



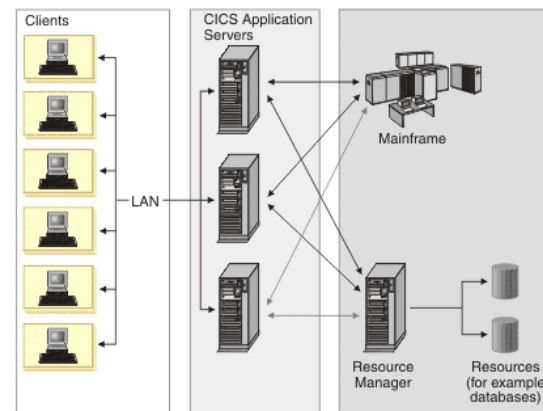
第一章 背景与概念

§1.1 背景与概念

§1.2 架构与模型

§1.1 背景与概念

- Distributed computing
- Internet and grid computing
- Cloud Computing



云计算诞生

- In the late 1990s, Salesforce.com
- In 2002, Amazon.com 启用 Amazon Web Services (AWS) 平台
- 2006, “云计算” 这一术语才出现在商业领域
- Amazon 推出其弹性计算云(EC2)
- 2009, Google 应用引擎(GAE)
- 随后成为热点和主流



Cloud Computing is an Evolution in IT

Grid Computing

- Solving large problems with parallel computing
- Made mainstream by Globus Alliance



Utility Computing

- Offering computing resources as a metered service
- Introduced in late 1990s



Software as a Service

Network-based subscriptions to applications

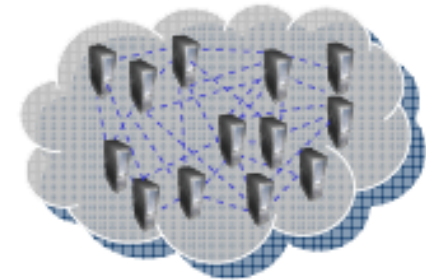
Gained momentum in 2001



Cloud Computing

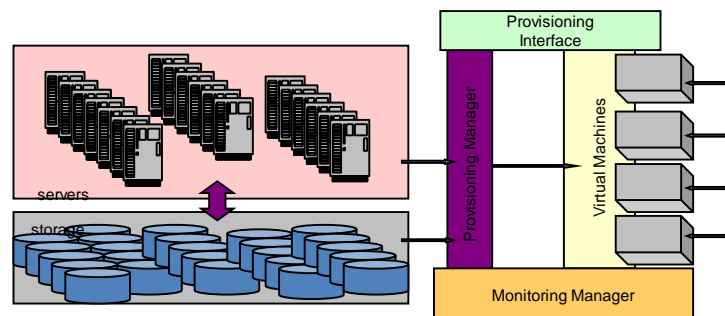
Next-Generation Internet computing

Next-Generation Data Centers



云计算定义

- Gartner report
...一种计算方式，能通过Internet技术将可扩展的和弹性的IT能力作为服务交付给外部用户。
- Forrester Research
...一种标准化的IT性能（服务、软件或者基础设施），以按使用付费和自助服务方式，通过Internet技术进行交付。
- NIST
Cloud computing is the delivery of computing as a service rather than a product, whereby shared resources, software, and information are provided to computers and other devices as a utility (like the electricity grid) over a network (typically the Internet). –from NIST



基本要点

云计算 =

数据 * (软件 + 平台 + 基础设施) * 服务

- 云计算强调服务，用户按需使用服务，根据使用多少付费。
 - 1: 提供“资源”——包括计算、存储及网络资源。
 - 规模巨大的全球化的数据库及存储中心，能够实现“海量”的存储、出色的安全性和高度的隐私性和可靠性
 - 此外，它还应是高效的、低价的、节省能源的。
 - 2: 提供动态的数据服务。
 - 数据包括原始数据、半结构化数据和经过处理的结构化数据。
 - 3: 提供云计算平台 ——包括软件开发API、环境和工具。

云计算服务

云计算 = 数据 * (软件 + 平台 + 基础设施) * 服务

- 数据 (Data)
 - 爆炸增长 (传感器、物联网) : $1.2\text{ZB} = 10^{21}\text{B}$
 - 各个领域各个层面
- 软件 (Software)
 - 检索、发现、关联、处理和创造数据
- 平台 (Platform) :
 - “云计算”时代也会诞生自己的通用平台
- 基础设施 (Infrastructure)
 - 存储资源、计算资源等
- 服务 (Service)
 - IT服务化: 产品 → 服务 XaaS

商业驱动力

- 容量规划
 - 领先策略(Lead Strategy): 根据预期增加IT资源的容量
 - 滞后策略(Lag Strategy): 当IT资源达到其最大容量时增加资源容量
 - 匹配策略(Match Strategy): 当需求增加时, 小幅增加IT资源容量
- 购置和维护成本
- 组织灵活性

技术创新

- 集群化
 - 集群式一组互联的独立IT资源，以整体形式工作。
 - 固有的冗余和容错特性。
- 网格计算
 - 计算网格为计算资源提供了一个平台，使其能组织成一个或多个逻辑池。
 - 这些逻辑池统一协调为一个高性能分布式网格。
- 虚拟化
 - 虚拟化是一个技术平台，用于创建IT资源的虚拟实例。



云使能技术（Cloud-Enabling Technology）

- 带宽网络和Internet架构
- 数据中心技术
- （现代）虚拟化技术
- Web技术
- 多租户技术
- 服务技术

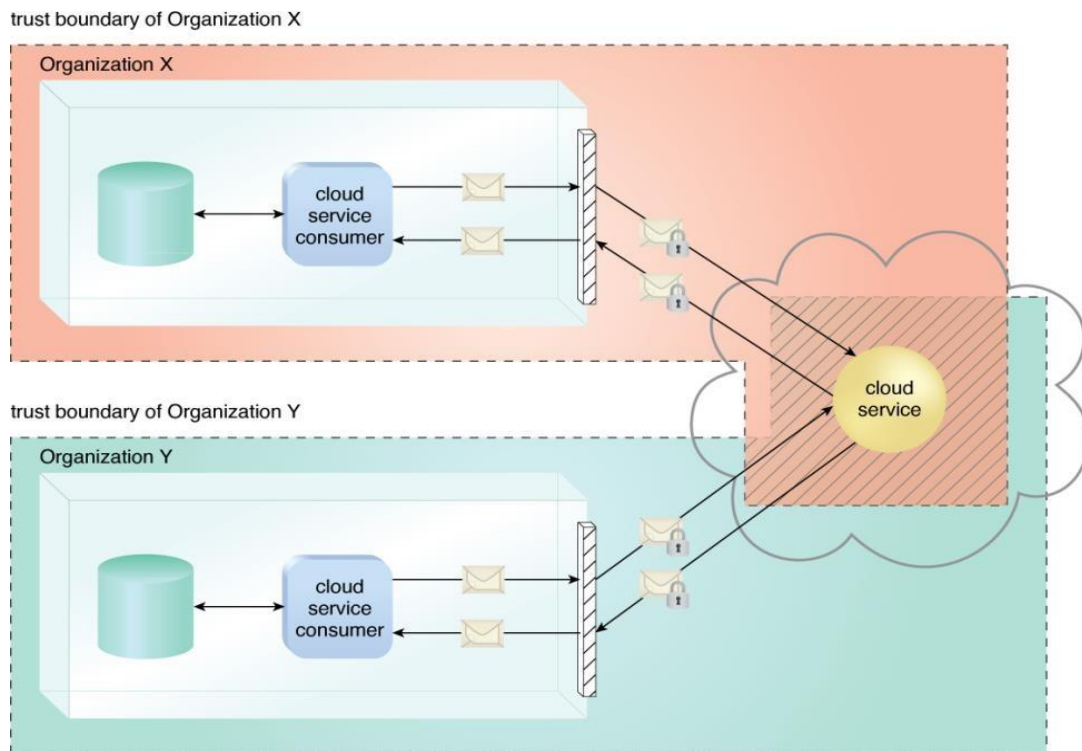


风险与挑战

- 网络空间安全风险
- 运营管理控制风险
- 软件移植性问题
- 跨地区法律风险

网络空间安全风险

- 云计算增加了安全风险
- 云用户将信任边界扩展到外部云，这就引入了漏洞
- 重叠的信任边界也为云提供者访问云用户数据提供了特权

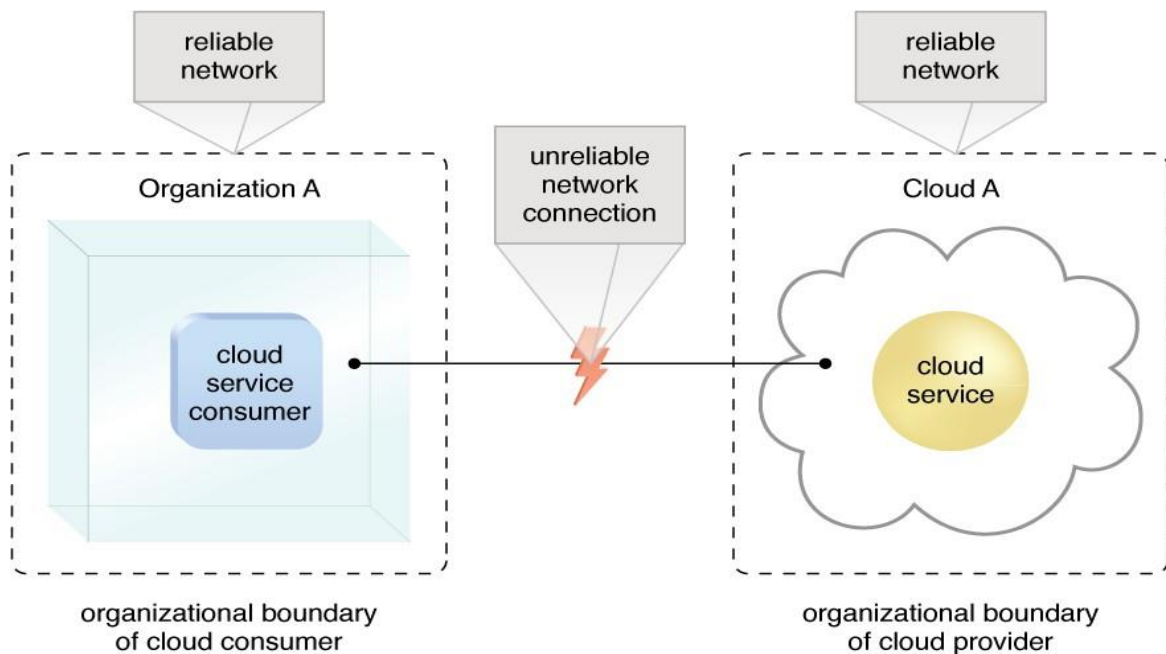


Copyright © Arcitura Education

信任边界重叠

运营管理控制风险

- 降低了运营管理控制能力
- 云用户对云资源的管理控制通常是低于对企业内部IT资源的管理控制的。
- 云用户和云提供者之间的地理、网络距离会导致延迟波动和可能的带宽受限。

