第1章 云计算概述

定义：云计算是一种信息技术（IT）的模式，它可以无所不在地访问共享的系统资源池和更高级别的服务，通常通过Internet，以最少的管理成本来快速提供服务。

基本属性（7个）：按需的：资源使用者可以按需获取计算服务；资源池化：云的提供商的计算资源是池化的；无处不在的网络访问：云服务和资源可通过异构网络访问获得；位置独立：使用者感知不到资源的位置；快速弹性：可以快速灵活地提供性能；按使用量付费：服务的使用者仅为使用的服务付费。多租户：应用程序和资源可以在多个使用者之间共享，并且使用者感知不到相互的存在。

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。水平扩展即增加或减少IT资源的数量。包括水平分配资源也称为向外扩展，和水平释放资源也成为向内扩展 。垂直扩展。即改变IT资源容量。当一个现有IT资源被具有更大或更小容量的资源所替代。被具有更大容量的IT资源替代，称为向上扩展，被具有更小容量的IT资源替代，称为向下扩展。

云服务：指任何可以通过云远程访问的IT 资源。

云服务用户：是一个临时的运行时角色，由访问云服务的软件程序担任。比如能够通过已发布的服务合同远程访问云服务的软件程序和服务，或运行某些软件的工作站、便携电脑和移动设备。

交付模式：软件作为服务SaaS、平台作为服务PaaS、基础设施作为服务IaaS

经济合理性：规模经济，设备利用率，多租户

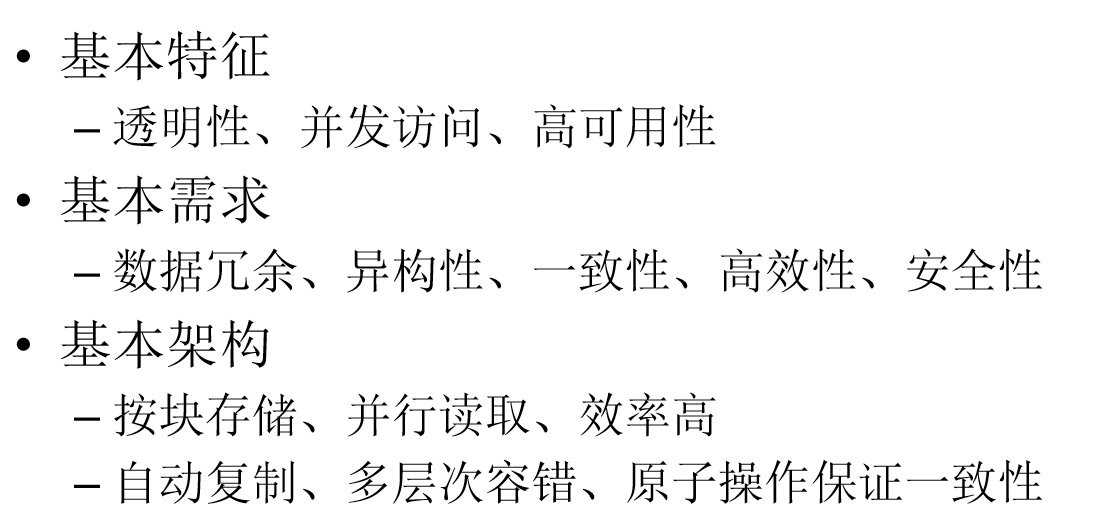
基本机制：虚拟机：具有与其他虚拟机隔离的地址空间，从应用程序角度看像一台裸机，分配了IP地址并具有网络功能，可以加载在主机处理器的任何操作系统或应用程序上。监控器：用于创建和管理虚拟机的操作系统。文件系统：每个虚拟机都可以访问文件系统Hadoop分布式文件系统HDFS：（客户端，主节点，从节点）。容错性。数据备份：权衡数量和可用性。

云架构：安全（非故意的信息共享，虚拟机逃跑，拒绝服务攻击）：性能：（保证性能的两种方法：负载平衡，自动扩展）可用性：（策略：冗余和备份）

第2章 云计算与大数据处理

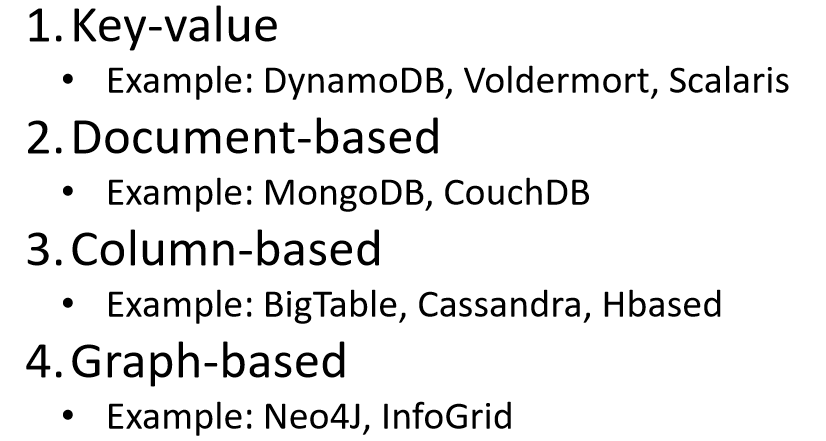
并发控制（锁：共享型锁和排他型锁、基于时间戳协议：读操作与写时间比较（读后不能写）。写操作与读写操作比较，均不能在两者前）

分布式文件系统（HDFS）：

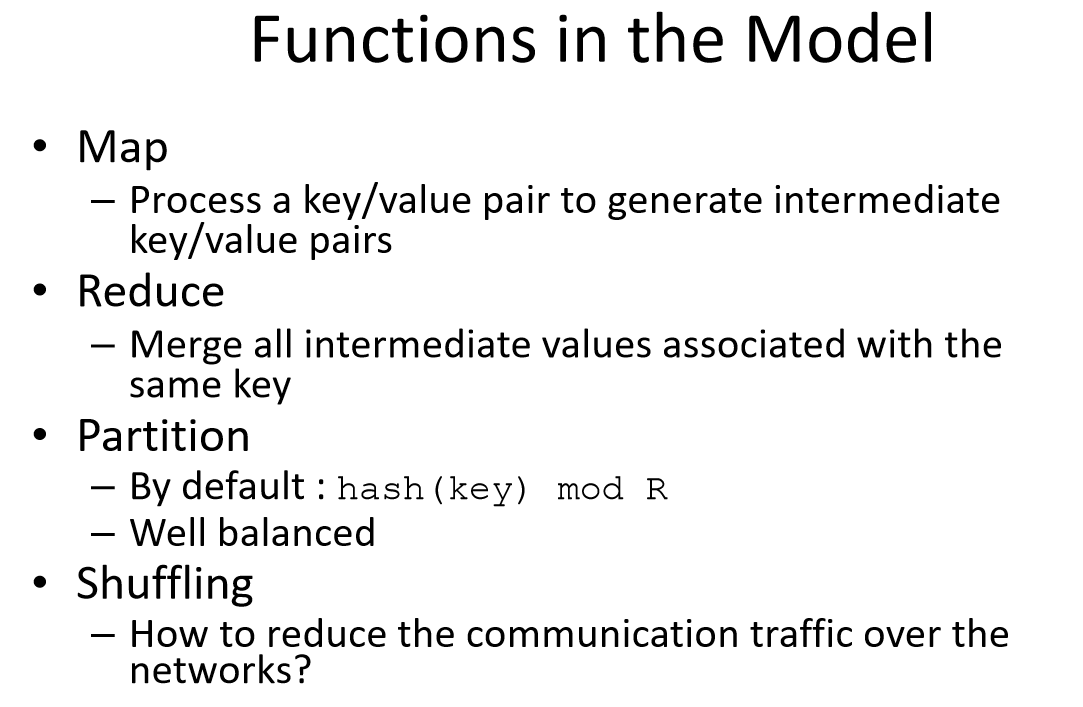


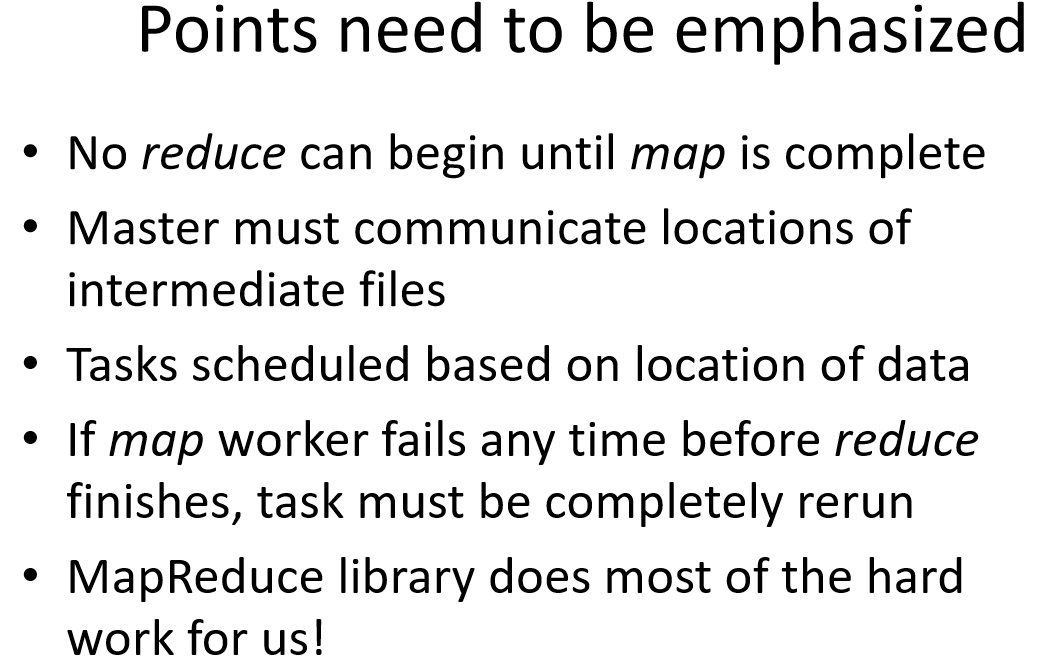
NoSQL系统：优点：支持非结构化数据。数据具有可用性可靠性。容易实现水平扩展。成本低，易于实现。保障大量数据写的性能。支持非结构数据的访问。缺点：不支持关系特征。不支持声明式查询语句。更少的保障（ACID）。不易于和支持SQL的应用集成。

