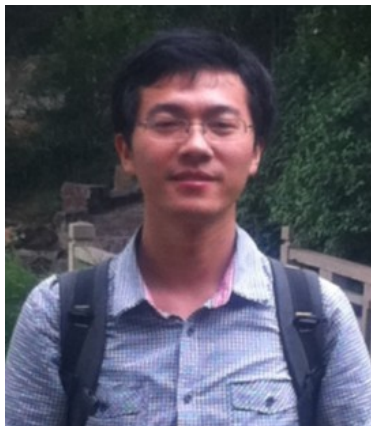


盛大云平台架构设计和实现

SNDR 盛大网络

杜 海

盛大云计算 架构师
duhai2008@hotmail.com



杜海

2010-至今 盛大创新院 云计算架构师 高级研究员

2008-2010 百度 虚拟化平台研发

2005-2008 虚拟化技术研究

毕业于复旦大学，关注linux内核，Xen，分布式存储，IaaS等主题，热衷于虚拟化技术的研究和企业级应用。

- 1. 盛大云平台综述
- 2. 盛大云主机架构
- 3. 分布式镜像系统
- 4. 网络安全系统
- 5. 资源调度系统
- 6. 实践中的问题

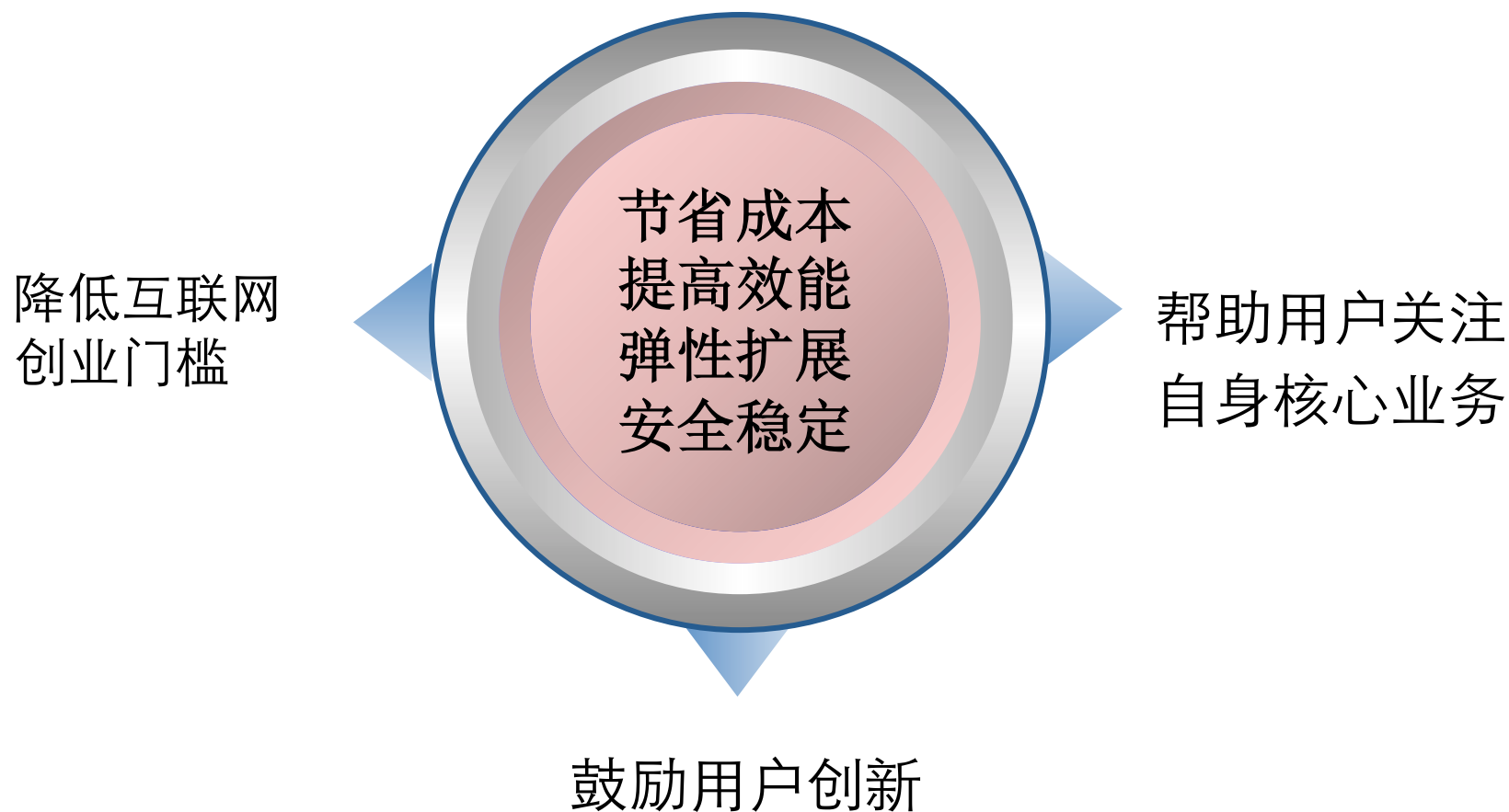
■ 公有云平台

- 提供基础资源即服务（IAAS）

■ 类似国际领先的亚马逊AWS的服务模式

- 被证明成功的商业模式

We are in the business of selling infrastructures





- 互联网创业团队和个人
- 物联网企业
- 政府、科研、教育机构
- 国外公司在中国的分部



[首页](#)[产品](#)[支持](#)[社区](#)[案例](#)[账户](#)

盛大云 Grand Cloud

- 云主机
- 云存储
- 云硬盘
- 云监控
- 数字分发
- MongoIC



免费注册盛大通行证

立即注册



产品入口 Products

[云主机](#)[云存储](#)[云硬盘](#)[云监控](#)[数字分发](#)[MongoIC](#)

新闻公告

媒体报道

近期活动

MORE

- ▶ 2011-10-21 "盛大云"MongoIC上线全球首次支持数据库恢
- ▶ 2011-10-20 盛大云—MongoIC 上线公告
- ▶ 2011-10-18 云计算论坛盛大开启 专家解读行业热点
- ▶ 2011-10-13 盛大云邀请码B类升级A类功能上线公告
- ▶ 2011-10-11 盛大云邀请码A类和B类
- ▶ 2011-10-11 "盛大云" 系统全面升级 双线云主机上线
- ▶ 2011-09-29 "盛大云" 荣获2011中国云计算最佳产品奖
- ▶ 2011-09-28 盛大云 - 双线云主机上线公告

云计算优势 Advantage

采用盛大云搭建应用, 您将具有如下优

■ 2011.7.22 – 盛大云(Grand Cloud)平台公测

- 云主机(Compute Cloud)
- 云硬盘(Elastic Block Service)
- 云存储(Cloud Storage)
- 数字分发(Digital Delivery System)
- 云监控(Cloud Monitoring Service)
- 云数据库(Mongo in Cloud)

