



淘宝网
Taobao.com

淘我喜欢!

高可用分布式数据库系统 架构实践

淘宝网
Taobao.com

陈吉平
Oracle ACE Director
《构建oracle高可用环境》一书作者

2009年7月



- 淘宝网站发展历程
- 淘宝数据库的发展历程
- 分布式数据库体系架构案例
- 总结

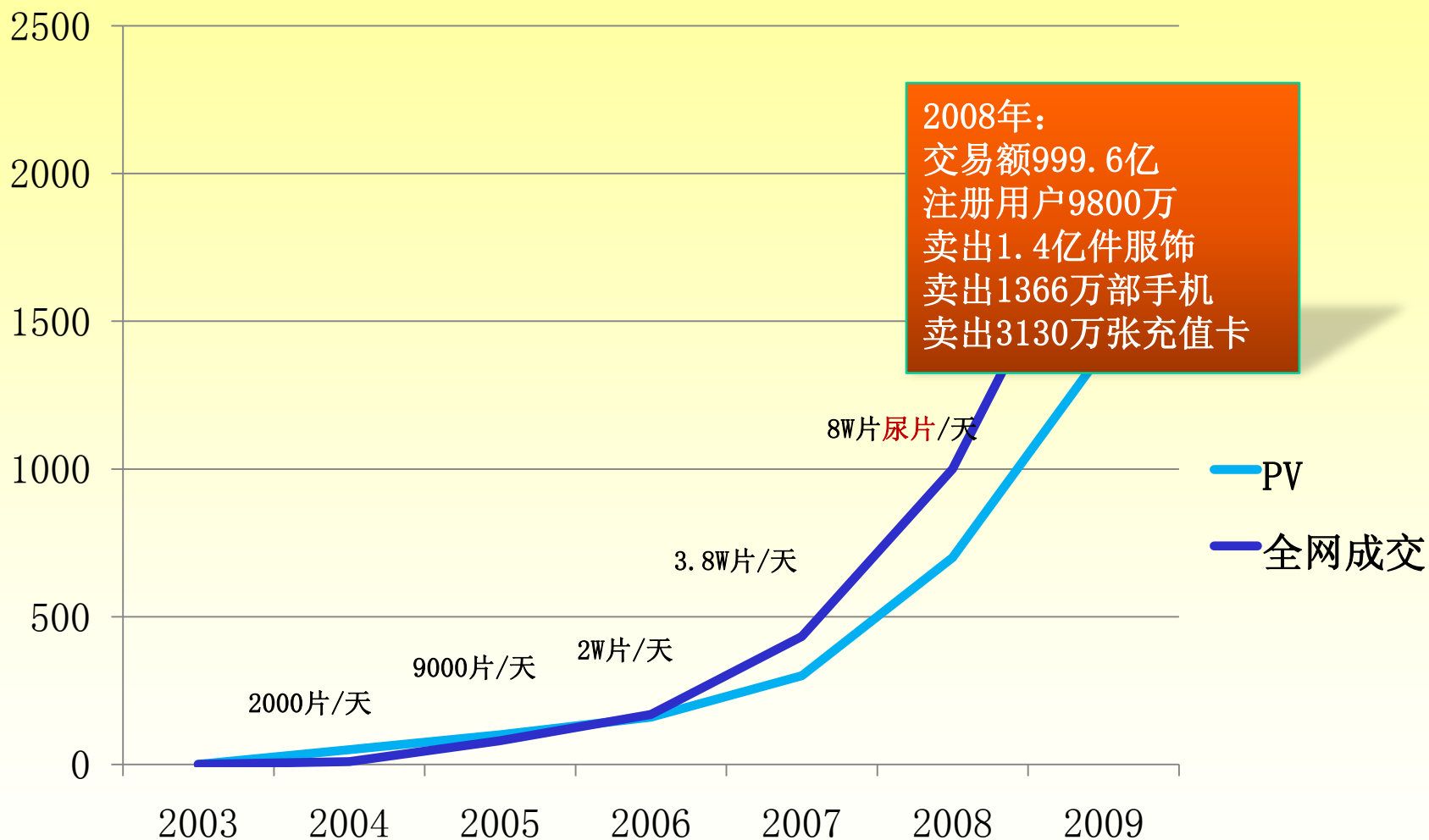




淘宝网站发展历程

淘宝网
Taobao.com

淘我喜欢!





