

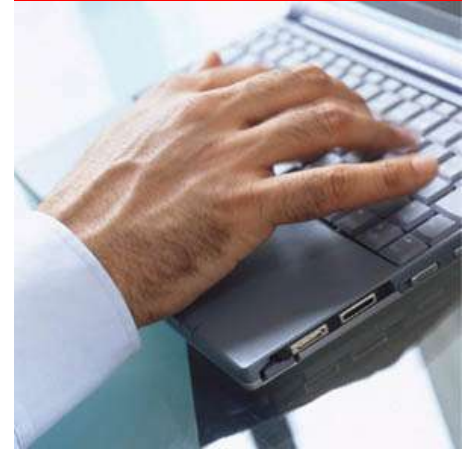


ORACLE®

Oracle Real Application Clusters Overview

Agenda

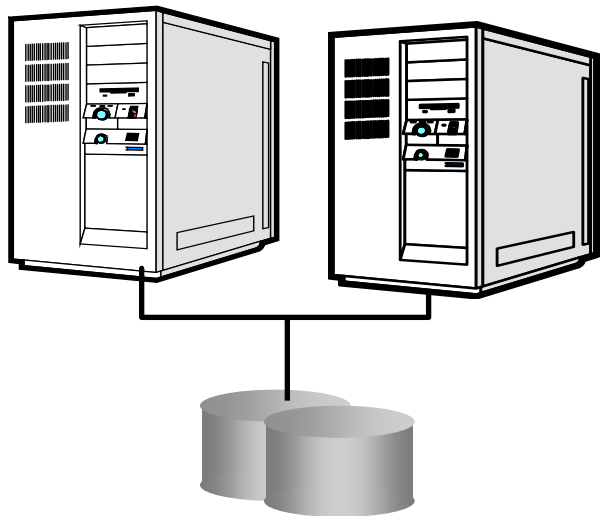
- 什么是RAC?
- RAC的存储管理
- RAC的可扩展性
- RAC的高可用性
- Oracle集群软件与RAC的集成
- 网格计算的基础架构
- RAC案例



RAC真正应用集群的比喻

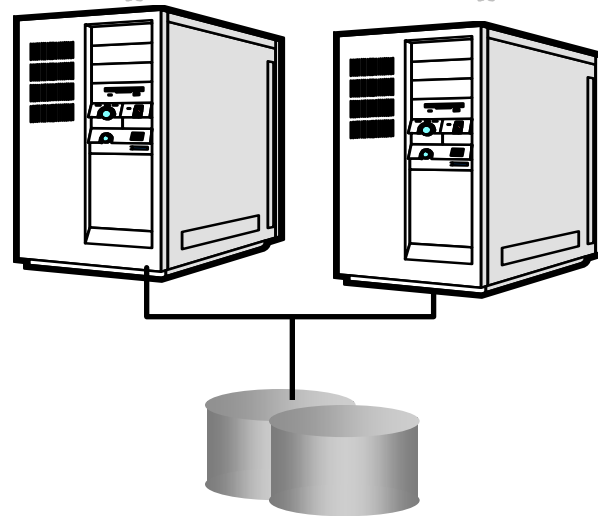
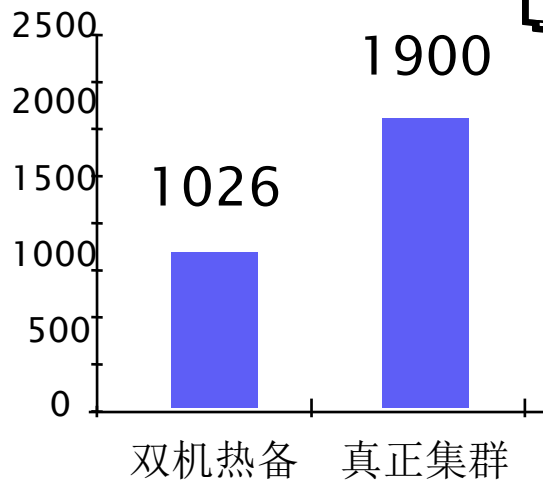