自主开发的拥有100%知识产权的国产云平台



1. 盛大云平台综述

■ 2. 盛大云主机架构

3. 分布式镜像系统

4. 网络安全系统

5. 资源调度系统

6. 实践中的问题

公有云平台面临的挑战

■ 部署规模&效率

- 盛大镜像系统

■ 安全

- 网络安全系统
- 云安全

■ 资源管理

- 资源调度系统

■ 数据可靠性

- 云硬盘
- 云存储

■ 镜像系统

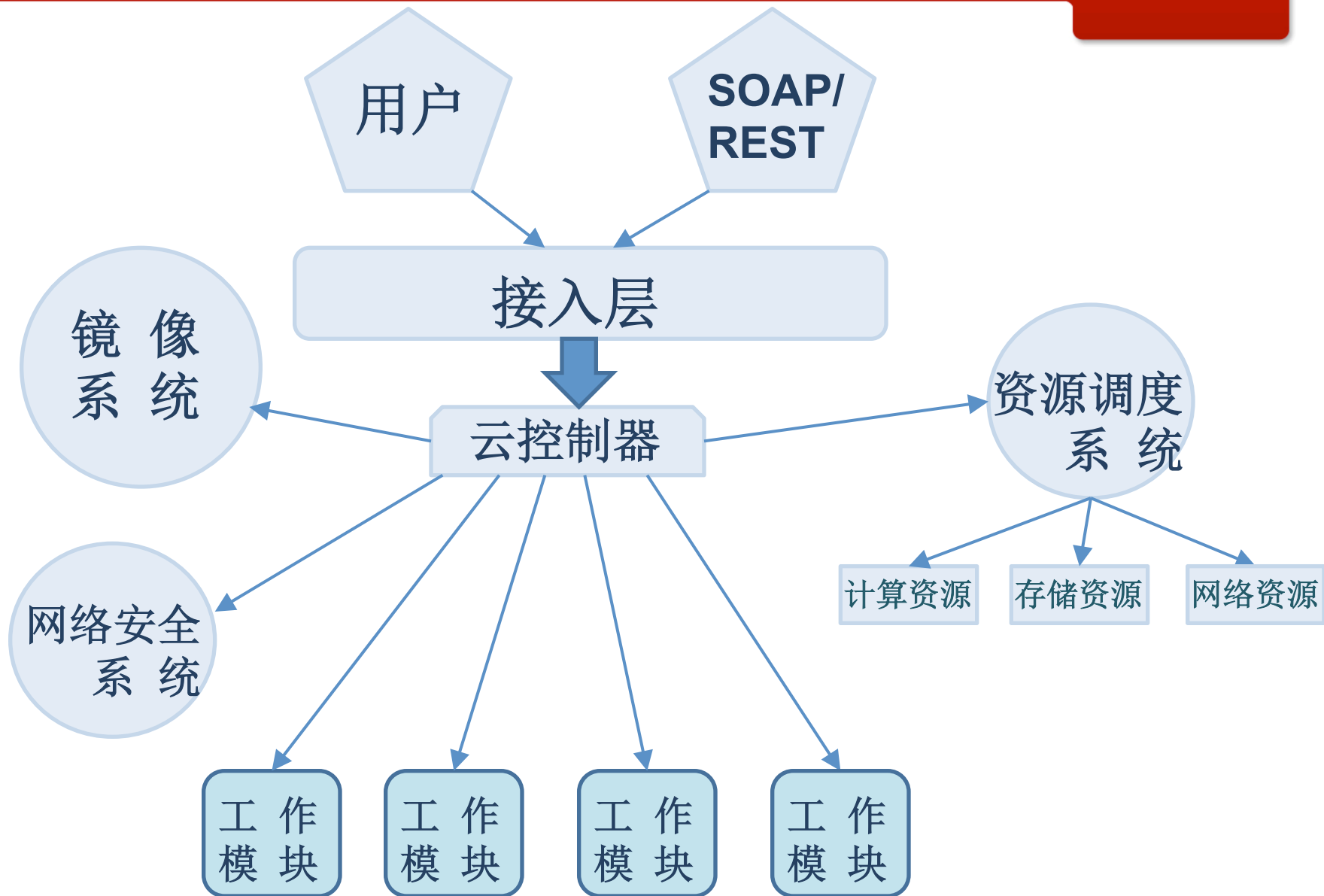
- 大规模并发条件下快速创建云主机
- 云主机的备份和恢复

■ 网络安全系统

- 用户组隔离

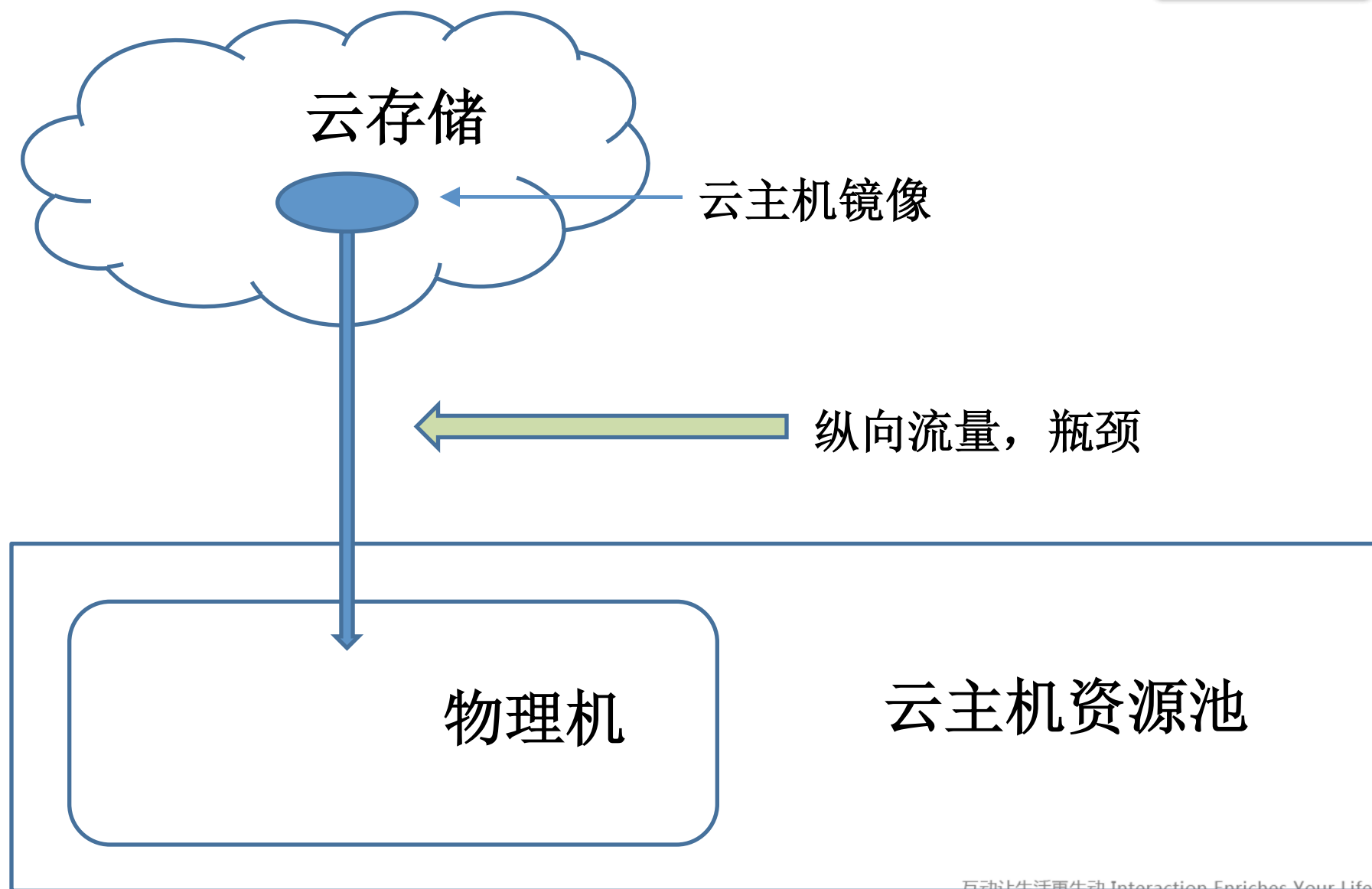
