

第五章云使能技术

§ 5.1 宽带网络和Internet架构

§ 5.2 数据中心技术

§ 5.3 虚拟化技术

§ 5.4 Web技术

§ 5.5 多租户技术

§ 5.6 服务技术

§ 5.7 云计算数据处理架构

§ 5.8 分布式文件系统

§ 5.9 分布式数据库

§ 5.10 MapReduce技术

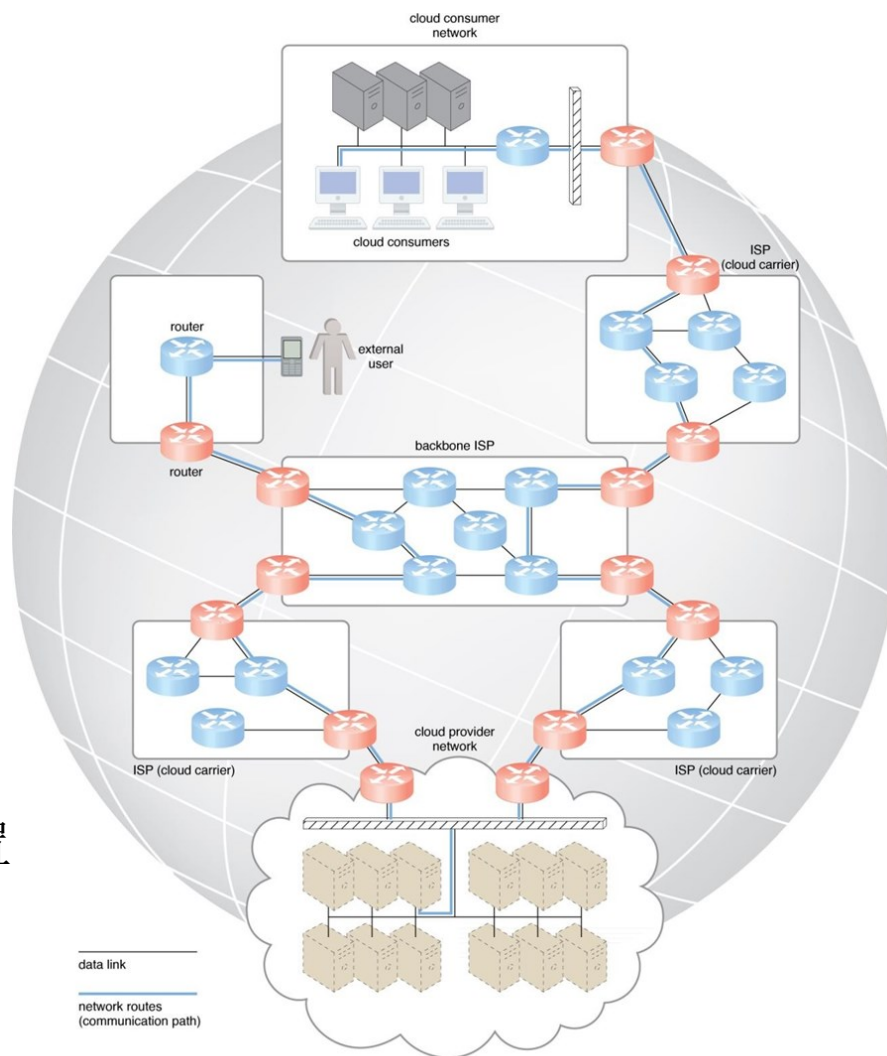


§ 5.1 宽带网络和Internet架构

- Internet服务提供者（ISP）
 - 无连接分组交换(connectionless packet switching)（数据报网络）
 - 基于路由器的互联(router-based interconnectivity)
- 技术和商业考量
 - 连接性问题
 - 网络带宽和延迟问题
 - 云运营商和云提供者选择



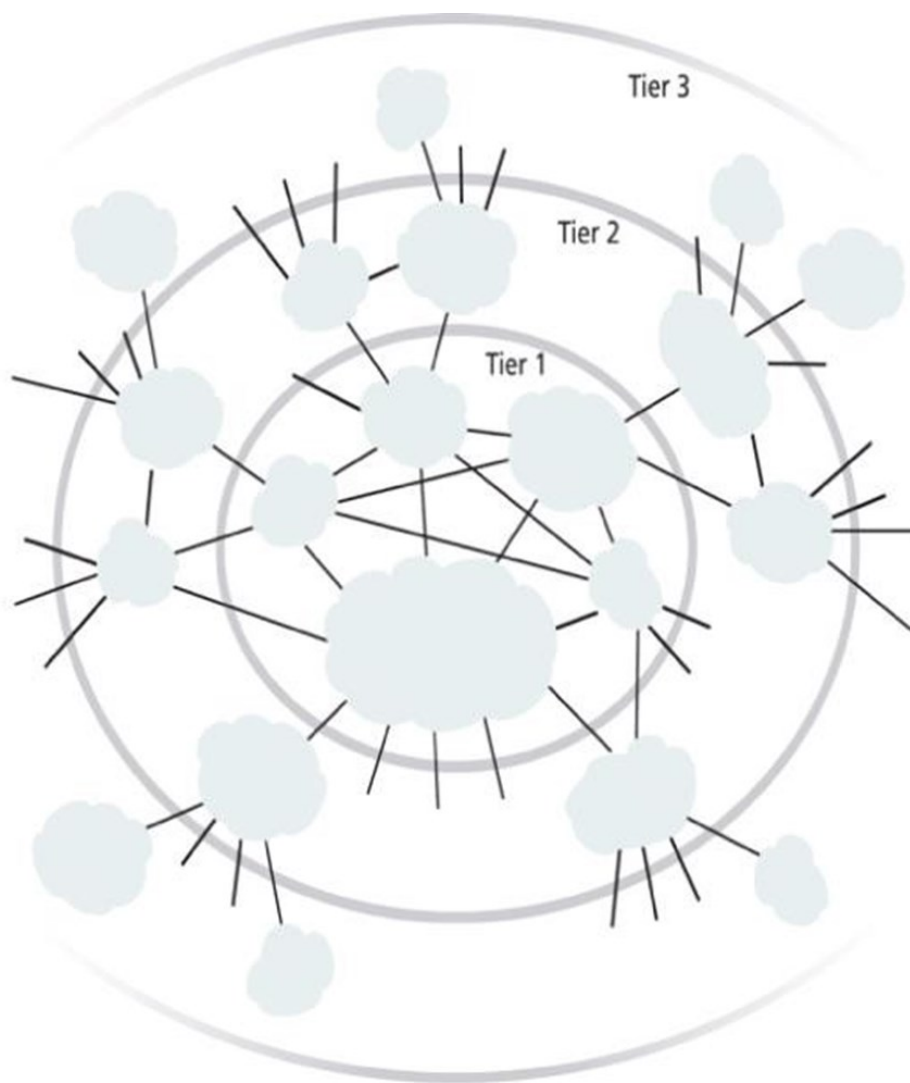
Internet服务提供者 (ISP)