为什么要使用真正应用集群



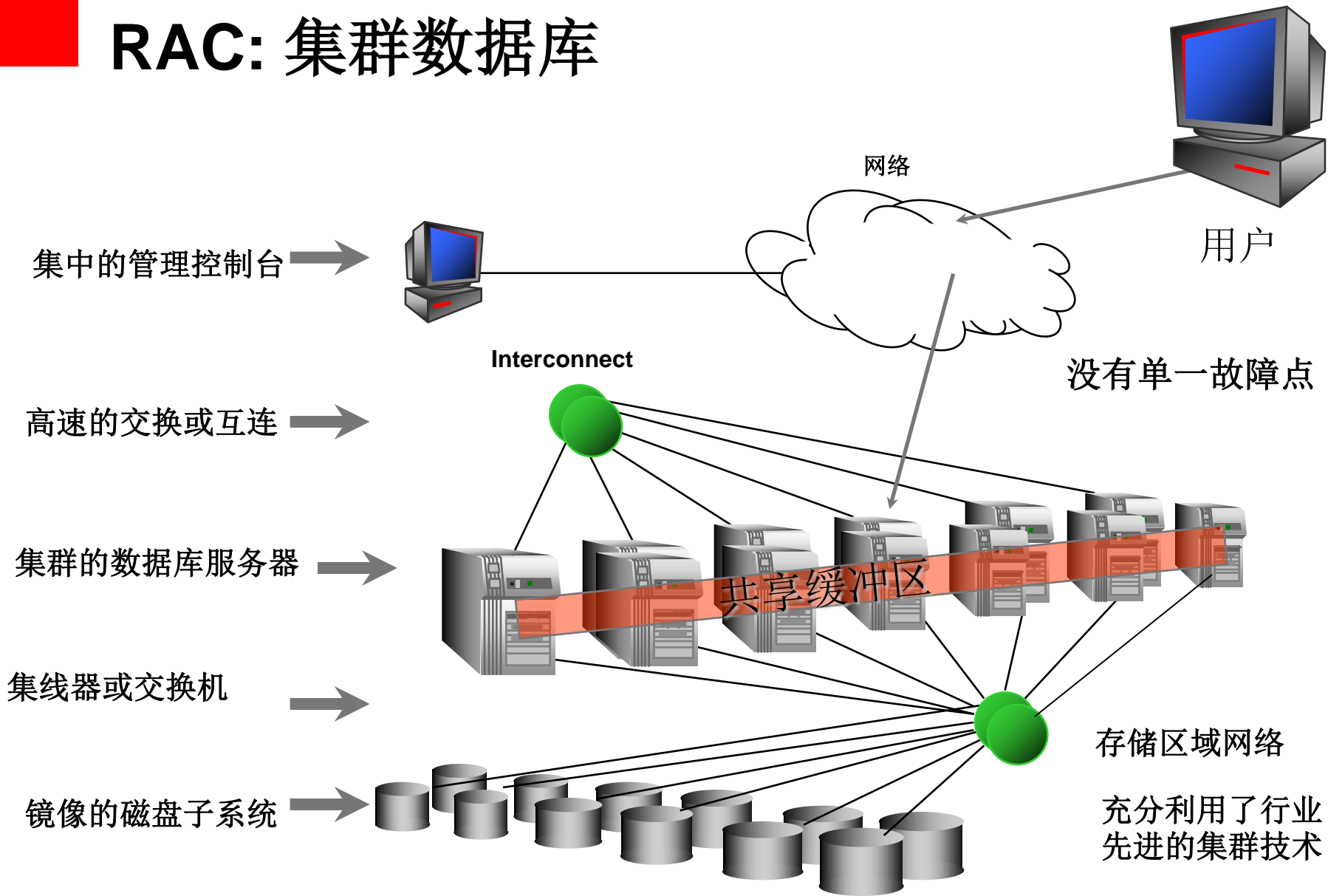
双机热备

用户数

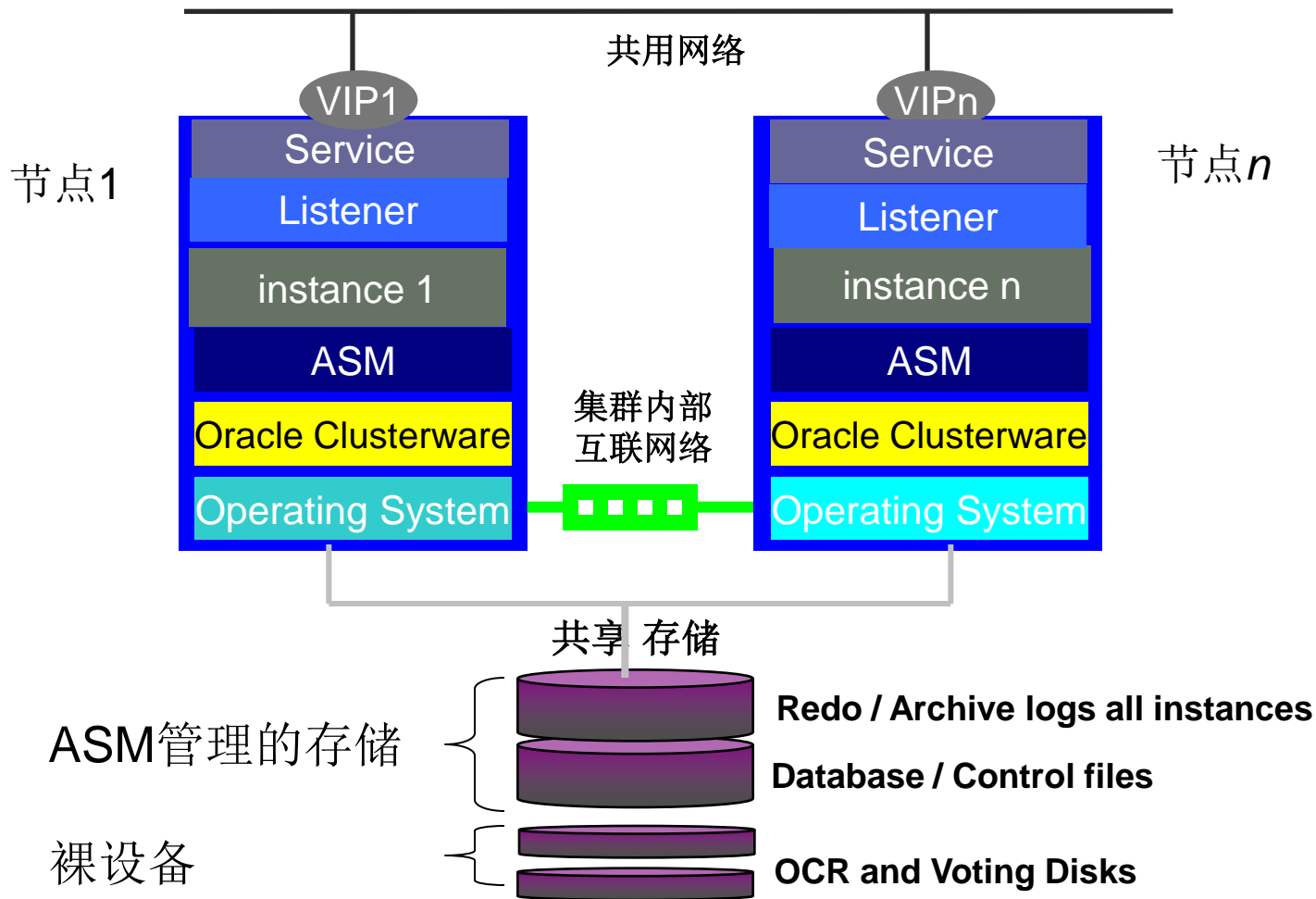


真正集群

RAC: 集群数据库



RAC架构



自动存储管理

11g新功能

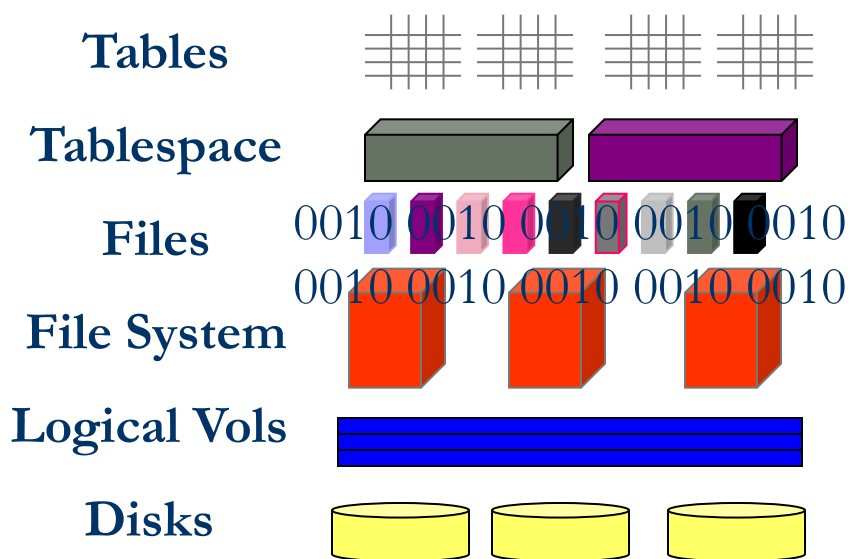
- **快速磁盘再同步** – 当磁盘offline时只对变化进行跟踪，当磁盘online时只对相关变化进行同步
- **快速镜像读** – 自动找到最近的镜像磁盘进行读取，提高读取速度
- **超大数据库支持** – 对大数据文件，自动使用大扩展段
- **滚动升级** – 一个或多个在ASM集群中的instance可以进行独立滚动升级

RAC的存储管理

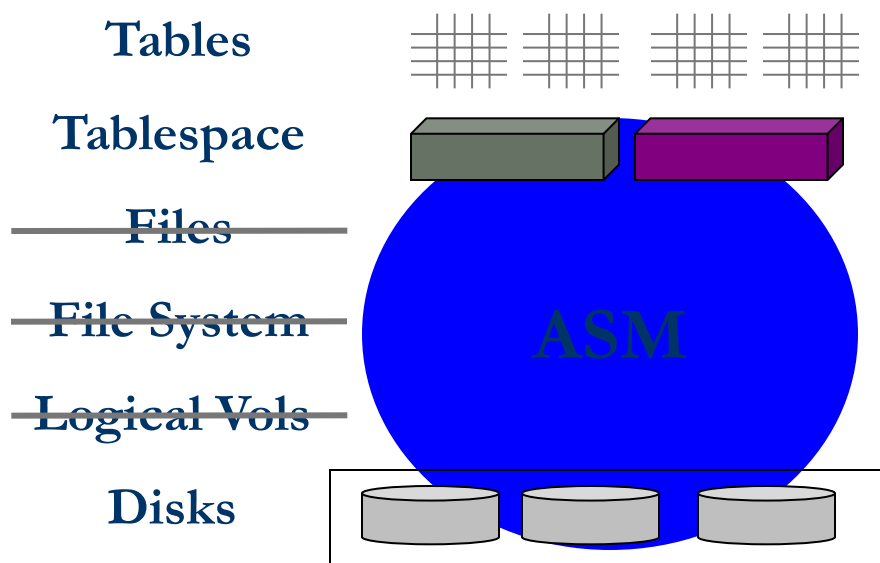


自动的存储管理 (ASM)

传统的方法



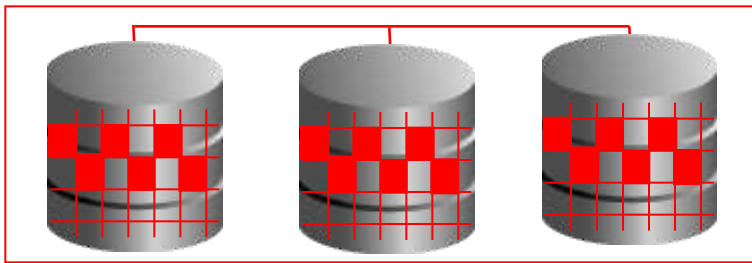
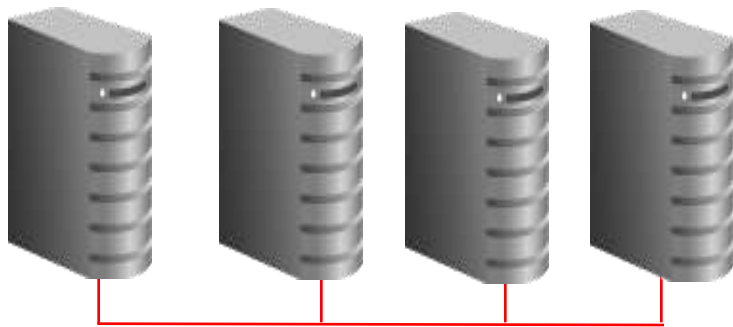
Oracle ASM



