

高等院校信息技术规划教材

云计算及应用

唐国纯 著



清华大学出版社

高等院校信息技术规划教材

云计算及应用

唐国纯 著

清华大学出版社
北京

内 容 简 介

本书系统总结了作者近几年在云计算应用方面的研究成果,是目前国内介绍云计算技术在不同行业领域应用较深刻的一部云计算著作。本书深入介绍了云计算技术在不同领域的应用,以云计算的领域应用为主线,论述了云计算的基础理论、云制造、教育云、环保云、物流云、云安全应用研究、移动云计算开发技术以及开源云计算平台 OpenStack 和 CloudStack,进而分析了多个领域的云计算深入应用。本书的特点是系统分析、架构设计和实际应用相结合,将云计算理论应用于制造行业、教育行业、环保行业、物流行业、信息安全和移动领域,知识涉及面广,形成了云制造、教育云、环保云、物流云、云安全、移动云计算开发技术等知识的纵横深入跨度体系。

本书补充了云计算和 Android 移动开发的基础知识,深度论述了云计算的不同领域应用,自成体系,既可作为信息科学的高年级本科生和研究生的教材,也可作为云计算领域的研究人员与工程人员以及其他行业信息化建设研究者的参考书。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

云计算及应用/唐国纯著. —北京:清华大学出版社,2015

高等院校信息技术规划教材

ISBN 978-7-302-40908-3

I. ①云… II. ①唐… III. ①计算机网络—高等学校—教材 IV. ①TP393

中国版本图书馆 CIP 数据核字(2015)第 169118 号



责任编辑:白立军 薛 阳

封面设计:常雪影

责任校对:李建庄

责任印制:刘海龙

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社 总 机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

课件下载: <http://www.tup.com.cn>, 010-62795954

印 刷 者:清华大学印刷厂

装 订 者:北京市密云县京文制本装订厂

经 销:全国新华书店

开 本:185mm×260mm

印 张:16.5

字 数:407千字

版 次:2015年8月第1版

印 次:2015年8月第1次印刷

印 数:1~2000

定 价:35.00元

产品编号:065237-01

前言

foreword

云计算是传统计算机技术和网络技术发展融合的产物,也是引领未来信息产业创新的关键战略性技术和手段。近年来,云计算已成为 IT 业界最热门的研究方向之一。几乎所有的主流 IT 厂商都在谈论云计算,既包括硬件厂商(IBM、英特尔等)、软件开发商(微软等),也包括互联网服务提供商(Google、Amazon、阿里巴巴、百度、腾讯等)和电信运营商(AT&T、中国移动等)。这些企业覆盖了整个 IT 产业链,构建了一个完整的云计算生态系统。云计算技术的兴起,提供了一种适应于各行业领域信息化发展需要的解决方法。云计算的出现为信息技术领域和企业信息化建设带来了新的挑战和机遇。然而,真正系统、深入、全面地阐述云计算概念和技术及领域应用的图书不多。本书作为一本全面、系统、深入论述云计算概念、技术和架构、领域应用的云计算专著,可以帮助对云计算领域应用感兴趣的读者理清相关的知识、理论和实践应用。本书论述了云计算的基础理论、云制造、教育云、环保云、物流云、云安全应用研究、移动云计算开发技术以及开源云计算平台 OpenStack 和 CloudStack。在写作过程中力求普及云计算的多领域应用、理论与实践相兼顾,力求给广大读者一个完整、正确和深入的云计算领域应用知识体系,作为大家日常学习、工作的案头参考书。本书结合作者多年对云计算领域应用的研究成果,对各章节结构做了精心的设计和安排,有较强的逻辑性、系统性、全面性、专业性和实践性。此外,本书还参考了其他作者发表在期刊、会议论文和网络日志等中的一些重要成果,以及一些公司提供的云计算应用解决方案,对此深表感谢。

本书相关研究获得海南省自然科学基金“基于模糊集和粗糙集的云安全综合评价模型”(614247)和海南省高等学校优秀中青年骨干教师的经费资助。谨向帮助、支持和鼓励我完成本书工作的我的家人和所有朋友致以深深的敬意和诚挚的感谢,感谢清华大学出版社为本书的出版所提供的大力帮助!