一个ISP互联网络配置

Internet服务提供者 (ISP)

Internet互联结构的抽象示意图



网络互联的基本机制

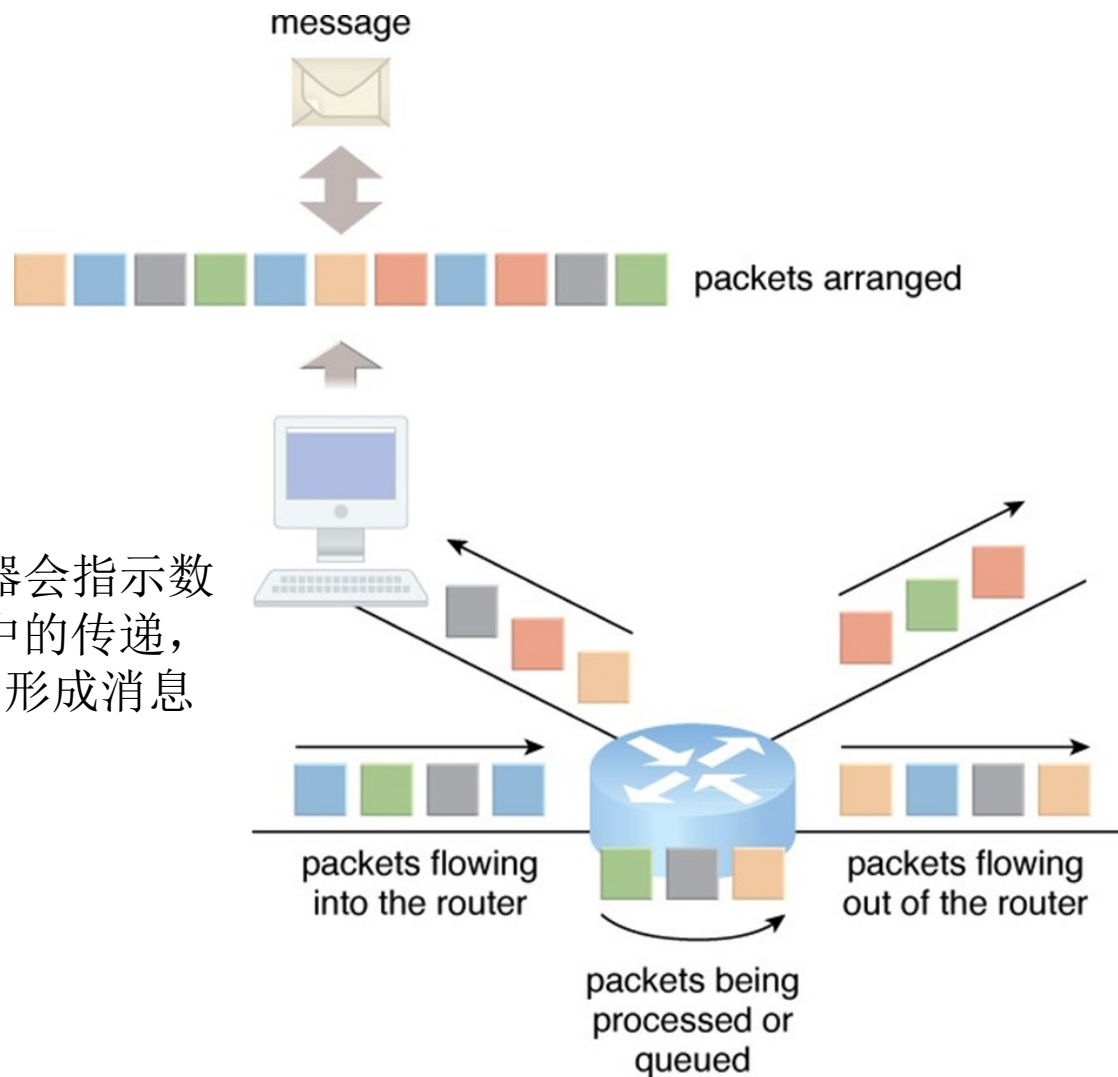
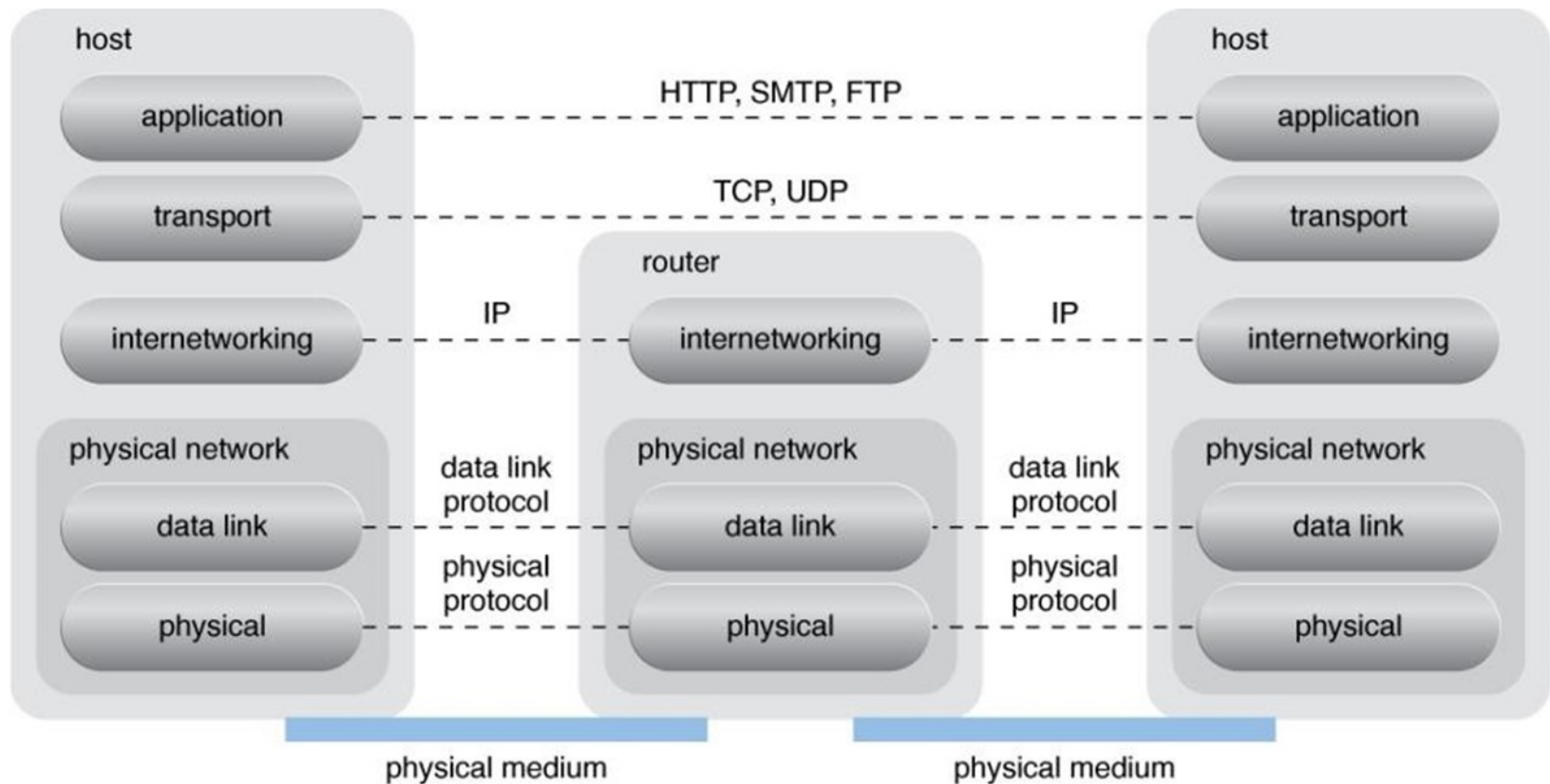


