火龙果•整理
uml.org.cn

ORACLE

ORACLE THIRTY YEARS 30

高性能数据库一体机 -- Oracle Exadata V2

孙戎
甲骨文(中国)软件系统有限公司

火龙果•整理
uml.org.cn

Agenda

- 企业信息化面临的问题
- 传统的解决方案
- 高性能 & 高智能的理想架构
- Oracle为企业信息化提供的解决之道
- Q&A

火龙果•整理
uml.org.cn

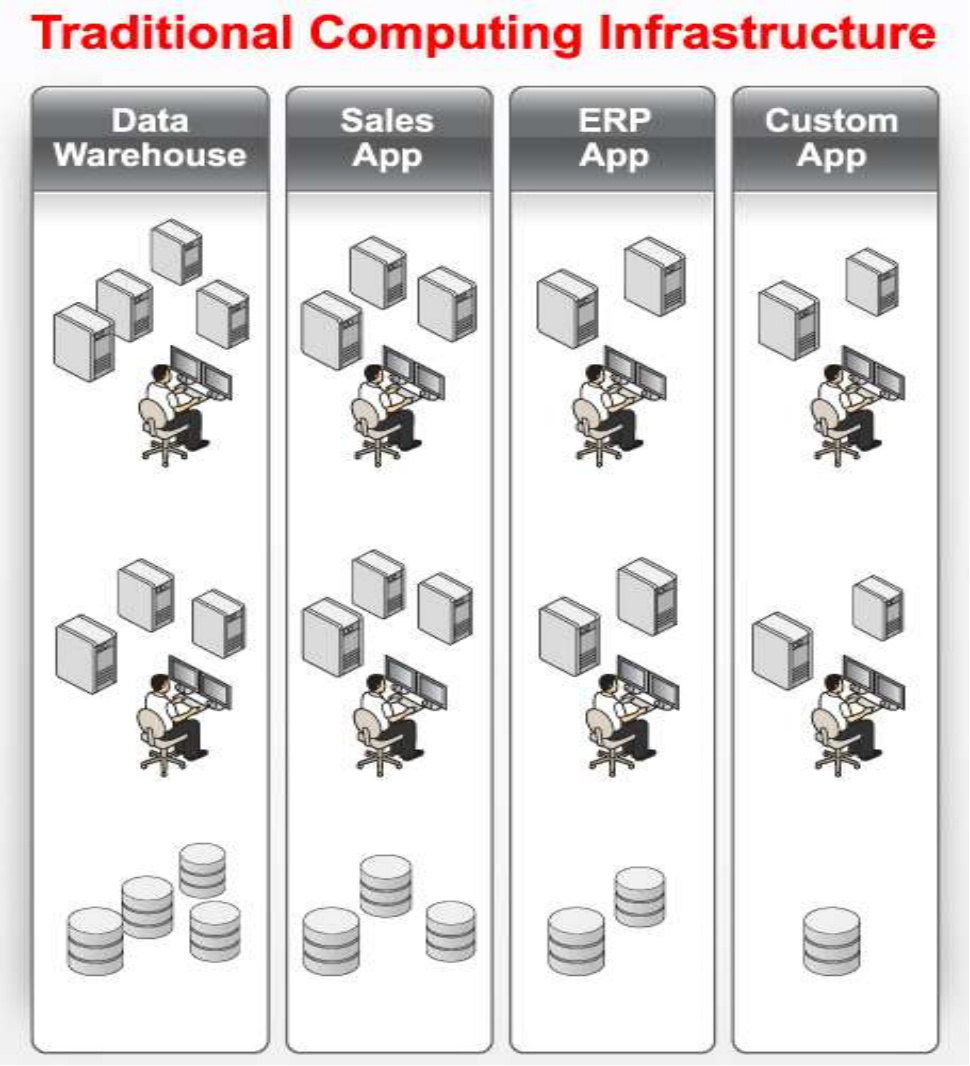
ORACLE

ORACLE THIRTY YEARS 30

企业信息化面临的问题和关注点

企业信息化面临的问题

昂贵的、低效的

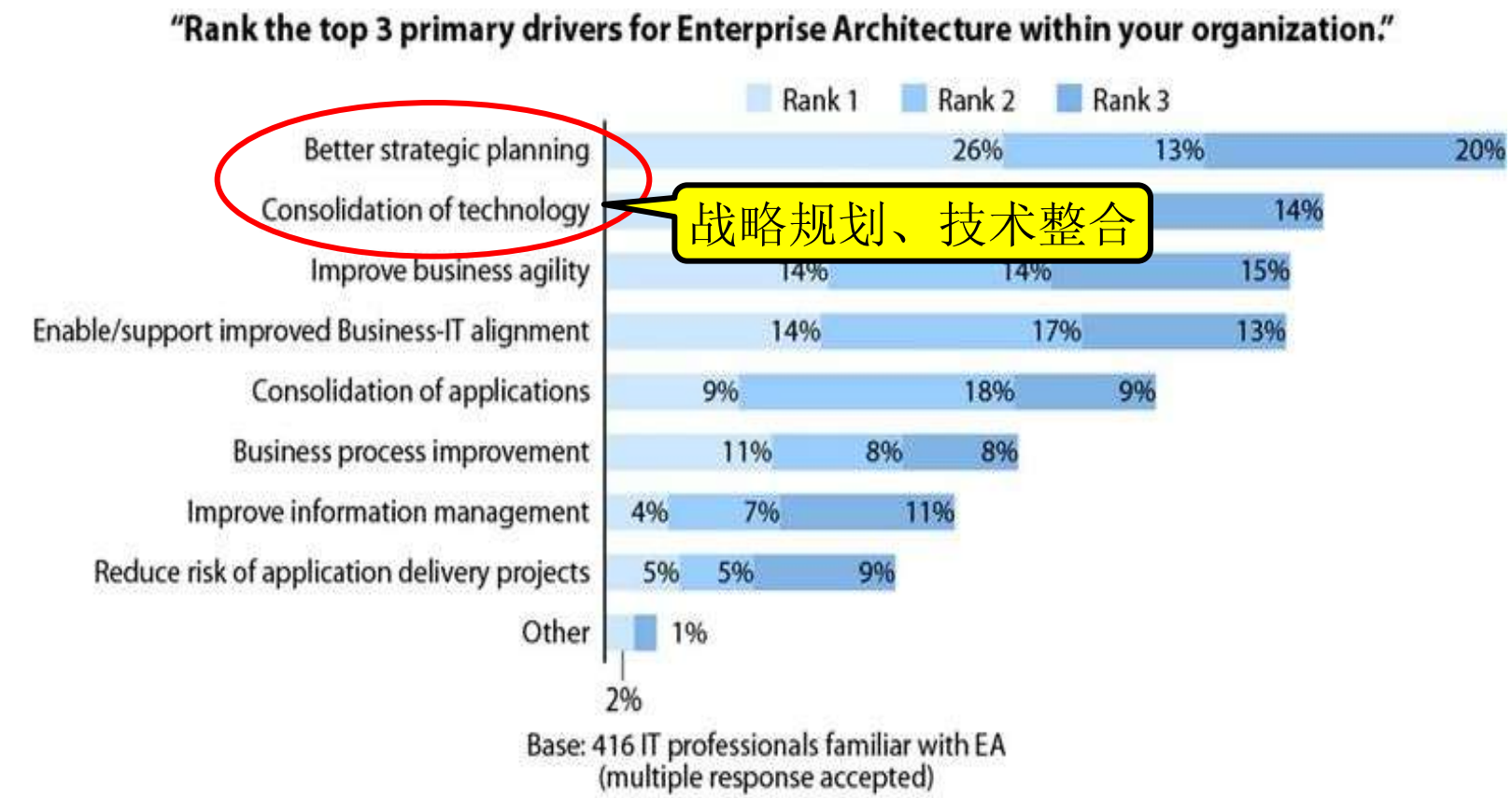


- 信息孤岛导致资源浪费
- 每个系统都是最高性能配置
- 性能仍然不能满足要求
- 很难进行扩展