■ 调度系统

- 池资源分配---计算资源、存储资源、网络资源
- 性能控制和隔离

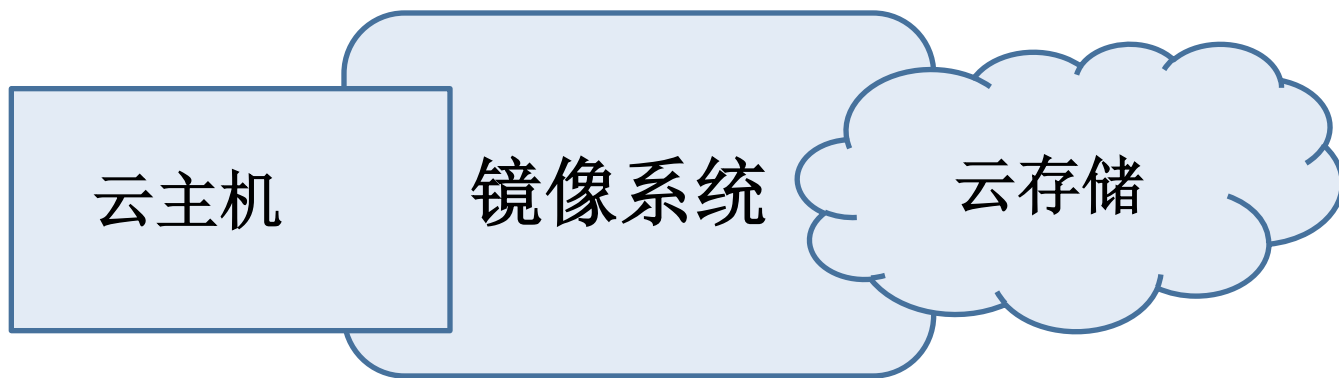


1. 盛大云平台综述
2. 盛大云主机架构
- 3. 分布式镜像系统
4. 网络安全系统
5. 资源调度系统
6. 实践中的问题

云主机磁盘镜像生成过程



- 介于计算资源池和存储资源池之间
- 每个计算节点同时划出一部分资源作为镜像系统的存储节点



■ 设计目的

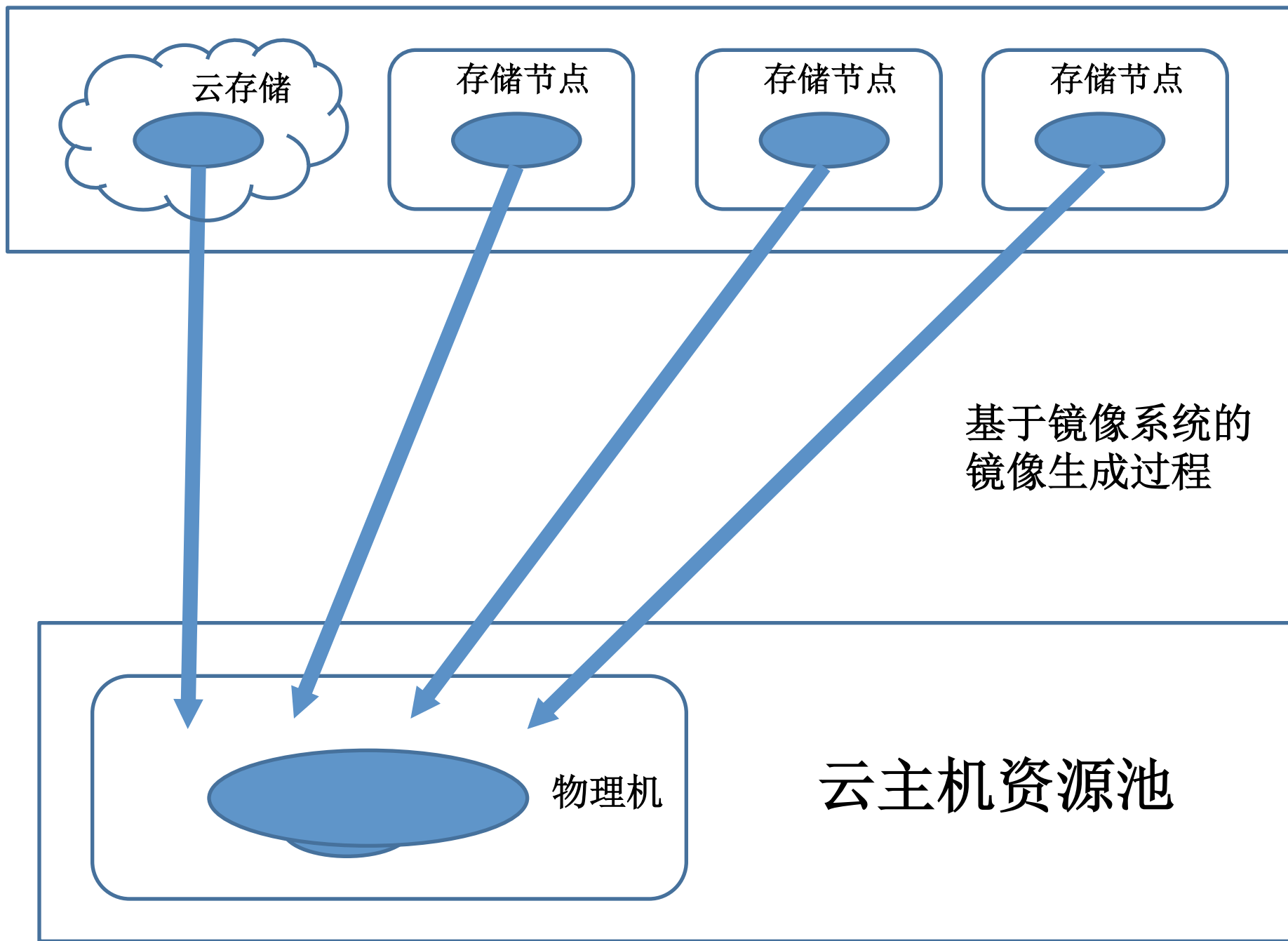
- 支持大规模并发、高效的云主机部署
- 云主机的备份，克隆
- 云主机的快速恢复

