

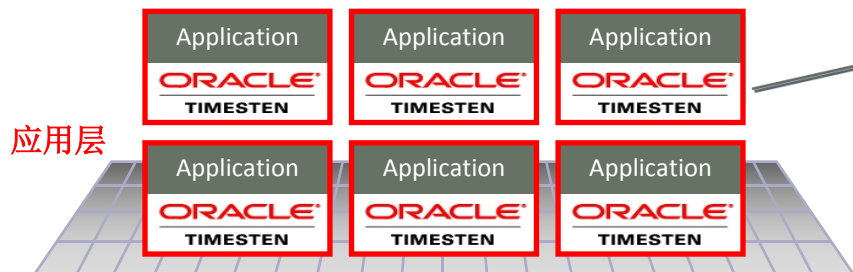
ORACLE®

Oracle TimesTen 内存数据库分布式演进之路

Jason Yang

Oracle TimesTen 数据库首席产品经理

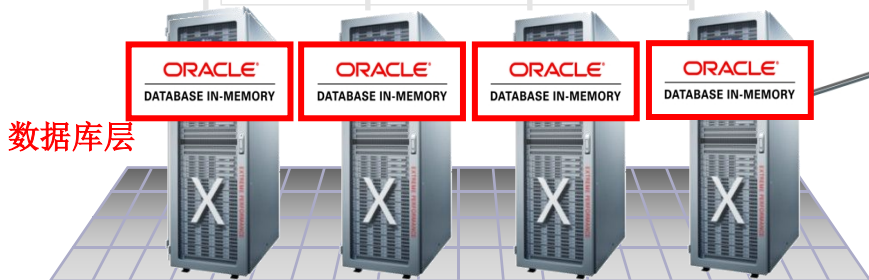
内存行存储技术



TimesTen 内存数据库

- 应用层嵌入式内存数据库解决方案
- 主要用途: **High Velocity (高速处理)** OLTP
- 微秒级的响应时间
- OLTP 应用使用最广泛的内存数据库产品

内存列存储技术



Oracle Database In-Memory Option 组件

- 双模内存数据库
- 主要用途: : **实时分析、OLTP / 海量数据分析混合事务**
- 每秒十亿条数据分析处理效率
- 透明的存储分层: 物理内存, 闪存, 磁盘

The Forrester Wave™: In-Memory Databases

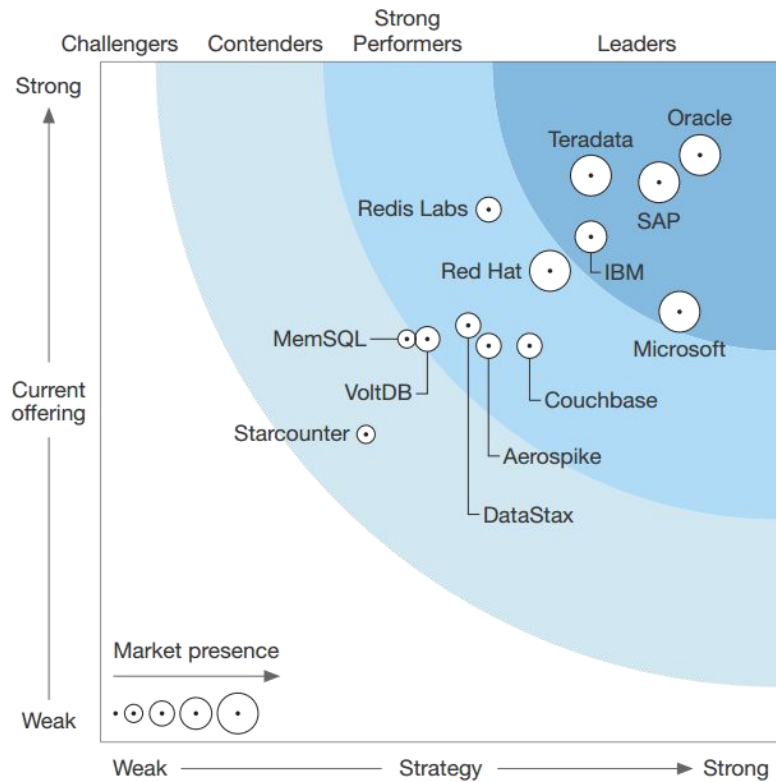
CN SOUG
中国南方ORACLE用户组

IT大咖说
知识共享平台

Oracle In-Memory Databases
Scored Highest by Forrester
on both Current Offering
and Strategy

<http://www.oracle.com/us/corporate/analystreports/forrester-imdb-wave-2017-3616348.pdf>

The Forrester Wave™ is copyrighted by Forrester Research, Inc. Forrester and Forrester Wave™ are trademarks of Forrester Research, Inc. The Forrester Wave™ is a graphical representation of Forrester's call on a market and is plotted using a detailed spreadsheet with exposed scores, weightings, and comments. Forrester does not endorse any vendor, product, or service depicted in the Forrester Wave. Information is based on best available resources. Opinions reflect judgment at the time and are subject to change.



TimesTen: 主要用途

高速处理应用系统

- 什么是高速处理应用系统？
 - 事件驱动，难以成批次
 - 例如 电话呼叫，股票交易，信用卡授权，需要**立即**处理
 - 常见于轻量级事务 – 行数不多，超高的交易量
 - TimesTen 是这些场景的理想选择
- TimesTen 轻量&超快
 - 运行在应用系统内：不需要网络
 - 为延迟关键型应用系统带来 **30 倍**性能提升
 - 上千家客户的选择