Figure 5.3 – 路由器会指示数据包在Internet中的传递，还会把它们排列形成消息

Internet参考模型和协议栈

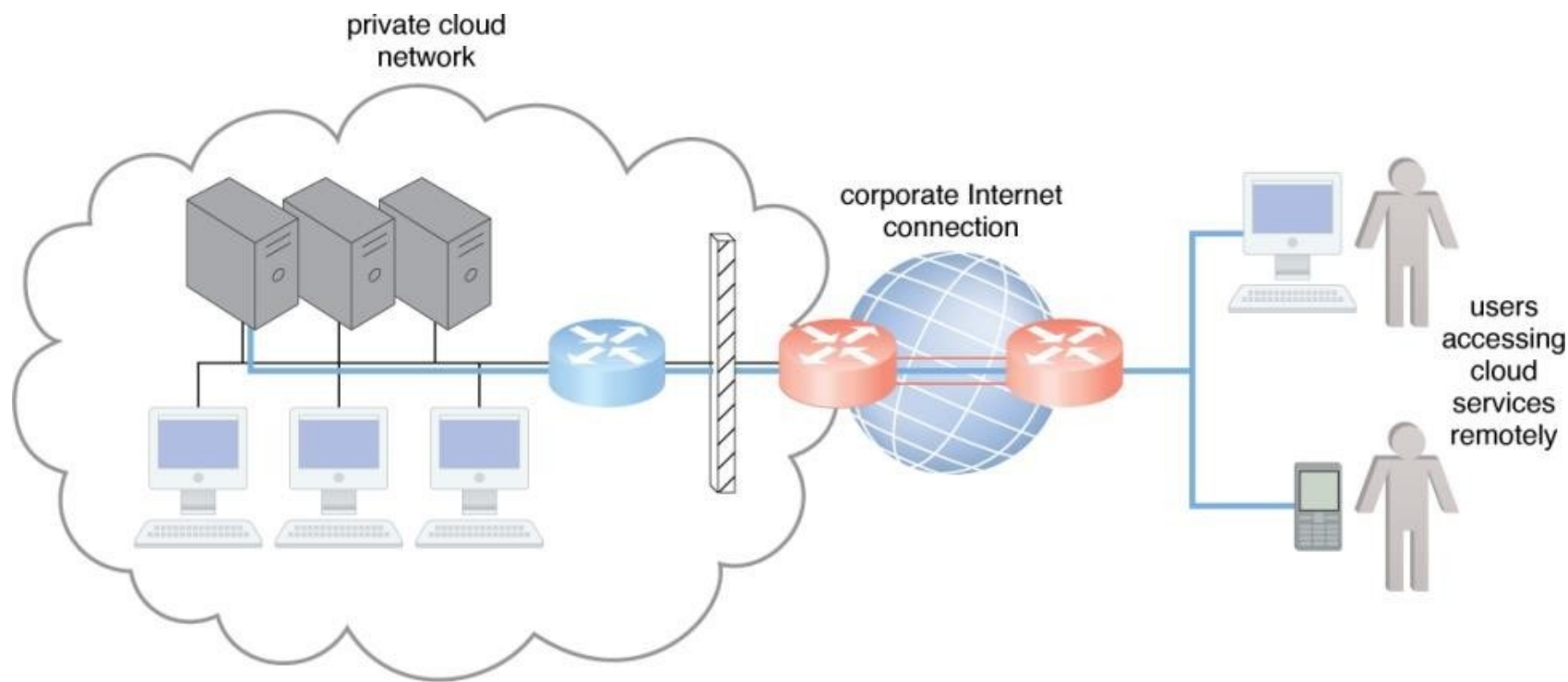


Copyright © Arcitura Education

Figure 5.4 – 通用Internet模型和协议栈示意图



连接性问题



Copyright © Arcitura Education

Figure 5.5 – 一个私有云的网络互联架构。
构成云的物理IT资源位于组织内部，并在其内进行管理



连接性问题

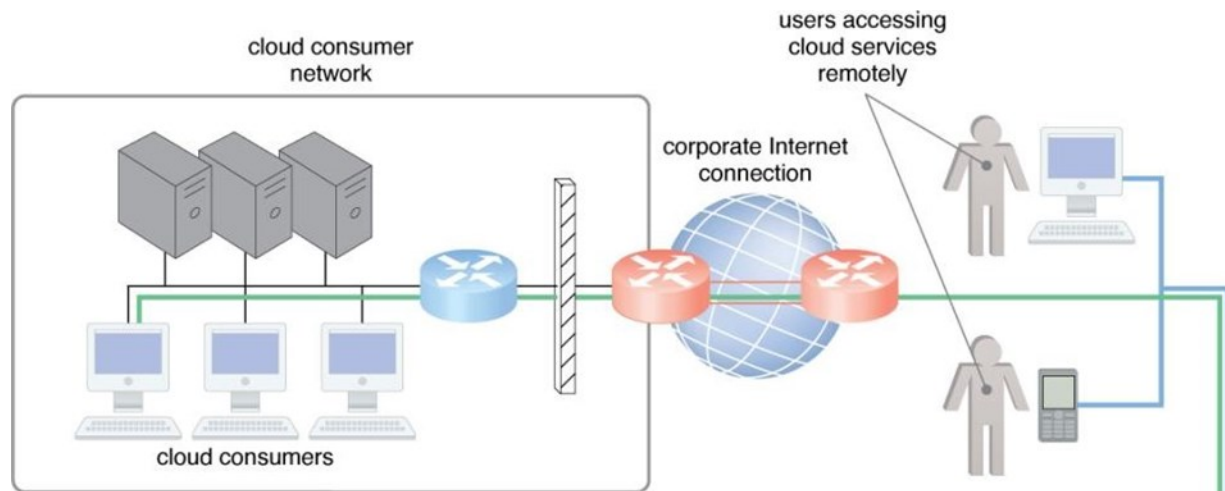
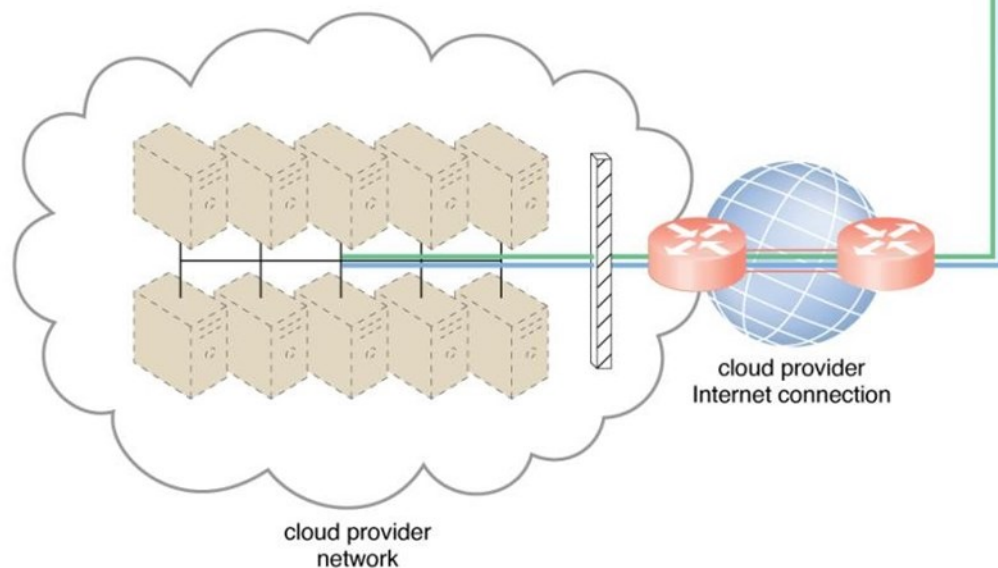


Figure 5.6 – 基于Internet的云部署模型的网络互联架构。
Internet是不直接相连的云用户、漫游终端用户和云提供者之间的网络连接代理



企业内部和基于云的网络互联的比较

企业内部IT资源	云IT资源
内部终端用户通过 企业网络 访问企业IT服务	内部终端用户通过 Internet连接 访问企业IT服务
内部用户在外网漫游时，通过 企业Internet连接 访问企业IT服务	内部用户在外网漫游时，通过 云提供者的Internet连接 访问企业IT服务
外部用户通过 企业Internet连接 访问企业IT服务	外部用户通过 云提供者的Internet连接 访问企业IT服务



关键点小结

- 云用户和云提供者通常利用**Internet**进行通信。**Internet**以**无中心的供给和管理模型**为基础，不受任何集中式实体的控制。
- 网络互联架构的主要组件是使用网络路由器和交换机的**无连接分组交换与基于路由器的互联**。
- **网络带宽和延迟**是影响**QoS**的因素，而网络拥塞对其有巨大影响。



