

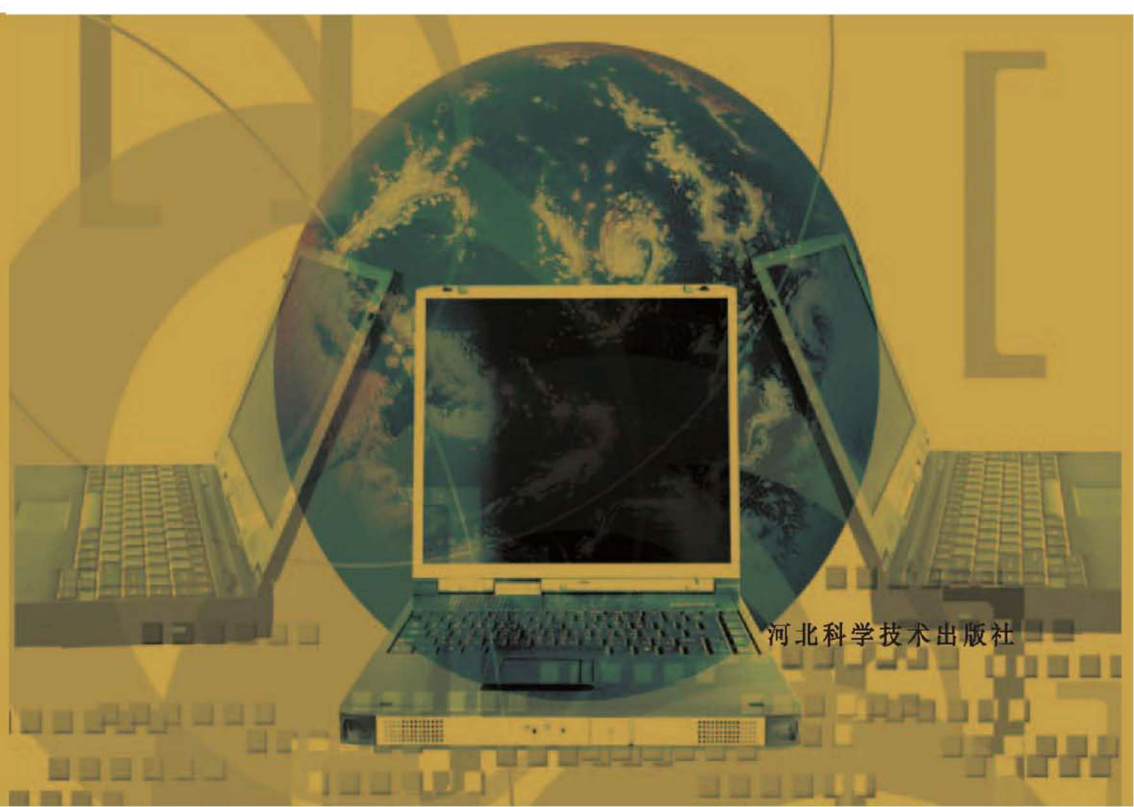
国家科技支撑计划项目，编号：2012BAH08B02

国家自然科学基金项目，编号：61272513

赵晓永/著

# 面向云计算的 数据存储关键技术研究

分析了云计算环境下数据存储的关键技术，针对小文件存储、副本放置策略、数据去重方法和系统节能等关键技术进行了重点研究，并建立了数学模型，提出了相应的优化策略、机制和算法。

The cover features a central globe with a laptop screen in front of it, all set against a background of server racks and abstract geometric shapes. The color palette is primarily yellow and green.

河北科学技术出版社

国家科技支撑计划项目，编号：2012BAH08B02

赵晓永/著

国家自然科学基金项目，编号：61272513

# 面向云计算的 数据存储关键技术研究

封面



河北科学技术出版社

## 图书在版编目(CIP)数据

面向云计算的数字存储关键技术研究 / 赵晓永著.  
—石家庄:河北科学技术出版社,2014.12

ISBN 978-7-5375-7380-1

I. ①面… II. ①赵… III. ①计算机网络-应用-数据  
管理-研究 IV. ①TP274

中国版本图书馆 CIP 数据核字(2014)第 281979 号

图片提供:(c)IMAGEMORE Co., Ltd.