Agenda

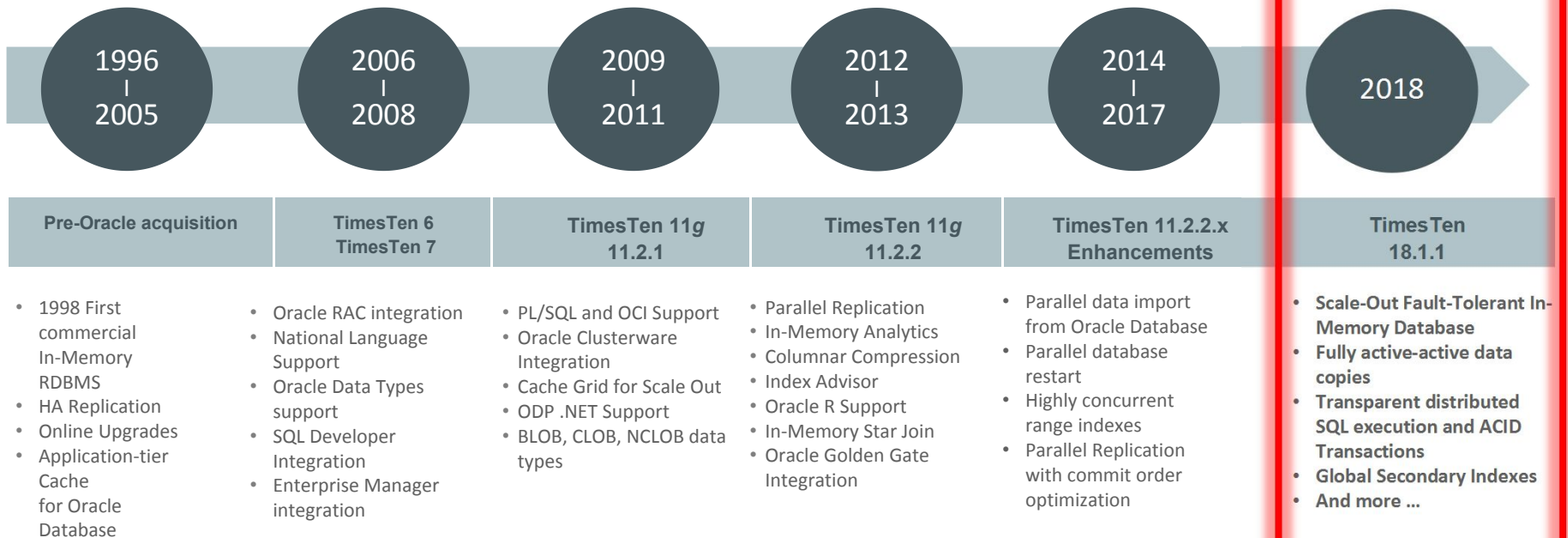
- 1 TimesTen 当前产品
- 2 TimesTen 分布式内存数据库新产品
- 3 Demo
- 4 Q & A

Oracle TimesTen – 关系型内存数据

CN'SOUG
中国南方ORACLE用户组

IT大咖说
知识共享平台

20+ 年极致性能体验

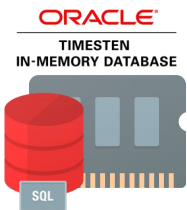


Agenda

- 1 TimesTen 当前产品
- 2 TimesTen 分布式内存数据库新产品
- 3 Demo
- 4 Q & A

Oracle TimesTen 内存数据库

应用层微秒级响应速度

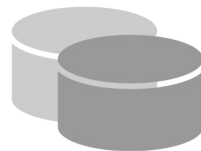


关系型数据库

- 纯粹内存计算
- 兼容 ACID
- 标准 SQL
- 整个数据库驻留物理内存

持久性和可恢复性

- 数据库和事务日志永久存放于磁盘和闪存存储 (TimesTen 本地存储)



极速性能

- 微秒级响应速度
- 超高的吞吐量



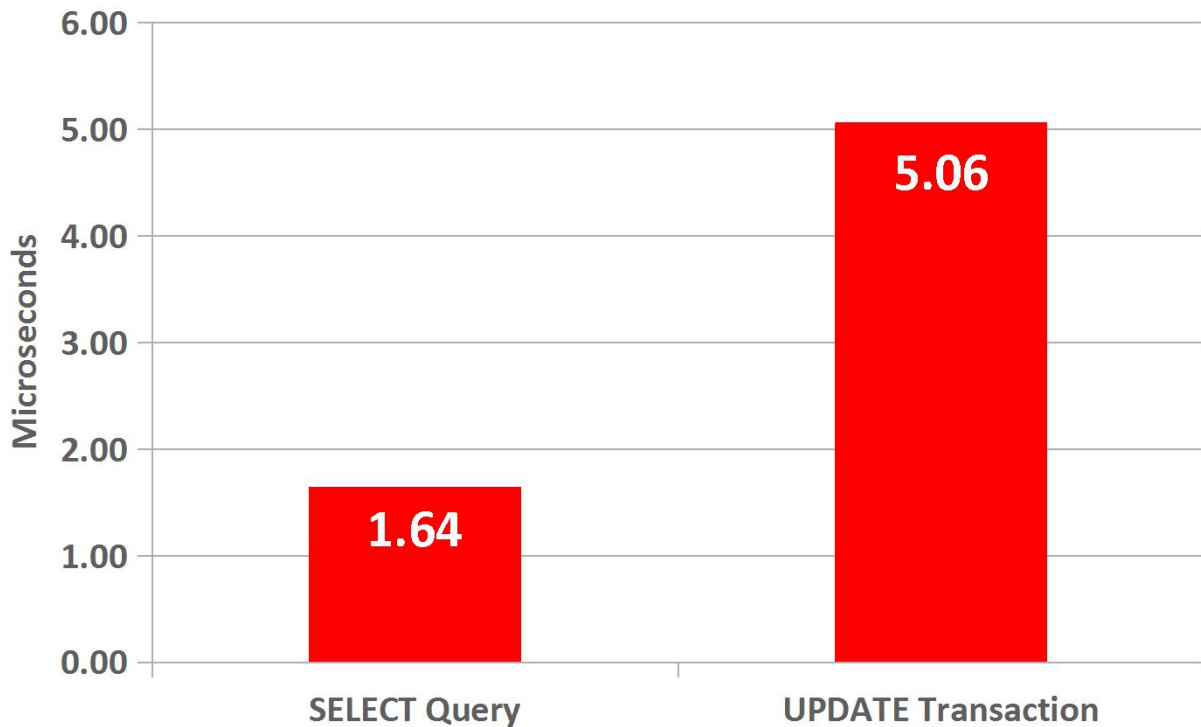
兼容 Oracle 数据库

- 数据类型, PL/SQL, OCI, ODP.NET, PHP, R 编程语言
- 集成 RAC, Data Guard, GoldenGate, Enterprise Manager, SQL Developer 等