降低管理成本的最佳方法就是降低复杂性

自动存储管理

Automatic Storage Management



- 不需要文件系统和卷管理软件
- 按需分配容量
 - 在线添加删除磁盘
- 自动的IO负载均衡
 - 数据条带化
 - 最大化IO输出
- 自动镜像
- 简化管理

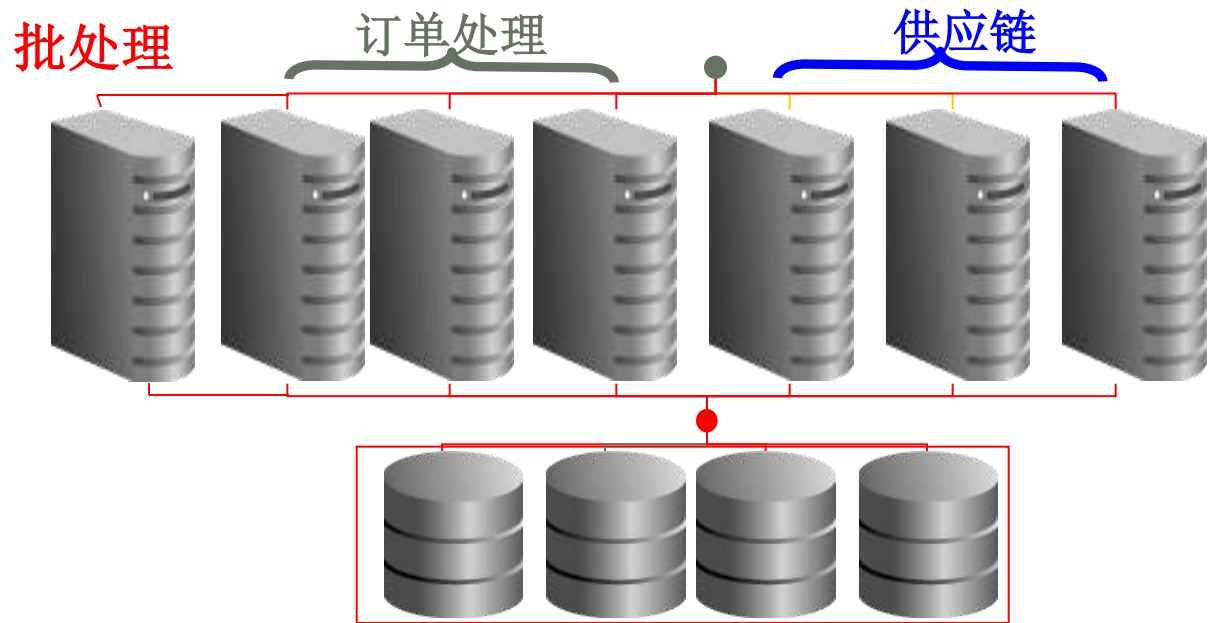
RAC的可扩展性



Real Application Clusters

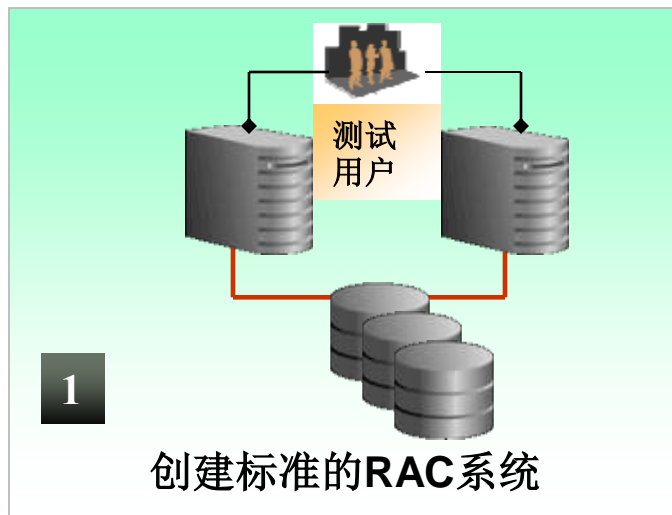
灵活的可扩展性

- 应用程序可在各节点间实现负载均衡
- 随着负载增加，可通过添加新节点扩展系统的处理能力
- 11g中缓存融合的性能提高

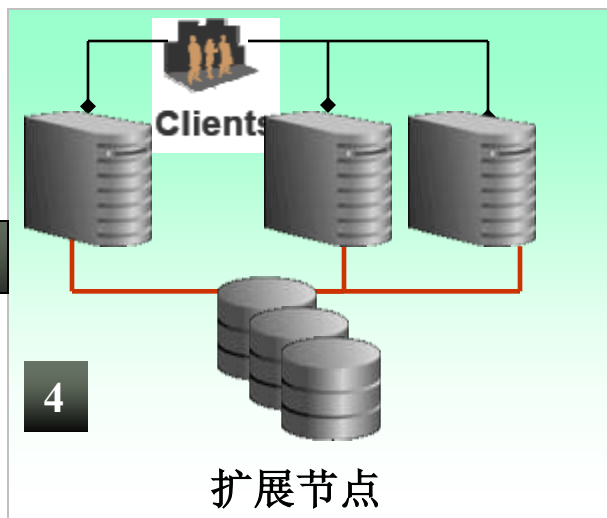
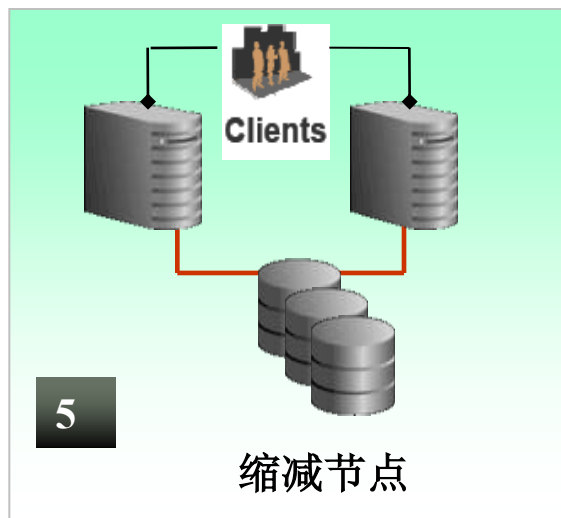


RAC部署周期

测试阶段



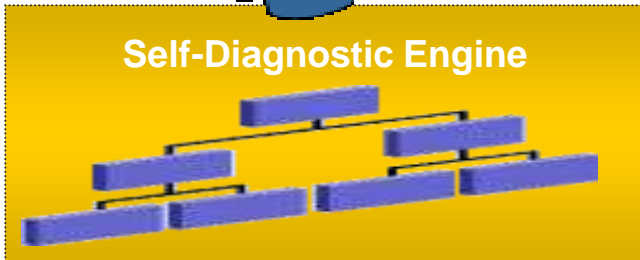
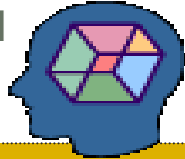
Production