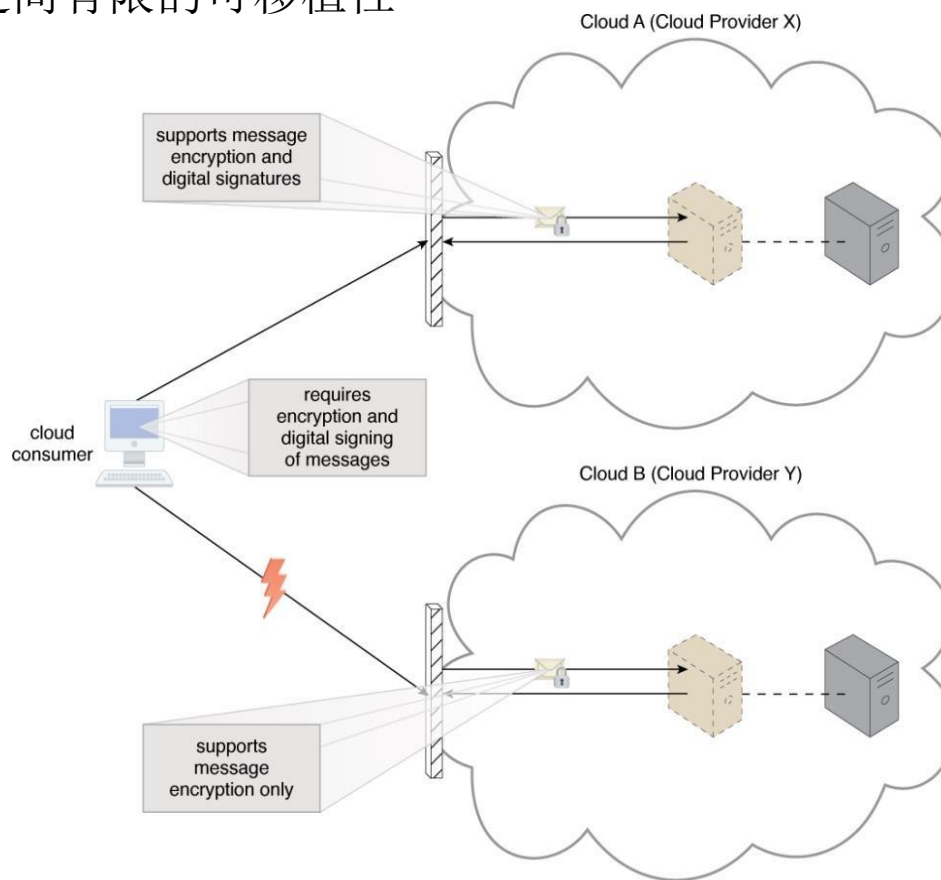


Copyright © Arcitura Education

不可靠的网络连接会影响通信质量

软件移植性问题

○ 云提供者之间有限的可移植性



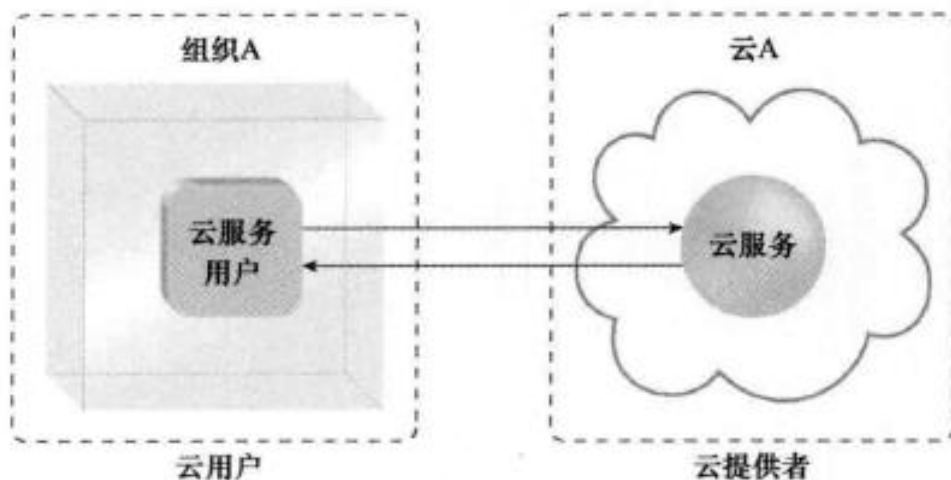
Copyright © Arcitura Education

跨地区法律风险

- 数据存放和处理
 - 当云用户的IT资源和数据被公共云处理的时候，他们通常不会意识到这些资源和数据的位置。
 - 这可能会造成严重的法律问题，因为这关系到规定了数据隐私和存储政策的行业或政府法规。
- 数据的获得和公开
 - 即使数据是由外部云提供者处理的，最终也是云用户组织对它们自己的数据的安全性、完整性和存储负责。

基本概念和术语——云用户和云提供者

- 提供基于云的IT资源的一方称为云提供者（cloud provider）
- 使用基于云的IT资源的一方称为云用户（cloud consumer）。
- 这两个术语通常代表的是与云及相应云供应合同相关的组织所承担的角色。



云用户（组织 A）与来自云提供者（拥有云 A）的云服务进行交互。在组织 A 内，使用云服务用户来访问云服务

基本概念和术语——可扩展性

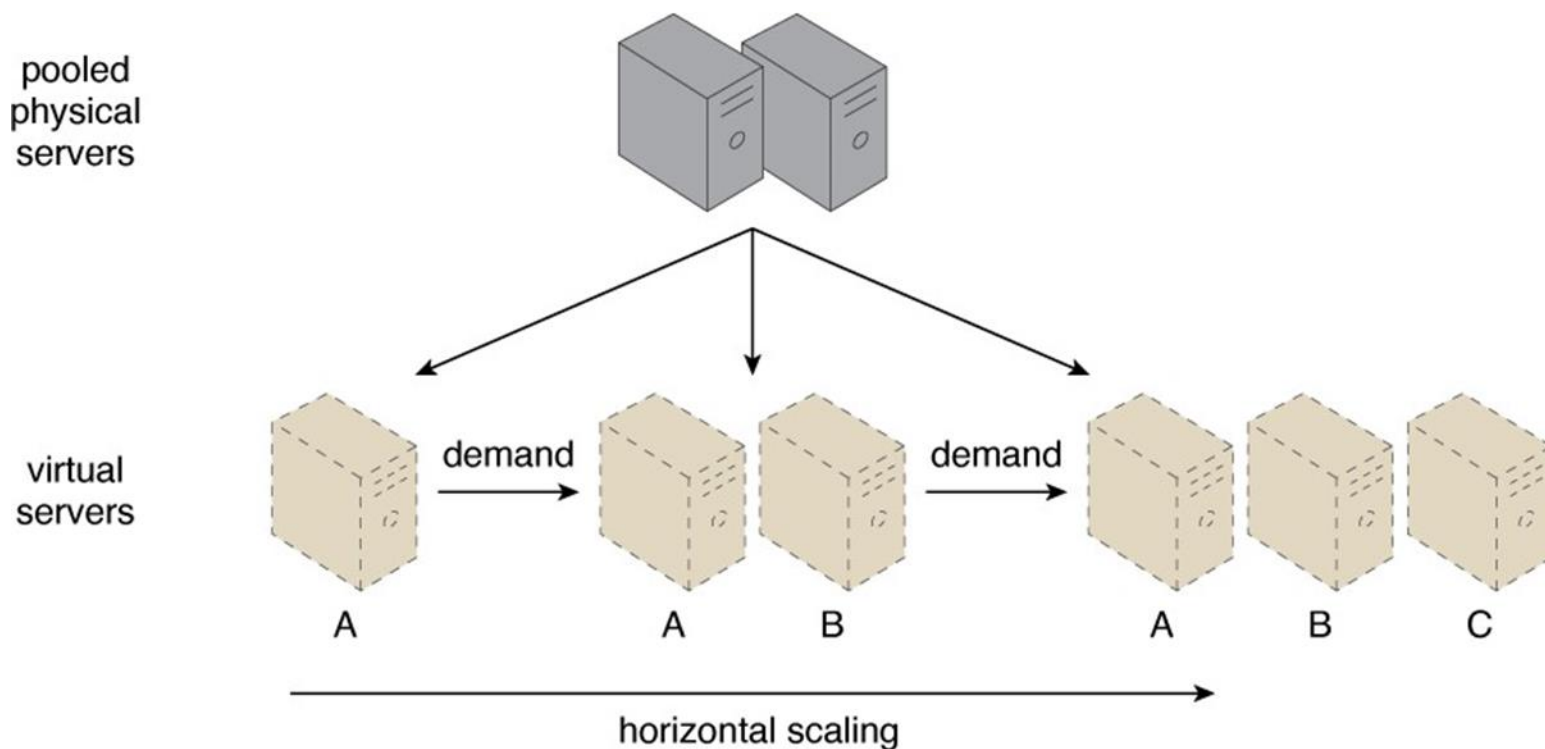
从IT资源的角度来看，可扩展是指IT资源可以处理增加或减少的使用需求的能力。

水平扩展（**horizontal scaling**）。水平分配资源也称为向外扩展（**scaling out**），水平释放资源也成为向内扩展（**scaling in**）。

