

分布式计算技术概述

葛 澎

(海军装备部驻西安军事代表局, 陕西 西安 710054)

摘 要: 分布式计算已经广泛而又深入的参透到各行各业的科技应用中. 本文对几种主要的分布式计算技术, 如网格、云计算、普适计算等, 做了分析和比较, 同时对分布式计算目前所面临的问题做了总结. 最后对分布式计算技术未来的发展方向做了展望与预测.

关键词: 分布式计算技术; 网格; 云计算; 普适计算

中图分类号: TP301

文献标识码: A

文章编号: 1000-7180(2012)05-0201-04

A Brief Overview on Distributed Computing Technology

GE Peng

(Equipment Department of the Navy in Xi'an Military Representative Bureau, Xi'an 710054, China)

Abstract: Distributed computing technology has now been widely applied in the many fields of the scientific research. This paper made a complete analysis and a comprehensive comparison to the key technologies in the area of distributed computing, such as grid computing, cloud computing, pervasive computing, and also made a summary of the current problems facing. Finally made a prediction to the future development of distributed computing technology.

Key words: distributed computing; grid; cloud computing; pervasive computing

1 引言

分布式计算是一种新的计算方式, 它研究如何把一个需要非常巨大的计算能力才能解决的问题分成许多小的部分, 再将这部分分配给多个计算机进行处理, 最后把这些计算结果综合起来得到最终的结果^[1-2]. 计算用服务端负责将计算问题分成许多小的计算部分, 然后将这些部分分配给许多联网参与计算的计算机(客户端)进行并行处理, 最后经综合评定得到最终结果. 目前一些较大的分布式计算项目的处理能力已经达到甚至超过目前世界上速度最快的巨型计算机.

2 分布式计算的发展格局

分布式计算技术起源于美国的 Intel 公司^[3], 目前几乎所有的分布式计算项目都是由欧美国家发起

的. 在分布式计算方面, 欧美国家处于领先地位, 而我国在此方面的研究和实施都与欧美以及其它发达国家存在一定差距. 以世界上最大的分布式计算项目 SETI@home^[4]为例, 它通过使用联网的计算机下载程序分析射电望远镜所收到的讯号, 来搜索地球外的生命迹象. 根据美国加州大学伯克利分校的开放式网络计算平台(Berkeley Open Infrastructure for Network Computing, BOINC)的统计数据来看, 中国参与人数的国际排名目前为 23, 是日本参与人数的九分之一, 只完成了日本用户完成的工作数量的六分之一.

3 典型分布式计算技术

3.1 中间件技术

中间件(Middleware)属于可复用软件的范畴^[5], 处于操作系统软件与用户应用软件中间. 中间

件在操作系统、网络和数据库之上、应用软件之下,其作用是为处于上层的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。

3.2 移动 Agent 技术

移动 Agent 是一个能在异构网络中自主地从一台主机迁移到另一台主机,并可与其它 Agent 或资源交互的程序。移动 Agent 具有自治性、移动性、智能性。它首先是一个软件 Agent,满足 Agent 的目标驱动特征,具有智能性和自治性。其次移动 Agent 可以在网络节点中移动,可以自主地转移到不同的地址空间中执行,在转移过程中它的执行状态被保持,转移到目标后的执行是持续的。从实现技术的角度看,移动 Agent 系统主要由移动 Agent 平台和移动 Agent 组成。

移动 Agent 技术是分布式技术 Agent 技术相结合的产物,它除了具有智能 Agent 的最基本特性如自主能力、社交能力、适应能力和一致主动性外,还具有移动性、可靠性和安全性等特性。移动 Agent 技术在实际中得到了广泛的应用,主要应用于电子商务、分布式信息检索、无线通信服务、入侵检测和网络管理等方面。

3.3 P2P 技术

P2P(peer-to-peer)是指由硬件形成网络连接后的信息控制技术^[2],是一种强调节点之间逻辑对等的新型计算模式。P2P 是由若干互连协作的计算机构成的系统,该系统依存于边缘化(非中央式服务器)设备的主动协作,每个成员直接从其他成员而不是从服务器的参与中受益;系统中成员同时扮演服务器与客户端的角色;系统应用的用户能够意识到彼此的存在,构成一个虚拟或实际的群体。