- 数据平台不能统一

火龙果•整理
uml.org.cn

火龙果•整理
uml.org.cn

企业信息化的关注点.....



Source: September 2009 Global Annual State of Enterprise Architecture Online Survey

Source: Forrester Research, Inc.

火龙果•整理
uml.org.cn

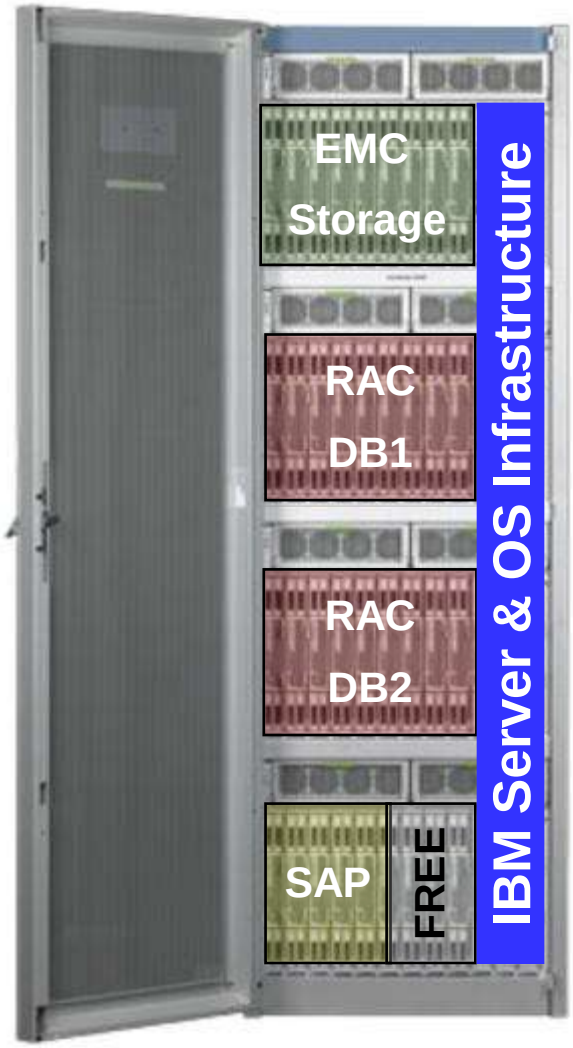
ORACLE

ORACLE THIRTY YEARS 30

传统的解决方案

传统SMP架构组成

火龙果•整理
uml.org.cn



- 硬件架构
  - CPU
  - Memory
  - Disk Array
- 软件架构
  - OS
  - Database (DW、OLTP)
  - Application

