

# The Vertica 下一代分析数据库解决方案介绍

Chen Hua-ping  
HP CMS China

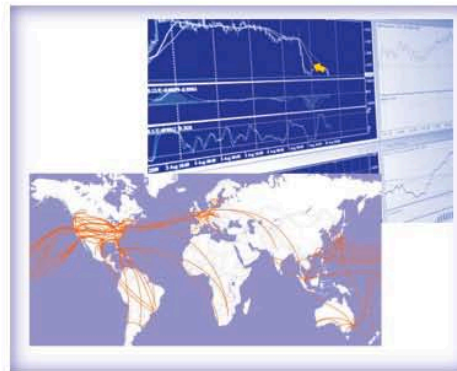
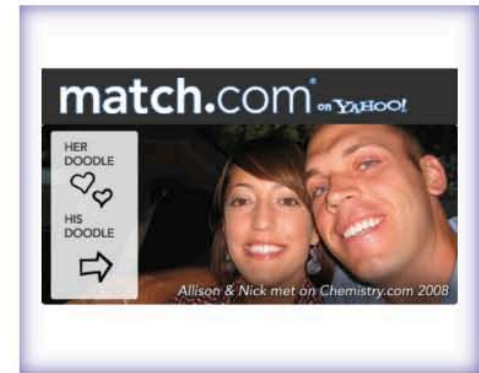


# Agenda

- 数据分析市场
- Vertica解决方案
- Vertica案例分析
- 总结

# We Live in an Analytics World – 我们生活在数据分析世界

- Gartner 近日在一份新闻稿中指出：“数据仓库将在 2011 年迎来它自问世后的一个最重要的转折点。Gartner 坚信数据仓库“理想”正在发生变化，它将朝向一种可以应对更加极端类型信息资产的新型数据仓库发展。



# Analytics Everywhere Brings Challenges 数据分析的增长也带来新的挑战



- More data, and it comes in continuously 越来越多的数据，而且是持续增长
- No more overnight batch loading 不再允许夜间批量处理
- Mixed workloads and user variety accessing 混合的数据负载和用户多变的访问方式
- Must retain long history of data for compliance and analysis 为满足合规性和分析数据保留更长的历史时间

**这些挑战也带来巨大的机会！**

# The New Math: Return on Information (ROI) 一种新的数学公式：信息回报

聪明的公司已经发现其中的价值：

$$\begin{array}{c} \text{数据量} \\ \text{Data Volume} \end{array} \times \begin{array}{c} \text{分析深度} \\ \text{Depth of Analytics} \end{array} \times \begin{array}{c} \text{用户数} \\ \text{\# of Users} \end{array} = \begin{array}{c} \text{数据回报} \\ \text{Data Return} \end{array}$$

延迟  
Latency

延迟

=

ROI

总成本

$$\begin{array}{c} \text{硬件} \\ \text{Hardware} \end{array} + \begin{array}{c} \text{许可} \\ \text{License} \end{array} + \begin{array}{c} \text{运维和支持} \\ \text{Tuning/Support} \end{array} = \begin{array}{c} \text{Total Cost} \end{array}$$

zynga

verizon

GUESS

Syniverse  
Technologies

mozilla

NOVARTIS

Bank of America

Aol.