云计算是一个比较新的领域,由于作者自身知识水平有限,书中难免有疏漏和不当之处,敬请读者批评指正。

编 者

2015 年 5 月

目录

Contents

第 1 章 云计算理论研究综述	1
1.1 云计算相关概念	1
1.1.1 云计算的概念	1
1.1.2 云计算服务形式	2
1.1.3 云计算的部署配置模式	3
1.1.4 云计算的特征	4
1.1.5 云计算应用存在的主要问题	5
1.2 云计算的主要技术	6
1.2.1 Google 云计算的技术架构	6
1.2.2 亚马逊云计算 AWS	11
1.2.3 微软云平台体系架构	13
1.2.4 开源云计算系统 Hadoop	14
1.2.5 虚拟化技术	15
1.2.6 代表性云计算方案及服务比较	17
1.3 云计算体系结构中的多层次研究	19
1.3.1 云计算的定义与特点	19
1.3.2 云应用系统的服务架构层次	19
1.3.3 云计算中的 QoS 与 SLA 层次	24
1.3.4 云计算的运维管理	27
1.3.5 云主机层次	29
1.4 传统的架构设计过程模型	30
1.4.1 传统面向组件的软件体系结构	30
1.4.2 采用 RUP 传统的架构设计过程模型	30
1.4.3 软件架构质量要求	32
1.5 基于 MDA 软件设计方法研究	34
1.5.1 模型、软件架构和框架的定义及其使用模型 开发的优点	35

1.5.2	MDA 的内涵及其开发过程	35
1.5.3	模型在软件架构中的层次和设计地位	36
1.5.4	MDA 的框架开发模型	38
1.6	5 层驱动模型的 SaaS 架构设计	40
1.6.1	从模块编程到面向对象,再到 SOA、SaaS 发展历程	40
1.6.2	SaaS 的成熟度模型及其演化过程	41
1.6.3	基于 5 层驱动模型的 SaaS 架构设计	42
1.7	大数据	44
1.7.1	大数据的概念与特点	44
1.7.2	建立以数据为中心的云计算应用	45
	参考文献	47
第 2 章 云制造——云计算在制造行业中的应用		50
2.1	云制造的概念	50
2.1.1	云制造的研究现状	50
2.1.2	云制造内涵	51
2.1.3	云制造设备的服务化封装与云端化接入模型	51
2.2	云制造体系结构	53
2.2.1	云制造的运行与应用模式	53
2.2.2	云制造虚拟资源构建	53
2.2.3	云制造平台总体结构	56
2.2.4	云制造中的关键技术	56
2.3	基于云计算的 PLM 与 ERP 的集成方案研究	58
2.3.1	PLM 系统与 ERP 系统集成分析	58
2.3.2	PLM 系统与 ERP 系统集成方式	58
	参考文献	61
第 3 章 教育云——云计算在教育行业中的应用		63
3.1	教育云的概念及其优势	63
3.1.1	教育云的概念	63
3.1.2	教育云的优势	64
3.2	教育云的体系结构及核心特征	64
3.2.1	教育云的体系结构	64
3.2.2	教育云的核心特征	66
3.3	教育云的应用模式与关键技术	67
3.3.1	教育云的应用模式	67
3.3.2	教育云的关键技术	68