分类：基于键-值，基于文件，基于列，基于图。

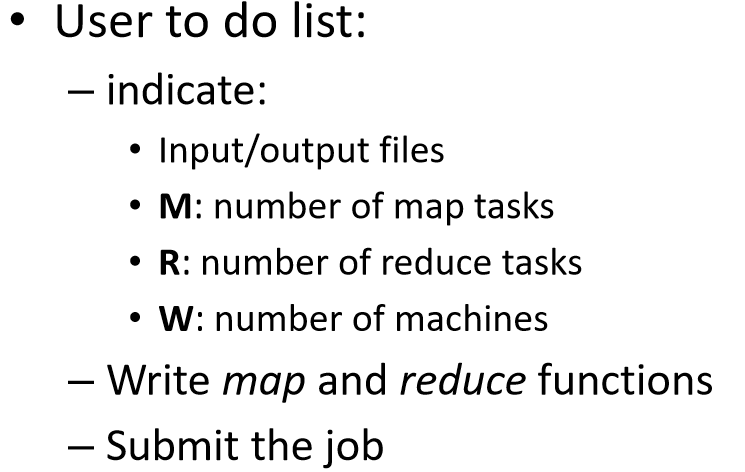


MapReduce：源于谷歌。一个简单的编程模型。功能模型。用于大规模的数据处理平台。



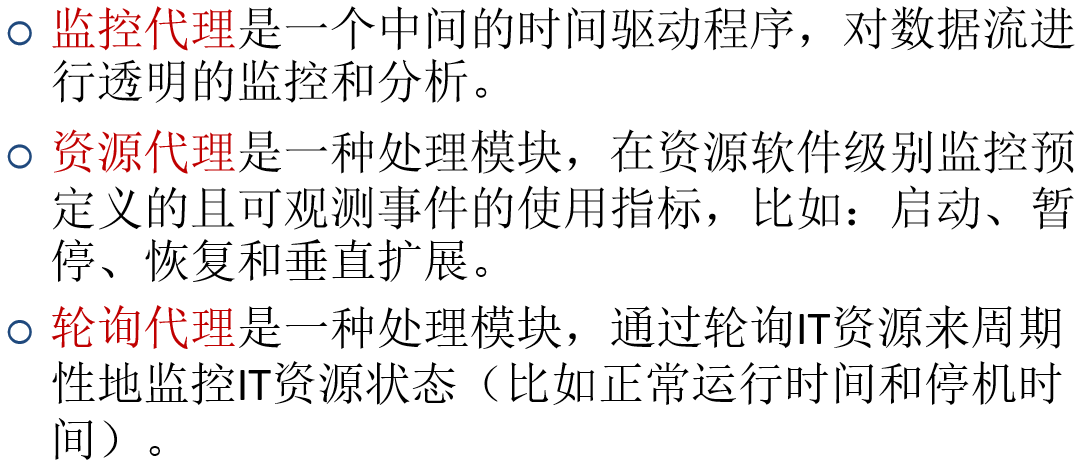


数据局部性，容错性，

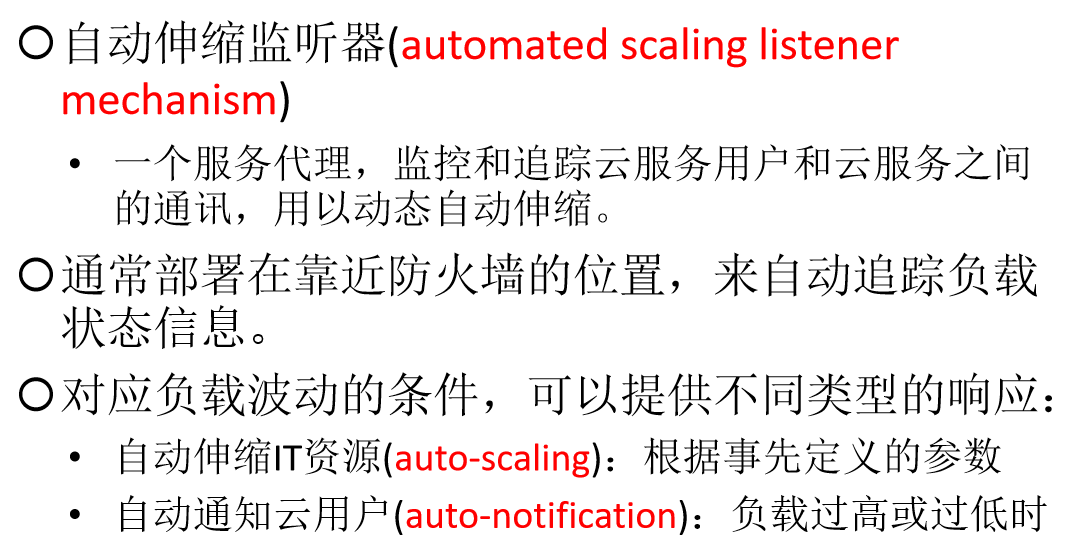


第3章 特殊云机制

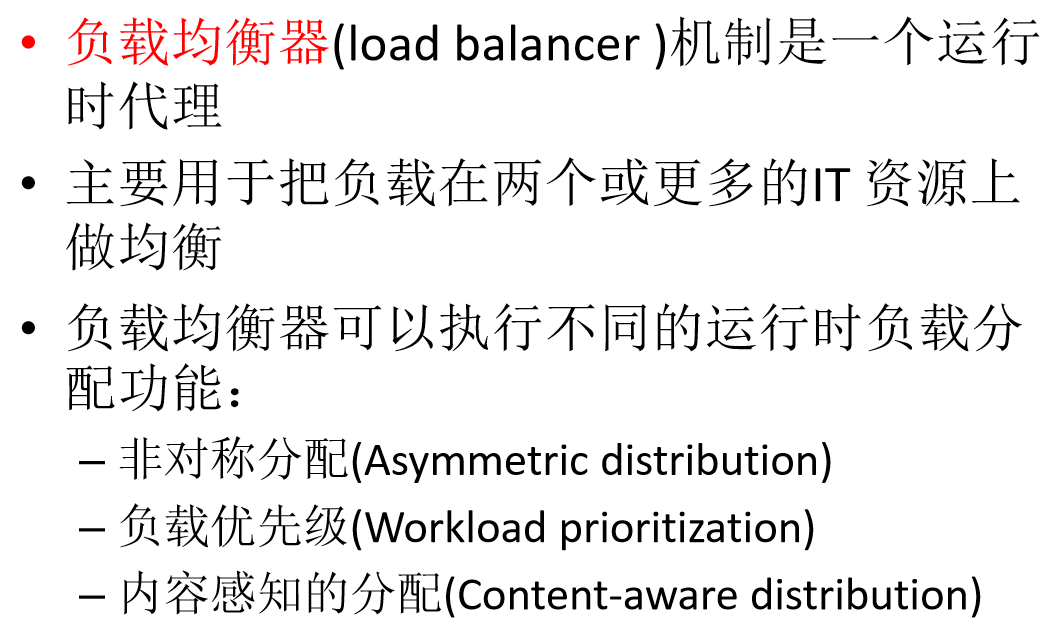
云使用监控（三种实现方式：监控代理，资源代理，轮询代理）

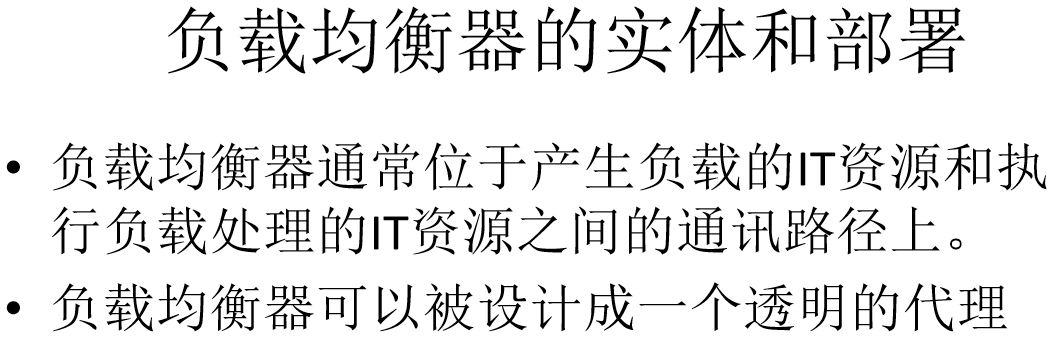


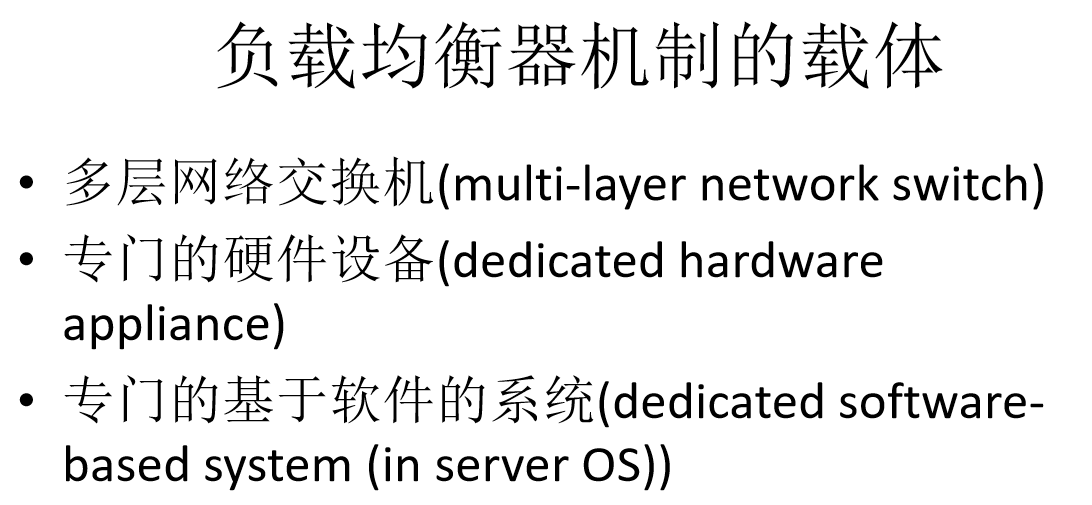
自动伸缩监听器：一个服务代理，监控和追踪云服务用户和云服务之间 的通讯，用以动态自动伸缩。



