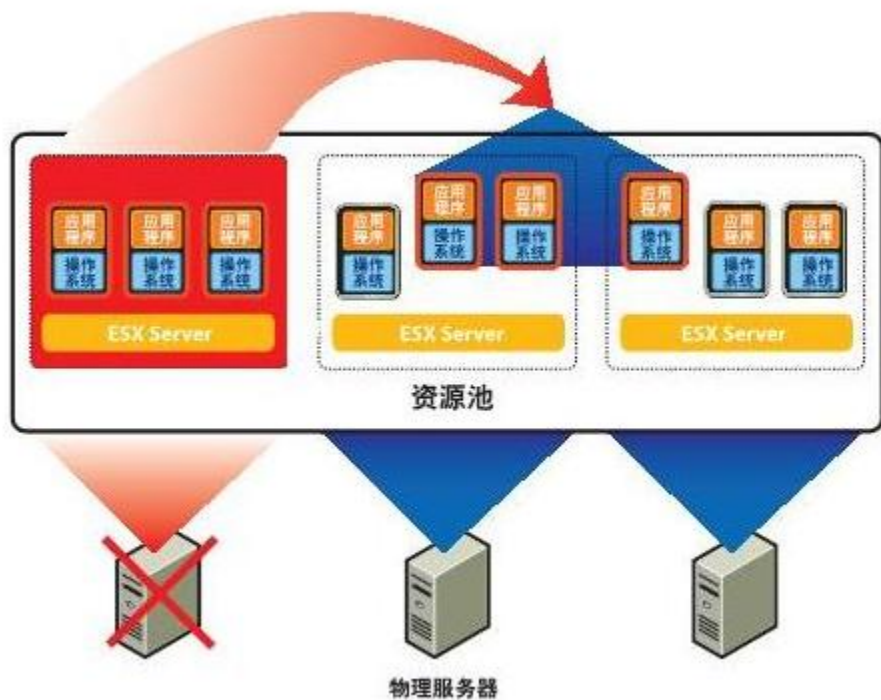




## 当新技术遭遇传统难题

徐景春



ORACLE  
甲骨文



SHANDA GAMES

# 数据库与虚拟化



## 现状- *The Situation*



## 目标 – *The Objective*

### 建立集中的敏捷Web运行平台

按访问量平衡  
处理能力资源调度

全局共享  
存储资源调度

虚拟化Web服务器

虚拟化数据库服务

### 提供高性能全局Web应用接入

1G 灵活配置的带宽

CDN支持下的TCP  
连接管理

CDN支持下的静态  
内容就近缓冲

### 满足开放平台和快速反应要求

15分钟就绪的Web  
应用环境

100%原位恢复的  
故障处理

100%生产场景重  
现的应用调试





## <<Oracle\_数据库可以在VMware产品上完美运行的10大理由>>

在此，我将向大家介绍一下确保我们获得最佳数据库性能的 10 大重要因素。以下是其中几项性能优势：

- 接近本机系统的性能：Oracle 数据库与物理系统的性能几近相似
- 超强的数据库 I/O 扩展性：VMware ESX Server 的精简虚拟化管理程序层每秒能够驱动 63,000 次数数据库 I/O 操作
- 多核扩展：使用 SMP 虚拟机和多个数据库实例进行纵向扩展
- 巨大的内存容量：可扩展的内存容量 - 每个数据库 64 GB；每台主机 256 GB

### 数据库性能认识误区

- 虚拟数据库具有很高的系统开销：就延迟和吞吐量而言，虚拟数据库的运行速度能等于或接近物理系统。对于常见的数据库而言，虚拟化开销很小 – 在 VMware ESX Server 中所测量到的 CPU 开销低于 10%。
- 数据库的 I/O 操作过多无法进行虚拟化：通常情况下，数据库会产生大量小的、随机的 I/O 操作，其理论值可能达到虚拟化管理程序层的 I/O 数量极限。然而，VMware ESX 的精简虚拟化管理程序层每秒可以驱动 63,000 多次数据库 I/O 操作，等同于 600 多个硬盘的 I/O 吞吐量。这样的 I/O 极限值，即使是用于 x86 系统上最大的数据库也绰绰有余。
- 虚拟化只能用于较小的、无关紧要的应用程序：ESX 虚拟化管理程序非常强健 - 很多客户的基于 ESX 的系统已经连续正常运行了两年多的时间。此外，ESX 虚拟化管理程序一直非常稳定，即便在资源过度使用时亦是如此。

数据库是否适合虚拟化?