P2P 的主要特征如下:

- (1) 去中心化,取消或者弱化了集中控制概念;
- (2) 对等性,逻辑上各节点功能上对等,即任意两台 PC 互为服务器,又是客户机;
- (3) 自组织性,各节点以自组织的方式互连成一个拓扑网络,能适应节点的动态变化;
- (4) 资源共享,相互连接的各节点以资源共享为目的。

3.4 网格

网格(Grid)是近年来国际上兴起的一种分布式计算技术^[6]。目前,网格技术已作为一个重要的研究领域,它区别于传统的集中式大规模资源共享、分布式计算以及高性能计算等技术。

所谓的“网格”,即在动态的一组个体、机构和资源的虚拟组织中实行灵活、可靠、可调整的资源共享环境。在此环境中,网格所需解决的问题包括:唯一性认证、资源访问、资源发现的方式等。网格供应商向用户提供高性能计算环境,而信息系统无需购买昂贵的计算设备,而只需从网格中获取所需的计算能力。

网格计算的实现方式是利用互联网或专用网络把地理上广泛分布的各种计算资源互连在一起。这些计算资源包括超级计算机和计算机集群、大型资料库、贵重仪器、传感设备、宽带骨干及网格软件等部分。

3.5 Web Service 技术

Web Service 是一种新的 web 应用程序分支,可以执行从简单的请求到复杂商务处理的任何功能。Web Service 使用标准互联网协议,像超文本传输协议(HTTP)和 XML,将功能纲领性地体现在互联网和企业内部网上。

Web Service 的基本架构由三个参与者和三个基本操作构成。三个参与者分别是:服务提供者、服务请求者和服务代理,而三个基本操作分别为:发布、查找和绑定。Web Service 基本架构如图 1 所示。服务提供者将其服务发布到服务代理的一个目录上;当服务请求者需要调用该服务时,首先利用服务代理提供的目录去搜索该服务,得到如何调用该服务的信息;然后根据这些信息去调用服务提供者发布的服务。当服务请求者从服务代理得到调用所需服务的信息之后,通信是在服务请求者和提供者之间直接进行,无须经过服务代理。Web Service 体系使用一系列标准和协议实现相关的功能,如:使用 WSDL(web service description language)来描述服务,使用 UDDI(universal description, discovery, integration)来发布、查找服务,而 SOAP(simple access protocol)被用来执行服务调用。

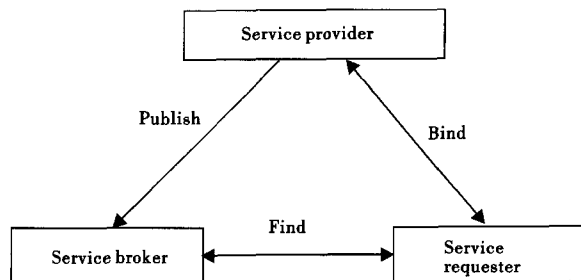


图 1 Web Service 的基本架构

在 Web Service 架构的备模块间以及模块内部,消息以 XML 格式传递;实现 Web Service 技术的流行方法是使用 .NET 和 Java 两种技术。

3.6 普适计算

普适计算(Pervasive computing 或 Ubiquitous computing)是指在普适环境下使人们能够使用任意设备、通过任意网络、在任意时间都可以获得一定质量的网络服务的分布式计算技术^[6]。它是在网络技术和移动计算的基础上发展起来的,其重点在于提供面向客户的、统一的、自适应的网络服务。

普适环境主要包括网络、设备和服务。网络环境包括:Internet、移动网络、电话网、电视网和各种无线网络等;设备包括:计算机、手机、汽车、家电等能够通过任意网络上网的设备;服务内容包括:计算、管理、控制、资源浏览等。

普适计算具有以下环境特点:在任何时间、任何地点、任何方式的方便服务,不同的网络(不同协议、不同带宽)、不同的设备(屏幕、平台、资源)、不同的个人偏好等。实现普适计算的目标需要以下关键技术:场景识别、资源组织、人机接口、设备无关性技术、设备自适应技术等。

