

列式数据库

加速企业分析应用，提升企业价值

SAP

SYBASE是谁？

Sybase, SAP的子公司, 是业界领先的企业级和移动软件公司, 致力于信息的管理、分析和移动。

SYBASE® | An  Company

数据管理

我们的数据管理解决方案安全地管理任务关键型数据并使之流动。

管理

分析

移动

数据管理

企业级

专为任务关键型与高性能的事务处理环境而优化的数据管理平台

ASE

远程与嵌入式

专为任务关键型的前端应用与远程/嵌入式计算而设计的企业级数据管理

SQL Anywhere

数据流动

双向无线数据同步

实时异构数据复制

MobiLink
Replication
Server

工具

专为业务处理与应用而设计的建模、设计和数据管理工具

PowerDesigner
PowerBuilder
WorkSpace

- 支撑60%以上的华尔街交易
- 为全球90%的证券公司提供动力
- 30,000+ 家使用数据管理解决方案的客户

3 - Sybase Confidential - Gartner 位于 Gartner 数据建模魔力象限图的领导者象限

SYBASE®



数据分析

Sybase 是引领颠覆性能的列式分析服务器的先锋。

管理

分析

移动

分析

列式、专用的分析服务器颠覆了可用性、可扩展性和性能的释义。

Sybase IQ

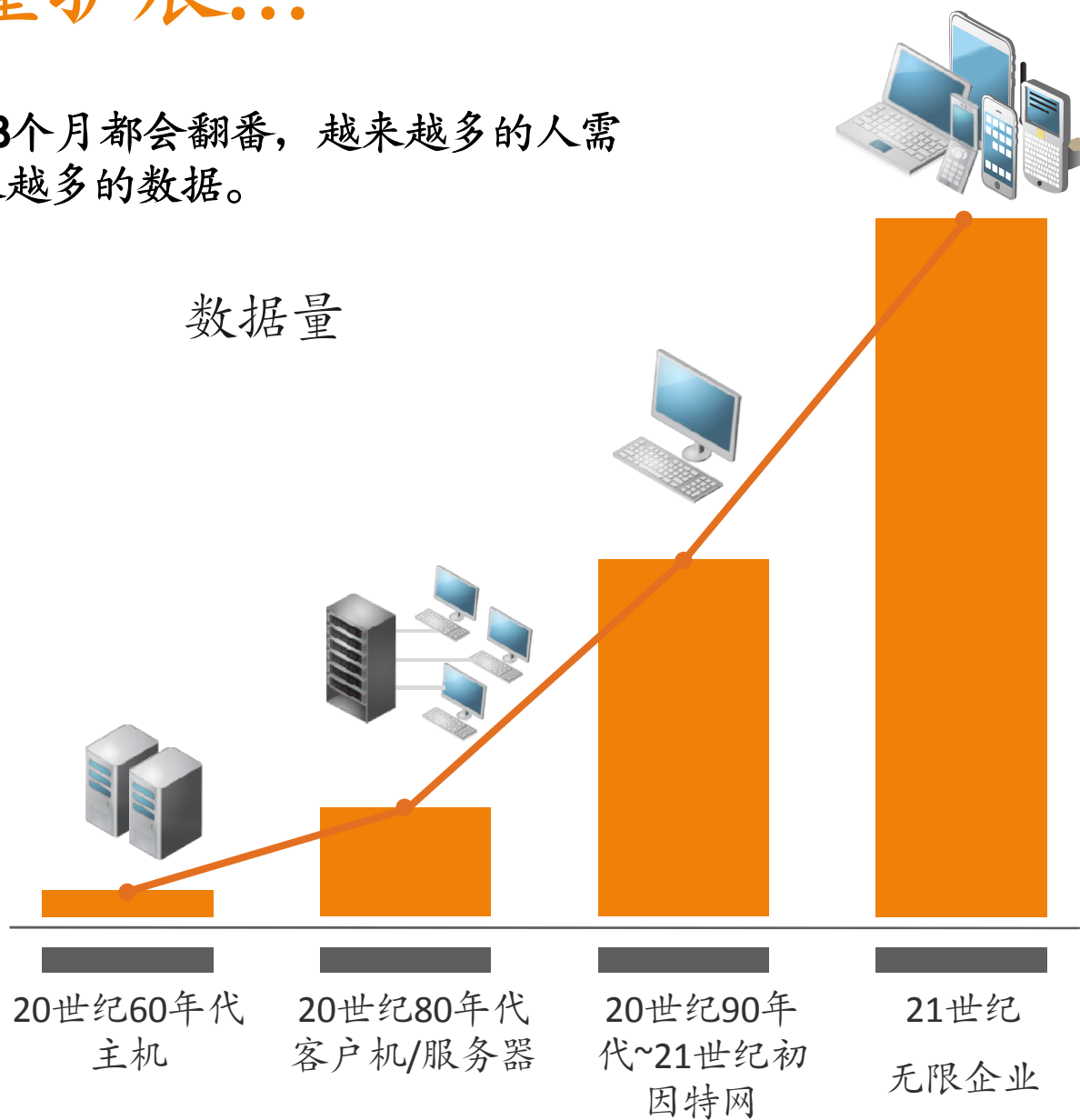
针对资本市场的统的市场分析平台可在完整的交易周期内更好地支持交易和投资决定。

CEP – 实时数据流分析
RAP – 交易版

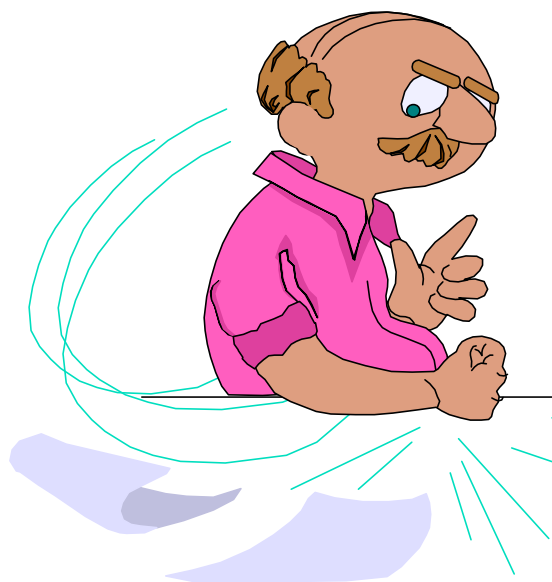
- #1 列式分析服务器
- 全球3,000名客户安装在1,400多个站点上
- 位列2010年Gartner数据仓库DBMS魔力象限图“领导者”象限

数据量扩展...

数据量每**18**个月都会翻番，越来越多的人需要访问越来越多的数据。



企业在开拓业务上的挑战



“查询时间太长了”

“我们又错过了报表窗口”

“除非大量重新调优，否则我们不能用新方法分析数据”

“数据膨胀使我们的数据仓库变成一个噩梦”