■ 实现

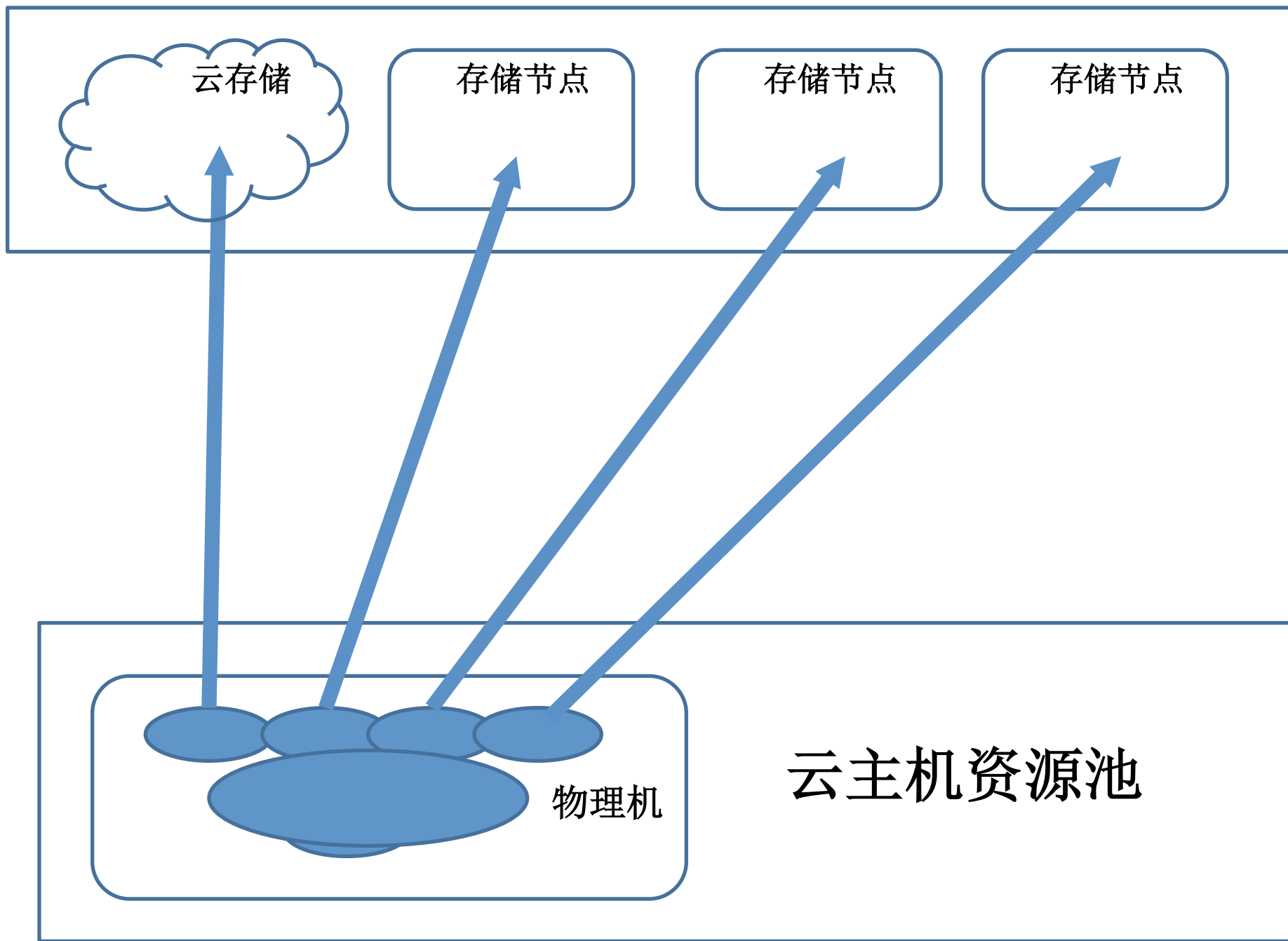
- 基于Key-Value的定制化分布式存储系统

■ 跟云存储的区别

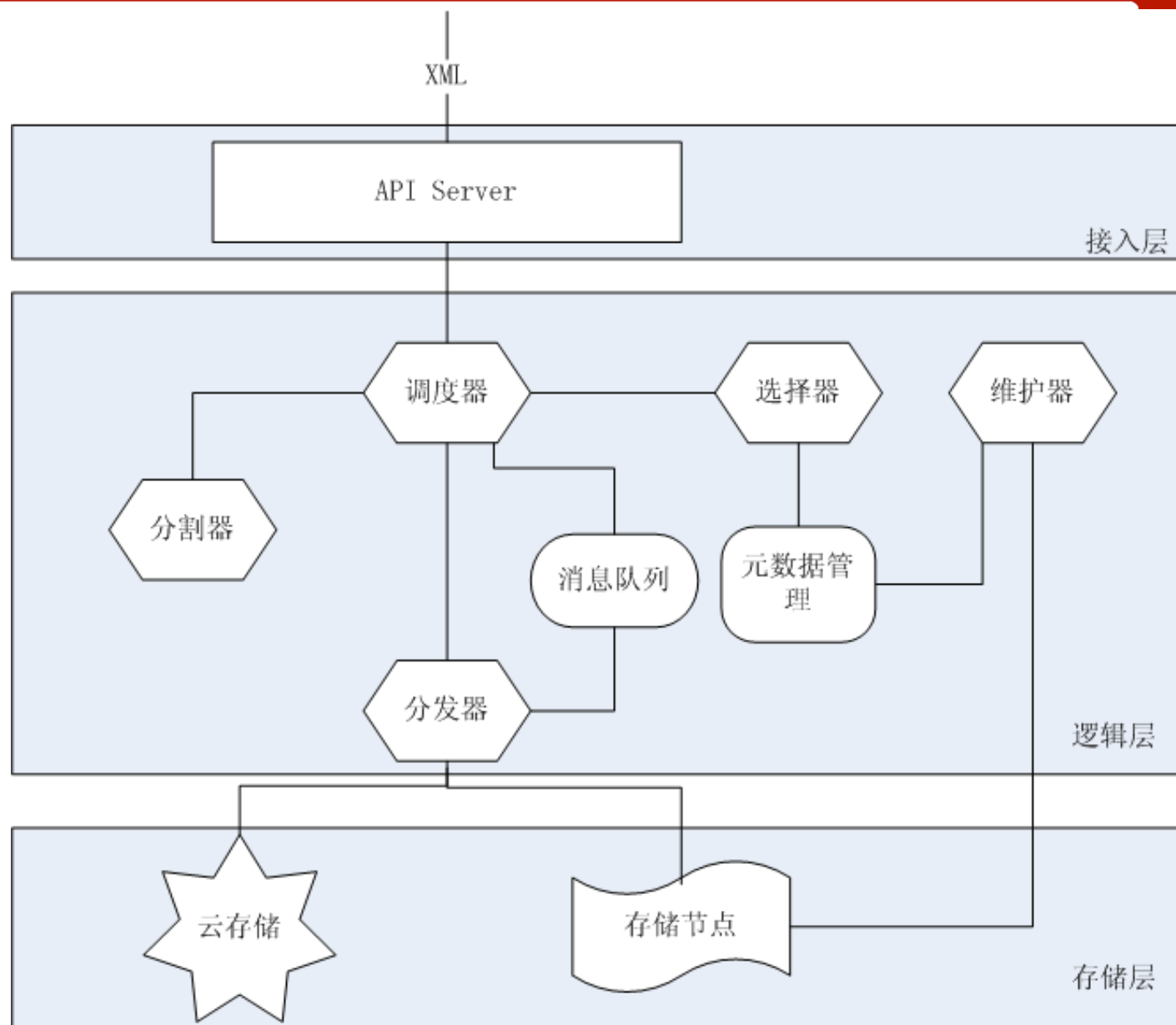
- 云存储：可靠性；海量数据；
- 镜像系统：效率；备份功能；



- 校验
- 切片
- 元数据调度
- 分发



系统架构



■ Key-value

- Key: 文件名+CRC32检验值

■ 元数据写: 9份

■ 元数据读: 5份, 判断时间戳

■ 镜像缓存备份数动态调整

- 镜像热度：前5分钟内下载次数
- 高热度镜像：扩大备份数
- 存储空间满：减少最冷镜像备份数；LRU算法；

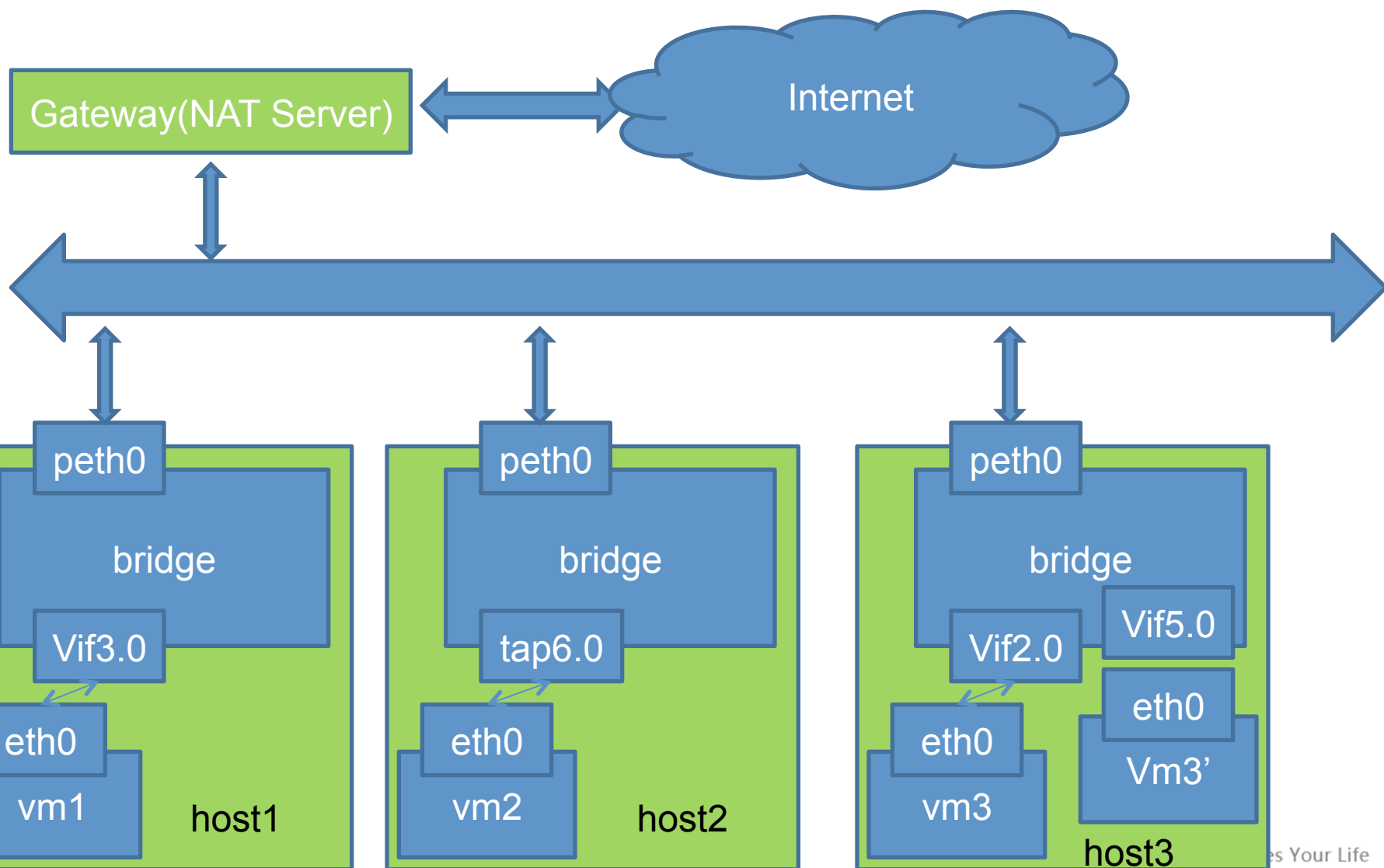
1. 盛大云平台综述
2. 盛大云主机架构
3. 分布式镜像系统
- 4. 网络安全系统
5. 资源调度系统
6. 实践中的问题