火龙果•整理
uml.org.cn

传统数据中心

硬件架构

- 如何优化配置资源？

  --- 安装配置

  --- 维护管理

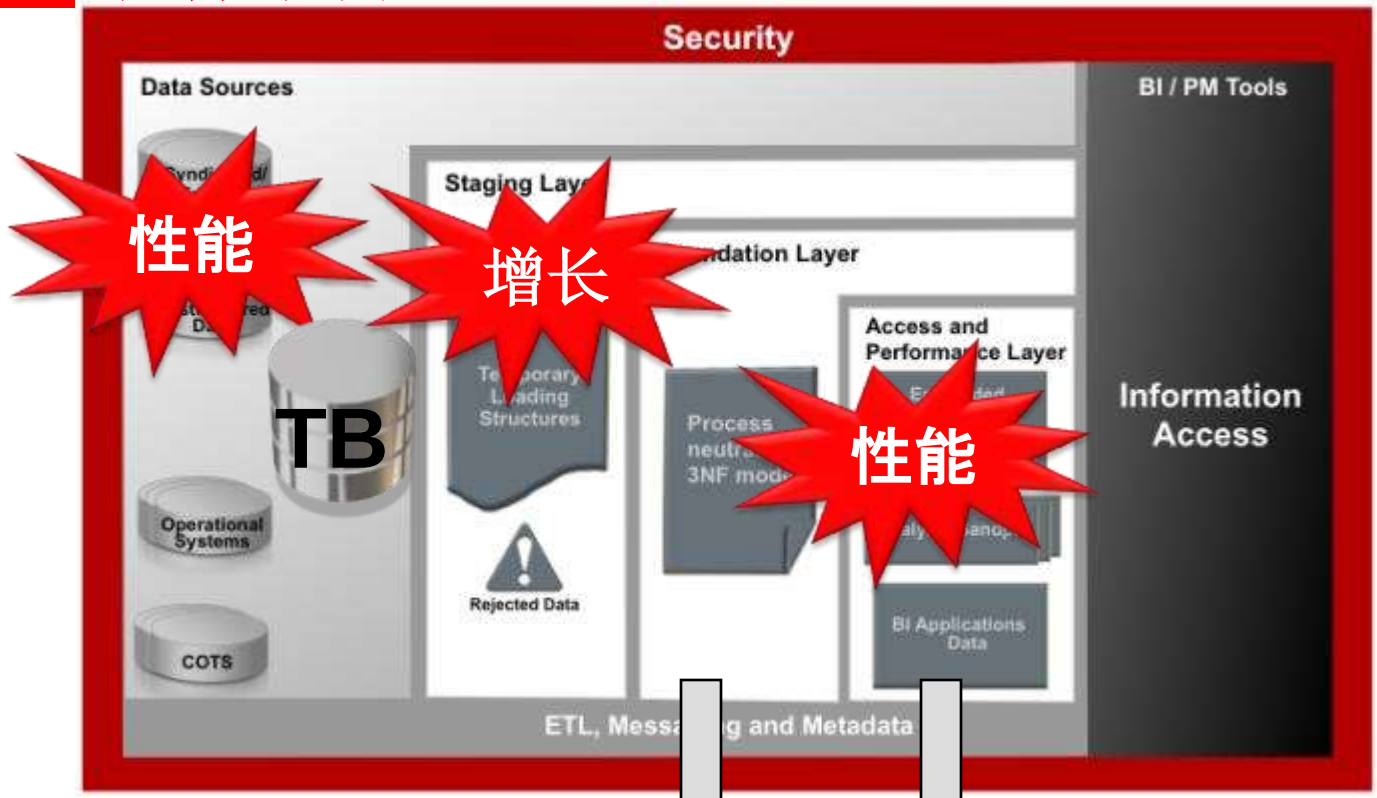
  --- 性能优化

  --- 监控管理

  --- 服务器扩展

传统SMP架构

软件架构-DW系统



IO-bandwidth intensive workloads

Random-IO intensive workloads

火龙果•整理
uml.org.cn

传统SMP架构

软件架构-OLTP系统

Memory

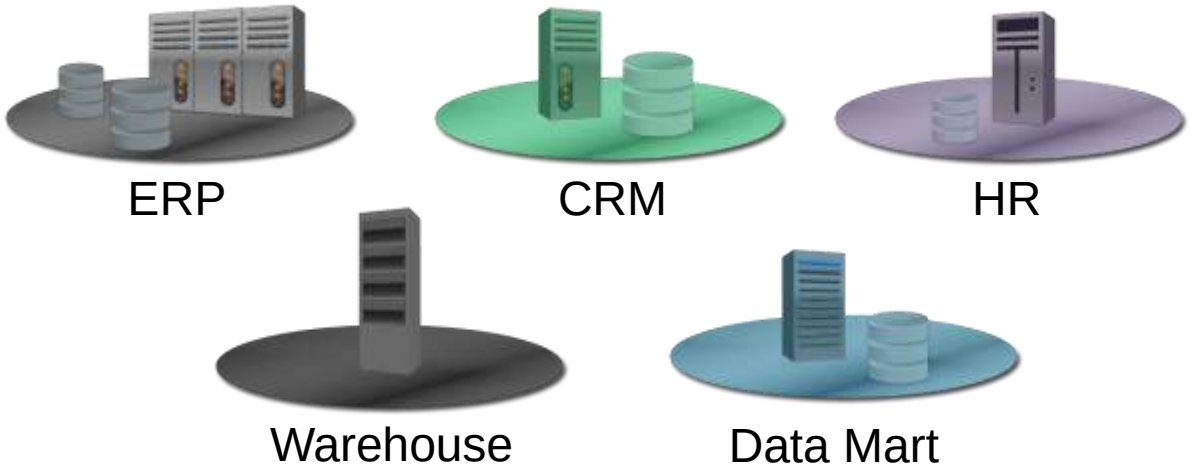
CPU是否还是100%，查询仍然很慢

CPU

火龙果•整理
uml.org.cn

混合系统如何优化、分配资源？

软件架构- Application系统



- 系统之间资源不能充分利用
- 单一硬件成本在不断增长
- 系统的响应时间、性能在下降
- 异构系统之间的数据访问压力在增加

火龙果•整理
uml.org.cn

火龙果•整理
uml.org.cn

ORACLE

ORACLE THIRTY YEARS 30

高智能 & 高性能的理想架构

火龙果•整理
uml.org.cn

构建高性能IT基础平台的理想架构

- 保障系统中没有单点故障（SPOF-Single Point of Failure), 支持业务连续性
- 在系统中消除性能瓶颈 （Performance Bottle-Neck）, 构建一个平衡的系统
- 能有效地进行系统的扩展， 满足不断增长的数据和用户需求
- 智能管理，提高整体的业务性能和业务管理
- 统一数据，更实时、有效地整合数据资源

