



云背景下

涂云杰 著

数据库安全性

与数据库完整性研究



中国水利水电出版社
www.waterpub.com.cn

云背景下数据库安全性与 数据库完整性研究

涂云杰 著



中国水利水电出版社
www.waterpub.com.cn

内容提要

全书共分八章。阐述了云计算、云数据库、数据库安全性以及数据库完整性等，集中阐述了数据库安全性控制的必要性、常用方法、技术及策略以及数据库完整性的定义、检查及违约处理等，并对图书资料智能管理系统进行数据库安全性与完整性的设计。

图书在版编目（C I P）数据

云背景下数据库安全性与数据库完整性研究 / 涂云杰著. — 北京：中国水利水电出版社，2014. 10
ISBN 978-7-5170-2582-5

I. ①云… II. ①涂… III. ①数据库系统—安全技术
IV. ①TP311.13

中国版本图书馆CIP数据核字(2014)第228523号

策划编辑：韩 光 责任编辑：李 炎 封面设计：李 佳

书 名	云背景下数据库安全性与数据库完整性研究
作 者	涂云杰 著
出版发行	中国水利水电出版社 (北京市海淀区玉渊潭南路1号D座100038) 网址: www.waterpub.com.cn E-mail: mchannel@263.net (万水) sales@waterpub.com.cn
经 售	电话: (010) 68367658 (发行部)、82562819 (万水) 北京科水图书销售中心(零售) 电话: (010) 88383994、63202643、68545874 全国各地新华书店和相关出版物销售网点
排 版	北京万水电子信息有限公司
印 刷	三河市天润建兴印务有限公司
规 格	170mm×240mm 16开本 7.5印张 137千字
版 次	2014年11月第1版 2014年11月第1次印刷
定 价	29.00元

凡购买我社图书，如有缺页、倒页、脱页的，本社发行部负责调换
版权所有·侵权必究

前 言

由于数据库系统中存放着大量的数据，且为许多用户所共享，因此安全性问题是必须首先要解决的。云计算的优势是明显的，它的发展也非常迅速，但是仍然有人对云计算的前景担忧，主要出于对安全问题的忧虑。云计算时代已经来临，完全依赖云是危险的，但是认为云存在一定的不安全因素而放弃使用云也是不可取的。我们要正确地认识云，云安全技术推出的时间还比较短，所以很多技术尚需完善，我们可以看到，很多人正在为此而努力。在云技术飞速发展的今天，云安全体系也要相应地发展，使得我们可以更快更好地进入云的世界。云计算在国内的发展还处于初级阶段，缺乏相应的标准和法规，云安全问题需要特别重视，在云计算发展过程中需要逐步予以解决。

云计算在欧美等国得到政府的大力支持和推广，云计算安全和风险问题也得到各国政府的广泛重视。2010年11月，美国政府CIO委员会发布关于政府机构采用云计算的政府文件，阐述了云计算带来的挑战以及针对云计算的安全防护，要求政府及各机构评估云计算相关的安全风险并与自己的安全需求进行比对分析。并指出，由政府授权机构对云计算服务商进行统一的风险评估和授权认定，可加速云计算的评估和采用并降低风险评估的费用。2010年3月，参加欧洲议会讨论的欧洲各国网络法律专家和领导人呼吁制定一个关于数据保护的全球协议，以解决云计算的数据安全弱点。欧洲网络和信息安全局（ENISA）表示，将推动管理部门要求云计算提供商通知客户有关安全攻击状况。日本政府也启动了官民合作项目，组织信息技术企业与有关部门对于云计算的实际应用开展计算安全性测试，以提高日本使用云计算的安全水平，向中小企业普及云计算，并确保企业和个人数据的安全性。