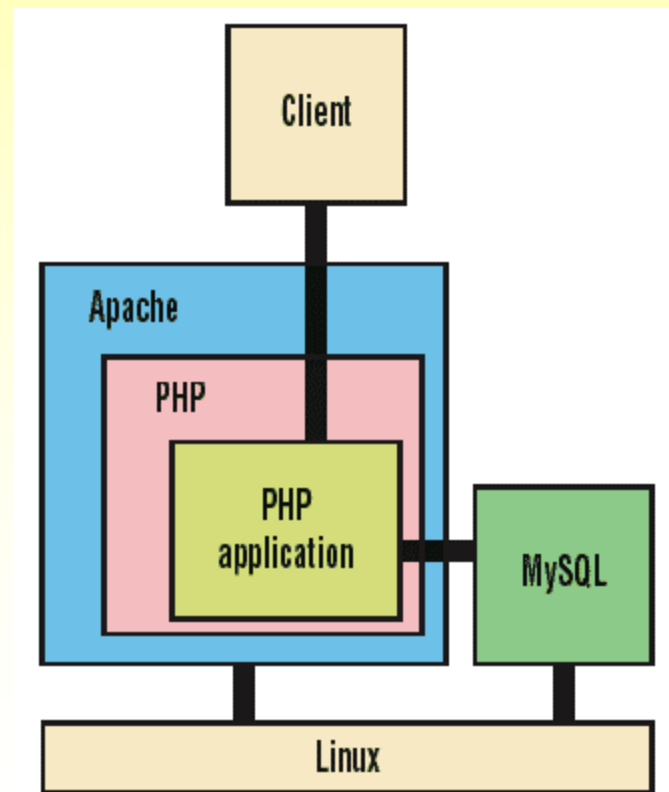
V1.0:最早的淘宝网

淘宝网
Taobao.com

淘我喜欢!

- 2003年非典时期
- 使用**LAMP**架构 (Linux, Apache, MySql, Php)
 - 业界流行的免费开源组合
- 使用**Mysql**数据库
 - 很少的表, 少量的数据
 - 用户, 交易, 商品, 其他

简单的结构, 但符合当时需求

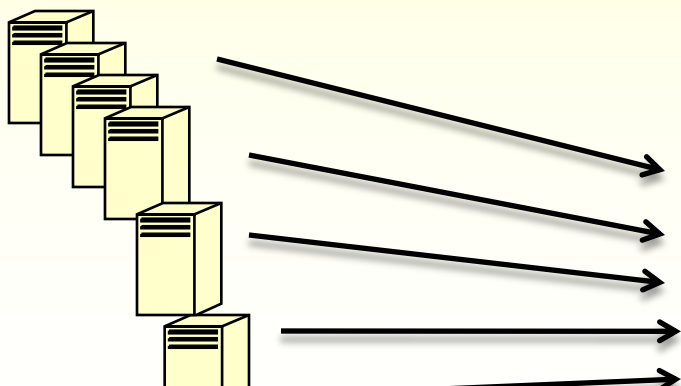




- 集中的数据库
 - IBM小型机+EMC存储
 - Oracle数据库
 - Scale up扩展方式
- 宝贵的数据库资源
- 单点，不方便扩展

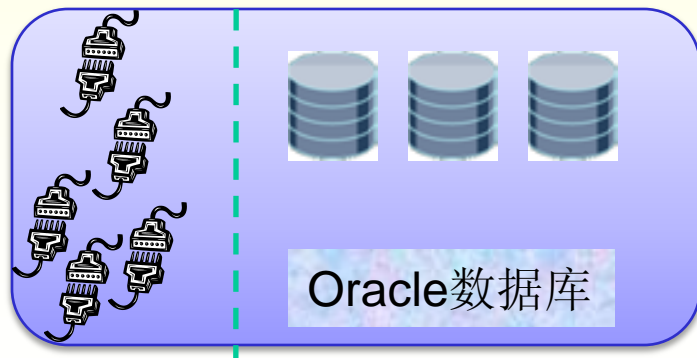
集中的数据库，简单快速

太多的应用机器



有限的链接池

集中的数据库





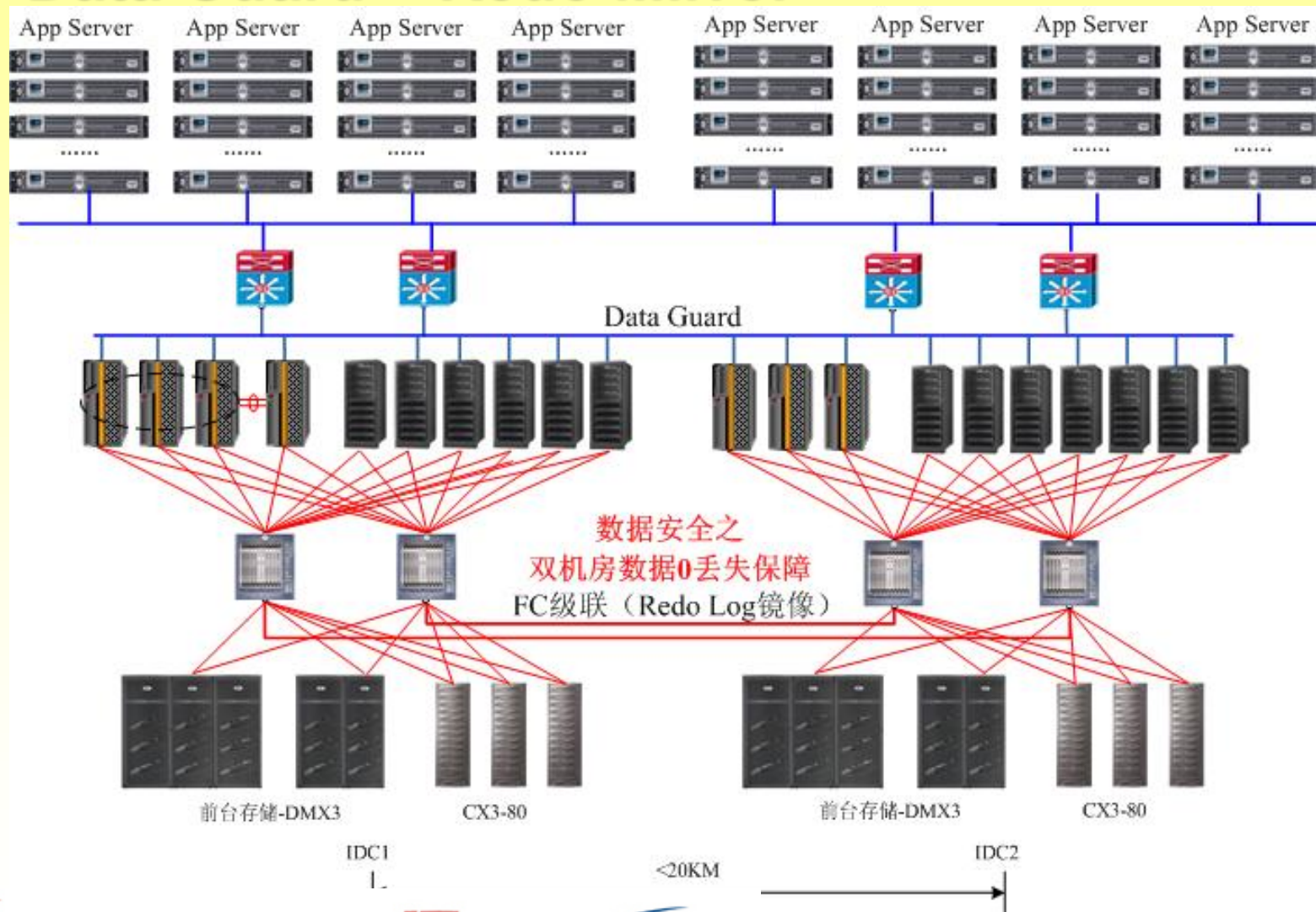
- 分布式存储(TFS)
 - 解决海量,非结构化数据的存储
 - Key-Value数据, 如图片
- 分布式Cache(Tair)
 - Cache才是王道



