



IBM Software Group

IBM 面向SOA的信息管理及基础平台DB2 9

邓宏

IBM 软件部信息管理市场经理



ON DEMAND BUSINESS™

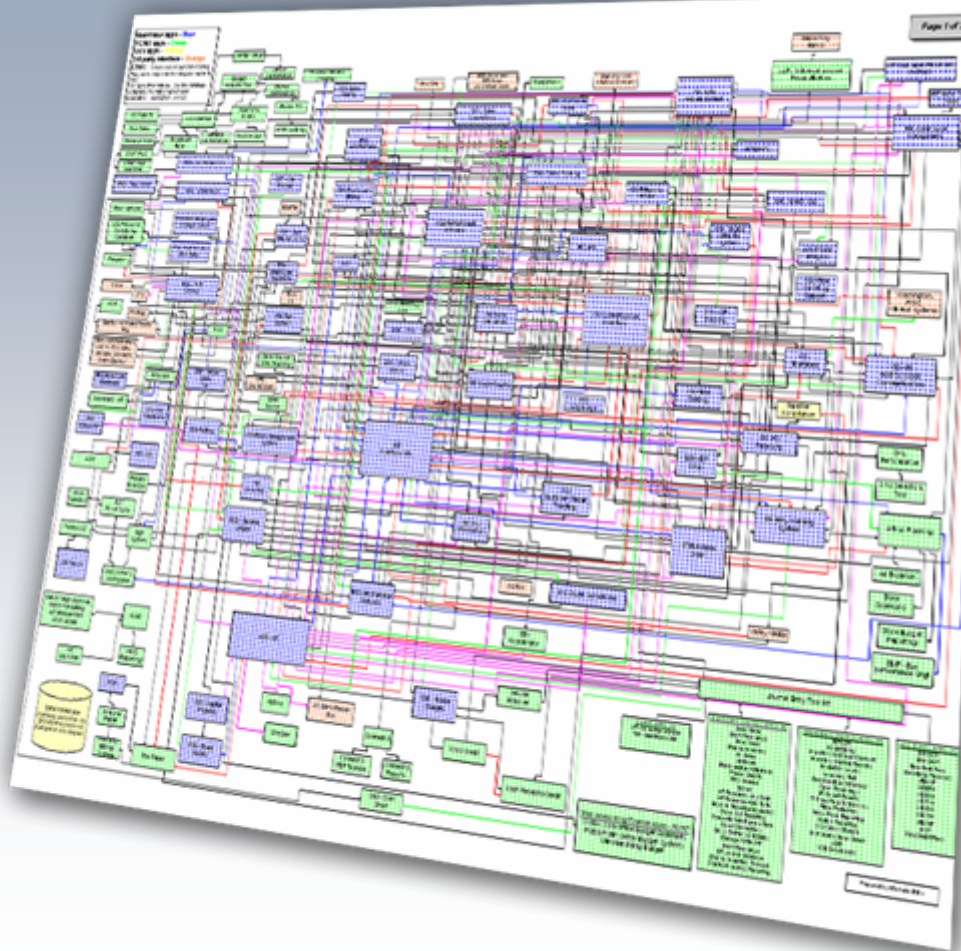
议程

- ✍ IBM 信息管理面向SOA的发展战略
- ✍ IBM 数据服务平台
- ✍ DB2 9 的新特性

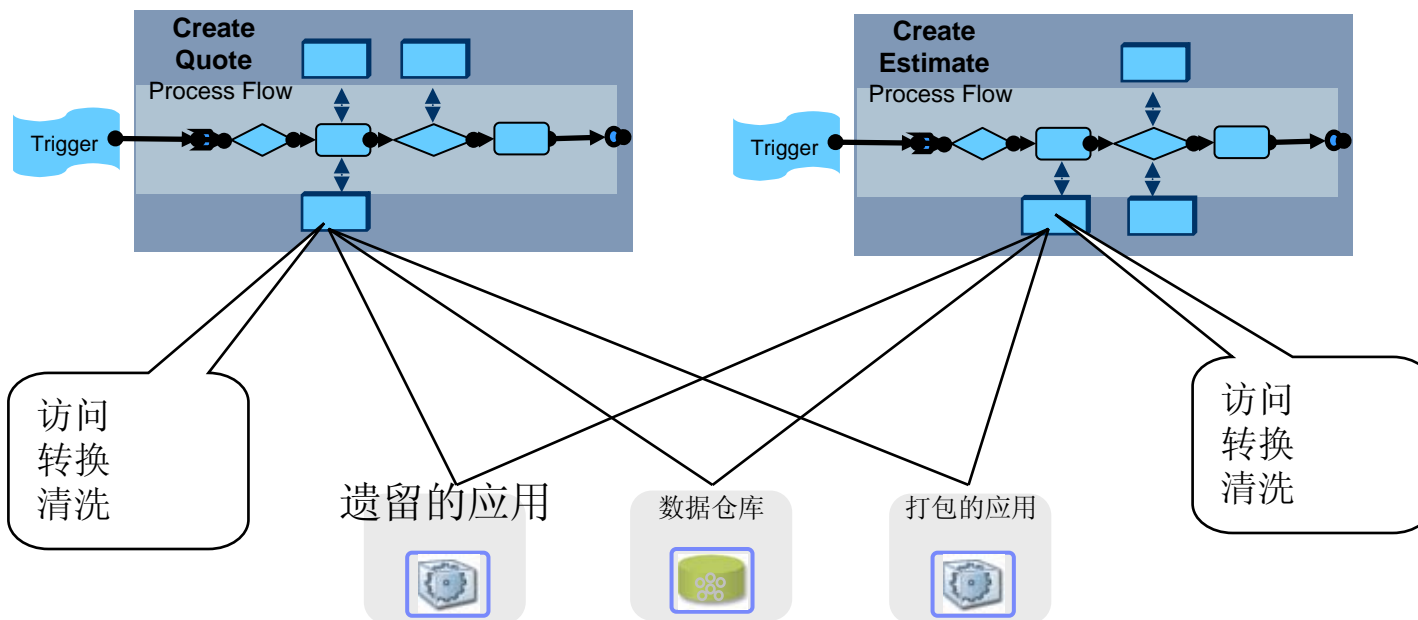


实现业务的灵活性、可重用的困难

- 缺乏标准
- 信息分散
- 不一致的信息
- 外购的应用
