

# 云计算及其关键技术研究\*

高林<sup>1,2</sup>, 宋相倩<sup>1</sup>, 王洁萍<sup>2</sup>

- (1. 桂林电子科技大学 计算机与控制学院, 广西 桂林 541004;  
2. 中国电子技术标准化研究所 信息技术研究中心, 北京 100007)

**摘要:** 在分析现有定义的基础上对云计算提出了较全面的参考性定义, 并给出了云计算的六大特征、四种部署模式和三种服务模式。此外, 通过对云环境中三种角色的责任分析和对主流云计算产品的研究, 抽象出了支撑云计算的关键技术, 最后指出云计算对未来软件市场发展的影响。

**关键词:** 云计算; 云服务; 云部署;

中图分类号: TP

文献标识码: A

文章编号: 1674-7720(2011)10-0005-03

## Research on cloud computing and key technologies

Gao Lin<sup>1,2</sup>, Song Xiangqian<sup>1</sup>, Wang Jieping<sup>2</sup>

- (1. School of Computer and Control, Guilin University of Electronic Technology, Guilin 541004, China;  
2. Conformance Test Center for Information Technology Standards, China Electronics Standardization Institute, Beijing 100007, China)

**Abstract:** Analyzing the existing definitions, this paper proposes a comprehensive definition, and discusses the six essential characteristics, four deployment models and three service models. What's more, researching on the mainstream products and analyzing the responsibility of three parts in cloud computing, we has abstracted the key technologies of cloud computing involve virtualization, distributional data storages, huge data processing, security etc. At last, influences of cloud computing in the further IT industry are illustrated.

**Key words:** cloud computing; cloud service; cloud deployment

作为一种 IT 基础设施交付和使用模式、一种信息服务交付和使用模式、一种基于互联网通过虚拟化方式共享信息资源的新型计算模式, 云计算为数据计算、存储和管理提供了虚拟资源空间和超强计算能力, 可使用户低成本、高效率、灵活地分享信息技术的发展成果, 更好地获取和使用知识, 减少数字鸿沟, 加强科技创新, 提高公共服务水平。

20 世纪 60 年代 John McCarthy 提出“计算能力在未来将成为公共设施”, 这是云计算概念的起源<sup>[1]</sup>。20 世纪 80 年代末诞生了网格计算, 将大量的异构计算机集成起来进行复杂问题(通常是科学问题)的求解。20 世纪 90 年代, 虚拟化的概念已经从虚拟服务器发展到更高层次的抽象, 从虚拟平台到虚拟应用程序。20 世纪末 21 世纪初, Web Service 和 SOA 等面向服务的概念和技术迅速发展。云计算正是在上述各项技术的基础上, 经过数十载的发展演化逐渐成熟起来。

### 1 云计算概念

美国加州大学伯克利分校在指出云计算既指在互

联网上以服务形式提供的应用, 也指在数据中心中提供这些服务的硬件和软件, 而这些数据中心中的硬件和软件则被称为云<sup>[2]</sup>。美国国家标准与技术研究院指出云计算由一个可配置的共享资源池组成, 该资源池提供网络、服务器、存储、应用和服务等多种硬件和软件资源<sup>[3]</sup>。资源池具备自我管理能力和用户只需少量参与就可以方便快捷地按需获取资源。云计算提高了资源可用性。

Wikipedia 和 NIST 都强调云计算中资源的按需提供。在对云计算范围的界定上, UCB 认为云计算不包括私有云, 而 NIST 则认为云计算包括私有云。本文认为 NIST 的定义主要从云计算提供者的角度出发, 比较全面地涵盖了云计算的本质特征, 但是缺乏对云计算使用者和云计算开发者的阐述。

综上所述, 本文提出自己的参考性定义: 云计算是一种以提高资源利用率、降低 IT 成本为驱动的计算模式, 在云计算中包括云服务使用者、云服务提供者和云服务开发者三类角色。

云服务使用者可在不具备专业知识的情况下利用网络通过交互技术以自服务的方式访问云端资源。

\* 基金项目: 国家科技重大专项资助项目(2009ZX01045-005-002)

云服务提供者作为联系服务使用者和服务开发者的桥梁,以按需使用、按量计费的方式通过网络提供动态可伸缩资源,云服务提供者负责安全管理、运营支撑管理、服务平台以及资源平台,其中服务平台包括软件即服务 SaaS (Software as a Service)、平台即服务 PaaS (Platform as a Service)、基础设施即服务 IaaS (Infrastructure as a Service) 以及这三种服务的组合和管理。资源平台包括资源池、虚拟资源管理、物理资源管理等。