TimesTen 内存数据库

低延迟 – 微秒级响应时间

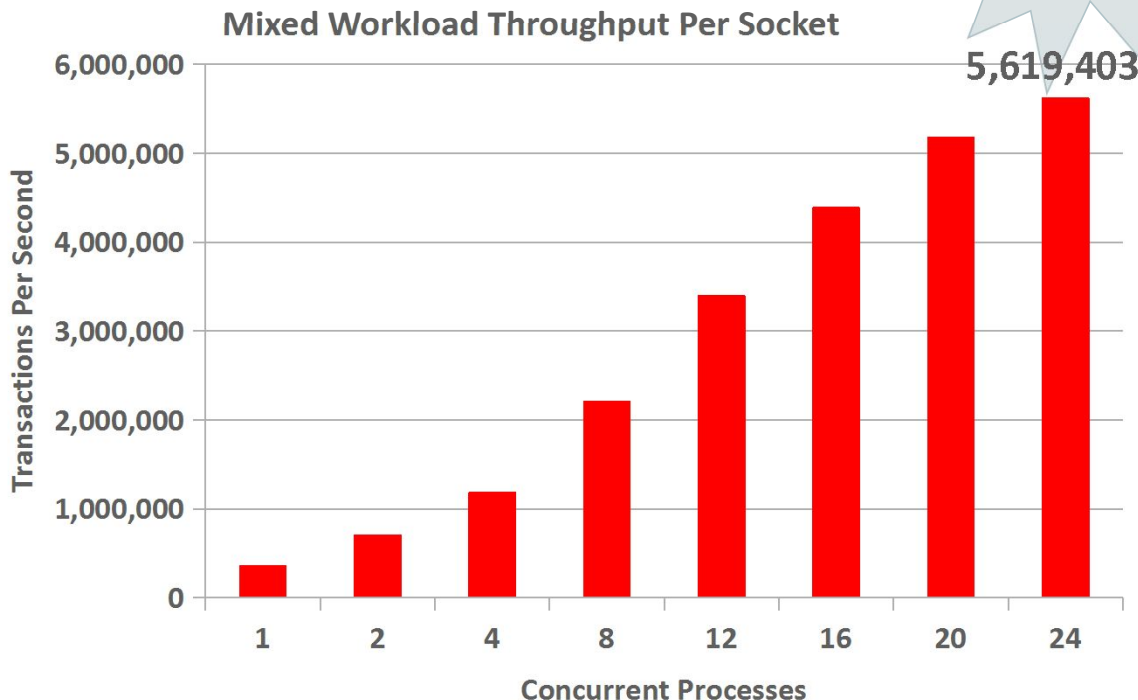


TPTBM Read and Update
E5-2699 v4 @ 2.20GHz
2 socket, 22 cores/socket,
2 threads/core
TimesTen 11.2.2.8.0
(100M rows, 17GB)

TimesTen 内存数据库

每个处理器每秒处理五百六十万笔事务

百六十万笔
事务

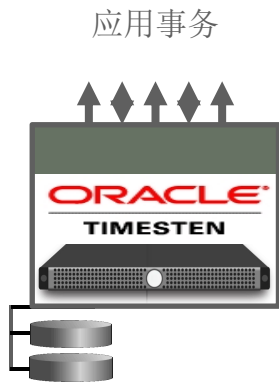


TPTBM 100% Mixed
Workload (80-10-5-5)
E5-2699 v4 @ 2.20GHz
2 socket, 22 cores/socket,
2 threads/core
TimesTen 11.2.2.8.0
(100M rows, 17GB)

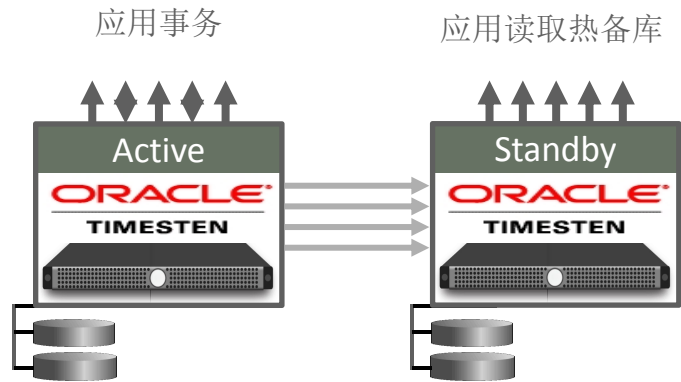
80-10-5-5 Workload = 80% select, 10% updates, 5% inserts, 5% deletes