- 基础架构建设缺乏合理规划

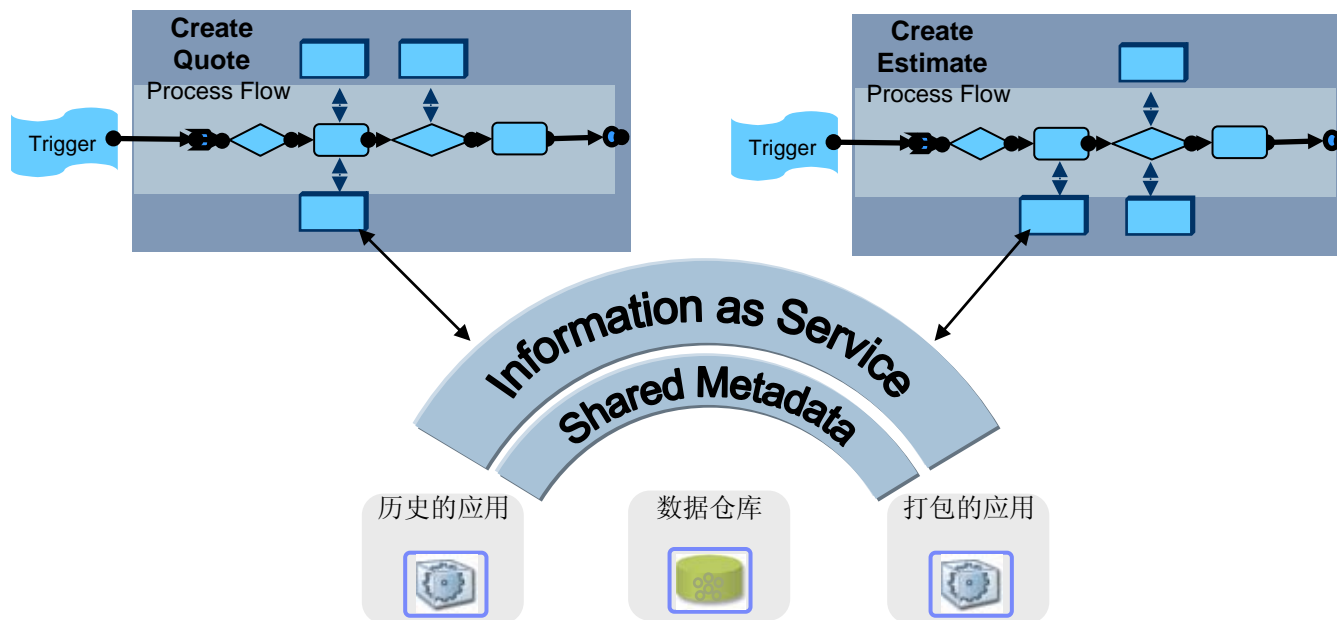


紧耦合的数据束缚了工作的流程



- 数据“见解”的不一致
- 数据来源和处理过程不一致
- 数据的定义不一致
- 多处维护相同的数据
- 变更信息来源和格式不够灵活

信息服务改善灵活性

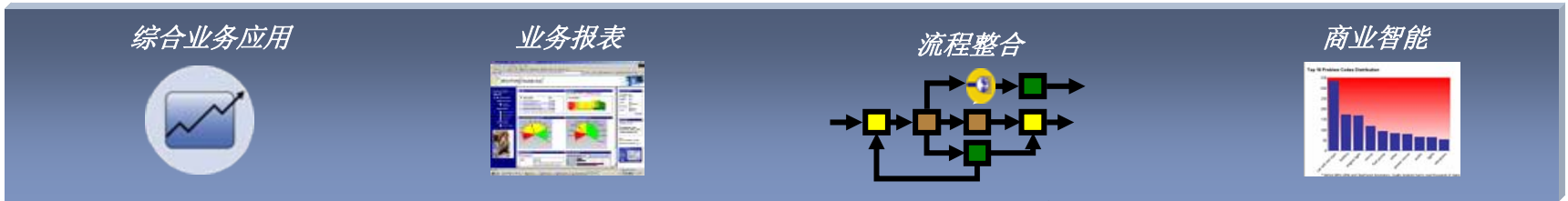


- 对数据封装的统一定义
- 统一对元数据关系的理解
- 对数据应用一致的规则
- 集中的控制和维护数据
- 可灵活改变信息来源和格式

信息作为服务

Information As a Service

破除数据孤岛，提供灵活的信息基础设施



基于标准: XQuery, JSR170, JDBC, Web Services...

实时: 在线帮助, 同步的主数据...

萃取的: Basel II, 业务优化...

异构的应用和信息



信息服务和加速器

Information On Demand的核心



- 虚拟化
- 背景相关 (**context**)
 - 安全
 - 角色
 - 任务

- 基于标准
 - SQL, XQuery, JSR170, XML, Web Services, JDBC, UIMA...

