

# 云计算与计算虚拟化

[www.huawei.com](http://www.huawei.com)





# 前言

- 许多公司使用虚拟技术作为建设云计算的基础技术，用以提高硬件资源的利用率，进行灾难恢复、提高办公自动化水平
- 虚拟化技术经历了漫长的发展时期，早在20世纪70年代，大型计算机就一直在同时运行多个操作系统实例，每个实例也彼此独立，当今，软硬件方面的技术取得了巨大的进步，使得虚拟化技术在基于行业标准的大众化x86服务器上得以支撑电信和企业关键应用的部署，虚拟化技术逐步得到了广泛的认可，从而得以普及。



# 培训目标

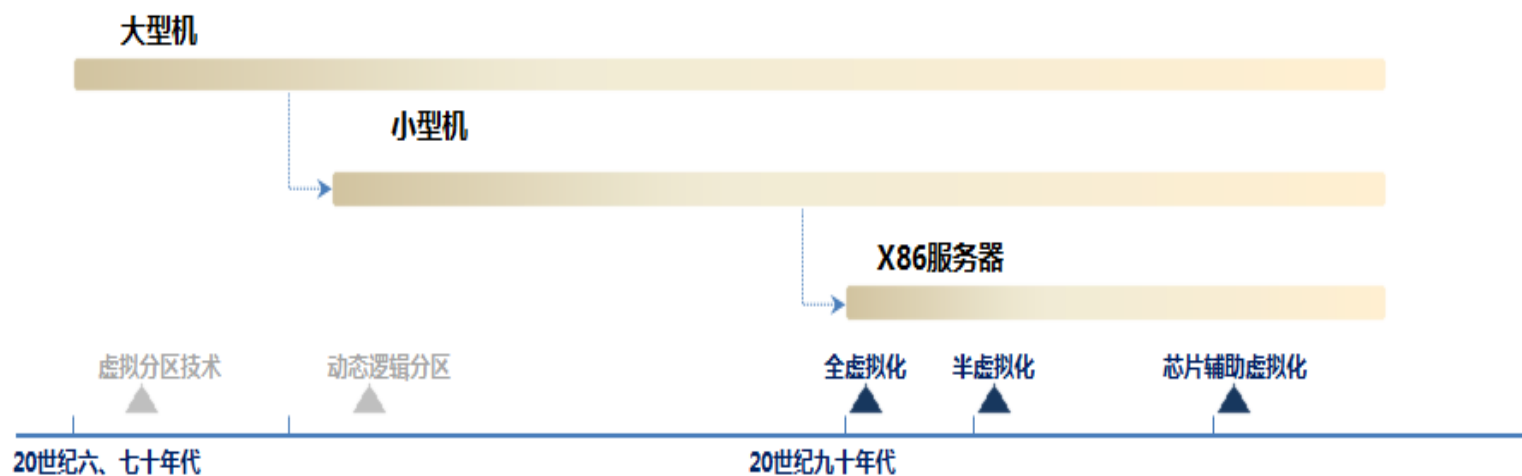
- 学完本课程后，您应该能：
  - 了解计算虚拟化核心引擎
  - 了解跨服务器的计算资源调度算法
  - 了解计算高可靠性保障



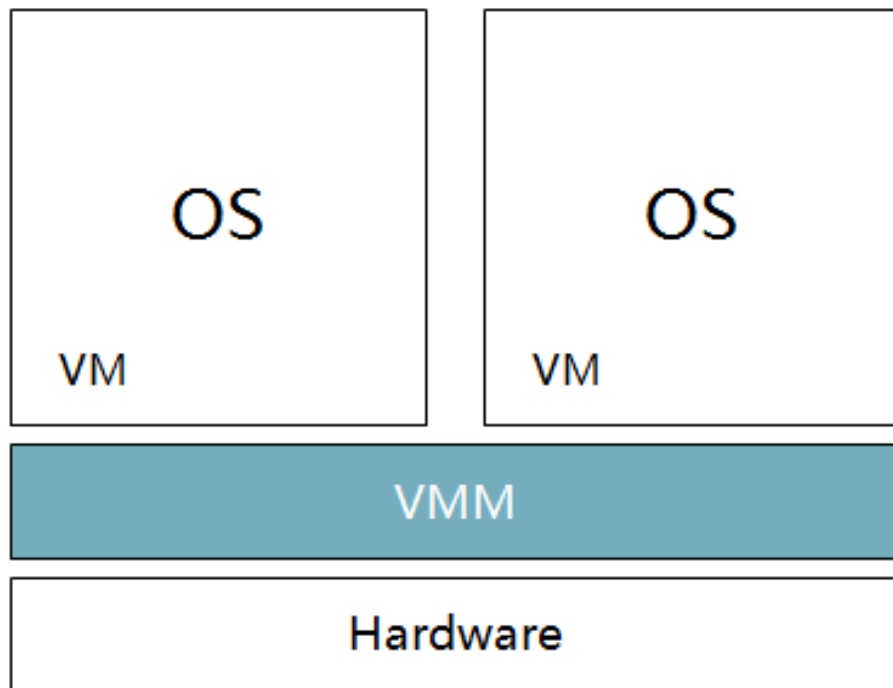
# 目 录

- 计算虚拟化核心引擎：Hypervisor介绍
- 跨服务器的计算资源调度算法
- 计算高可靠性保障

# 虚拟化技术的历史

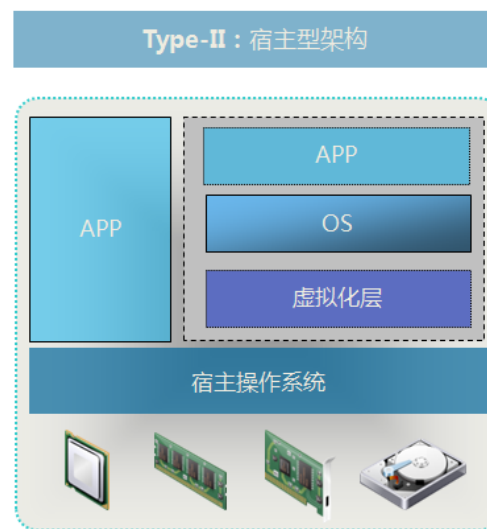
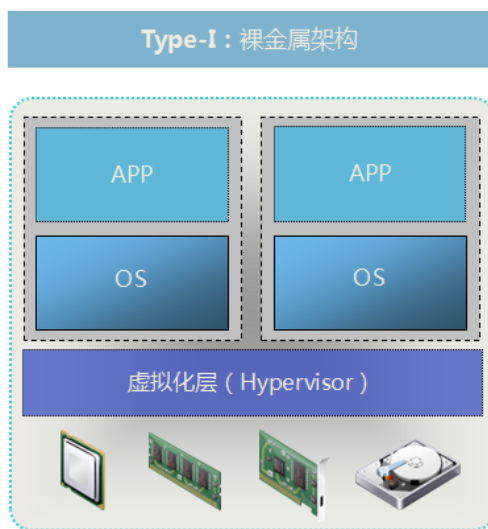