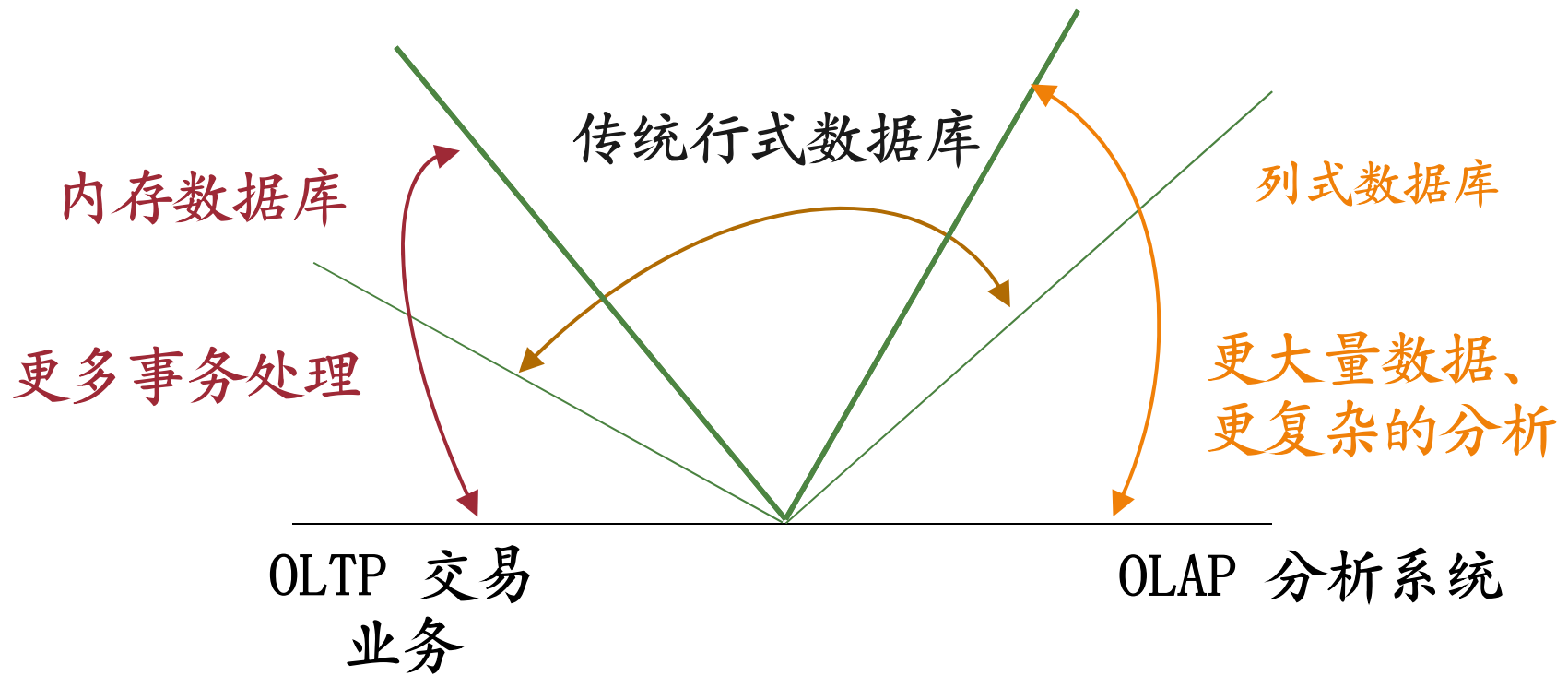
“我们需要一个自服务的模式，服务于上千的用户”

“我们需要实时决策”

“数据爆炸 — 不仅削弱系统性能，而且增加大量存储成本”

“当我们增加用户，我们的性能变成无法接受”

数据库市场的细分格局



列式数据库是革命性的

传统行式数据库

	C 1	C 2	C 3	C 4	C 5	C 6	C 7	C 8	C 9	...
r1										
r2										
r3										
r4										
r5										

- 数据是按行存储的
- 没有索引的查询使用大量I/O
- 建立索引和物化视图需要花费大量时间和资源
- 面对查询的需求，数据库必须被大量膨胀才能满足性能要求

列式数据库

	C 1	C 2	C 3	C 4	C 5	C 6	C 7	C 8	C 9	...
r1										
r2										
r3										
r4										
r5										

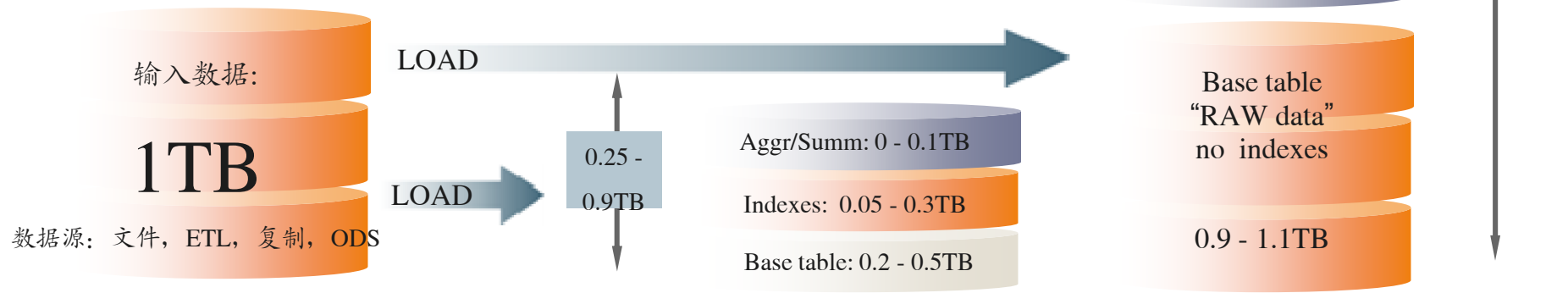
- 数据按列存储 - 每一列单独存放
- 数据即是索引
- 只访问查询涉及的列 - 大量降低系统IO
- 每一列由一个线索来处理 - 查询的并发处理
- 数据类型一致，数据特征相似 - 方便压缩