§ 5.2 数据中心技术

○ 数据中心基本资源

- 云化的计算集群

- 硬件

- 计算设备（标准化商用服务器）

- 存储设备（磁盘阵列和存储虚拟化）

- 网络设备（高速互联网络：以太网、光网等）

- 配套设备（电源、散热.....）

- 软件

- 操作系统

- 系统软件

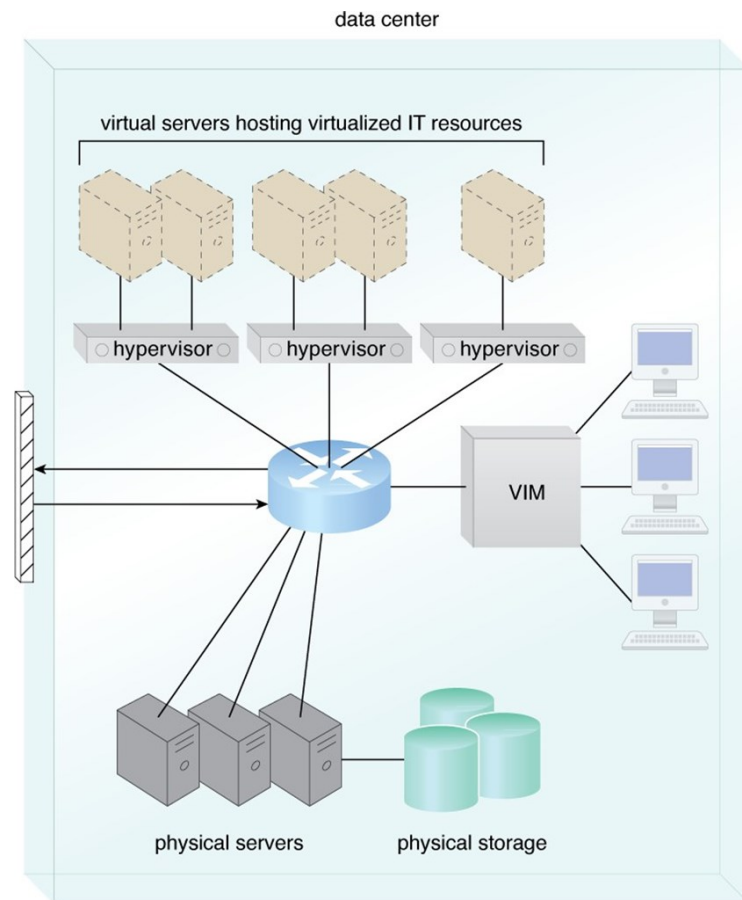
- 应用软件

-



数据中心基本技术

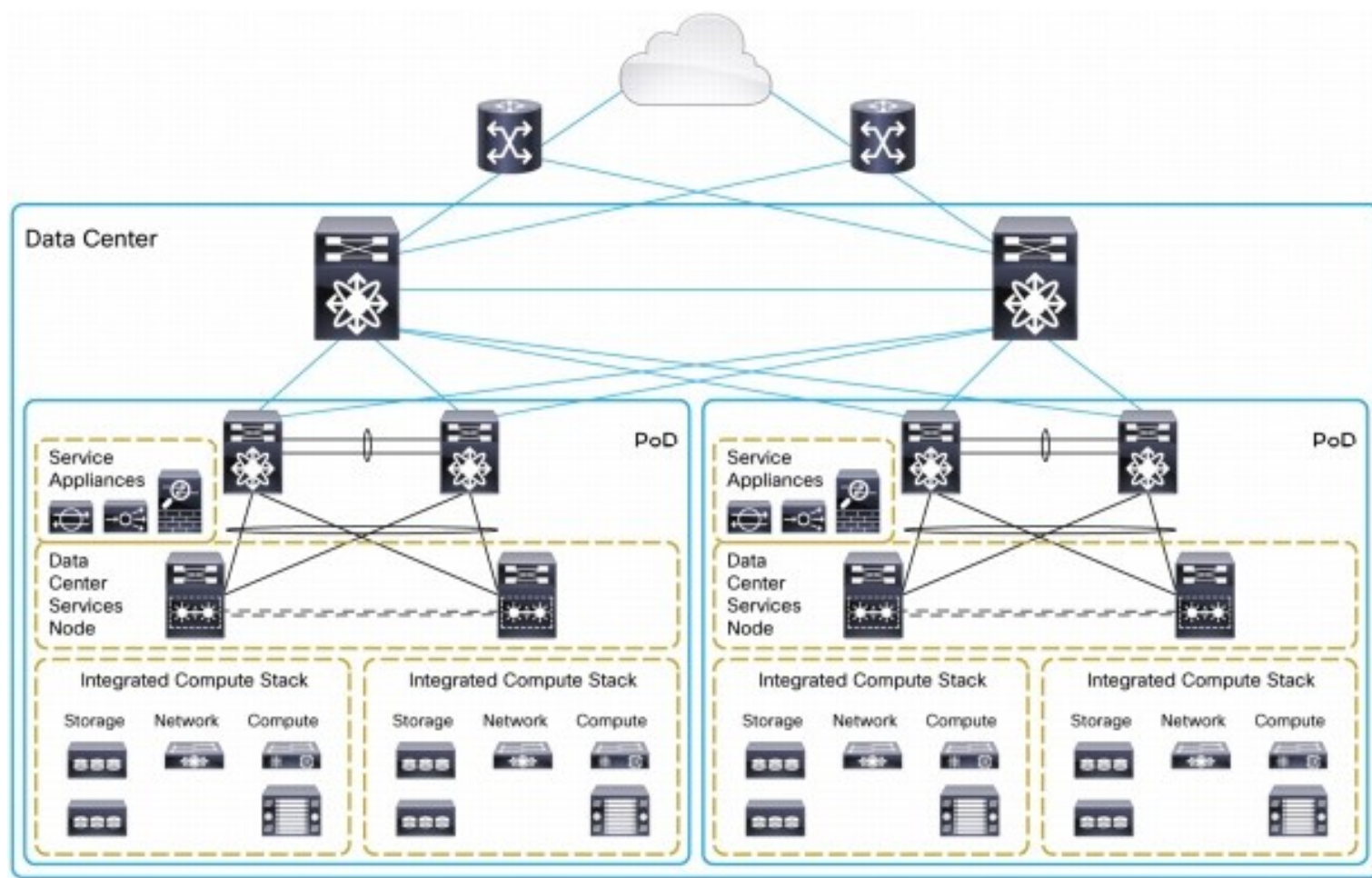
- 网络互连技术
- 虚拟化技术
- 标准化与模块化
- 自动化（自配置和自恢复）
- 远程操作与管理
- 高可用性（采用高冗余度）



Copyright © Arcitura Education



