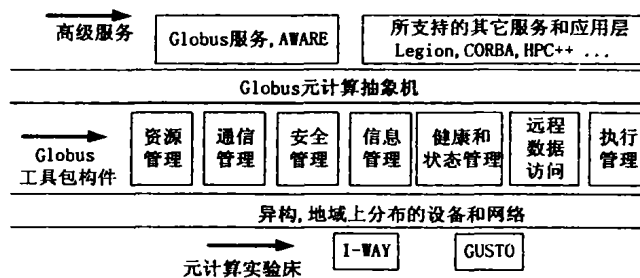


图5 Globus 的体系结构

Globus 提供一些高级服务,可实现对下一层服务的调用,它的最终目标是形成“适应性广域范围资源环境”(AWARE)。同时也支持基于消息传递的 MPI、C++、远程文件访问系统 Remote I/O 等。Globus 工具包支持这些高级应用来调用 Globus 的核心服务。

网格计算可以合理而有效地将远程资源高效地组织在一起,形成“网络虚拟计算机”,以获得超强的计算能力;由于网络环境的异构、可扩展和动态自适应等特性,这类系统在开发软件、并行工具、应用程序等方面都存在较大挑战。Globus 的目标是 AWARE,虽然取得了一些研究成果,但目前仍有大量问题需要进一步的研究。

(下转第100页)

当 k 不能整除 $n-2$, 且 $n \geq k+2$ 时, 令 $\alpha = (n-2) \bmod k$, 即 $n-2$ 除以 k 的余数。由于 $n \geq k+2$, 所以 $T'_k(n-2)$ 至少有 $k+2$ 个结点, 也就是说至少有两个叶子结点, 我们分别称为第一叶子结点和第二叶子结点, 并且定义第一个叶子是以自上而下同层自左而右方式排列的第一个结点, 第二个叶子是以同样方式排列中的第二个结点。首先构造 k 棵树, 每棵树都是 $T'_k(n-\alpha)$, 把剩下的 α 个处理器分别用 $P_0, P_1, P_2, \dots, P_{\alpha-1}$ 编号, 然后再根据一定的算法把余下的 α 个结点加到每棵树上, 确保在所有的 k 棵树中, 每个处理器最多有 k 个孩子结点。那么, 当 k 不能整除 $n-2$, 且 $n \geq k+2$ 时, k 棵树的最大高度是 $h_2 = 1 + f(n-1-\alpha+2k, k)$ 。

对于 $n < k+2$ 的情况 (由于 $n=1, n=2$ 的情况非常简单, 我们这里讨论 $2 < n < k+2$ 的情况), 以上述方式定义的第二叶子结点不一定存在, 因此重新定义第二个叶子结点为第一个叶子结点的第一个孩子, 仍然用上述增加 α 个结点的算法, 很明显在这种算法中, 树的高度最多为 3。

根据 $f(n, k)$ 的表达式可以看出, $f(n, k)$ 是 n 的单调递增函数, 而 $(n-1-\alpha+2k) > (n-1)$, 所以只要得到 k 不能整除 $n-2$, 且 $n \geq k+2$ 情况下高度的上界, 也就是 k 可以整除 $n-2$, 并且 $n \geq k+2$ 情况的界值。

$$\therefore f(n, k) = \lceil \log_k(nk - n + 1) \rceil - 1$$

$$\therefore h = 1 + f(n-1-\alpha+2k, k)$$

$$= \lceil \log_k(n-1-\alpha+2k)(k-1) \rceil + 1 \leq \lceil \log_k(n+2k)k \rceil \leq \lceil \log_k(n+2k) \rceil + 1$$

因此 k 不能整除 $n-2$, 且 $n \geq k+2$ 时, 算法的时间复杂度的最大值是 $\lceil \log_k(n+2k) \rceil + \lceil m/k \rceil$; 而 k 可以整除 $n-2$, 且 $n \geq k+2$ 时, 由于 $h_1 < h_2$, 因此算法的时间复杂度最大值也是 $\lceil \log_k(n+2k) \rceil + \lceil m/k \rceil$; 在 $n < k+2$ 的时候, 有 $h \leq 3$, 所以时间复杂度的最大值是 $\lceil m/k \rceil + 3 - 1 = \lceil m/k \rceil + 2$ 。

综上所述, 对 $k \geq 2$ 和任何值 n , 其时间复杂度都可以表示

为 $\lceil m/k \rceil + \max(\lceil \log_k(n+2k) \rceil, 2)$, 而最小值是 $\lceil \log_{k+1} n \rceil + \lceil m/k \rceil - 1$, 因此该算法不是最优的。

