

# 关系数据库的挑战和机遇

阳振坤

[yangzhenkun@gmail.com](mailto:yangzhenkun@gmail.com)

北京五洲皇冠国际酒店

2014-4-10

# Agenda

关系数据库现状

分布式数据库

数据库与现代硬件

小结

# 关系数据库的前景堪忧？

## ● NoSQL/NewSQL风起云涌

- NoSQL和NewSQL已经超越了关系数据库
- 云时代NoSQL当道，关系数据库日薄西山
- NewSQL使传统关系数据库黯然失色
- 云计算使得关系数据库逐渐落伍
- 关系数据库的辉煌日子已经远去
- 关系数据库行将就木.....

## ● 关系数据库

- 几十年风平浪静
- 单机系统，扩展能力差，大数据量、高并发则捉襟见肘
- 高端服务器/高端存储，性价比低，曲高和寡

# 关系数据库：业务需求强劲！

- ACID(原子性/一致性/隔离性/持久性)无法取代
  - 如果业务需要ACID，那么底层弱化ACID的结果是上层应用付出更大的代价
- 电子商务高速发展、互联网金融遍地开花
  - ACID
  - 高可用
  - 高可扩展(性能+容量)
  - 高性价比

# 关系数据库的现状

✓ ACID

- 高可用

- 主库+备库

- 高可扩展

- 单机系统：高性能服务器，共享存储

- 高性价比

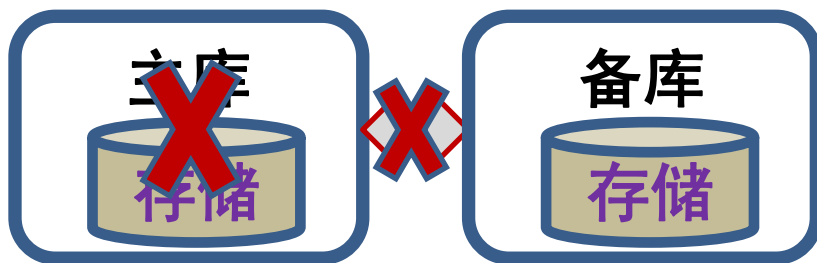
- 高端专用服务器，高端专用存储

# 主库+备库：没有共享存储

## ● 主备同步模式

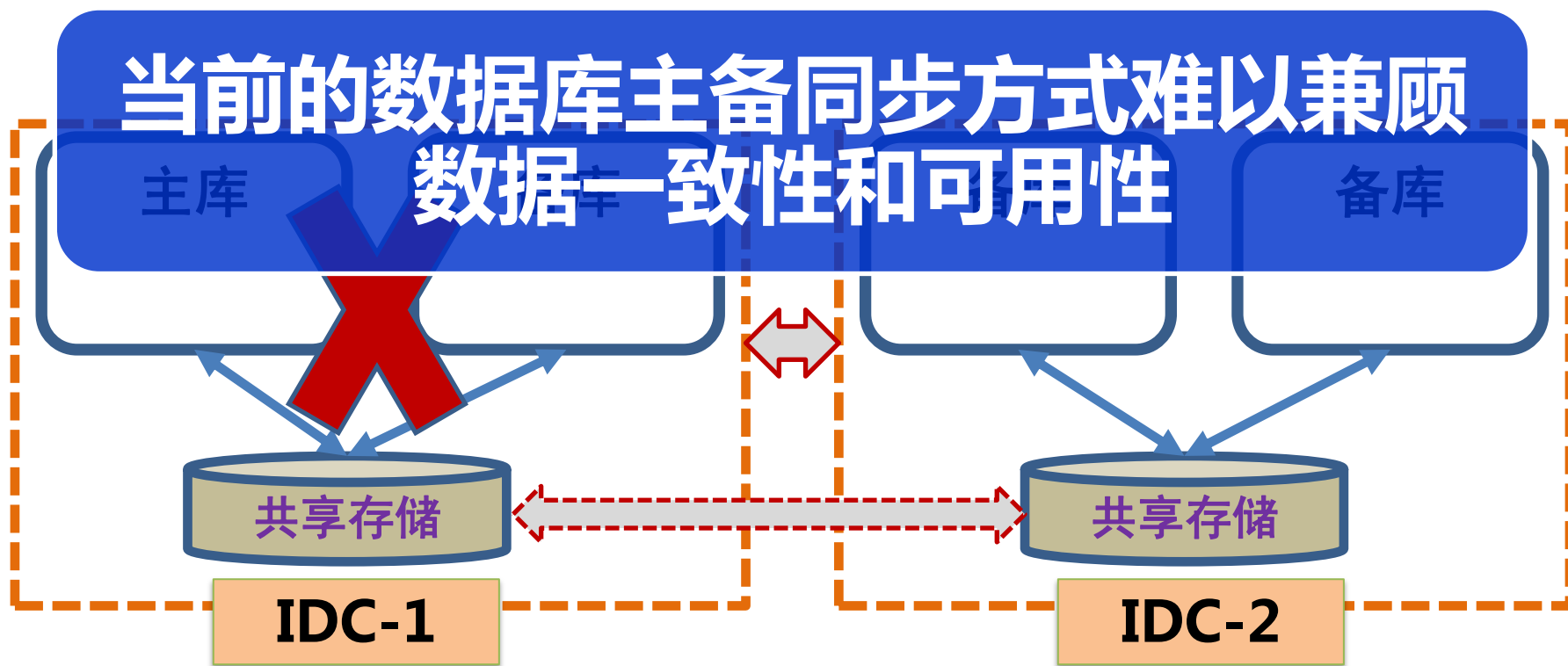
- 最大保护：事务同步到备库才算成功；主备库完全一致，但备库异常或主备间的网络异常则数据库停止工作
- 最高性能：主库完成事务即为成功，日志尽快同步到备库；性能高，但主库异常会丢数据
- 最大可用性：事务同步到备库或同步超时为成功；备库或主备之间的网络异常不影响业务，但主库异常可能丢数据

## ● 牺牲可用性，或者牺牲数据一致性



# 主库+备库：有共享存储

- 数据库的数据(包括redo log等)保存在共享存储上
- 共享存储数据的跨IDC同步：准实时
- 问题：主库IDC故障通常有数据丢失



# 数据库 vs. 现代硬件

- 数据库

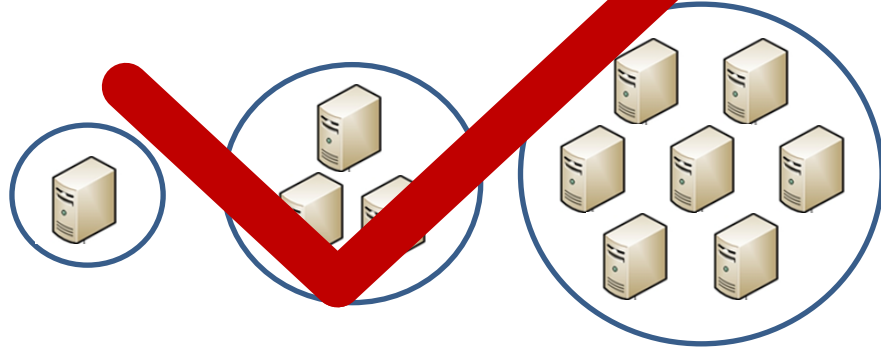
- 完全的数据一致性、极高的可用性(99.999%)