火龙果•整理
uml.org.cn

ORACLE

ORACLE THIRTY YEARS 30

Oracle为企业信息化提供的解决之道

既然成熟的技术，为什么性能会如此之

火龙果•整理
uml.org.cn

硬件和软件，集成设计、卓越性能

Hardware and Software

ORACLE

Engineered to Work Together

- 工程调校
  - CPU主频越快，主机性能就越好？
  - 发动机越好，排量越大，汽车就跑的越快？

Exadata v2 – 高性能体系结构

火龙果•整理
uml.org.cn

智能数据库一体机

最快的数据仓库

最快的OLTP

最强的混合系统

Oracle Exadata数据库机：“数据库云”的

火龙果•整理
uml.org.cn

- **软硬件互为优化，提供极限性能**
  - 提供更宽、更多的IO通道，成千倍减少IO传输量
  - 1,500,000次的IOPS，每秒75GB的IO吞吐量
  - 数据加载 12TB/小时
- **OLTP、OLAP混合负载**
  - 基于任务的系统资源管理
- **统一智能管理，运维简单**
  - QOS、SLA
- **绿色数据中心**
  - 电耗、散热、承重低

Exadata V2

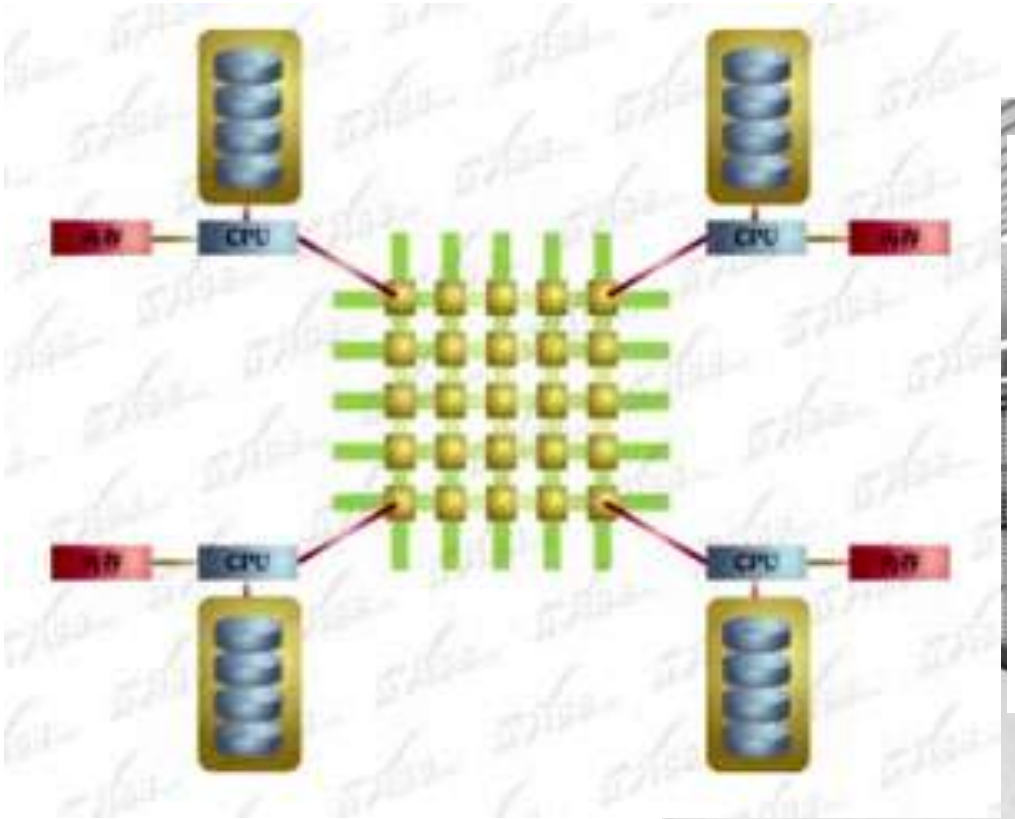
火龙果•整理
uml.org.cn

- 完整的, 预配置的, 严格测试的超级性能系统
  - 数据库服务器
  - **Exadata**存储服务器
  - **InfiniBand**交换机
  - **Ethernet**交换机
  - 出厂完成线缆连接
  - **Keyboard, Video, Mouse (KVM) hardware**
  - **Power Distribution Units (PDUs)**
- 随付即用
  - 插电
  - 联网
  - 启动数据库