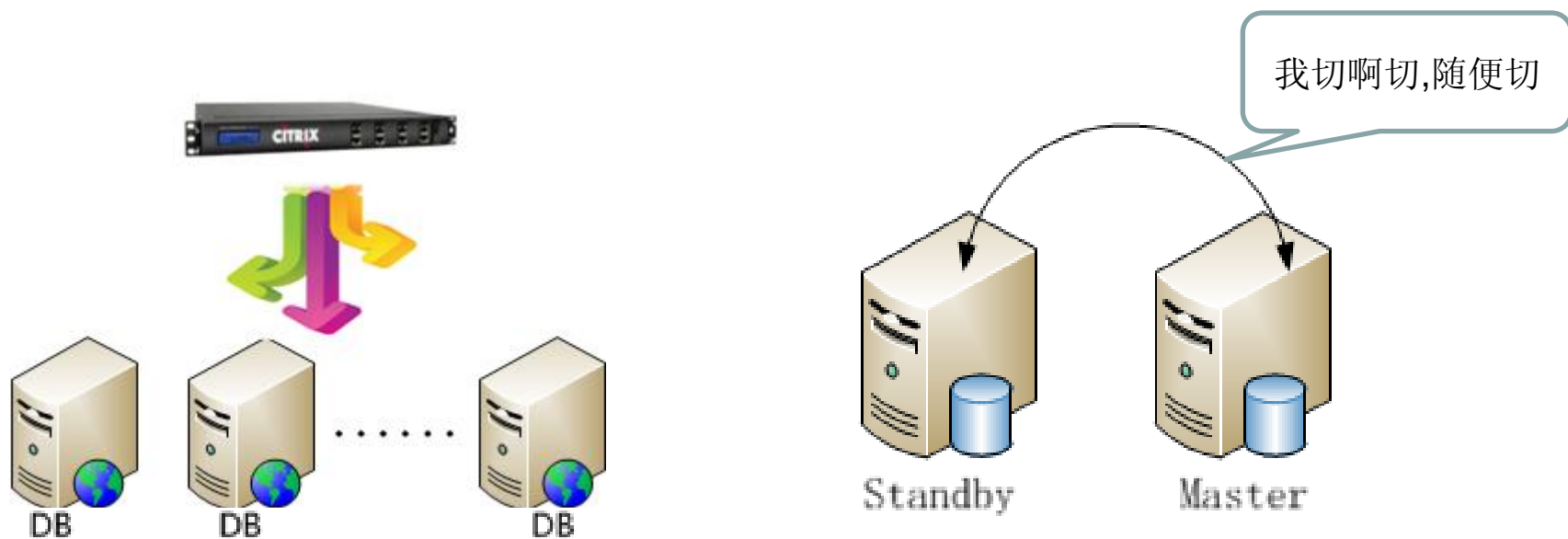


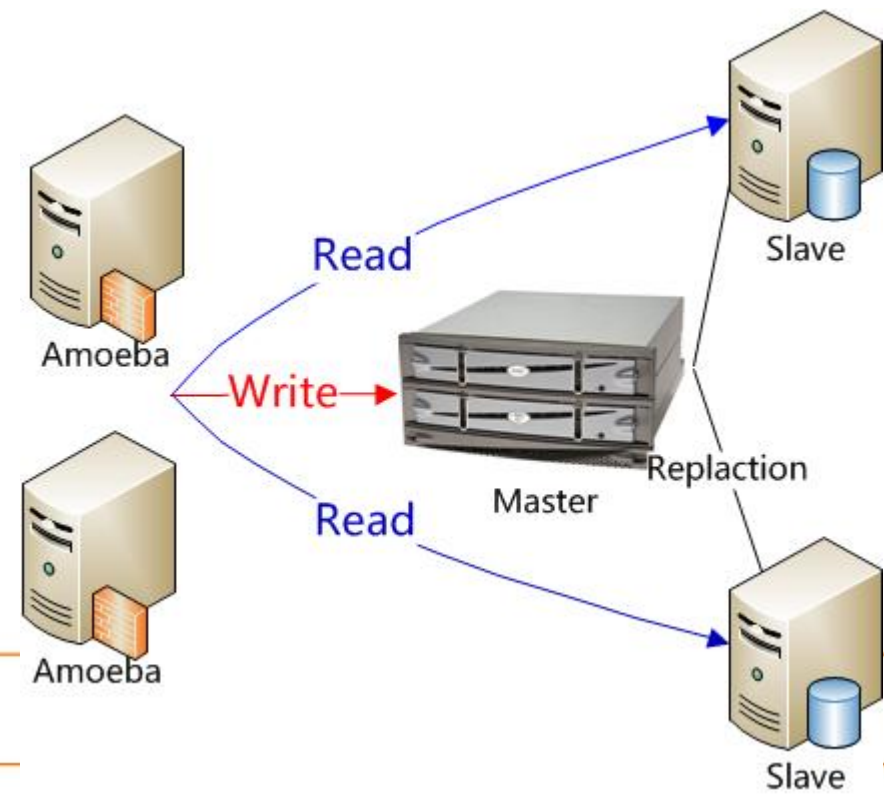
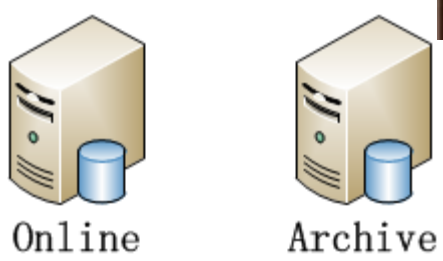