comcast

# Value of Time – Vertica让我们重新认识时间的价值

TERADATA  
ORACLE

NETEZZA

VERTICA

Days 天

Hours 小时

Seconds 秒

传统的数据库解决方案

今天的数据分析需求

Proprietary 专用产品

Commodity 开放产品

Expensive 昂贵

Cost-Effective 成本效益

Centralized, Monolithic 集中式的

Distributed 分布式

Process-Heavy 重流程的

Self-Service 自服务

Batch 批量

Real-Time 实时

Slow 慢

Agile 敏捷



# People Want to Question Everything..... 然而传统数据分析解决方案却存在系列疑问...

*Even their choice of hardware.....*

*甚至包括硬件 .....*



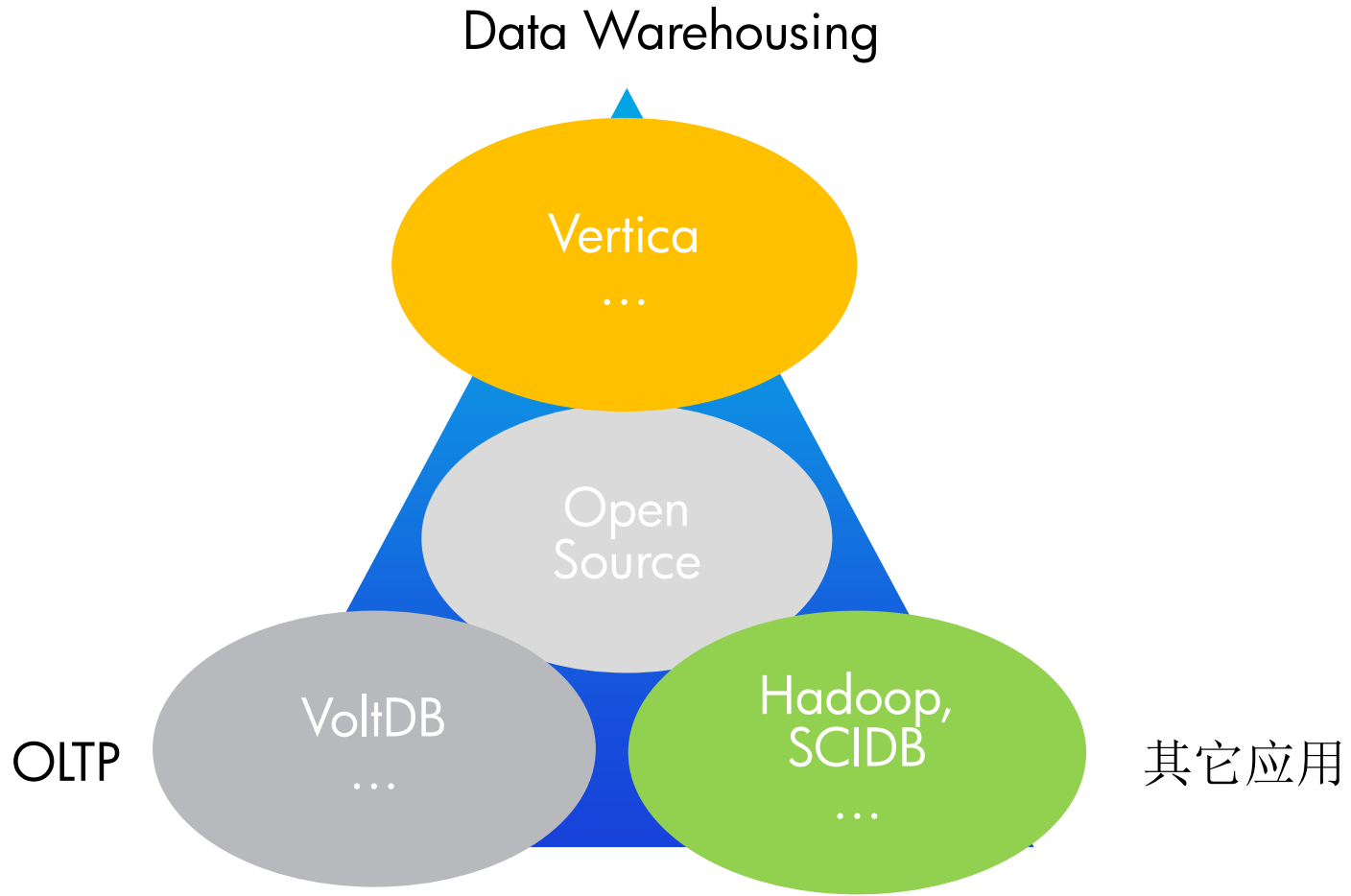
- ▶ **Proprietary hardware** 专用硬件
- ▶ **Row-oriented, not ideal for DW** 面向行结构
- ▶ **No elastic scale-out capabilities** 缺乏弹性
- ▶ **Batch oriented** 批处理模式
- ▶ **Disk I/O bound** DISK I/O 约束
- ▶ **Large footprint/energy draw** 大量日志
- ▶ **Requires teams to install/setup**



- ▶ **Simple to install (minutes)- embeddable** 简单易用
- ▶ **Elastic re-provisioning/sizing** 弹性开通
- ▶ **Column-oriented, purpose built** 列结构
- ▶ **Elastic MPP scale-out, just add blades** 弹性增加
- ▶ **Flexibility- PowerEdge racks or blades and Equallogic/PowerVault** 适合刀片
- ▶ **Real-time load and querying** 实时加载和查询
- ▶ **Superior density with advanced compression – ideal for the Cloud** 高密度存储和高压压缩, 适合云服务模式