断电怎么办？

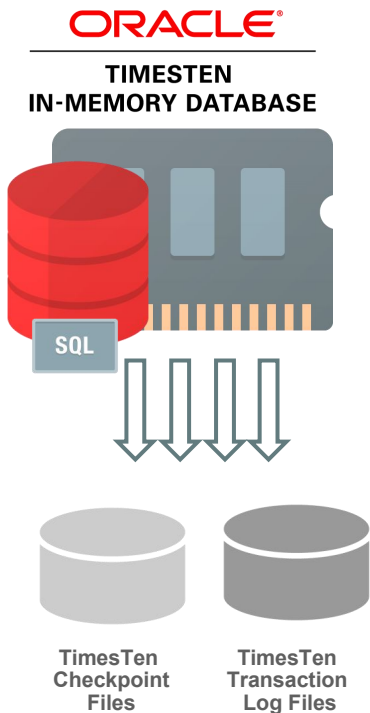
方案一：本地持久化



方案二：内存复制高可用



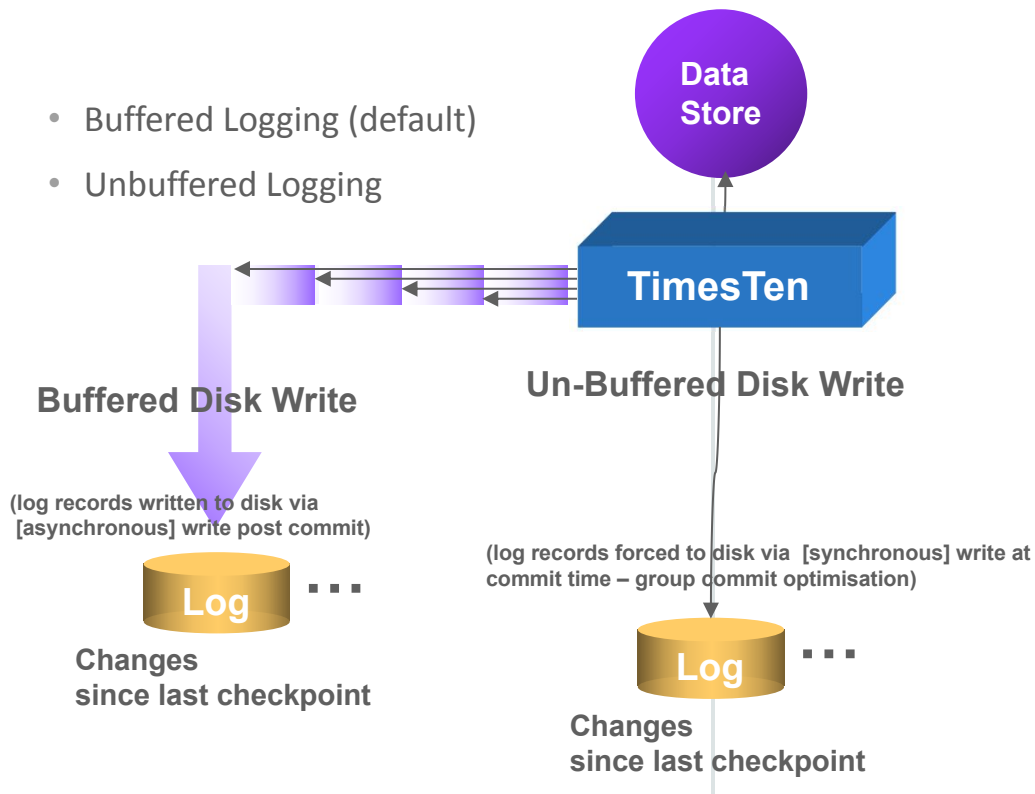
TimesTen 内存数据库 持久性和恢复



- TimesTen 数据库持久性可以配置到
 - 闪存, 固态硬盘, 磁盘存储设备
- 所有事务都被记录并持久化
 - Redo, undo 和 recovery
- 双数据库检查点文件
 - 数据库重启
 - 数据库发生异常后的恢复

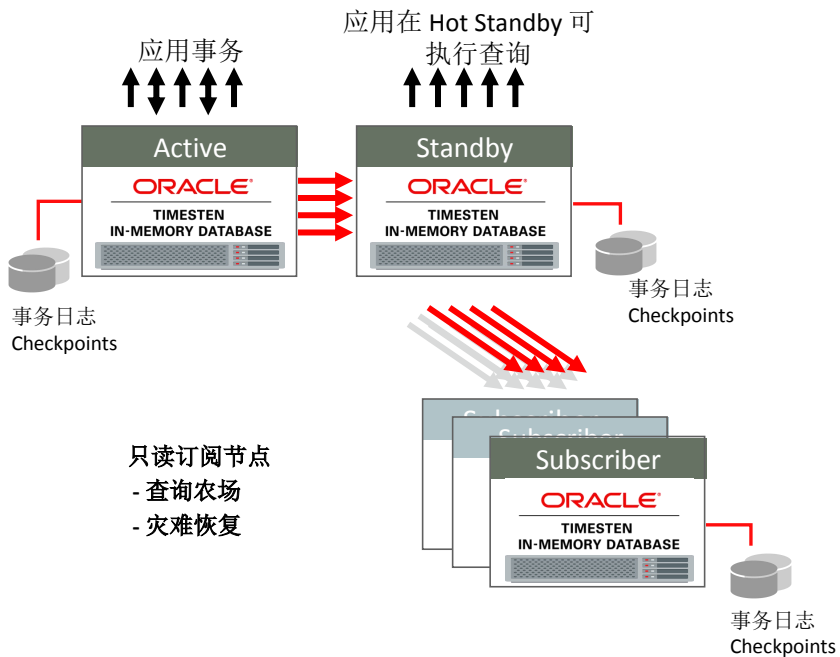
TimesTen Logging 机制

- Buffered Logging (default)
- Unbuffered Logging



实时事务复制

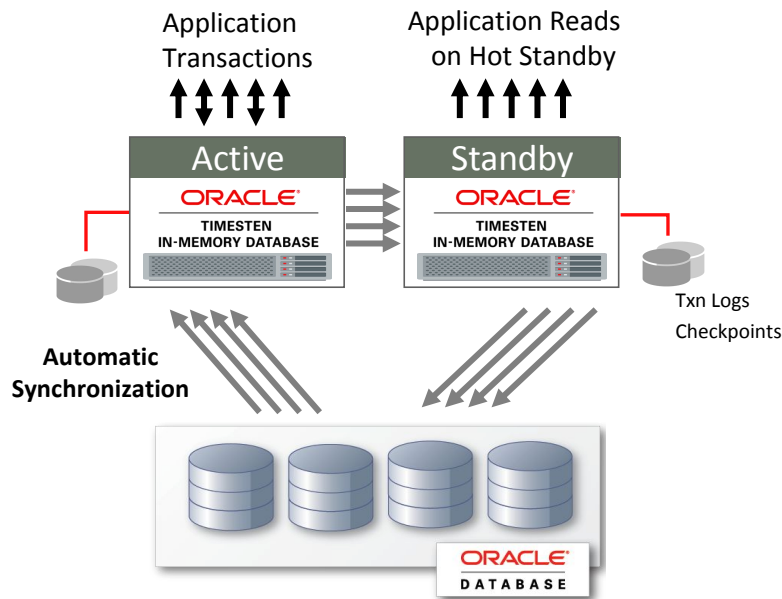
高可用性和灾难恢复



- 高性能
 - 同步 / 异步
 - 并行发送日志流
 - 并行应用变化到 Standby 和 Subscribers
- 支持高可用性和灾难恢复
- 在线升级
 - 零宕机
 - 跨版本复制
- 集成 Oracle Clusterware