- 现代硬件

- 高可靠服务器/存储：可靠性高，但成本高、单机容量受限
- 主流服务器及磁盘：可靠性低，但性价比高、单机容量十分有限

不可靠硬件是否能  
构建可靠、可水平扩展的数据库？

- 当前主要关系数据库基本都是垂直扩展的单机系统

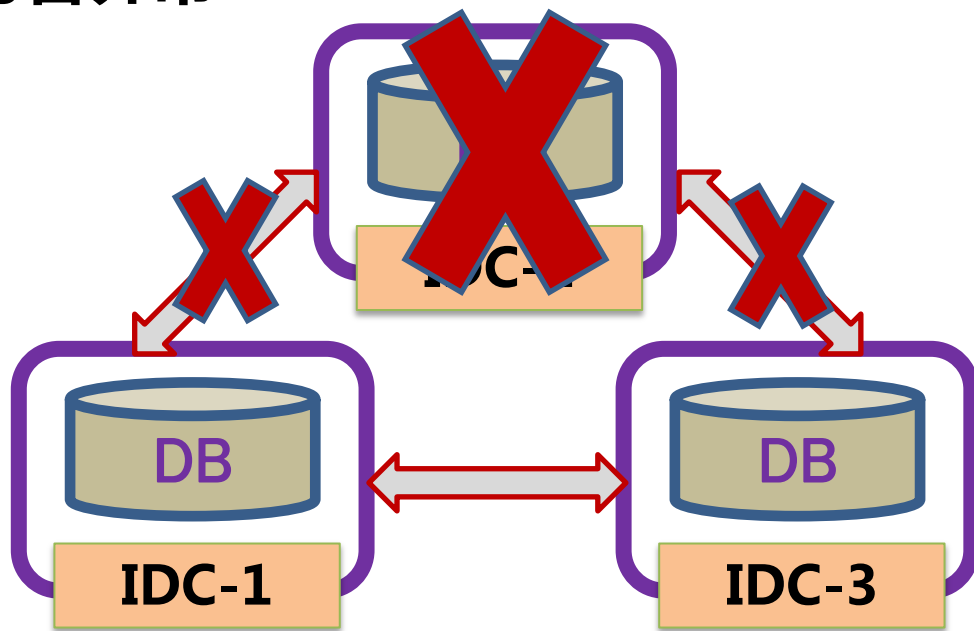




# 低可靠部件 vs. 高可靠服务

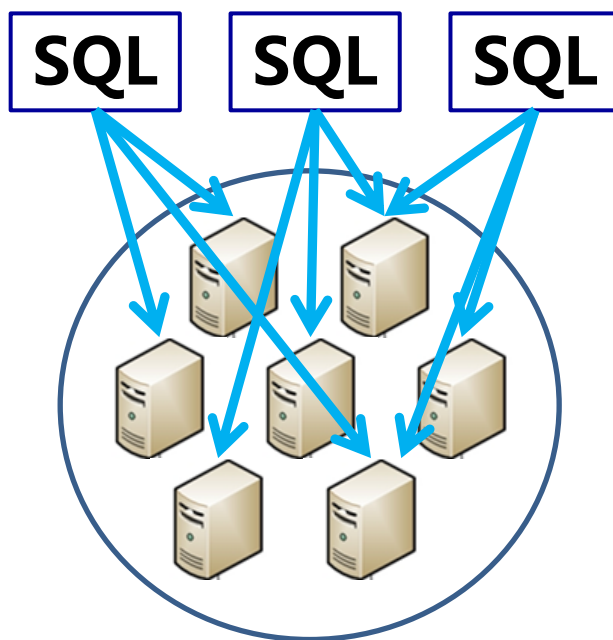
## ● 云计算技术

- 多副本+跨IDC：抵御服务器、网络、IDC故障
- 分布式投票：超过半数才成功，3/5/7个参与者抵御1/2/3个参与者异常



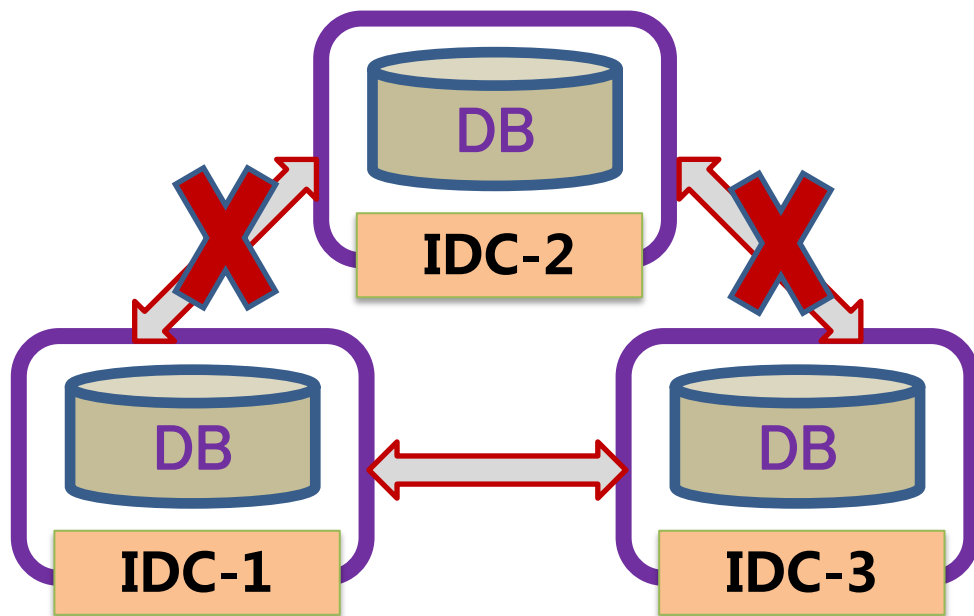
# 分布式关系数据库

- 事务的响应时间：适当的sharding
  - 例如，很多操作以用户为维度
- 系统的性能(RT)：适度的异步化
  - 例如，转账付款



# 一致性、可用性与网络分区(CAP)

- 折衷：数据强一致、较高的可用性(几十秒不可用)、一定的网络分区容忍(三个IDC中容忍一个被分区)



# 固态硬盘(SSD)

- 固态硬盘(SSD)

- 主流SSD：单盘几万+的随机读IOPS(机械磁盘：几百)
- 高端SSD：单盘几万+的随机写IOPS(机械磁盘：几百)
- 寻址与访问：块(4KB~64KB)

- 作为数据库的存储设备

- 只需随机读IOPS：主流SSD
- 需要随机读IOPS和随机写IOPS：高端SSD

# 内存(RAM)

- 单机内存：几十GB/百GB/TB
- 全内存的数据库
  - 性能高、性价比低、功耗大、容量小
  - 适合需要极高性能的业务
- 内存磁盘混合的数据库
  - 修改数据在内存、延后写入磁盘
  - 未修改数据在磁盘(SSD)：读性能好
  - 性能略逊、性价比高、容量大

# Non-Volatile Memory

- 带电容/电池的RAID/SAS卡
  - 加速fsync→事务commit加速
- FRAM (Ferroelectric RAM)
  - 4Mb容量(2013)
- MRAM (Magnetoresistive RAM)
  - 64Mb容量(2012)

# 小结

## ● 现状

➤ NoSQL 和 NewSQL 风起云涌，数据库几十年波澜不惊

➤ 扩展能力和性价比是两大软肋

对于关系数据库，  
云计算不是的终结者，  
而是一个新时代的开启者！

● 需求强劲：商务和互联网时代都繁荣

➤ ACID + 高可用

● 关系数据库突破的瓶颈

➤ 云计算：低可靠部件 → 高可靠服务，水平扩展能力

➤ 分布式关系数据库

## ● 现代硬件技术

➤ 固态硬盘(SSD)成为数据库的主要存储设备

➤ 内存磁盘混合型数据库是主流

➤ NVM还有待时日

# Thanks

**开源关系数据库OceanBase：**  
<http://alibaba.github.io/oceanbase/>

**明天下午第一场**

**杨传辉(日照)**

**OceanBase：可扩展的分布式数据库**



开源关系数据库OceanBase :  
<http://alibaba.github.io/oceanbase/>