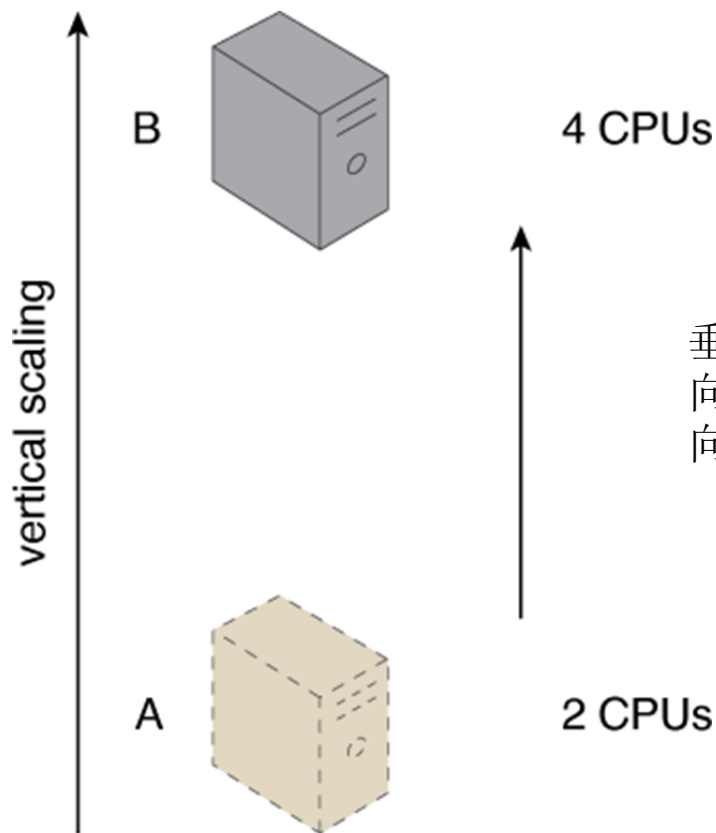
当一个现有IT资源被具有更大或更小容量的资源所替代，则称为垂直扩展（**vertical scaling**）。被具有更大容量的IT资源替代，称为向上扩展（**scaling up**），被具有更小容量的IT资源替代，称为向下扩展（**scaling down**）。

基本概念和术语——可扩展性

水平扩展（horizontal scaling）：改变IT资源的数量。
水平分配资源也称为向外扩展（scaling out）。
水平释放资源也成为向内扩展（scaling in）。



基本概念和术语——可扩展性



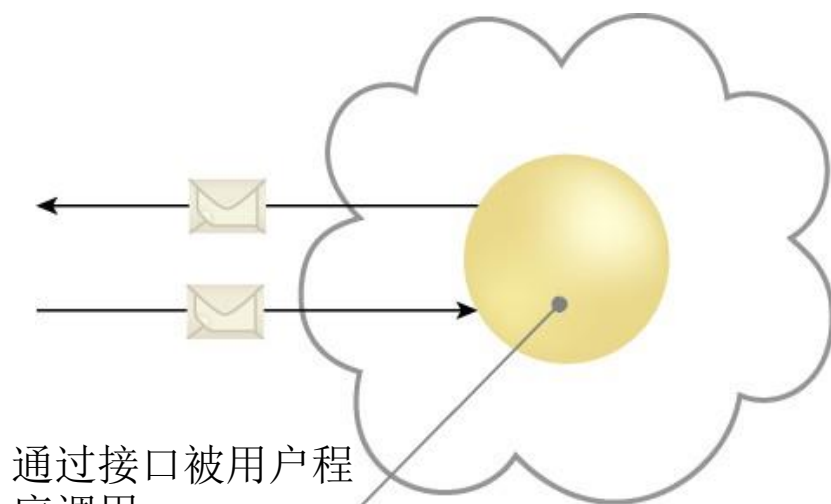
垂直扩展（vertical scaling）：改变IT资源容量。
向上扩展（scaling up）：分配更大容量的资源。
向下扩展（scaling down）：分配更小容量的资源。

Copyright © Arcitura Education

基本概念和术语——云服务

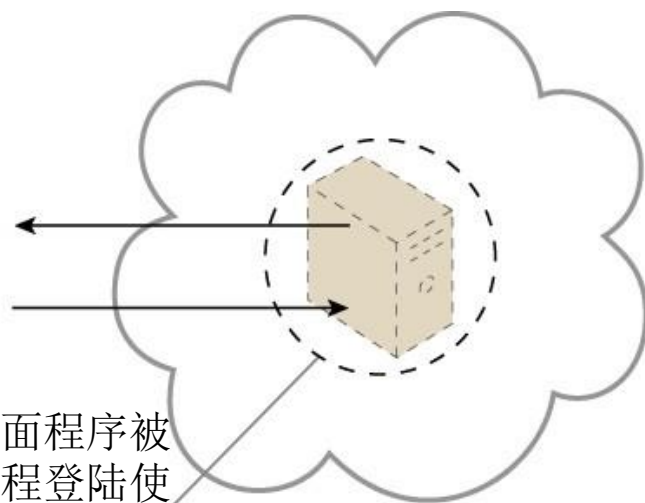
- 云服务（**cloud service**）是指任何可以通过云远程访问的IT资源。
- 云服务中的“服务”与其他IT领域中的服务技术（比如面向服务的架构，**SOA**）的“服务”含义更为宽泛。
- 并非云中所有的IT资源都可以被远程访问，其中有公开发布的**API**的软件程序可以专门部署为允许远程客户访问。

基本概念和术语——云服务



通过接口被用户程序调用。

remotely accessed Web service
acting as a cloud service



通过界面程序被
用户远程登陆使用。

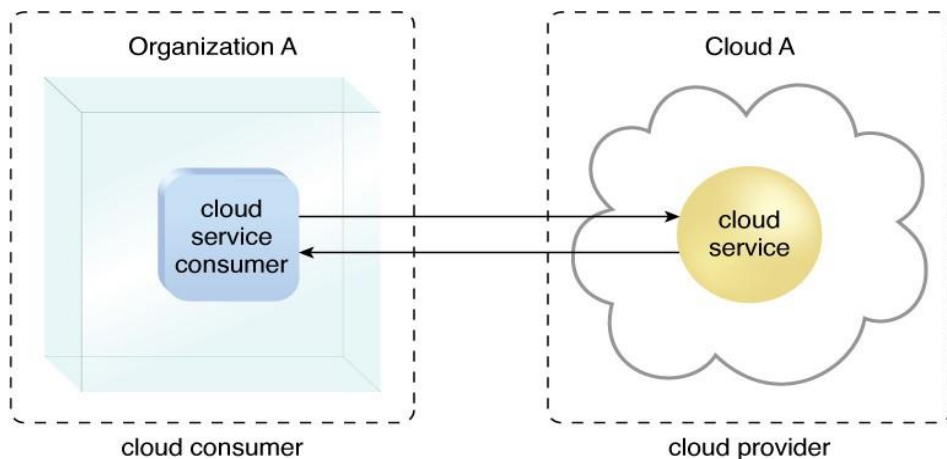
remotely accessed virtual server
acting as a cloud service

Copyright © Arcitura Education

具有已发布技术接口的云服务被云外用户访问（左）。
作为云服务的虚拟服务器也可以被云边界之外访问（右）。

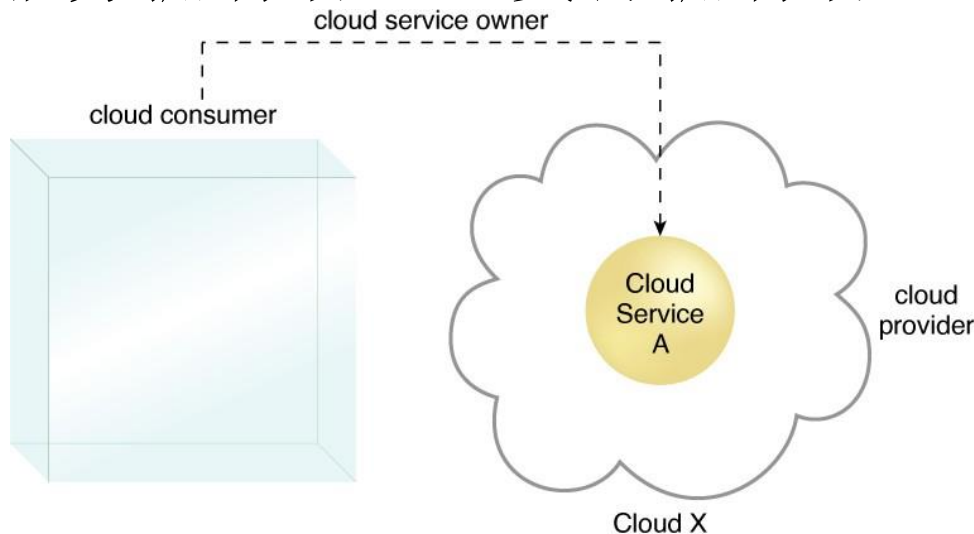
角色与组织

- 云提供者Cloud Provider
 - 提供基于云的IT资源的组织或个人就是云提供者
 - IT资源可以是云提供者自有或者租赁的
- 云用户Cloud Consumer
 - 云用户是指通过合约协定使用云的IT资源的人或者组织
 - 云用户可以通过云服务中的用户角色来访问云服务



云服务拥有者

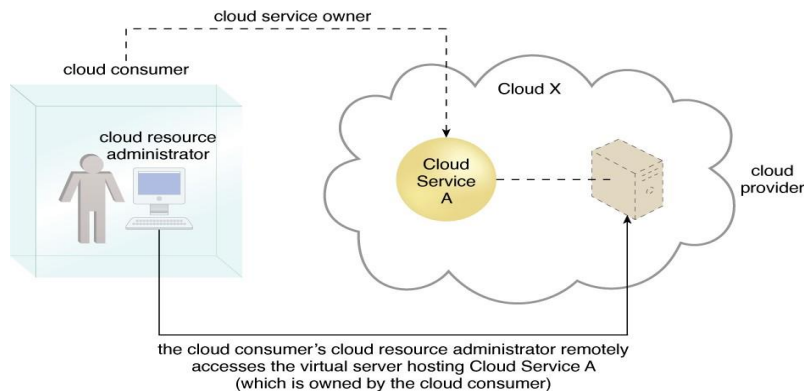
- 拥有云服务的个人或者组织称为云服务拥有者
- 云服务拥有者可以是云用户或者是云提供者本身
- 云服务拥有者 \neq 云资源拥有者