使用PC Server做存储媒介



• Data Guard + Redo Mirror





- 基于服务的构架体系
 - 用户中心，商品中心等
 - 每个业务有独立的应用与数据库环境
 - 可单独部署与扩展
- 非核心数据从Oracle迁移MySQL
 - E.g., 收藏夹
 - 节省费用

松耦合，服务导向的架构

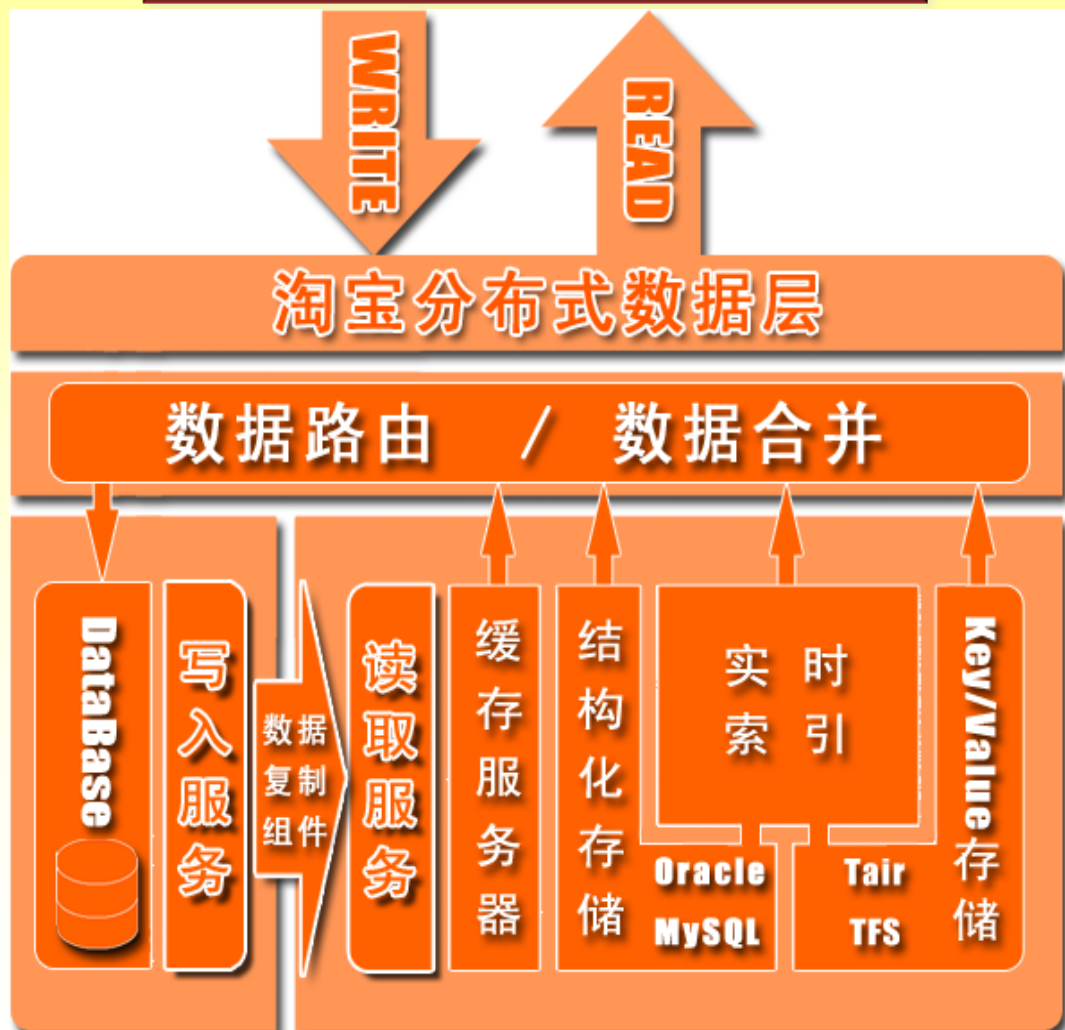


Copyright © 2004-2006 淘宝网 版权所有



从Scale up 到Scale out的转变

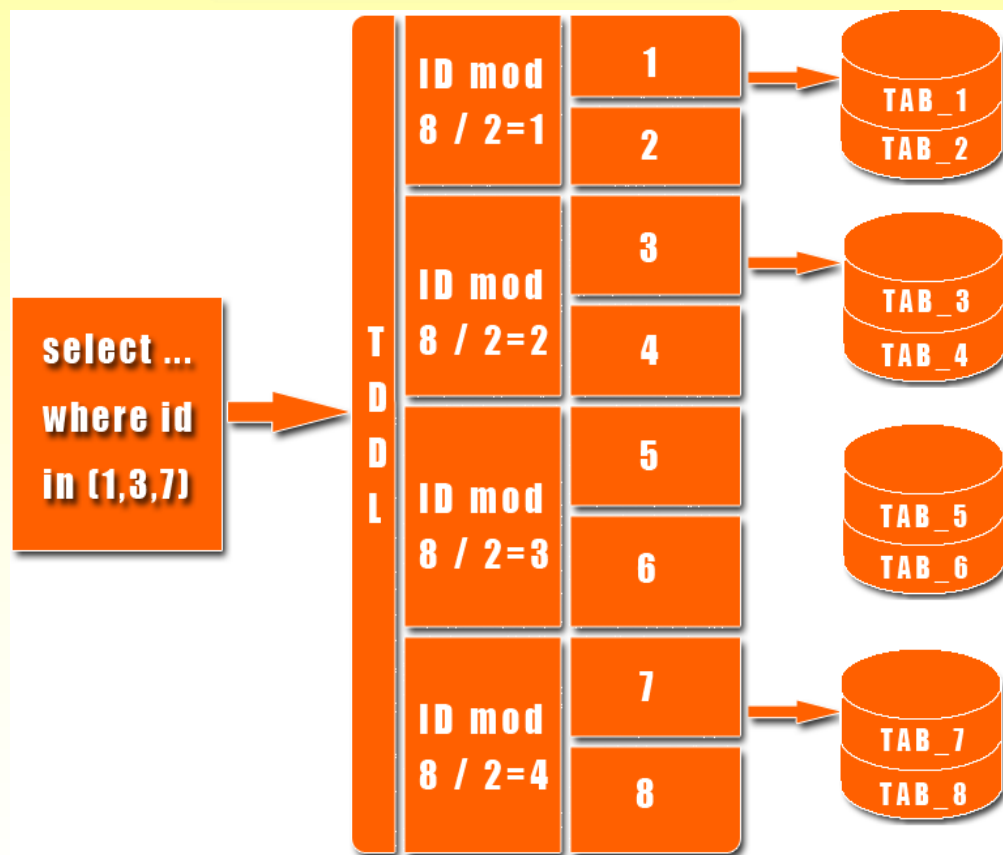
- 水平扩展构架体系
 - Scale out的解决方案
 - 缓解主库的压力
 - 独特的消息中间件同步方式 (Notify)
- 读库可以是Oracle或者是Mysql。
 - 读库可扩展
 - 坏掉任何一个读库，不影响业务