# 虚拟化原理



核心思想：计算虚拟化技术的实现形式是在系统中加入一个虚拟化层，将下层的资源抽象成另一形式的资源，提供给上层使用。

# 虚拟化架构



# 计算虚拟化技术细分

- 全虚拟化 (Full Virtualization)
- 超虚拟化 (Paravirtualization)
- 硬件辅助虚拟化 (Hardware-Assisted Virtualization)
- 部分虚拟化 (Partial Virtualization)
- 操作系统级虚拟化 (Operating System Level Virtualization)



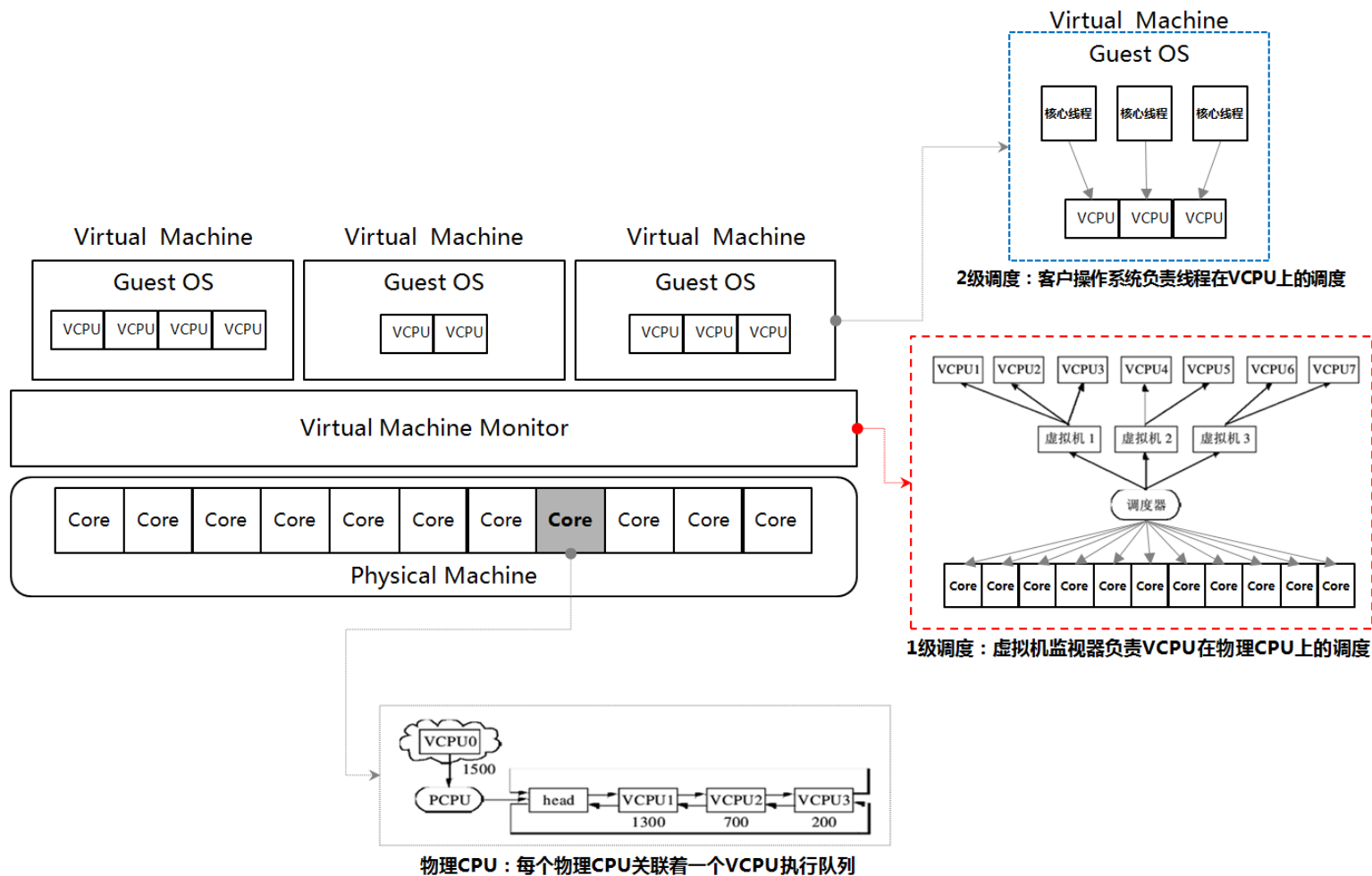
# 满足电信和企业关键应用的计算虚拟化

- 精细化CPU调度技术
- NUMA架构感知的调度技术
- 内存复用技术
- IO调度中断优化技术

# 精细化CPU调度技术

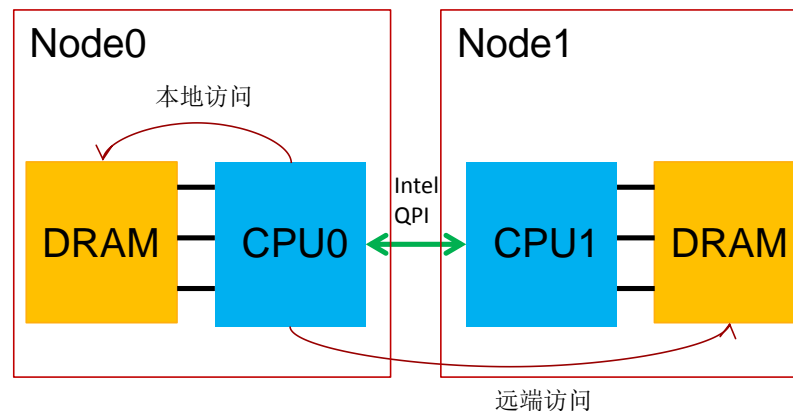
- 精细化CPU调度技术主要指的是CPU上下限配额及优先级调度技术
- X86指令集中有若干条指令是需要被VMM捕获的敏感指令，但是却不是特权指令
  - (1)访问或修改机器状态或虚拟机状态的指令
  - (2)访问或修改敏感寄存器或存储单元的指令，比如访问时钟寄存器和中断寄存器。
  - (3)访问存储保护系统或内存、地址分配系统的指令（段页之类）。
  - (4)所有I/O指令。