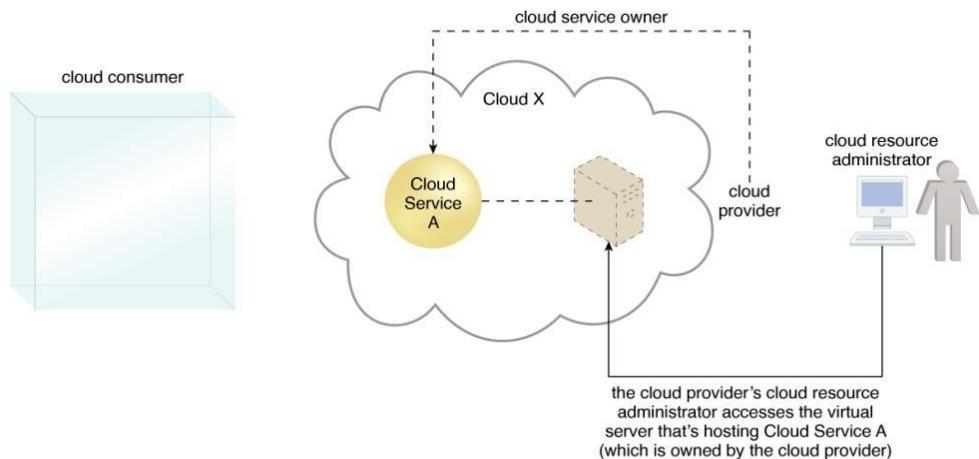
Copyright © Arcitura Education

云资源管理者

- 负责管理基于云的IT资源（包括云服务）的人或者组织。
- 云资源管理者可以是
 - 这个云的用户、
 - 云提供者或者是
 - 受委托的第三方组织

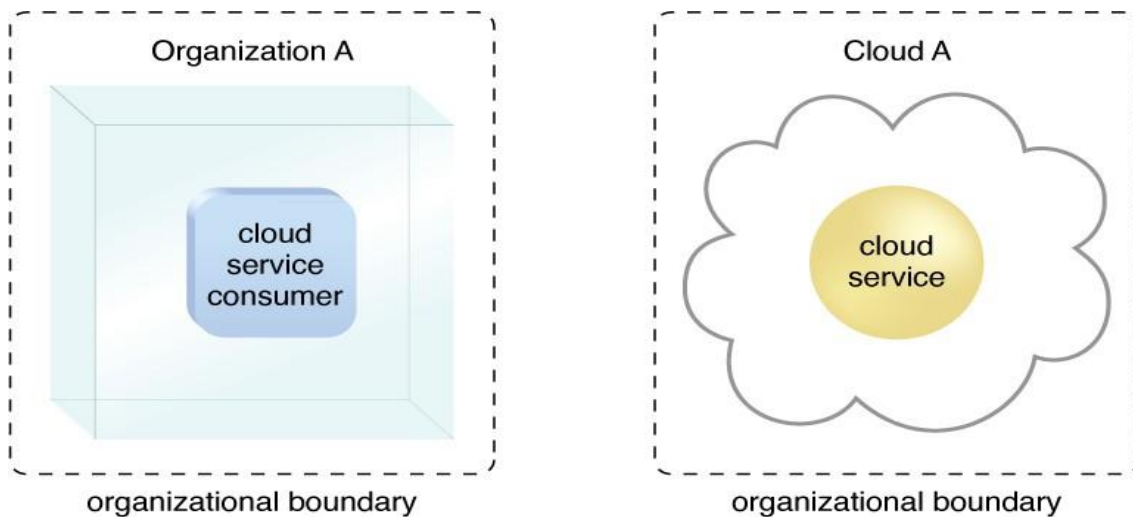


Copyright © Arcitura Education



组织边界

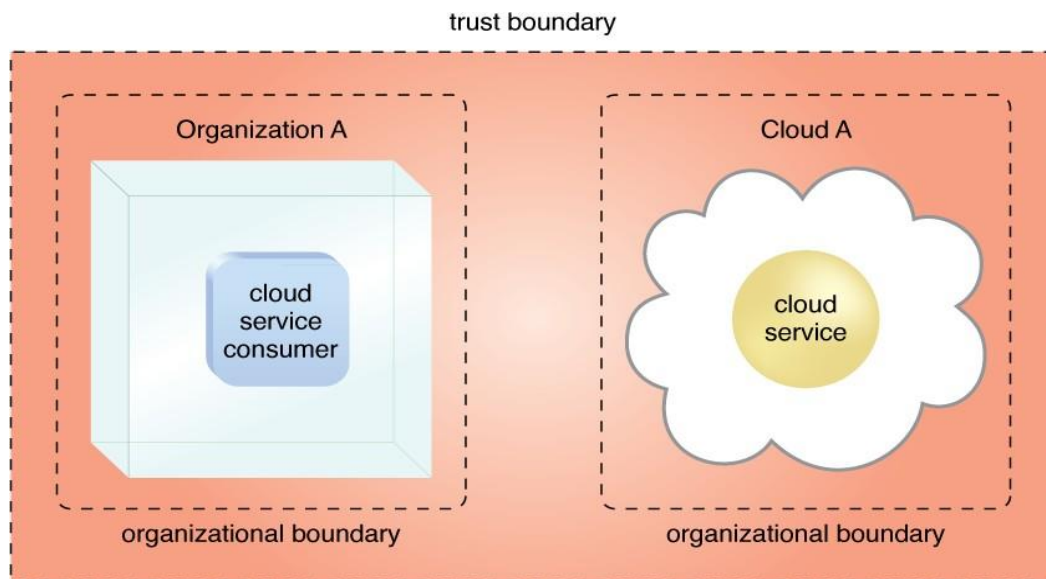
- 组织边界是一个物理范围，是一个组织拥有和管理的IT资源的集合。
- 组织边界不变时组织边界不代表组织实际的边界，只是该组织的IT 资产和IT资源。
- 云组织边界：



Copyright © Arcitura Education

信任边界

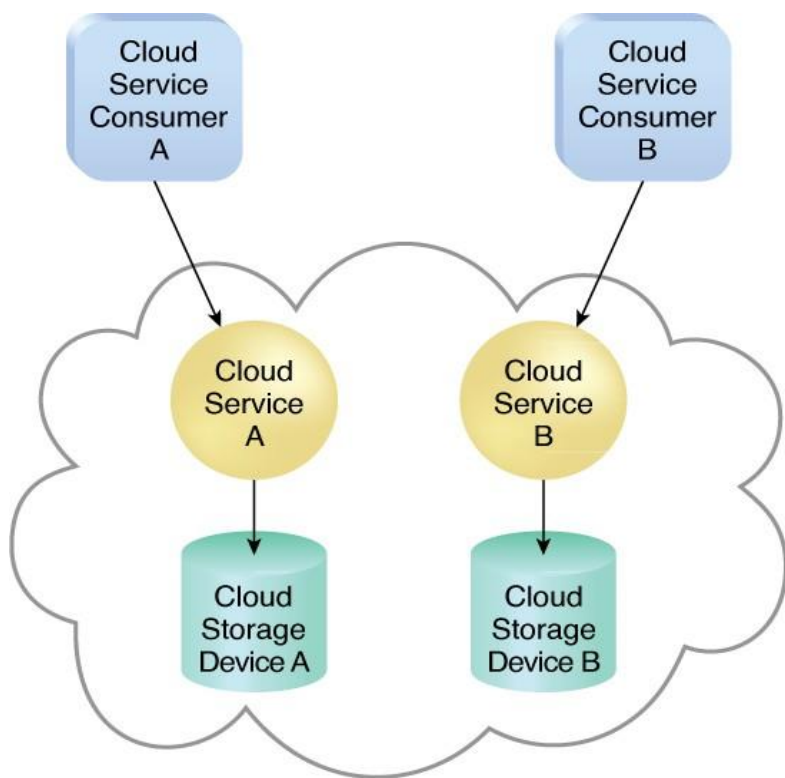
- 信任边界是一个逻辑范围，通常会跨越物理边界，表明IT资源受信任的程度。
- 云用户要访问基于云的IT资源时需要将信任扩展到该物理边界之外，把部分云环境包括进来。
- 在分析云环境的时候，信任边界最常与作为云用户的组织发出的信任关联到一起。



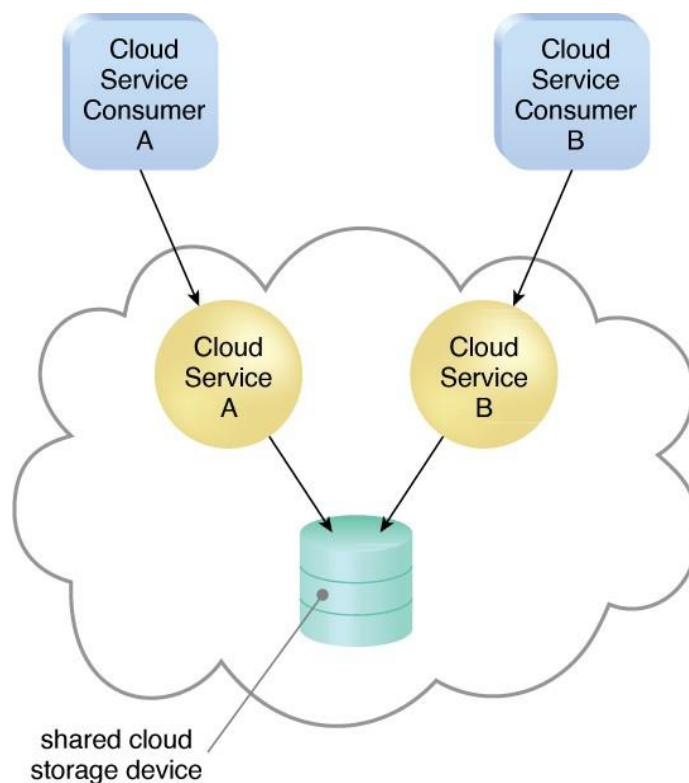
云特性

- 按需使用On-demand usage
 - 基于服务的特性和使用驱动的特性。
- 随处访问Ubiquitous usage
 - 需要支持一组设备、传输协议、接口和安全技术。
- 多租户Multitenancy (and Resource Pooling)
 - 多个云用户共享软件和实例。
 - 可以根据云服务用户的需求动态分配IT资源。
- 弹性Elasticity
 - 与降低投资和与使用比例的成本这些好处紧密地联系在一起。
- 可测量的使用Measured Usage
 - 云平台对云用户使用的IT资源使用情况的记录能力。
- 可恢复性Resiliency
 - 是一种故障转移的形式。
 - 在多个物理位置分放IT资源的冗余实现。