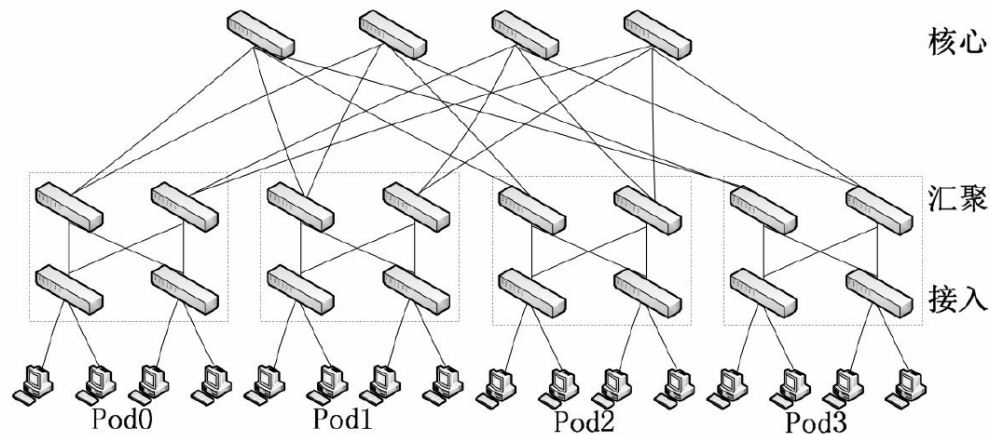
数据中心基本架构





数据中心网络架构

- 改进型胖树网络架构 (FatTree)
- ✓ 三层网络：核心交换机、汇聚交换机、接入交换机
- ✓ 连接方法：
 - ◆ 每台核心交换机有 k 个端口，每个端口与一个Pod相连，共 k 个Pod
 - ◆ 每个Pod有 $k/2$ 个汇聚交换机和 $k/2$ 个接入交换
 - ◆ 每个汇聚交换机有 $k/2$ 个端口连接各个核心， $k/2$ 个端口连接接入
 - ✓ 因此需要 $(k/2) * (k/2)$ 个核心交换机
 - ◆ 每个接入交换机有 $k/2$ 个端口连接同Pod的各个汇聚， $k/2$ 个端口连接主机