议程

- ✍ **IBM** 信息管理面向**SOA**的发展战略
- ✍ **IBM** 数据服务平台
- ✍ **DB2 9**的新特性



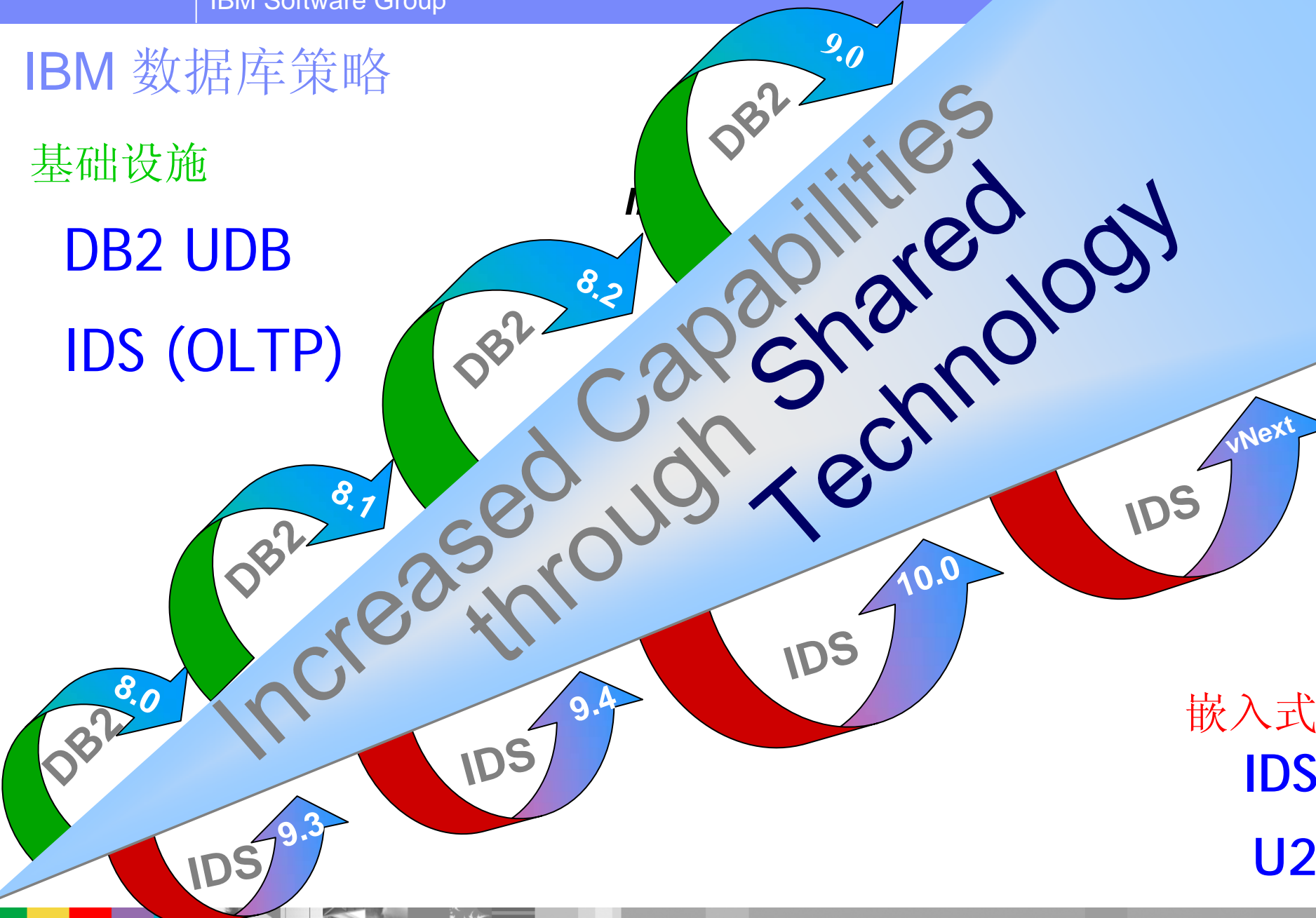
IBM 数据库策略

基础设施

DB2 UDB

IDS (OLTP)

Increased Capabilities
through Shared
Technology



嵌入式

IDS

U2

IBM在数据库领域的技术实力

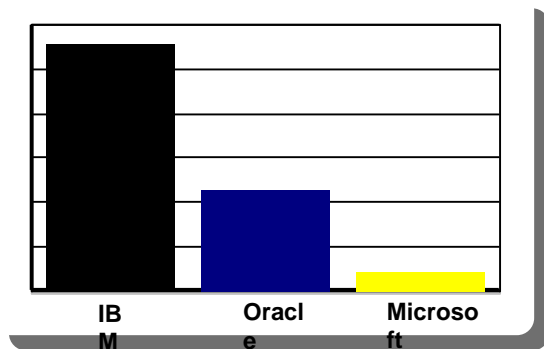
- 关系数据库的发明者
- 持续的研发投入-每年超过**10**亿美元
- **2001**年收购**Informix**，与**Informix**进行技术互补。
- **2004**年推出企业信息整合解决方案
- **2005**年收购**Ascencial**，增强数据处理方面的能力
- **2006**年**IBM**发布**DB2 UDB V9**,其混合层次与关系数据处理能力开创数据库发展新纪元