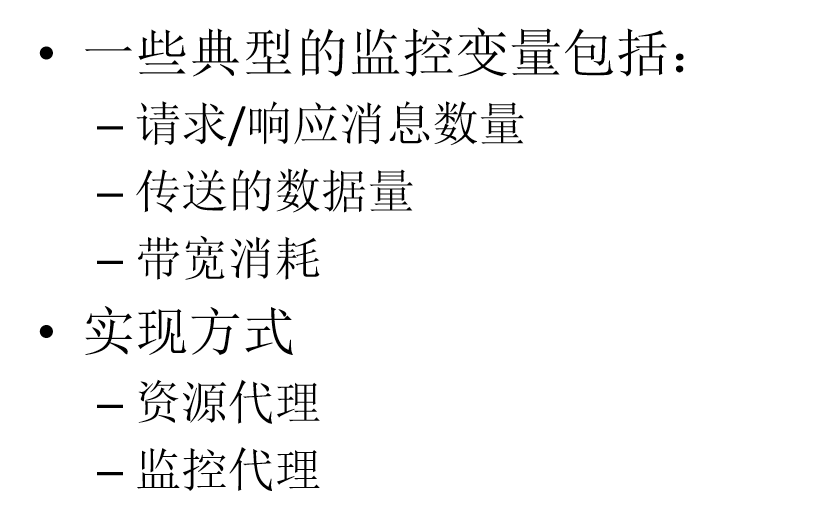
负载均衡器：



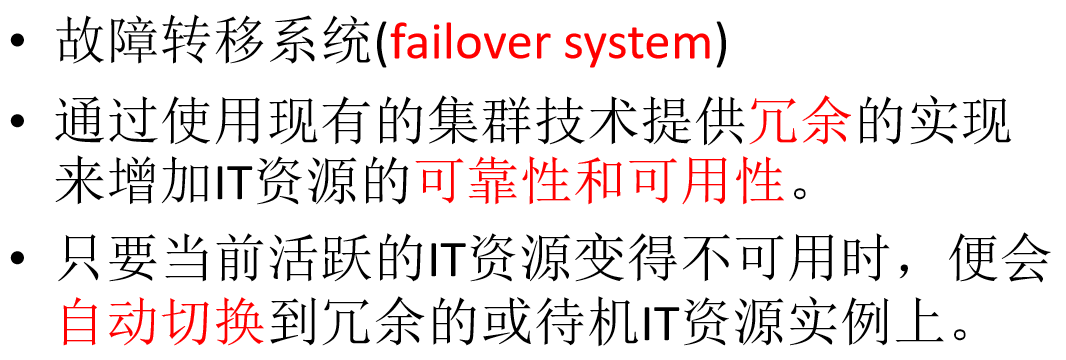


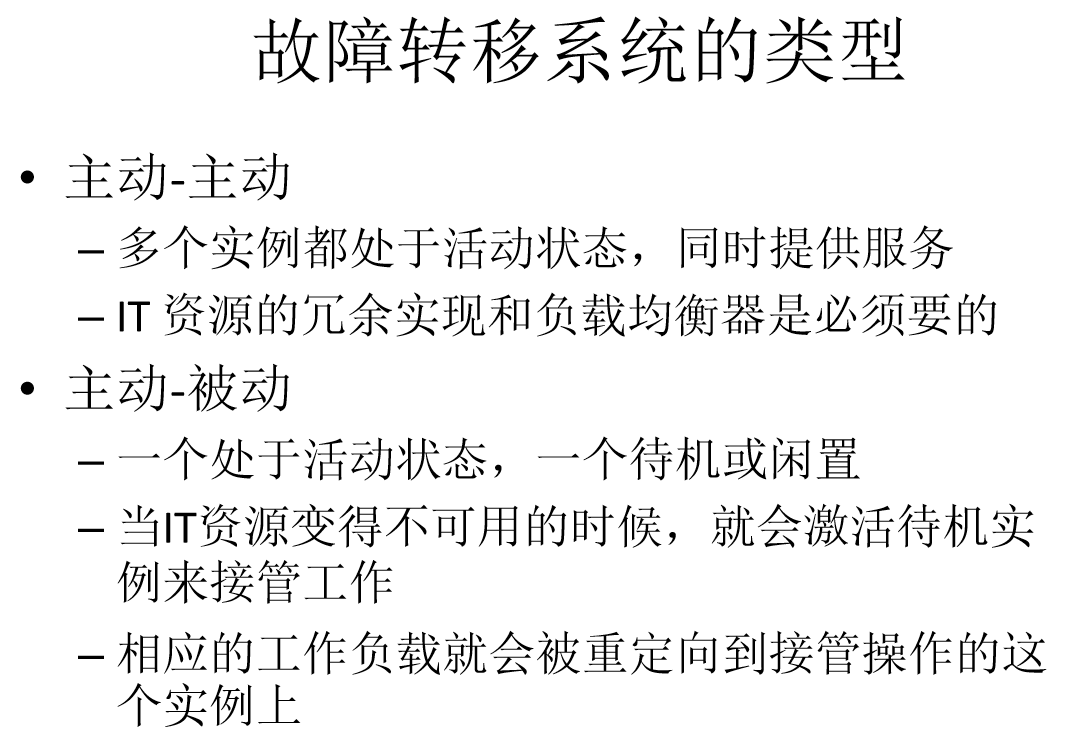


按使用付费监控器：

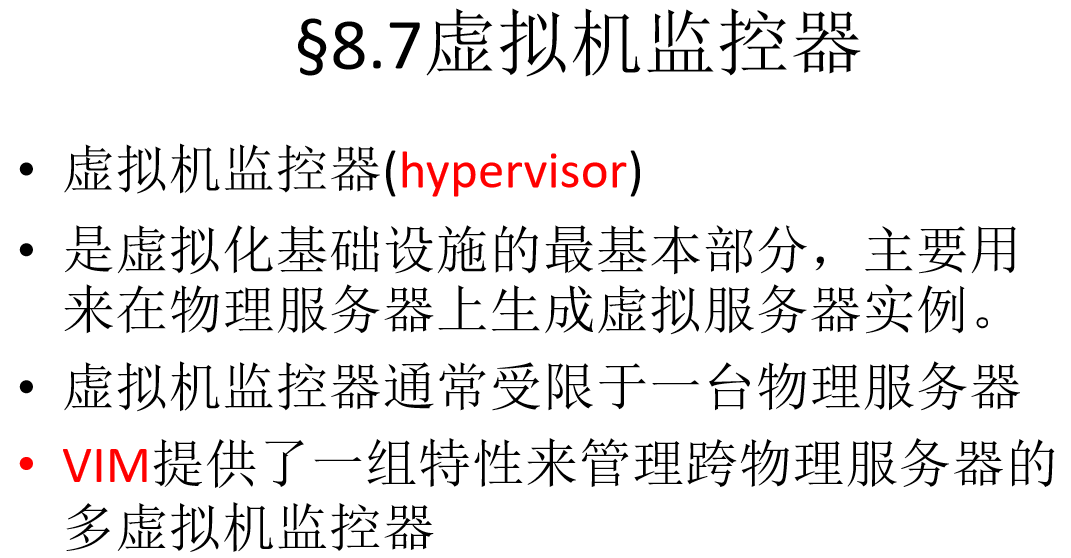


故障转移系统：

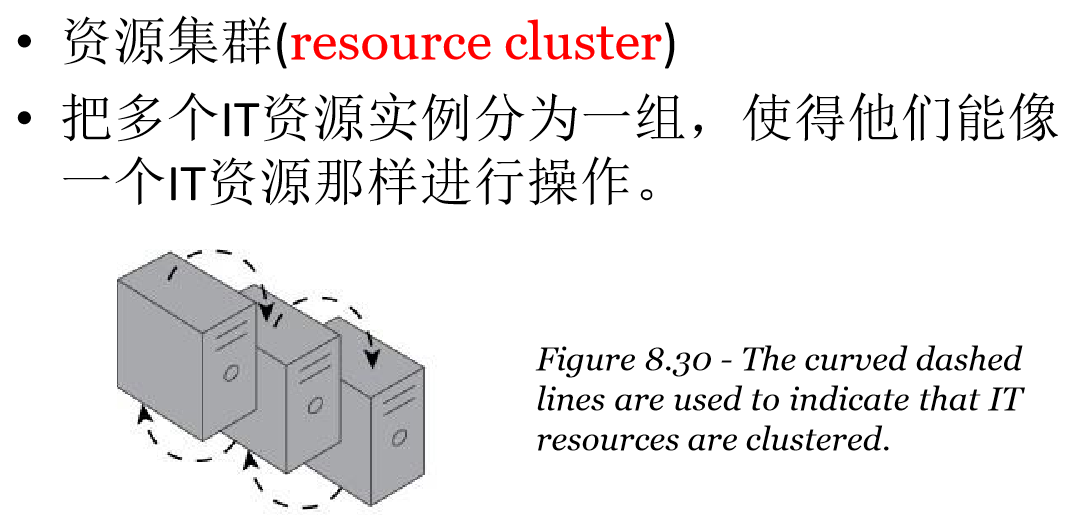


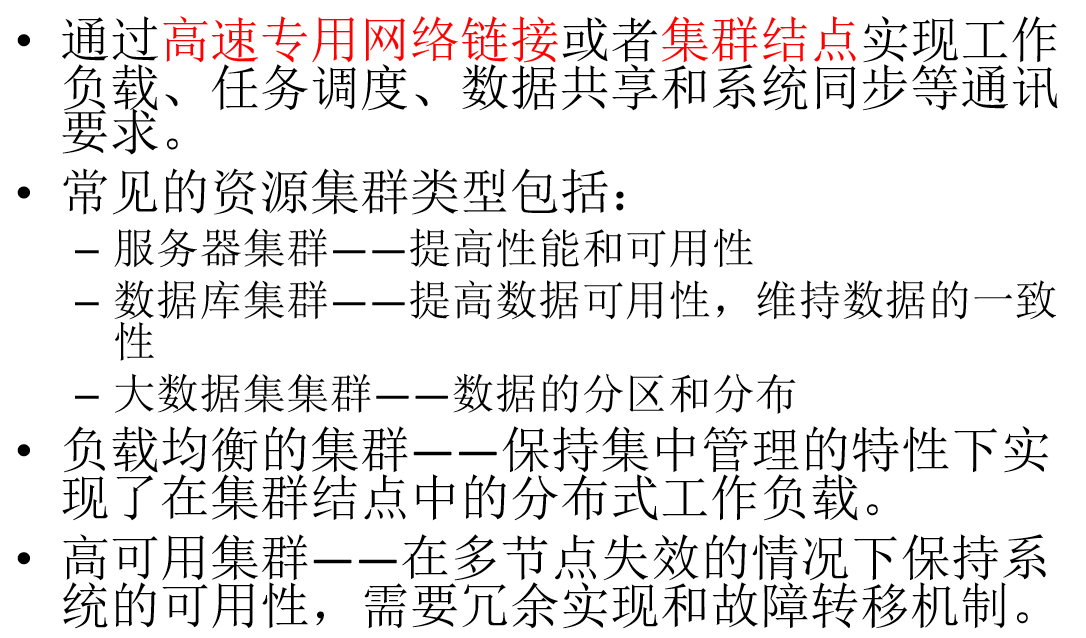


虚拟机监控器

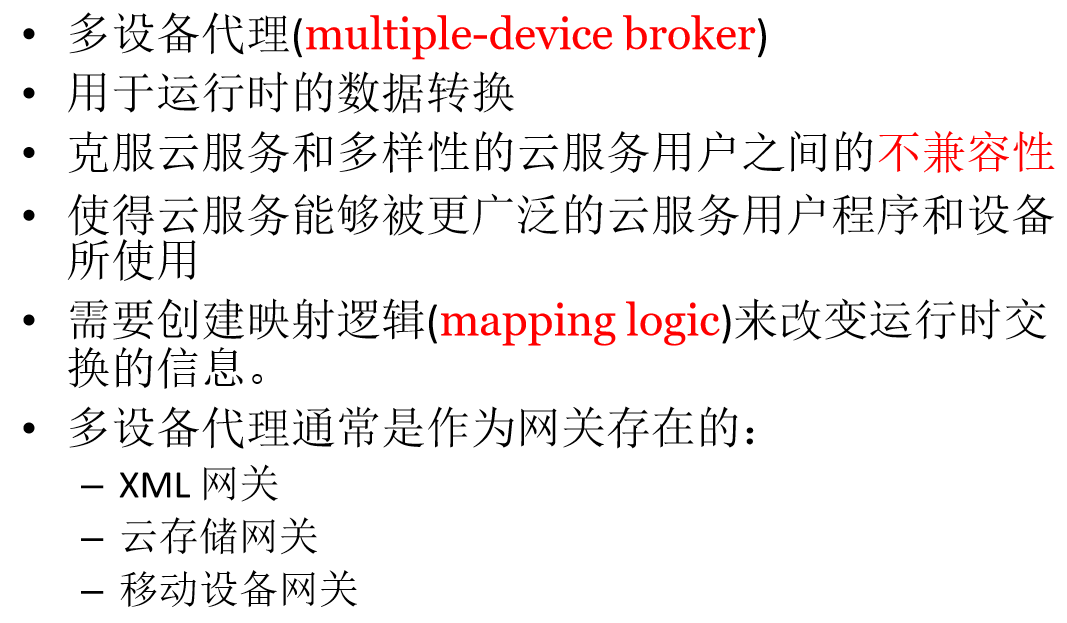


资源集群：

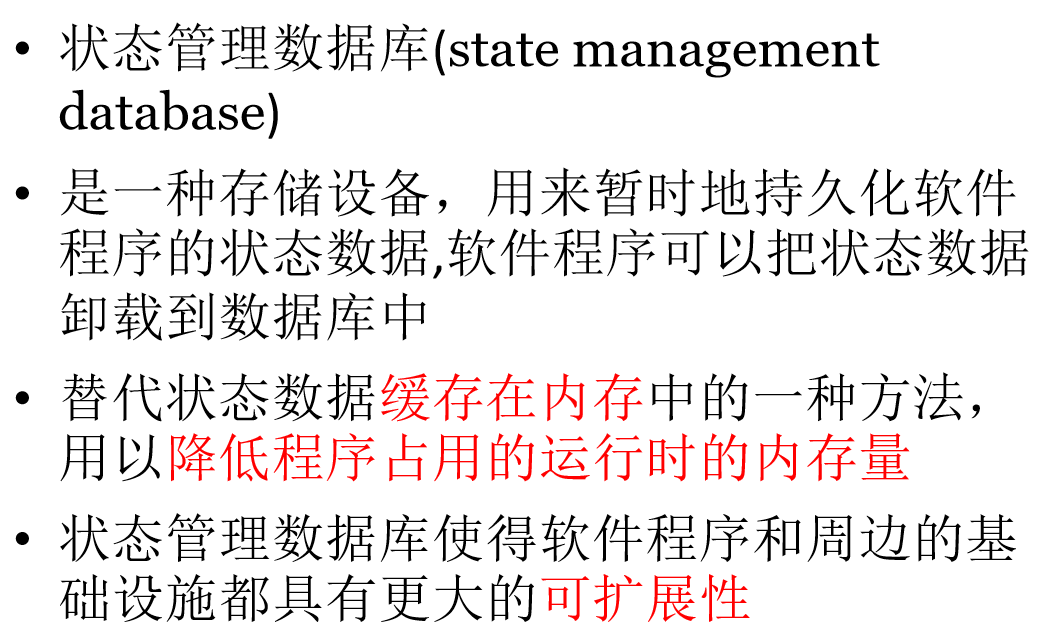


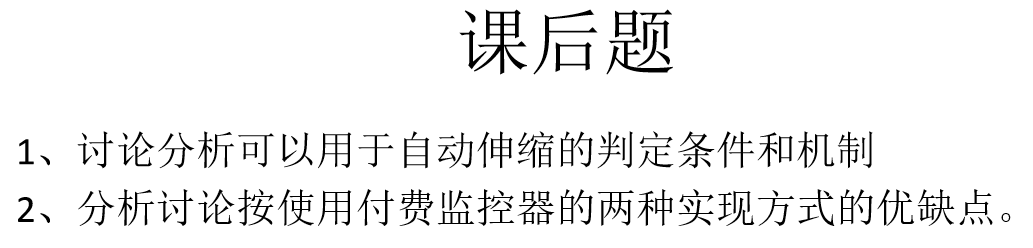


多设备代理：



状态管理数据库：





1：根据资源的使用率。根据负载/流量。监控代理和资源代理。

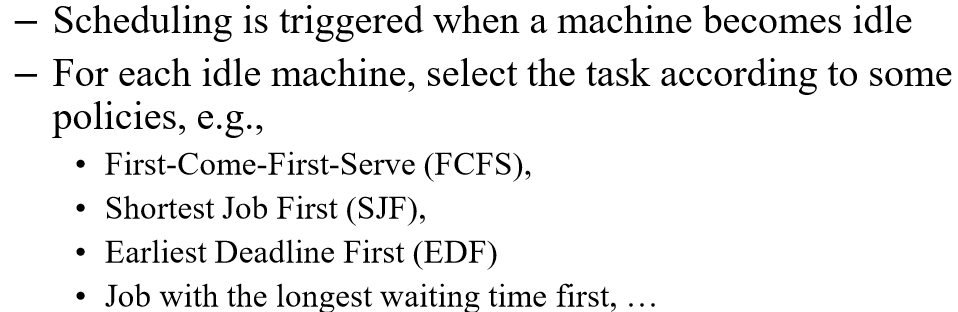
2：监控代理，资源代理。

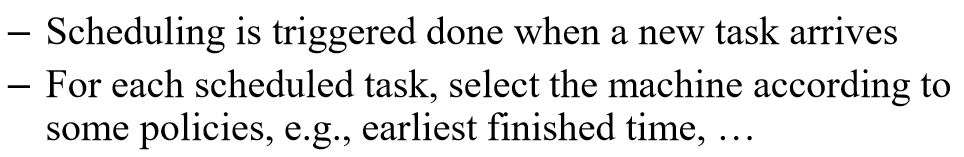
第4章

常见调度问题：当出现对资源竞争的情况下就需要资源调度。处理器调度，缓冲调度，网络调度。三维框架（三要素）：任务-资源-目标。

调度方法：在线调度方法：

基于机器的调度：