云背景下的数据库安全问题的研究在解决云用户和云服务商之间的纠纷、化解双方的信任危机、防止重大事故发生而造成的数据丢失等方面有重大意义。云计算的安全问题涉及面广、问题复杂，需要更多学者的进一步研究。在云技术飞速发展的今天，云安全体系也要相应地发展，使得我们可以更快更好地进入云的世界，对更好地保护数据、保护机密与隐私起到基础性服务作用。

数据库完整性指的是数据库中数据的正确性和相容性。为维护数据库的完整性，DBMS的完整性控制机制必须具有提供定义完整性约束条件（完整性规则）的机制、提供完整性检查的方法和违约处理机制。

全书共分八章。阐述了云计算、云数据库、数据库安全性以及数据库完整性等，集中阐述了数据库安全性控制的必要性、常用方法、技术及策略以及数据库完整性的定义、检查及违约处理等，并对图书资料智能管理系统完成了数据库安全性与完整性的设计。

本书的第一章概述了云计算的产生背景、云计算的定义、云计算的特点、云计算的形式、云计算的原理和云计算发展现状与趋势。通过对云计算的介绍使大家对云计算有个大概了解。

第二章给出了云数据库的基本知识，包括云数据库的定义、云数据库的特点、云数据库的发展现状与发展趋势、云数据库的重要性和中国云数据库的发展前景。

第三章集中阐述了数据库安全性控制的常用方法和技术，还给出了常用的数据库安全标准。

第四章详细讨论了云背景下数据库安全性研究的必要性、云背景下数据库安全发展现状、云背景下数据库安全新特点、云背景下数据库安全问题和云背景下数据库安全性策略研究。

第五章介绍了国内著名的云计算创新机构。百度已成为当前我国云计算产业创新发展的重要力量，其开放云平台能够帮助中小企业显著降低信息化建设成本和创新创业成本。同时，百度通过向开发者开放核心云能力，推动了移动云计算的发展和成熟。还介绍了华为的云平台、上海市云计算创新基地、国富安电子商务安全认证有限公司、曙光云计算中心、腾讯云平台等。

第六章阐述了图书资料智能管理系统中的数据库安全性的设计与实现。根据本图书资料智能管理系统的设计要求，设计了3类管理员：超级管理员、读者管理员和图书管理员。他们拥有不同的操作权限。另外为了保证数据库的正常使用，避免意外带来的损失，数据库用户需要选择某种备份方式进行数据库的备份，以防患于未然。本章还结合图书资料智能管理系统阐述了备份与还原的灵活使用，完成了图书资料智能管理系统的安全性的设计。

第七章研究数据库完整性，给出了数据库完整性有关的概念，并介绍了触发器的功能与分类等，结合实例阐述了数据库完整性的定义、检查与违约处理机制。

第八章研究了数据库完整性在图书资料智能管理系统中的应用，实现了图书资料智能管理系统中完整性的设计与定义，保证了图书资料智能管理系统中数据的正确性和相容性，避免错误数据的出现。

本书的出版得到内蒙古自治区高等学校科学研究项目“云背景下的数据库安全性研究”(NJZY14308)的支持。在这里感谢内蒙古自治区以及学院领导对科研工作的大力支持，是他们的高瞻远瞩为学院的科研发展奠定了坚实的基础，为把

呼伦贝尔学院建设成为一所国内有影响、区内有地位，具有地方特色、民族特色的本科院校奠定了基础。与此同时还要感谢一直支持我工作的同事、朋友和亲人，他们的无私关爱、默默奉献与真心帮助，给了我工作的动力，使我能将更多的精力投入到工作与科研中来。

