面向过程感知的泛在网络服务资源调度

摘要

关键词

Abstract

Keywords:

1. 绪论

# 研究背景及意义

早在1983年，太阳电脑（Sun Microsystems）就已经提出“网络就是计算机” （“The Network is the Computer”）。经过这么多年，网络技术得到了飞速发展，使得人们接受网络中的服务变得越来越方便。在当今泛在网络的环境中，各种设备都可以无缝的接入网络，享受网络带来的方便。而对服务和资源的需求，引发了许多计算技术的诞生。

网格计算诞生于上个世纪90年代，其目标是将分布在世界各地的异构的资源，用网络连接在一起，形成一个大的资源池。用这些网络中自治的计算资源进行计算，实现网络资源的共享，高效的协同计算。

经过了十几年的发展，网格计算取得了一定的发展，但是，仍然是政府和科研机构为投资主体在使用，没有得到商业的认可。

2006年8月9日，Google首席执行官埃里克·施密特（Eric Schmidt）在搜索引擎大会（SES San Jose 2006）首次提出“云计算”（Cloud Computing）的概念。后来，“云计算”技术得到了业界的一致认可。许多科研机构和商业公司都迅速的做出了反应。Google，IBM，MicroSoft，亚马逊等等分别推出了云计算产品。

然而，云计算一直没有一个统一的定义。美国国家标准与技术研究院（NIST）给出的定义：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，或与服务供应商进行很少的交互。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。

在2011年国际云计算研讨会上，中国工程院院士张尧学认为：云计算的目标是“通过对资源的集中使用和共享，为用户提供不知不觉、无处不在的个性化服务”。终极目标就是说，不管用户使用PC机也好，iPhone也好，都能跨平台地获得“不知不觉”、“无所不在”个性化的服务。实现这一目标的方式，就是通过网络、通过服务器把计算能力集中起来使用和共享。

显然，云计算模式在泛在网络时代里能够很好的提供商业服务，利用这些数据中心的服务器给各种终端提供优质的服务。所以，现在已经有许多公司开发出了基于云计算的泛在网络应用。

目前，对云计算的研究主要集中在以下几个方面：云计算中心所提供的软硬件服务和服务方式，数据中心的资源管理，大数据挖掘等方面。

当前云计算数据中心在资源管理方面存在以下缺陷：

（1）硬件基础设施廉价

数据中心的硬件基础设施大多是由大规模的廉价服务器集群构成，特别是x86架构的服务器使用较多。例如Amazon公司的Amazon EC2就是由上百万台廉价的服务器通过网络连接而成为用户提供计算和存储服务。并且连接这些服务器的网络一般是普遍的千兆以太网。

（2）系统负载不均衡

数据中心负载不均衡体现在以下几点：

第一，同一个节点内不同类型的硬件资源使用不均衡。例如，CPU利用率很高，但网络带宽利用率很低。这是因为云计算服务器上运行着各种应用，这些应用对硬件资源的需求不同，例如计算应用需要CPU，内存资源多一些，而Web应用不需要太多的CPU，内存等资源，而只需要更多的带宽。因此应该按照应用的需求来分配服务器的硬件资源。

第二，同一应用不同节点之间负载不均衡。例如Map/Reduce架构中使用了Master节点和大量的Worker节点。Master节点首先把用户提交上来的作业划分成若个干任务，然后把这些任务分配给Worker节点来完成。因此在执行作业前，Master节点需要根据任务的请求压力分配Worker节点。

第三，时间不均衡。用户对服务的使用具有一定的时间规律性。例如，对于网络游戏来说，晚上的负载比白天大，周末的负载比平常工作日大。因此数据中心的节点压力随着时间段发生变化。

（3）节点易失效

数据中心由大规模的集群构成，例如Google公司的搜索引擎架构在100万台服务器之上，这些服务器分布在全球200多个地点。存在如此规模巨大且廉价的节点，节点的失效将不可避免。根据调查报告，数据中心单个节点可以无故障工作1万多个小时，因此在一个拥有10万台服务器的数据中心中，平均每6分钟就有一台服务器发生故障。