MIANXIANG YUNJISUAN DE SHUZI CUNCHU GUANJIAN JISHU YANJIU

面向云计算的数字存储关键技术研究

赵晓永 著

---

出版 河北科学技术出版社  
地址 石家庄市友谊北大街 330 号(邮编:050061)  
经销 新华书店  
印刷 河北新华第二印刷有限责任公司  
开本 700 毫米×1000 毫米 1/16  
印张 8  
字数 120 千字  
版次 2014 年 12 月第 1 版  
印次 2014 年 12 月第 1 次印刷  
定价 25.00 元

---

# 前 言

随着整个社会信息化程度的不断提高，特别是在物联网和移动互联网的大力推进下，各行业的应用信息处理和数据服务的规模越来越大，各类数据密集型应用层出不穷，数字化信息规模的爆炸式增长推动了人们对信息存储的巨大需求。如此海量的数据最终分布式地存储在云中各个不可信节点之上，这就对云中数据的分布式存储模型提出了新的要求，如何高效地对云存储系统中的海量数据进行存储和管理，已经成为云计算时代面临的新的挑战和考验。

本书在全面分析云计算环境下分布式数据存储相关技术基础之上，重点针对小文件存储机制、副本放置策略、数据去重方法和系统能耗优化等关键问题进行了深入研究，建立了数学模型，提出了相应的优化策略、机制和算法；此外，为有效支撑本书所提出的模型、算法并进行合理验证，本书设计并搭建北京科技大学校园云存储原型系统，作为理论研究的有益补充，具有重要现实意义。

本书主要研究内容可概括为以下几个方面：

(1) 针对云中海量小文件带来的元数据管理复杂、内存占用过高、访问性能及吞吐量下降问题，本书提出一种基于元数据聚类的小文件存储机制 SFMC (Small Files Metadata Clustering)。在全面分析云存储中小文件问题的基础上，充分利用文件自身包含的丰富的描述元信息，通过元数据谱聚类算法，将逻辑相关性强的小文件合并存储到同一物理文件中，大幅减少了文件系统

中的文件数量；同时引入一级索引和扩展元数据管理节点，较好地解决了海量小文件的管理复杂性和访问性能问题，在适用性和复杂度之间取得了较好的平衡。

(2) 针对目前云计算存储能耗增加、环境资源压力大，而已有关于副本放置的研究对能耗关注不够的问题，本书提出一种面向节能的副本放置方法 ECRP (Energy Conservation Replica Placement)。基于硬件部件的三个主要不同能耗级别的工作状态，将数据节点划分为热区 (Hot Zone)、温区 (Warm Zone) 和冷区 (Cold Zone) 三个逻辑区域，各区域采用不同的电源管理策略和副本放置策略，并在构建出基于模拟退火 BP 神经网络的文件活性预测模型基础上，根据文件的活性不同将文件分别存储于不同区域的数据节点上，周期性计算各区域文件的活性，当文件活性超过某个阈值时，文件即开始在热区、温区和冷区之间逐级进行双向流动。实验表明，这种方法较好地解决了副本放置与能耗的优化问题，具备良好的节能效果。

(3) 针对目前的数据去重方法对图片、音乐和视频类型的多媒体文件处理时不够健壮，不能有效识别和去除经过变换后的文件，导致去重率不高，且不适应海量数据在线处理要求的问题，本书结合多媒体文件的感知相似特性，提出一种基于感知相似的海量文件在线近似去重方法 NPDF (Near Perception Fingerprint Deduplication)，显著地提高了去重率，同时加快了平均文件处理时间，减少了网络带宽占用。实验表明，该方法具备良好的时间和空间效率，能够有效支撑海量数据处理，具有良好的实用价值。