ADDM for RAC

11g新功能

数据库级别
ADDM
11g



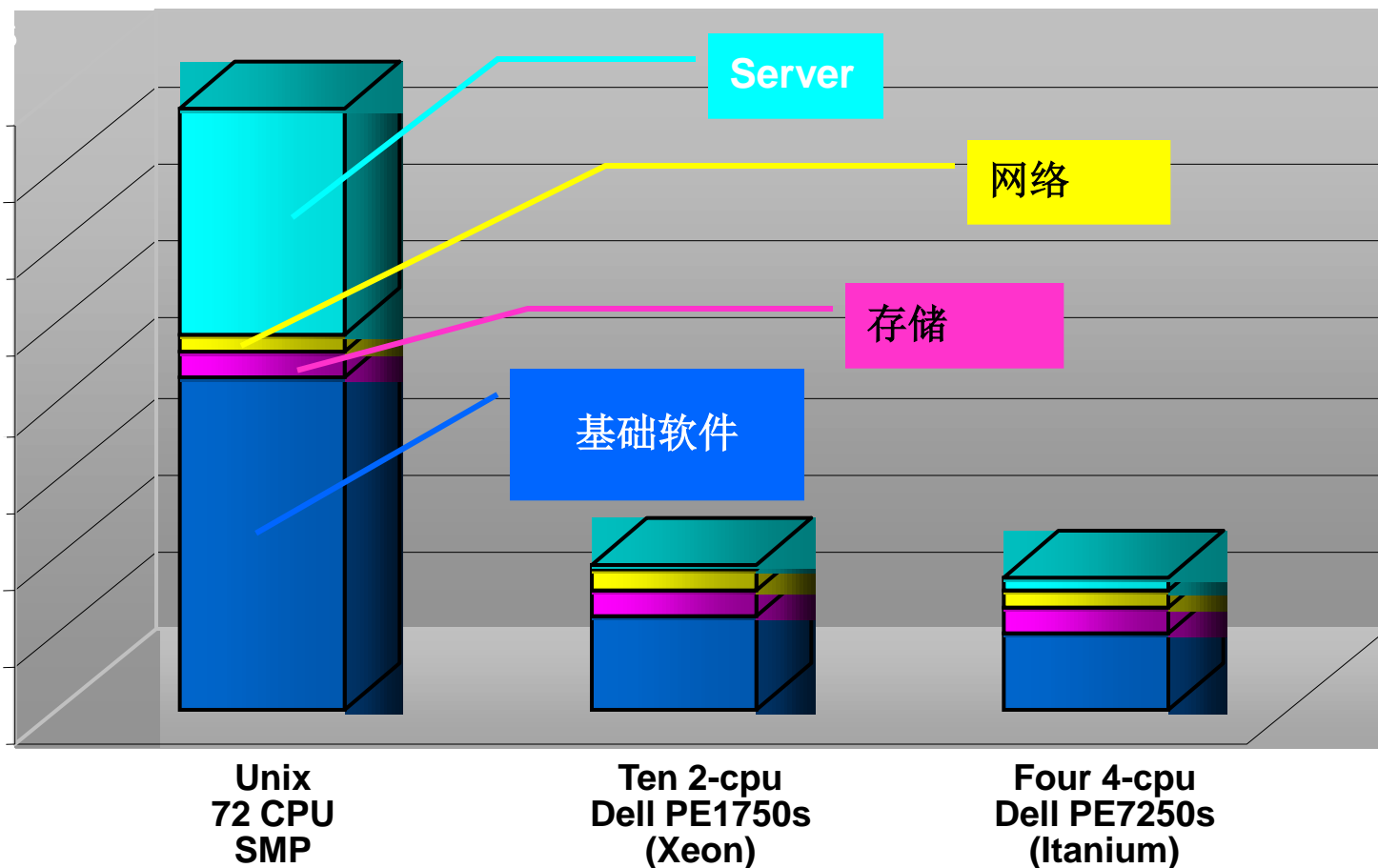
实例级别 ADDM



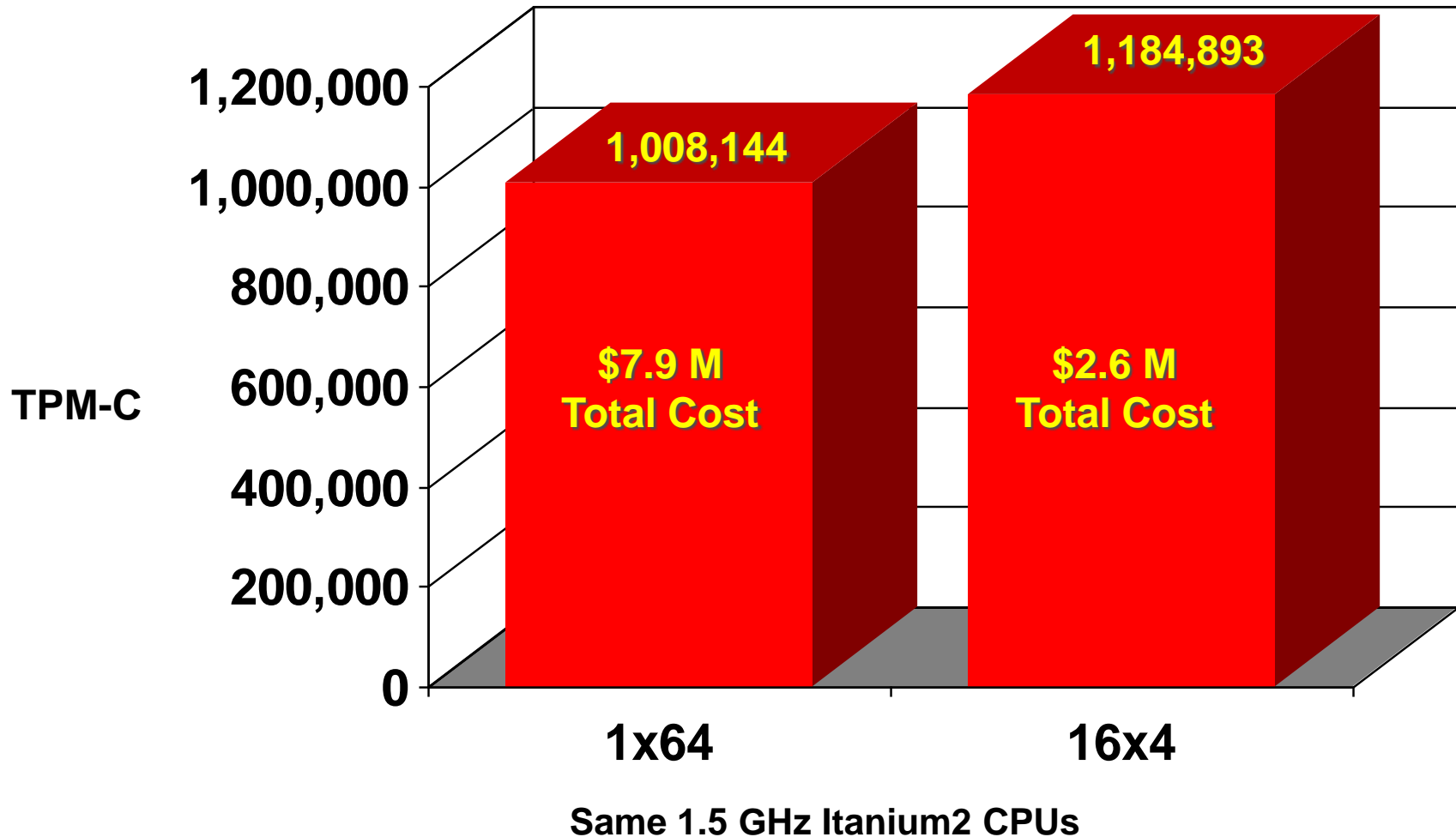
- 一个内在的性能专家
- 能够从整个RAC级别从全局的角度定位性能问题
- 数据库级的分析：
 - 全局缓存的内部连接问题
 - 全局资源竞争, 比如IO带宽, 热块等
 - 全局的TOP SQL
 - 实例响应时间的不均衡
- 缺省每小时自动收集

成本比较: SMP 与 集群(Grid)