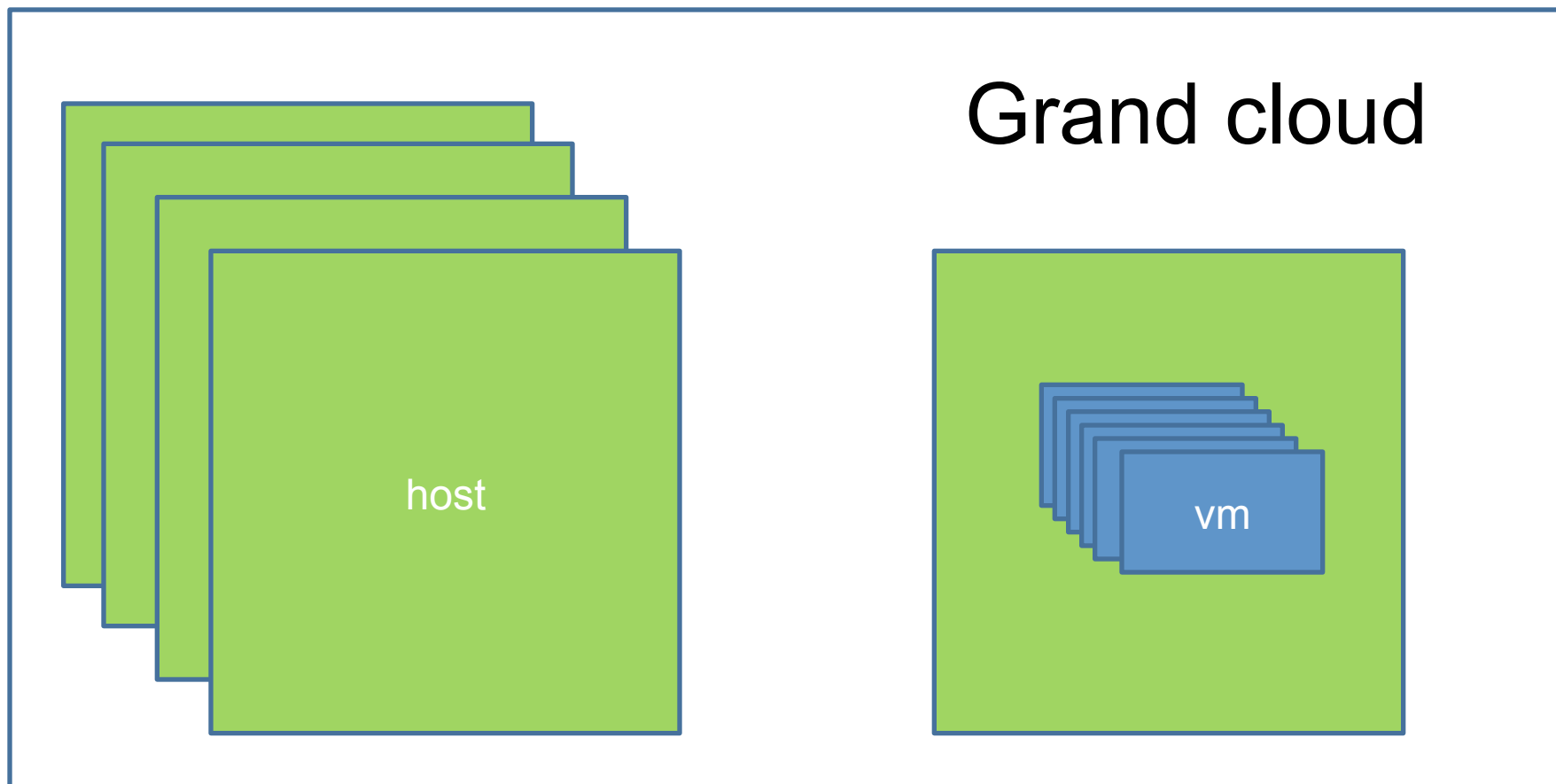
- 保证云主机安全

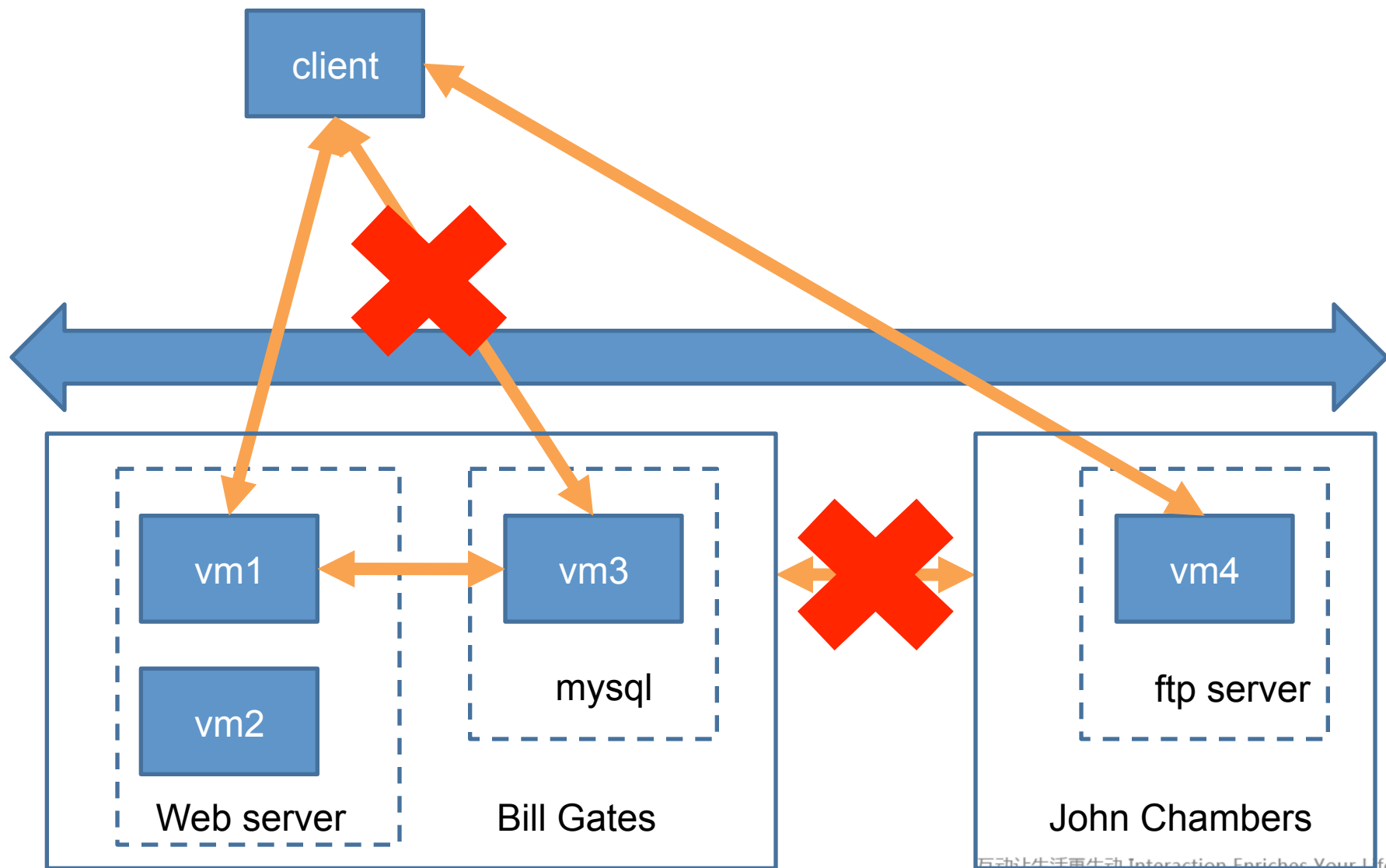
- 端口访问控制

- 集群权限控制

当前盛大云网络体系结构







■ 限制云主机可以发送的包内容

- 禁止云主机发送修改过MAC或IP的包

■ ARP攻击防御

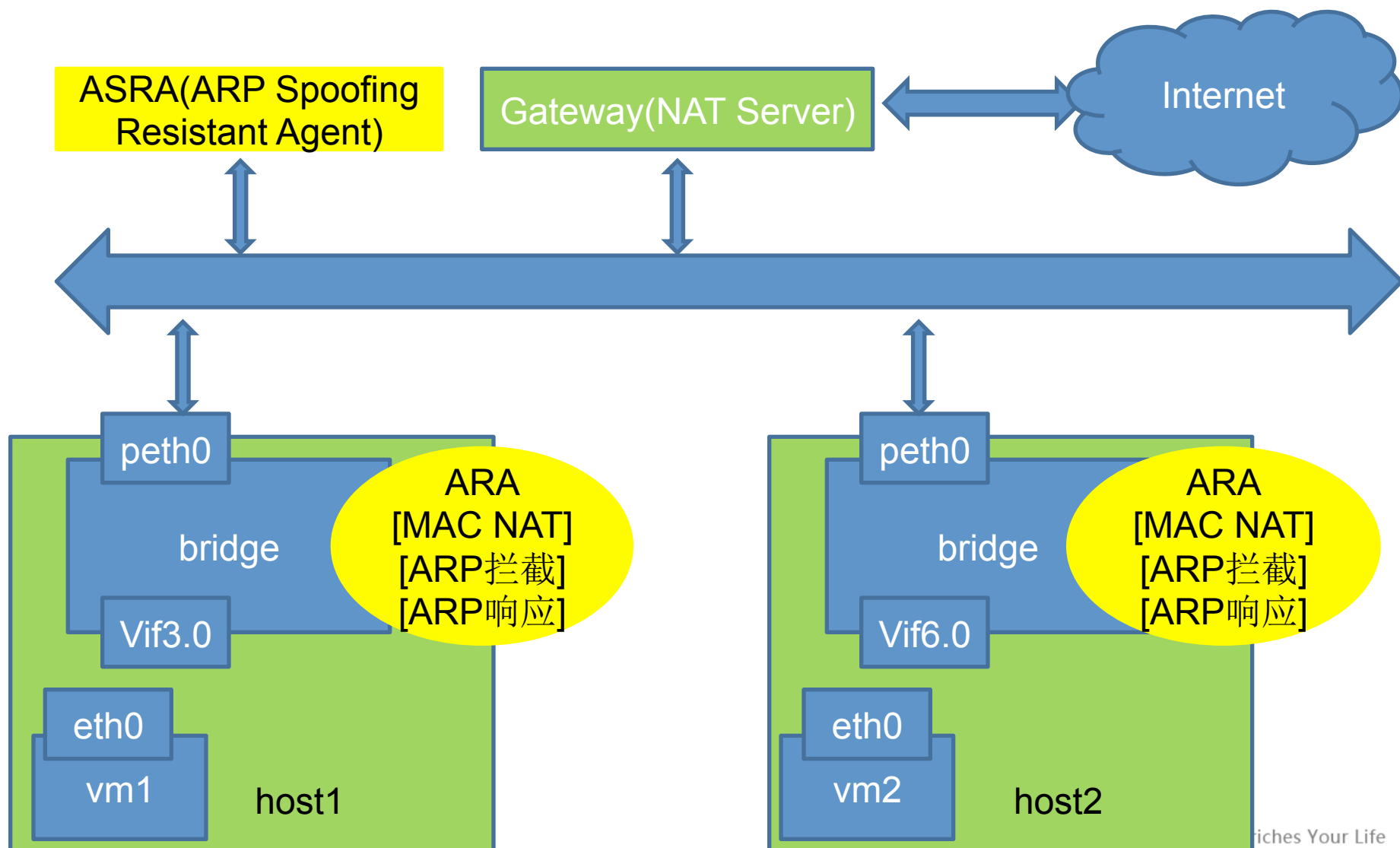
■ 以太网上出现的MAC:

- » Host1Peth0, Host2Peth0, 所有host
- » ~~vm1...~~, ~~vm2...~~, 所有云主机
- » GW

■ VM看到的MAC—只有一个(mac masquerade):