(4) 基于 Hadoop 平台设计并搭建了校园云存储原型系统，在该原型系统中实现并验证本书提出的小文件存储机制、节能的副本放置策略和感知相似的近似去重方法，通过北京科技大学校内一定时期的实践，已初步体现出了系统优越性，验证了本书提出的各项方法的可靠性及可行性。

综上所述，本书围绕云计算环境下数据存储的关键技术问题展开研究，给出了有效的解决方法和策略，设计了合理有效的模型及算法，并最终通过

原型系统予以支撑及验证。

本书得到国家自然科学基金（项目编号：61272513）、国家科技支撑计划项目（项目编号：2012BAH08B02）和北京市属高等学校人才强教计划资助项目（项目编号：PHR201106133）的资助，在此一并表示感谢。

本书适合从事云计算、大数据及分布式系统研究的相关人员参考，也可以作为高等学校高年级本科生和研究生的学习参考书，以及教师教学的参考资料。

由于作者水平有限，书中的不当和疏漏之处，敬请广大读者不吝指正。

赵晓永

2014年8月

# 目 录

1	引言 .....	1
2	绪论 .....	3
2.1	概述 .....	3
2.1.1	云计算 .....	3
2.1.2	云数据存储 .....	9
2.2	发展现状与关键技术 .....	12
2.2.1	发展现状 .....	12
2.2.2	关键技术 .....	22
2.3	研究内容及创新点 .....	30
2.3.1	主要研究内容 .....	30
2.3.2	主要创新点 .....	31
2.4	本书的组织 .....	32
2.5	本章小结 .....	35
3	基于元数据聚类的小文件存储机制 .....	36
3.1	研究现状 .....	36
3.2	聚类算法 .....	38
3.2.1	聚类算法概述 .....	39
3.2.2	K 均值算法 .....	40
3.2.3	谱聚类算法 .....	42

3.3 基于谱聚类的元数据聚类算法 .....	45
3.3.1 算法描述 .....	45
3.3.2 算法复杂度 .....	46
3.4 基于元数据聚类的小文件存储机制 .....	46
3.4.1 一级索引 .....	47
3.4.2 文件写入 .....	48
3.4.3 文件访问 .....	48
3.4.4 扩展元数据管理 .....	48
3.5 实验结果和分析 .....	49
3.5.1 实验环境 .....	49
3.5.2 实验结果 .....	50
3.6 本章小结 .....	52
<b>4 面向节能的副本放置方法 .....</b>	<b>53</b>
4.1 副本放置研究现状 .....	53
4.2 模拟退火 BP 神经网络 .....	57
4.2.1 BP 神经网络 .....	57
4.2.2 模拟退火 BP 神经网络 .....	60
4.3 文件活性模型的建立 .....	61
4.3.1 选取样本数据 .....	61
4.3.2 设计 BP 网络结构 .....	62
4.3.3 文件活性预测模型 .....	65
4.4 面向节能的副本放置方法 .....	68
4.4.1 区域划分 .....	69
4.4.2 正向迁移算法 .....	71
4.4.3 反向迁移算法 .....	72
4.5 实验结果和分析 .....	73