3.4 基于 SOA 架构的教育软件开发平台框架的研究	69
3.4.1 教育软件的概念及其数字化校园信息的建设内容	69
3.4.2 基于 SOA 架构的数字化校园信息平台的框架	71
3.4.3 基于 SOA 架构的教育软件开发平台的框架和运营模式	72
3.5 云计算中面向服务组合的资源库建设研究	74
3.5.1 云计算环境下面向服务组合模式构建资源库的内涵	74
3.5.2 云计算环境下面向服务组合模式构建资源库的架构	77
3.5.3 区域教育云系统的部署模式与结构	77
3.6 浪潮教育云应用	79
3.6.1 浪潮区域教育云解决方案	79
3.6.2 教育数据中心解决方案	80
参考文献	80
第 4 章 环保云——云计算在环保行业中的应用	84
4.1 环保云的相关概念及其优势	84
4.1.1 环保云的概念	84
4.1.2 实施环保云的优势	85
4.2 环保云的体系结构及技术架构	86
4.2.1 环保云的平台结构	86
4.2.2 环保云的功能结构	87
4.2.3 环保云 SOA 架构	87
4.2.4 环保云的信息资源层次框图	87
4.3 基于云计算和物联网的环境保护云平台及关键技术	89
4.3.1 环保云的管理平台	89
4.3.2 环保云的关键技术	89
4.4 环保云应用方案	91
4.4.1 联想 PM2.5 云监测平台	91
4.4.2 浪潮环保云数据中心	93
参考文献	95
第 5 章 物流云——云计算在物流行业中的应用	97
5.1 物流云的相关概念及要求	97
5.1.1 物流云的概念	97
5.1.2 物流云的业务架构、运作模型与特征	98
5.2 物流云公共信息平台的体系结构	100
5.2.1 物流公共信息平台功能设计	101
5.2.2 物流云公共信息平台 subsystem 设计	102

5.2.3	物流云的公共信息平台体系结构	103
5.3	物流云管理平台的体系结构及关键技术	105
5.3.1	物流云管理平台的体系结构	105
5.3.2	物流云服务的关键技术	106
5.4	物流园区管理平台的云服务系统结构	107
5.4.1	物流服务模式	107
5.4.2	物流园区物流云服务系统体系结构	107
5.5	物流云技术方案架构示例	109
5.5.1	方案架构示例一：公有云方案	109
5.5.2	方案架构示例二：混合云方案	110
5.5.3	铁路“智能物流”架构	110
	参考文献	111

第6章 云安全应用研究

113

6.1	云安全的体系结构及关键技术研究	113
6.1.1	云计算实施要点及管理平台的主要功能	113
6.1.2	云服务信息安全框架	115
6.1.3	计算安全和安全云的安全原理与关键技术	116
6.2	基于SLA的云安全模型研究	118
6.2.1	云服务水平协议结构	119
6.2.2	云服务信息安全框架与等级保护基本安全要求	119
6.2.3	基于云安全作为一种服务的云服务系统的定价与收费	121
6.3	基于云计算的物联网安全研究	122
6.3.1	基于云计算的物联网系统架构	123
6.3.2	基于云计算的物联网安全模型	123
6.3.3	物联网统一的云安全管理解决方案	125
6.3.4	基于云计算的物联网安全薄弱环节分析与建议	126
6.4	基于多层次模糊集的云安全评估模型	127
6.4.1	模糊综合评判法	127
6.4.2	云安全评价指标选取与指标权重确定	128
6.4.3	云安全模糊综合评价	132
6.5	基于粗糙集理论的云安全评估模型研究	134
6.5.1	粗糙集理论的基本概念与综合评价流程	135
6.5.2	云安全评估的粗糙集模型	136
6.5.3	决策表的属性约简	137
	参考文献	140

第 7 章 移动云计算开发技术	143
7.1 移动云计算的发展概况	143
7.1.1 移动云计算的概念	143
7.1.2 移动云计算的案例	147
7.1.3 移动应用面临的问题	148
7.1.4 Android 系统概述	149
7.2 Android 系统下 HTML 5 的应用开发	151
7.2.1 HTML 结构与 CSS	151
7.2.2 HTML 5 编写规范	154
7.2.3 新增的主体结构元素	156
7.2.4 新增的非主体结构元素	160
7.2.5 网页编排示例	162
7.2.6 Android 网页编排示例	162
7.3 云视野下电子书包的开发应用	169
7.3.1 电子书包的现状	169
7.3.2 云学习资源的特征与开发要求	170
7.3.3 云学习视角下的电子书包系统架构	171
7.4 Android 系统下电子书包开发常用技术	177
7.4.1 Web 程序与 Android 应用程序的交互	177
7.4.2 使用 Android 应用调用 Webservice	183
7.4.3 Android 平台下的数据分享	190
7.4.4 Android 电子书翻页效果的实现	195
7.5 Android 平台下消息的云推送	232
7.5.1 消息推送基础原理	232
7.5.2 消息云推送解决方案概述	232
7.5.3 AndroidPN 平台的信息推送示例	237
参考文献	240
第 8 章 开源云计算平台 OpenStack 和 CloudStack	242
8.1 开源云计算平台现状	242
8.1.1 开源云计算平台列举	242
8.1.2 成功实施 OpenStack 和 CloudStack 项目建议	243
8.2 OpenStack	244
8.2.1 OpenStack 介绍	245
8.2.2 OpenStack 安装	247