TimesTen 缓存高可用技术

High Availability



- 读写缓存
 - 并行的主备库间事务复制
 - 并行的事务写入到Oracle 数据库
- 只读缓存
 - 多数据流的刷新来自Oracle数据库的事务
 - 将刷新的事务并行复制到备库
- 即便Oracle 数据库发生异常，应用仍正常工作

电话坐席分发系统

应用概览

- 行业： 金融
- 业务： 保险、银行、投资
- 应用： 坐席分发系统
 - 基于预定义规则自动分发来电
 - 手动重新指派新的代理

挑战

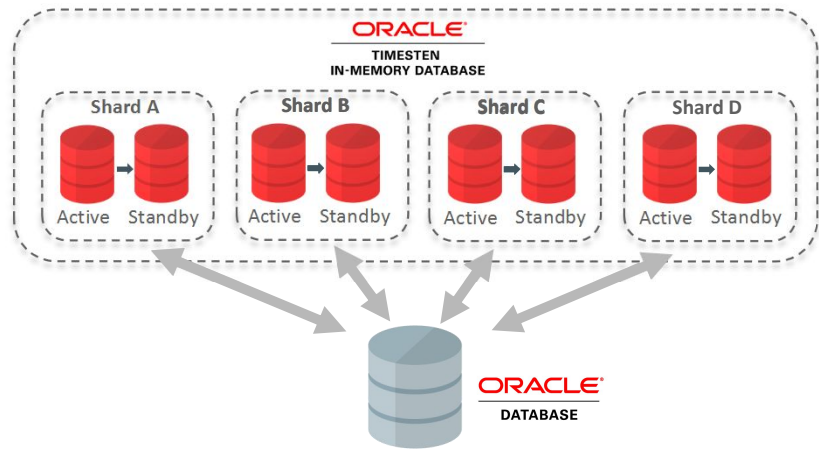
- 超大并发下的数据库扩展能力严重影响端到端响应时间
- 维护客户满意度
- 对现有架构和应用进行最少的应用改造
- 必须具有高可用性

解决方案

- Oracle TimesTen Application-Tier Database Cache
- TimesTen Replication for High Availability
- Oracle Database

为什么选择 TimesTen?

- 提供低且一致的响应时间；获得四十倍的吞吐量和响应时间的性能提升
- 数据可以在 TimesTen 和 Oracle 数据库之间自动同步
- 具备内置的高可用性、支持自动的 failover 和 switchover