和基于任务的调度：



列调度（ List Scheduling ）算法：

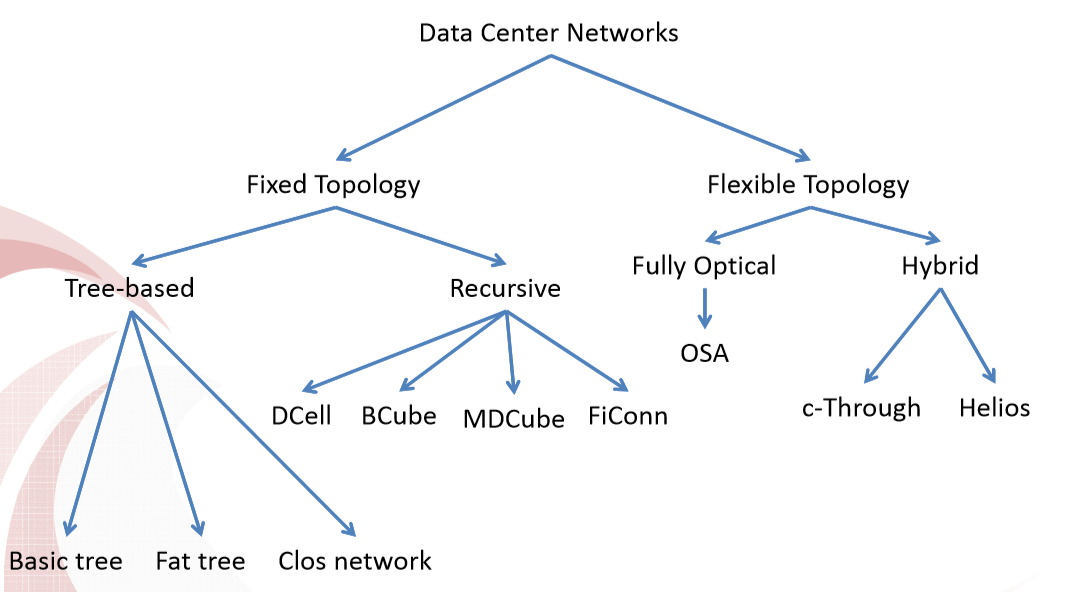
Step1：任务选择，根据任务优先级选择任务，step2：选择处理器最小化预定义的成本函数，重复step1，2直到所有任务被调度。

第5章 数据中心网络

设计目标与需求：目标：敏捷性（任何服务，任何服务器）。将服务器变成单个大型可替代池（按需动态扩展和收缩其占用空间）好处（提高服务开发人员的生产力，成本更低，实现高性能和可靠性。需求：均匀的高容量（服务器间的容量只受NICs限制，增加服务器时不需要考虑拓扑结构）。性能隔离（服务的流量不受其他服务的影响）。易于管理。对可扩展性，易于管理，容错性和高效的数据中心网络（DCN）的要求（任何VM都可以迁移到任何物理机，而无需更改其IP地址。管理员无需在部署之前配置任何交换机。任何终端主机都应通过任何可用路径与其他终端主机进行有效的通信。没有发送循环。快速有效的错误检测）。对网络协议的意义。

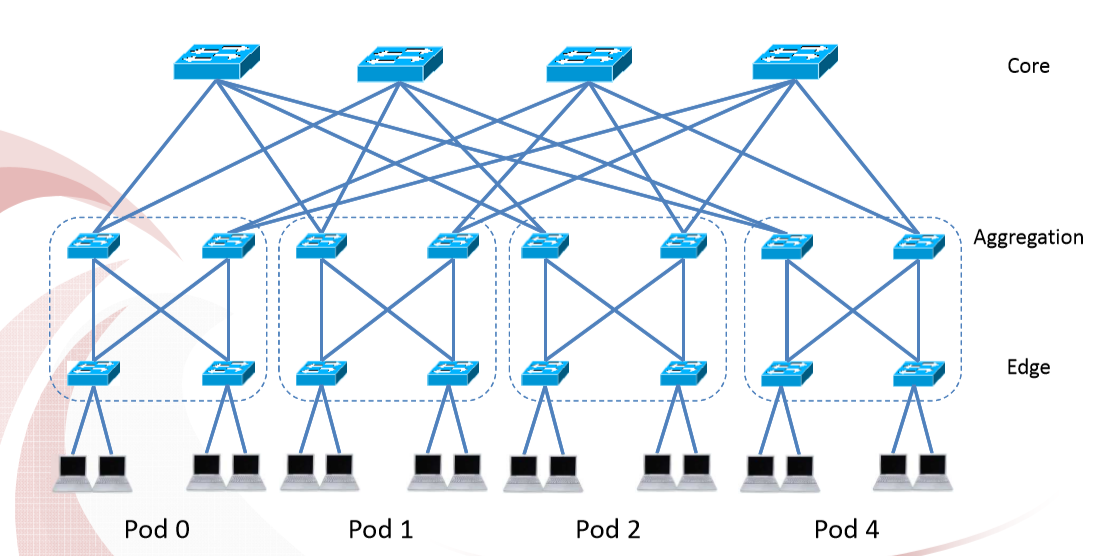
成本：

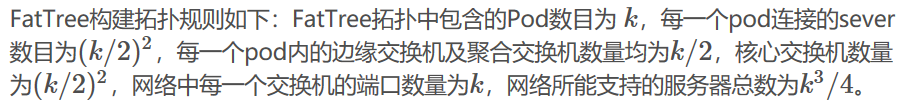
数据中心拓扑结构（分类）



常见拓扑结构（基本树型网络，胖树型网络，Clos 网络：n阶Clos和折叠Clos），重点掌握 FatTree 架构（三层：边缘层（edge）、汇聚层（aggregate）和核心层（core），其中汇聚层交换机与边缘层交换机构成一个pod）