涂云杰

2014年7月29日于呼伦贝尔学院

目 录

前言	
第一章 绪论	1
1.1 云计算的产生背景	1
1.2 云计算的定义	2
1.3 云计算的特点	3
1.4 云计算实现形式	5
1.5 云计算的原理	6
1.6 云计算发展现状与趋势	7
1.7 云计算发展面临的主要问题	13
1.8 本书的主要内容	14
第二章 云数据库	16
2.1 云数据库的定义	16
2.2 云数据库的特点	16
2.3 云数据库的发展现状	17
2.4 云数据库的重要性	21
2.5 中国云数据库的前景	22
2.6 结论和讨论	23
第三章 数据库安全性	25
3.1 数据库安全性概述	25
3.2 实现数据库安全性控制的常用方法和技术	25
3.2.1 用户身份认证	25
3.2.2 访问控制	28
3.2.3 视图技术	30
3.2.4 数据加密	32
3.2.5 安全审计	38
3.2.6 数字水印	40
3.2.7 防火墙	41
3.3 安全标准	42
3.3.1 TCSEC/TDI 标准	42
3.3.2 CC 标准	44
3.4 结论和讨论	44

第四章 云背景下数据库安全性	46
4.1 云背景下数据库安全性研究的必要性	46
4.2 云背景下数据库安全发展现状	47
4.3 云背景下数据库安全新特点	48
4.4 云背景下的数据库安全问题分析	50
4.5 云背景下的数据库安全性策略研究	53
4.6 结论和讨论	58
第五章 国内著名的云计算创新机构	60
5.1 百度云服务平台	60
5.2 华为云平台	61
5.3 上海市云计算创新基地	62
5.4 国富安公司	63
5.5 山丽网安	64
5.6 曙光云计算中心	66
5.7 腾讯云平台	67
5.8 盛大云平台	67
5.9 结论和讨论	67
第六章 图书资料智能管理系统中的数据库安全性的研究与实现	69
6.1 图书资料智能管理系统权限的提出	69
6.2 图书资料智能管理系统权限的实现	70
6.2.1 基本操作语句介绍	70
6.2.2 界面方式分配图书管理权限	72
6.2.3 命令方式分配图书管理权限	76
6.3 数据库的备份与还原	77
6.3.1 数据库的备份	77
6.3.2 数据库的恢复	84
6.4 结论和讨论	89
第七章 数据库完整性	90
7.1 有关概念	90
7.2 数据库完整性的分类	91
7.2.1 实体完整性	91
7.2.2 参照完整性	91
7.2.3 用户定义的完整性	93
7.2.4 数据库安全性与完整性的区别与联系	94
7.3 触发器	94

7.4	结论和讨论	96
第八章	数据库完整性在图书资料智能管理系统中的应用	97
8.1	图书表中完整性的实现	97
8.2	读者类别表完整性的实现	98
8.3	读者表完整性的实现	99
8.4	借阅表中完整性的实现	99
8.5	完整性的检查与违约处理	100
8.6	触发器在图书资料智能管理系统中的应用	102
8.6.1	“借阅”表上的 INSERT 触发器	102
8.6.2	“归还”功能实现	102
8.6.3	“图书”表上的 INSERT 触发器	103
8.6.4	“图书”表上的 DELETE 触发器	103
8.6.5	“图书”表上的 UPDATE 触发器	103
8.7	结论和讨论	104
参考文献		105

第一章 绪论

1.1 云计算的产生背景

由于互联网的急速发展，存储、计算机能量消耗，数据急剧增长，成本也随之升高，原始的互联网系统与服务设计已经不能解决上述种种问题，互联网急需新的解决方案。同时，大型企业必须充分研究数据资源，才能支持商业行为，数据的收集与分析必须建立在一种新的平台上。于是，2007年，一个称为“云计算”的概念首次被 Google 提出，这是一个美丽的网络应用模式。随后，这一 IT 技术风暴席卷了整个 IT 界，为全球 IT 界带来了一场全新的变革。