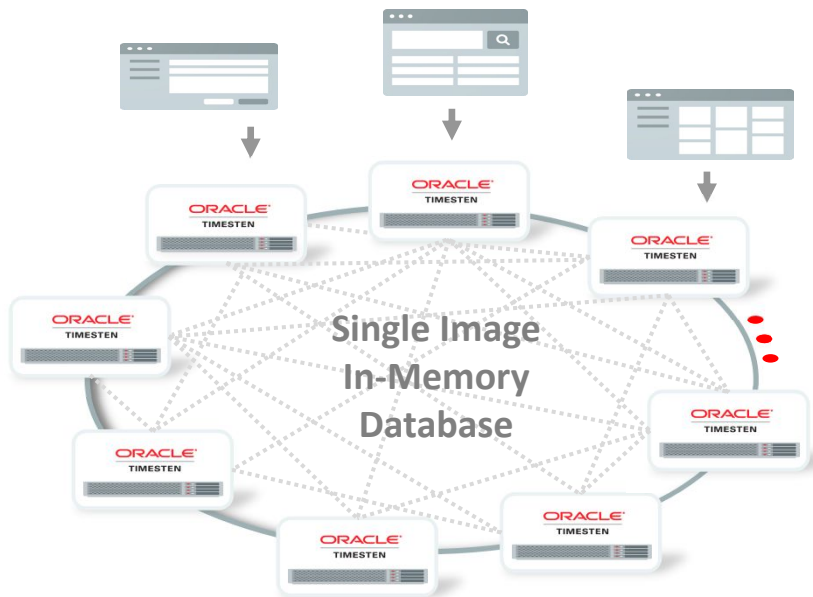


Agenda

- 1 TimesTen 当前产品
- 2 TimesTen 分布式内存数据库新产品
- 3 Demo
- 4 Q & A

Oracle TimesTen Velocity Scale Database

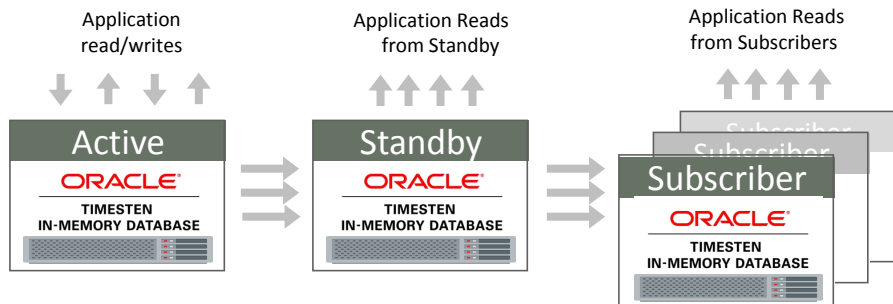
一款高效、无共享、横向扩展的关系型数据库



Velocity Scale 集单实例和分布式技术优点于一身

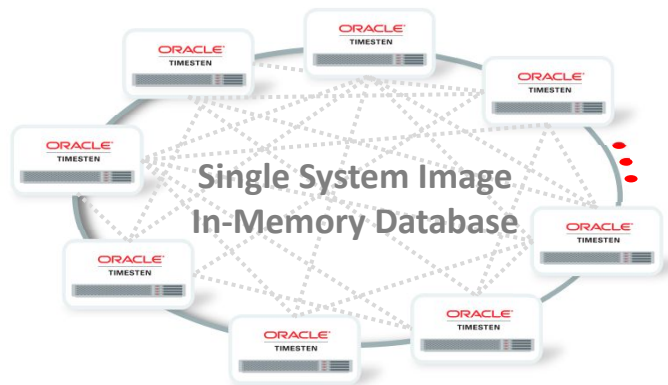
一个产品两种部署方式

TimesTen 传统模式



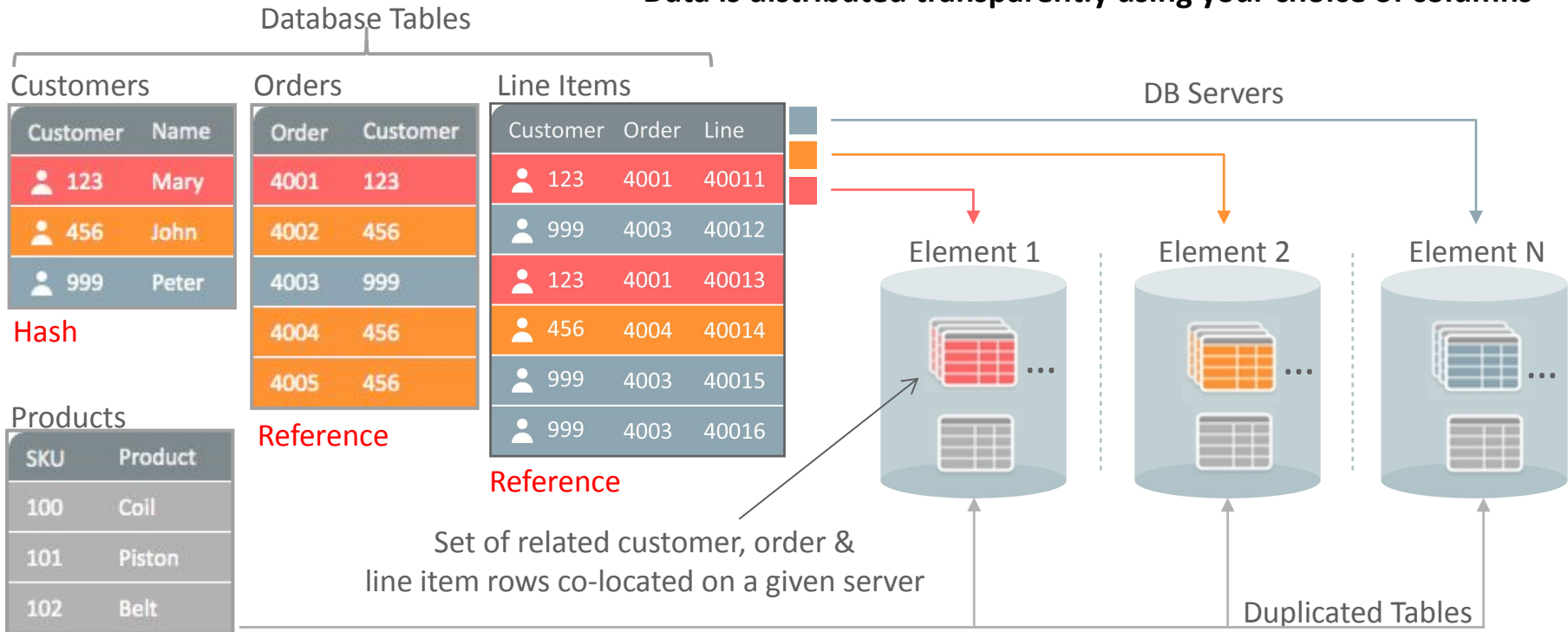
- 关系型内存数据库的复制技术
- 超低延时的读写能力
- 高可用性
- 查询可扩展、但写操作无法扩展