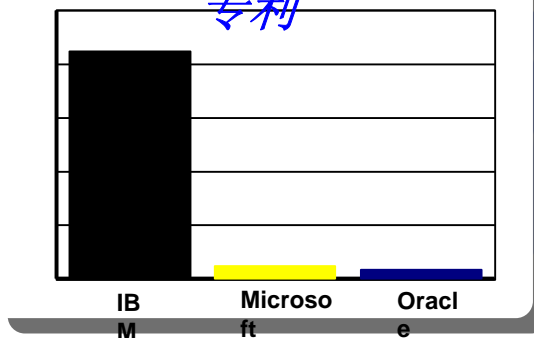
IBM DB2数据库 先进性

- 技术领先
 - 发明了关系数据库和SQL...
 - 拥有7倍领先于行业对手的专利技术
 - 拥有2倍领先于行业对手的标准
 - 第一个混合数据库(融合关系型与层次型存储)
- 广泛认同
 - 用户超过400,000家公司
 - 超过16,000家合作伙伴

SQL标准建议



技术专利



IBM DB2 UDB—开放平台

- HW : Intel & RISC, Palm Device --> Massively Parallel
- O.S.: Windows, Linux, AIX, HP-UX, Sun Solaris
- Information: structured & unstructured
- Size : single user -- small & medium -- very large
- Scale: one partition → 1000 partitions

Everyplace

- Linux
- Windows
- PalmOS
- EPOC-32
- Neutrino



Express

- Linux
- Windows



Workgroup, Workgroup Unlimited

- Linux
- Windows
- AIX
- Solaris
- HP-UX



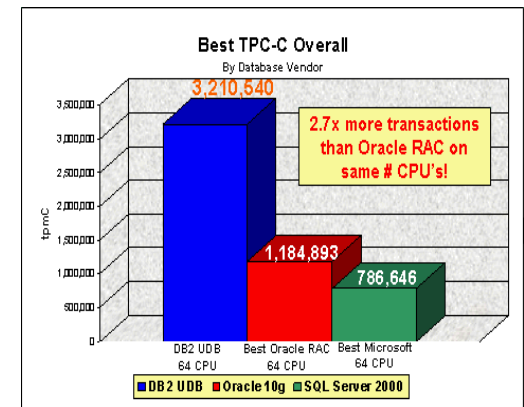
Enterprise

- Linux 32-bit (Intel, AMD)
- Linux 64-bit (Intel, AMD)
- Linux pSeries
- Windows
- Windows 64-bit
- AIX
- HP-UX
- Solaris



DB2 Connect

- Linux – all platforms
- AIX
- Windows
- HP-UX



IBM Informix产品的发展

Informix产品得到了空前的发展，短短三年多的时间，发行了两个重要的版本。

10

专注于关键业务联机事务处理的企业级数据库系统

9.40

性能更上一层楼
高达128PB的扩展能力
更高的数据安全性
灵活的管理能力

9.30

共享SQL语句缓存
模糊检查点
动态日志技术
动态锁管理
数据刀片技术

7.x

高端数据库系统...

- 数据分片, 并行数据库查询, 高可用数据复制和企业级数据复制, 支持大量并发用户等.

6.x

增强的功能和性能...

- 多线程动态可伸缩体系结构(DSA), 异步 I/O, 完善的备份管理等.