多租户 (multitenant)

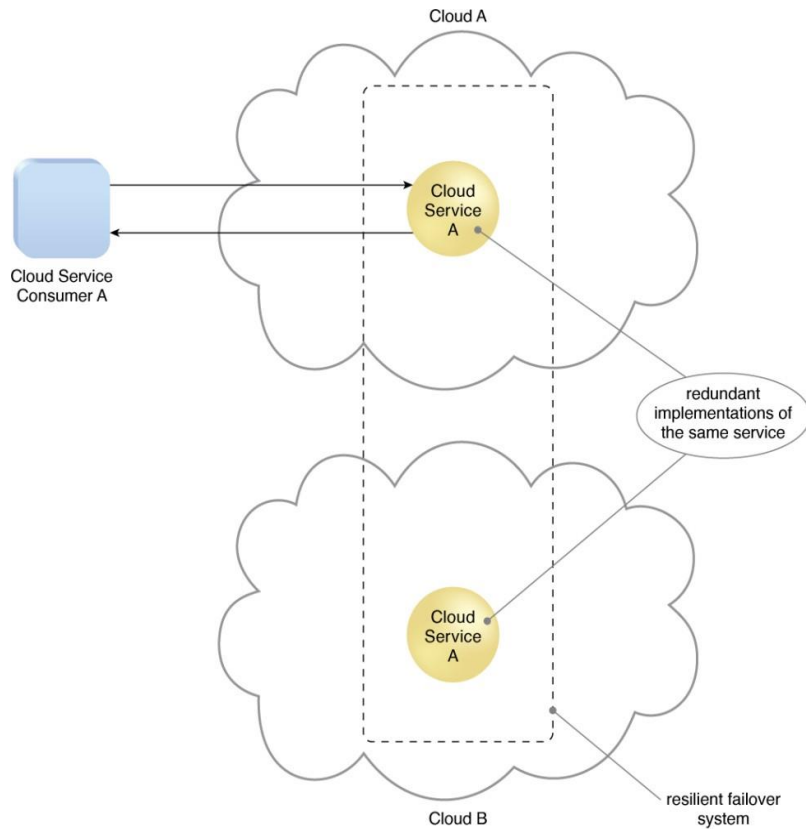


Copyright © Arcitura Education



Copyright © Arcitura Education

Resiliency



Copyright © Arcitura Education





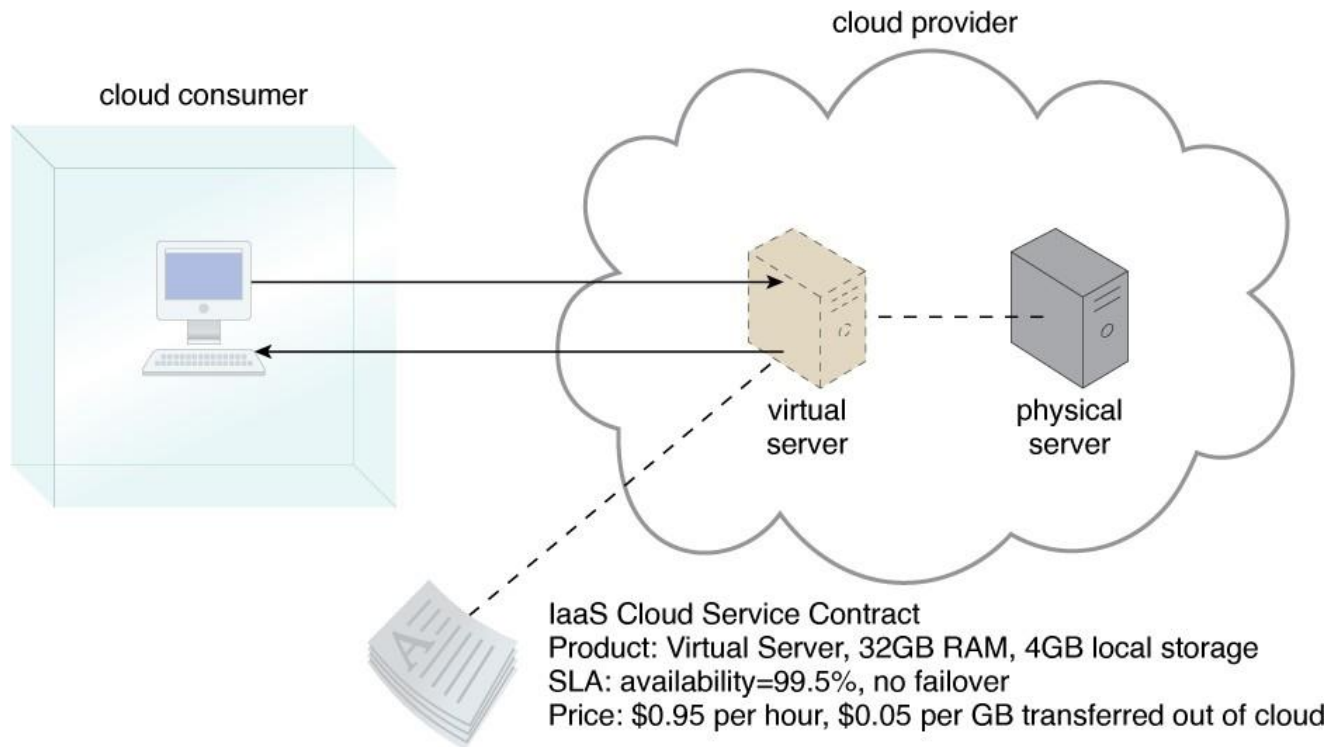
§ 1.2 架构与模型

- 交付模型：服务交付给云用户的模式和方式
- 基础设施作为服务 (IaaS)
- 平台作为服务 (PaaS)
- 软件作为服务 (SaaS)
- 由不同IT资源组合构成的这三种基本云交付模型的变种，例如：
 - 存储作为服务storage-as-a-service
 - 数据库作为服务database-as-a-service
 - 安全作为服务security-as-a-service
 - 通信作为服务communication-as-a-service
 - 集成作为服务integration-as-a-service, etc.

IaaS

- 基础设施作为服务Infrastructure-as-a-Service (IaaS)
- IaaS交付模型是一种自我包含的IT环境，由以基础设施为中心的IT资源组成，可以通过基于云服务的借口和工具访问和管理这些资源。
- IaaS 环境包括硬件、网络、连通性、操作系统以及其他一些原始的IT资源。
- 在 IaaS中, IT 资源通常是虚拟化的并打成包。

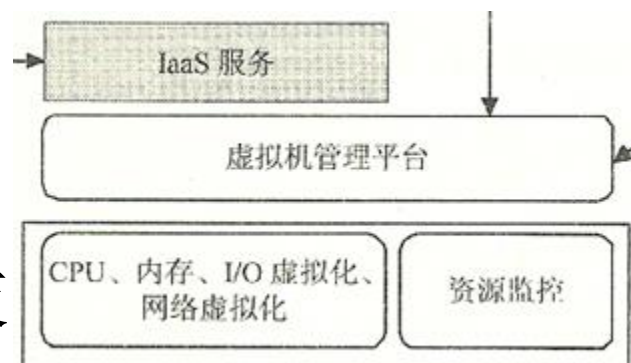
IaaS



Copyright © Arcitura Education

IaaS

- Infrastructure as a Service
- 硬件资源作为服务提供给用户
- 主要技术
 - 虚拟化技术
 - 资源动态管理与调度技术
- 典型产品
 - EC2 from Amazon
 - ECS from Aliyun