胖树是一种特殊的Clos网络：K叉胖树

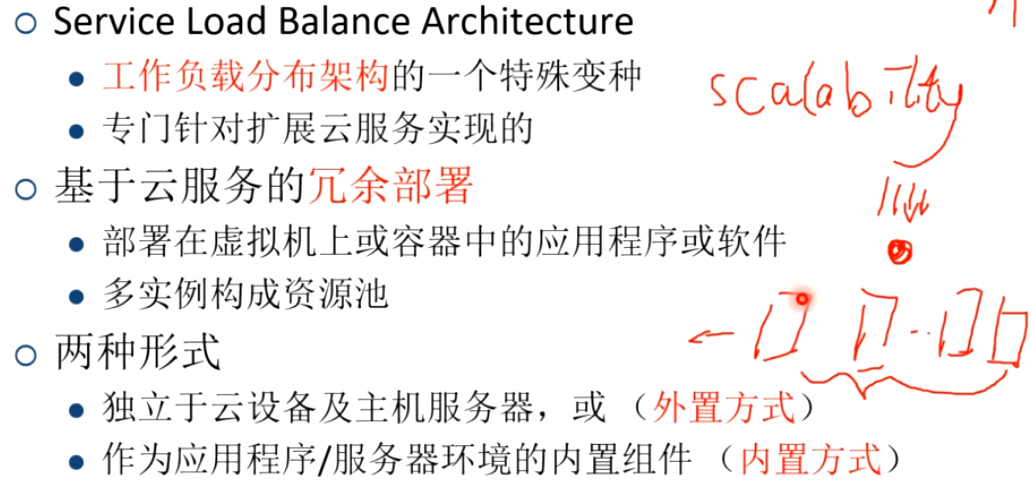




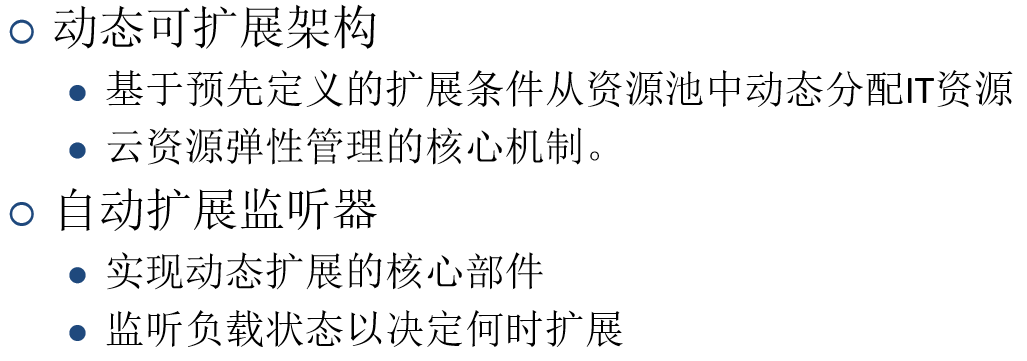
胖树结构属性：在任何二等分点都具有相同的带宽。每层具有相同的聚合带宽。可以使用具有统一容量的廉价设备来构建（每个端口与终端主机支持的速度相同。如果数据包分布均匀则所有设备都可以以线路速度传输）。出色的可扩展性。