- » FE:FF:FF:FF:FF:FF

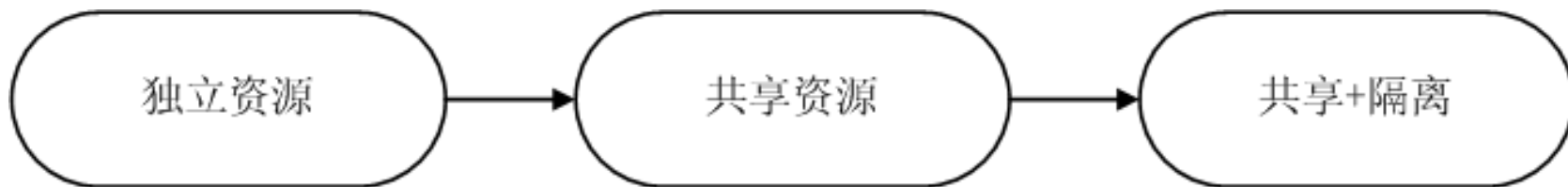
ARP 攻击防御



1. 盛大云平台综述
2. 盛大云主机架构
3. 分布式镜像系统
4. 网络安全系统
- 5. 资源调度系统
6. 实践中的问题

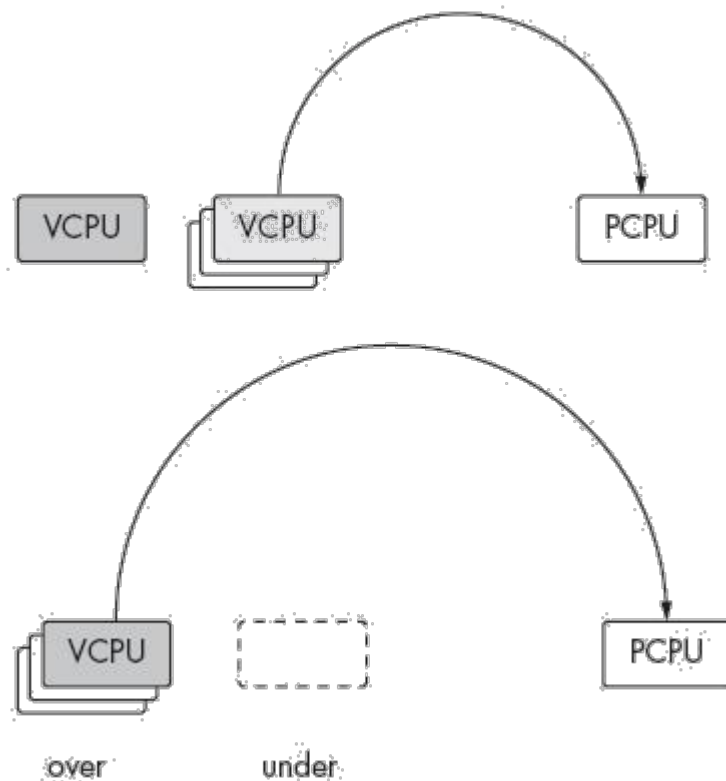
■ 资源池

- 隔离性
- 公平性
- 稳定性
- 高性能

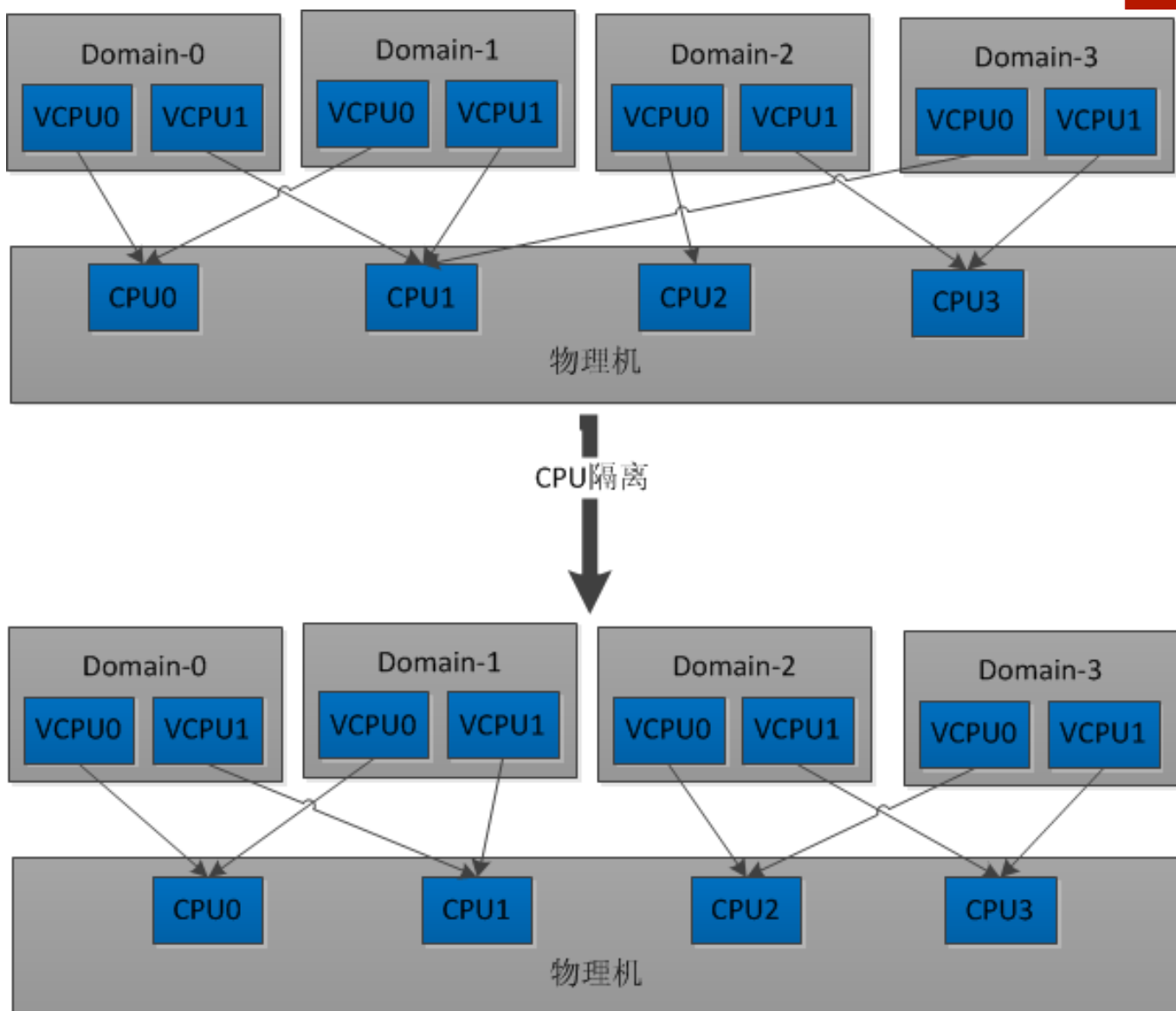


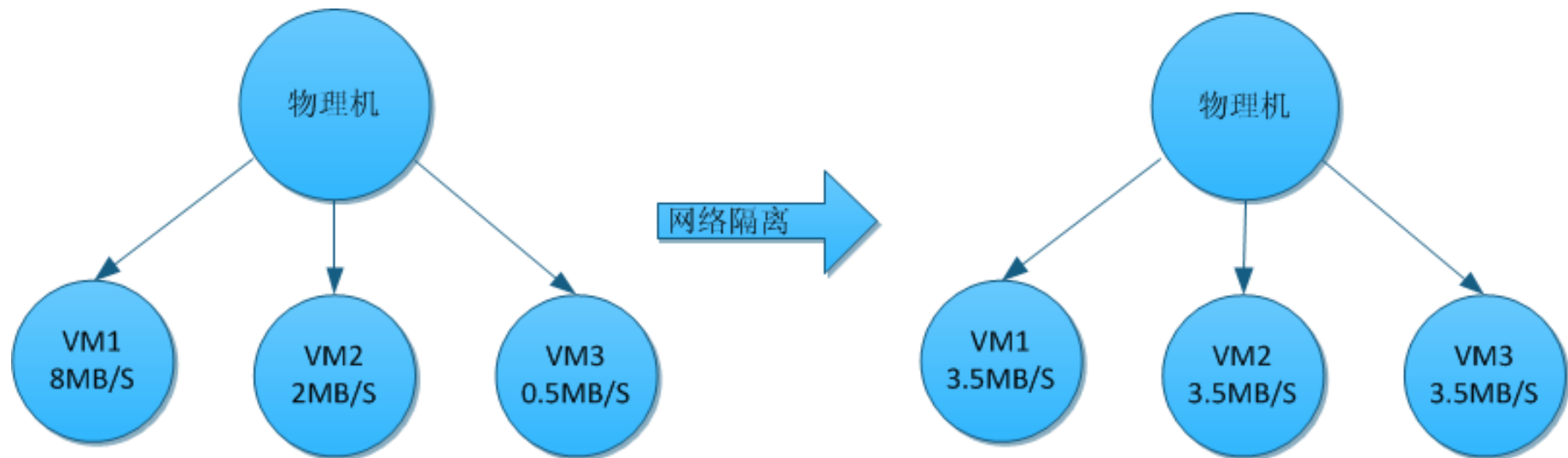
- CPU隔离
- 网络隔离
- 磁盘I/O隔离

- 为所有虚拟机的VCPU绑定物理CPU