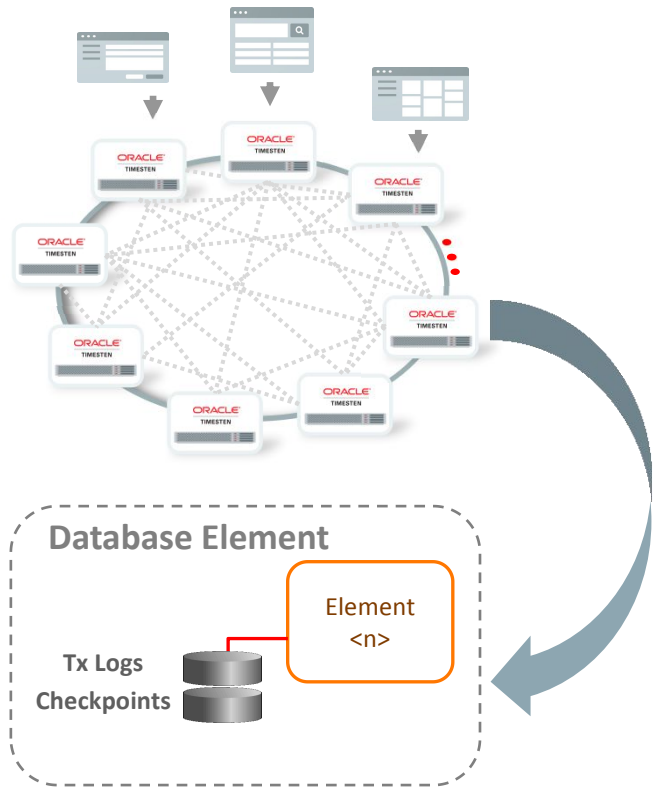
TimesTen 分布式模式



- 横向扩展的关系型内存数据库
- 超高吞吐量的读写能力
- 全新的高可用
- 读、写操作均可扩展

Data is distributed transparently using your choice of columns

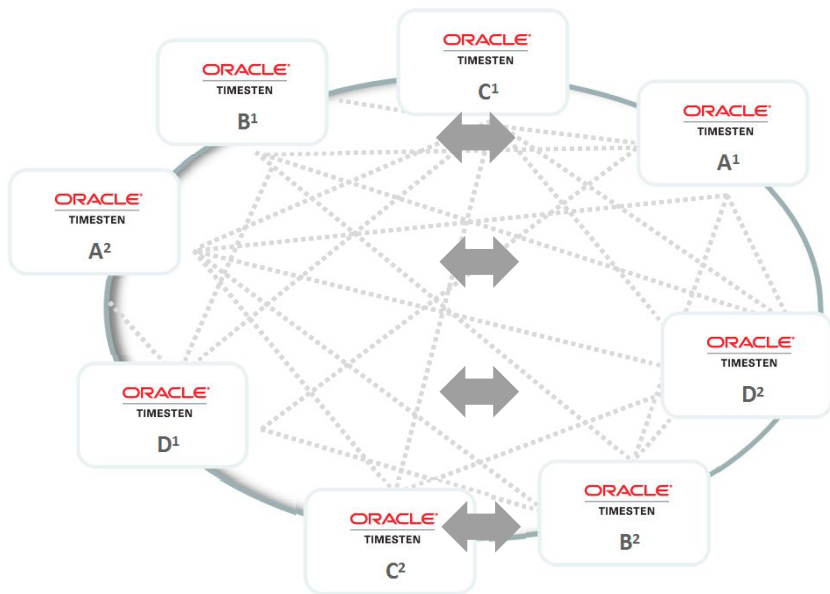




- 每个element有其自己的用于持久化的检查点文件和事务日志文件
- Element是数据库用于灾难恢复和高可用的最小单元

高可用性和最大吞吐量

K-Safety, 多活技术



- 通过多数据副本技术实现内置高可用(K-safety)
 - 自动保持同步
- 所有副本都是active的，且都可以执行读操作和写操作
- 可以在任意一个副本发起和执行事务操作

Replica Sets 和 Data Spaces

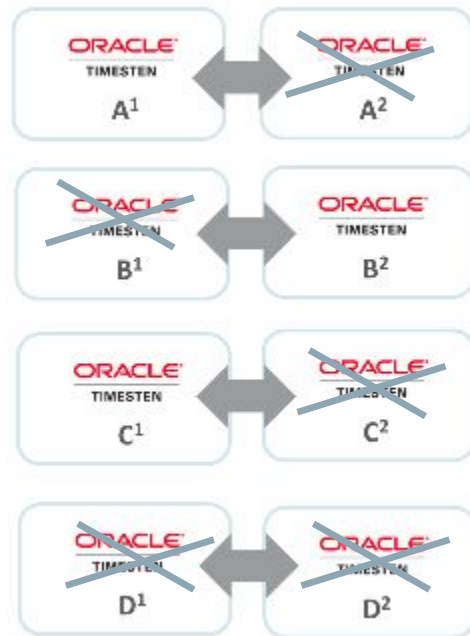
- 一个 Replica Set 是一组具有相同数据集的 element
- 一个 Data Space 是包含一个完整数据库副本的 element 集合
- Data spaces 正交于 Replica Sets
- Data Space 的数量等于 K 的系数
- Data Spaces 用于定义“容灾组”

	机柜 1	机柜 2
	Data Space 1	Data Space 2
Replica Set A	Element A ₁	Element A ₂
Replica Set B	Element B ₁	Element B ₂
Replica Set C	Element C ₁	Element C ₂
...
Replica Set Z	Element Z ₁	Element Z ₂

数据库容错能力 – 应用无宕机

提供一套完整可用数据库副本

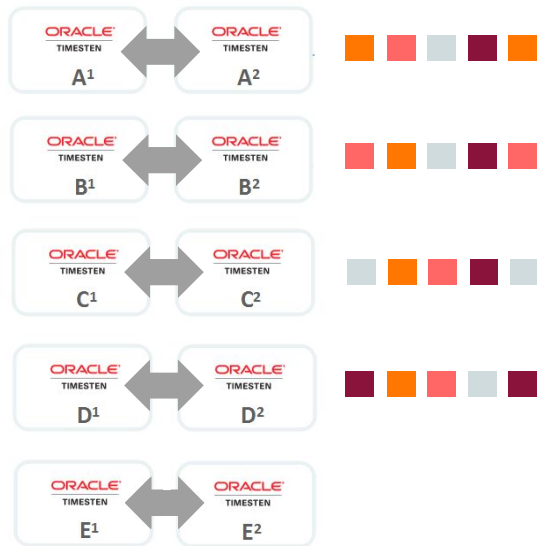
- 即便多个element发生异常，只要有一套完整可用副本，应用无影响
- 修复异常后，自动恢复技术
- 如果整个replica set 发生异常，应用仍可以显示接受余量结果集