# vCPU调度分配机制

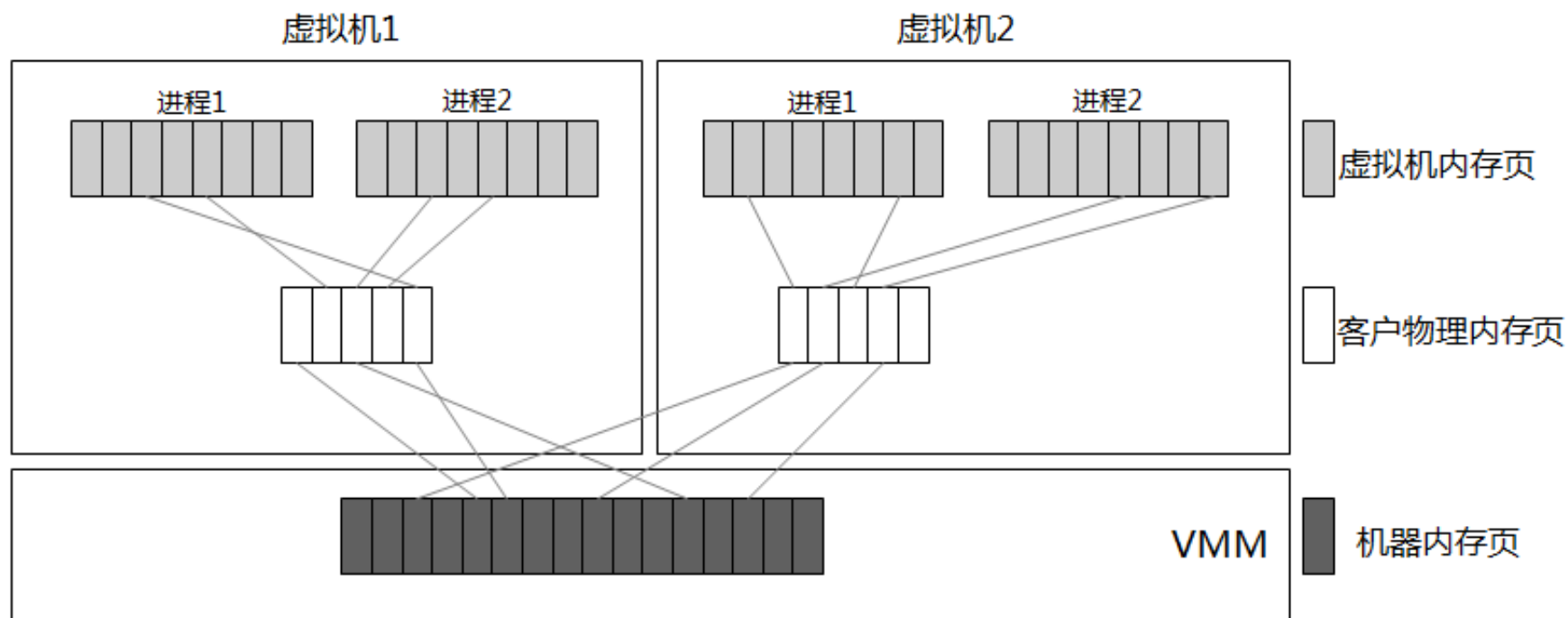


# NUMA架构感知的调度技术

- NUMA：非一致性内存架构(Non-uniform Memory Architecture)解决了SMP系统中的可扩展性问题。NUMA将几个CPU通过内存总线与一块内存相连构成一个组，整个系统就被分为若干个Node
- 通过虚拟化软件的Host NUMA技术，可以显著的提高虚拟机的性能，降低处理时延

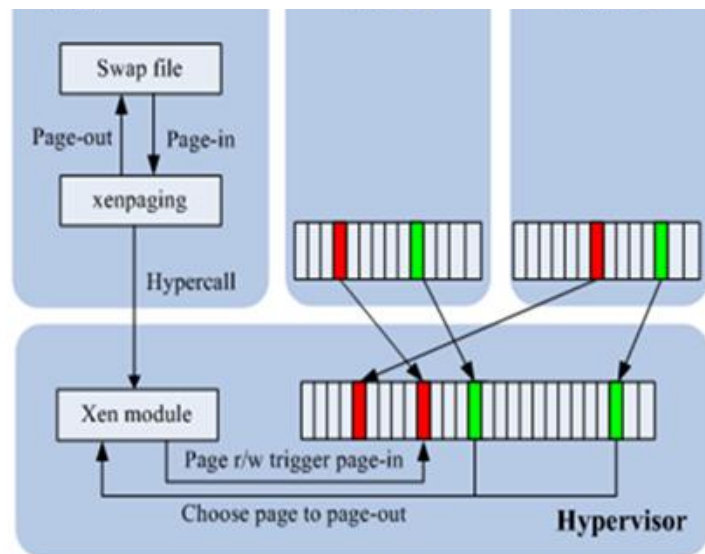
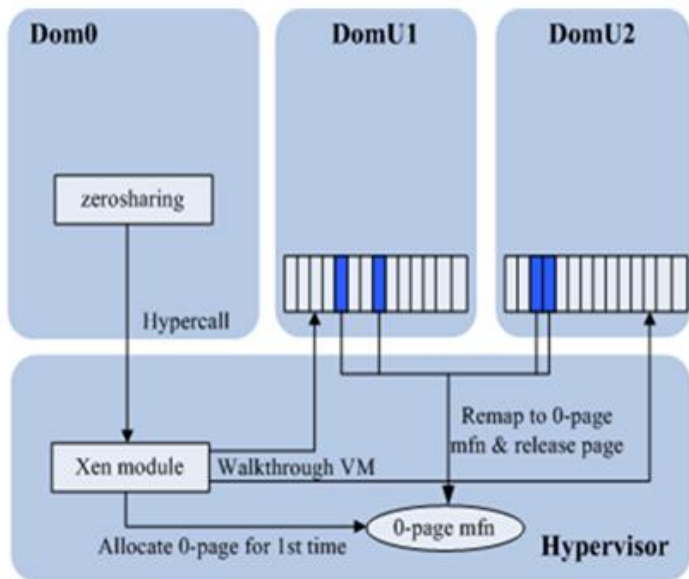
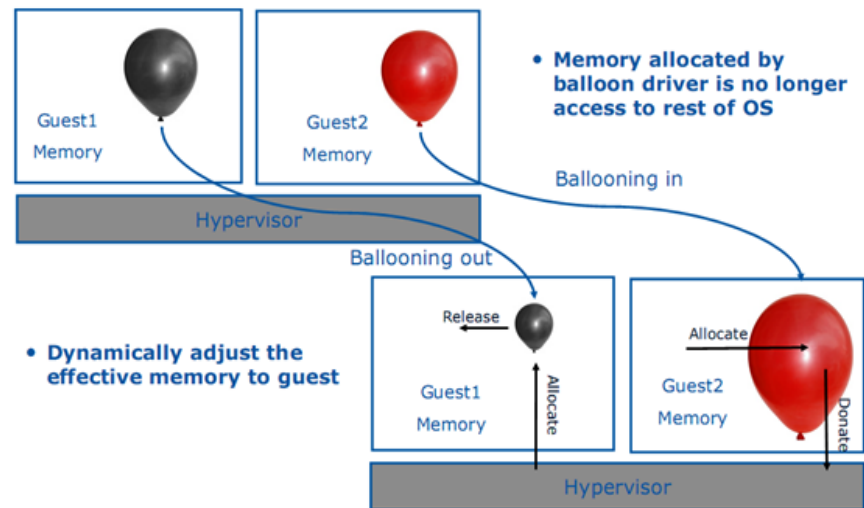