- 为Domain-0绑定物理CPU
- 为所有VCPU设定权值，以便多个VCPU共享一个物理CPU时可以公平调度

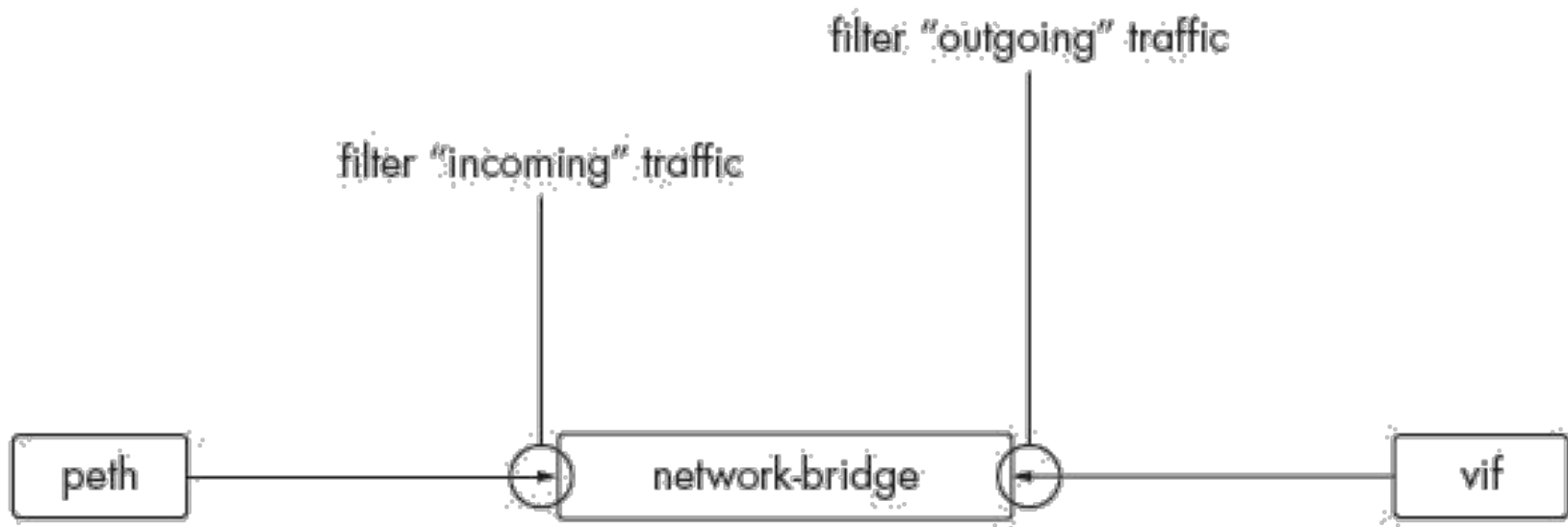


CPU资源隔离示意图

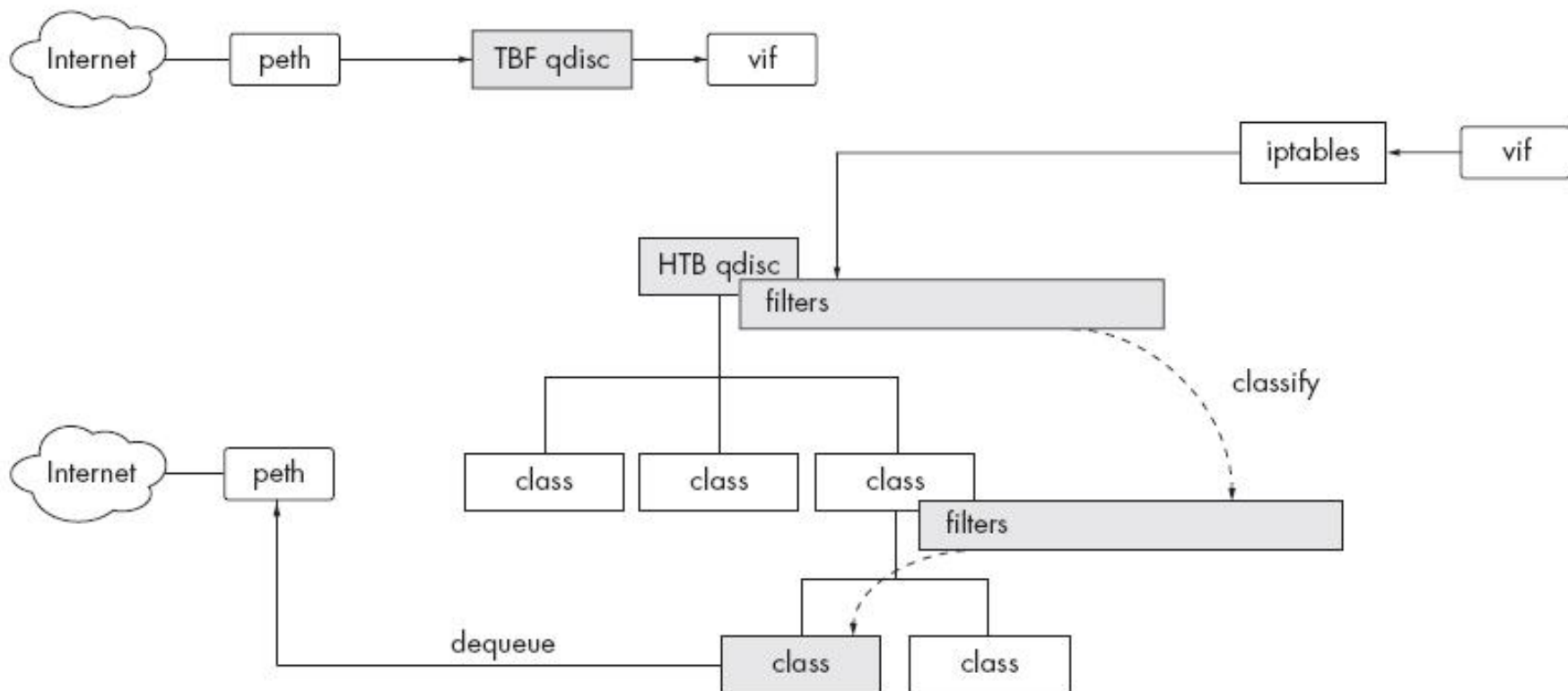




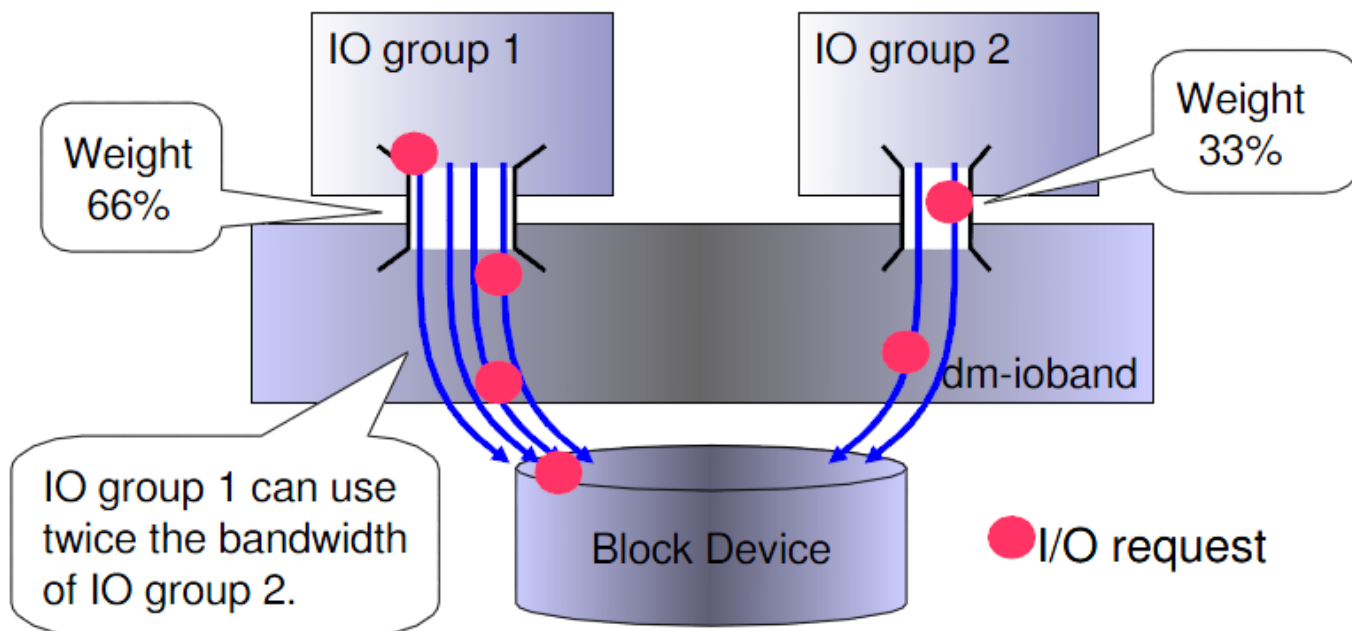
- 使用iptables+tc进行网络隔离
- 入口流量：一般不作限制，但也可以通过策略限制
- 出口流量：对网桥设备使用HTB控制。



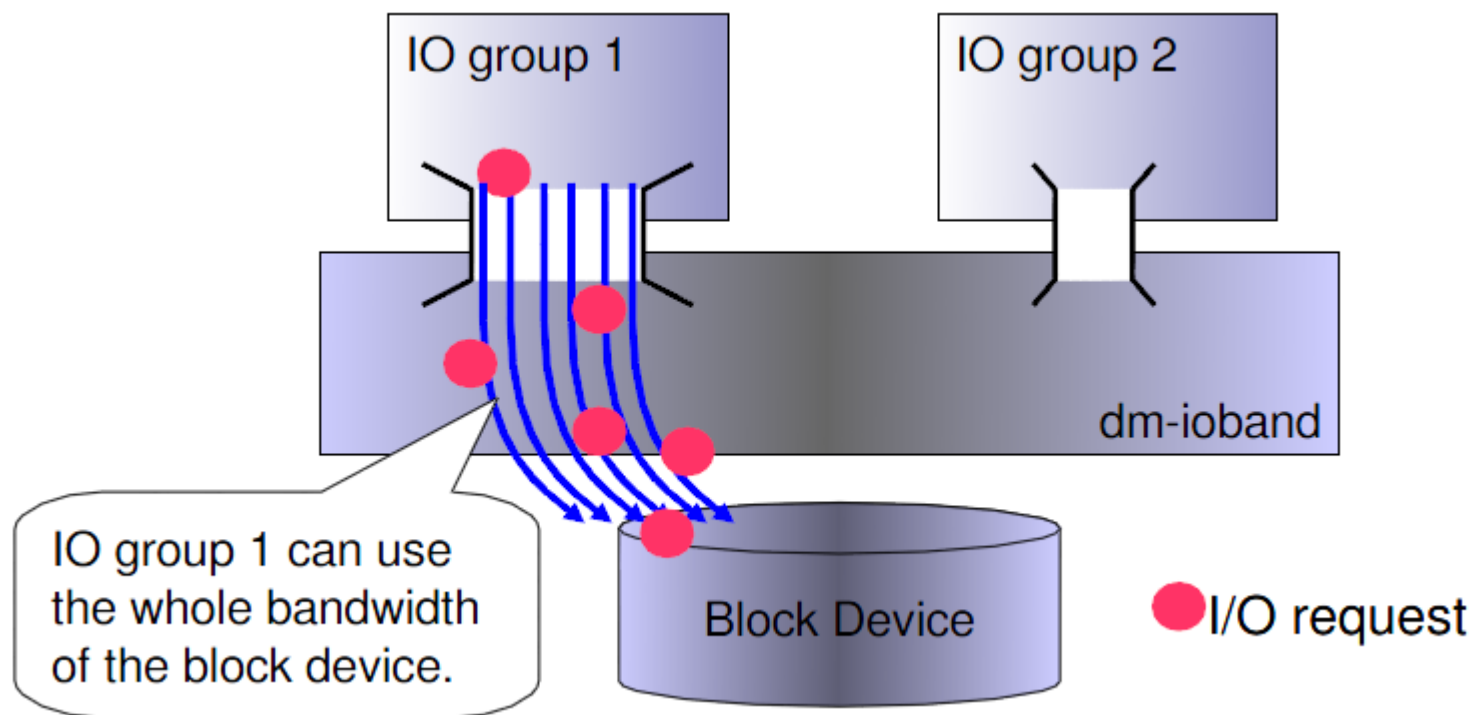
■ 网络控制图示



- dm-ioband
- 不设置绝对上限
- 只设置比例



- 非活跃组的带宽被活跃组使用



1. 盛大云平台综述
2. 盛大云主机架构
3. 分布式镜像系统
4. 网络安全系统
5. 资源调度系统
- 6. 实践中的问题

功能问题（简单）

- 双线
- Etc.

技术问题（困难）

- 性能隔离
- 安全体系
- 大规模快速部署
- Etc.

其他问题（不定）

- 域名备案
- Etc.

谢谢!

Q & A



北京站 · 2012年4月18~20日
www.qconbeijing.com (11月启动)

QCon杭州站官网和资料
www.qconhangzhou.com

全球企业开发大会

INTERNATIONAL
SOFTWARE DEVELOPMENT
CONFERENCE