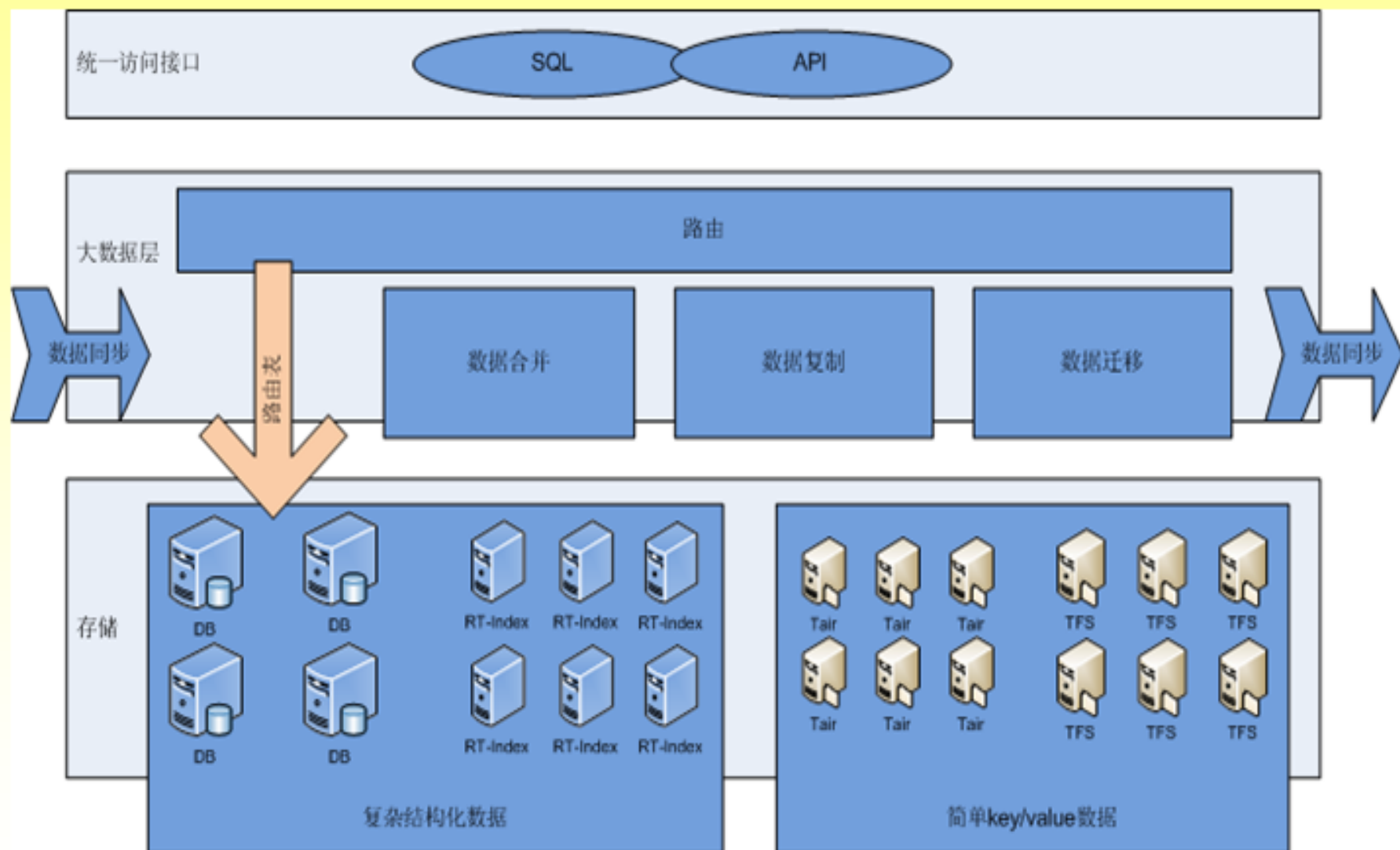




- 水平扩展构架体系
 - Scale out的解决方案
 - 对业务基本透明
 - 可动态扩展
- 支持任何数据库
- 未来支持多主结构
 - 坏掉任何一个主库，不影响业务
- 未来支持压力动态均衡
 - 数据可以动态分布
 - 可以方便的扩展/减少数据库主机

解决单库天花板问题







未来：多数据中心分布

淘宝网
Taobao.com

淘我喜欢！

高可用性 高稳定性





• 平台服务

- 如Amazon的S3, Ec2, Simple Db, 存储服务和标准的开放
- 持久层不一定需要数据库





Case 1: 读写分离

淘宝网
Taobao.com

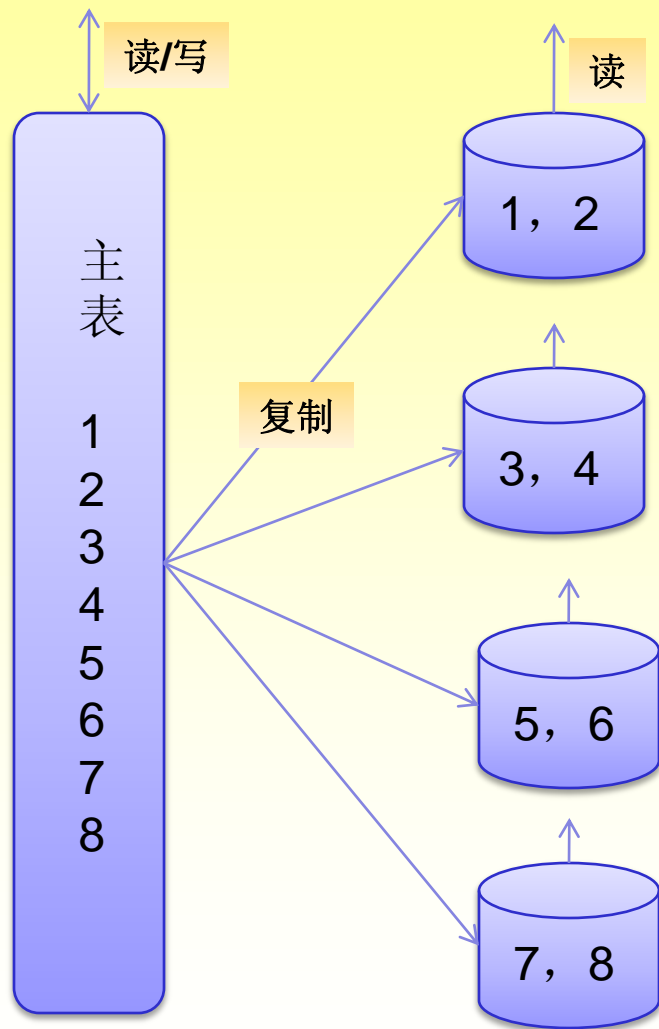
淘我喜欢!