云时代来临了。云计算（Cloud Computing）所倡导的“按需使用、按使用付费”与“像用水和用电一样用 IT”如出一辙。谁有独到的眼光，谁能够预见未来的发展趋势，他就可能引领这个产业继续发展，成为 30 年后的王者^[1,9]。传统模式下，企业建立一套 IT 系统不仅仅需要购买硬件等基础设施，还要买软件的许可证，需要专门的人员维护。当企业的规模扩大时还要继续升级各种软硬件设施以满足需要。对于企业来说，计算机等硬件和软件本身并非他们真正需要的，它们仅仅是完成工作、提供效率的工具而已。对于个人来说，想正常使用电脑需要安装许多软件，而许多软件是收费的，对不经常使用该软件的用户来说购买是非常不划算的。可不可以有这样的服务？能够提供我们需要的所有软件供我们租用。这样我们只需要在用时付少量“租金”即可“租用”到这些软件服务，为我们节省许多购买软硬件的资金。

我们每天都要用电，但并不是每家都自备发电机，而是由电厂集中提供；我们每天都要用自来水，但不是每家都有水井，而是由自来水厂集中提供。这种模式极大地节约了资源，方便了我们的生活。面对计算机给我们带来的困扰，我们可不可以像使用水和电一样使用计算机资源？这些想法最终导致了云计算的产生。

云计算是在分布式计算（Distributed Computing）、网格计算（Grid Computing）、并行计算（Parallel Computing）等发展的基础上提出的一种新型计算模型，是一种新兴的共享基础框架的方法。它以互联网为基础，借助于互联网高速的传输能力将大量的数据和计算资源分布在大量计算机构成的资源池上，将数据的处理过程从个人计算机或服务器移到互联网上的计算机集群中，把存储和计算作为一种

服务提供给云用户。云计算是虚拟化（Virtualization）、效用计算（Utility Computing）、IaaS（基础设施即服务）、PaaS（平台即服务）、SaaS（软件即服务）等概念混合演进并跃升的结果。^[1,2]

云通过网络把多个计算机整合成一个具有强大计算能力的完美系统，并借助 SaaS、PaaS、IaaS 等先进的商业模式把强大的计算能力分布到终端用户手中。云计算的核心理念就是通过不断提高“云”的处理能力，进而减少用户终端的处理负担，最终使用户终端简化成一个单纯的输入输出设备，并能按需享受“云”的强大计算处理能力。

云计算是一种新的计算模式，是分布式处理、并行处理和网格计算、网络存储、虚拟化、负载均衡等传统计算机技术和网络技术发展融合的产物。云计算将计算资源分布在由大量计算机构成的资源池上，而非本地计算机或远程服务器中，用户根据需求通过网络访问计算机和存储系统。在远程的数据中心里，成千上万台计算机和服务器连接成一片计算机云。利用这些“云”，用户通过终端接入数据中心，按自己的需求进行运算。

1.2 云计算的定义

云计算的概念最早由 Google 提出，是一种网络应用模式。狭义的云计算是指 IT 基础设施的交付和使用模式，通过网络以按需和易扩展的方式获得所需的资源；广义的云计算是指服务的交付和使用模式，通过网络以按需和易扩展的方式获得所需的服务，这种服务既可是和软件、互联网相关的，也可是任意其他的服务，具有超大规模、虚拟化、可靠安全等独特功效。目前，云计算没有统一的定义，这也与云计算本身特征很相似。维基百科对云计算的定义是：云计算是一种基于互联网的计算新方式，通过互联网上异构、自治的服务为个人和企业提供按需即取的计算。由于资源是在互联网上，而互联网通常以云状图案来表示，因此以云来类比这种计算服务，同时云也是对底层基础设施的一种抽象描述。云计算的资源是动态扩展且虚拟化的，通过互联网提供，终端用户不需要了解云中基础设施的细节，不必具有专业的云技术知识，也无需直接进行控制，只关注自身真正需要什么样的资源以及如何通过网络来获得相应的服务。