为什么快？ -- 硬件MPP体系架构

火龙果•整理
uml.org.cn

- 全球TOP500超级计算机排行榜于2010年11月15日正式公布
  - 国家超级计算天津中心的天河一号A获得了冠军

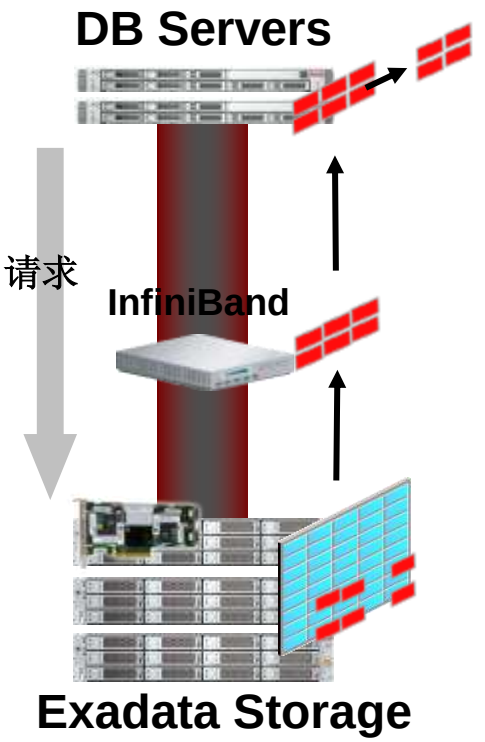


- 20多万颗处理器核心
  - 14336颗Intel Xeon X5670 2.93GHz六核心CPU
  - 7168块NVIDIA Tesla M2050 GPU
  - 2048颗我国自主研发的飞腾FT-1000八核心处理器
- InfiniBand专有互联网络

为什么快？ -- Exadata革命性的设计

火龙果•整理
uml.org.cn

Exadata



Hardware and Software
Engineered to Work Together

高性能理由①

“Smart Scan”： 计算负载部分卸载至并行智能存储层，并只传输经筛选的有用数据

高性能理由②

“综合列压缩(EHCC)”： 技术获得10x－50x 超高压缩比，并提高磁盘I/O效率

高性能理由③

“Smart Flash Cache”： 高达5.3TB 的 Flash Cache，随机I/O性能高达1M IOPS

高性能理由④

“Infiniband ”：提供40Gb/s 端口带宽，聚合带宽高达880Gb/s , 端口延时(0.1um ）小于以太网的 1/1000

Smart Scan、EHCC、Smart Flash Cache等是Oracle Exadata 独有的关键性能突破技术

智能存储层 (Disk Array)
Exadata Smart Scan 处理技术

火龙果•整理
uml.org.cn

Compute Intensive Processing

ORACLE DATABASE

Bandwidth Intensive Searches

ORACLE EXADATA

- Database Server
  - 计算和内存密集型的数据处理在数据库服务器执行
  - 完全并行和聚集
- Exadata Storage Server
  - IO-bandwidth 密集型搜索在存储服务器上执行
    - Exadata Smart Scans and Exadata Storage Indexes 可以直接在存储层过滤数据

数据库服务器和 Exadata 存储互相协调执行SQL

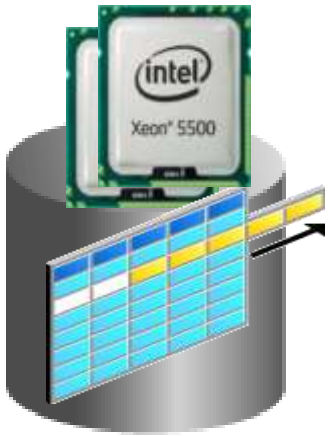
•Exadata cell 是智能存储，不是一个完整的数据库节点

火龙果•整理
uml.org.cn

Exadata的存储端数据库处理

- ExadataStorage Server 在存储端进行数据密集型处理
  - – 使用 “where”谓词筛选行
  - – 列筛选
  - – 联接筛选
  - – 增量备份筛选
  - **--I/O prioritization**
  - – 存储索引
  - – 扫描加密数据
  - – 数据挖掘模型评价

New



Exadata Intelligent Storage Grid

- 通常可将发送到数据库服务器的数据减少 10 倍

- 无需更改应用程序
  - –处理是自动和透明的
  - –即使在查询期间单元或磁盘发生故障

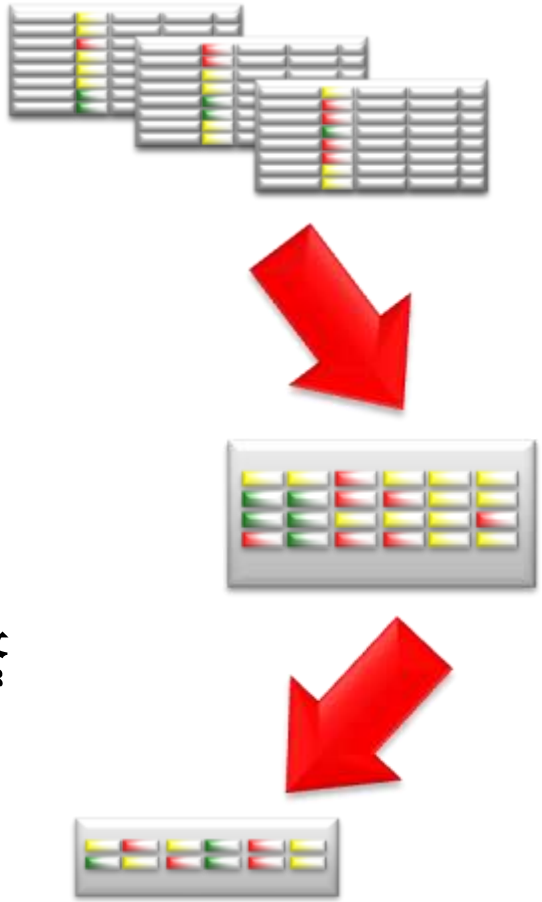
智能存储层(Disk Array)
Exadata Hybrid Columnar 压缩技术

火龙果•整理
uml.org.cn

- 数据以字段形式存储并压缩

- 查询模式 –针对data warehousing
  - 最优的速度
  - 典型的可达到10X的压缩率
  - 按压缩比率改进扫描

- Archival Mode –针对频繁的数据访问
  - 减少空间
  - 典型的可达到15X的压缩率
  - 有些可达到50X



Up To
50X

智能存储层 (Disk Array)