# One Size Does Not Fit All

- 我们知道没有通治百病的药，不如对症下药、分而治之



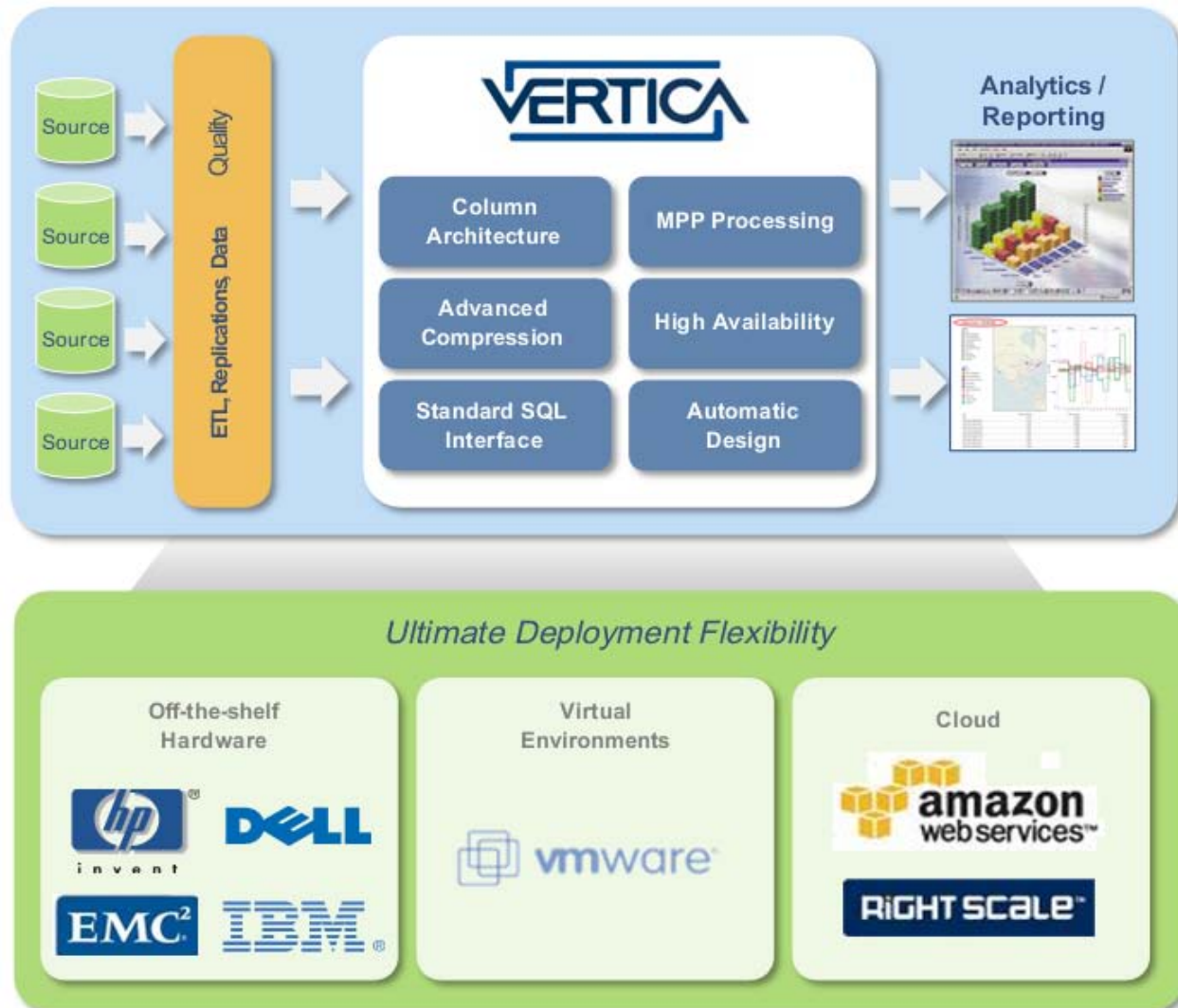


# Vertica 面向数据分析的数据仓库系统解决方案



- 基于列存储和MPP等先进技术
- 平均性能提高50x-1000x倍
- 高可扩展性(TBs ~ PBs)
- 节省高至90%的存储空间
- 快速集成ETL/BI解决方案
- 高性能运行于电信级硬件
- 快速和灵活部署

# Vertica 支持成熟开放的硬件环境



# Agenda

- 数据分析市场
- Vertica 解决方案
- Vertica 案例分析
- 总结

# Vertica 分析数据库 - 关键创新

- 从一个开发者的视角来看，VERTICA符合标准，支持SQL, ACID, JDBC, ODBC，可以和流行的ETL, BI报表集成。
- 面纱之后是一个完全不同的故事：Vertica原生(Native)支持网格（分布式）计算，极大节省 Disk I/O
- Vertica 设计成支持获取和查询T级别的CDR, IPDR, SNMP及电信分析数据的新一代的解决方案
- Vertica其划时代的功能包括：基于列的存储架构，高效的数据压缩，多个“Projections”方式存放，Shared-nothing机制，基于K-Safe的高可用性，自动物理数据库设计，混合(hybrid)存储模型, 灵活的部署方式

# Vertica 分析数据库 – 关键创新



# Vertica 关键创新1 – 基于列的存储架构

- 在 Vertica 每一列数据独立存储在磁盘上的连续块上。
- 查询数据时，Vertica 只需要取得那些需要的列，而不是被选择的行的所有的列数据
- 由于大多数的决策分析系统只是列的子集，Vertica垂直分区的方法极大地节省了Disk I/O,

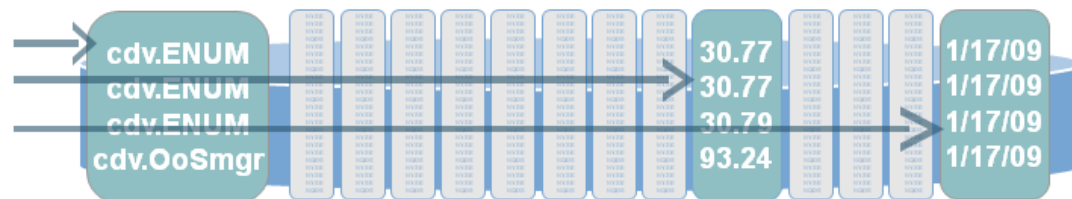
Ideal for read-intensive workloads

Reduces disk IO by orders of magnitude

*SELECT reading  
FROM readings  
WHERE nodetype = 'cdv.ENUM' and timestamp = '1/17/2009'*

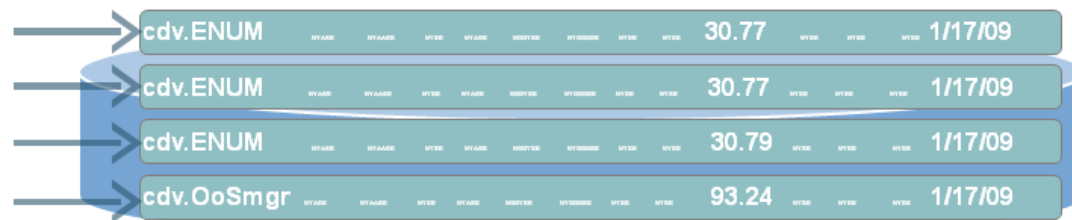
Column Store

Reads 3 columns



Row Store

Reads all columns





# Vertica 关键创新 2- 主动压缩技术

- Vertica 根据数据类型，基数，排序自动进行数据压缩，最小化每列占用的空间。
- Vertica 支持**12+**压缩算法：
  - Run length Encoding
  - Delta value Encoding
  - Integer packing for integer data
  - Block-based dictionary encoding for Character data
  - Lempel-Ziv compression
- Vertica会根据样列数据自动选择最佳压缩算法，基于列的压缩由于同样数据类型和相同取值范围，通常会大幅度提高压缩效果
- Vertica 分析数据库尤其支持直接访问压缩的数据值，这样节省在数据查询期间的CPU开销，也是和其他数据库压缩技术的差异化优势

Node Type	Timestamp	Reading
<i>cdv.ENUM, 16</i>	0000001	700.00 75.05 36.53 145.88 283.35 31.40 23.27 344.44 21.30 23.92 50.22 38.22 27.92 74.26 152.49 89.23
<i>cdv.ENUM</i>	0	23.21
<i>cdv.ENUM</i>	2	344.44
<i>cdv.ENUM</i>	2	21.30
<i>cdv.ENUM</i>	4	23.92
<i>cdv.ENUM</i>	10	50.22
<i>cdv.ENUM</i>	10	38.22
<i>cdv.ENUM</i>	19	21.92
<i>cdv.ENUM</i>	25	74.26
<i>cdv.ENUM</i>	49	152.49
<i>cdv.ENUM</i>	50	89.23
<i>cdv.ENUM</i>	51	
<i>cdv.ENUM</i>	52	
<i>cdv.ENUM</i>	67	
<i>cdv.ENUM</i>	68	
<i>cdv.ENUM</i>	70	