中国云计算网为云计算定义如下：云计算（cloud computing）是一种基于因特网的超级计算模式，在远程的数据中心里，成千上万台电脑和服务器连接成一片电脑云。中国网格计算、云计算专家刘鹏定义如下：“云计算将计算任务分布在大量计算机构成的资源池上，使各种应用系统能够根据需要获取计算力、存储空间和各种软件服务”。^[1,4,5,6,10]

“云”是一些可以自我维护 and 管理的虚拟计算资源，通常为一些大型服务器集群，包括计算服务器、存储服务器、宽带资源等。云计算将所有的计算资源集中起来，并由软件实现自动管理，无需人为参与。这使得应用提供者无需为繁琐的细节而烦恼，能够更加专注于自己的业务，有利于创新和降低成本。

之所以称为“云”，是因为它在某些方面具有现实中云的特征：云一般都较大，云的规模可以动态伸缩，云的边界是模糊的。云在空中飘忽不定，你无法也无需确定它的具体位置，但它确实存在于某处。之所以称为“云”，还因为云计算的鼻祖之一亚马逊公司将曾经大家称作网格计算的东西，取了一个新名称“弹性计算云”（EC2），并取得了商业上的成功^[1]。

有人打了个比方：这就好比是从古老的单台发电机模式转向了电厂集中供电模式。它意味着计算能力也可以作为一种商品进行流通，就像煤气、水电一样，取用方便，费用低廉。最大的不同在于，它是通过互联网进行传输的^[10]。

1.3 云计算的特点

云计算是一种基于互联网、大众参与的计算模式，包含互联网上的应用服务及提供这些服务的软硬件设施，具有虚拟化、动态性、可伸缩的计算资源等特点。谷歌（Google）、亚马逊（Amazon）、IBM 和微软（Microsoft）等公司纷纷推出各自的云计算平台、服务和产品。云计算使真正具有实用意义的海量数据挖掘成为可能。云计算就是把一些应用放到服务器端，这样对普通用户机器的配置要求就不那么高了，但由于应用在服务器端，所以对网络的要求就高了。

1. 灵活性

目前云计算技术主要是虚拟化，就是服务器端的机器用虚拟化技术虚拟出多个客户机可以使用，就想你真的有一台机器一样。搜狗的云输入法就是字库在服务器那里，本机不用安装搜狗输入法但却能在网页中使用。

用户可以在任意时间、任意地点，通过自己的终端（例如：手机，电脑等）连接云计算服务端获取自己所需要的服务。它并不是一个可形状化的实体，也不具有固体的位置，用户甚至不需要了解“云”的具体运行机制。

云计算实际上是服务器虚拟化技术和基础设施即服务（IaaS）两者的结合。其核心是将某一或某几个数据中心的计算资源虚拟化之后，向用户提供以租用计算资源为形式的服务。而这种提供计算资源的服务实际上并不是新技术，而是业务模式上的创新。简单地讲，就是将很多位于同一地点或不同地点的计算机集合起来为企业或个人用户提供各种服务。云计算是一个新兴的 IT 部署和交付模式，用来通过网络实时提供产品、服务和解决方案^[102]。

2. 经济性

云计算采用虚拟化技术, 弱化了对于终端设备的要求, 以较低的成本获得较高的处理能力, 使用起来很方便。用户无需了解云计算的具体运行机制, 只需要一个简单的网络连接设备就可以获得自己需要的服务, 这不仅在很大程度上实现了资源的共享, 而且节省了硬件资金的投入。

3. 共享性

它可以轻松实现不同设备间的数据与应用共享。云存储通过集群应用、网络技术、分布式处理等功能将网络中的设备集中起来协调工作。云将终端与庞大的数据中心相联, 用户可以通过统一的、标准的应用接口登录云, 获取云资源。人们因此得以突破资源的限制, 在近乎无限量的资源上以近乎无限快的速度获得信息, 实现数据与应用共享。用户在极低的成本下可以获得极高的运算处理能力, 从而实现了将计算作为一种公用设施来提供的梦想^[3,10]。