为了克服云计算数据中心在资源管理方面存在的缺陷，保证云计算为用户提供更好的服务QoS。本文对云计算环境中的虚拟机资源调度进行研究，并提出了适用于云计算环境的基于蚁群优化算法的虚拟机资源调度算法（其中包括基于改进蚂蚁系统的虚拟机资源调度算法和基于改进蚁群算法的虚拟机资源调度算法）和节能及信任驱动的虚拟机资源调度算法。基于蚁群优化算法的虚拟机资源调度算法与节能及信任驱动的虚拟机资源调度算法在用户QoS方面略有不同。基于蚁群优化算法的虚拟机资源调度算法主要关注用户任务的分配时间和执行时间。而信任驱动的虚拟机资源调度算法则侧重于关注用户的信任需求以及对数据中心的节能要求。这两个调度算法在一定程度上满足了用户QoS对云计算系统的不同要求，因此本课题的研究对云计算的实际应用具有重要意义。

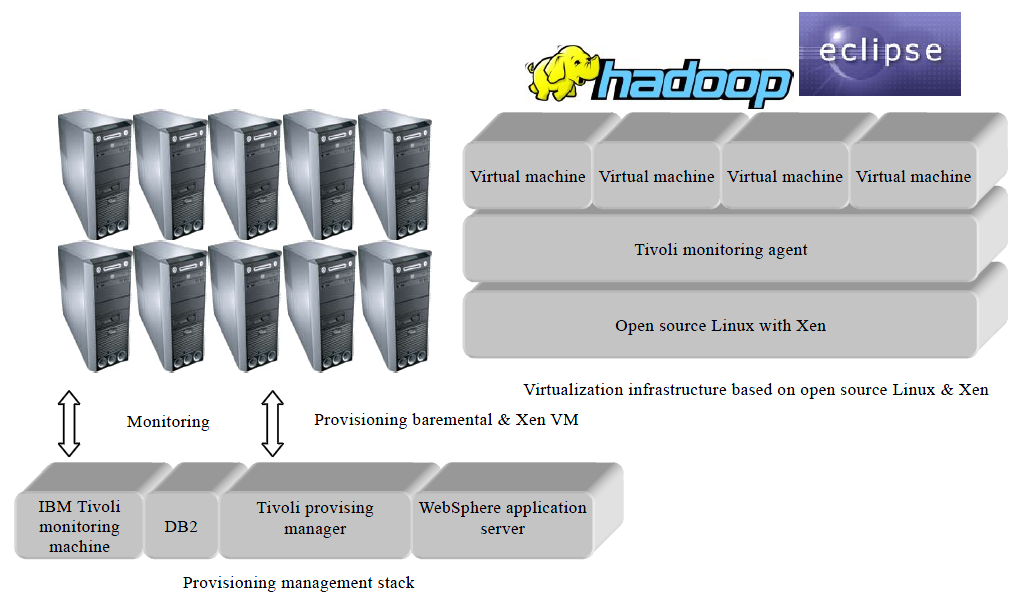
# 研究现状

自从云计算概念被提出来后，云计算技术就受到了各大 IT 厂商和学者专家们的广泛关注。虽然，对于云计算技术的研究才过去短短几年，但是现在已经涌现出许多优秀的云计算系统，其中比较典型的云计算系统有：IBM“蓝天”云计算平台、Amazon 弹性计算云、Windows Azure、GoogleApp Engine 平台以及清华大学透明计算平台。

## 国外研究现状

1. IBM“蓝天”云计算平台

IBM 的“蓝云”计算平台是一套软、硬件平台，将Internet 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。



图为IBM“蓝天”云计算平台架构

可以看出，IBM 的“蓝云”计算平台由一下几个部分组成：基于 X86 架构的服务器集群、IBM Tivoli 监控软件（IBM Tivoli monitoring，ITM）、IBM 部署管理软件（Tivoli