电信级别的负载 – 550,000 transactions/hour



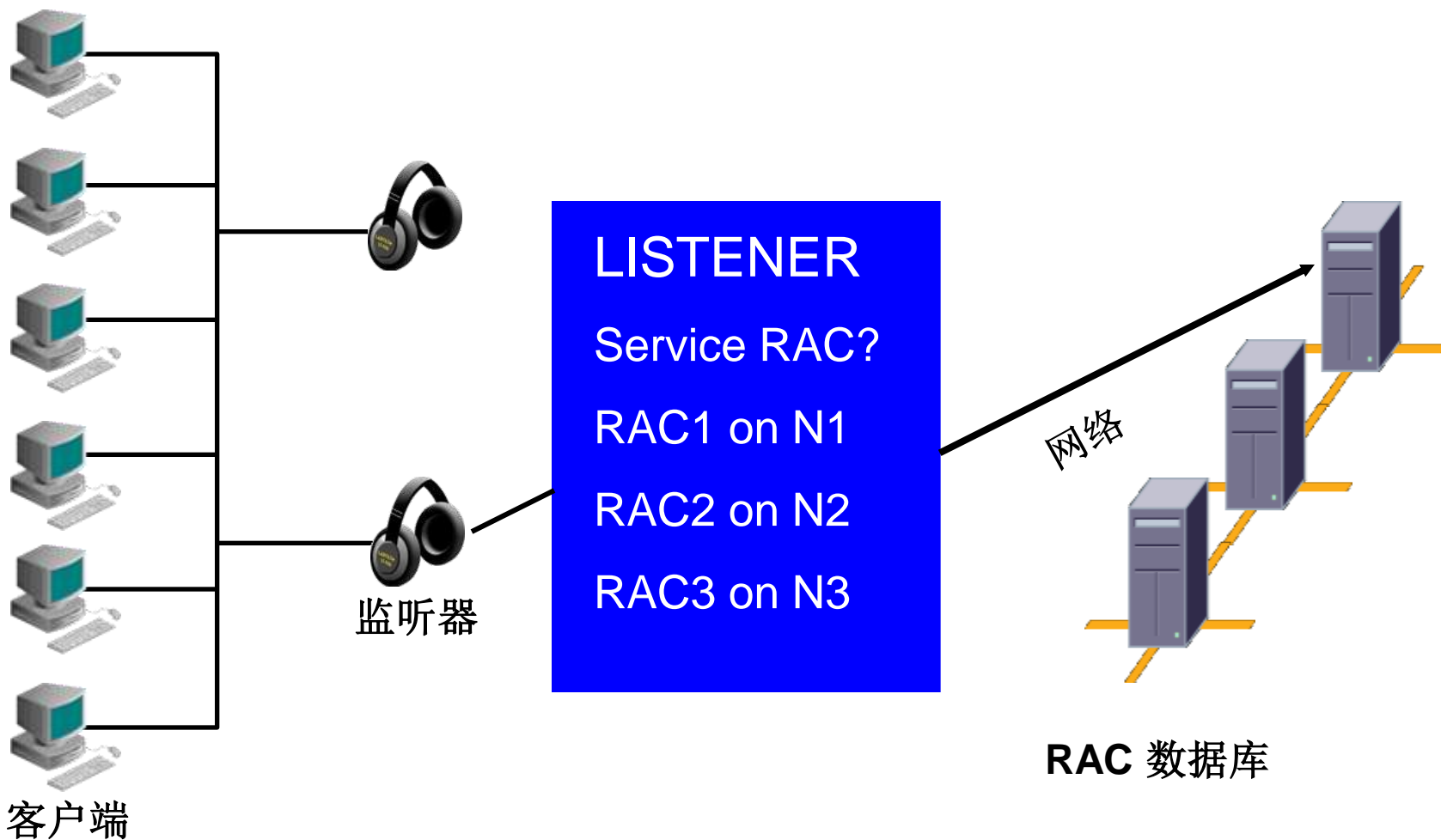
低成本服务器 - HP TPC-C Benchmark



ORACLE

Automatic Workload Management

服务与连接的负载均衡



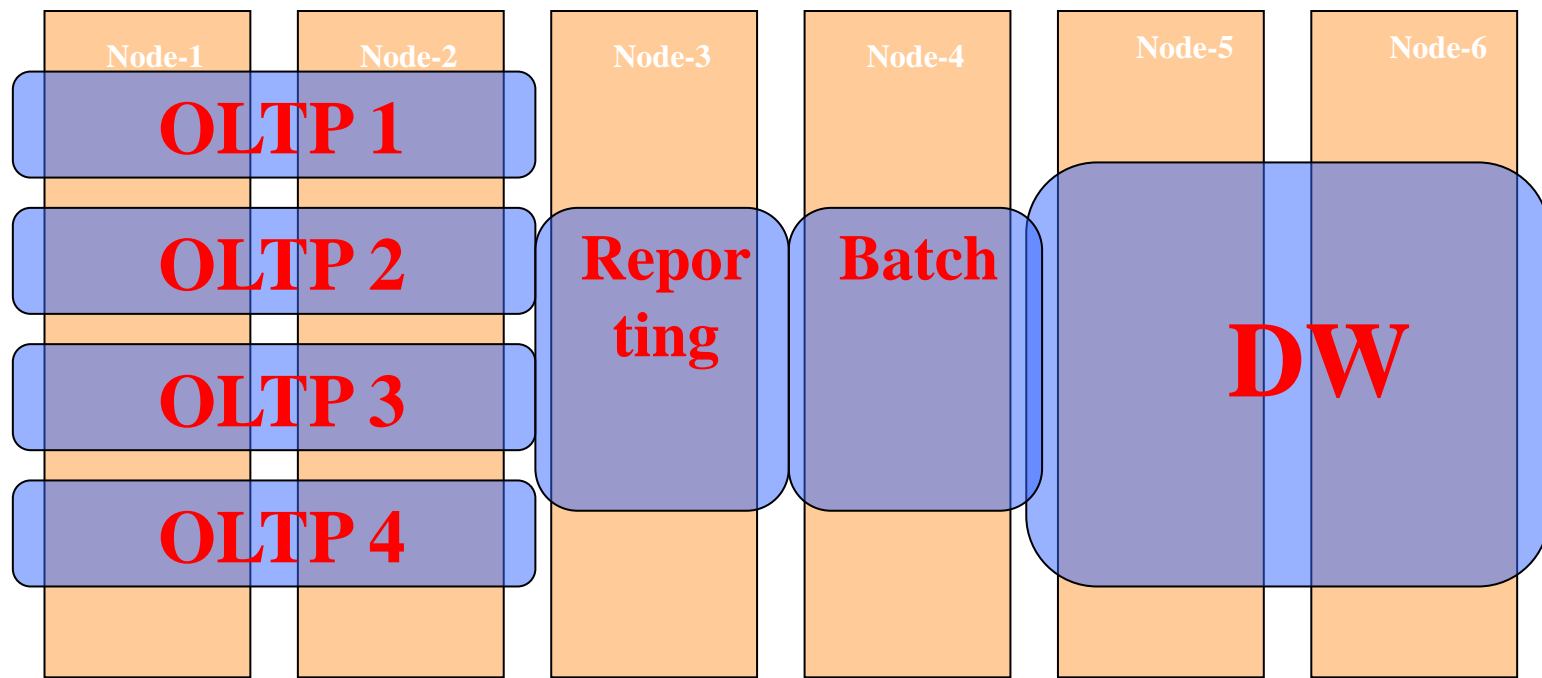
自动负载管理

负载均衡建议器

- 负载均衡建议器可为RAC实例之间的负载均衡提供建议
- 负载均衡建议
 - 对所有应用程序可用
 - 将工作直接分配到执行完好和有空闲资源的节点
 - 调节不同节点间的任务分配、优先级别、负载类型和变更需求
 - 停止将任务发送到慢的、挂起的或有问题的节点