第6章 云架构

负载均衡架构：

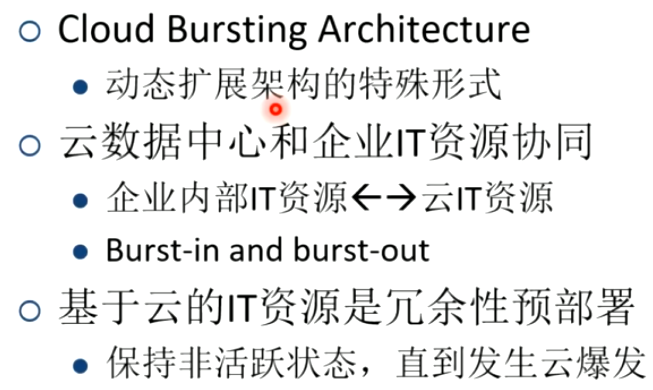


动态可扩展架构（水平、垂直、重定位）：



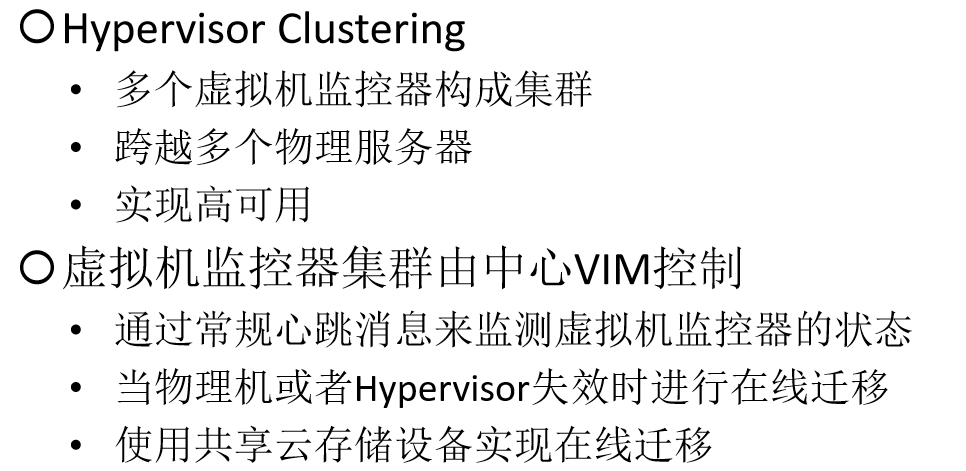


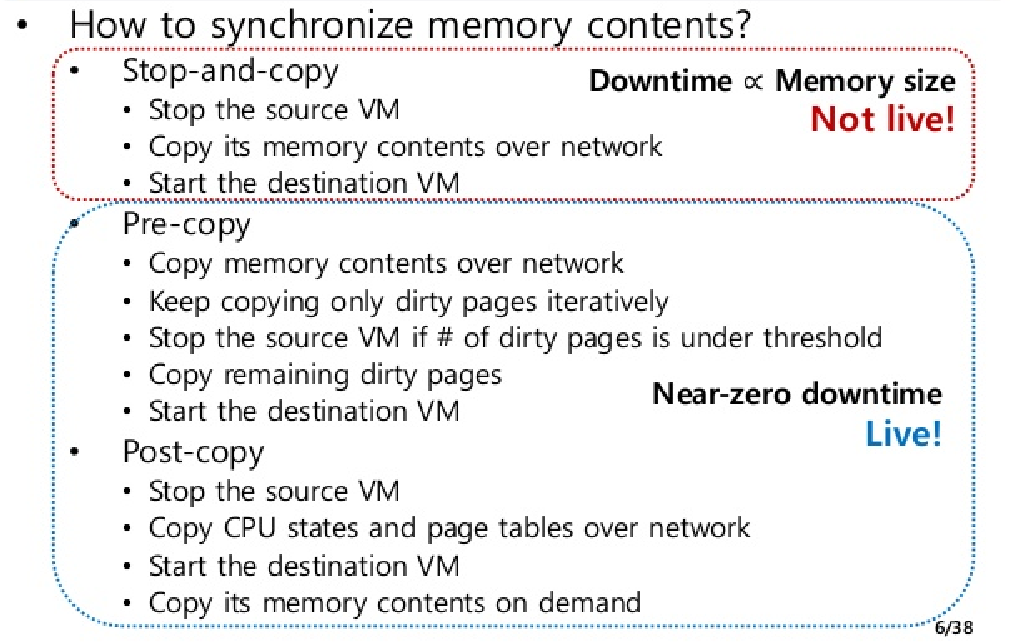
云爆发架构：

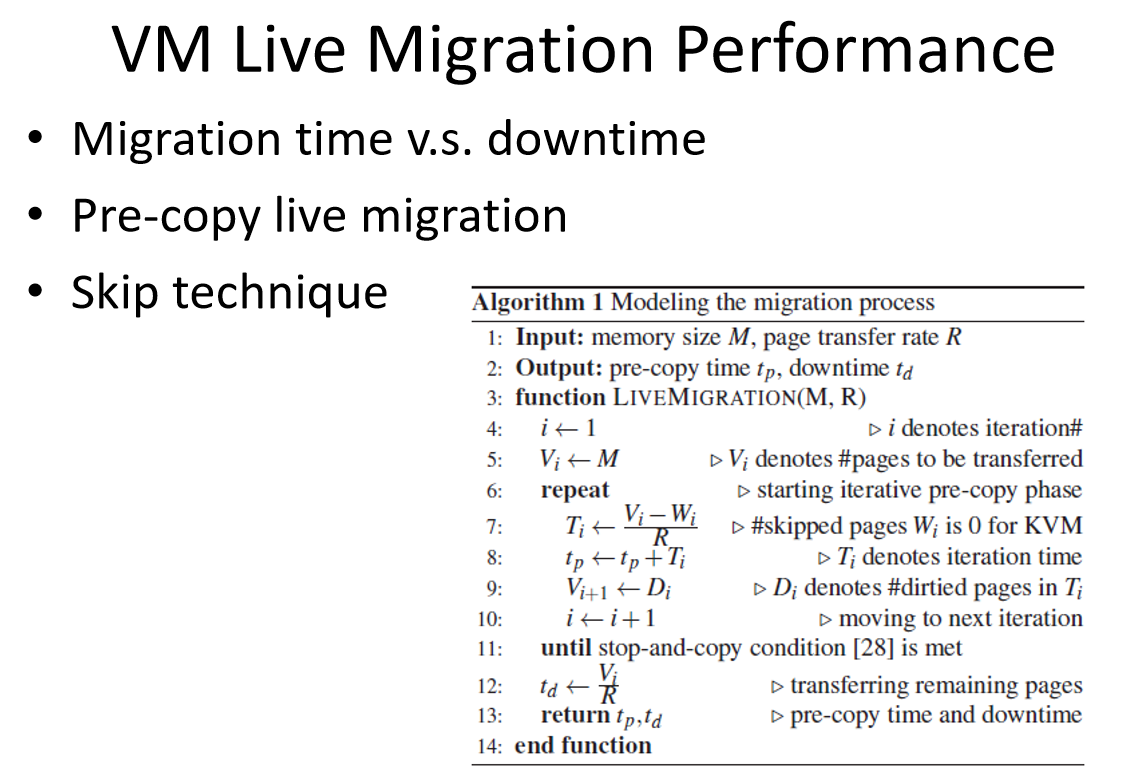


弹性磁盘供给架构：

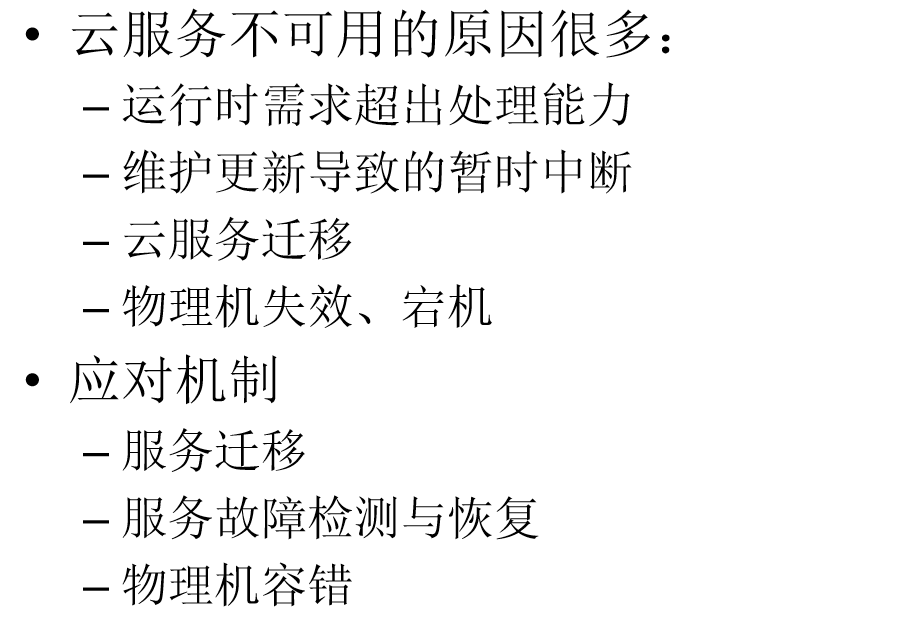
虚拟机（架构，在线迁移机制，性能指标及函数）：

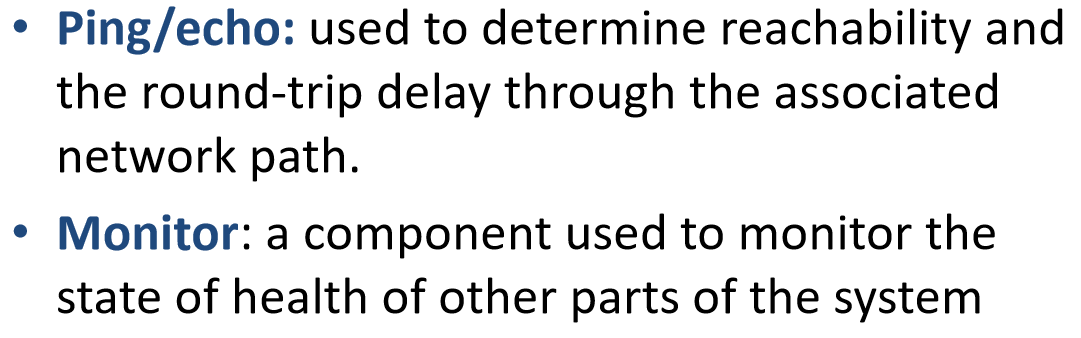