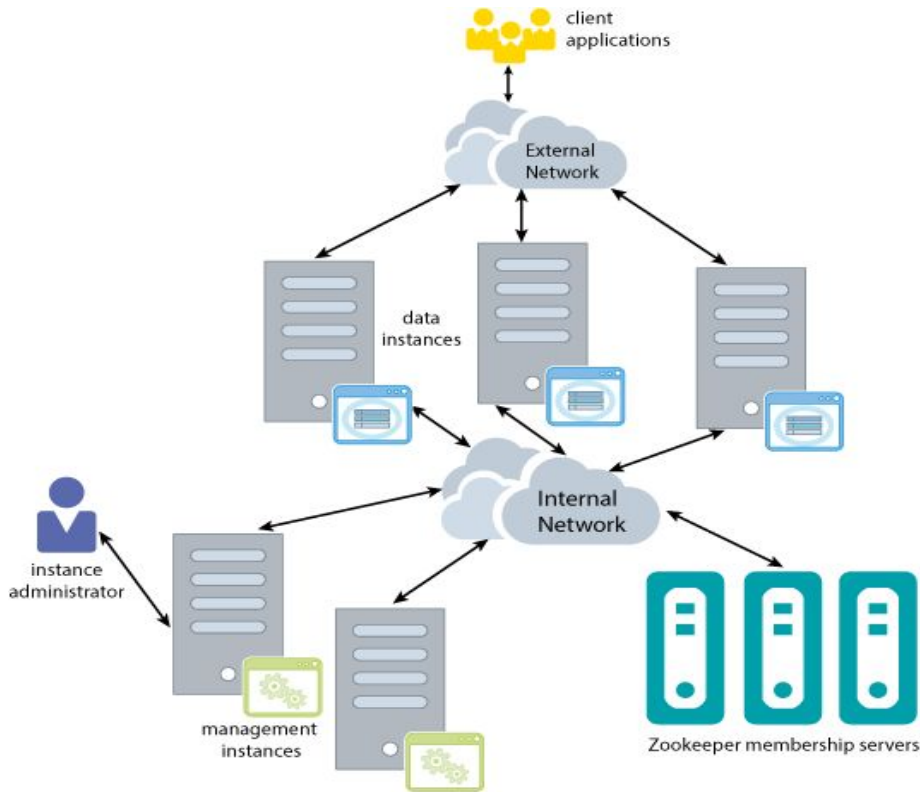


弹性扩展

随业务需要而扩展或收缩数据库

- 向网格中添加或删除 elements
 - 数据自动重新分布到新的 elements 中
 - 工作负载自动使用新的 elements
 - 连接会开始使用新的 elements
 - 吞吐量随新增加的计算资源而增长





一组数据实例

- 包含有 tables, indexes, views 等
- 一个主机可以创建多个实例
- 一个实例可以创建多个 elements

• 管理实例

- 用于在一处统一管理网络
- 使用一组 Active Standby Pair 实现高可用
- 单个数据库可以用于开发和测试
- 管理实例数据库存放网络对象信息
- 定义网络状态
- 网络对象版本化

• 一组 Zookeeper 服务器

- 用作成员服务管理
- 生产环境建议使用3到5个用作高可用
- 单个ZK服务器可以用于开发和测试

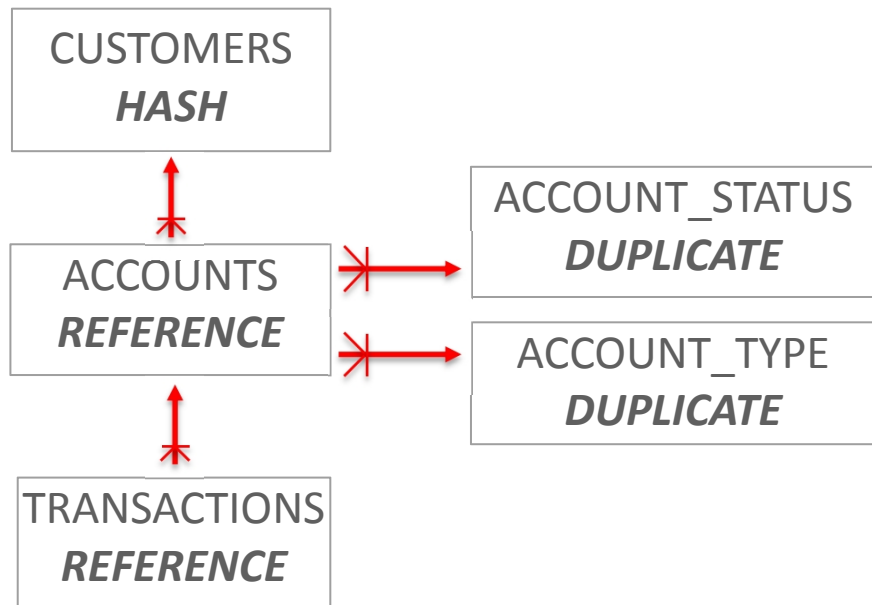
Agenda

- 1 TimesTen 当前产品
- 2 TimesTen 分布式内存数据库新产品
- 3 Demo
- 4 Q & A

演示内容

- **GUI Admin and Development Tool**
 - 网格配置、监控和管理图形界面化
 - Schema 创建、查询以及开发
- **Application Transparency**
 - 连接任意 element, 访问所有数据
- **High Availability and Fault Tolerance**
 - 内置高可用
- **Elasticity Scalability**
 - 添加 elements 实现弹性可扩展
 - 数据重分布

Demo Data Model



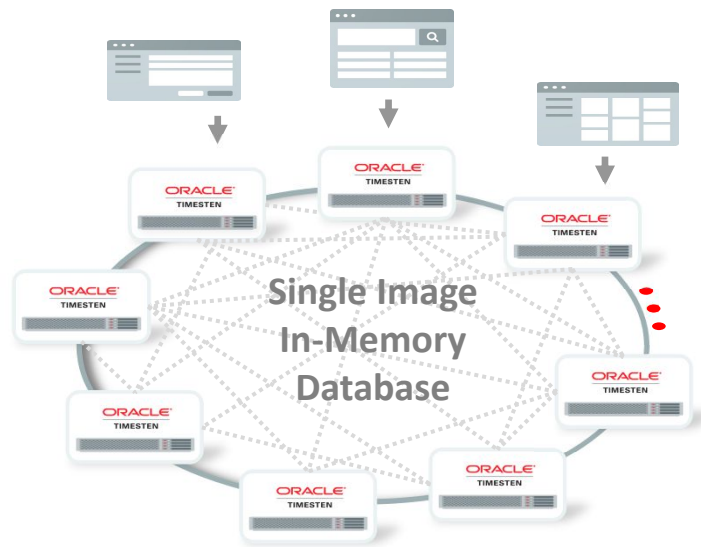
演示现在开始！

Agenda

- 1 TimesTen 当前产品
- 2 TimesTen 分布式内存数据库新产品
- 3 Demo
- 4 Q & A

Oracle TimesTen Velocity Scale Database

- 横向扩展、无共享架构设计开发
- 数据位置透明
- 内置 K-safety 高可用技术
- 弹性伸缩能力
- 易部署、易管理
- 易开发 - 面向OLTP的应用
- 内部或云上部署



Q & A

QUESTIONS & ANSWERS

Hardware and Software Engineered to Work Together

ORACLE®