Exadata Flash 解决RANDOM I/O瓶颈

火龙果•整理
uml.org.cn

New

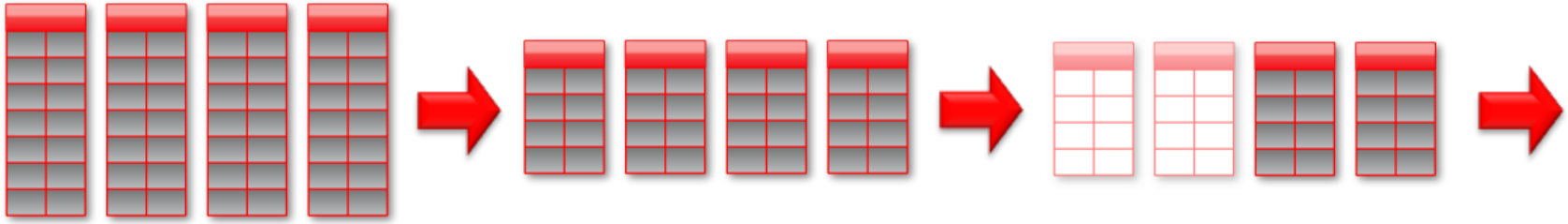
5X More I/Os than 1000 Disk Enterprise Storage Array

- **Exadata有5+ TB 的闪存存储**
- **Exadata Smart Cache 技术缓存热点访问的数据**
  - 可以根据业务表规则进行优化
- 极大地降低成本
  - 同样的I/O能力，超过10x倍的磁盘节省
  - 低能耗