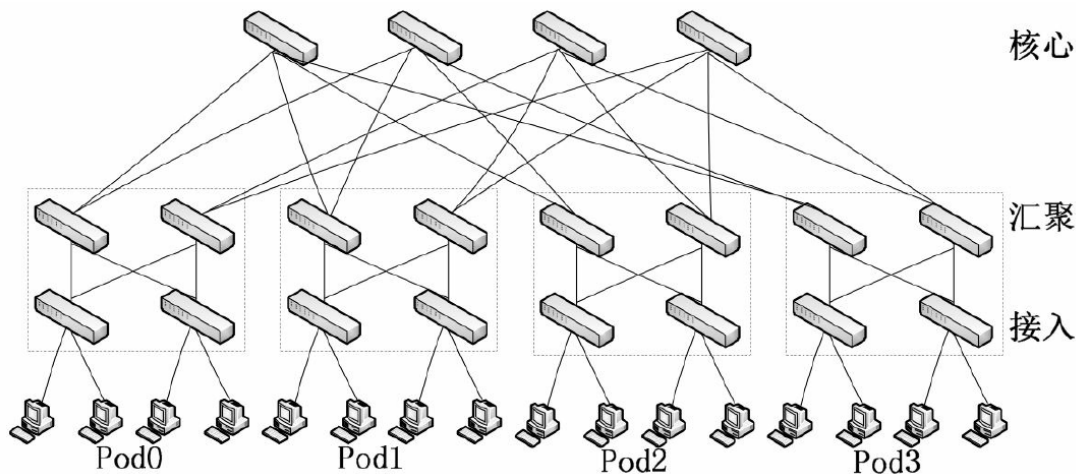




数据中心网络架构

■ 改进型胖树网络架构优缺点分析

- ✓ 每台核心交换机与每个Pod相连
 - ◆ 只要一台核心交换机正常，整个网络就不会断开
- ✓ 同一Pod内，每个汇聚交换机与所有接入交换机连通
 - ◆ 同一Pod内的流量内部消化
- ✓ 扩展性
 - ◆ 最大规模为 $k * (k/2) * (k/2)$ ，受限于交换机的端口数
 - ◆ 即：Pod数 * 每个Pod内的接入交换机数 * 每个接入用来连接主机的端口数

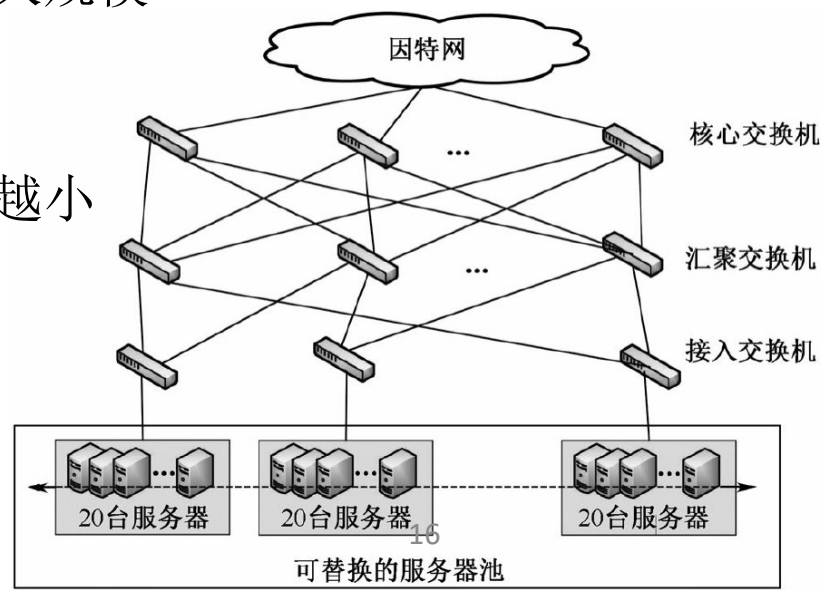




数据中心网络架构

■ 微软网络架构（VL2）

- ✓ 三层网络：核心交换机、汇聚交换机、接入交换机
- ✓ 连接方法：
 - ◆ 若干台主机（例如20台）连接到一个接入交换机
 - ◆ 每台接入与两个汇聚相连
 - ◆ 每台汇聚与所有核心相连：核心交换机端口数限制网络规模
- ✓ 优缺点分析
 - ◆ 节省汇聚和接入的端口用来扩大规模
 - ◆ 假定交换机端口数 k ， n 个核心
 - ✓ 则规模为 $k * (k - n) / 2 * (k - 2)$
 - ◆ 核心越多，健壮性越高，规模越小



数据中心网络架构

○ 递归层次结构之Dcell

○ 连接方法：

- 采用递归方法构建网络第0层由一个交换机连接n个服务器
- 第1层由n+1个第0层的节点构成：为什么是n+1？
 - 对于一个特定的0层，它的n台服务器分别与其他n个0层中的一台服务器相连
- 递归关系总结
- 第k层服务器数为 S_k ： $S_k = (S_{k-1})S_{k-1}$

○ 优缺点分析

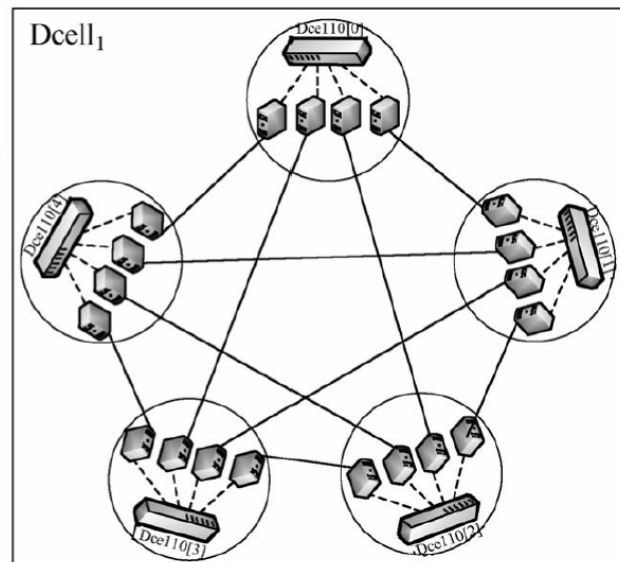
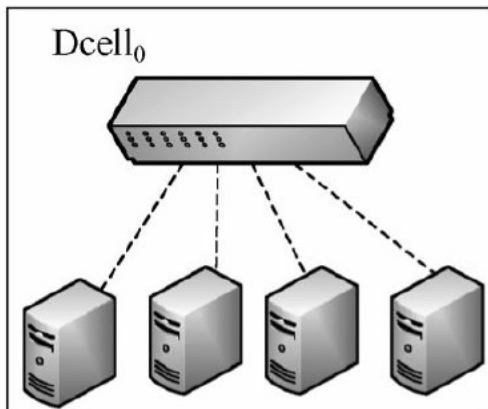
- 布线复杂
- 层数受限于端口数
- 扩展性好