云服务开发者负责将各种软硬件资源封装成服务,并负责服务的创建、发布和维护。

## 2 云计算特征

为了全面理解云计算,本文进一步分析云计算的关键特征,具体如下:

(1)网络接入。云计算通过网络提供服务能力,支持各种标准设备接口。

(2)按需提供。云计算可以根据用户需求的改变,快速动态地分配和回收资源,既保证忙时用户需求的及时满足,又保证闲时资源的及时回收。

(3)按量计费。云计算提供对计算、存储、网络和其他各类资源的测量能力,能够统计某个用户在某段时间对某类资源或者服务的使用情况。

(4)用户自服务。用户可以自服务的方式配置和调用云服务。例如,配置服务器时间和网络存储容量。

(5)多租户。云计算提供者可以同时为多个使用者提供服务,利用规模经济提高资源利用率。

(6)可靠性。云计算通过异地备份、多副本备份等策略增强服务可靠性。

这些关键特征一方面可以帮助理解云计算的内涵和外延,另一方面可以作为区别云计算和其他各类计算模式的依据。

## 3 云计算的服务模式

云服务<sup>[3]</sup>主要包括 SaaS、PaaS、IaaS 以及三种服务的组合。如图 1 所示。

(1) SaaS 是指将某些特定应用软件功能封装成服务,对于服务使用者而言购买的是软件的使用权而不是其所有权。如 Salesforce 公司提供的在线客户关系管理 CRM 服务。

(2) PaaS 是指将一个完整的应用开发平台,包括应用设计、应用开发、应用测试和应用托管都作为一种服务提供给客户。典型的 PaaS 包括 Google App Engine<sup>[4]</sup>、Microsoft Windows Azure。

(3) IaaS 将硬件设备等基础资源,包括计算、存储和网络等,封装成服务供用户使用,典型的如亚马逊的弹性计算云 EC2 和简单存储服务 S3<sup>[6]</sup>。

## 4 云计算的部署模式

云计算有四类典型的部署模式(如图 2 所示):“公共云”、“私有云”、“社区云”和“混合云”。具体描述如下:

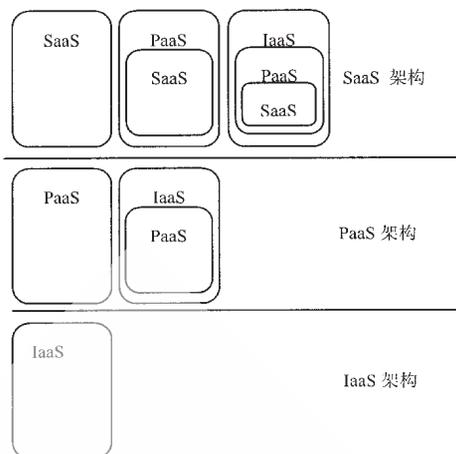


图 1 云计算的服务模式



图 2 云计算的部署模式

(1)公共云。云基础设施对公众或某个很大的业界群组提供云服务。

(2)私有云。云基础设施特定为某个组织运行服务,可以是该组织或某个第三方负责管理。

(3)社区云。云基础设施由若干个组织分享,以支持某个特定的社区。

(4)混合云。云基础设施由两个或多个云组成,独立存在,但是通过标准的或私有的技术绑定在一起,这些技术可促成数据和应用的移植性。

## 5 云计算关键技术

本文通过对云计算三种角色的责任研究和典型云计算产品的分析,抽象出了支撑云计算的七大关键技术。

### 5.1 系统虚拟化

系统虚拟化<sup>[5]</sup>是指将一台物理计算机系统虚拟化为一台或多台虚拟计算机系统。通过虚拟化层的模拟,虚拟机中的操作系统认为自己仍然是独占一个系统在运行。这个虚拟化层被称为虚拟机监控器。

### 5.2 虚拟化资源管理

虚拟化资源是云计算中最重要的组成部分之一,它对虚拟化资源的管理水平直接影响着云计算的可用性、可靠性和安全性。云资源池中应用的需求不断改变,在线服务的请求经常不可预测,这种动态的环境要求云计算的数据中心或计算中心能够对各类资源灵活、快速、动态地按需进行调度。云计算中的虚拟化资源与以往的网络资源相比,具有数量更为巨大、数据分布更为离散、数据调度更为频繁的特征<sup>[6]</sup>。