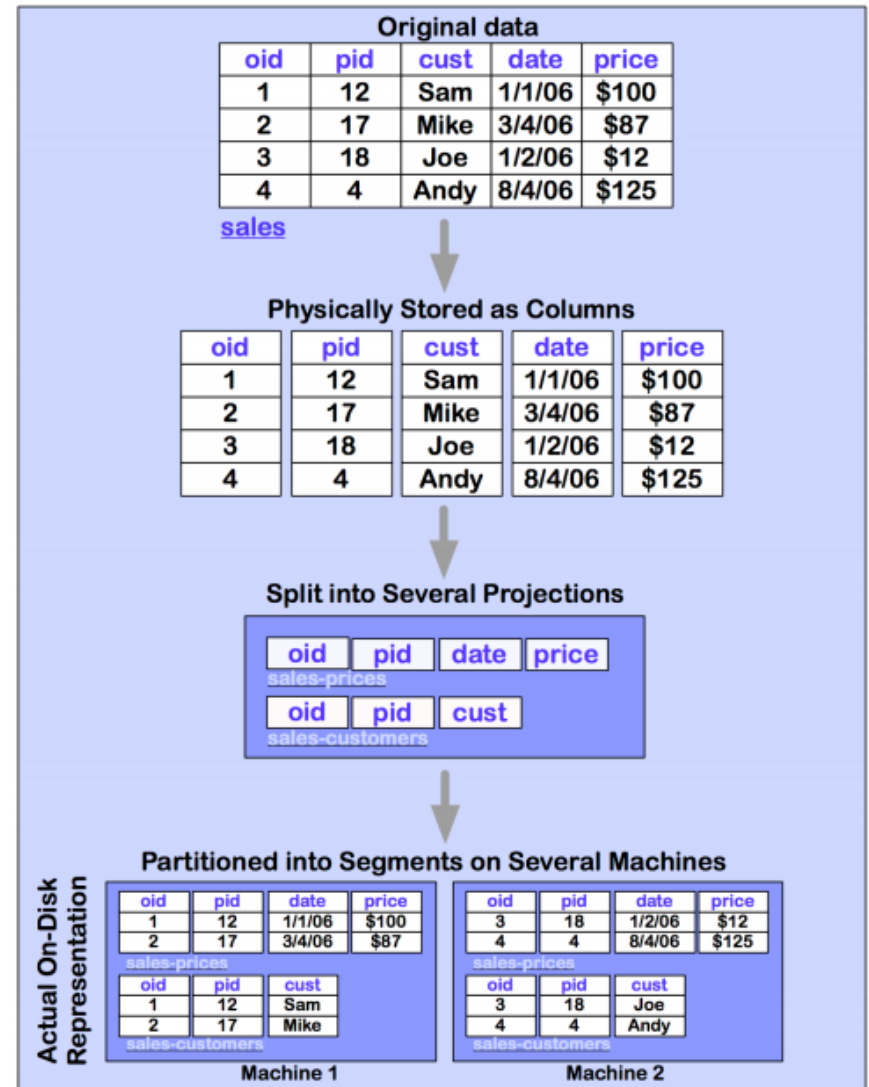
Run-length Encoding (Few values, sorted)	Delta Encoding (Many values, sorted)	Float Compression (Many values, unsorted)
---	---	--

典型数据压缩比率：

Call Detail records **CDR**: 8:1 (节省**87%**)  
Network Logging: 60:1 (节省**98%**)  
Switch-Level SNMP 13:1 (节省**92%**)  
Weblog and click-stream 10:1 (节省**90%**)

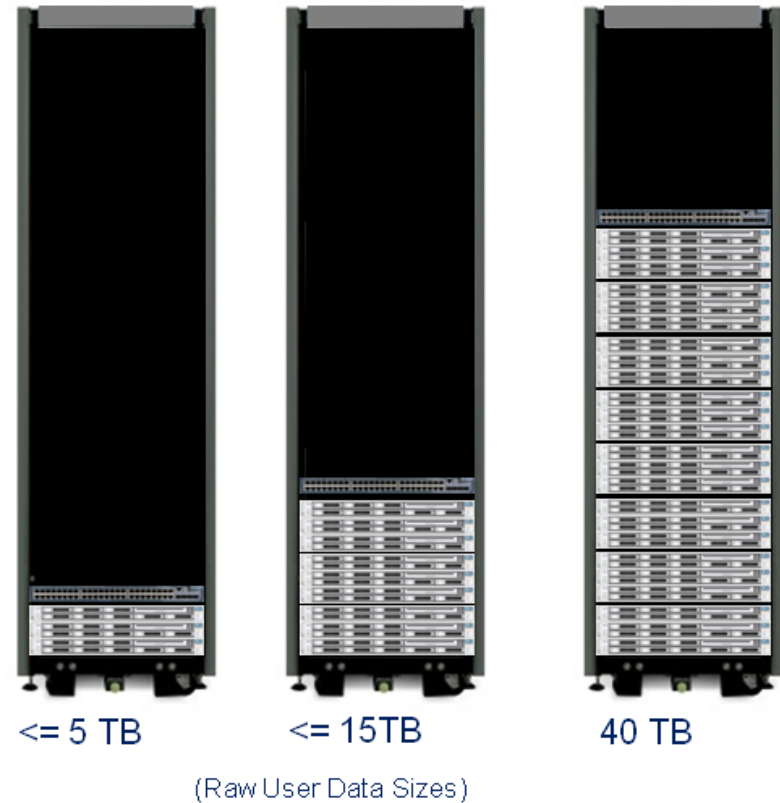
# Vertica 关键创新3 – 多映射“Projections” 存储

- Vertica 物理存储一个表的多个views, 定义为映射“**Projections**”, 每一个 Projections 是一个特定次序排序的表的列的子集。
- Projections也支持来自多个“tables”的 列 (materializing joins)
- 为了支持“ad hoc”查询, 每个数据元素保证至少在一个映射projections中出现
- DB designer 会根据数据样本和查询样本将逻辑schema映射为多个Projections(物理 schema), 优化查询需求。将一个行映射拆分为多个行映射, 同时保证每个查询的 from部分都只来自一个映射, 极大提高查询性能
- Vertica 通常维护多个不同排序的有重叠的映射, 以提高大范围的查询性能 (包括Ad hoc)
- Vertica利用重叠的映射来实现数据冗余和高可用性