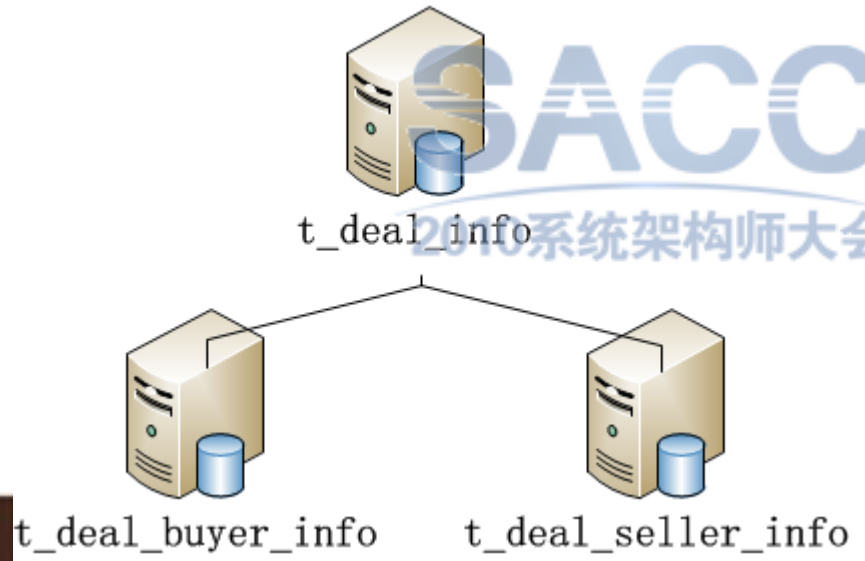
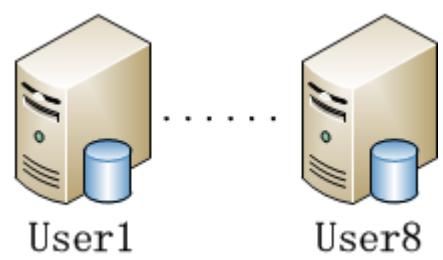
# 数据库与分布式