在云计算的大趋势下, 整个 IT 服务业将进行大的调整, 人们依据云实现虚拟化, 让用户可以脱离技术上的复杂性而直接获得应用。

4. 动态可扩展性

云计算可以大规模地扩展其子结点, 甚至可以再促使几千个结点同时处理不同的应用。伴随着用户规模的增长以及应用规模的增加, 它具有动态的伸缩性。

在理想状态下, 它能够支持无限的并发用户, 提供永不枯竭的数据应用资源。

5. 规模大

云计算规模非常大, 例如, Google 云计算拥有上百万台服务器, 同时像 Amazon、IBM、微软、Yahoo 等的“云”也拥有几十万台服务器。对于一般企业而言, 云计算一般拥有成百上千台个服务器。“云”所带来的计算能力对用户是前所未有的^[3,10]。

6. 透明性

云计算提供的是软、硬件的服务。一般情况下, 服务的实现机制对用户来说是透明的。用户无需了解云计算的具体机制, 就可以获得所需要的服务, 具有很强的易用性^[3,10]。

7. 按需服务

按需付费是目前各类云计算服务中不可或缺的一部分。对用户而言, 云计算不但省去了基础设备的购置运维费用, 而且能根据企业成长的需要不断扩展订购的服务, 不断更换更加适合的服务, 提高了资金的利用率^[1,3,4,9,10]。

8. 潜在的安全危险

由于云计算不同于现有的以桌面为核心的数据处理和应用服务都在本地计算

机中完成的使用习惯，而是把这些都转移到“云”中，它将改变我们获取信息、分享内容和相互沟通的方式，于是，随之产生的是客户的重要数据和应用服务在云中的安全问题^[103]。

任何以互联网为基础的应用都存在着一定的危险性，云计算也不例外。正是由于用户能够简单动态地享受云计算提供的各项服务，使得如何在频繁的存储与计算服务中保证云环境数据库的安全问题成为我们面临的主要问题。

云计算服务除了提供计算服务外，还必然提供存储服务。云计算存储服务当前垄断在私人机构手中，而他们仅仅能够提供商业信用。对于政府机构、商业机构，例如银行等，在选择云计算服务时，应保持相当高的警惕，以防止数据泄露给机构和用户带来的巨大损失。身处信息社会，信息无疑是至关重要的。另外，云计算中的数据对于数据所有者以外的其他用户是保密的，但对于提供云计算的商业机构却毫无秘密可言，这就像常人不能监听别人的电话，但是在提供通讯服务的机构内部，他们的相关人员却可以随时监听任何电话。所有这些潜在的危险，是商业机构和政府机构选择云计算服务，特别是国外机构提供的云计算服务时，不得不考虑的一个重要前提。

1.4 云计算实现形式

1. SaaS（软件即服务）^[12]

这种类型的云计算通过浏览器把程序传给成千上万的用户。在用户眼中，这样会省去在服务器和软件授权上的开支。从供应商角度而言，只需要维持一个程序即可，这样能够减少成本。Salesforce 是迄今为止提供这类服务最为著名的公司，SaaS 在人力资源管理程序和 ERP 中比较常用。

2. 效用计算（Utility Computing）

效用计算是一种提供模型的服务，最近在 Amazon、Sun、IBM 和其他提供存储服务和虚拟服务器的公司中获得新生。效用计算是为 IT 行业创造虚拟的数据中心，使其能够把内存、I/O 设备、存储和计算能力集中起来，成为一个虚拟的资源池为整个网络提供服务。

3. 网络服务

同 SaaS 关系密切，网络服务提供商们提供 API，让开发者能够开发更多基于互联网的应用，而不是单机程序。

4. PaaS（平台即服务）

平台即服务（Platform as a Service, PaaS）是一种无需下载或安装，即可通过

因特网发送操作系统和相关服务的模式。由于平台即服务能够将私人电脑中的资源转移至网络云，所以有时也被称为“云件”（cloudware）。平台即服务是软件即服务（Software as a Service）的延伸。软件即服务是将软件部署为托管服务并通过因特网提供给客户^[27]。