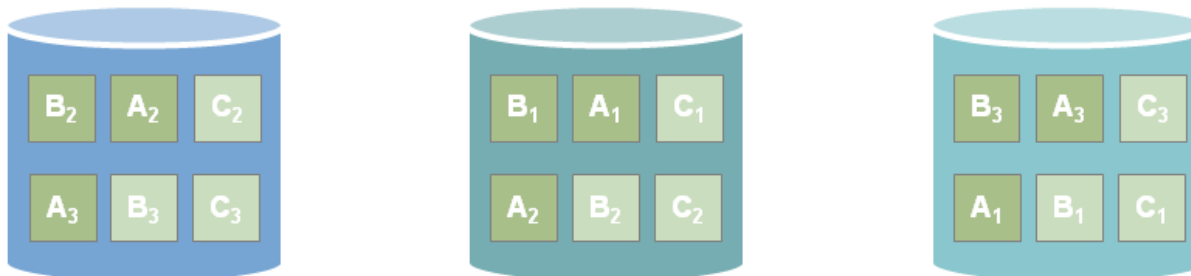
# Vertica 关键创新4 – 非共享大规模并行处理

- Vertica 分析数据库是一个非共享，MPP 大规模并行处理的数据库系统，设计运行在基于网络的Linux server
- Nodes 节点支持多个多核的商用CPU 每个CORE配置2-4GB的内存（不需要专门的或定制化的硬件产品）
- 可以直接使用机内盘直接attached到每个节点，由于其分布式技术，所以不影响可用性（可以大量节省硬件成本），也支持SAN存储
- Vertica 支持对数据在集群范围内进行自动分区
- 例如：fact table将hash分段到集群的多个节点，dimension复制到每个节点，对于大规模的dimension也用和fact标一样的key进行分段。



# Vertica 关键创新5 – 基于K-Safe的高可用性

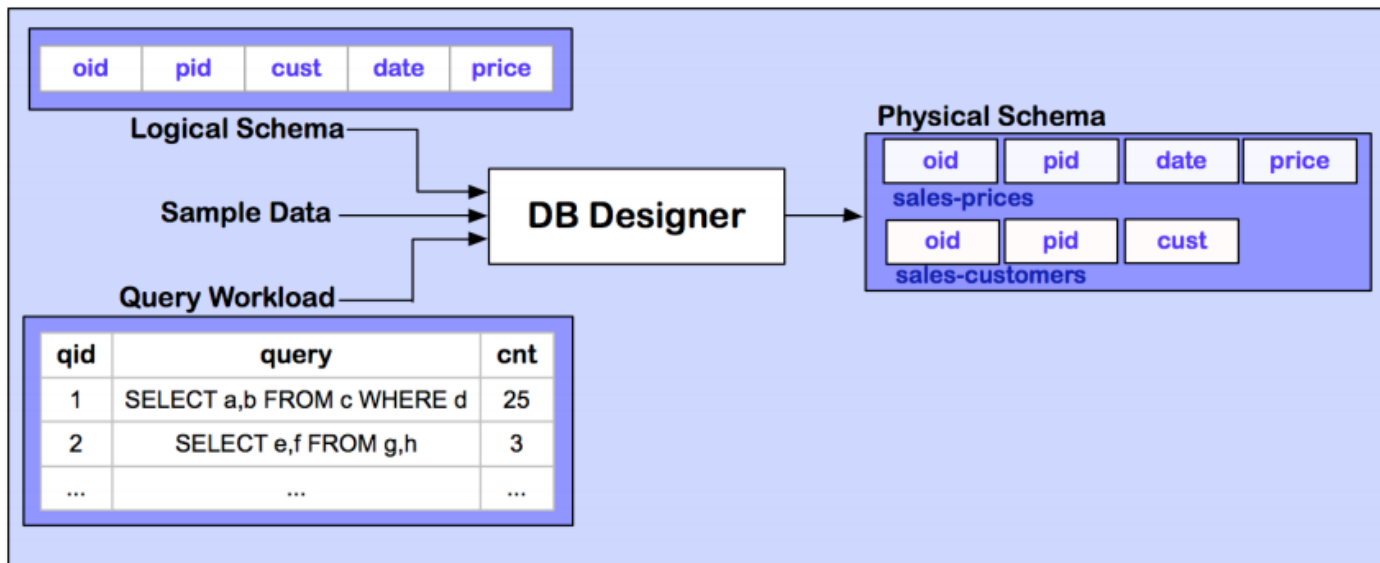
- 传统数据库系统为了实现Recover，通常通过LOG和TWO-PHASE COMMIT包保证交易的完整性及可恢复性，这些手段带来性能的降低和增加磁盘空间
- Vertica用分布式的可恢复性技术来避免这些问题：
  - **k-safe** 技术：保证在Vertica集群内，每一个数据库的每一个表的每一列被存储在至少K+1台机器上，这样当K台机器不可用，仍然有一个完整的数据库备份可用。
  - 意味着如果K台或少于K台机器不可用，则任意一台Crashed的机器可通过从系统中的其他机器上复制Crashed期间最新提交的数据来实现恢复
  - 用这种方法可避免写日志，因为“nodes replicating the data”可用确保一个正在恢复的机器获得另外一份（正确的）数据复制。
  - Vertica 的K-SAFE也意味着高可用性highly available，在集群里面任意K台机器不可用不影响查询处理
  - K值配置数量需要平衡硬件成本和高可用性保障，当然增加K值，不仅增加高可用性，也增加了性能，因为每一个COLUMN复制了K次，但每一份的COPY的排序不一样，大大增加了查询性能
  - 传统的STAND-BY技术，只是增加了硬件成本，但不能提高性能，由于VERTICA是K-SAFE的，所以也支持“hot-swapping”，当移除一个节点，数据库可以继续查询（速度慢一些），同样的，当增加一个节点，数据库将自动增加节点到集群中，这样可以自动提高查询等性能。