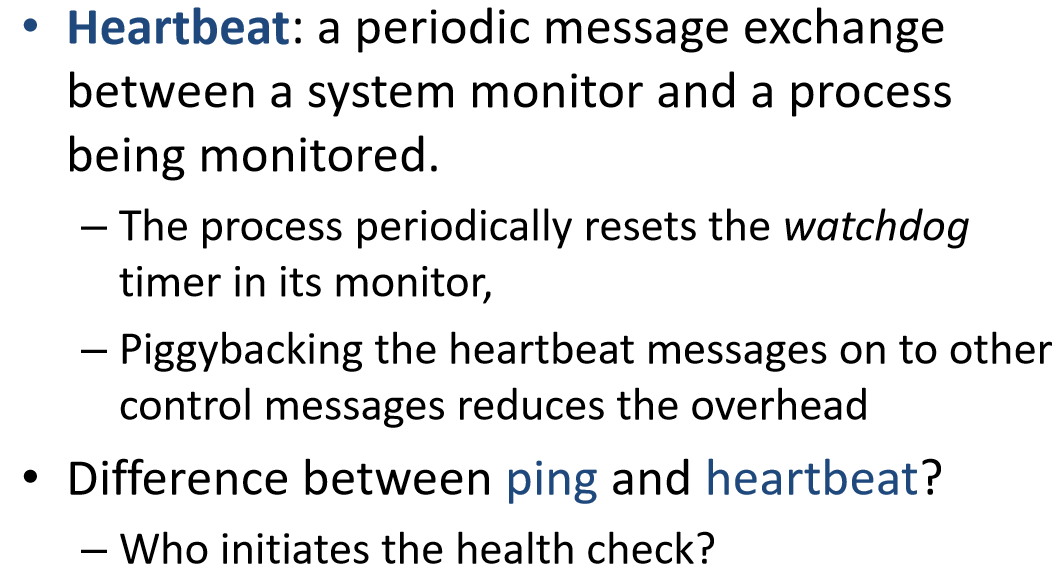




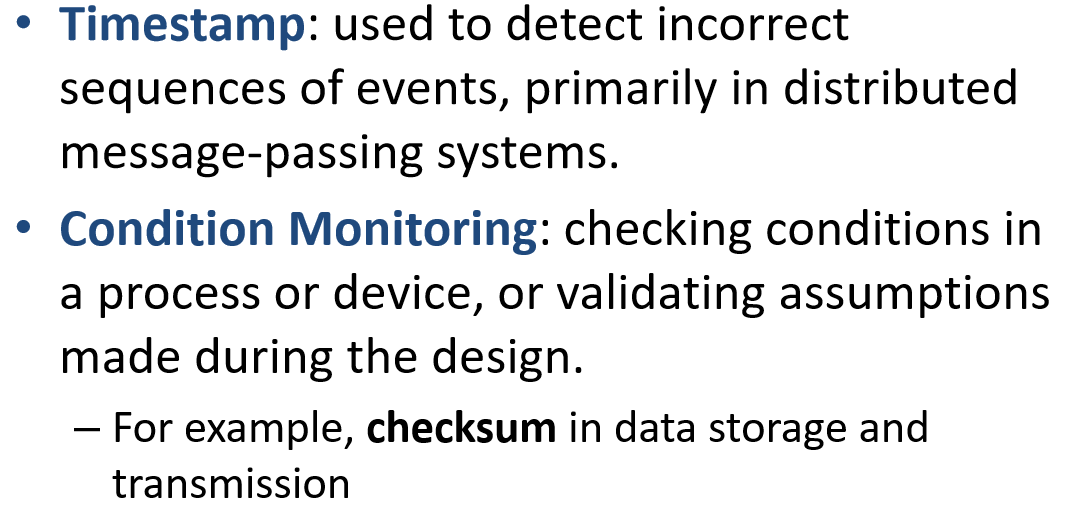
云服务容错架构以及容错机制

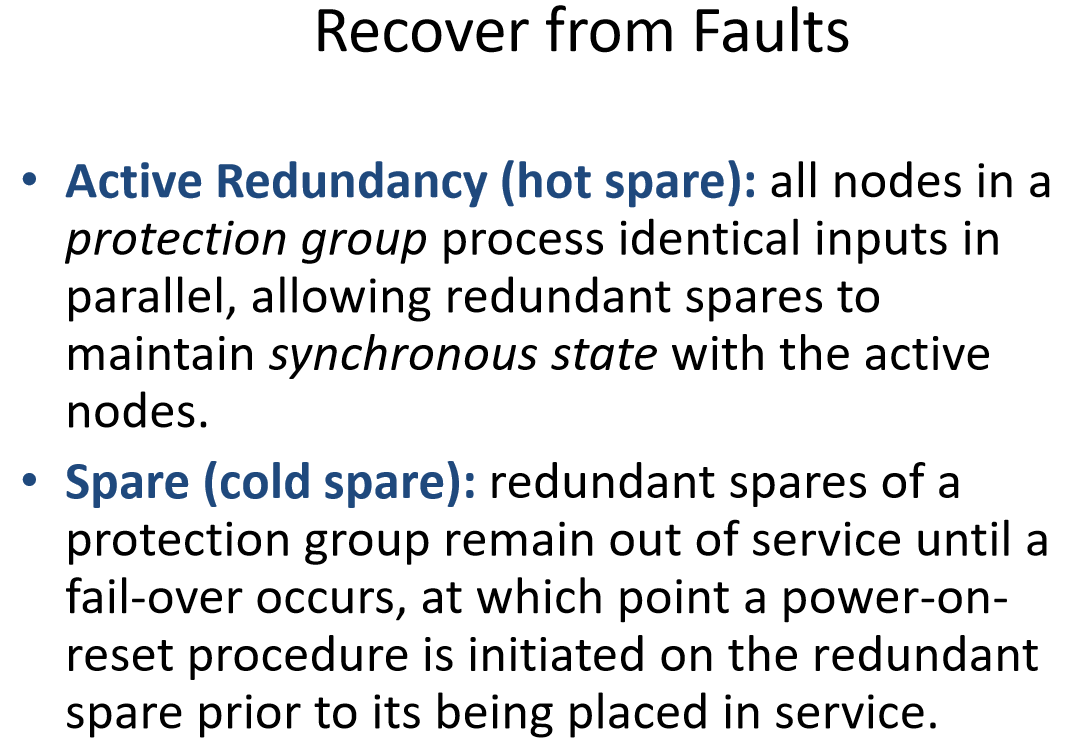


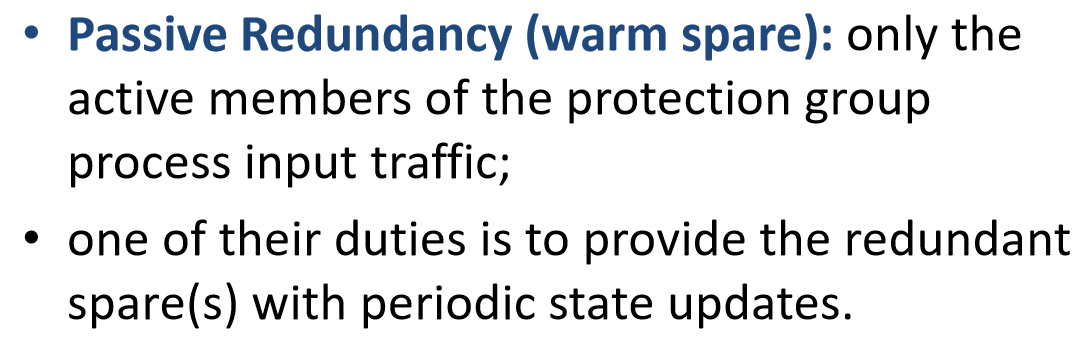


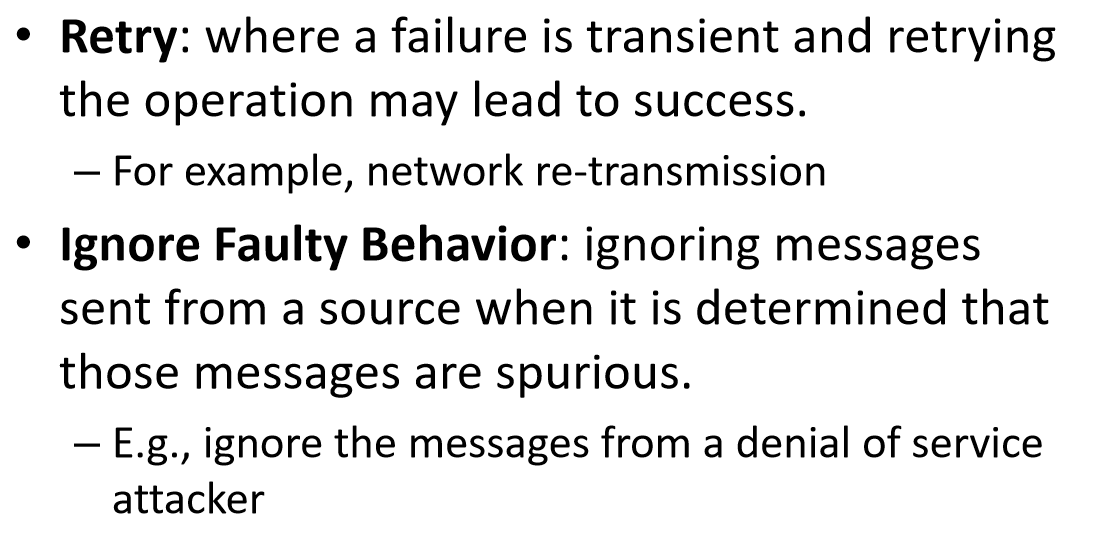


Ping是用户主动发起，heartbeat是被监控对象主动发起。





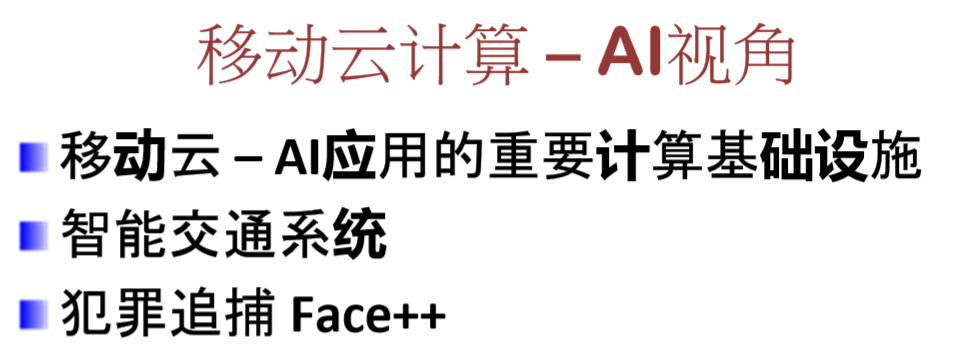