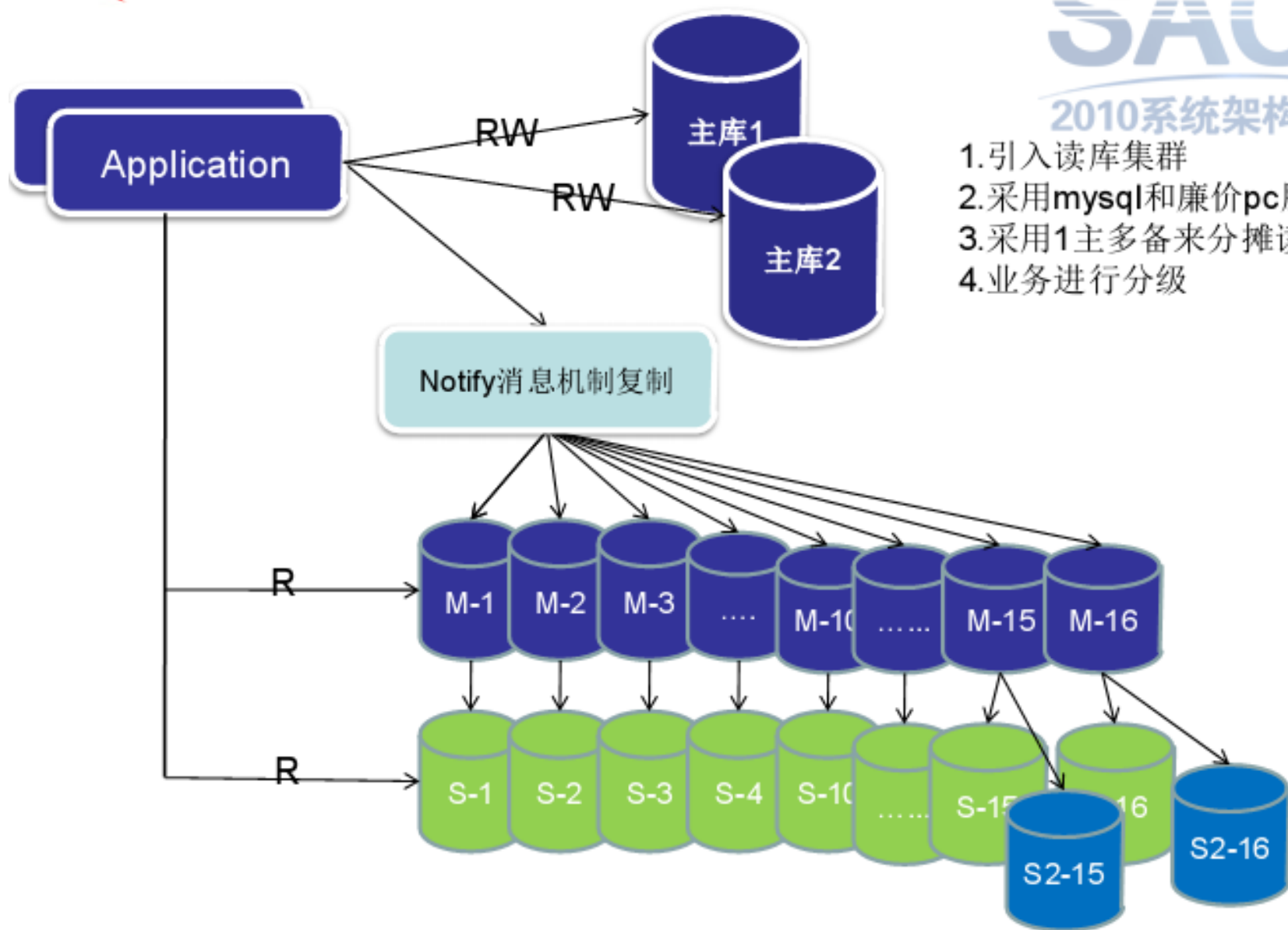


## 数据库难念的经





1. 引入读库集群
2. 采用mysql和廉价pc服务器
3. 采用1主多备来分摊读压力
4. 业务进行分级





### ◆规模效应

更为低廉的成本

“人”多力量大

### ◆良好的扩展性

### ◆良好的可用性,良好的容错性

### ◆良好的并发性能



- ◆大文件
- ◆大部分写操作是insert,最忌讳随机update
- ◆大部分情况是insert后,文件只读





## DFS能适合DB吗？

CloudEx

 **Alibaba.com**<sup>®</sup>  
Global trade starts here.™

  
北京美地森科技有限公司  
Beijing MDS Technology Co., Ltd.



SHANDA  GAMES



## Storage

Amazon Simple Storage Service (S3)

Amazon Elastic Block Storage (EBS)

AWS Import/Export

### Amazon EBS Volume Performance

The latency and throughput of Amazon EBS volumes is designed to be significantly better than the Amazon EC2 instance stores in nearly all cases. You can also attach multiple volumes to an instance and stripe across the volumes. This is one way to improve I/O rates, especially if your application performs a lot of random access across your data set.

The exact performance will depend on the application (e.g. random vs. sequential I/O or large vs. small request sizes), so the best measure is to benchmark your real applications against the volume. Because Amazon EBS volumes require network access, you will see faster and more consistent throughput performance with larger instances.

This is particularly helpful for database style applications that frequently encounter many random reads and writes across the dataset.



inspur 浪潮



## 整合

- 支持NAS/IPSAN/FCSAN/IB各种存储网络的统一架构