第0层：4节点

第1层：4*5=20节点

第2层：20*21=420节点

第3层：420*421=176820



数据中心网络架构

■ 递归层次结构之FiConn

✓ 连接方法：采用递归方法构建网络

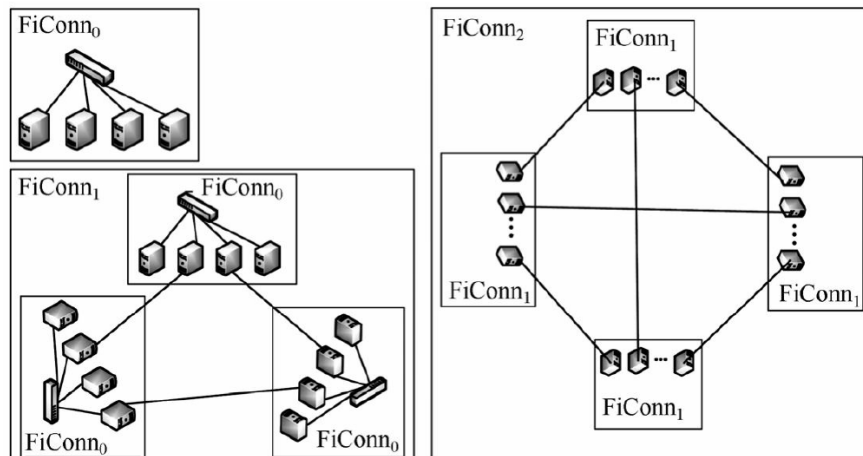
- ◆ 第0层由一个交换机连接 n 个服务器，每个服务器除了跟该交换机连接外，还有一个备用端口待用
- ◆ 第1层构建：对于每个 $FiConn_0$ ，拿出其中一半还没使用的备用端口，与其它的 $FiConn_0$ 连接
- ◆ 第2层构建：对于每个 $FiConn_1$ ，拿出其中一半还没使用的备用端口，与其它的 $FiConn_1$ 连接
- ◆ 递归关系总结
 - ✓ 第 k 层服务器数为 S_k ，空闲备用端口 B_k
 - ✓ 则 $S_{k+1} = S_k(\frac{B_k}{2} + 1)$ ， $B_{k+1} = \frac{B_k}{2}(\frac{B_k}{2} + 1)$

第0层：4节点， $B_0=4$

第1层： $4 * (2+1) = 12$ 节点， $B_1=6$

第2层： $12 * (3+1) = 48$ 节点， $B_2=12$

第3层： $48 * (6+1) = 336$ 节点， $B_3=42$



数据中心网络架构

■ 递归层次结构之BCube

✓ 连接方法：采用递归方法构建网络

◆ 第0层由一个交换机连接n个服务器

◆ 第1层构建：n个n端口交换机与n个0层相连

◆ 第k层由n个k-1层通过 n^k 个n端口交换机组成

◆ 递归关系总结

✓ 第k+1层由n个第k层组成， $S_{k+1} = nS_k$ ， $S_k = n^{k+1}$

◆ 第k层为什么需要 n^k 交换机

✓ 因为n个k-1层的所有服务器、一共 n^{k+1} 个，都要与本层的交换机相连，每个交换机n个端口

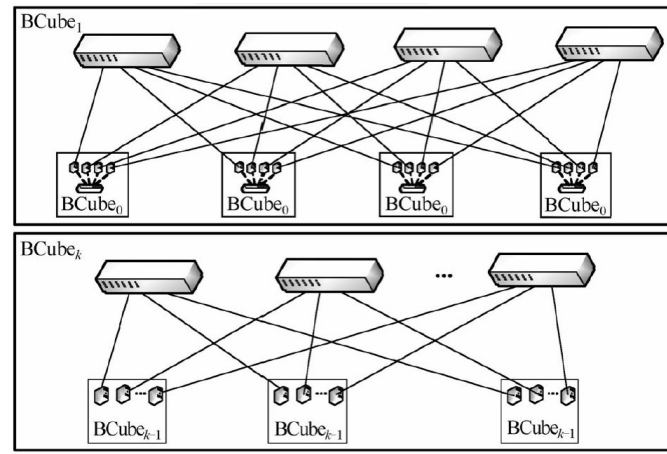
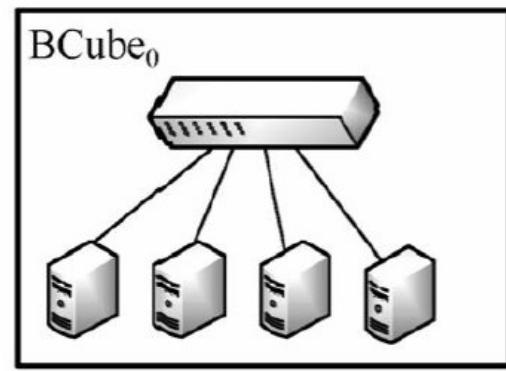
✓ 优缺点分析

◆ 同时利用服务器和交换机的转发

◆ 链路资源丰富，冗余度高

◆ 但是，每个服务器需要k+1的端口

✓ 因为每个服务器需要与每层交换机相连



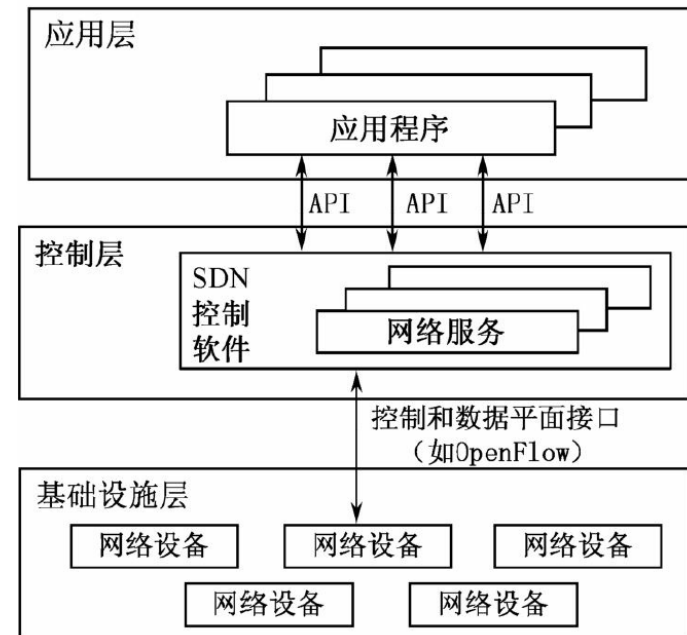


数据中心网络架构

- 软件定义网络（SDN, Software Defined Networking）
 - ✓ 一种新型的网络技术
 - ✓ 将网络的控制与数据转发进行分离
 - ✓ 网络智能地被抽取到一个集中式的控制器中，数据流的接入、路由等都由控制器来控制
 - ✓ 交换机只是按控制器所设定的规则进行数据分组的转发
 - ✓ 通过开放可编程软件模式来实现网络的自动化控制功能