Oracle 是第一个闪存优化的数据库系统

Exadata V2 智能存储带来性能增倍

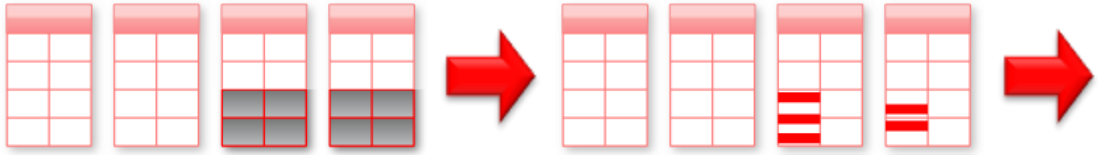
火龙果•整理
uml.org.cn



10 TB 的用户数据
需要 10 TB 的 IO

1 TB
（压缩）

100 GB
（分区修剪）



20 GB
（存储索引）

内存或闪存上的 5 GB
智能扫描

在 Database Machine 上实现
亚秒级

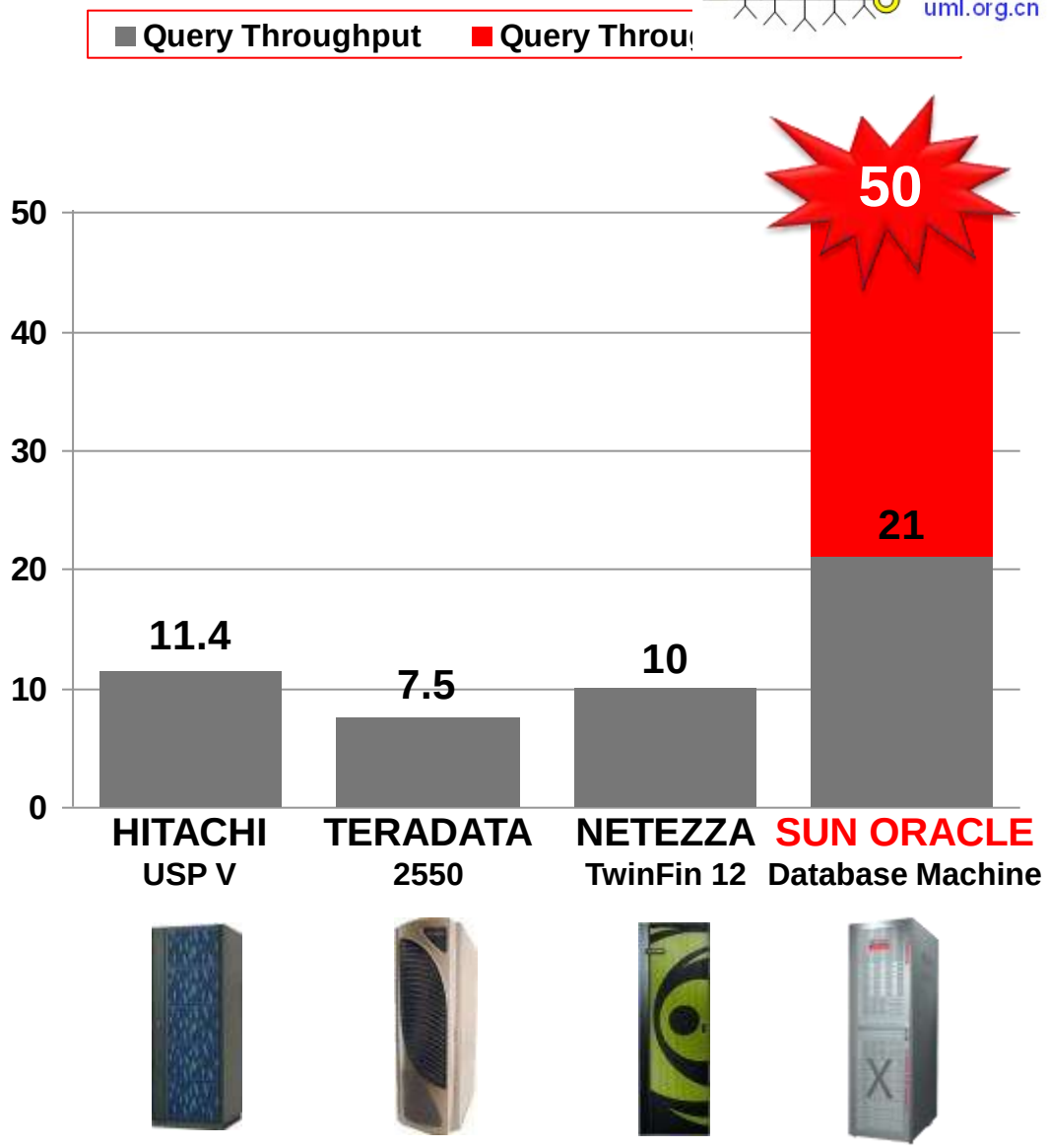
数据减少 10 倍，扫描加快 2000 倍

火龙果•整理
uml.org.cn

Exadata V2 性能

为什么Oracle更快?

- ✓ 数据在存储层处理
- ✓ 更高的压缩率 (10x)
- ✓ Smart Flash Cache
- ✓ Faster Interconnect (40Gb/sec)
- ✓ 更多的磁盘
- ✓ 更快的磁盘 (15K RPM)



Exadata 中国区用户POC测试结果

火龙果•整理
uml.org.cn

序号	行业及项目	配置	测试结果	测试日期
1	Telco: 经营分析	V2, ¼ Rack	168倍	2010.11.26 - 28
2	Telco: 财务系统	V2, ¼ Rack	20倍	7天, 2010
3	Telco:经营分析	V2, ½ Rack	235倍	2010.9.8 - 15
4	保险: 数据仓库	V2, ¼ Rack	25倍	21天, 2010.5
5	MRD: ETC (DW)	V2, ¼ Rack	300倍	14天, 2010.5
6	MRD: ODS/DW	V2, ¼ Rack	62倍	14天, 2010.5
7	DATES & DW	V2, ¼ Rack	10 - 450倍	2010.7.5 -16
8	MRD: 集运运输(OLTP)	V2, ¼ Rack	10倍	2010.8.9 - 18
9	电力: 低压集超(DW)	V1	10 倍	7天, 2010.1

* 测试结果为**Exadata**相当于用户原有环境或第三方对比测试方案的性能提升倍数

Exadata V2 灵活扩展

火龙果•整理
uml.org.cn

- 扩展性
  - 单个机器可以扩展至8 RACK
    - 更多的扩展通过InfiniBand switches实现
  - 可以扩展成千上万个存储服务器
    - Multi-petabyte databases

四分之一配

半配

满配

Exadata容量选择依据，IO和磁盘

火龙果•整理
uml.org.cn

		基本系统	四分之一机架	半个机架	整个机架	2-8 整个机架
Database Servers		**1**	**2**	**4**	**8**	**16-64**
Exadata Storage Servers		**1**	**3**	**7**	**14**	**28-112**
Total Disk Capacity		**7.2 TB**	**21 TB**	**50 TB**	**100 TB**	**200 – 800TB**
User Data (uncompressed)		**2 TB**	**6 TB**	**14 TB**	**28 TB**	**56 – 224 TB**
I/O Throughput (disks)		**1.5 GB/sec**	**4.5 GB/sec**	**10.5 GB/sec**	**21 GB/sec**	**42 - 168 GB/sec**
I/O Throughput (flash)		**3.6 GB/sec**	**11 GB/sec**	**25 GB/sec**	**50 GB/sec**	**100 - 400 GB/sec**
I/O per Second (IOPS)		**75,000**	**225,000**	**500,000**	**1,000,000**	**1M – 8M**
Raw Disk[1]	SAS	7.2 TB	21 TB	50 TB	100 TB	
	SATA	24 TB	72 TB	168 TB	336 TB	
Raw Flash[1]		384 GB	1.1 TB	2.6 TB	5.3 TB	
User Data[2] (assuming no compression)	SAS	2 TB	6 TB	14 TB	28 TB	
	SATA	7 TB	21 TB	50 TB	100 TB	

火龙果•整理
uml.org.cn

Exadata V2 更节省成本

与传统SMP对比

项目	Exadata V2	传统SMP系统
体系结构	云架构(基于廉价PC)	传统SMP结构
初始投资	可以少量节点，投资少	大架子，少内涵，投资高
扩容投资	按需扩容，以PC节点为单位，廉价	按专用的CPU板、IO板扩容，困难，昂贵
管理	PC级的管理，高度一致	众多的SMP型号，不一致
资源共享方式	云化服务器池的方式，资源动态调度	有限程度共享，不同系统难以共存
智能存储	大量的CPU在存储层，提供计算能力 如EHCC，存储索引可提高效率10倍	非智能存储，智能垂直切割共享
机房空间	2架	4+8架
总体成本	<50%	1

火龙果•整理
uml.org.cn

Exadata V2 总结

提供全面解决方案的高性能一体机

- 适用于数据仓库应用的最好的数据库服务器（Best for Data Warehousing）
  - 基于10x压缩表的Smart scan
  - 基于内存数据的并行查询
  - 整体上比11.1版本快5倍
- 适用于OLTP系统的最好的数据库服务器（Best for OLTP）
  - 唯一基于网格技术扩展的数据库
  - Smart flash cache 可达到20x快的IOPS,或者节省20x的磁盘
  - 对于归档数据可达到50x的压缩率
  - 安全, 容错
- 适用于混合负载的最好的数据库服务器（Best for Consolidation）
  - 唯一的支持所有负载类型的database machine
  - 多个数据库,多个应用,多个用户环境都能提供可预测的响应时间