结束语 随着广播技术的发展, 广播算法越来越成熟, 广播理论的总结完善也变得不可忽视。在这篇文章里, 我们对单消息广播和多端口模式下多消息广播进行了比较, 针对文[1]中算法理论上的不足, 在满 k 叉树、完全 k 叉树的基础上, 提出了 k 扫描的数据模型, 并且根据算法的介绍和性能的分析, 归纳得到两条广播定理, 而且这两条定理普遍适用于多端口模式下的广播算法。

事实上, 要使广播性能更加理想, 单从算法入手是比较困难的。我们可以通过对网络的拓扑结构、网络的互联协议以及网络设备等基本入手, 从根本上提高网络的性能, 由此可以达到更加理想的广播效果。

参考文献

- 1 Amotz Bar-Noy, Ching-Tien Ho. Broadcasting Multiple Messages in the Multiport Model. IEEE Transactions On Parallel And Distributed Systems, 1999, 10(5): 500~507
- 2 Ko H, Latifi S, Srimani P K. Near-Optimal Broadcast in All-Port Wormhole-Routed Hypercubes Using Error-Correcting Codes. IEEE Transactions On Parallel And Distributed Systems, 2000, 11(3): 247~259
- 3 李宏, 马建峰, 李学干. 网络系统的广播技术和广播算法. 航空计算技术, 1999, 29(2): 35~39
- 4 严蔚敏, 吴伟民编著. 《数据结构》第二版, 清华大学出版社, pp122~123
- 5 Akl S G. The Design and Analysis of Parallel Algorithm. Prentice Hall, Eaglewood Cliffs, NJ, 1989
- 6 Dimakopoulos V V. On Single-Port Multinode Broadcasting. IEEE Transactions On Parallel And Distributed Systems, 2001, 12(4): 75~78

(上接第97页)

从图1中可以看出: 操作系统+TCP/IP 形成了网络操作系统, 再加上网格功能, 来进一步管理资源, 就形成了网格操作系统。从图1中也可看出: 从 Internet 走向网格操作系统要比从 Internet、软环境走向网格操作系统难得多, 这也可以从 J2EE 支持 Javalin 这一网格研究得到证实。因此, 我们不仅要重视网格的研究, 更应当重视 Internet 软环境的研究。

结束语 1989年, NCFC 的建设奠定了我国 Internet 1 的基础。Internet 2 试验网的建设, 更应当促进我国 Internet 2 应用的发展。全光网络的研究会为我国实现信息化创建更为先进、高速、安全的信息基础设施, 为实现计算机、电信、电视三网合一奠定基础。

网络与经济相结合是网络发展的必由之路。Internet 不仅要能支持科学和教育等应用领域, 更应当与经济相结合, 成为经济全球化、信息化发展的基础。为此, 我们必须高度重视 Internet 软环境的研究和应用的发展, 促进我国的软件研究的产业化、规模化。网格应用将会充分利用网、机资源, 极大地提高应用水平和效率。新技术的涌现为我们带来了机遇, 我们更应当在跟踪的基础上实现创新。

加强 Internet、Internet 软环境和 Internet 上的各种应用的研究, 会极大地促进我国的 IT 产业的迅猛发展, 加速实现信息化带动工业化的宏伟目标。

参考文献

- 1 Rutkowski A M. Internet, Critical Infrastructure and Government: How dose it fit together; can it fit together. <http://www.chaos.com/rutkowski.html>
- 2 Rutkowski A M. A Strategic Perspective: Next Generation Internet <http://www.ngi.org>
- 3 Over view of next Generation Internet Developments, International ISP Forum Workshop, Cannes, 26 Oct 1998. <http://www.ngi.org/public-presentations.html>
- 4 Technology Central: CORBA 3 Release Information. <http://www.omg.org>
- 5 Technology Central: CORBA, IDL, and IIOP. <http://www.omg.org>
- 6 Distributed Multitiered Applications. The J2EE Tutorial. <http://java.sun.com/j2ee/tutorial>
- 7 J2EE connector architecture. <http://java.sun.com/j2ee/white>
- 8 Architecting and designing J2EE[tm] Application. <http://suned.sun.com/us/catalog>
- 9 Introducing microsoft.net. <http://www.microsoft.com/net/>
- 10 White papers. <http://www.microsoft.com/net/whitepapers.asp>
- 11 Computational Grid. Grid research. <http://www.globus.org/research/papers>
- 12 张怡, 胡建平. 网格计算. 北京: 计算机世界报, 2000, 12, 4
- 13 李真文, 杨学良. J2EE 在网站建设中的应用. 计算机科学, 2001, 28(5)
- 14 徐泽同. Internet 软平台. 计算机系统应用, 2000(9)
- 15 Zetong Xu, et al. Technology of Shared LANs & Switched LANs in internetworking. In: Proc. Intl. Conf. on CADEM'95. 2-16 Oct. 1995