# 内存复用技术



# 虚拟化软件的内存复用技术

- 内存气泡技术(Ballooning)
- 零页共享技术
- 内存交换技术



# IO调度中断优化技术

- 影响电信和企业关键应用的运行时响应时延的因素很多，其中，中断处理就是一个在虚拟化条件下的关键因素。
- CPU在处理IO访问请求的中断时，实际处于等待的状态。
- VMM 通过I/O虚拟化来复用有限的外设资源，其通过截获 Guest OS对I/O设备的访问请求，然后通过软件模拟真实的硬件。
  - 设备接口完全模拟
  - 前端／后端模拟
  - 直接划分



# 目 录

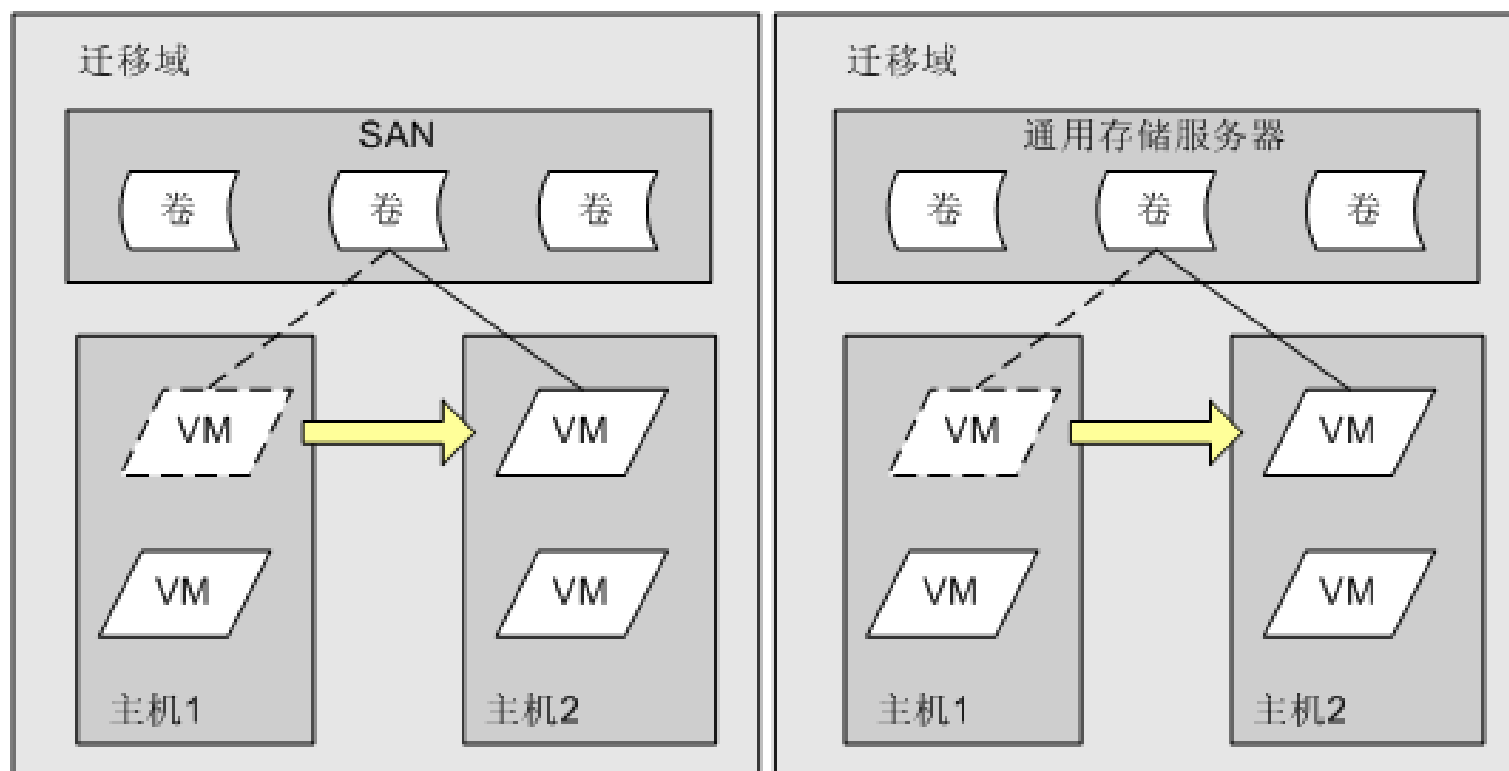
- 计算虚拟化核心引擎：Hypervisor介绍
- **跨服务器的计算资源调度算法**
- 计算高可靠性保障