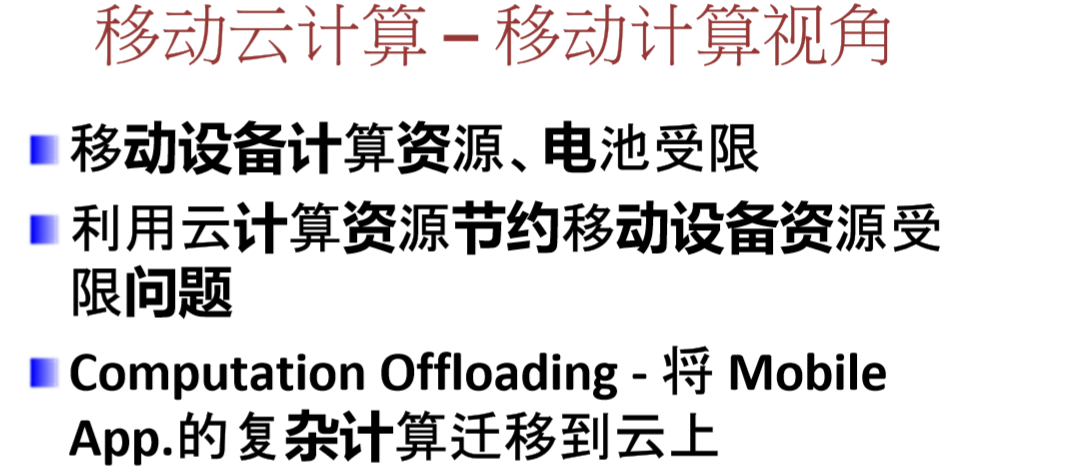
第7章 移动云计算

定义：从三个视角定义

三个视角：

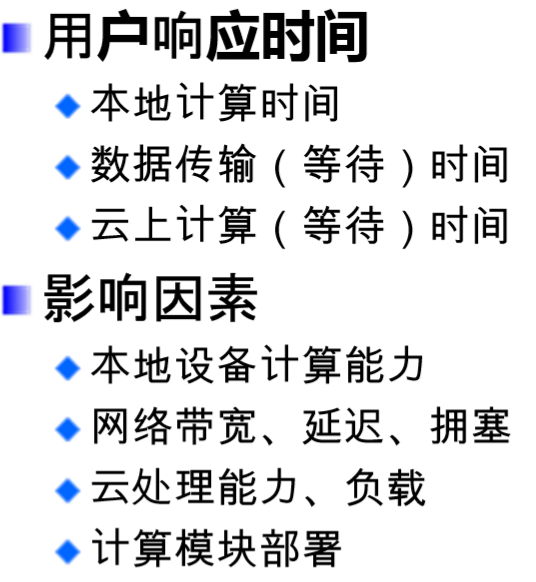






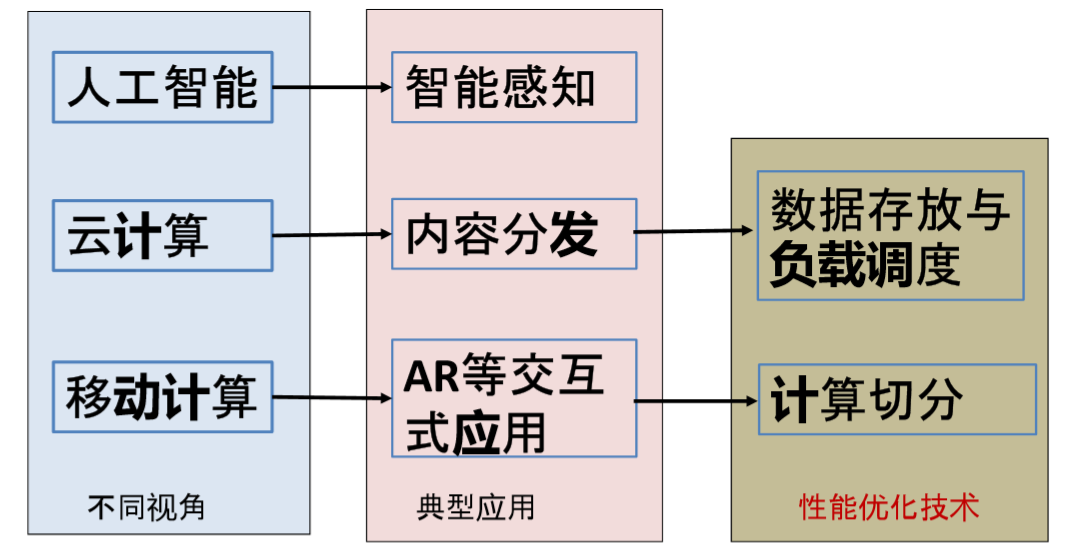
性能建模与优化：

分析出系统性能指标是什么，比如响应时间，吞吐率等。分析出影响系统性能的因素有哪些，比如网络端，服务器端，设备端。然后建立因素和性能之间的关系模型。



性能建模方法：系统分析和数据挖掘

优化技术：



计算切分：将应用程序的工作负载分解为较小的工作负载，并确定哪些部分在本地执行，哪些部分卸到云上执行。