并行查询与Service的集成

- 并行查询将只在Service宿主活动的节点上运行
- 不需要设置instance_groups



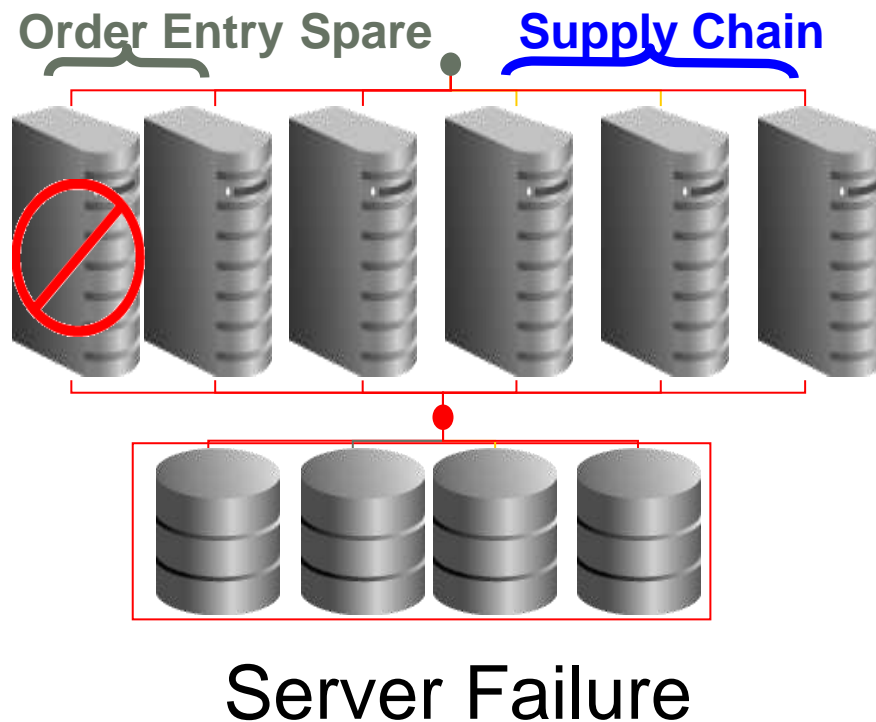
RAC的高可用性



高可用性

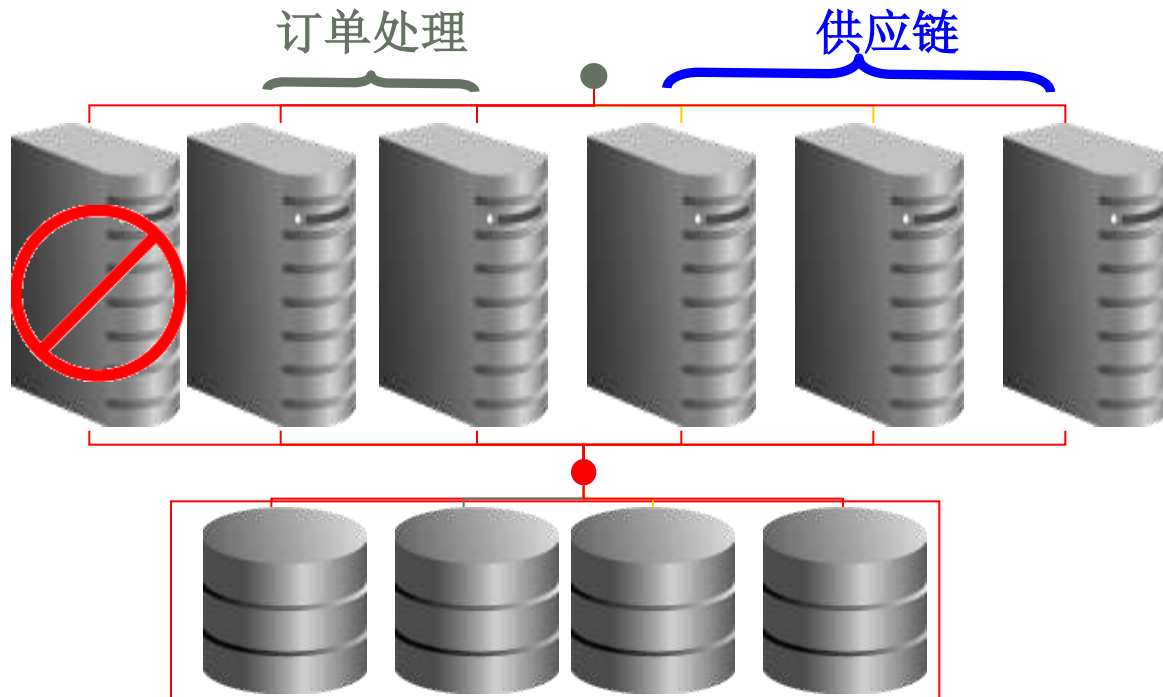
消除单点故障

- 是Oracle最大可用性架构(MAA)的一部分
- 能够实现底层的故障对应用程序透明



自动负载管理

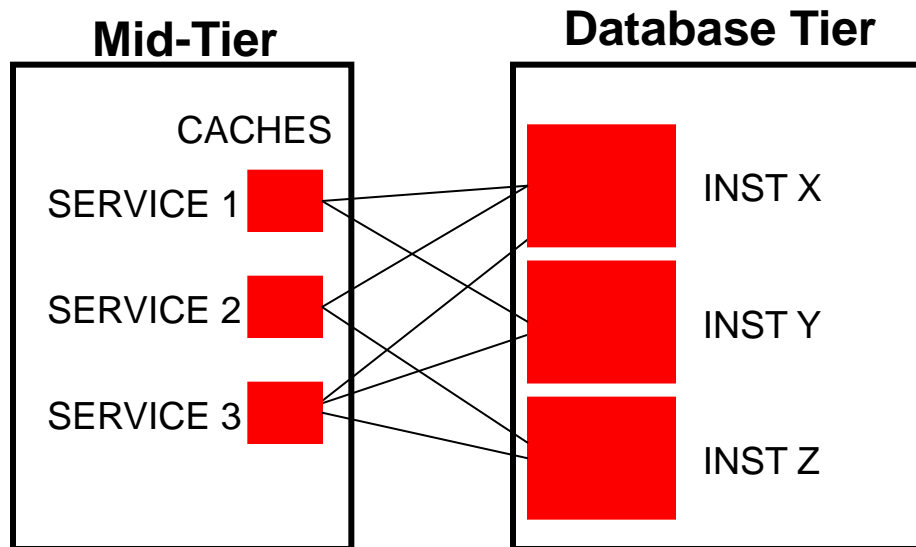
应用自动切换到可用节点



自动分配订单处理应用

快速连接切换

- 支持多个连接缓存
- 每一个缓存的数据源被定义到一个服务
- 对每个连接进行跟踪
- 当错误发生时立刻清除连接
- 将新任务分配到活动的实例

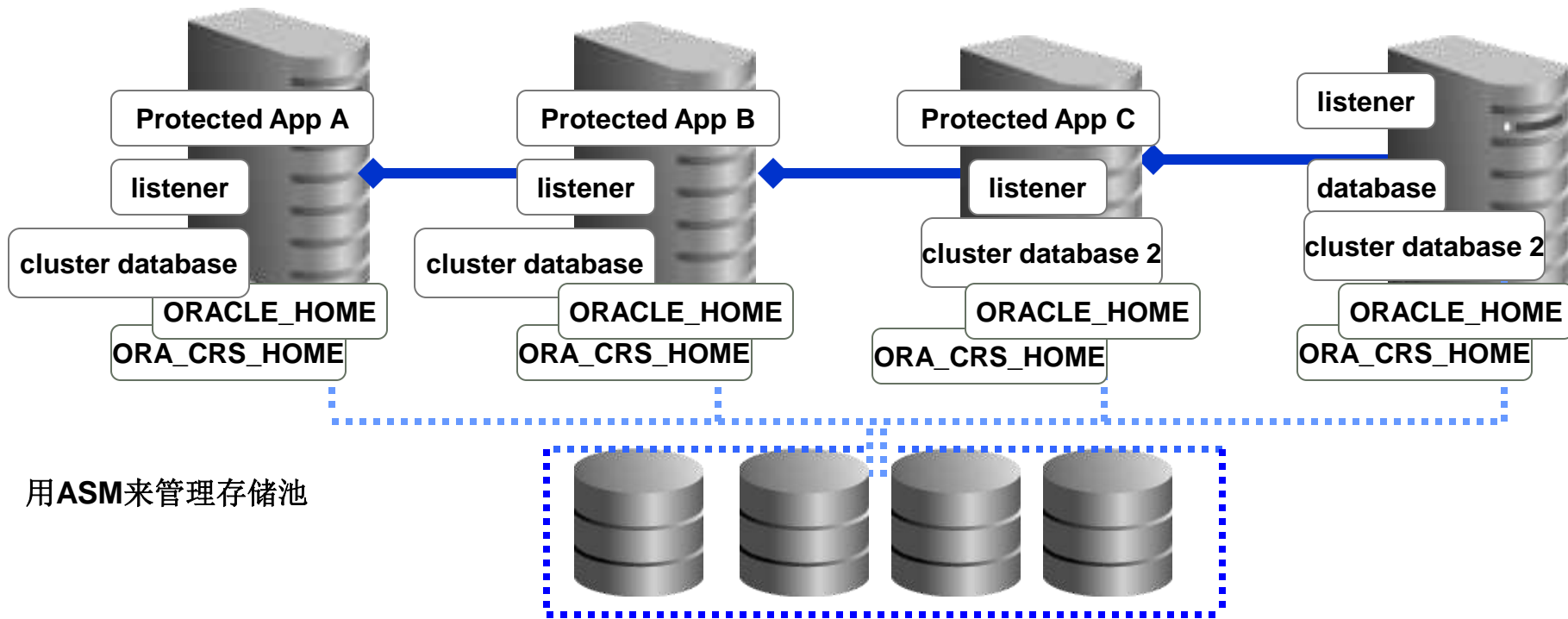


JDBC, ODP.NET, OCI

Oracle 集群软件与 RAC 的集成



集群集成



- 仔细规划适合要求的容量

RAC的关键价值

- 降低宕机后的切换时间 → 提高系统的可用性(99.999%，一年允许停机5分钟)
 - 双机热备需要重新启动和恢复数据库，可能需要10分钟以上，意味着应用系统有10分钟以上不能提供服务