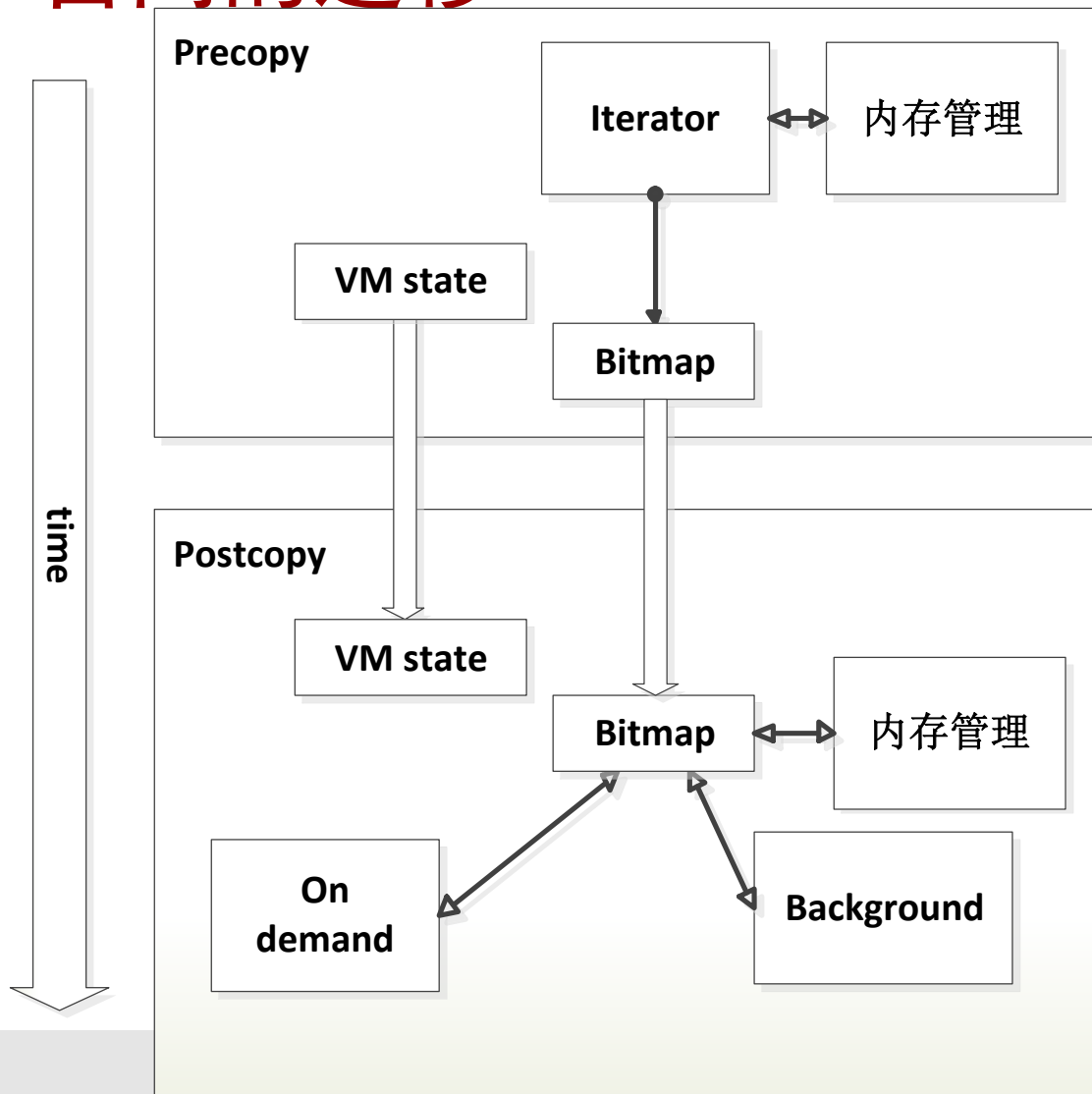
# 虚拟机热迁移的作用

- 降低客户的业务运行成本
- 保证客户系统的高可靠性
- 硬件在线升级
- 一个虚拟化系统，至少需要在下列场景下支持虚拟机热迁移功能：
  - 根据需要按照迁移目的手动把虚拟机迁移到空闲的物理服务器；
  - 根据资源利用情况将虚拟机批量迁移到空闲的物理服务器。

# 高性能、低时延的虚拟机热迁移机制



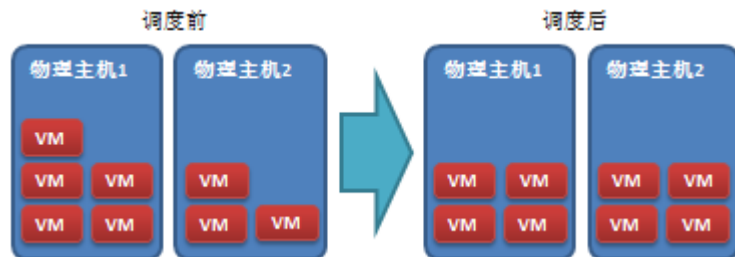
# 电信和企业关键应用领域虚拟网元在硬件平台间的迁移



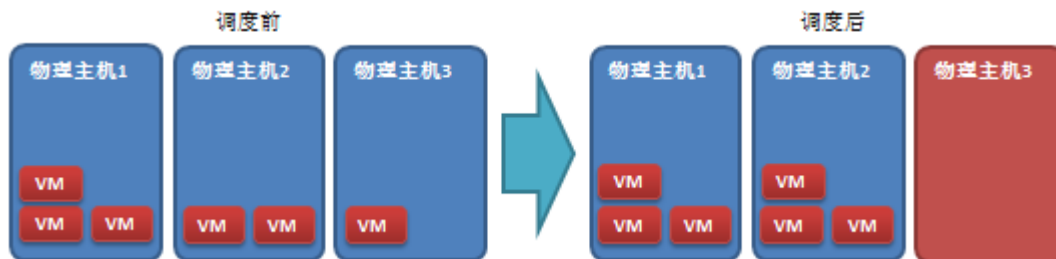
# 计算资源池的动态资源调度管理和动态能耗管理

- 动态资源调度管理，简称为DRS（Dynamic Resource Schedule）
- DRS周期性监控集群下物理主机和虚拟机负载（CPU和内存），如果主机负载的不均衡度（标准方差）超过集群配置的阈值，则触发虚拟机迁移，使得集群范围内的负载区域平衡。
- DPM（Distributed Power Management 分布式电源管理）支持在集群负载低（CPU和内存负载小于集群配置的阈值）时，主动对某些主机进行下电，达到节能减排的目的，同时在集群负载高（负载大于集群配置的阈值）时，重新启动一些主机，满足业务需要。

# DRS的几种基本的应用场景

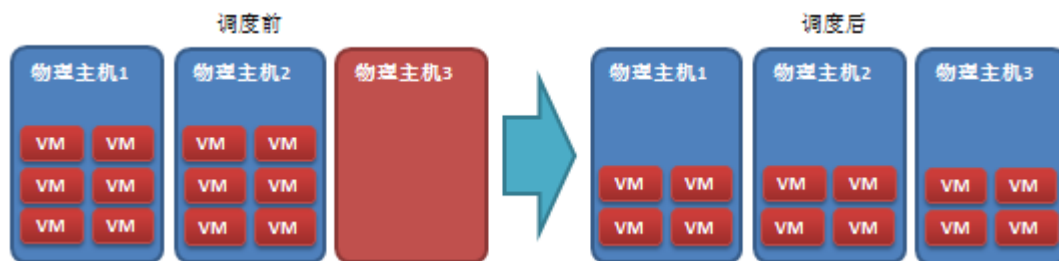


集群内主机间的负载不均衡

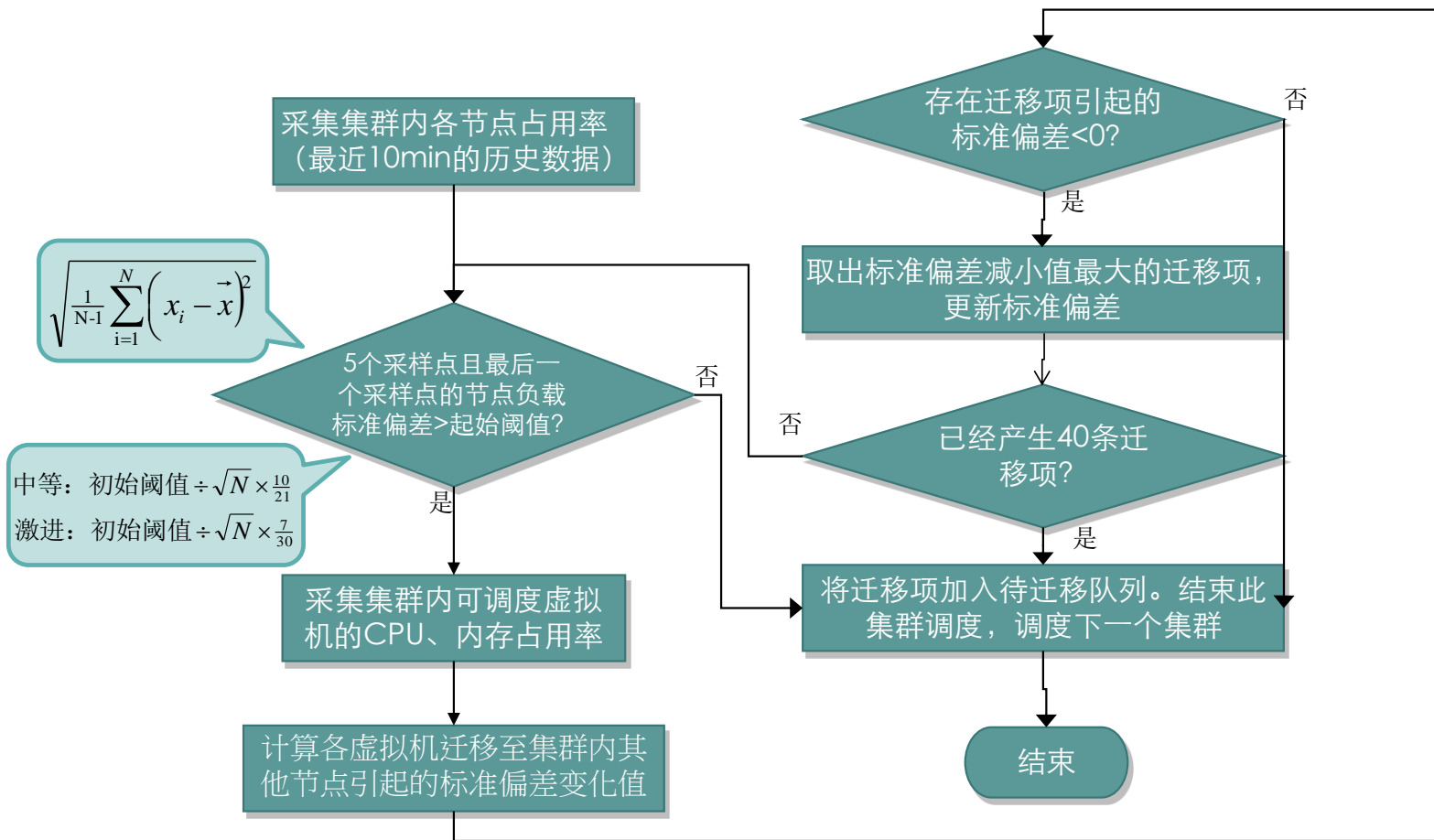


集群内主机的  
负载较低

集群内主机的负载均很高



# 一种动态负载调整的通用算法



# Local Regression算法

负载预测和动态资源调整的相互作用关系

## Local Regression算法

$$T(u) = \begin{cases} (1 - |u|^3)^3 & \text{if } |u| < 1, \\ 0 & \text{otherwise,} \end{cases}$$

$$w_i(x) = T\left(\frac{\Delta_i(x)}{\Delta_{(q)}(x)}\right),$$

计算权值,  $x$ 是时刻,  
 $\Delta(x)$ 是 $x$ 到当前的时间,  
越久远的数据, 权值越  
小, 大于 $q$ 的为0

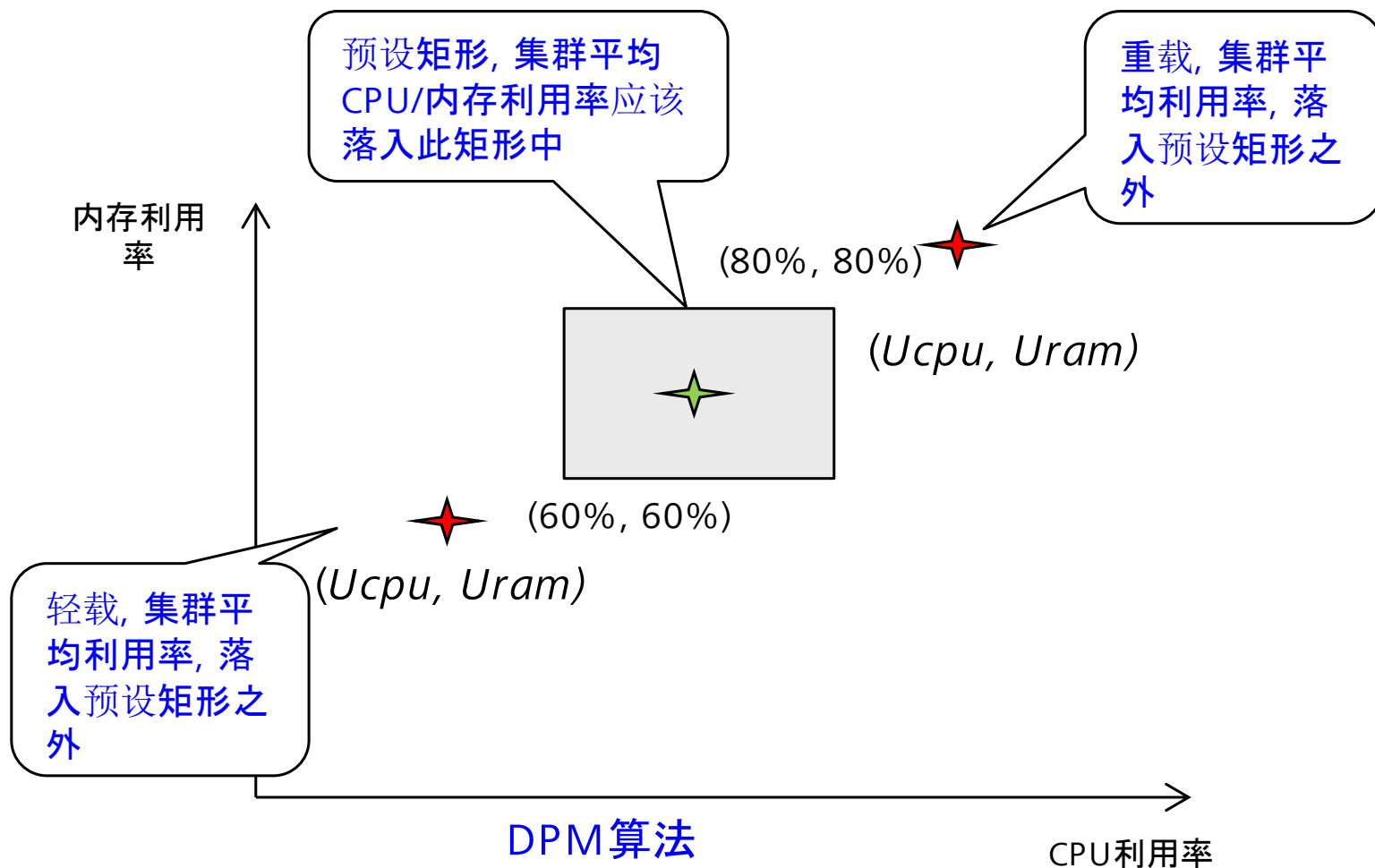
$$\sum_{i=1}^n w_i(x)(y_i - a - bx_i)^2.$$