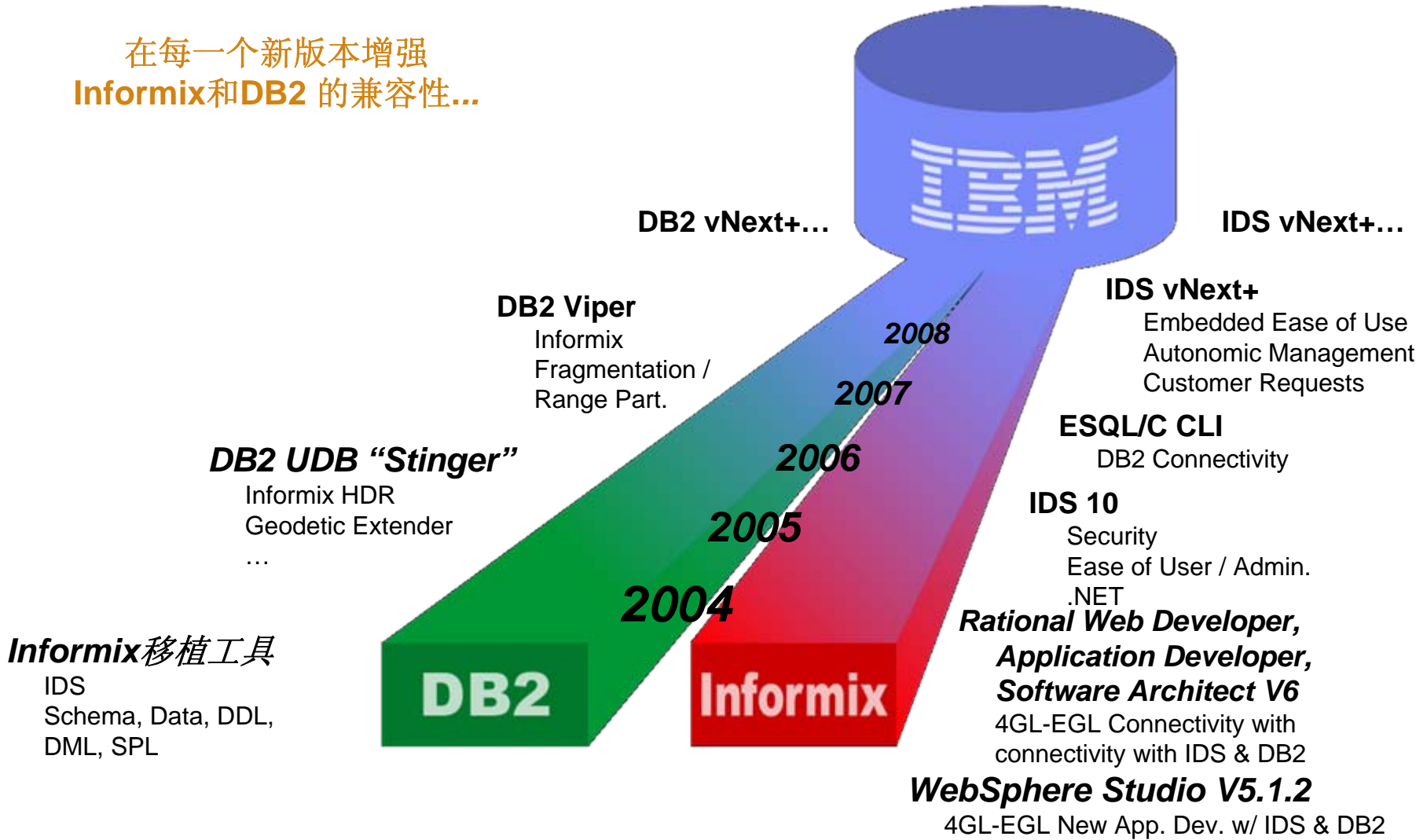
5.x

低端数据库系统, 具有...

- 基于成本的优化器, 两阶段提交, 参照完整性等.

IBM数据库发展策略

在每一个新版本增强
Informix和DB2 的兼容性...



议程

- ✍ **IBM** 信息管理面向**SOA**的发展战略
- ✍ **IBM**数据服务平台
- ✍ **DB2 9** 的新特性



IBM 领导数据管理走向新的高峰

DB2 9: 26年来信息管理划时代的技术创新!

IBM信息管理创新里程碑

1968

第一个层次型的
数据服务器

- IBM 为Apollo太空计划设计IMS层次型数据库

1980

第一个关系型的
数据服务器

- IBM 为 System/38 发布第一个关系模型的数据库系统

2006

第一个混合结构型的
数据服务器

- DB2 Viper 成为第一个同时支持关系型数据和内置XML数据支持的数据库

IBM 持续的创新能力

XML 将改变关系数据库

IBM Moves the Database Goalposts...



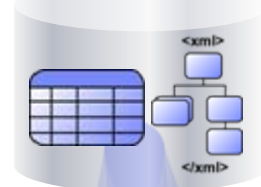
■ 更深入的XML集成，简化和统一基础设施

- ▶ 统一数据和内容
- ▶ 扩展资产利用
- ▶ 利用已有技能

■ 数据库支持“原生”XML文档存储

- ▶ XML文档保真度
- ▶ 优化性能
- ▶ 灵活性
- ▶ 利用成熟的数据库服务

混合处理能力
灵活的基础设施



- XQuery 和/或 SQL
- XML 和/或关系数据

性能、可伸缩性、可用性和安全...