### 5.3 分布式数据存储

分布式数据存储技术包含分布式文件存储系统、分

## 软件天地 Software Technology

布式对象存储系统和分布式数据库技术。下面分别阐述这三方面的技术:

### (1) 分布式文件存储系统

为了存储和管理云计算中的海量数据,Google 提出分布式文件系统 GFS<sup>[7]</sup>(Google File System)。Google GFS 的结点由廉价不可靠 PC 构建,因而硬件失败是一种常态而非特例。

### (2) 分布式对象存储系统

对象存储系统是传统的块设备的延伸,具有更高的“智能”:上层通过对象 ID 来访问对象,而不需要了解对象的具体空间分布情况。Amazon 的 S3 就属于对象存储服务。

### (3) 分布式数据库技术

云计算环境下,大部分应用不需要支持完整的 SQL 语义,而只需要 Key-Value 形式或略复杂的查询语义。Google 的 BigTable<sup>[8]</sup>是一个典型的分布式结构化数据存储系统。

## 5.4 海量数据处理

并行计算模型是提高海量数据处理效率的常用方法。云计算环境下的并行计算模型属于面向互联网数据密集型应用的并行编程模型,云计算下把海量数据分布到多个结点(通常是廉价不可靠的 PC 机)上,将计算并行化,利用多机的计算资源,加快数据处理的速度<sup>[9]</sup>。如 Google 的 MapReduce 模型、微软的 Dryad 模型。

## 5.5 用户交互技术

随着云计算的逐步普及,浏览器已经不仅仅是一个客户端的软件,而逐步演变为承载着互联网的平台。浏览器与云计算的整合技术主要体现在两个方面:浏览器网络化与浏览器云服务。浏览器都将网络化作为其功能的标配之一,主要功能体现在用户可以登录浏览器,并通过自己的帐号将个性化数据同步到服务端。而目前的云端服务,主要体现在 P2P 下载、视频加速等单独的客户端软件中。

## 5.6 安全管理

安全问题是用户是否选择云计算的主要顾虑之一。传统集中式管理方式下也有安全问题,云计算的多租户、分布性、对网络和服务提供者的依赖性,为安全问题带来新的挑战<sup>[10]</sup>。主要的数据安全问题和风险包括:数据存储及访问控制、数据传输保护、数据隐私及敏感信息保护、数据可用性、依从性管理;相应的数据安全管理工作包括:数据保护及隐私、身份及访问管理、可用性管理、日志管理、审计管理、依从性管理等。

## 5.7 运营支撑管理

为了支持规模巨大的云计算环境,需要成千上万台服务器来支撑。如何对数以万计的服务器进行稳定高效的运营管理,成为云服务被用户认可的关键因素之一。

下面从云的部署、负载管理和监控、计量计费、SLA 服务水平协议(Service Level Agreement)、能效评测五个方

面分别阐述云的运营管理。

### (1) 云计算的部署

云计算的部署是一个系统工程,涉及到从机房建设、网络优化、硬件选型、软件系统开发和测试到运维等各方面的技术。为了保证服务的健壮性,需要将云计算以一定冗余部署在不同地域的若干机房。为了应对规模的不断增长,云计算要具备便利的、近乎无限的扩展能力,因而从数据存储层、应用业务层到接入层都需要采用相应的措施。为了保护云计算及其应用的安全,需要建立起各个层次的信息安全机制。除此之外,还需要部署一些辅助的子系统,如管理信息系统、安全系统、监控和计费系统等,使云计算的部署和运营管理达到高度自动化和智能化的程度。

### (2) 负载管理和监控

云计算的负载管理和监控是一种大规模集群的负载管理和监控技术<sup>[11]</sup>。对于结点粒度,它需要能够实时地监控集群中每个结点的负载状态,报告负载的异常和结点故障,对出现过载或故障的结点采取既定的预案。对于集群整体粒度,通过对单个结点、单个子系统的信息进行汇总和计算,近乎实时地得到集群的整体负载和监控信息,为运维、调度和成本提供决策。