- 支持iSCSI、FC、万兆、InfiniBand多种主机连接
- 全面支持SSD/FC/SAS/SATA硬盘
- 支持异构主机平台（Windows/Linux/Unix）



两台客户端:

[root@sn02 workloads]# filebench\_ops

27788: 6.024: Eventgen rate taken from variable

27788: 6.024: Usage: set \$dir=<dir> defaults to /mnt/client

27788: 6.024: set \$eventrate=<value> defaults to 0

27788: 6.024: set \$filesize=<size> defaults to 65536

27788: 6.024: set \$iosize=<value> defaults to 2048

27788: 6.024: set \$nthreads=<value> defaults to 80

27788: 6.024: set \$workingset=<value> defaults to 0

27788: 6.024: set \$directio=<bool> defaults to 0

27788: 6.024: run runtime (e.g. run 60)

client--1:

rand-wtrate 0ops/s 0.0mb/s 0.0ms/op 0us/op-cpu

rand-write1 619ops/s 1.2mb/s 123.2ms/op 281836us/op-cpu

rand-rdrate 0ops/s 0.0mb/s 0.0ms/op 0us/op-cpu

rand-read1 14223ops/s 27.8mb/s 0.4ms/op 6723us/op-cpu

27788: 77.222: IO Summary: 891407 ops, 14842.145 ops/s, (14223/619 r/w),

29.0mb/s, 249us cpu/op, **5.5ms latency**

Client-2:

rand-wtrate 0ops/s 0.0mb/s 0.0ms/op 0us/op-cpu

rand-write1 582ops/s 1.1mb/s 98.9ms/op 223977us/op-cpu

rand-rdrate 0ops/s 0.0mb/s 0.0ms/op 0us/op-cpu

rand-read1 15009ops/s 29.3mb/s 0.4ms/op 4901us/op-cpu

27997: 111.908: IO Summary: 936410 ops, 15591.446 ops/s, (15009/582 r/w),

30.4mb/s, 237us cpu/op, **4.1ms latency**

	单客户端	两客户端
4K	32000ops/s	29500ops/s
32K	29000ops/s	29000ops/s
64K	32800ops/s	29500ops/s

低延迟、高元数据操作速度，适用计算密集型、虚拟化等应用



## 通用DFS



祥云



DVD





数据库能在分布式上应用吗？





Thank You!