8.3 CloudStack	247
8.3.1 CloudStack 部署架构与软件架构	248
8.3.2 CloudStack 安装	250
8.3.3 CloudStack 和 OpenStack 的比较	250
参考文献	251

云计算理论研究综述

据 Occams Business 研究与咨询公司的预测,到 2020 年,全球云市场价值将从 2013 年的 900 亿美元增长到 6500 亿美元,复合增长率将达 29%。在云服务中,全球的平台即服务市场增长率最高,预计到 2020 年复合增长率达 39%。在地理位置上,目前亚太地区是增长最快的地区,复合增长率为 35%。该报告在服务模型、部署模型和区域基础之上讨论了全球云计算市场未来 5 年的发展趋势。随着大数据管理必要性的提升,刺激着云计算的需求呈指数形式增长。云计算技术大部分用于在线媒体应用,如 Dropbox、Gmail、Face book、Evernote 和 Skype 等。云计算是一种基于互联网的技术,作为互联网所连接的远程设备的中心数据资源。它是一种网络,所有的程序或应用都运行在一台服务器上,同时跨多台设备共享,如 PC、平板、智能手机。

1.1 云计算相关概念

1.1.1 云计算的概念

“云计算”这个全球关注度最高的 IT 词汇,谷歌、微软、亚马逊、IBM 等业界领导厂商对其分别有着不同的阐述,还有二十多位专家的诠释,“仁者见仁智者见智”的表述依然令人们不知所措。现阶段广为接受的是美国国家标准与技术研究院(NIST)的定义:云计算是一种按使用量付费的模式,这种模式提供可用的、便捷的、按需的网络访问,进入可配置的计算资源共享池(资源包括网络、服务器、存储、应用软件、服务),这些资源能够被快速提供,只需投入很少的管理工作,或服务供应商进行很少的交互。从技术的角度来讲,云计算是一种颠覆性的交付模式,一体化的信息平台 and 运营平台。它将企业所有的服务器、存储等基础设施以及网络整合到统一的云平台上。在“云的世界”里,将技术和业务结合起来交付给用户使用,企业的运营管理决策分析都将基于云平台展开。从细分角度来看,云由云计算平台和云服务应用两个层面组成。企业可以将基础设施包括传统的服务器、操作系统、存储运维等都统一部署在一个平台上,这是一个技术的层面,企业可以不必过多地关注这个平台本身,而是关注应用。政府、企业和个人可以根据不同的需求部署成不同的应用,形成个性化的交付模式,形成一种云服务。一个是云技术层面,一个是云服务层面。

1.1.2 云计算服务形式

从用户体验的角度来讲,云计算可以分为以下几种交付模式:基础设施即服务(IaaS)、开发平台即服务(PaaS)、软件应用即服务(SaaS)、数据存储即服务(DaaS)和后端即服务(BaaS)。

IaaS(Infrastructure as a Service):基础设施即服务。消费者通过 Internet 可以从完善的计算机基础设施获得服务。比如 IBM 的无锡云计算中心、世纪互联的 CloudEx 云主机等。

PaaS(Platform as a Service):平台即服务。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。例如,软件的个性化定制开发,比如 Google 的 APP Engine 和 Salesforce 的 Force.Com 等。

SaaS(Software as a Service):软件即服务。它是一种通过 Internet 提供软件的模式,用户无须购买软件,而是向提供商租用基于 Web 的软件,来管理企业经营活动。例如,阳光云服务器。

DaaS(Data as a Service):数据存储即服务。这是网络上提供虚拟存储的一种服务方式,客户可以根据实际存储容量来支付费用。例如,亚马逊的 EC2、中国电信上海公司与 EMC 合作的“e 云”等。