### (3) 计量计费

云计算的主要商业运营模式是采取按量计费 (pay-as-go) 的收费方式<sup>[12]</sup>。为了精确地度量“用了多少”,就需要准确、及时计算云计算上的每一个应用服务使用了多少资源,这称为服务计量。在计量的基础上,选取若干合适的维度组合,制定相应的计费策略并进行计费。

### (4) SLA

SLA 是在一定开销下为保障服务的性能和可靠性,服务提供商与用户间定义的一种双方认可的协定。对于云服务而言,SLA 是必不可少的,因为用户对云服务的性能和可靠性有不同的要求。从用户的角度而言,也需要从云服务提供商处得到具有法律效力的承诺,来保证支付费用之后得到应有的服务质量。

### (5) 能效评测

云计算提出的初衷是将资源和数据尽可能放在云端,降低能源消耗。但是在实际应用中,为解决大型数据中心的散热问题造成了大量的能源消耗。如何有效降低能源消耗构建绿色数据中心成为云服务提供商迫切需要解决的问题之一。显性的能耗测试评价可以参照传统数据中心的评价体系,具体包括:能源效率指标、IT 设备的能效比、IT 设备的工作温度和湿度范围、机房基础设施的利用率指标。隐性能耗测试评价包括云计算服务模式节省了多少社会资源,由于客户需求的不同,云计算系统吞吐量的变化节省了 IT 设备的投资和资源的重复建设。

亚马逊、谷歌、雅虎、IBM 和微软等跨国公司纷纷开展云计算研究,并在相关领域开展部署,部分应用已取

得成功。我国华为、中兴、浪潮等网络设备提供商也积极投入相关设备研发。在电信领域、基础运营商和增值服务商也相继开展了相关的尝试，启动了相关研究工作，这些都为我国云计算产业的创新发展提供了众多有利的探索经验。伴随着云计算产业的发展，云计算将为信息产业带来广阔的市场空间，催生大量的市场发展新机遇，对IT产业链更是影响重大。

#### 参考文献

- [1] Wikipedia. Cloud Computing [EB/OL]. [2010-07-16]. <http://en.wikipedia.org/wiki/Cloudcomputing>
- [2] ARMBRUST M, FOX A. Above the clouds: a berkeley view of cloud computing[R/OL]. (2009-02-10)[2009-05-15]. <http://www.grid.pku.edu.cn/cloud/Berkeley-abovethe-clouds.pdf>
- [3] MELL P, GRANCE T. The NIST definition of cloud computing. [EB/OL]. (2009-10-7)[2010-4-9]. [Http://csrc.nist.gov/groups/SNS/cloud-computing/index.html](http://csrc.nist.gov/groups/SNS/cloud-computing/index.html).
- [4] GOOGLE. 开发人员指南. [EB/OL]. [2010-05-15] <http://code.google.com/intl/zh-CN/appengine/docs/>
- [5] 金海. 计算系统虚拟化: 原理与应用[M]. 北京: 清华大学出版社, 2008.
- [6] PATERSON M. Dynamic resource distribution across clouds.V00214440[R]. Victoria, University of Victoria Faculty of Engineering Winter 2010 Work Term Report. 2010.
- [7] GHEMAWAT S, GOBIOFF H, Leung Shuntak. The google file system. In: Proc of the 19th ACM SOSP, New York, 2003.
- [8] CHANG F, DEAN J, GHEMAWAT S, et al. Bigtable: a distributed storage system for structured data. In: Proc of the 7th USENIX Symp on OSDI, Berkeley, 2006.
- [9] LAMMEL R. Google's Map Reduce programming model Revisited. [EB/OL]. [Http://www.cs.vu.nl/~ralf/MapReduce/papa.pdf](http://www.cs.vu.nl/~ralf/MapReduce/papa.pdf), 2007.
- [10] Cloud Security Alliance. Top threats cloud computing V1.0.[EB/OL]. <http://www.cloudsecurityalliance.org/topthreats/csathreats.v1.0.pdf>
- [11] RANGLES M, LAMB D, et al. A comparative study into distributed load balancing algorithms for cloud computing. [C].//Proc of IEEE24th International Conference on Advanced Information Networking and Applications Workshops. 2010. (收稿日期: 2011-02-16)

#### 作者简介:

高林,男,1973年生,博士,高级工程师,主要研究方向:智能计算,云计算。

宋相倩,女,1986年生,硕士,主要研究方向:云计算。

王洁萍,女,1976年生,博士,工程师,主要研究方向:数据库,云计算。