# Vertica 关键创新6 – 物理数据库自动设计

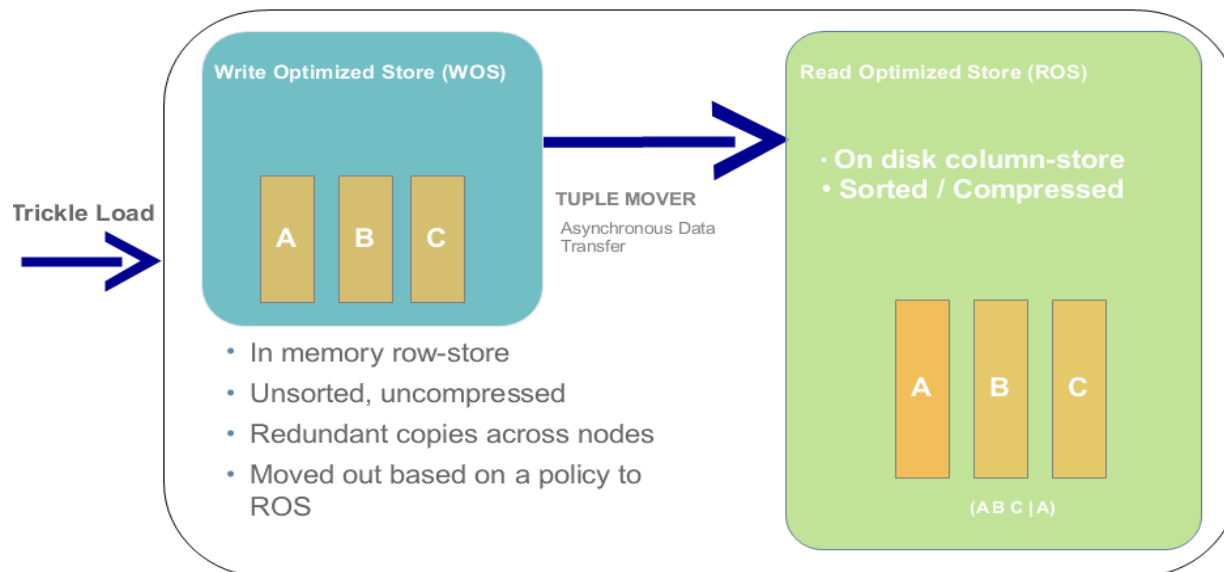
通过简化数据库的设计，让设计和管理人员更多的关注业务需求和逻辑模型设计

- 用Vertica, 用户仅需要设计一个逻辑的数据库Schema, 对于一个逻辑Schema, Vertica的DB Designer工具自动产生一个合适的物理数据库设计, 这个设计基于Schema, 样例数据集和查询及数据库空间估算。
- DB Designer通过确保所有的查询数据至少出现在一个projections, 从而保证任何一个有效的查询能有效命中。
- DB Designer 根据每个列的数据类型和数据值自动选择优化的压缩算法, 并决定为了优化某个查询样例而构建“Projections”, 决定包含哪些columns, 以及如何排序。
- DB Designer 选择每个projections的合适的segmentation key, 保证这个projections满足具体的K-Safety要求



# Vertica 关键创新7 – 混合存储模型 实现高速连续加载并且不影响高速查询

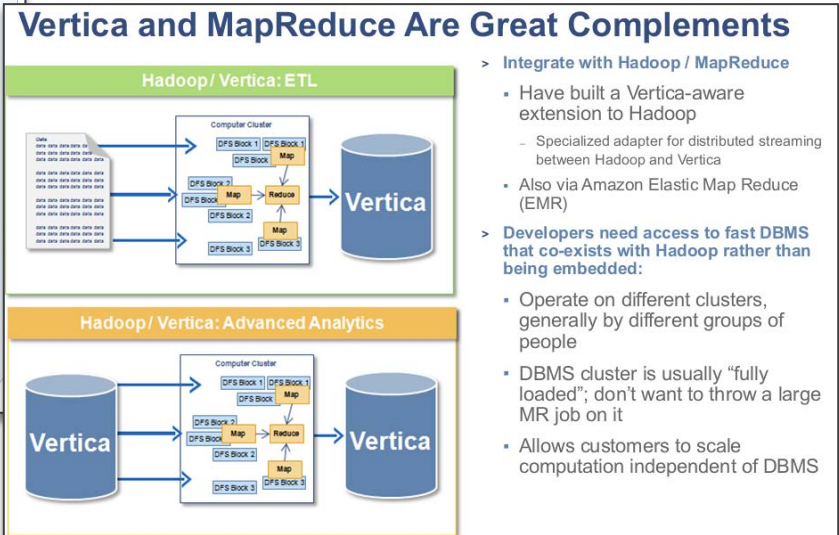
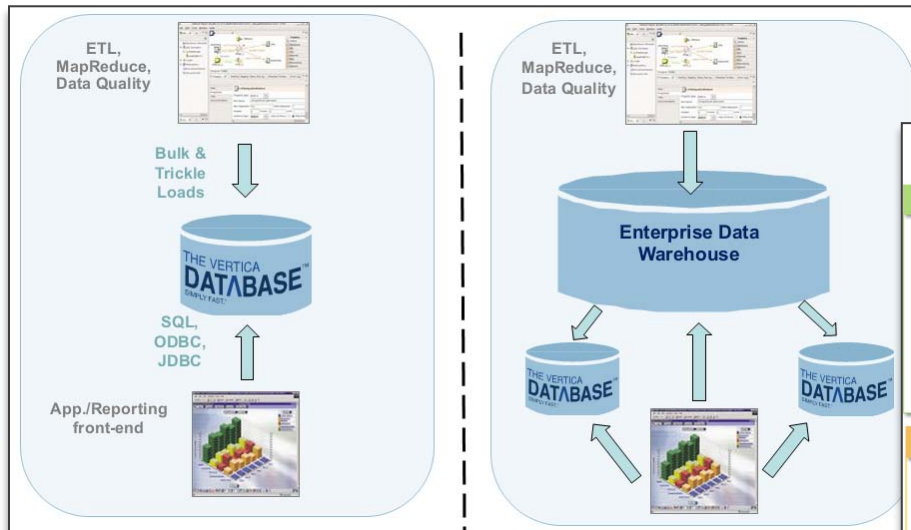
- VERTICA提供一种“continuous load”技术，即不需系统offline,也能提供非常好的“insert”性能，通过下面技术完成
- 1) “snapshot isolation”, VERTICA的“ready only”操作通过快照隔离模式完成，读取最近的并且“consistent”的数据，这些数据是“ready only”,并且不能被并发的“update”和“delete”，这样“read only”查询不需要BLOCK任何并发插入和更新操作，从而查询和更新操作可以完全并行而且相互不影响
- 2) VERTICA通过“WOS”实现非常好的连续加载性能，WOS的数据在内存且不排序和加索引的，通过“TUPLE MOVER”后台进行数据往ROS更新（同时完成排序和索引），由于后台用“BULK”方式更新，性能非常好（类似传统数据库BULK，但不用离线)