BaaS(Backend as a Service):后端即服务。这是一种新型的云服务,起源于 MBaaS(Mobile Backend as a Service),即移动后端即服务,旨在为移动和 Web 应用提供后端云服务,包括云端数据/文件存储、账户管理、消息推送、社交媒体整合等。BaaS 是垂直领域的云服务,随着移动互联网的持续火热,BaaS 也受到越来越多的开发者的青睐。它作为应用开发的新模型,可以降低开发者成本,让开发者只需专注于具体的开发工作。随着移动互联网的发展,移动行业的分工也会像其他行业一样逐渐细化,后端服务就是这样被抽象出来,它统一向开发者提供文件存储、数据存储、推送服务等实现难度较高的功能,以帮助开发者快速开发移动应用。BaaS 的特征包括以下 4 个方面。

(1) 必须是与网络相关的服务。有些 B2D 服务与网络无关,那么就不是 BaaS,比如游戏引擎,使用大型 3D 引擎开发游戏一般需要付费授权,但这个并不需要联网,所以它就不是 BaaS。

(2) 必须嵌入在终端应用中,间接地为消费者提供服务。因此云测试虽然也依靠于云,但并不属于 BaaS。

(3) 必须是弹性可扩充的。这其实是云服务的特征,所以,非云的网络服务也不是 BaaS。

(4) 按使用量计费。这个其实就是云时代的典型商业模式。这些服务就像水和电一样,给钱就能用,不给就停掉,就这么简单。

在国外,BaaS 已经受到巨头的重视。2013 年 4 月,Facebook 收购 Parse;2014 年 6 月,苹果发布了 CloudKit;2014 年 10 月,Google 收购了 Firebase。Parse、CloudKit、Filrebase 都是国外知名的 BaaS 类产品。苹果和谷歌通过 BaaS 可以更好地完善其生态

圈,Parse 也可以帮助 Facebook 建立它在移动端的地位。从巨头们在 BaaS 方面的布局也可以看出 BaaS 的价值。在国内,提供 BaaS 的厂商也有很多,典型的代表有 APICloud、Bmob、友盟,主要提供的功能包括社会化媒体集成、数据/文件存储、数据分析、消息推送、支付。大多数 BaaS 从本质上是属于 SaaS 的,因为所提供的 API 也算是软件服务,但如云存储则又涉及 PaaS,所以说 BaaS 是针对特殊领域的综合性服务,和 SaaS 等是不同的划分关系,是云服务的一个分支。

无论是 SaaS、PaaS 还是 IaaS,其核心概念都是为用户提供按需服务。于是产生了“一切皆服务”(Everything as a Service,EaaS 或 XaaS)的理念。基于这种理念,以云计算为核心的创新型应用不断产生。企业和研究机构最终会把高级别的计算任务交给全球运行的服务器网络,也就是云。以下领先者在这一领域占据着主导位置,如表 1-1 所示。

表 1-1 几个云计算领先者的比较

Google	Yahoo	IBM	Microsoft	Amazon
唯一以硬件起家的搜索公司。每年在数据中心的投入超过 20 亿美元。成为云计算领域难以超越的领跑者和极力推动者	规模和资金比 Google 稍逊一筹,开发的软件与云计算兼容不够。但是作为 Hadoop 的首要资助方,可能后来居上	商业数据计算的龙头和传统超级计算机的绝对领导者。与 Google 合作后立足云计算一方。为越南政府开发了飞行员“云”系统试点。并在无锡成立了数据中心	现在只能与自身开发的软件结合,这可能成为它的软肋。但是在“云”科学基础理论中扮演重要的角色。正在伊利诺伊州和西伯利亚建立大型数据中心	第一个将云计算作为服务出售的公司。规模小于其他竞争者,但是是在该领域的专业性为这家零售商在下一代网络服务方面从零售到传媒业的转型助了一臂之力

1.1.3 云计算的部署配置模式

云计算的部署配置模式包括公有云、私有云、混合云、社区云。社区云也可以称为行业云,即以行业为中心,将供应链上的所有产业群围绕云共享服务,展开商业活动。

公有云(Public Cloud): 简而言之,是由云服务提供商运营,为最终用户提供从应用程序、软件运行环境,到物理基础设施等各种各样的 IT 资源。在该方式下,云服务提供商需要保证所提供资源的安全性和可能性等非功能性需求,而最终用户不关心具体资源由谁提供、如何实现等问题。