3.7 云计算

云计算(Cloud Computing)是网格计算、虚拟化等核心技术在网络时代的发展和商业实现^[7],它是一种动态的、易扩展的、且通常通过互联网提供虚拟化的资源计算方式。云计算包括:基础设施及服务(IaaS)、平台及服务(PaaS)和软件及服务(SaaS)以及其他依赖于互联网满足客户计算需求的技术趋势。云计算主要提供通用的通过浏览器访问的在线商业应用、软件和数据存储等服务。人们可以通过传统的计算机、移动设备、甚至是在网络中的嵌入式芯片来获取这些服务。

云计算的基本原理是:通过使计算分布在大量的分布式计算机上,而非本地计算机或远程服务器中,企业数据中心的运行将更与互联网相似。这使得企业能够将资源切换到需要的应用上,根据需求访问计算机和存储系统。

3.8 各种技术的比较

Web Service 技术的体系结构与基于中间件分布式系统的体系结构相比非常相似,可将体系结构中的 Web 程序看作中间件。从结构上来看,Web 服务只是从侧面对中间件平台技术进行革新,虽然所有服务之间的通信都以 XML 格式的消息为基础。但调用服务的基本途径主要还是 RPC,而且具体实

现并没有提供一种全新的编程模式。

网格计算与基于中间件的分布式计算技术相比,它依然以“中间件”为技术核心,在实现形式上并没有太大的改变。然而经过一系列的技术革新,网格系统中的技术内涵已经发生了深刻的变化。(1)基于中间件的分布式计算技术的资源主要是数据和软件,而网格计算的资源已经延伸到所有用于共享的实体,包括硬件、软件,甚至分布式文件系统、缓冲池等;(2)在 Internet 上,网格中间件层提供了与 web 服务一样优秀的扩展功能,打破了 P2P 技术的部分局限。

网格计算、Web Service 等技术在异构平台上构筑了一层通用的、与平台无关的信息和服务交换设施,从而屏蔽了 Internet 中千差万别的差异,使信息和服务畅通无阻地在计算机之间流动。网格计算与 Web Service 技术的共同载体是 Internet。但两者的不同之处在于,网格系统连接物理上分散的硬件资源,形成虚拟计算组织,从而使计算资源得到充分共享;而 Web 服务则是以商务应用为背景,是基于网格系统之上的。网格系统为 Web 服务提供一个与硬件无关的虚拟计算机;而 Web 服务是架构在虚拟计算机平台上,与环境、语言无关的应用集成平台。

云计算与网格技术是各有所长,而不是相互取代关系。网格技术主要解决分布在不同机构的各种信息资源的共享问题,而云计算主要解决计算能力和存储空间的集中共享使用问题。云计算正是普适计算的开始,它可使网络从一个文字的、社区性的一个平台变为一个可以在任意时间、任意地点提供无限应用的、汇集众多服务的互联网时代。

总体来说,尽管各种分布式计算技术在理念、规范和实现等方面都有较大的差异,但同时也相互交融。因此,各种分布式计算技术可以共同存在,它们的相互结合也是非常有意义和现实的。

4 分布式计算的优势和面临的问题

4.1 分布式计算的优势

分布式计算比起传统计算模式具有以下几方面的优势:

- (1) 稀有资源可以共享;
- (2) 通过分布式计算可以在多台计算机上平衡计算负载;
- (3) 可以把程序放在最适合运行它的计算机上。

4.2 分布式计算的面临的问题

分布式计算技术是计算机网络的产物,也是计算机网络应用未来的发展方向,虽然已经产生了大量的技术产品,但同时出现了一些问题:

(1) 标准问题

目前,几乎所有的分布式计算技术都没有完整的统一的标准,标准的缺乏使得资源的共享和社会化程度低下. 分布式计算之间缺乏有效的交互、协作与协同,从而在很大程度上制约了其发展.