/品商



ECS

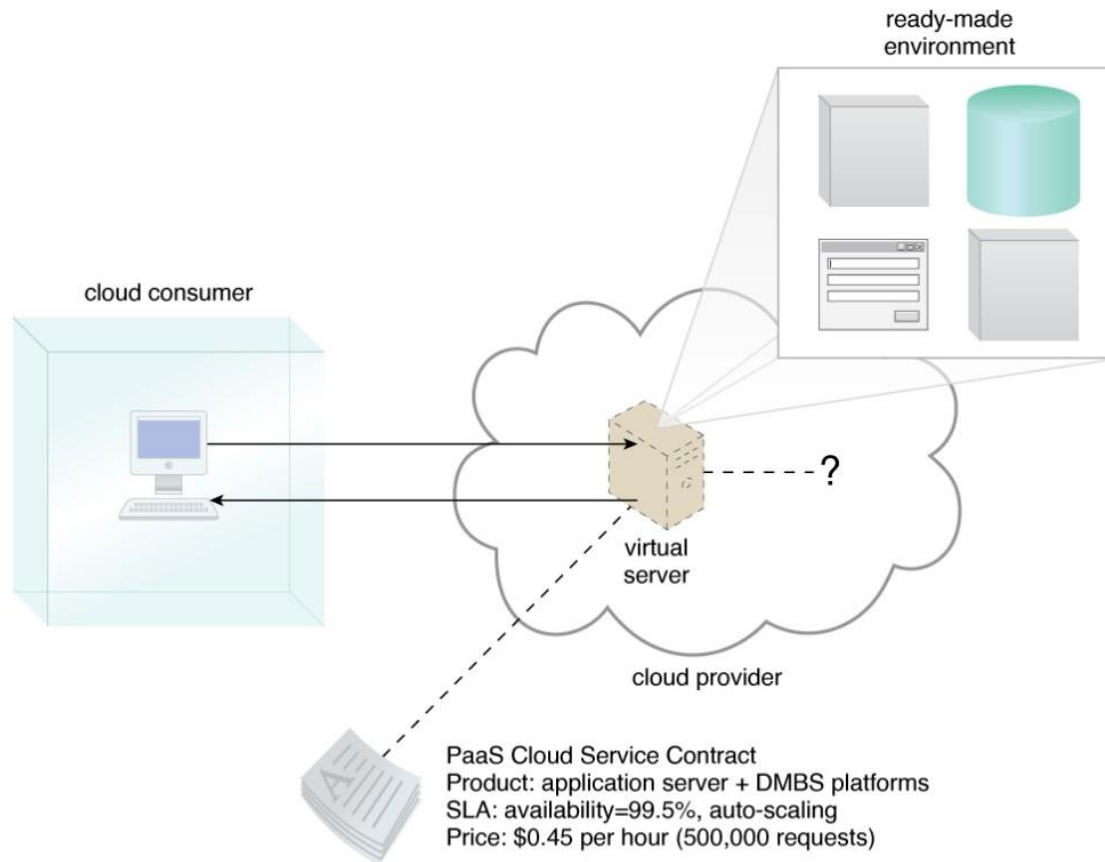
- 不同规格配置、不同计费方式
- 包月（元/月）、按量付费（元/小时）

| 规格族 | 实例规格 | 杭州 | 青岛 | 香港 | | 美国硅谷 | | 亚太（新加坡） | |
|-----|-----------|------|------|-------|------|-------|------|---------|------|
| | | 北京 | | | | | | | |
| | | 深圳 | | | | | | | |
| | | 上海 | | 非 Win | Win | 非 Win | Win | 非 Win | Win |
| n1 | 1 核 1GB | 62 | 62 | 66 | 86 | 60 | 72 | 70 | 91 |
| | 1 核 2GB | 88 | 88 | 136 | 177 | 170 | 204 | 200 | 260 |
| | 2 核 4GB | 196 | 196 | 272 | 354 | 340 | 408 | 400 | 520 |
| | 4 核 8GB | 412 | 412 | 544 | 708 | 680 | 816 | 800 | 1040 |
| | 8 核 16GB | 844 | 844 | 1088 | 1415 | 1360 | 1632 | 1600 | 2080 |
| | 16 核 32GB | 1708 | 1708 | 2176 | 2829 | 2720 | 3264 | 3200 | 4160 |
| | 32 核 64GB | 3436 | 3436 | 4352 | 5658 | 5440 | 6528 | 6400 | 8320 |

PaaS

- 平台作为服务Platform-as-a-Service (PaaS)
- PaaS交付模型是预先定义好的“就绪可用”的环境，一般由已经部署好和配置好的IT资源组成
- PaaS依赖于使用已就绪环境，设立好一套预先打包好的产品 and 用来支持定制化应用的整个交付生命周期的工具
- 例如，GAE提供了基于Java和Python的环境

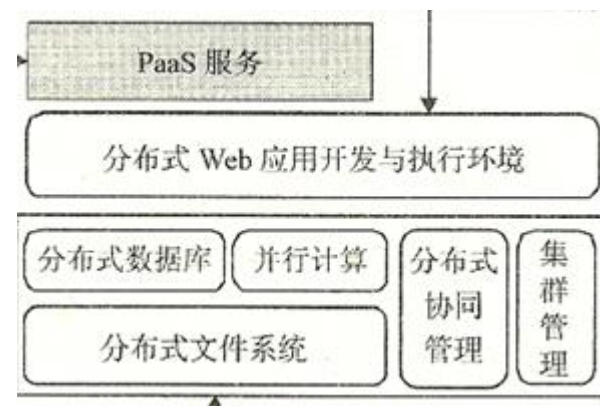
PaaS



Copyright © Arcitura Education

PaaS

- Platform as a Service
- 提供应用软件的开发、测试、部署和运行环境
 - 运行平台，如应用服务器
 - 辅助系统软件，如数据库
- 关键技术
 - 分布式文件系统
 - 分布式数据库
 - 并行计算技术
- 典型产品
 - GAE – Google
 - SAE – Sina
 - ACE – Aliyun
 - Windows Azure



Google
App Engine



ACE



新浪云计算
e.sae.sina.com.cn

中山大学吴维刚

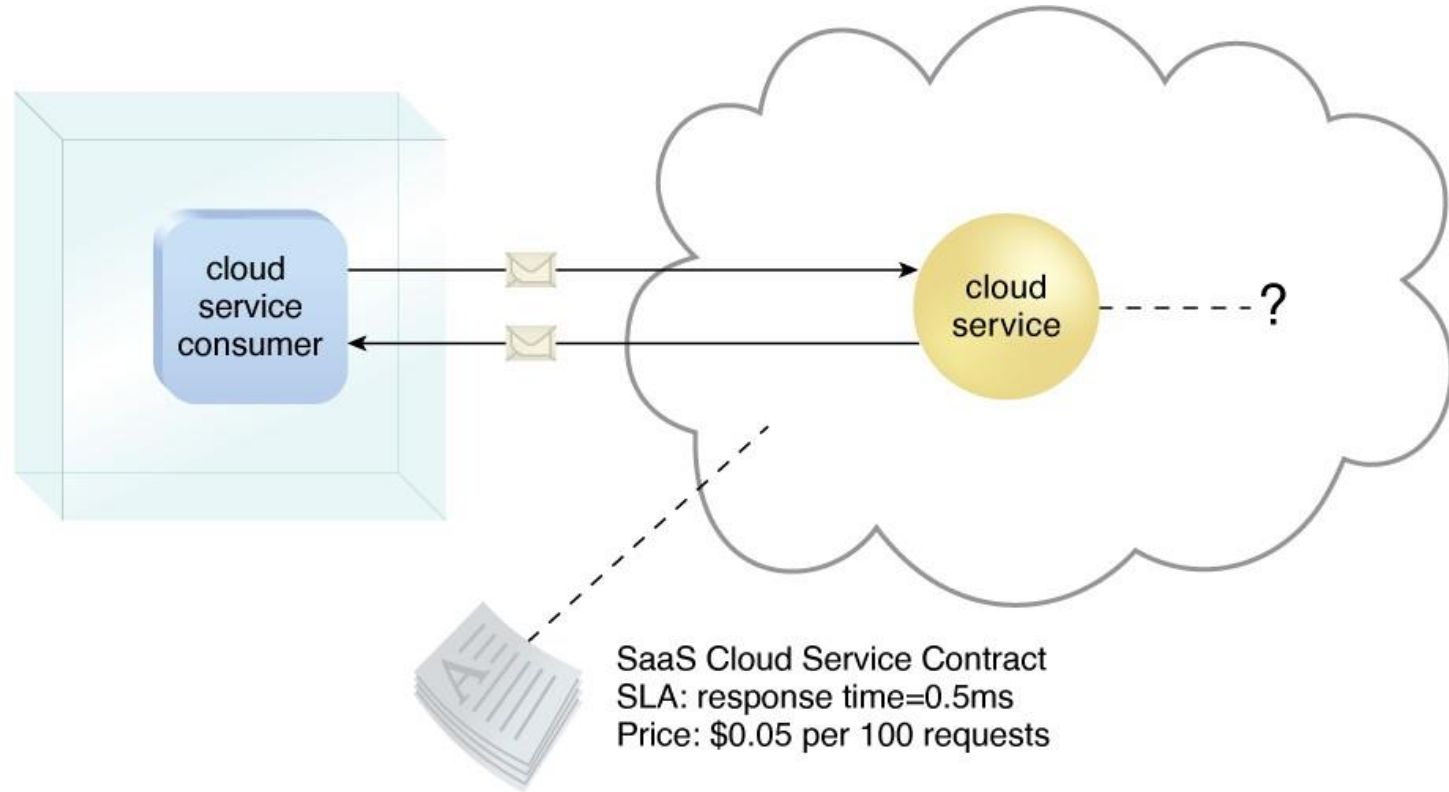


Windows Azure₇

SaaS

- 软件作为服务Service-as-a-Service (SaaS)
- SaaS通常是把软件程序定位成共享的云服务，作为“产品”或通用的工具进行提供。
- 通常，云用户对SaaS实现的管理权限非常有限。

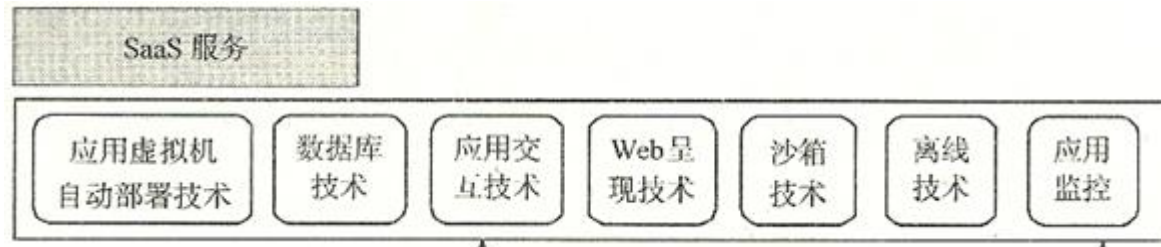
SaaS



Copyright © Arcitura Education

SaaS

- Software/Application as a Service
- 将运行于云中的应用软件的功能交付给用户。
- 关键技术
 - 呈现技术
 - 多租户技术（Multitenancy）
- 典型产品
 - Salesforce 的CRM服务



交付模型比较

| 云交付模型 | 赋予云用户的典型控制等级 | 云用户可用的典型功能 |
|-------|--------------|------------------------------------|
| SaaS | 使用和与使用相关的配置 | 前端用户接口访问 |
| PaaS | 有限的管理 | 对于云用户使用平台相关的IT资源的中等级别的控制 |
| IaaS | 完全的管理 | 对虚拟化的基础设施相关的IT资源以及可能的底层物理IT资源的完全访问 |