私有云(Private Cloud): 私有云是由企业自建自用的云计算中心,具备许多公有云环境的优点,例如弹性、适合提供服务。两者的差别在于私有云服务中,数据与程序皆在组织内管理,且与公有云服务不同,不会受到网络带宽、安全疑虑、法规限制影响;此外,私有云服务让供应者及用户更能掌控云基础架构、改善安全与弹性,因为用户与网络都受到特殊限制。

混合云(Hybrid Cloud): 混合云是把“公共云”和“私有云”结合在一起的方式。在这个模式中,用户通常将非企业关键信息外包,并在公用云上处理,但同时掌控企业关键服务及数据。

社区云(Community Cloud): 社区云由众多利益相仿的组织掌控及使用,例如特定安全要求、共同宗旨等。社区成员共同使用云数据及应用程序。

云计算归根结底是一种 IT 服务提供模式,不论是公有云还是私有云(以 IT 设备的归属不同分类),其本质都是 IT 的最终使用者可以随时随地并且简便快速地获取 IT 服务,并以获取服务的层次分为 IaaS(仅获取虚拟的硬件资源)、PaaS(获取可编程的环境)、SaaS(直接获取软件应用服务)。云计算服务通过公有、私有和混合云模式部署。目前,目前,公有云引领着云市场,占据着大量的市场份额。采用公有云的一个主要原因是“按需支付”的成本效益模型。另外,它还通过优化运营、支持和维护服务给云服务供应商带来了规模经济。私有云是使用第二多的模式,因为它相对来说更安全。混合云模型目前市场中占有的份额最少,但未来的发展空间巨大。用户可以根据其需求,选择适合自己的云计算模式。

1.1.4 云计算的特征

云计算(Cloud Computing)是分布式计算(Distributed Computing)、并行计算(Parallel Computing)、效用计算(Utility Computing)、网络存储(Network Storage Technologies)、虚拟化(Virtualization)、负载均衡(Load Balance)、热备份冗余(High Available)等传统计算机和网络技术发展融合的产物。被人们普遍接受的云计算具有如下特点。

(1) 超大规模。“云”具有相当的规模,Google 云计算已经拥有一百多万台服务器,Amazon、IBM、微软、Yahoo 等的“云”均拥有几十万台服务器。企业私有云一般拥有数百上千台服务器。“云”能赋予用户前所未有的计算能力。

(2) 虚拟化。云计算支持用户在任意位置、使用各种终端获取应用服务。所请求的资源来自“云”,而不是固定的有形的实体。应用在“云”中的某处运行,但实际上用户无须了解、也不用担心应用运行的具体位置。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需要的一切,甚至包括超级计算这样的任务。

(3) 高可靠性。“云”使用了数据多副本容错、计算节点同构可互换等措施来保障服务的高可靠性,使用云计算比使用本地计算机可靠。

(4) 通用性。云计算不针对特定的应用,在“云”的支撑下可以构造出千变万化的应用,同一个“云”可以同时支撑不同的应用运行。

(5) 高可扩展性。“云”的规模可以动态伸缩,满足应用和用户规模增长的需要。

(6) 按需服务。“云”是一个庞大的资源池,可按需购买。消费者无须同服务提供商交互就可以自动地得到自助的计算资源能力,如服务器的时间、网络存储等(资源的自助服务)。

(7) 极其廉价。云系统对服务类型通过计量的方法来自动控制和优化资源使用(例如存储、处理、带宽以及活动用户数)。资源的使用可被监测、控制以及对供应商和用户提供透明的报告(即付即用的模式)。由于“云”的特殊容错措施可以采用极其廉价的节点来构成云;“云”的自动化集中式管理使大量企业无须负担日益高昂的数据中心管理成本;“云”的通用性使资源的利用率较之传统系统大幅提升,因此用户可以充分享受“云”