列式数据库

数据压缩, 降低存储空间和维护成本

数据压缩 – 显著降低存储空间和维护成本

加载相同数据: “行式数据库”
是列式数据库的3—6倍



列式数据库 vs. 行式数据库

	行式数据库	列式数据库	比率 (倍)
5000万条纪录带索引加载	7111.33"	265"	26.8
1亿条纪录带索引加载	14463.8"	1161.44"	12.5
数据存储总占用空间	183.51(G)	27.5(G)	6.7
测试大表count、sum聚合操作,排序	30'57.24"	7'34"	4.1
测试利用索引过滤纪录后对大基数字 段group by ,sum操作	3'32.54"	1'33.3"	2.3
测试大表count、sum聚合操作	10'50.37"	4'23.87"	2.5
测试利用索引过滤纪录后对小基数字 段group by ,sum操作	13.65"	1.27"	13.8
子表操作	30'57.24"	7'34"	4.1

给企业带来的好处



企业领导可以:

- 动态获取全貌
- 历史数据的实时汇总, 随机分析
- 历史明细数据
- 趋势分析
- 热点提示
-



列式数据库的用户

- 尼尔森媒体研究：5020亿条数据(2005年全球第一名) -- 15年的收视数据
- 美国税务局(IRS)：全美国所有报税人7年报税记录(及原始文档) (15亿条记录)
- 花旗银行：Sybase IQ保存6年所有交易记录，HR数据及文档
- 联邦快递(FedEx)：全球所有送递品的海关报关单
- 中国农业银行：“银行卡统计分析系统” 获得世界IT精英组织 (COMPUTERWORLD HONORS PROGRAM —— 计算机世界荣誉奖励计划，简称CHP) 授予“2008年Computerworld荣誉桂冠”
- 铁道部：客票系统分析
- 中国移动：短信/彩信统计分析，报表系统

成功案例 中国农业银行



业务背景

- 随着中国农业银行业务系统（ABIS）数据的上收，数据不断积累，以及市场竞争的加剧，管理部门和人员对经营管理信息化的要求越来越高，数据的价值被充分认识；
- 全行每天交易增量数据 300G, 年增量约1000T。

业务挑战

- 经营分析需要从各业务系统中统一、方便地获取基础数据
- 基础数据的分散收集和逐级上报方式越来越不能适应当今业务数据集中处理方式；基础数据缺乏统一的标准和规范，数据不能共享，影响了管理分析的科学性和有效性
- 为提高数据的可用性和共享性，农行提出了建立分析类信息系统基础数据平台系统的需求