利用最小二乘拟合 $a$ 和 $b$   
值,  $x$ 是历史时刻,  $y$ 是历  
史负载

$$y_{i+1} = a + bx_{i+1}$$

预测下一时刻 $x_{i+1}$ 的负载  
 $y$

# 节能降耗调度算法







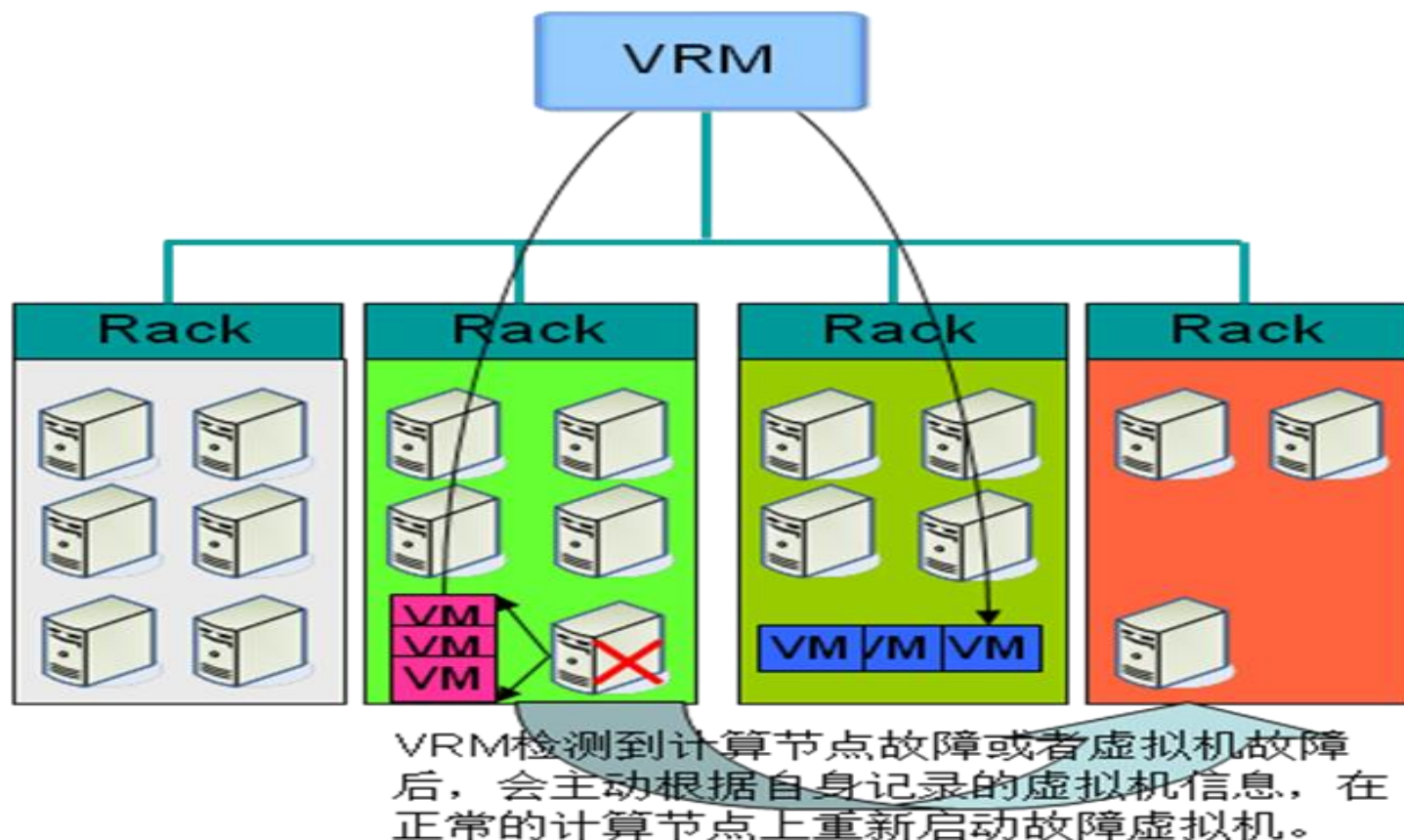
# 目 录

- 计算虚拟化核心引擎：Hypervisor介绍
- 跨服务器的计算资源调度算法
- 计算高可靠性保障

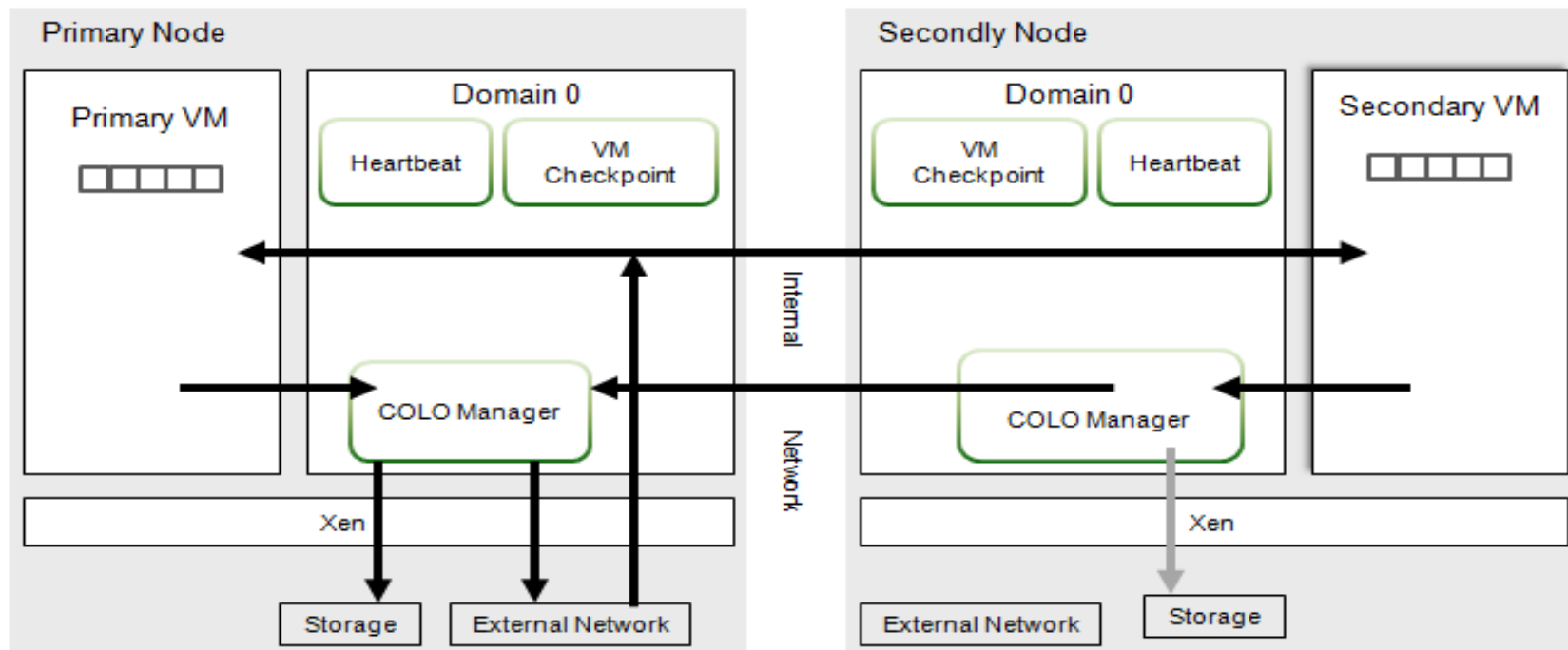
# 计算高可靠性保障

- 基于冷备机制的虚拟化HA保护，这种方式主要提供了在物理硬件故障的条件下，选择资源池中的其他健康物理主机重新部署虚拟机，在原有数据不丢失的条件下，尽快恢复虚拟机业务的方法；