---

4.5.1	实验环境 .....	73
4.5.2	实验结果 .....	74
4.6	本章小结 .....	78
<b>5</b>	<b>基于感知相似性的数据去重方法 .....</b>	<b>79</b>
5.1	研究现状 .....	79
5.2	感知相似度模型 .....	81
5.2.1	感知相似 .....	81
5.2.2	局部敏感哈希 .....	82
5.2.3	感知相似度模型 .....	84
5.3	感知近似去重方法 .....	85
5.3.1	布隆过滤器 .....	85
5.3.2	算法描述 .....	87
5.3.3	算法复杂度 .....	89
5.4	实验与分析 .....	89
5.4.1	实验环境 .....	89
5.4.2	实验参数 .....	90
5.4.3	实验步骤 .....	90
5.4.4	实验结果 .....	90
5.5	本章小结 .....	93
<b>6</b>	<b>校园云存储原型系统的设计实现 .....</b>	<b>94</b>
6.1	校园云存储原型系统概述 .....	94
6.1.1	系统总体架构 .....	95
6.1.2	系统逻辑架构及相关技术 .....	96
6.2	系统功能介绍 .....	98
6.2.1	系统功能结构 .....	98
6.2.2	系统主要功能介绍 .....	100

6.3 本章小结 .....	102
<b>7 总结与展望 .....</b>	<b>103</b>
7.1 主要结论 .....	103
7.2 研究展望 .....	104
<b>参考文献 .....</b>	<b>106</b>

# 1 引言

如今，IT 行业进入到了一个关键的变革阶段，这场变革正在改变人们对于 IT 资源的使用习惯和方式，改变人们对于信息的获取、分享和交流方式，云计算正是驱动这一全新浪潮的最大动力源泉。云计算是一种基于互联网的计算方式，是分布式计算（Distributed Computing）、并行计算（Parallel Computing）和网格计算（Grid Computing）不断发展的结果，它提供了一种全新的计算模式与商业实现<sup>[1]</sup>。

随着整个社会信息化程度的不断提高，特别是在物联网和移动互联网的大力推动下，各行业的应用信息处理和数据服务的规模越来越大，各类数据密集型应用层出不穷，搜索引擎、各类互联网服务、企业的信息系统、航天、新能源、核能模拟、地球物理、气候预测以及生物信息、太空探索等科学计算类应用所产生的数据总量呈现几何级数增长，IDC 的研究报告指出，2010 年全球产生了 1.2ZB（1ZB=1 万亿 GB）数据<sup>[2]</sup>，数字化信息规模的爆炸式增长推动了人们对信息存储的巨大需求。图灵奖获得者 Jim Gray 提出了一个存储界的新摩尔定律，即每 18 个月新增的存储量等于有史以来存储量之和，UC Berkley 研究显示，未来 3 年内所产生的数据将超过过去 4 万年中产生数据的总和，且 93% 新生成的信息将以数字形式存在。庞大的数据量在为人们提供更加智能和便捷服务的同时，也极大地增加了人们发现有用知识的难度。

云计算环境下，海量的数据存储云中，云的数据存储能力和计算能力以资源服务的方式呈现给用户，而云中的存储和计算资源则是由大量不可信

节点动态组合构建而来，云中海量数据的存储最终分布式地落在各个不可信节点上，这就对部署在不可信节点上的分布式存储模型提出了新的挑战。如何高效地对云存储系统中的海量数据进行存储和管理，最终以资源服务的方式呈现给用户，已经成为云计算时代面临的新考验，也成为当前热点研究领域之一。基于上述分析，本书提出“面向云计算的数据存储关键技术研究”，对云计算环境下数据存储的关键技术问题进行研究。

本书所涉课题以国家自然科学基金项目“网络环境下的服务管理的基础理论研究”（No. 60873192）、“动态网络环境下的服务组合、重建与优化的研究”（No. 61070182）、“面向下一代互联网的网络服务建模基础理论研究”（No. 61170209）和“可编程网络环境下多粒度服务与服务组合的机制研究”（No. 61272508）为依托，是立足于云计算环境下对数据存储关键技术问题的研究。本书在对云计算环境下数据存储关键技术分析的基础上，针对小文件存储机制、副本放置策略、数据去重方法和系统能耗优化等关键技术进行了重点研究，对云计算环境下数据存储的关键问题建立了数学模型，提出了相应的优化策略、机制和算法，设计搭建北京科技大学校园云存储原型系统并对所提出的模型、机制及算法进行验证。本书所涉课题对于进一步开展云计算环境下数据存储的理论研究和深入开发云计算数据存储实用技术具有深远的意义和重要价值。