 - 真正应用集群切换只需要几十秒钟，因为无需重启数据库
- 提高硬件资源的利用率 → 提高投资回报率
 - 双机热备意味着有一台机器不工作或承担少量工作（例如充当应用服务器）
 - 真正应用集群充分利用每一台机器的资源
- 降低整体成本
 - 使用真正应用集群可以采购比原来配置更低的服务器硬件
 - RAC 10g/11g自带Clusterware, 可以节省购买第三方集群软件的费用

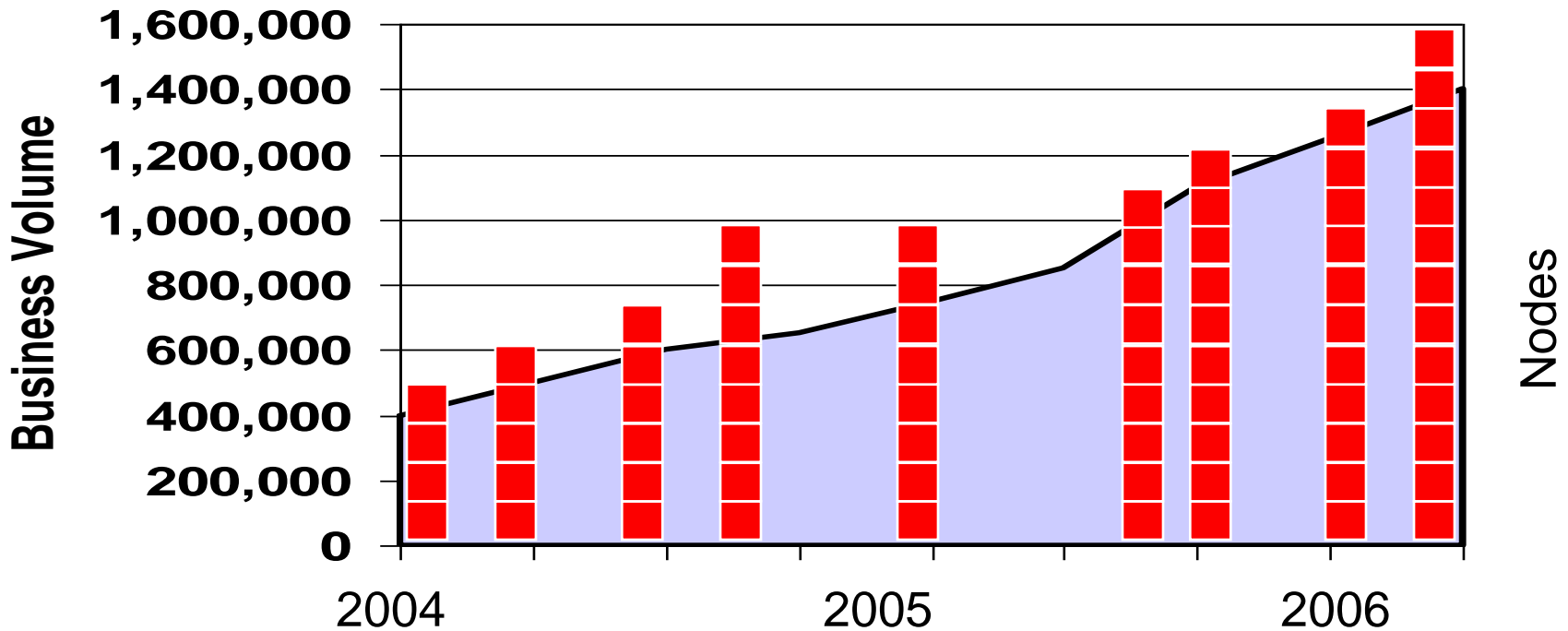
RAC应用案例





Mercado Libre

- eBay拉丁美洲分公司
- 基于RAC的在线市场应用
- 随业务发展动态添加RAC节点



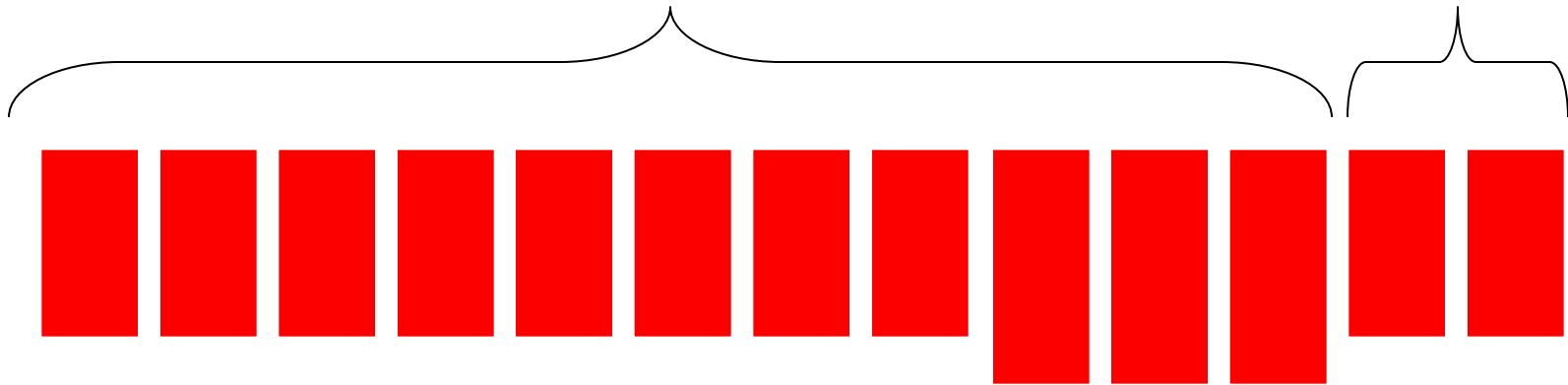


Mercado Libre

交易分布

在线市场应用
搜索, 列表, 交易, 报表

客户服务



- 新节点有更多的内存和更快的处理能力
- 在线市场应用的负载没有按节点划分

亚马逊的RAC解决方案

	Query	ETL
#Nodes x #cpus	16 x 4	8 x 4
Total DB	61 TB	10 TB
Data	51 TB	7 TB
Index	2 TB	1-2 TB
Disk	71 TB	36 TB