- 基于虚拟机热备机制的虚拟机运行态镜像冗余方案，这种方式主要是在不同的物理主机上，提供虚拟机业务运行的镜像冗余，在一台物理主机发生故障的时候，自动由另外一台主机上的虚拟机业务运行镜像接管进行业务处理，从而保证虚拟机的业务不发生中断。

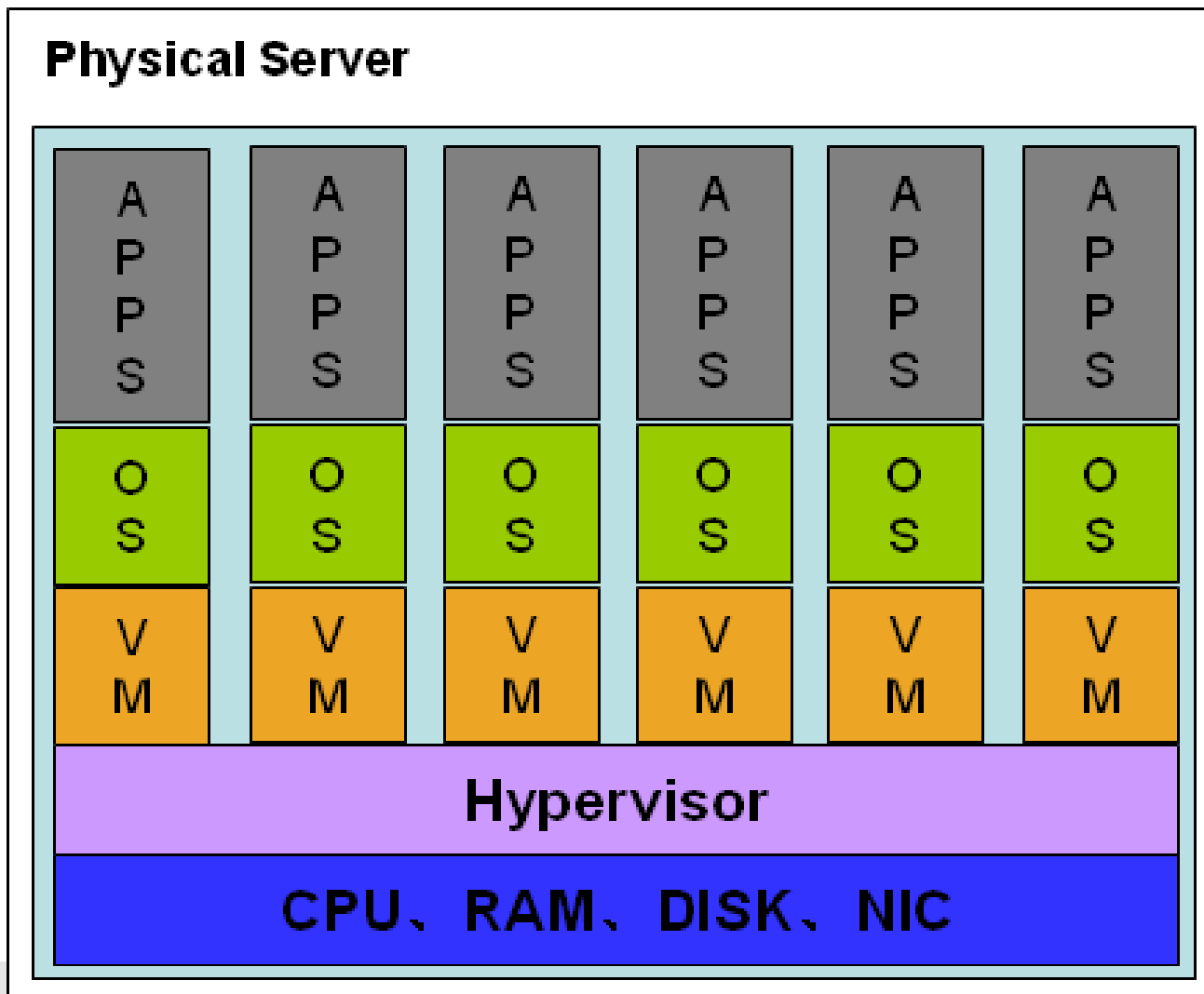
# 基于冷备机制的虚拟机HA保护



# 基于热备机制的虚拟机运行业务镜像冗余



# 无状态计算及物理机可靠性保障



## 思考？

- 1 CPU调度技术采用什么方法？
- 2、动态资源调度管理的原理是什么？



# 总 结

- 1、计算虚拟化的方式
- 2、负载均衡的调度算法
- 3、高可用计算能力的保障

谢谢

[www.huawei.com](http://www.huawei.com)