平台即服务使得用户能将云基础设施部署与创建至客户端，或者借此获得使用编程语言、程序库与服务。用户不需要管理与控制云基础设施，包含网络、服务器、操作系统或存储，但需要控制上层的应用程序部署与应用代管的环境^[28]。

所谓 PaaS 实际上是指将软件研发的平台作为一种服务，以 SaaS 的模式提交给用户。因此，PaaS 也是 SaaS 模式的一种应用。但是，PaaS 的出现可以加快 SaaS 的发展，尤其是加快 SaaS 应用的开发速度。PaaS 之所以能够推进 SaaS 的发展，主要在于它能够为企业提供定制化研发的中间件平台，同时涵盖数据库和应用服务器等。PaaS 可以提高在 Web 平台上利用的资源数量，用户或者厂商基于 PaaS 平台可以快速开发自己需要的应用和产品。同时，PaaS 平台开发的应用能更好地搭建基于 SOA 架构的企业应用。此外，PaaS 对于 SaaS 运营商来说，可以帮助他们进行产品多元化和产品定制化^[29]。

5. MSP（管理服务提供商）

MSP 是最古老的云计算运用之一。这种应用更多的是面向 IT 行业而不是终端用户，常用于邮件病毒扫描、程序监控等。

6. 商业服务平台^[12]

SaaS 和 MSP 的混合应用，该类云计算为用户和提供商之间的互动提供了一个平台。比如用户个人开支管理系统，能够根据用户的设置来管理其开支并协调其订购的各种服务。

7. 互联网整合^[12]

将互联网上提供类似服务的公司整合起来，以使用户能够更方便地比较和选择自己的服务供应商。

1.5 云计算的原理

云计算的最终目标是将计算、服务和应用作为一种公共设施提供给公众，使人们能够像使用水、电、煤气和电话那样使用计算机资源^[13]。

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。这可是一种革命性的举措。打个比方，

这就好比是从古老的单台发电机模式转向了电厂集中供电模式。它意味着计算能力也可以作为一种商品进行流通,就像煤气、水电一样,取用方便,费用低廉。最大的不同在于,它是通过互联网进行传输的。云计算的蓝图已经呼之欲出:在未来,只需要一台笔记本或者一个手机,就可以通过网络服务来实现我们需要的一切,甚至包括超级计算这样的任务。从这个角度而言,最终用户才是云计算的真正拥有者^[14]。

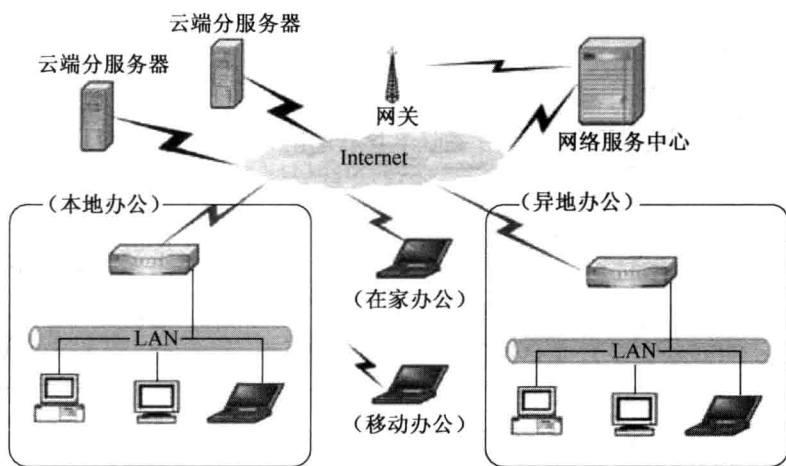


图 1-1 云计算的原理

1.6 云计算发展现状与趋势

由于云计算技术范围很广,目前各大 IT 企业提供的云计算服务主要根据自身的特点和优势来实现。国外的云计算平台以 Google、IBM、Amazon 为例,国内的云计算平台得到成果的例子主要有清华大学透明计算平台和云快线科技的 CloudEX 云计算服务平台。