# Vertica 关键创新8 – 灵活部署

- 可以运行于Linux server (Red Hat Enterprise Linux 5, SuSE 10 or Fedora Core 6)，并且可以预装在惠普的刀片机
- 也可以运行于云服务平台上，作为数据分析云服务方式提供服务，例如：Amazon的EC2上（按月收取license)
- Vertica 可以根据客户需求，作为独立的数据仓库系统进行部署，也可以部署成数据集市的数据仓库系统
- Vertica的架构非常适合云的特征，包括虚拟化，分布式多节点运行等，并且可以和Hadoop/MapReduce进行集成



# Vertica 关键创新 — 总结

- Vertica 支持混合型的数据存储，最新的插入数据被添加到一个写优化的区域，从而支持持续的、高性能的插入操作的同时也支持并发的查询，以支持实时的数据分析
- 大多数数据库厂商提供的是基于行的存储系统，对于大多数是“读”的负载工作，在同样的磁盘空间情况下，这些系统一般只能提供Vertica 的1/100的性能，或者在10倍多的磁盘空间情况下提供1/10的VERTICA的性能（例如提供大量物化视图）
- 也存在一些Column-oriented系统，但是缺乏主动压缩（Aggressive compression），混合存储优化设计(Hybrid WOS/ROS)，缺乏VERTICA所特有的复制和恢复策略。
- 一般来说，相对Column-oriented系统，VERTICA只需3/4的磁盘空间提供10倍性能

# Agenda

- 数据分析市场
- Vertica解决方案
- Vertica案例分析
- 总结

# Vertica – 创造数据加载世界记录

- 2008年12月，Syncsort, HP 和 Vertica 一起工作打破了一下数据仓库数据加载的世界记录，
  - 在**57分钟21秒**的情况下加载了**5.4TB**的数据进入**Vertica**数据库，系统运行在**HP的C7000**刀片机架上。
  - 早先微软的记录是1小时2.36TB



# 案例分析 1: Vertica 电信客户实际数据分析

## – 分析数据:

- 1.2TB的话单, Vertica使用3个Dell服务器(2个双核,16GB内存, 1TB的本地磁盘), 某传统的数据库使用24CPU的小型机。
- VERTICA的加载时间为5.5小时, 而某数据的加载时间为82小时,
- VERTICA的数据库大小仅为220GB, 而某数据库的大小为4TB
- VERTICA的查询时间为5.4分钟, 而某数据库的时间为2小时

- 相对传统的数据库, Vertica硬件节省超过300万人民币, 查询速度快22倍, 加载速度快15倍, 数据库空间节省1/18

CDR Data Warehouse	Vertica® Analytic Database	Row-Store DBMS	Vertica Advantage
Benchmark Data Set	1.2 TB Call Detail Records		
Production Data Set	50TB Call Detail Records		
Benchmark Hardware	3 Dell servers – 2 x dual core Opteron 2220, 16GB RAM, 1TB local disk	24-CPU SMP server, large SAN	Over \$500K hardware savings
Mean Query Time (4 queries)	5.4 minutes	2 hours	22x faster
Load Time	5.5 hours	82 hours	15x faster
Database Size (1.2TB Raw Data)	220GB	4TB	18x less storage (82% compression)

# 案例分析2: Vertica 电信客户实际数据分析

- 相对于PostgreSQL而言，相同的硬件条件，Vertica 管理的数据为PostgreSQL的25倍，但查询速度为41倍。（网络监控应用数据SNMP）

Network Performance Reporting Appliance	Vertica® Analytic Database	PostgreSQL	Vertica Advantage
Benchmark Data	Test 1: 6 hours of network data (425 million SNMP records) Test 2: 6 days of network data (10 billion SNMP records)		
Production Data	1 month of network data (over 30 billion SNMP records)		
Benchmark Hardware	(1) Dell 2850 Dual Xeon, 4GB RAM, 1TB local disk		Manages 25x more data on the same hardware
Test 1–6 hours of data (mean query time)	1.08 seconds	303.77 seconds	<b>304x faster</b>
Test2–6 days of data (mean query time)	7.47 seconds	Too lengthy to run	Queries 6 <u>days</u> of data 41x faster than PostgreSQL queries 6 hours' worth