- 三个层

- 基础设施层
 - 表示网络的底层转发设备
- 控制层
 - 集中维护网络状态
 - 通过南向接口获取底层基础设施信息
 - 同时为应用层提供可扩展的北向接口
- 应用层





数据中心网络架构

■ SDN主要优势

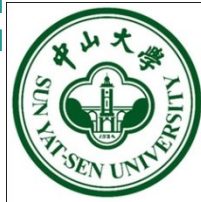
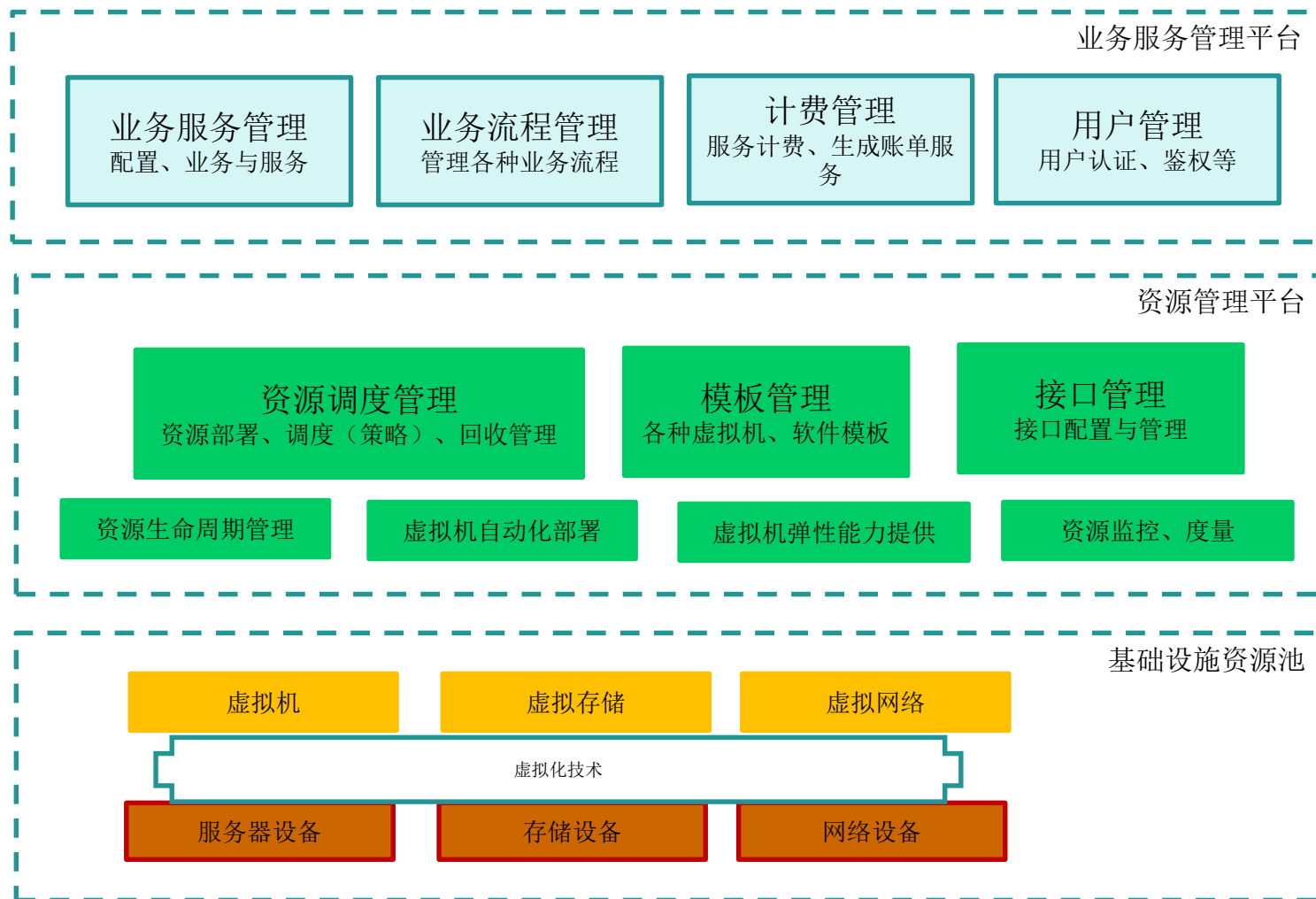
- ✓ 网络的运行维护仅需要通过软件的更新来实现网络功能的升级，网络管理者无须再针对每一个硬件设备进行配置或者等待网络设备厂商硬件的发布，从而加速了网络部署周期（软件实现硬件功能）
- ✓ SDN降低了网络复杂度，使得网络设备从封闭走向开放，底层的网络设备能够专注于数据转发而使功能简化，有效降低网络构建成本（网络设备功能简化）
- ✓ SDN通过软件来实现集中控制，使得网络具备集中协调点，因而能够通过软件形式达到最优性能（集中控制）

■ 典型标准

- ✓ OpenFlow



数据中心管理软件层次



自动化部署

○ 初次部署

- 选择物理服务器：
 - 尽量不启动新服务器
 - CPU和I/O互补
- 挂接系统模板
 - 配置主机名、IP

○ 迁移

- 虚拟机操作系统复制
- 动态迁移：边运行边迁移



弹性能力提供

- 向上/下扩展（**Scale Up/Down**）
 - 根据负载高低调整资源分配
 - 多用于小型系统
- 向内/外扩展（**Scale In/Out**）
 - 根据负载升降增加或减少虚拟机数量
 - 负载均衡是关键
 - 多用于大型系统



资源监控

○ 功能

- 状态监控
- 性能监控
- 容量监控
- 安全监控
- 使用度量

○ 常用方法

- 日志分析
- 报嗅探
- 探针采集



资源调度

- 需求分析和预测
- 网络状况
- 负载均衡
 - 负载类型的互补
- 能耗