1. Google 的云计算平台

作为最大的云计算使用者,Google 的搜索引擎建立在 200 多个地点、超过 100 万台服务器的支撑之上,且这些设施的数量还在迅猛增长^[21]。Google 以其具有的硬件条件优势、大型的数据中心和搜索引擎的支柱应用,促进 Google 云计算迅速发展。Google 的云计算主要由 MapReduce、Google 文件系统 (GFS)、BigTable 组成。它们是 Google 内部云计算基础平台的 3 个主要部分。Google 还构建其他云计算组件,包括领域描述语言以及分布式锁服务机制等。Sawzall 是一种建立在

MapReduce 基础上的领域语言,专门用于大规模的信息处理。Chubby 是一个高可用、分布式数据锁服务,当有机器失效时,Chubby 使用 Paxos 算法来保证备份^[13,30]。

2. IBM “蓝云” 计算平台^[13,30]

“蓝云” 解决方案是由 IBM 云计算中心开发的企业级云计算解决方案。该解决方案可以对企业现有的基础架构进行整合,通过虚拟化技术和自动化技术,构建企业自己拥有的云计算中心,实现企业硬件资源和软件资源的统一管理、统一分配、统一部署、统一监控和统一备份,打破应用对资源的独占,从而帮助企业实现云计算理念。

IBM 的“蓝云” 计算平台是一套软、硬件平台,将因特网上使用的技术扩展到企业平台上,使得数据中心使用类似于互联网的计算环境。“蓝云” 大量使用了 IBM 先进的大规模计算技术,结合了 IBM 自身的软、硬件系统以及服务技术,支持开放标准与开放源代码软件。

“蓝云” 基于 IBM Almaden 研究中心的云基础架构,采用了 Xen 和 PowerVM 虚拟化软件、Linux 操作系统映像以及 Hadoop 软件 (Google File System 以及 MapReduce 的开源实现)。IBM 已经正式推出了基于 x86 芯片服务器系统的“蓝云” 产品。

“蓝云” 计算平台由一个数据中心、IBM Tivoli 部署管理软件 (Tivoli provisioning manager)、IBM Tivoli 监控软件 (IBM Tivoli monitoring)、IBM WebSphere 应用服务器、IBM DB2 数据库以及一些开源信息处理软件和开源虚拟化软件共同组成。“蓝云” 硬件平台环境与一般的 x86 服务器集群类似,使用刀片的方式增加了计算密度。“蓝云” 软件平台的特点主要体现在虚拟机以及对于大规模数据处理软件 Apache Hadoop 的使用上。

“蓝云” 平台的一个重要特点是虚拟化技术的使用。虚拟化的方式在“蓝云” 中有两个级别,一个是在硬件级别上实现虚拟化,另一个是通过开源软件实现虚拟化。硬件级别的虚拟化可以使用 IBM p 系列的服务器,获得硬件的逻辑分区 LPAR (logic partition)。逻辑分区的 CPU 资源能够通过 IBM Enterprise Workload Manager 来管理。通过这样的方式加之在实际使用过程中的资源分配策略,能够使相应的资源合理地分配到各个逻辑分区。p 系列系统的逻辑分区最小粒度是 1/10 颗 CPU。Xen 则是软件级别上的虚拟化,能够在 Linux 基础上运行另外一个操作系统。

虚拟机是一类特殊的软件,能够完全模拟硬件的执行,运行不经修改的完整的操作系统,并保留了一整套运行环境语义。通过虚拟机的方式,在云计算平台上可获得如下一些优点: