

云计算简介

📅 2017-10-25 | 📁 [软件基础设施](#)

本文源自一次内部科普，目标是让非IT人员能够对云计算有一个大致的了解。

我比较倾向于基于云计算的内涵，狭义的去定义云计算，因此可能与网上别人的观点不完全一致。

概念篇

云计算是什么

出于 不同理解 商业目的，云计算的定义不下20种，解释不下100种。

我的理解：

- 云(Cloud)
在IT领域，云通常指资源的集合。其特点是规模庞大，细节不重要，可动态伸缩。
- 计算(Computing)
这里应该指计算能力。计算能力取决于可用于计算的资源，包括软件运行需要的各种平台软件、操作系统、CPU能力、内存大小、存储空间、网络带宽等。
- 云计算(Cloud Computing)
通过分布在 **大量** 计算机上的资源构成资源池，为 **大量** 应用系统提供所需的
计算力、存储空间、网络带宽和其他服务。应用系统不需要知道支持其运行的各种资源的底层细节情况。

云计算的特征

基于上述定义，可以得出云计算在 **提供资源方面** 的特征：

- 多对多
用大量的计算机，为大量的应用提供所需的资源。
- 可动态伸缩
云的规模可以根据需要随时扩大或减小，而不影响上面运行的应用系统。
- 成本更低
通过动态的分配和调度，提高资源使用效率，使得单个应用系统运行所需的成本更低。
- 灵活使用
应用系统可以根据需要，随时获取、按需使用资源，而且可以做到按使用计费。
现在有的云计算平台甚至可以做到按分钟、按每M流量计费。

常见 误区 误导

- 云计算不一定依赖互联网
不依赖互联网的不一定不是云，比如私有云。
依赖互联网的不一定是云，很可能是伪云。
- 云计算是服务于开发/运维人员
对于应用系统的用户来说，其实并不关心其是否运行在“云”上。

什么不是云计算

- 云应用：基于云计算的应用
如前所述，用户只关心应用的功能和性能，不关心应用如何实现，如何部署。
通常，大部分应用不在云平台上也能运行。
- 伪云：把简单的互联网应用套在云计算上
 - 搜狗云输入法 --> 网络输入法（词库放在服务器上）
 - 有道云笔记 --> 网络笔记（笔记内容放在服务器上）
 - 网易云音乐 --> 网络音乐（音频文件放在服务器上）
 - 百度云网盘 --> 网盘（各种文件放在服务器上）

原理篇

分层体系架构

分层体系架构(Layered Architecture),是IT领域的重要思想。

分层体系架构将系统的组件分隔到不同的层中，每一层中的组件保持内聚性，并且大致在同一抽象级别(**高内聚**)；每一层都通过接口访问下一层，不关心其具体实现。下层不会调用上层(**低耦合**)。这样把问题划分开来各个解决，易于控制，易于延展，易于分配资源(**分而治之**)。

分层架构可以带来很多好处：

- 每层专业分工
- 层间接口固定，替换某一层不影响其它层
- 有利于通用组件的复用
- 利于系统的水平扩展

本文相关的分层架构包括：

- 开发角度的分层架构

展现层
逻辑层
存储层

- 运维角度的分层架构

应用软件
平台软件
操作系统
硬件

硬件不足与集群

之前很长一段时间，计算机硬件的能力不足以满足软件的运行。

主要的解决手段是集群和负载均衡技术。与升级硬件相比，这种扩充能力的方式成为“水平扩展”。

- 集群(Cluster)

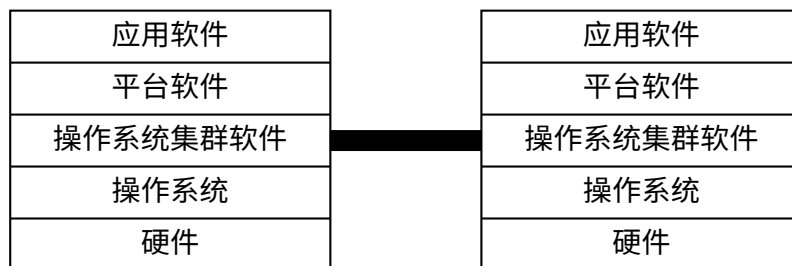
将多个计算机连接到一起，像一台机器那样工作。这样可以增加处理能力，或者提高稳定性。

- 负载均衡(Load Balancing)

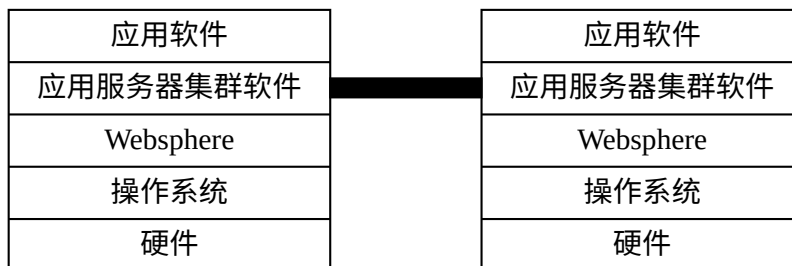
按照一定的规则，将负载分配到集群中的各个计算机上进行处理。

集群和负载均衡技术，可以应用在不同的层级上面。其上的各层，都运行在集群环境。

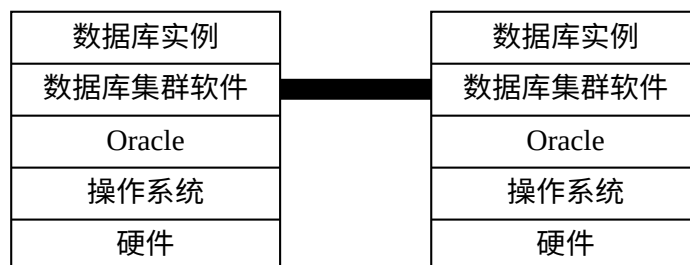
- 操作系统集群



- 应用服务器集群



- 数据库集群

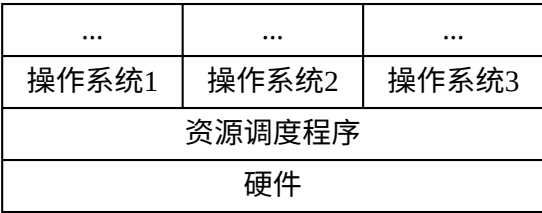


硬件过剩与云计算

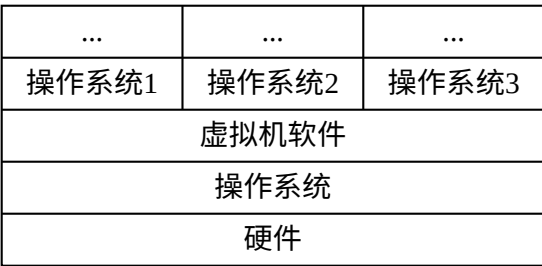
随着硬件技术的飞速发展，现在计算机的性能远远高于软件的需要。也就是说，硬件资源过剩了。为了提高使用效率，产生了一些新的技术。

虚拟化

一台计算机只运行一个操作系统，资源可能闲置。通过虚拟化，可以在一台计算机上运行多个操作系统，这些操作系统共享硬件资源。这里面的关键是资源调度程序，或者叫“虚拟化操作系统”。

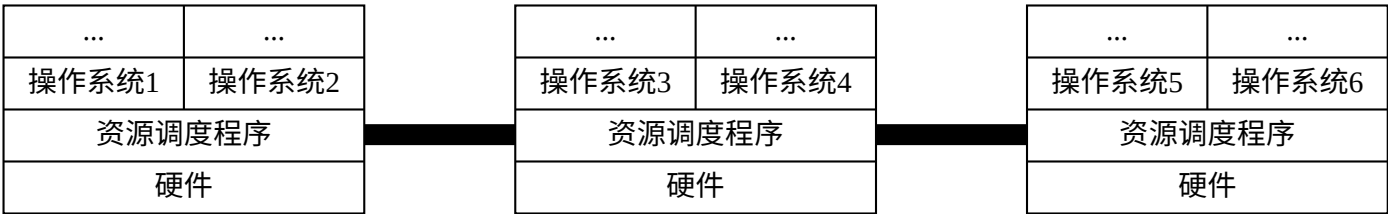


可能有人运行过单机的“虚拟机软件”，比如vmware。虚拟机和虚拟化的本质不同在于：资源调度程序运行在不同的层上。



虚拟化+集群

单个的计算机上面能够运行的虚拟机个数还是有限的。要想进一步增加容量，很容易想到可以把前面提到的集群技术用到虚拟化的资源调度程序上：



这样就构建了一个大的资源池。

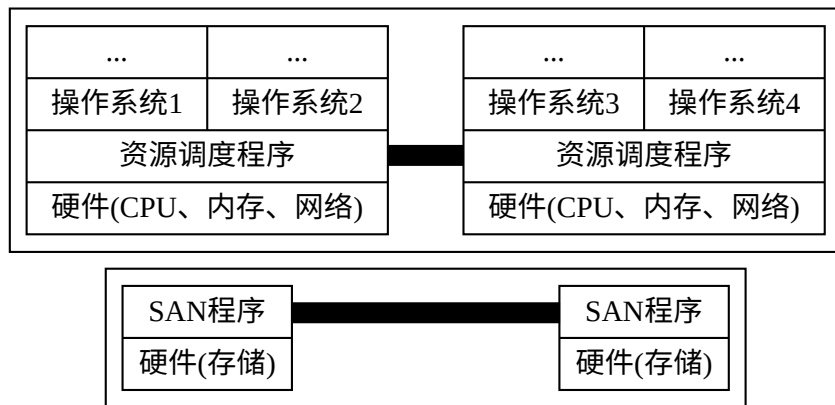
云存储

上面的架构中，虚拟机还不能无缝从一个主机迁移到另一个主机运行。关键在于：虚拟机是存储在硬盘上的，一个主机无法访问另一个主机的硬盘。

最简单的办法当然是所有的虚拟机都存储在一个存储设备(比如EMC存储)上，所有的主机都能访问这个存储设备。但这显然是多对一的关系，而且不可动态伸缩。

我们需要一种符合云计算特征的存储功能，不妨称作云存储(Cloud Storage)。

如果把“文件存储”看做一种特殊的应用，这个问题可以很容易的用前面的模式解决：



其中SAN(Storage Area Network，存储区域网络)，是一种服务器连接外部存储设备的架构技术。从操作系统的角度，访问外部的存储设备与访问本机的硬盘没有任何区别。

该技术可以采用光纤等设备，使得访问外部存储的速度达到甚至超过本机硬盘。

管理和调度

现在，已经有了足够的灵活性：可以创建虚拟机，可以为虚拟机分配和调整资源，可以指定虚拟机运行在哪一个主机上。但是还可以更加智能一些。

- 自动分配
创建新虚拟机时，根据需要的资源和各主机的情况，自动分配主机
- 自动平衡
监控各个主机的运行状况，当主机资源紧张时自动将虚拟机迁移到空闲的主机上运行，在此过程中对虚拟机访问不受任何影响。
- 自助服务
用户可以自助创建虚拟机，决定使用多少资源。并按照实际使用情况实时计费。
- 区域自治
对于某些用户，可以为其分配一个大的资源池（比如，5000GHz CPU、3TB 内存、3PB 存储空间、1TB网络带宽），这个资源池的管理员可以创建子用户，子用户使用池中的资源再创建自己的虚拟机。
- 更多功能
没有做不到，只有想不到！

XaaS

尽管业内搞死了SOA，但是服务(Service)这个词已经到处在使用，什么东西都可以是服务。

在云计算领域，根据提供资源的层次不同，可以分为 IaaS, PaaS, SaaS。

- IaaS (基础设施即服务)

前面的例子中，都是把CPU、内存、存储空间、网络带宽以及管理这些硬件的操作系统作为资源。

这些资源提供出去，就叫做基础设施即服务（IaaS, Infrastructure-as-a- Service）。

- PaaS（平台即服务）

PaaS把平台(Platform)作为资源。什么是平台，还要从分层架构说起。

开发一个应用系统，总有很多共同的功能，比如http服务，session管理，数据库连接池等等。如果每一个应用系统的开发都重新实现这些功

能，效率太低。于是有一些组织，将这些公共的功能提取出来，发布成一个软件，就叫做平台软件。
平台软件也叫做容器，可以支持很多应用系统的运行。平台软件在架构的层次中，介于应用系统和操作系统之间。

- SaaS(软件即服务)

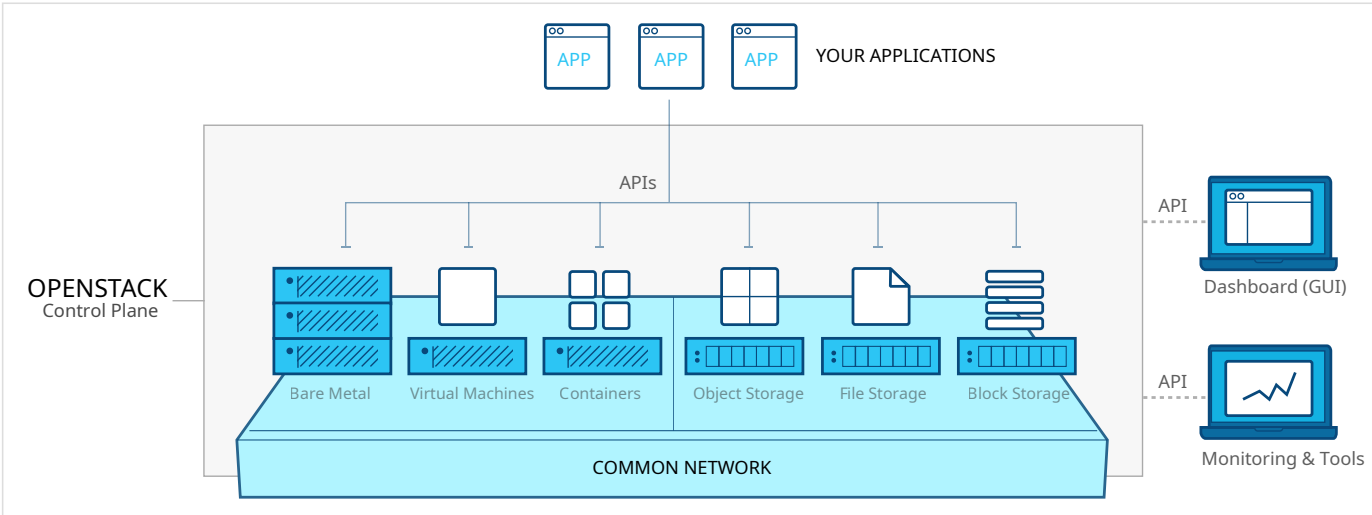
平台软件是通用的，业务无关的。但对于某些特定类型的应用，有一些共同的功能和逻辑。比如CDN、视频直播等、用户认证等。
把这些功能提取出来，作为一种资源在多个应用系统中复用，就可以称之为 SaaS。
这些服务的对象，仍然是开发人员。对于最终用户来说，这些服务并不能直接使用。

OpenStack一出，谁与争锋？

早期搭建云计算平台时，需要整合各个厂商的技术和软件，甚至在此基础上自习开发一些工具。所以搭建一个云平台的难度很大。

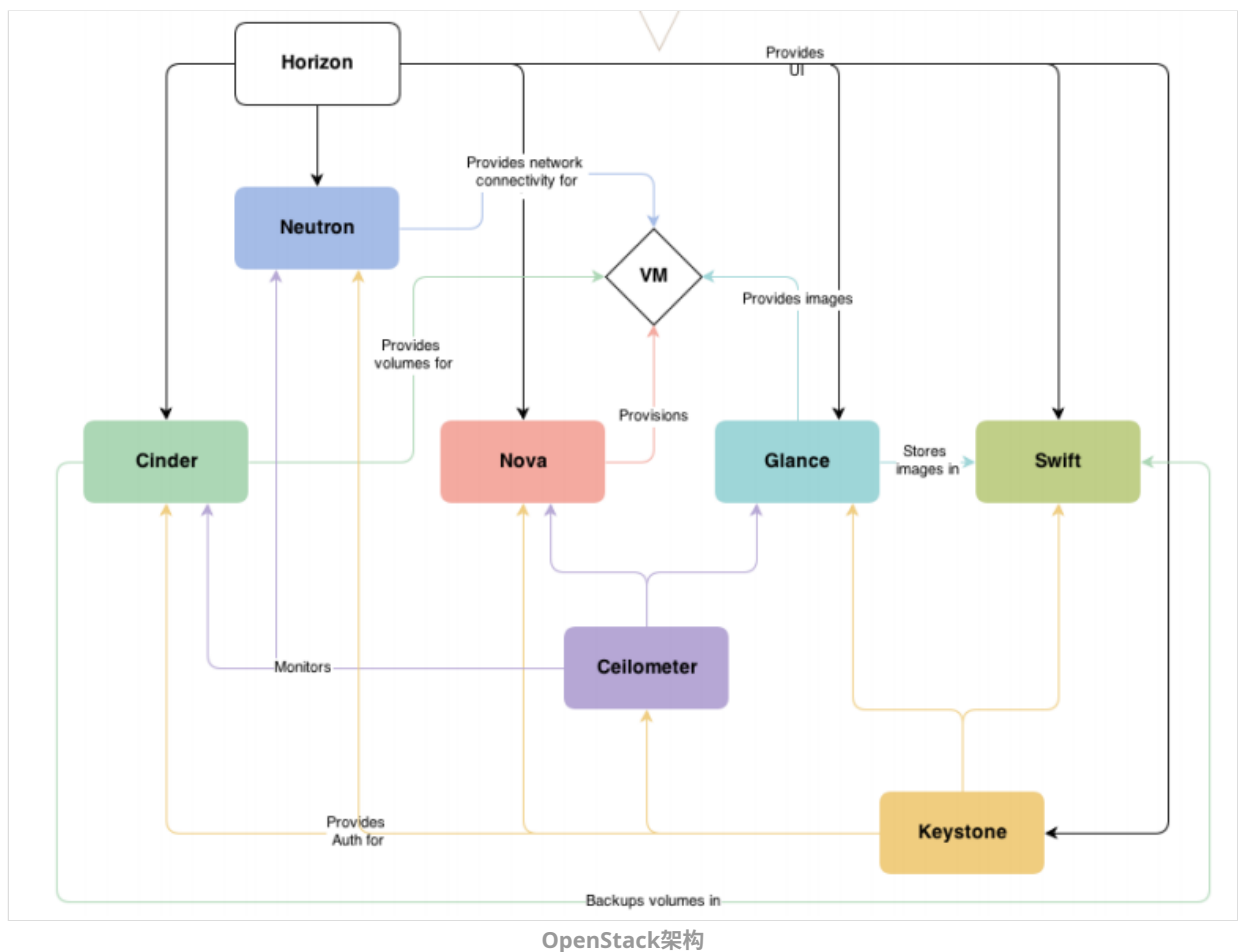
2010年7月，NASA伙同AMD、Intel、Dell等厂商，共同宣布“OpenStack”开放源代码计划，后来微软，思科等厂商也陆续加入。

OpenStack宣称自己是一个“云操作系统”，可以将大量的计算、存储和网络资源整合为一个数据中心，并提供统一的管理控制。而用户可以通过web界面定制自己需要的资源。



OpenStack概念图

OpenStack通过一系列的开源软件组成”战斗群“，实现了搭建一个云平台的所有功能：



其中：

- VM: 虚拟机，所有的功能都围绕虚拟机展开。OpenStack可以管理常见的各类虚拟机，如开源的Xen、KVM，和商业的Hyper-V、VMware等
- Nova：管理 VM 的生命周期，是 OpenStack 的核心
- Neutron：提供虚拟网络和物理网络连接
- Glance：管理 VM 的启动镜像，Nova 创建 VM 时将使用 Glance 提供的镜像
- Cinder：为 VM 提供虚拟硬盘
- Swift：提供对象存储服务
- Keystone：认证和权限管理
- Ceilometer：监控和计量服务，为报警、统计或计费提供数据。
- Horizon：为用户提供一个 Web 的自服务 Portal。

通过OpenStack，可以搭建任何类型的云平台，比如：

- Web应用
- 大数据平台
- 电子商务
- 视频直播和内容分发
- 高性能计算(High Throughput Computing)
- 虚拟主机
- 公有云
- 数据库云(DBaaS)

行业篇

云计算简史

云计算从概念，到技术研究，到现在的广泛使用，大概经历了以下几个阶段：

- 概念的建立

- 1983年，Sun 提出 The Network is the Computer
- 2006年3月，Amazon 推出弹性计算云（Elastic Compute Cloud；EC2）服务
- 2006年8月，Google 首次提出“云计算”（Cloud Computing）的概念

- 试验田

- 2007年10月，Google 和 IBM 开始在美国大学校园推广云计算的计划
目标：降低分布式计算技术在学术研究方面的成本
通过数百台服务器提供服务，学生可以通过网络开展以大规模计算为基础的研究
- 2008年1月，Google 宣布在台湾启动“云计算学术计划”

- 吃螃蟹

这个阶段，主要提供面向个人开发者和小企业的PaaS服务。

- 2007年6月，Heroku 对外提供服务
- 2008年4月，google发布GAE(Google App Engine)
- 2009年8月，新浪开始开发SAE(Sina App Engine),11月正式推出第一个Alpha版本

- 抢市场

- 2008年2月，IBM 宣布在无锡科教产业园为中国的软件公司建立全球第一个云计算中心
- 2008年7月，雅虎、惠普和英特尔宣布推出云计算联合研究计划，在全球建立6个云计算的数据中心
- 2008年8月，戴尔申请“云计算”（Cloud Computing）商标
- 2009年8月，李彦宏提出"框计算"

- 抢占制高点

- 2010年3月，Novell与云安全联盟（CSA）共同宣布一项 供应商中立计划，名为“可信任云计算计划（Trusted Cloud Initiative）”。
- 2010年7月，美国国家航空航天局和包括Rackspace、AMD、Intel、Dell等支持厂商共同宣布“OpenStack”开放源代码计划
- 2010年10月，微软表示支持OpenStack与Windows Server 2008 R2的集成
- 2011年2月，思科系统正式加入OpenStack，重点研制OpenStack的网络服务

- 商业化服务

目前比较适合商业化的，还是面向企业级客户的IaaS服务。

- 2011年7月，阿里云上线，开始大规模对外提供云计算服务(2009年创立公司)
- 2013年7月，青云上线

价值

云计算，能够极大降低IT开发、运行的成本。目前主要是通过IaaS、PaaS降低运行成本。随着SaaS的积累，开发成本也会降低。

比如，有了公有云，就不需要自建机房，操心供电、空调、消防等一系列问题，也不需要购买服务器、网络设备和这些设备的维护和折旧。

只需要支付资金就可以获得所需的一切，甚至比自建的成本还要低。

即使出于种种原因不能使用公有云，搭建私有云也能带来很多好处。比如，传统的一个应用系统上线，需要采购服务器、服务器上架、安装操作系统、安装平台软件，然后才能部署应用系统，这个过程少则一周，多则有可能一个月。而采用私有云技术，只需10分钟，部署应用系统的基础环境就可以准备好。

云计算另一个巨大的潜力是，由于云的伸缩能力，可以为应用系统提供近乎无限的资源，可以用廉价的方式达到以前中型机甚至大型机的能力，这就使得原来很昂贵的某些应用，现在可以从实验室走出来，变成商业应用。典型的就大数据类型的应用。

参与者类型

云计算在技术上已经比较成熟了，在整个产业链中，有多类参与者：

- 云技术

搭建、管理私有云/公有云所需要的产品和技术。

一个重要的趋势是，软件开源、免费。厂商支持开源软件的开发，通过这种方式推广自己的“标准”。

厂商通过解释自己的标准(以解决方案、技术支持的形式)赚钱，不再赚实施的辛苦钱。

如果硬件厂商针对云计算推出了新的硬件设备，也可以归为此类。

- 云服务

这里主要指提供公有云服务，供使用者租用。竞争的方式主要是通过规模效应拼成本。当然也要看个性化的服务功能。

从云服务的提供者和使用者的关系上，可以分为：

- 私有云

- 公有云

- 混合云

- 共享云

- 云应用

与“伪云”不同，有些应用必须在云计算的环境下运行，比如基于大数据的应用。

这些应用能够带来与传统应用不同的新的价值。

相关企业

- 技术提供商

虚拟化技术主要有VMWare, Citrix的Xen, Redhat的KVM，微软的hyper-v等。

实现云平台的技术有很多可以选择，但只有OpenStack是全面解决方案。

国内在底层技术上能提供的很少，阿里有一些项目可以开源使用。

- 服务提供商

- Amazon是最早的公有云提供商

- heroku依然把重点放在PaaS领域

- 由于众所周知的原因，google现在的情况未知

- 国内的公有云提供商主要有阿里云，青云，内容涵盖了IaaS, PaaS和 SaaS。青云价格更低，甚至实现了“按秒计费”

- 国内的IDC也在积极搭建公有云，如首都在线、世纪互联，主要把原来的主机托管和虚拟主机租用客户转移到IaaS上面，同时也逐步推出一些SaaS服务

- 百度云和新浪云，啥都有。但定位和目标未知

- 一些新秀，祝他们好运。比如又拍云，为视频点播、直播类的应用提供了很多SaaS组件

可能的新模式

目前，主要的业务模式是公有云通过规模效应建立成本优势，个人认为会越来越向大公司集中。

未来可能出现新的模式：

- 效用计算

1961年，人工智能之父麦肯锡在一次会议上提出了“效用计算”这个概念。目标是整合分散在各地的服务器、存储系统以及应用程序来共享给多个用户，根据用户的使用量来付费。

未来云计算也可能会出现真正意义上的共享云，但还有很多技术问题未解决。

已经出现过类似的模式是 P2P 下载。

- 网格计算模式

网格计算研究如何把一个需要非常巨大的计算能力才能解决的问题分成许多小的部分，然后把这些部分分配给许多低性能的计算机来处理，最后把这些计算结果综合起来攻克大问题。重点放在解决“计算力”上面。

从技术上，不是所有的计算问题都能分解，或者很难分解。从业务上，需要如此强大计算力的应用目前看来还很少。

已经出现过类似的模式是“在家寻找外星人”项目。

云计算

◀ 机器学习总结

用gitbook制作电子书 ▶

© 2018  Holbrook

由 [Hexo](#) 强力驱动 | 主题 — [NexT.Gemini](#) v5.1.4