交付模型比较

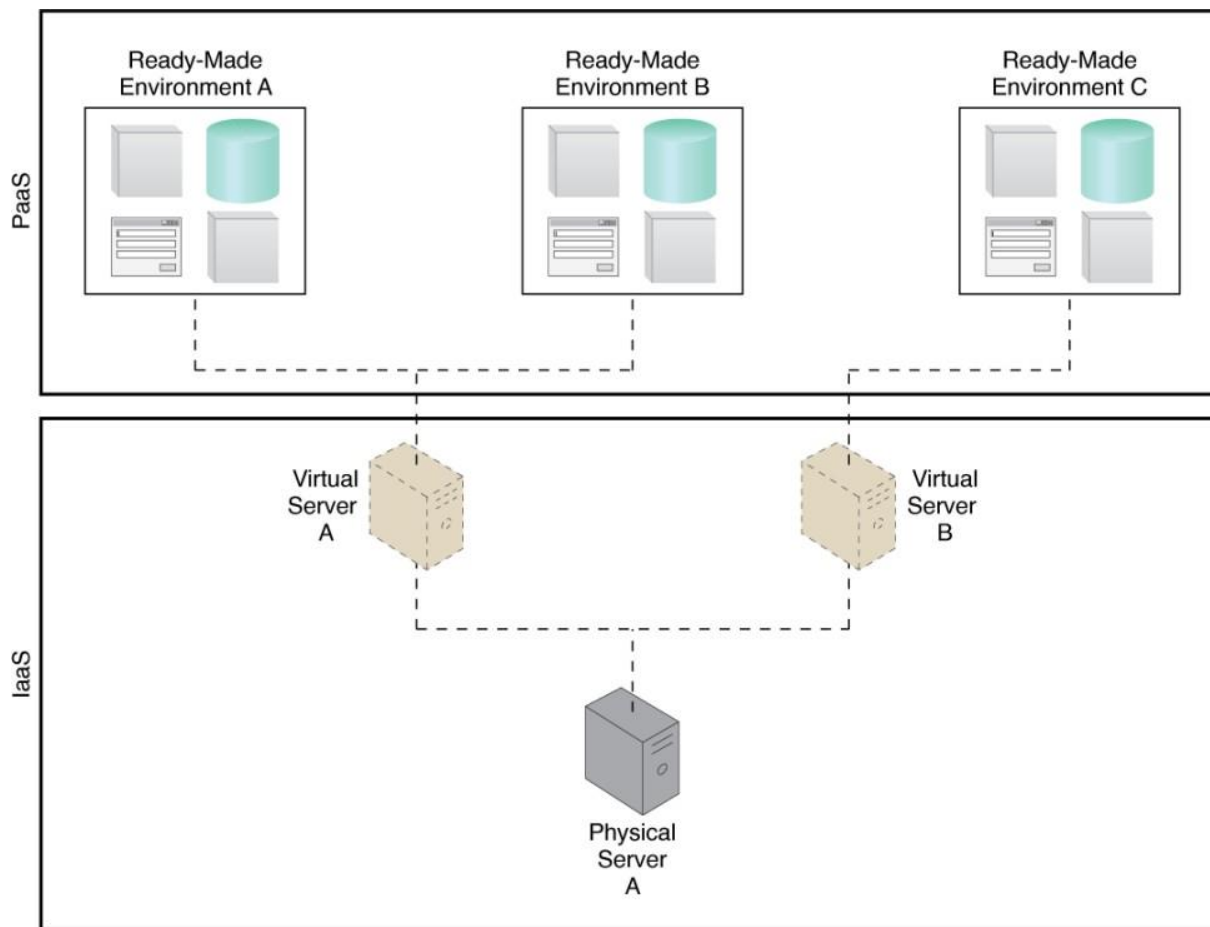
| 云交付模型 | 常见的云用户行为 | 常见的云提供者行为 |
|-------|---------------------------|--|
| SaaS | 使用和配置云服务 | 实现、管理和维护云服务 监控云用户的使用 |
| PaaS | 开发、测试、部署和管理云服务以及基于云的解决方案 | 实现配置好的平台和在需要时提供底层的基础设施、中间件和其他所需的IT资源 监控云用户的使用 |
| IaaS | 建立和配置裸的基础设施、安装、管理和监控所需的软件 | 提供和管理需要的物理处理器、存储、网络 和托管监控云用户的使用 |



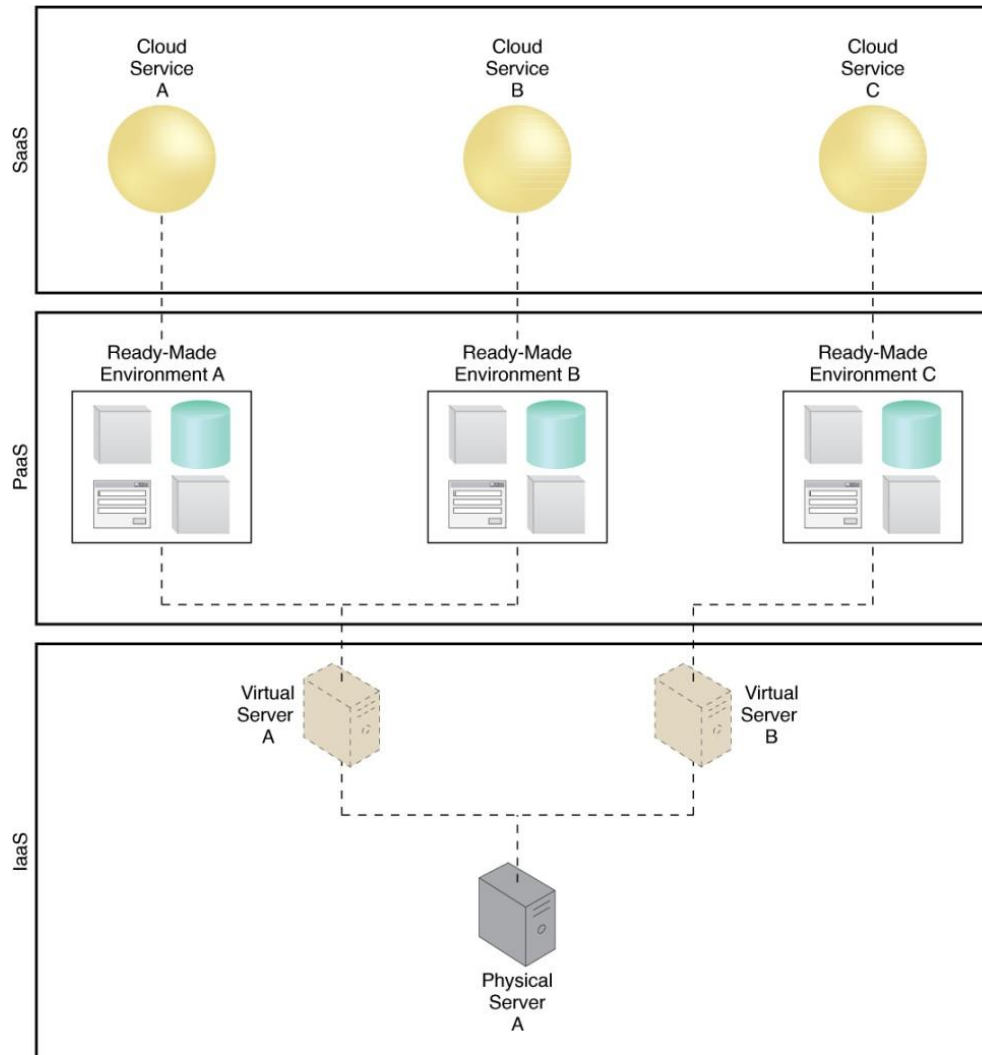
交付模型组合

- Combining Cloud Delivery Models
- IaaS + PaaS
 - 一个提供PaaS环境的云提供者选择租用另一个云提供者的IaaS环境
- IaaS + PaaS + SaaS
 - 云用户可以利用PaaS环境提供的已就绪环境来开发和部署他自己的SaaS云服务
- PaaS + SaaS?!
- IaaS + SaaS?!