provisioning manager，TPM）、IBM DB2 数据库、IBM WebSphere 应用服务器以及一些信

息处理软件和虚拟化软件。云计算的相关技术，例如虚拟化、软件即服务、SOA、RDP（按

需部署资源）等都是 IBM 的优势所在。IBM 通过三种方式帮助合作伙伴及其用户实现云

计算环境：IBM 帮助合作伙伴及客户创建供其使用的内、外部云计算环境；IBM 直接为合

作伙伴及客户提供云服务；IBM 正在创建一个由合作伙伴、开发人员、大学和厂商组成的

云计算生态系统，大力推动云计算标准和技术的制定。

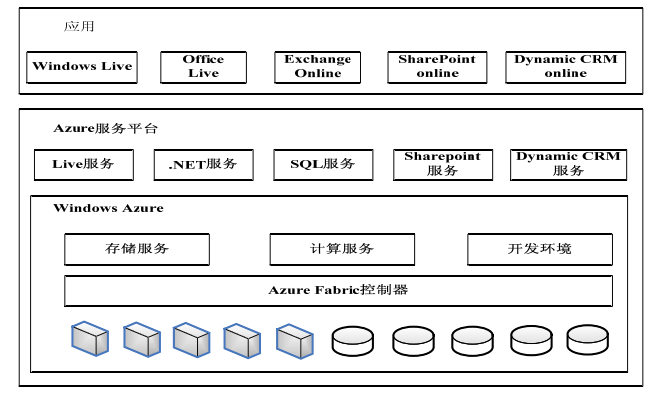
1. Amazon 弹性计算云

Amazon EC2全称为 Amazon Elastic Cloud Computing，是亚马逊公司为用户提供的一项 Web 服务，其名称为 Amazon 弹性计算云。该 Web 服务实际上提供的是 Amazon 公司数据中心中的 Linux 虚拟机资源，这些虚拟机资源又被称为实例。根据其大小不同，可以分为三种：小型实例、大型实例和极大型实例。用户可以使用应用程序编程接口或者 Web 工具运行基于 Linux 的应用程序。Amazon EC2 可以根据实际情况自动调整其容量。例如当负载达到高峰时，它可以自动添加更多的虚拟机以保障服务性能。2006 年 Amazon 公司发布了简单存储服务 S3（simple queue service，简称 SQS），用户可以把创建好的 AMI 镜像上传到 SQS 上，然后利用 Amazon 公司提供的 Web 服务接口启动、监控和停止 AMI 镜像。Amazon 为用户指定了虚拟机实例的收费标准，用户可以根据实际使用情况向 Amazon 公司付费。

1. Windows Azure

The Azure Services Platform (Azure)是一个互联网级的运行于微软数据中心系统上的云计算服务平台，它提供操作系统Windows Azure和可以单独或者一起使用的开发者服务。Azure是一种灵活和支持互操作的平台，它可以被用来创建云中运行的应用或者通过基于云的特性来加强现有应用。它开放式的架构给开发者提供了Web应用、互联设备的应用、个人电脑、服务器、或者提供最优在线复杂解决方案的选择。

Windows Azure以云技术为核心，提供了软件+服务的计算方法。 它是Azure服务平台的基础。Azure用于帮助开发者开发可以跨越云端和专业数据中心的下一代应用程序，在PC、Web和手机等各种终端间创造完美的用户体验。Azure能够将处于云端的开发者个人能力，同微软全球数据中心网络托管的服务，比如存储、计算和网络基础设施服务，紧密结合起来。这样，开发者就可以在“云端”和“客户端”同时部署应用，使得企业与用户都能共享资源。



图为Windows Azure架构图

1. Google App Engine

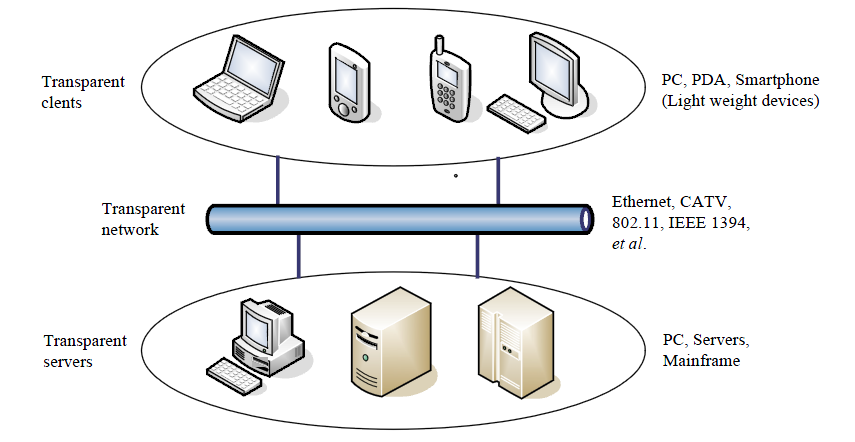
谷歌 App Engine是利用谷歌公司的数据中心部署和托管Web服务的平台即服务（platform as a service，简称 PaaS）的云计算平台。它跨越多个服务器来虚拟化 Web 服务。Google App Engine 向用户提供一定的资源时是免费的，只有当用户使用更多的 CPU、存储空间、带宽等资源时 Google App Engine 才会收取一定的费用。Google App Engine 由 GAEWeb 服务基础设施、应用开发套件（SDK）、分布式存储服务（Datastore）、应用程序运行时环境（Application Runtime Environment）和管理控制台（Admin Console）五部分组成。GAE Web 服务基础设施提供了可伸缩的服务接口，保证了 GAE 对网络和存储等资源的灵活使用和管理。用户可以将自己编写的应用程序在 GoogleApp Engine 上面运行，由 GoogleApp Engine 负责资源的维护。当前用户只能使用 Python 一种语言编写应用程序，以后可能会实现使用其他语言编写其应用程序。

GAE 的目标不同于 Amazon 弹性云，Amazon 弹性云的目标是提供一个可伸缩的、分布式的、高可靠的虚拟机环境，而GAE 则更专注于提供一个开发简单、部署方便的 Web 应用运行和管理平台。GAE 使应用开发者只需要注意核心业务逻辑的实现，而不需要关心物理资源的分配、应用请求的路由、负载均衡、资源及应用的监控和伸缩等任务。

## 国内研究现状

1. 透明计算平台

张尧学教授领导的研究小组从1998 年开始就从事透明计算系统和理论的研究，到2004 年前后正式提出，并不断完善了透明计算的概念和相关理论。随着硬件、软件以及网络技术的发展，计算模式从大型机的方式逐渐过渡到微型个人计算机的方式，并且近年来过渡到普适计算上。但是用户仍然很难获得异构类型的操作系统以及应用程序，在轻量级的设备上很难获得完善的服务。而在透明计算中，用户无须感知计算具体所在位置以及操作系统、中间件、应用等技术细节，只需要根据自己的需求，通过连通在网络之上的各种设备选取相应的服务。下图显示了透明计算平台的3 个重要组成部分。用户的显示界面是前端的轻权设备，包括各种个人计算机、笔记本、PDA、智能手机等，被统称为透明客户端。透明客户端可以是没有安装任何软件的裸机，也可以是装有部分核心软件平台的轻巧性终端。中间的透明网络则整合了各种有线和无线网络传输设施，主要用来在各种透明客户端与后台服务器之间完成数据的传递，而用户无须意识到网络的存在。与云计算基础服务设施构想一致，透明服务器不排斥任何一种可能的服务提供方式，既可通过当前流行的PC 服务器集群方式来构建透明服务器集群，也可使用大型服务器等。



图为 透明计算的体系结构

当前透明计算平台已经达到了平台异构的目的，能够支持Linux 以及Windows 操作系统的运行。用户具有很大的灵活性，能够自主选择自己所需要的操作系统运行在透明客户端上。透明服务器使用了流行的PC服务器集群的方式，预先存储了各种不同的操作平台，包括操作系统的运行环境、应用程序以及相应的数据。每个客户端从透明服务器上获取并建立整个运行环境以满足用户对于不同操作环境的需求。由于用户之间的数据相互隔离，因此服务器集群可以选取用户相对独立的方式进行存储，使得整个系统能够扩展到很大的规模。在服务器集群之上进行相应的冗余出错处理，很好地保护了每个用户的透明计算数据安全性。

1. 其它

云计算在企业层面的运用也越来越广泛，它可以让中小企业不必费大力气购买服务器、雇佣运维人员，通过网络租赁云服务就可以达成目标。使用云服务按照实际使用量付费，用多少就付多少钱，硬件成本和人力支出都大大降低。

为了抢占云计算产业发展先机，深圳市政府从云计算产业分工、资源的合理利用等方面都做了明确规划。目前，深圳已经建立起云计算相关的国家级、省级和市级工程实验室8个，重点实验室11个，公共技术服务平台6个。2009年启动千万亿次国家超级计算深圳中心的建设，2010年依托超算中心成立了深圳云计算中心。