## 2 绪论

随着整个社会信息化程度的不断提高，特别是在云计算、物联网和移动互联网的大力推动下，IT 行业发展到了一个重要的转折阶段，各行业的应用信息处理和数据服务的规模越来越大，各类数据密集型应用层出不穷，数据总量呈现几何级数增长，数字化信息规模的爆炸式增长推动了人们对信息存储的巨大需求。云中拥有的海量数据在为人们提供更加智能、更加便捷服务的同时，也带来了数据应用的瓶颈和困惑。如何有效地对这些海量数据进行存储、加工和管理，成为云计算时代面临的新挑战，也成为诸多研究者探索云计算新概念、新方法和新技术的动力所在。本书所做的研究即是在这一具有无限前景和深厚背景的热点问题上开展的。

### 2.1 概述

#### 2.1.1 云计算

云计算的服务类型多种多样，简单的云计算在人们日常网络应用中随处可见，如 Google Docs、Microsoft Office 365 和 Zoho 的在线文字处理、电子表格、幻灯片制作等 Office 服务，网易邮箱的附件在线预览服务，腾讯 QQ 空间的在线 Flash 图片集制作服务，联图网的在线二维码生成服务，彩字秀的在线个性文字图片处理服务，等等。除了这些面向普通大众用户的服务外，还有

众多面向专业领域的服务，如百度和新浪均跟随 Google 推出的 AppEngine 应用平台服务、阿里云和盛大云推出的与 Amazon EC2 类似的弹性虚拟机服务等。针对各类层出不穷的云服务，美国国家标准和技术研究院明确了三种云计算服务模式：基础设施即服务（Infrastructure as a Service, IaaS）、平台即服务（Platform as a Service, PaaS）和软件即服务（Software as a Service, SaaS），三者之间的关系如图 2-1 所示<sup>[1]</sup>。

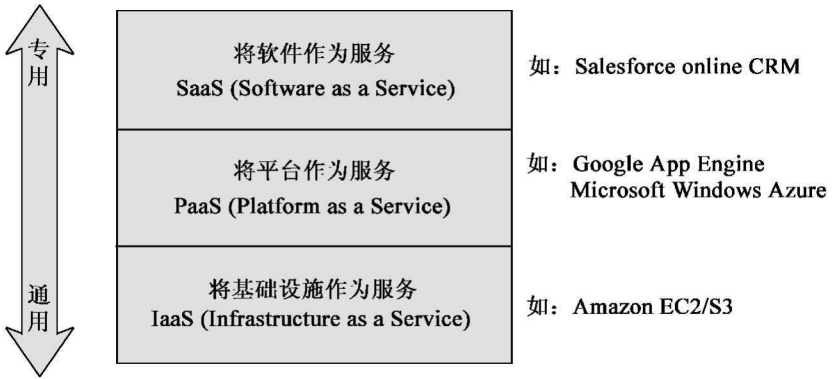


图 2-1 云计算的服务类型

基础设施即服务（IaaS）将处理能力、存储空间、网络组件或中间件等等“基础资源”封装成服务供用户使用，用户能掌控属于自己的操作系统、存储空间、已部署的应用程序以及防火墙、负载均衡器等网络组件，但不能掌控整个云的基础架构。典型的这类服务如 Amazon AWS（Amazon Web Services）云计算产品线中的 EC2 弹性计算云服务和 Rackspace 的云主机服务等。

平台即服务（PaaS）对应用程序的基础资源进行了进一步的抽象，用户能够控制运行应用程序的环境，也即平台，也可能拥有一部分主机的控制权，但不能完全控制操作系统、硬件或网络基础架构。典型的 PaaS 服务有：Vmware 的 Cloud Foundry、百度和新浪的 App Engine、Google App Engine 以及微软的 Microsoft Windows Azure 等。