• 2009年实现

- Oracle到Mysql
- 主库集中，读库水平分割
(主库的1/N)
- 写操作与部分重要的读操作回主库，其它的读操作回读库
- 读库故障可回主库

• 收到的成效

- 缓解主库的压力瓶颈
- 读节点可扩展
- 性能提升

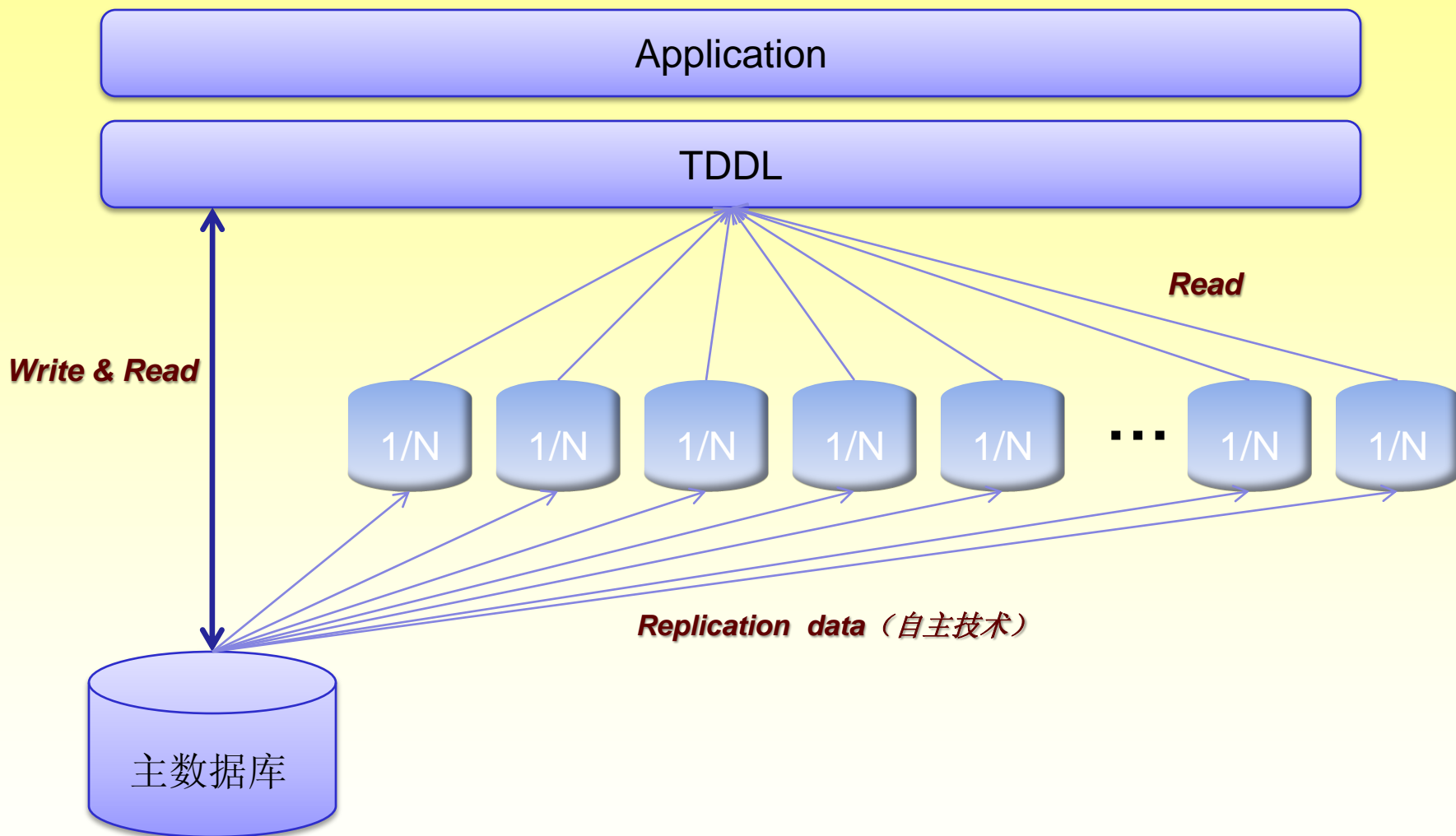




读写分离体系结构

淘宝网
Taobao.com

淘我喜欢!





Case 2: 水平分割

淘宝网
Taobao.com

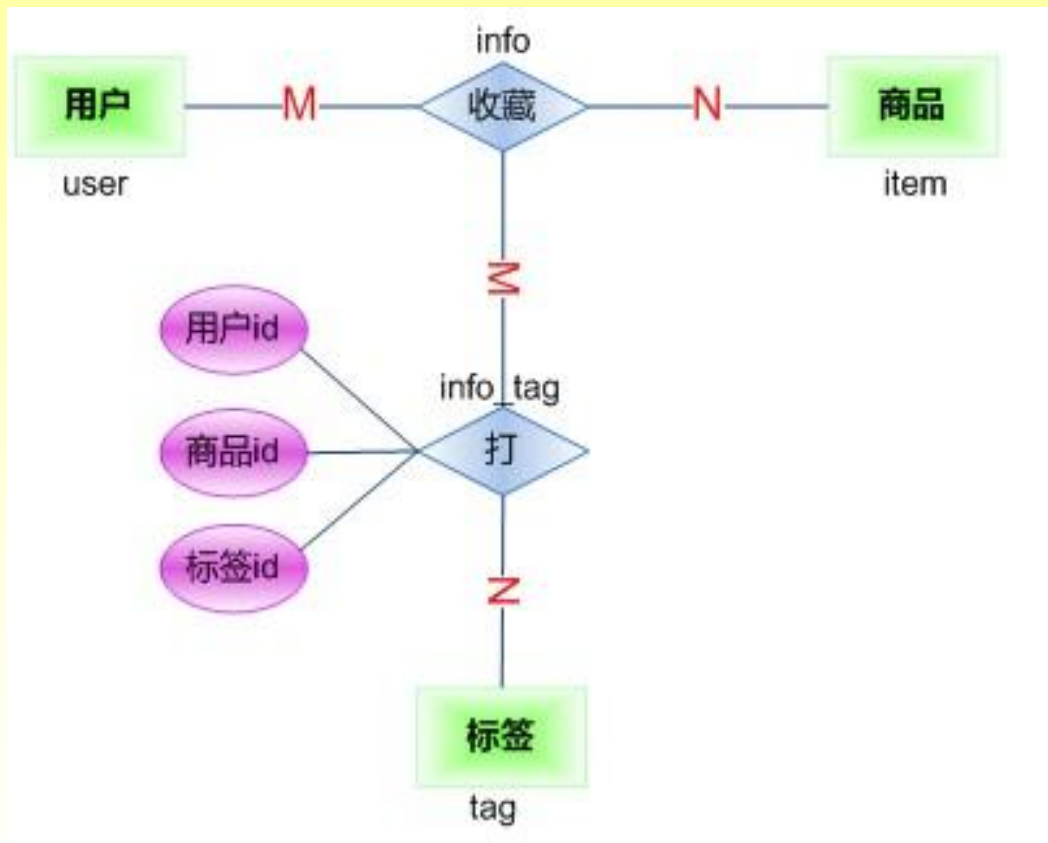
淘我喜欢!

- 2009年实现

- Mysql解决方案
- 水平分割
- 规划16个数据库，1024张表
- 分拆之前，单表16亿条记录

- 收到的成效

- 成本节省
- 速度更快
- 扩展性提高

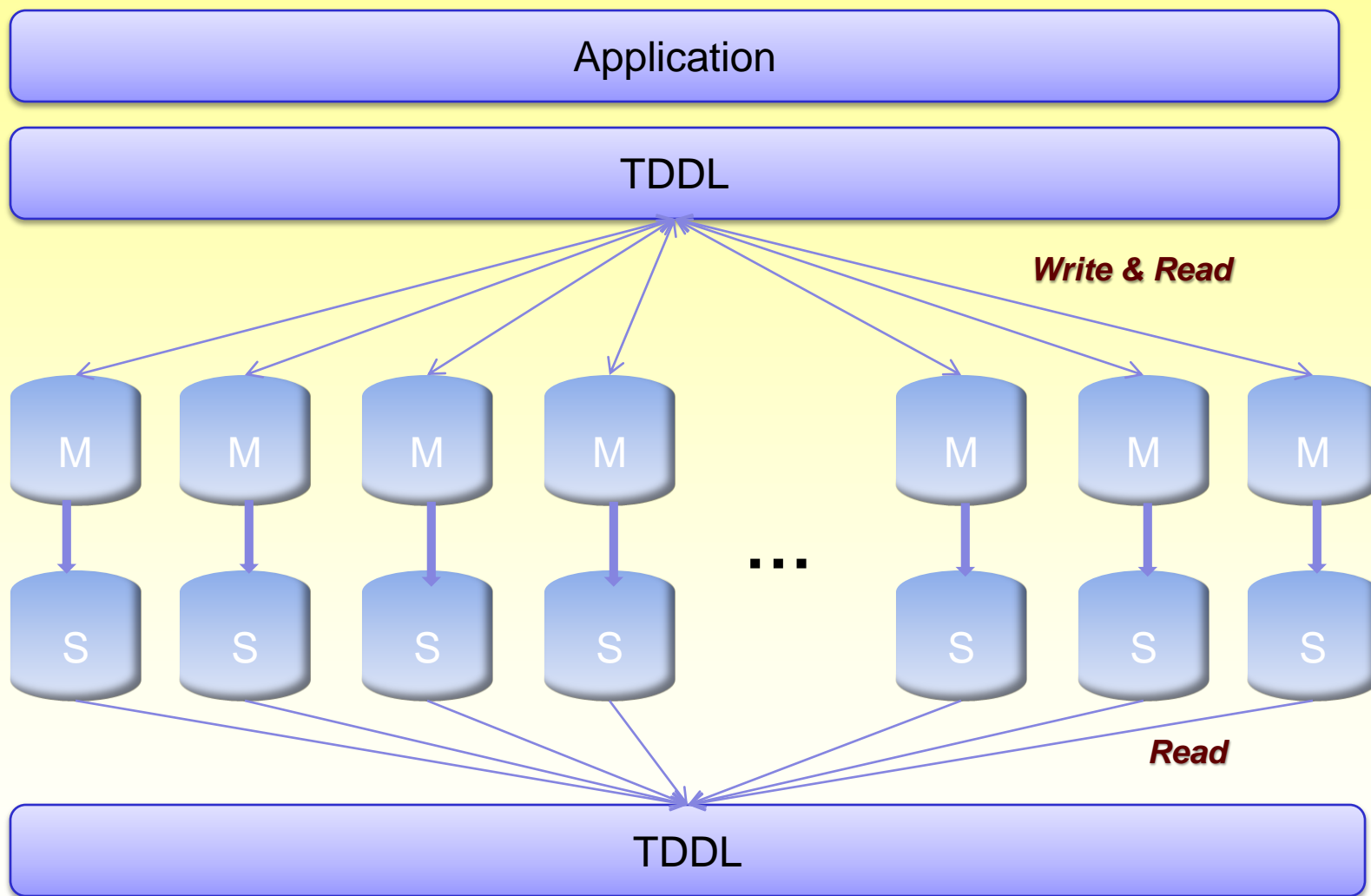




水平分割体系结构图

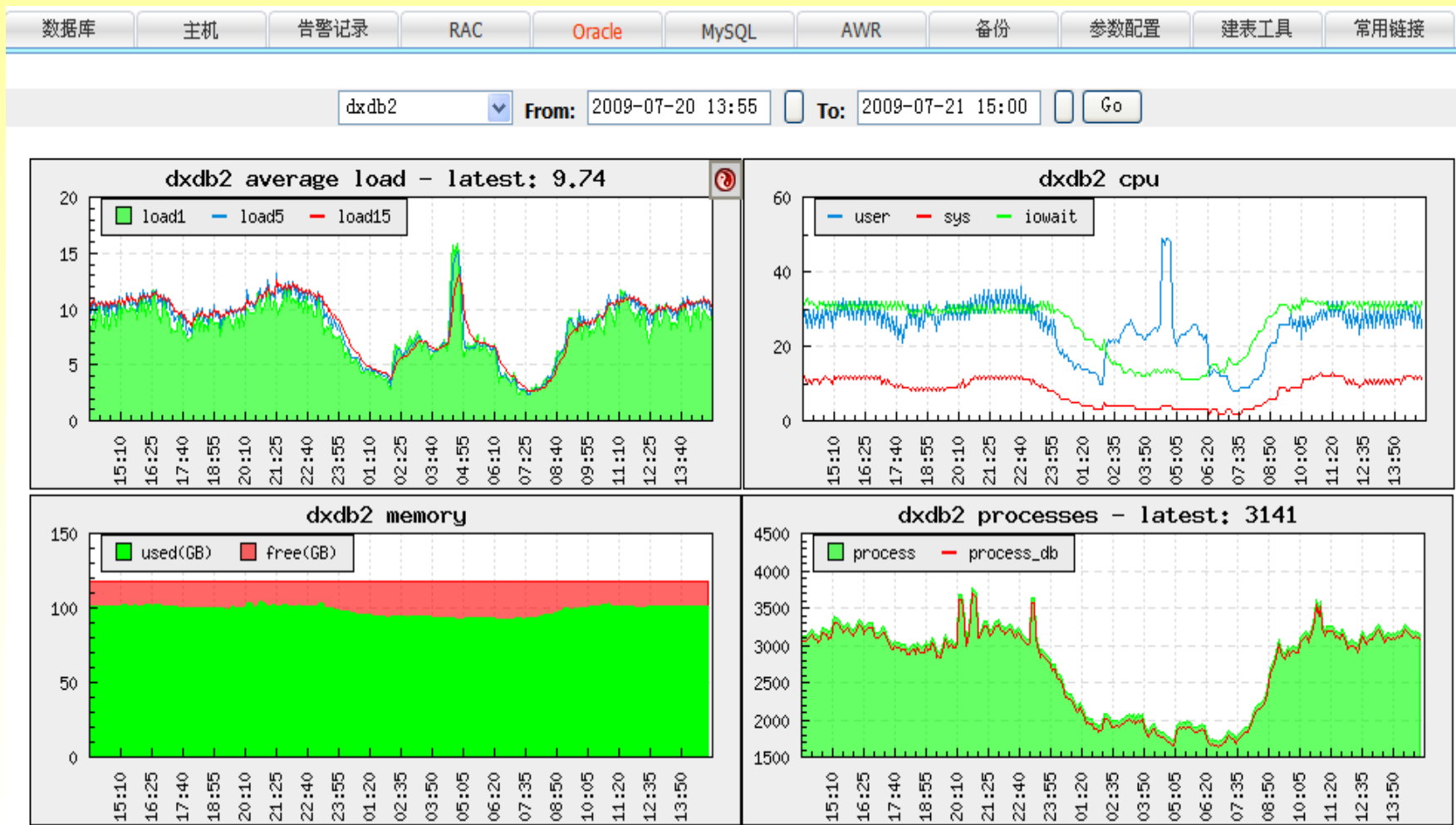
淘宝网
Taobao.com

淘我喜欢!





- 好的系统，需要好的监控





- 数据库技术发展的取决于业务与数据的规模
 - 没有最好的，只有最合适的
- 数据库技术需要结合整体构架一起考虑
 - 不是所有的东西都要用数据库
 - 配套的系统构架很重要
- 分布式环境中监控与工具的重要性
 - 工具&自动化，节省人力资源
 - 报警，了解最新状况
 - 标准化，减少错误的发生





谢谢!
Q&A