火龙果•整理
uml.org.cn

Exadata V2 成功案例

中国移动通信 CHINA MOBILE 移 动 信 息 专 家
Haier CHINA
MCC 中国中冶
RIM
LOYALTY
Garanti
Mtel
true
allegro
KnowledgeBase Marketing
CTC Challenging Tomorrow's Changes
MTN OFFICIAL SPONSOR FIFA
NAGASE 長瀬産業株式会社
OKsystem
AMTRAK
AA For the road ahead
HISCOM
NS Solutions
SoftBank
NTT DaTa
ヤマザキ
GIANT EAGLE
TUI Nederland

我们的客户说

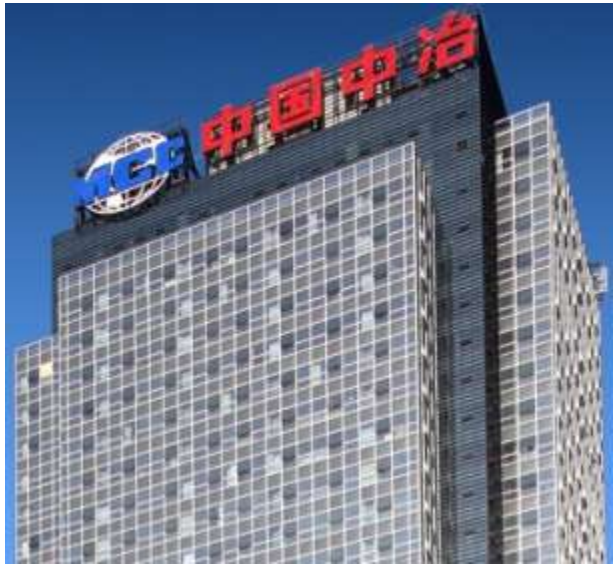
火龙果•整理
uml.org.cn

- AXIATA BANGLADESH (APAC) 孟加拉国亚通，孟加拉国第三大电信营运商，
  - 数据装载快了30倍(从1小时减少到6分钟),查询快了10倍
  - 针对于活跃和归档数据有2 到 10倍的压缩比，同时无任何性能损失；从而有能力扩展了在线数据时间窗口
- SOFTBANK (JAPAN) 日本电信运营商，替代原有36台Teradata，占地面积从36Racks减到3Racks，查询速度提升2~8倍
- TURKCELL (EMEA) 土耳其电信运营商
  - 数据压缩10倍(250TB->25TB)，查询速度提升10倍~400倍(50K报表/月),
  - 数据翻倍每年，1.5Billion CDR per day
  - 减少DC占地面积\电力、制冷11倍(11 racks => 1 rack)
- LINKSHARE (NORTH AMERICA) 北美最大的联属营销平台
  - 提升查询速度8倍，服务器和存储为原来的1/8, 减少运维OPEX**3.5x**
  - 减少DC占地面积\电力4倍
- COMMONWEALTH BANK OF AUSTRALIA (APAC) 澳大利亚联邦银行
  - 超过250+独立运行的数据库整合到Exadata 3个数据库中 部署实施时间提升30倍(4月->4天); 提供-数据库即服务云

中冶集团ERP招标单位

ORACLE E-BUSINESS SUI
火龙果•整理
uml.org.cn

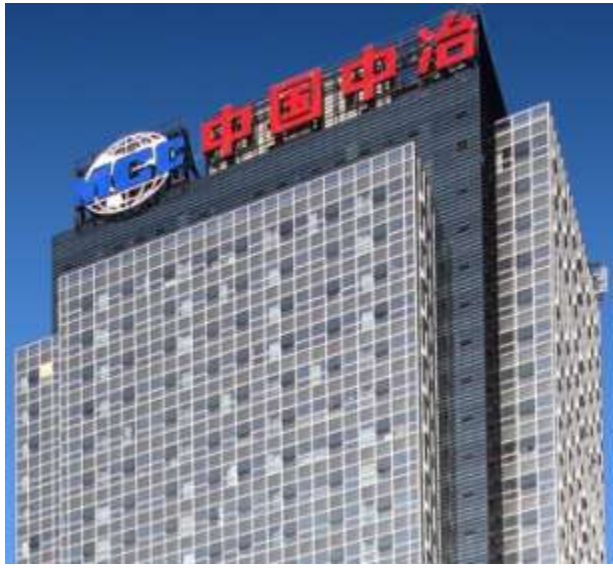
- 中冶集团ERP FY10 Q2签约Oracle EBS12的单位:
  - 中冶集团
  - 冶建总院
  - 中冶南方
  - 中冶焦耐
  - 中冶宝钢技术
  - 二冶
  - 十九冶
  - …



中冶集团招标单位应用系统的建设内

火龙果•整理
uml.org.cn

- ERP系统
  - 项目管理PM／财务管理FIN／采购管理PUR／库存管理INV／设备资产管理EAM／人力资源管理HCM
  - Oracle电子商务套件系统：EBS R12
- BI系统
  - 业务智能： Oracle BIApps系统
- 文档管理系统（中冶南方／焦耐）
  - 文档管理系统
  - AutoVue



Oracle MCC项目中价值点

火龙果•整理
uml.org.cn

- Oracle提供Total Solution，Oracle懂得最佳的解决方案
- Oracle提供Disk-to-Application Service，最好的服务保障
- Exadata配置好的服务器、存储、网络设备，性能最佳，降低了风险，一定程度上节省了成本
- 作为数据中心的平台核心，保障其长期发展

问答

火龙果•整理
uml.org.cn

火龙果•整理
uml.org.cn

ORACLE IS THE INFORMATION COMPANY

火龙果•整理
uml.org.cn

ORACLE®