软件即服务（SaaS）则更加贴近普通用户，从某种意义上来说，也是人

们最早接触和使用的云计算服务类型。在 SaaS 模式下，软件服务供应商将特定的应用软件功能封装成服务，以租赁的形式提供给用户使用，云用户无需在本地安装软件，只需购买该项服务，便可在任何地方、通过任何终端设备使用所需的软件功能，相关数据也存储在云端，用户无需担心更换电脑、重新安装操作系统而导致的软件版本和数据同步等问题，更不用关心云硬件 (IaaS) 和云平台 (PaaS)。典型的 SaaS 服务有 Salesforce 的 CRM 服务、Google 和微软的在线 Office 服务等。也许未来没有“操作系统”这个概念，只有相应的服务接口或 API<sup>[15]</sup>。

从用户角度来看，云计算环境下的服务相比传统模式，具有诸多优势<sup>[12]</sup>：

#### (1) 更充分的资源供给和更强的计算能力

云计算能够通过共享的方式利用闲置资源，使用户获得更好的性价比。同时，用户任务不再局限于一台或几台服务器来完成，而是可以依靠成千上万服务器的超强计算能力实现高性能的服务。

#### (2) 灵活可扩展的资源分配

可根据云应用的具体需求，利用虚拟化技术及有效的资源调度机制，弹性分配资源，通过“云”规模的动态伸缩，适应用户和应用规模的不断增长。

#### (3) 更便捷的服务方式

由于计算主要在云端完成，云计算服务对接入设备的要求很低，用户几乎可以在任何时刻、任何地点，通过多种方式，使用各种设备获取应用服务，而且使用这种服务模式，还能够轻松实现用户间的实时协作。

#### (4) 高可靠的数据存储

云计算提供了近乎无限的存储空间，“云”中数据的存储和处理都是分布式完成的，在部分节点失效的情况下，仍能保证数据的可靠性和可用性。

云计算的优势已被越来越多的研究机构、工业界和用户所认可，市场需求愈发旺盛，越来越多的传统 IT 企业巨头投入到云计算领域中，推进了云计

算的快速发展，目前已涌现出大量的商业云计算应用平台。Google、IBM 和亚马逊是云计算的先行者，微软、Yahoo、VMware、Salesforce、Facebook、SAP、EMC、Oracle、百度、阿里、腾讯、中国移动和华为等互联网领域和企业应用领域的众多巨头公司也都在云计算领域投入了大量的精力，图 2-2 中总结了云计算发展的主要历程。