IaaS + PaaS



IaaS + PaaS + SaaS

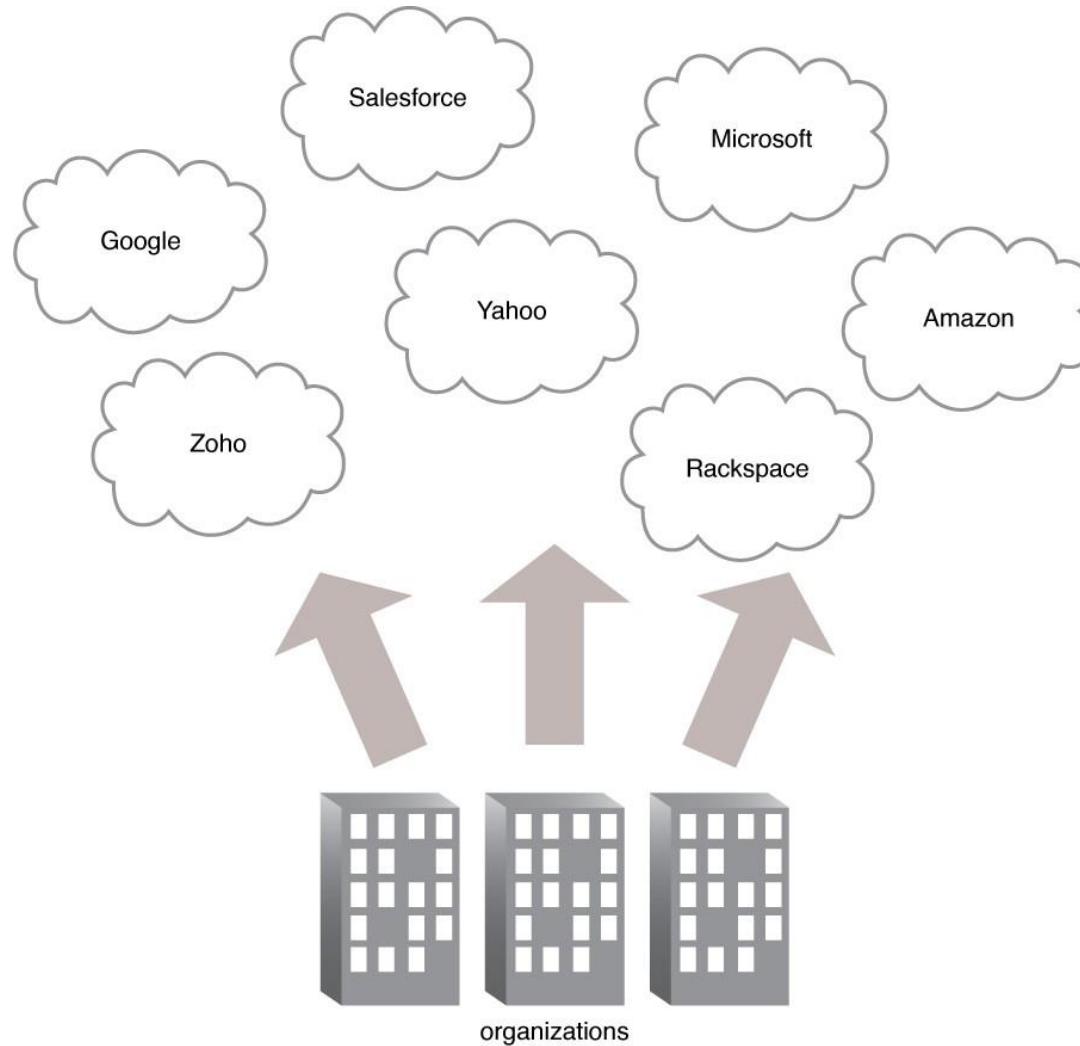




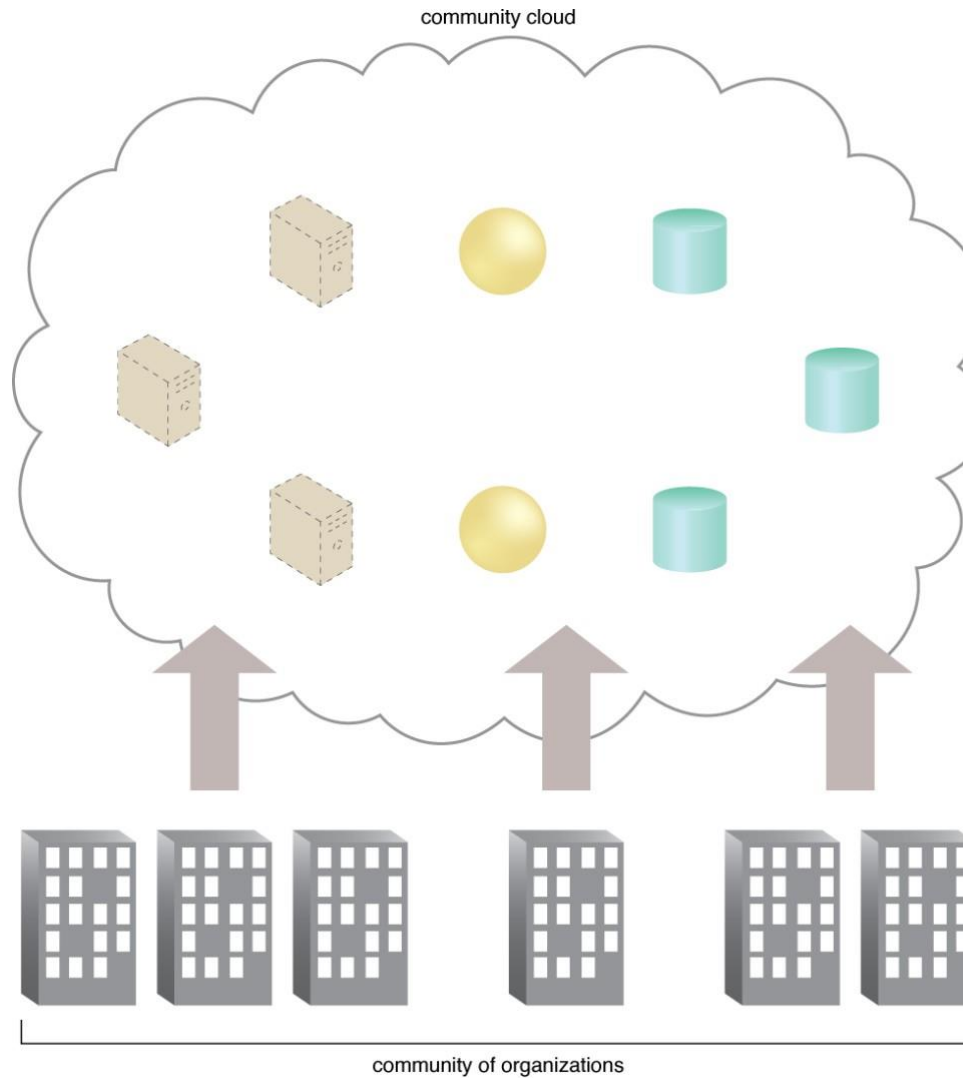
云部署模型

- 公有云Public (Figure 4.17)
- 社区云Community (Figure 4.18)
- 私有云Private (Figure 4.19)
- 混合云Hybrid (Figure 4.20)

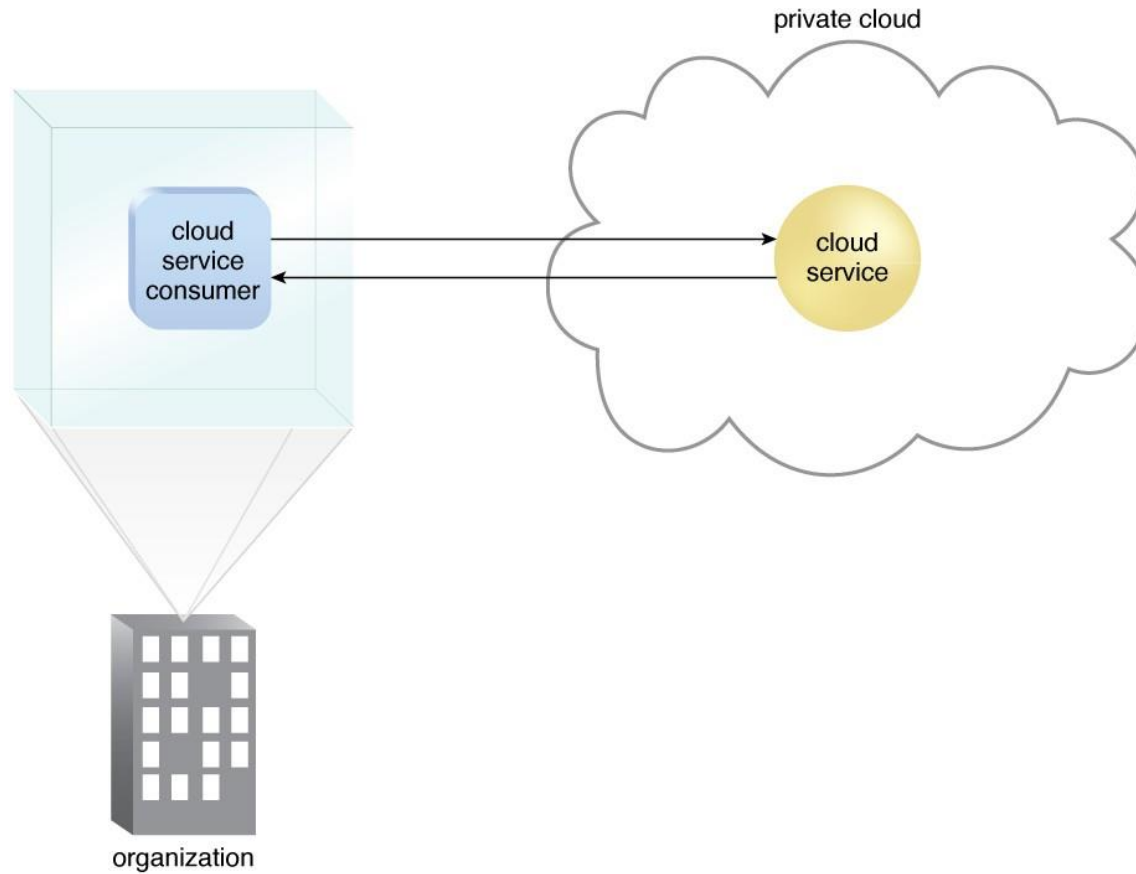
Public Cloud



Community Cloud

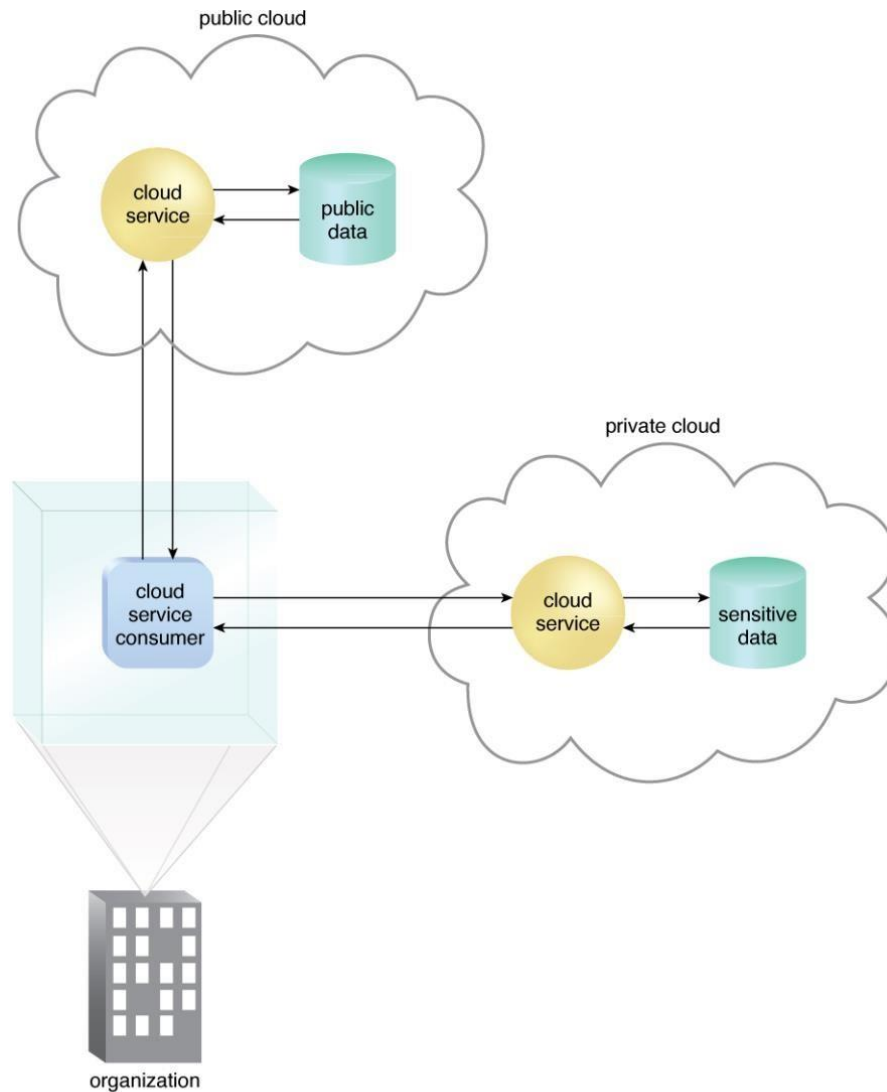


Private Cloud



Copyright © Arcitura Education

Hybrid Cloud





本章小结

- 云计算定义
- 技术特征
- 角色与边界
 - 用户、提供者、管理者...
- 云特性
 - 弹性资源利用、按需分配、高可靠高可用...
- 云交付模型
 - IaaS、PaaS、SaaS及组合
- 云部署模型
 - 公有、私有、混合、社区云