成功案例 中国农业银行

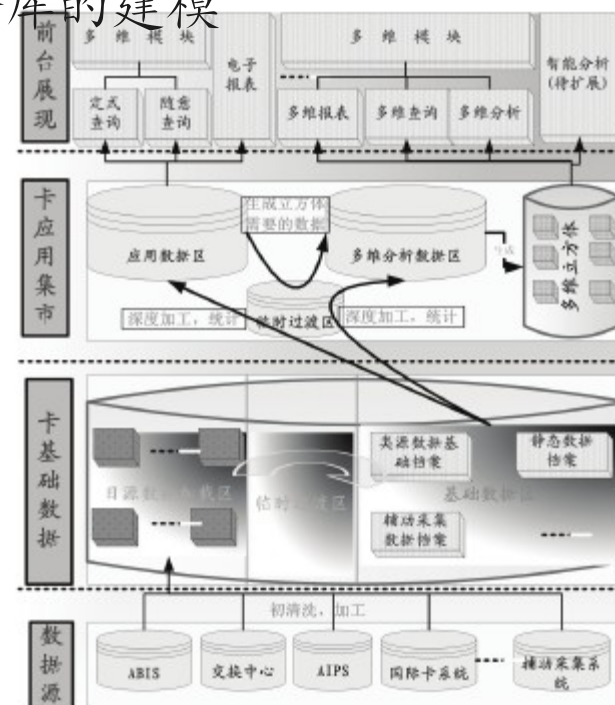


解决方案

- 对分析类系统基础平台分为数据获取层、数据存储层和数据访问层
- 使用Sybase IQ作为数据存储层的分析类数据存储数据库
- PowerDesigner用于数据访问层和数据仓库的建模

应用效果

- 有些分析类系统，如：审计系统已采用Sybase IQ，并取得性能上的大幅度提高
- 大大降低了运营和存储成本
- 大幅度提高查询、统计的效率



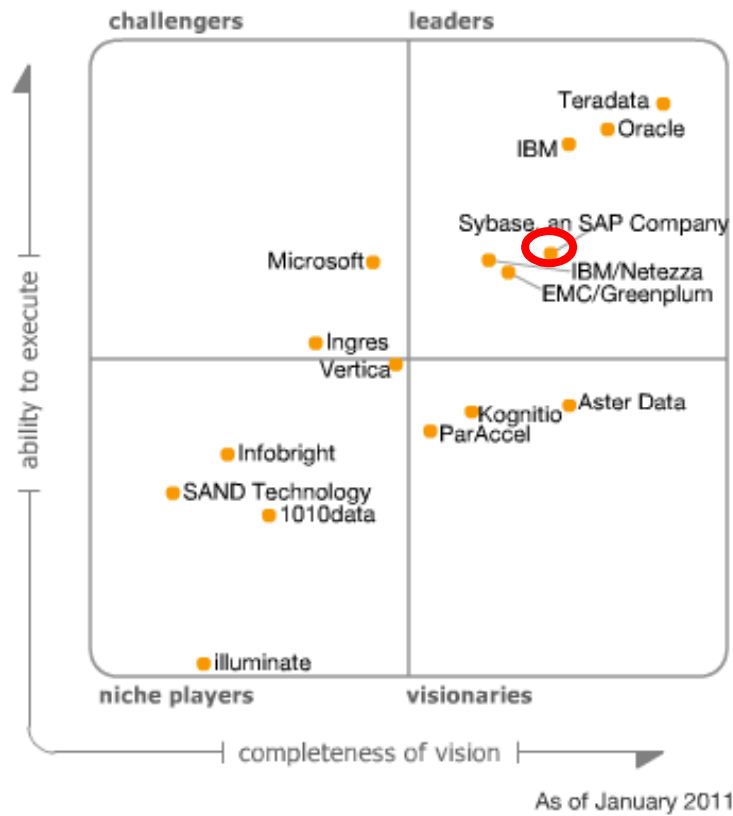
成功案例 中国农业银行

Sybase IQ核心价值

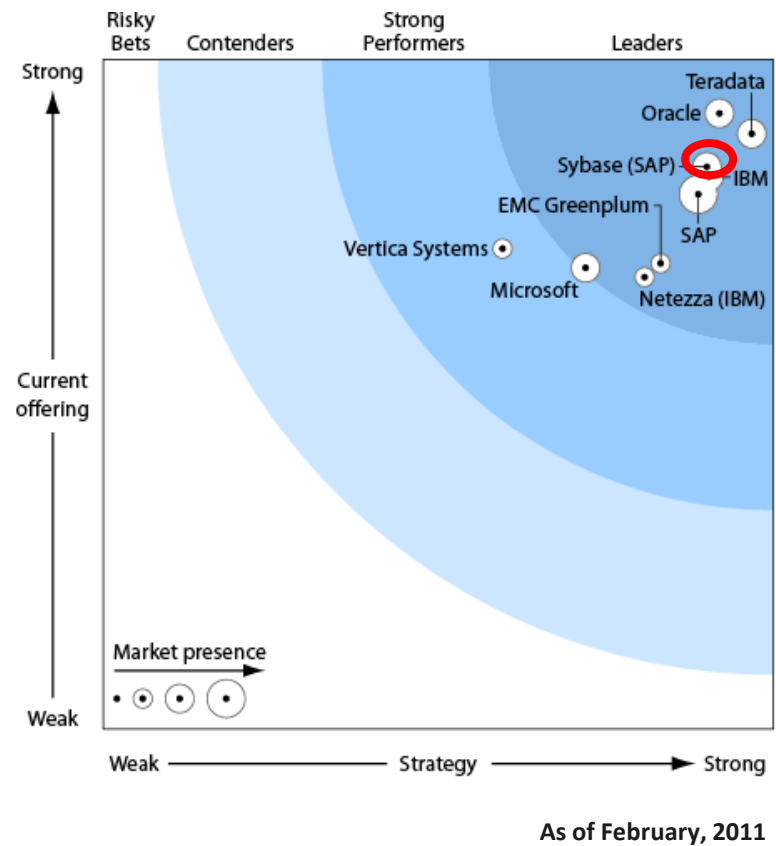
- 性能：在处理卡量余额数据项时，源表为三张五千万条记录级，进行关联操作，由原来的**25小时**缩短到**5小时**，而CPU平均使用率由**93%降到50%**，内存平均使用率由**50%降到30%**。
- 存储：以广东分行为例，每天35G的增量数据文件，在行式数据库上40G存储，而在IQ上只需18G，**50%的存储节省**。
- Sybase IQ在农业银行历史库中，单表记录数为300亿时，数据文件约为**7.8T**，IQ中实际数据存储仅为**4.1T**。

SYBASE IQ 15

业界认可的领导者



Gartner
数据库魔术象限



Forrester Wave
企业级数据仓库

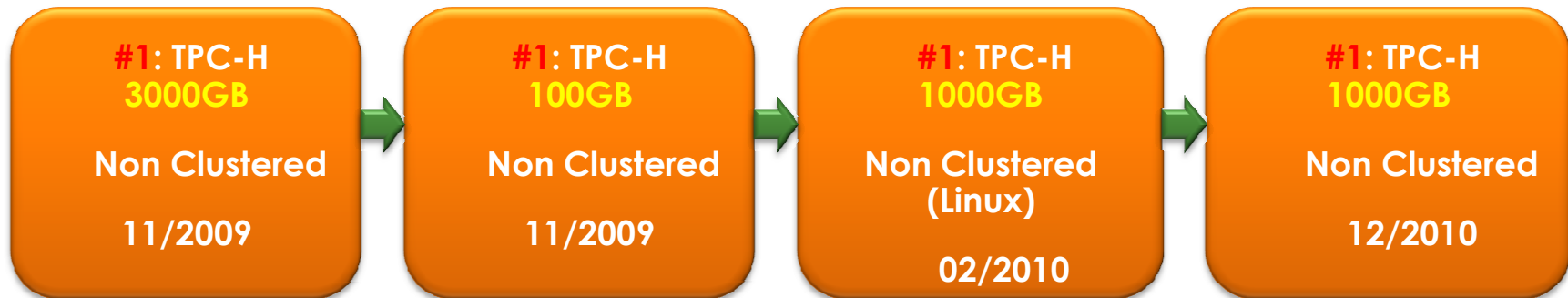
SYBASE®

SYBASE IQ 15

业界最好的BI数据库平台

TPC Transaction Processing
Performance Council

TOP TPC-H PERFORMANCE RESULTS



Sybase IQ :数据仓库市场的领导者

第一个 column-based 分析服务器

- 为分析型应用提供服务近20年
- 10项专利的领先技术
- 三个量级:
 - 存储省一个量级
 - 速度快一个量级
 - 硬件成本低一个量级
- 超过 4,000 独立安装点和 2,000 多客户 并且还在增长





其他网上资源

1. Sybase中国网站: <http://www.sybase.com.cn>
2. Sybase英文网站: <http://www.sybase.com/iq>
3. Sybase中国官方微博:
<http://weibo.com/sybasechina>
4. 新浪微群: “列式数据库”, “CEP”