# 本文的工作与组织结构

1. 相关技术介绍

大概思路：云概念介绍，SAAS,PAAS,IAAS介绍，Cloud Foundry介绍。

# 2.1 云计算的起源

当网络技术诞生后，人们对网络的需求呈爆炸式的增长，从而出现了许多网络应用。各种商业网站如雨后春笋般的出现了，人们方便的浏览Internet，在网络中分享图片，日志。这些服务基本都是运行在供应商的服务器上。随着各种智能终端的普及，给网络带来了更多的负载。然而，社会的信息量也快速的增长，如何处理这些大量的数据和提供优质的服务成为服务供应商需要面对的一道难题。

Internet技术的出现，为网络应用提供了一种通用的平台，产生了http,XML等开放式的技术标准，为各种web 服务提供了一个共同的机制。在这些机制下，网络服务可以根据根据业务流程进行动态的调整，动态的分析客户的需求。

为了提供web服务，服务供应商们购置了大量的服务器，组成服务器集群。然而，用户需要处理的信息量远远超过了这些集群的处理能力，这时候，网格技术应运而生。网格技术能够把分散在世界各地的计算资源集中起来。网格技术能够分配分布式的资源和透明地访问它们。网格技术的一个关键点就是实现了基于标准web service的协议，使得分布式的资源能够被发现，分配，监视，计费，并且能够作为一个虚拟的子系统进行管理。由于，这些分布式的计算资源数量巨大，依靠人工来管理是不现实的，所以，人们通过自动化的管理方式来管理这些集群和分布式的计算资源。

网格计算的发展，为云计算奠定了扎实的基础。云计算是基于数据中心的一种运算模式，其提供资源池，利用虚拟化技术，动态的分配资源给客户，并且是具有高可靠性的计算模式。它的商业模式中，以用户为中心，提供高质量的，安全的，便捷的网络服务，使得任何一个用户都能够将网络作为其计算中心和数据中心，从而使得用户能够随时随地的在网络中进行各种活动。

那么，什么是云计算呢？目前，业界还没有达成一致的意见，每个人都有自己的想法。但是，其运用到的技术主要有：网格计算（Grid computing）、数据中心自动管理、虚拟化技术、Web2.0等。其技术结构图如下。

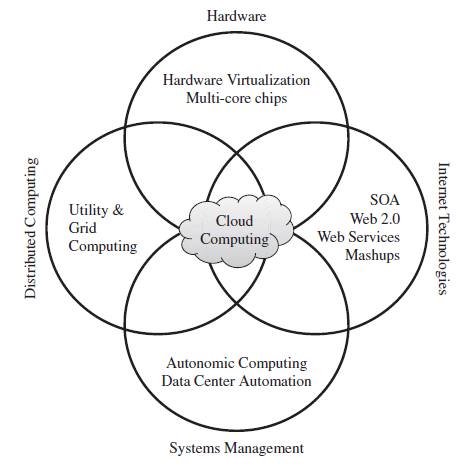


图 云计算的基础技术图

# 2.2 云计算特点

1、超大规模。“云”具有相当的规模，Google云计算已经拥有100多万台服务器，亚马逊、IBM、微软和Yahoo等公司的“云”均拥有几十万台服务器。“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”中某处运行，但实际上用户无需了解应用运行的具体位置，只需要一台笔记本或一个PDA,就可以通过网络服务来获取各种能力超强的服务。

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更加可靠。

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于变万化的应用，同一片“云”可以同时支撑不同的应用运行。

5、高可伸缩性。“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，用户按需购买，像自来水、电和煤气那样计费。

7、极其廉价。“云”的特殊容错措施使得可以采用极其廉价的节点来构成云;“云”的自动化管理使数据中心管理成本大幅降低;“云”的公用性和通用性使资源的利用率大幅提升;“云”设施可以建在电力资源丰富的地区，从而大幅降低能源成本。

# 2.3 云计算类型与层次

云计算是以客户为中心，向客户提供服务的一种计算模式。然而，供应商能够提供的服务差异性比价大。根据供应商所能够提供的服务的功能和模型，可以将云计算服务分为三个类别，分别为：基础设施即服务（IAAS），平台即服务（PAAS）和软件即服务（SAAS）。

## 2.3.1 IAAS

IAAS意思为：基础设施即服务（Infrastructure as a Service）。服务供应商将自己的硬件设备作为基础设施出租给用户，可以出租的设施有许多。比如：计算能力，网络，存储和其他计算资源。客户可以根据自己的需求，将自己的程序部署在供应商提供的基础设施中，让其他终端用户来使用。服务供应商根据客户的需求来收取一定的费用，并且为客户维护这些基础设施。这是一种依靠出租基础设施作为服务的模式。

亚马逊（Amazon）是这个市场的先驱者，亚马逊的弹性云主要提供的就是已经安装客户操作系统的虚拟机AMI（Amazon Machine Images），客户将自己的程序预安装在这些虚拟机中，并将这些虚拟机镜像提交给服务器，虚拟机镜像在客户的控制下运行。

## 2.3.2 PAAS

PAAS意思为：平台即服务（Platform as a Service）。服务供应商提供的是一个软件平台。研发人员在该软件平台上进行软件研发和测试，最终将其开发的软件部署在该平台中。PAAS开发套件一般是免费的，客户仅仅需要为托管程序支付一定的费用。目前，最受欢迎的PAAS平台有谷歌的AppEngine，SaleForce公司的Force.com；开源的PAAS平台—CloudFoundry也表现非凡。

PAAS将软件平台作为一种服务提供给客户，这种服务被用来构建高水准的软件。从服务供应商和客户的观点来看，对于PAAS的观点是不相同的。

1. 服务供应商需要对PAAS的平台进行设计，来满足自己需要提供的服务。供应商一般通过集成操作系统中的组件，应用程序和中间件，甚至也包括作为开发环境的开发套件来生成这个PAAS平台。比如：开发PAAS的人员有可能集成某些IAAS中的控制系统，利用该控制系统可以将客户部署在其中的程序运行在基础设施中；集成某些服务提供给客户使用，并且提供开发套件，让客户能够开发出无需移植的程序。
2. 客户使用PAAS的时候，使用供应商提供的开发套件来开发程序。客户可以调用API与PAAS平台进行交互。PAAS平台提供了许多服务和可编程元素，这些可编程元素可以是程序框架，也可以是业务逻辑，客户在开发自己的应用程序的过程中可以充分享受这种便利，为客户开发应用节省了部分资金。客户开发完成后将应用部署到PAAS平台中，这些应用将被部署在一个虚拟设备中。每一个虚拟设备成为一个实例（instance），在PAAS平台的控制下运行。

PaaS平台可以用来执行软件开发过程中的各个阶段的研发和测试，也可以专门用于某个特定的领域，如内容管理。PaaS平台中比较出名的商业示例有：Google的App Engine，它利用Google强大的基础设施来为客户提供应用程序服务；Microsoft的Azure平台，其整合了微软在.net和数据库方面的优势，为客户提供了一个基于windows操作系统的云平台。由于市面上的PAAS平台种类比较多，其能够提供的功能服务因供应商而异。

## 2.3.3 SAAS

SAAS意思为：软件即服务（Software as a Service）。供应商将应用软件统一部署在自己的云平台上，客户可以根据自己实际需求，通过互联网向厂商定购所需的应用软件服务，按定购的服务多少和时间长短向厂商支付费用，并通过互联网获得厂商提供的服务。用户不用再购买软件，而改用向提供商租用基于Web的软件，来管理企业经营活动，且无需对软件进行维护，服务提供商会全权管理和维护软件。有些软件厂商在向客户提供互联网应用的同时，也提供软件的离线操作和本地数据存储，让用户随时随地都可以使用其定购的软件和服务。对于许多小型企业来说，SaaS是采用先进技术的最好途径，它消除了企业购买、构建和维护基础设施和应用程序的需要。

在这种模式下，客户不再象传统模式那样花费大量投资用于硬件、软件、人员，而只需要支出一定的租赁服务费用，通过互联网便可以享受到相应的硬件、软件和维护服务，享有软件使用权和不断升级；公司上项目不用再像传统模式一样需要大量的时间用于布置系统，多数经过简单的配置就可以使用。

# 2.4 云计算的规则

## 2.4.1云计算的开发体系架构规则

(1)原则l——在云中集成生态系统管理

这个框架必须能够支持云计算中的生态系统管理。该系统包括让所有提出服务和解决方案的供应商、伙伴和终端用户能够在云计算环境中共享资源。云供应商提供内部操作、外部交互界面和产品开发能力。云伙伴向云供应商提供部件或者作为云供应商的代理向云客户提供增值服务。云客户是云服务的使用者，同时提供业务目标促使资源共享。

从框架的设计原理来看，云供应商操纵盘提供了将云供应商前台交互界面与后台操作相结合的想法。比如，在前台云供应商操纵盘可以支持销售业务发展活动，提供服务和客户支持。在后台设计解决方案、开发活动、以及主机环境都是用来支持前台的界面操作。由于大多数的云供应商不是单独工作，在云计算环境的价值链中它们需要与它们的伙伴合作。因此，云伙伴的操纵盘需要与云供应商和云客户打交道。例如，如果云伙伴作为云供应商的一部分，建立与云供应商的交互和与策略管理模块的合作是整个价值链的核心。在云计算中云客户或者终端用户可以分为两类：企业用户和消费者用户。云客户操纵盘为所有的与云计算服务交互的用户提供了一个接口。这个接口为用户提供了一个统一的框架，通过多种渠道为用户提供服务，如门户网站渠道、基于程序的业务和业务的经营合作渠道、或者手机用户渠道。在基于多价值策略和安全性以及软件服务的其他特征上，探索为企业和用户服务的软件服务框架仍然还有很多机会。因为企业用户和消费者用户在服务系统中是并存的角色，一个企业用户可以是由多个消费者用户所组成的。最终，它们只是云计算中不同级别的消费者。把所有这些操纵盘结合起来，云计算生态管理层(1A)提供了一个集成的登记过程和共同的设施管理来支持在云供应商、云伙伴和云客户之间的无缝合作和信息交流。

(2)原则2——虚拟化的云基础设施

在云计算中有两种支持虚拟化的方法。第一种方法是硬件虚拟化，将硬件设备设置为即插即用的管理模式。硬件设备可以添加或删除不影响系统中的正常操作和其他设备。当然，因为那些添加和删除的行动、性能和存储空间会动态地改变。第二种方法是软件虚拟化，即使用软件镜像管理或者软件代码虚拟化技术使软件共享。确切地说，软件镜像的创建可基于一套软件系统的可重用程度，包括操作系统、中间件和应用程序。其他的软件虚拟化技术还有动态代码的组装和执行，在这种情况下没有软件镜像，代码可以动态地从库中拷贝然后粘贴到业务逻辑中合适的地方。如在网络应用程序里，一些JavaScript代码可以动态地检索并正确地插入Ajax包中来创建客户端新的函数或功能。同样，根据可重用代码和实时编译技术，服务器端程序可动态地装配和执行。随着多核技术和并行程序设计的发展，这种动态地组装和执行程序将有极大的发展前景。特别是这种技术将会解决中间件和开发工具的镜像占用巨大存储空间这个难题。如今，在实际运用中虚拟技术可以并存而且可以相互支持。简而言之，云计算的IT基础设施管理模型包括软件镜像管理、硬件虚拟化和传统应用程序的包装。

(3)原则3——面向服务的共同可重用服务

如前面介绍，除了虚拟化特点，面向服务是另外一种驱动力使云计算进一步从资产可重用性、组合应用程序和混合服务中实现商业价值。有两种通用的云可重用服务：横向业务服务和垂直业务服务。

云横向业务服务包括各种平台服务，其隐藏了中间件、数据库和工具。在云计算环境中除了提供中间件和其他开发工具，也提供了一些通用的服务，如注册，安装、监控、计费的工具，或者还有一些跨行业的服务，如客户关系管理(CRM)和企业资源规划(ERP)等。云垂直业务服务包括所有的特定领域或工业领域中的公用设施服务，如航运和支4,-tN务等。