的低成本优势,经常只要花费几百美元、几天时间就能完成以前需要数万美元、数月时间才能完成的任务。云计算可以彻底改变人们未来的生活,但同时也要重视环境问题,这样才能真正为人类进步做贡献,而不是简单的技术提升。

(8) 潜在的危险性。云计算服务除了提供计算服务外,还必然提供了存储服务。但是云计算服务当前垄断在私人机构(企业)手中,而他们仅能够提供商业信用。政府机构、商业机构(特别像银行这样持有敏感数据的商业机构)对于选择云计算服务应保持足够的警惕。一旦商业用户大规模使用私人机构提供的云计算服务,无论其技术优势有多强,都不可避免地会让这些私人机构有机会以“数据(信息)”的重要性挟制整个社会。对于信息社会而言,“信息”是至关重要的。另一方面,云计算中的数据对于数据所有者以外的其他云计算用户是保密的,但是对于提供云计算的商业机构而言却是毫无秘密可言。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的前提。

1.1.5 云计算应用存在的主要问题

尽管使用云计算服务优点众多,但作为一项新生事物,云计算仍然面临一些问题,使人们对其仍然抱有怀疑和观望的态度。主要体现在以下几个方面。

(1) 安全。安全是云计算面临的首要问题。云计算意味着企业将把类似客户信息这类具有很高商业价值的数据存放到云计算服务提供商的手中,信息的安全性和私密性是用户非常关心的事情。对于严重依赖云计算的个人或企业用户,一旦服务提供商出现安全问题,他们存储在云中的数据可能会被长期遗忘在某个角落里甚至像石沉大海般消失得无影无踪。Google 等云计算服务提供商造成的数据丢失和泄漏事件时有发生,这表明云计算的安全性和可靠性仍有待提高。根据 IDC 的调查结果,将近 75% 的受访企业认为安全是云计算发展路途上的最大挑战。相当数量的个人用户对云计算服务尚未建立充分的信任感,不敢把个人资料上传到“云”中,而观念上的转变和行为习惯的改变则非一日之功。安全已经成为云计算业务拓展的主要困扰之一。

(2) 技术。建立云计算系统是一个技术挑战。必须购买或征用数百或数千台个人计算机或服务器并将它们连在一起进行并行管理,并且需要开发功能丰富的软件以提供 24×7 的 Web 应用。此外,目前众多云计算服务提供商各自云计算服务的技术和标准还不统一,用户在选择时面临不少困惑。

(3) 经济。建立云计算服务花费巨大,只有大公司才可能承担如此大规模的资源,那么这些提供云计算服务的公司如何获得足够的回报将是一个重要的问题。另外,收费模式和定价都是十分困难的事情,云计算将像其他新技术一样遇到盈利模式的问题。毕竟企业对于现有本地数据和业务往往已经建立了专有的数据中心,是否迁移以及如何以更低成本迁移到云计算平台之中是个不小的困扰。

(4) 网络。持久的宽带互联网接入是云计算成功运行的基本前提,但是目前接入是网络发展最主要的瓶颈,有赖于政府和企业投入更多的资源来提高接入的带宽和质量。

(5) 兼容性。用户已经习惯于现有的操作系统和文件系统,云计算要实现跨平台的服务,就必须保证现有文件格式与未来基于 Web 应用的文件格式能够兼容,否则很难使

大多数用户向云计算迁移。

1.2 云计算的主要技术

云计算系统运用了许多技术,包括编程模型、数据管理技术、数据存储技术、虚拟化技术、云计算平台管理技术等。

1.2.1 Google 云计算的技术架构

Google 的云计算应用均依赖于 4 个基础组件,即分布式文件存储 GFS、并行数据处理模型 MapReduce、分布式锁 Chubby 和结构化数据表 BigTable。Google 的云计算技术架构如图 1-1 所示。