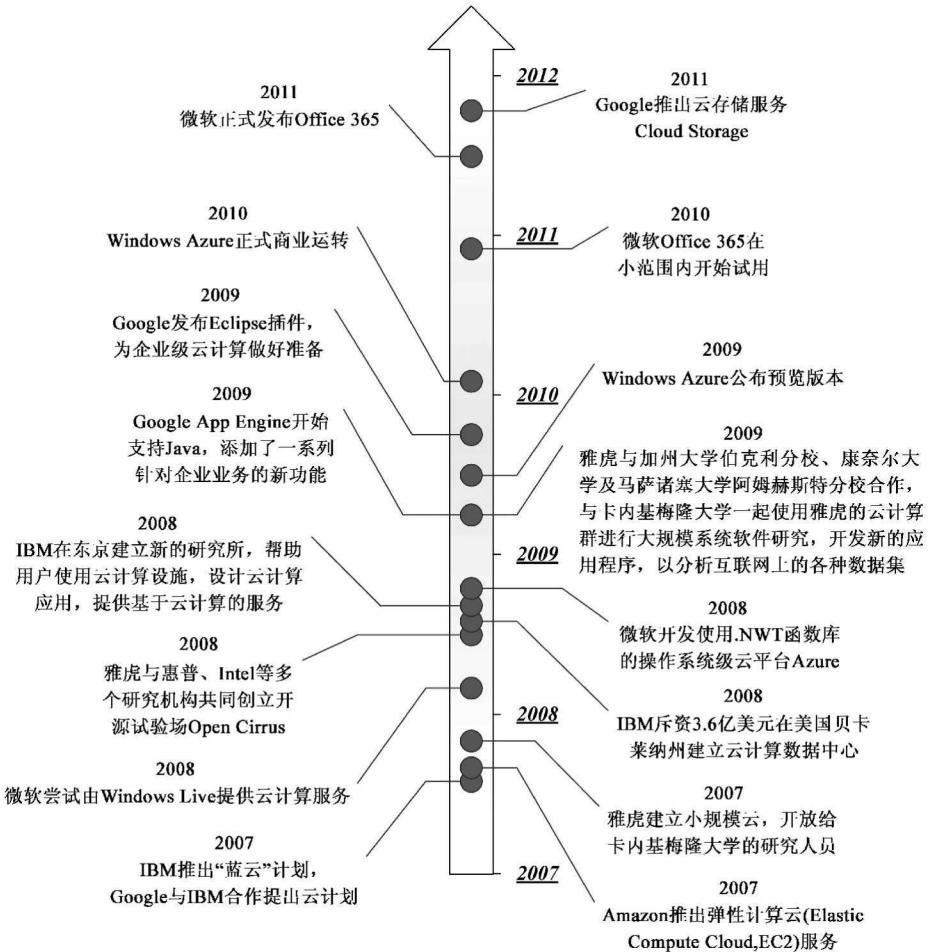


图 2-2 云计算发展



Google 是云计算技术的领导者，也是应用云计算技术最多的使用者。Google 新一代的搜索引擎就建立在云计算基础设施之上，这些节点分布在全球 30 多个地区，一共有 500 万台左右的服务器，而且数量仍在不断地增加。除了搜索服务外，Google 众多的应用平台和服务，包括 Google Earth、Google App Engine、Google Maps、Google Docs、Gmail、Google Plus 等都使用了这些云计算基础设施。Google 也提供了另外两种类型的云服务，Google Docs 就是非常成功的 SaaS 服务之一。用户使用 Google Docs 时，无需下载安装客户端软件，直接在线进行文档编辑、阅读和共享，文档保存在 Google 云中，之后也可在任何地方、任何一个能访问互联网的终端访问这些文档，十分便利。Google 的 PaaS 服务 Google App Engine 也成为后来者争相模仿的对象，它允许第三方开发者在 Google 的云计算环境中开发、部署和运行各类应用程序，而不用关心负载均衡、网络带宽等问题。目前已经有超过 200 万用户在 Google App Engine 部署了各自的应用服务。2003 年，Google 对云计算领域做出了一个重大贡献，它以学术论文的形式公开了其云计算三大核心技术：Google 文件系统 GFS、半结构化数据库 Bigtable 和并行计算框架 Map/Reduce，受到这三大技术的启发，模仿者不断，Hadoop 即是其中最受关注的、也最成功的开源项目，在 Yahoo 等公司的领导和积极推动下，目前已有大量基于 Hadoop 的商用解决方案出现。

Amazon 在 2006 年推出了面向公众的弹性计算云服务 EC2 (Elastic Computing Cloud) 和简单存储服务 S3 (Simple Storage Service)，也是最早为企业提供 IaaS 服务的供应商。服务一经推出，就受到了大量企业用户的肯定，短短两年时间内，Amazon AWS 服务的注册用户就多达 44 万人，目前 Amazon 的云服务类型越来越丰富，囊括了从计算、网络、存储、分发、数据库和应用平台等各类 IaaS 和 PaaS 服务，是公有云领域的领头羊。

2007 年 11 月 IBM 推出“蓝云”私有云计算平台解决方案，包括一系列自我管理和自我修复的虚拟化云计算软件，这标志了云计算技术正式进入企