(4)原则4---可扩展的供应和订阅云

扩展的供应服务为云计算系统的一个特色。没有可扩展性，云计算的供应部分只支持某些类型的资源共享。这意味着供应服务结构对于免费用户和付费用户都是一样的。这两种类型的用户都可以是服务的供应者或者消费者。这里强调一个关键问题，即在云计算结构中如何处理服务供应商的供应过程和消费者的订阅过程。我们可以把云核心分类为云供应服务和云订阅服务。云供应服务包括供应过程、角色定义和通知框架。云订阅服务包括订阅过程、角色定义和通知框架。角色属性定义和通知框架的定义在云供应和云订阅是可以共享的。

(5)原则5——可配置的云计算输出

云输出是云计算平台最终的产品或服务。由于所有的云输出要解决某些业务目标，云输出也被称为“云”的商业解决方案。运用SOA参考架构的可扩展性和可配置性，可以解决云计算平台和服务的可配置性。生态系统管理模块、虚拟化模块、面向服务模块和云核心形成了一个基础，以确保计算平台的可配置、可合成和可管理。

绝大部分的云输出都是通过web浏览器进行传输。例如，我们可以使用浏览器来上传图片、声音、视频文件和其他文件到存储云。在所有其它的云输出中，web接口已经提供了有效的管道使云客户能够与云伙伴和云供应商的在周期服务中交互。

(6)原则6——统一的信息交流和表达框架

云计算资源的信息表达和交换对于云计算各功能的协同工作是非常重要的。云计算资源包括所有的业务实体(如客户、伙伴、云供应商)和支持资源(如虚拟化相关模块、面向服务相关模块、云核心和云输出)。例如在云的生态管理系统中，国家名、网站名、组织结构名都与云计算的业务实体名相关联。角色是用来定义动态云计算生态系统中的功、能。项目、任务、文件、交易、业务流程、参考链接、注释、事件是在云计算环境中支持各项业务实体合作的潜在资源。

(7)原则7——云的质量和治理

从质量指标的角度看，服务质量参数可直接用于定义云实体的可靠性、响应时间、安全、诚信。诚信可以通过检查可追溯性和可验证性。安全是云质量的一个十分重要的方面，只有经过授权的商业实体或使用者才可以获得正确的云计算资源。云实体的访问控制、隐私、保护等方面形成了云计算的信任和安全基础。

从治理的角度看，云计算环境和服务可以借鉴SOA治理中许多优秀的实践。在项目执行阶段、监控、计量、结帐、异常处理要启用并且要协调跨系列业务的生产、经营和管理各种模块的云输出。

## 2.4.2 资源管理与调度原则

资源调度与管理通常情况下处于各种平台的核心地位。它深刻地影响着对于某个平台的评定标准的三方面：性能，功能和能耗。一个效率低的资源管理方法直接导致了平台的性能低下，能耗过大，间接的导致了系统的功能质量差；一个效率高的资源管理方法能够大大的提升平台的性能，并且能够降低能耗，从而，平台的功能也能够获得质量的保证。

云计算平台是一种拥有大量共享资源的复杂系统，其根据难以预测的请求和外部事件来管理资源，所以，云平台不能够按照硬性条约来对资源进行管理。云资源管理需要综合的策略。云资源管理可以概括为5大类。

1. 权限控制（Admission control）。权限控制的目标是防止未经授权的用户或者系统模块对资源进行操作。例如，某个未经授权的用户操作实例的运行是不被允许的。
2. 分配容量（Capacity allocation）。分配容量意思是为每一个独立的实例分配资源，以支持其运行。
3. 负载均衡（Load balancing）。由于云平台处理的请求是未知的，并且请求数量可能在某段时间内爆发，从而给云平台带来大量的负载，为了让云计算平台能够平稳的处理这些请求，需要支持负载均衡。
4. 能耗优化（Energy optimization）。对于一个大规模的云计算平台，其拥有的服务器数量是巨大的。每天消耗的能量也是惊人的，为了节约能源，同时也为了节省运维资金，对云平台进行能耗优化是非常有必要的。
5. 服务质量保障（Qos）。作为一个商业服务平台，云计算首要任务是为用户提供服务，那么，这些服务必须要有质量的保证。如果一个云平台不能够提供高质量的服务，那么这个平台不会受到客户的青睐。

# 2.5 本章小结