图 1-1 Google 的云计算技术架构

在这些组件中,其调用关系如图 1-2 所示。

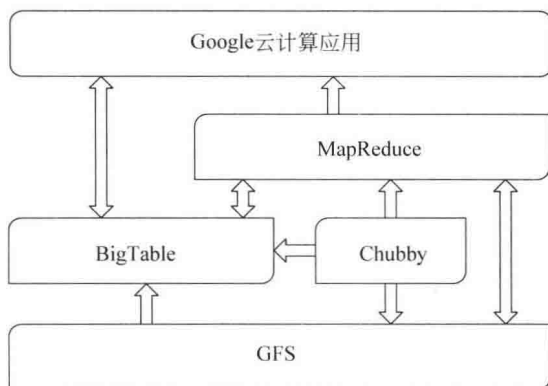


图 1-2 Google 的云计算组件调用关系

1. 分布式锁 Chubby

Chubby 系统本质上就是一个分布式的、存储大量小文件的文件系统。它是 Google 为解决分布式一致性问题而设计的提供粗粒度锁服务的文件系统,其他分布式系统可以使用它对共享资源的访问进行同步。根据它的设计目标,需要实现的特性有高可用性、高可靠性、支持粗粒度的建议性锁服务和支持小规模文件直接存储 4 个方面,而对高性

能和存储能力两个特性一般不做考虑。Chubby 中的锁就是文件,在 GFS 的例子中,创建文件就是进行“加锁”操作,创建文件成功的那个 Server 其实就是抢占到了“锁”。用户通过打开、关闭和存取文件,获取共享锁或者独占锁;并且通过通信机制,向用户发送更新信息。

Chubby 的作用如下。

- (1) 为 GFS 提供锁服务,选择 Master 节点;记录 Master 的相关描述信息。
- (2) 通过独占锁记录 Chunk Server 的活跃情况。
- (3) 为 BigTable 提供锁服务,记录子表元信息(如子表文件信息、子表分配信息、子表服务器信息)。
- (4) (可能)记录 MapReduce 的任务信息。
- (5) 为第三方提供锁服务与文件存储。

Chubby 的系统架构如图 1-3 所示。

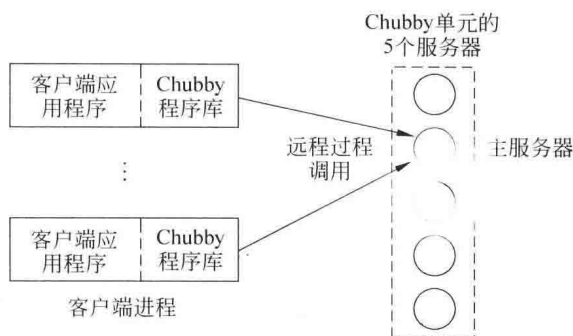


图 1-3 Chubby 的系统架构

2. 分布式文件存储 GFS

GFS 的容错机制包括 Chunk Server 容错和 Master 容错(影子节点热备)两个方面。分布式文件存储 GFS 的结构如图 1-4 所示。

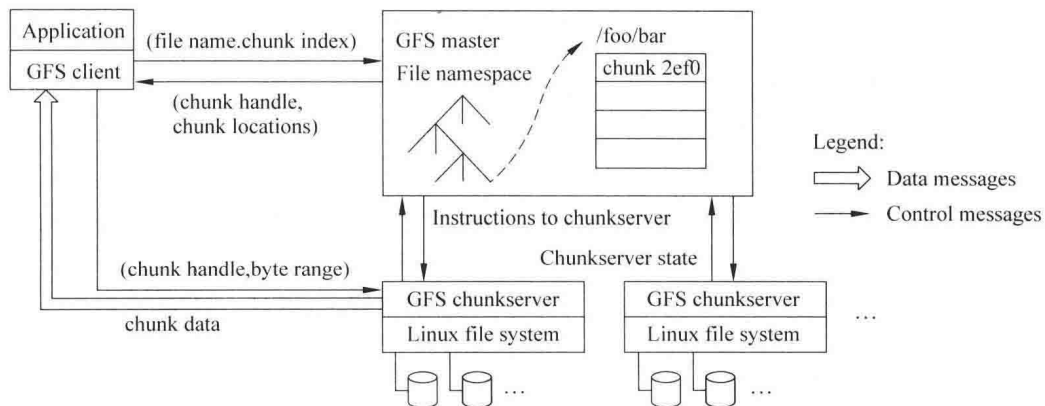


图 1-4 分布式文件存储 GFS 的结构