(2) 技术问题

分布式计算技术是分布式计算能够转化为工业化生产力的前提,也是分布式计算的质量、功能的保证. 可是,到目前为止,所有的分布式计算基数都或多或少存在没有解决的问题,没有哪一种技术能实现完全意义上的分布式计算,满足所有分布式计算的需求.

(3) 异构问题

现在的网络是一个异构的环境,分布式计算技术首先需要解决异构环境下的互操作问题. 而要解决异构环境下的互操作问题,首要任务是如何解决互相识别. 目前,不可能实现对所有的资源采用同一种描述方式,也无法智能地识别这些资源,这就导致分布式计算技术只能在一定的范围内使用.

(4) 安全性问题

分布式计算系统存在如下的安全问题^[8]: ①身份认证; ②数据传输安全; ③主机资源保护问题; ④恶意主机问题. 为了保证分布式计算的安全,传统的安全机制不是特别有效,表现在分布式计算中事务交互的任一方都无法判断对方是否真正的可信,如: 请求是否是病毒和木马所发起; 远程计算机的环境是否存在信息泄漏或者恶意欺骗等.

5 分布式计算技术的发展展望

5.1 分布式计算技术的有效组合

要创建可信的、松散的、健壮分布式系统,使分布式计算得到真实的应用和普及,必须加快标准的研制,从底层信号的传输到复杂业务的流程等各种不同的层次都要形成统一的标准. 此外,应该尽可能对分布式计算技术进行有效的组合,集中各种技术的优势,满足分布式计算的需求,这应该是分布式计算技术研究的方向. 可将云计算与网格技术结合,这是因为: 要使用云计算,就必须要将各种数据、系统、应用集中到云计算数据中心; 而很多现有信息系

统要改变运行模式,迁移到云计算平台上的难度和成本是较高的. 如果进一步要求将这些数据随时地传送到云中心,则对网络宽带的消耗也是巨大的和不经济的. 因此,还会有大量的应用系统处于分散运转状态,而不会集成到云计算平台上. 而网格技术可以实现在现有资源上集成,达到“物理分散、逻辑集中”的效果.

5.2 分布式计算技术与其它高新技术的结合

据国际 IDC 统计,2010 年最热门的四大技术分别是: 普适计算、云计算、3G、物联网 (Internet of Things)^[9]. 任何一门技术的发展都不是孤立的,所以我们不必局限于对分布式计算的各种技术之间进行组合,还可以将其同其他技术相结合,达到相互促进,共同为人们提供更舒适的生活服务. 如将云计算和物联网结合,一方面,物联网将使用数量惊人的传感器,采集到更加惊人的数据量,这些海量数据需要通过无线传感器、宽带互联网向某些存储和处理设施汇聚,而采用云计算承载这些任务则具有非常显著的性价比优势; 另一方面,物联网将成为云计算最大的用户,为云计算的成功奠定基石.

参考文献:

- [1] 周晓峰, 王志坚. 移动 Agent 计算研究[J]. 计算机时代, 2006, 25(7): 156-158.
- [2] 金培弘. 分布式系统概念与设计[M]. 北京: 机械工业出版社, 2004.
- [3] 中国分布式计算总站[EB/OL]. [2011-09-25]. <http://www.equn.com/>
- [4] Distributed Computing Stats from Free-DC [EB/OL]. [2011-09-25]. <http://stats.free-dc.org/stats.php?page=userbycpid&cpid=bd73e92b61829c17efec4d786894b51e>
- [5] 蒋雄伟, 马范援. 中间件与分布式计算[J]. 计算机应用, 2004, 22(4): 7-8.
- [6] 王伟. 面向未来的计算——普适计算[J]. 计算机学报, 2009, 15(2): 13-16.
- [7] 吴吉义, 平玲嫡. 云计算从概率到平台[J]. 计算机研究和发展, 2009, 25(12): 183-191.
- [9] 李晓用. 可信分布式计算环境的关键技术研究[D]. 北京: 北京交通大学, 2008.
- [9] 王宝云. 物联网技术研究综述[J]. 电子测量与仪器学报, 2009, 23(12): 1-7.

作者简介:

葛 澎 男, (1968—), 大校, 高级工程师. 研究方向为质量控制、并行处理.