2005年10月亚马逊将它35TB的Clickstream数据仓库和25TB的查询数据仓库合并到一个RAC数据库，这套系统很好的支持了当年年底的销售业务高峰

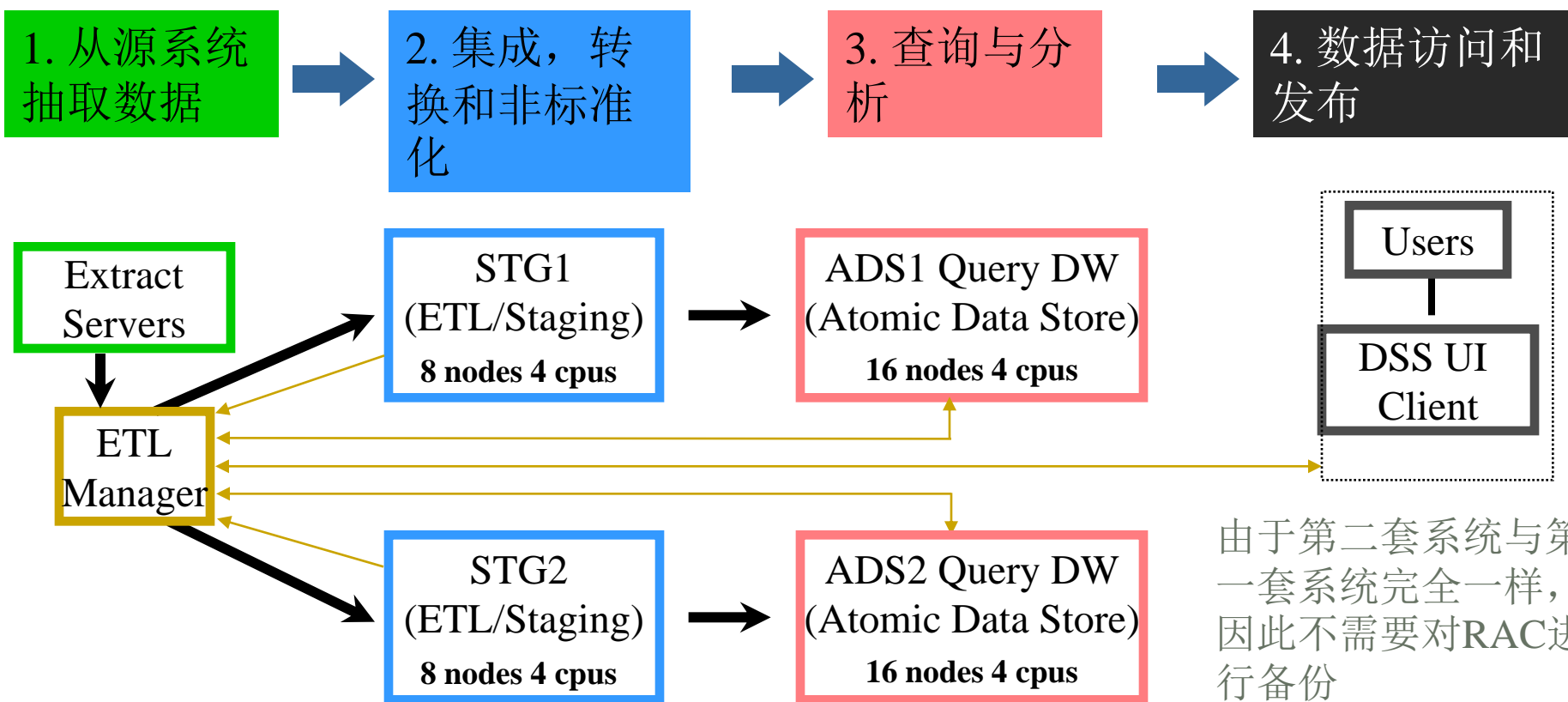
在当时这2个数据库都是全球最大的10个数据库仓库之一：

www.wintercorp.com

现在查询数据库已经超过
61 TB

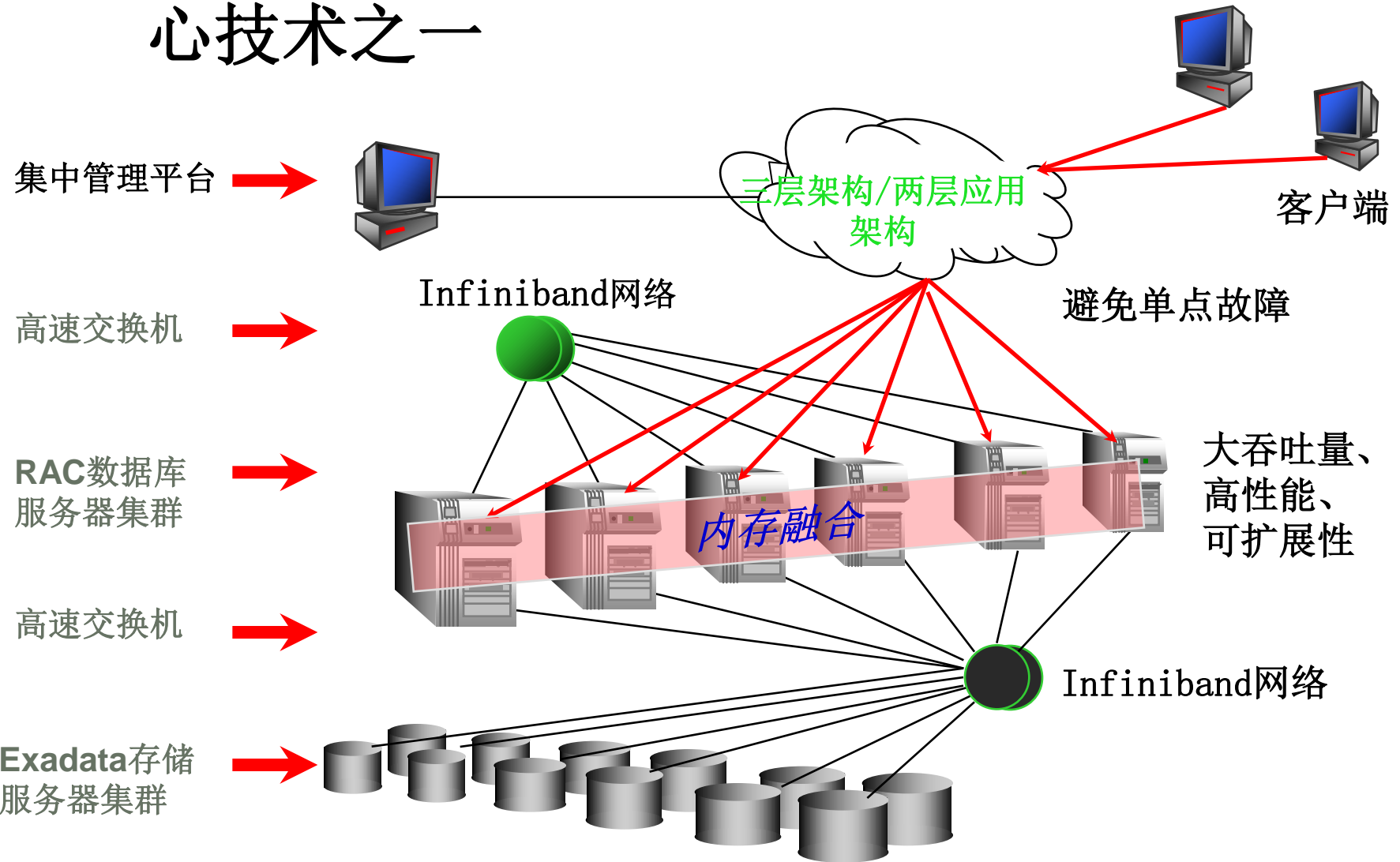
亚马逊模块化的RAC架构

亚马逊的RAC架构非常符合成本效益,它同时运行2套相同的系统但仍然能节省成本



由于第二套系统与第一套系统完全一样,因此不需要对RAC进行备份

RAC是Oracle Exadata数据库一体机的核心技术之一





**Q U E S T I O N S
&
A N S W E R S**

<http://otn.oracle.com/rac>

For More Information

<http://search.oracle.com>

Real Application Clusters



or

oracle.com/clusters