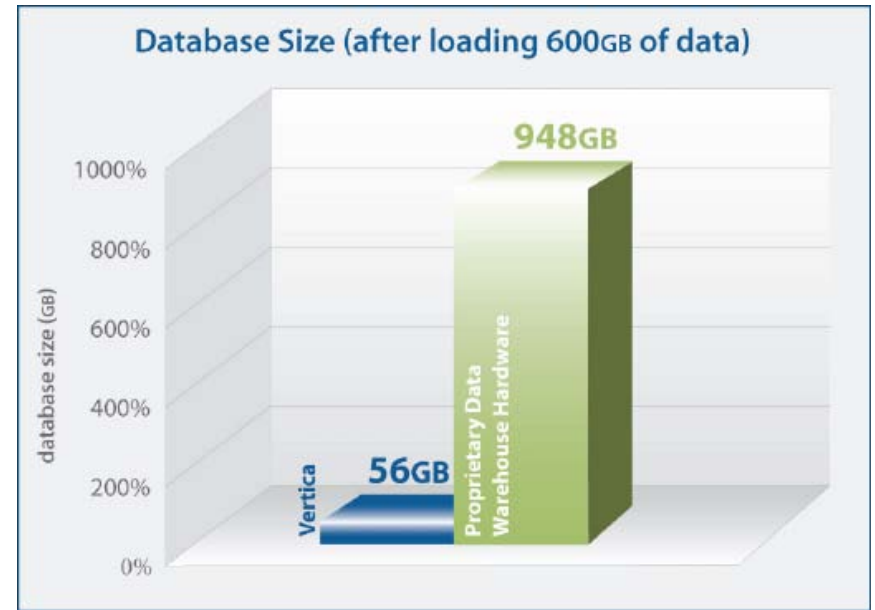
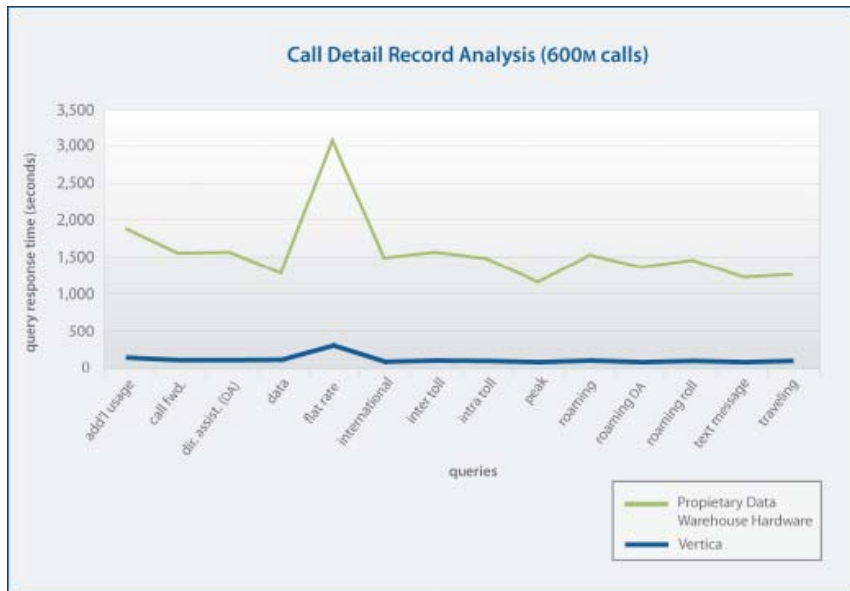
- Vertica 管理某WEB网站流量数据，性能比其它的传统数据仓库快33倍，成本节省约400万人民币。

Web Marketing Analytics SaaS	Vertica® Analytic Database	Proprietary Data Warehouse Hardware	Vertica Advantage
Data Set	25M Rows of Click stream data		
Benchmark Hardware	(4) servers – 2 x dual core x86 CPU, 8GB RAM, 7x146 HDD	Specialized Appliance Hardware - 112 CPU-disk pairs	>\$900K hardware savings; no proprietary hardware
Mean Query Time (7 queries)	7 seconds	233 seconds	Vertica is 33x faster



# VERTICA 电信相关性能数据 -相对于传统的数 据仓库产品

- Vertica queried 600 million CDRs 214x faster (查询6一条话单记录, 快214倍)
- Vertica loaded the 600GB of data 3x faster (加载600GB数据, 快3倍)
- Vertica can store 1.5 years of CDRs in the same space the proprietary data warehouse hardware used to store 90 days' worth (存储节省10倍, 可以存储1.5年或以上的CDR)
- Vertica ran on four, \$5,000 off-the-shelf servers costing 95% less than the \$50,000 proprietary data warehouse hardware. (相对专用硬件, 节省95%成本)



# Agenda

- 数据分析市场
- Vertica解决方案
- Vertica案例分析
- 总结

# 300+ 客户，并且不断增长



金融行业

通信行业

消费市场

在线网站和游戏

医疗

零售

# Vertica 的用户观点:



美国的开心农场, FaceBook最好的游戏

**Ken Rudin, VP of Analytics, Zynga**

超过4千万玩家, 每天新增3TB数据,  
230个节点分布在2个集群, Zynga使用  
基于VERTICAL的数据仓库, 不再让  
数据分析能力像 "windup toy" (发条  
玩具)。

评估几个可选方案后, 我们选择了  
Vertica, 因为它在极短时间内快速  
交付客户需求, 并且相对其他数据库  
节省大量成本。



**GUESS**

**Mike Relich, CIO, Guess**

# 专业评估机构观点：Gartner, Forrester

用户选择使用Vertica，因为它可以让用户有用更多更快的方式分析更多数据的自由，所以在竞争中能够脱颖而出，基于T级别的数据查询，Vertica速度快50到200倍)

Donald Feinberg and Mark A. Beyer, Gartner, Inc.,

**Gartner**



*Vertica...continues to separate itself from the pack. With more use case data, Vertica has begun to shift from the very technical sell of its first couple of years to a more "what problems we solve" approach.*

James Kobiulus, Forrester

**FORRESTER®**

# Vertica 分析数据库系统需求:

## – Vertica Analytic Database System Requirements:

- A cluster of 1 or more (at least 3 for production use) shared-nothing computers, each featuring:
- Dual-core 64-bit CPU (at least 1.6GHz)
- At least 2GB of RAM per CPU core
- 1TB of disk (e.g., 4x250GB drive, 10K RPM, SATA or SAS)
- Red Hat Enterprise Linux 5, SuSE 10 or Fedora Core 6