复杂模式演化序列和层次



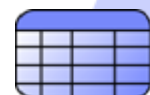
XML RDB Columns



XML 数据库

查询灵活性

内容存储



关系数据库



XML 文档, 交换

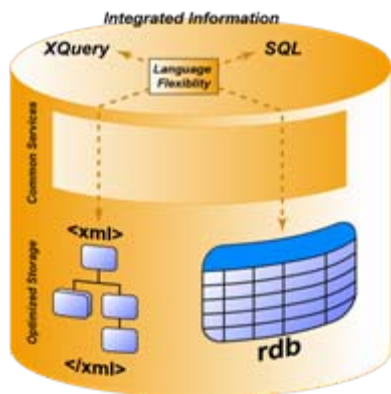
数字化内容

应用无关



层次数据库

文件



DB2 9 的新增功能

- 更快更强大的功能
- 有弹性低成本的运行
- 更快更方便的开发
- 降低管理风险
- 优化的基础架构, 降低管理成本
- 减少信息的复杂性

- 支持 **SOA / XML**
 - ▶ 更快更简单地访问 XML 数据
 - ▶ 无需外部的 XML 处理过程
 - ▶ DB2 Viper 成本 SOA 信息战略的核心
- 加强对**SAP**支持
 - ▶ 更紧密的合作关系(共同开发)
 - ▶ 针对**SAP**应用的优化特性
- 存储优化
 - ▶ 超强存储压缩技术,大大降低成本 (**50%**)
 - ▶ 存储管理简化,降低管理复杂性
 - ▶ 多种分区技术提升性能
 - ▶ 扩展对海量数据存储的支持
- 简化管理
 - ▶ 在易用性方面极大的增强
 - 自动对象管理
 - 自动收集统计信息
 - 资源的自动调节
- 加强安全
 - ▶ 基于标签的访问控制(LBAC)
 - ▶ 实现记录级安全性
- 优化成本
 - ▶ 业界领先的性能
 - ▶ 管理成本持续降低

What is XML?

XML 技术要点

- ▶ XML = Extensible Markup Language (可扩展标记语言)
- ▶ 自我描述的数据结构
- ▶ XML 使用Tags描述文档元素及其属性

XML技术优势 —— 灵活!

- ▶ 可扩展
 - 无固定语法或格式
 - 结构很容易变更
- ▶ 与平台无关
 - 不与任何平台、操作系统或开发语言绑定
 - 易于交换
- ▶ 完全兼容Unicode编码方式

```
<? xml version="1.0" ?>
<purchaseOrder id='12345' secretKey='4x%$^'>
  <customer id="A6789">
    <name>John Smith Co</name>
    <address>
      <street>1234 W. Main St</street>
      <city>Toledo</city>
      <state>OH</state>
      <zip>95141</zip>
    </address>
  </customer>
  <itemList>
    <item>
      <partNo>A54</partNo>
      <quantity>12</quantity>
    </item>
    <item>
      <partNo>985</partno>
      <quantity>1</quantity>
    </item>
  </itemList>
</purchaseOrder>
```

XML —— 数据模型

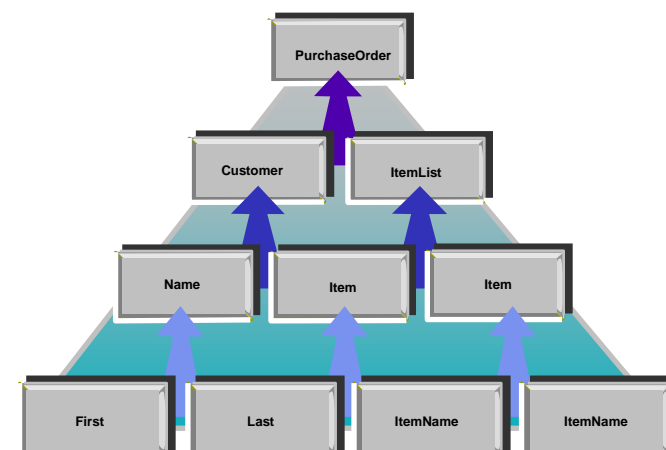
- 关系模型:
 - 关系(表)
 - 属性(列)
 - 基于集合操作
 - 严格的Schema限定

POID	CustomerID	ItemID
12	1	2
162	3	4
162	3	5

Id	LastName	FirstName	Street	City	State	Zip
1	Pirahesh	Hamid	1 Harry Rd	San Jose	CA	95141
3	Selinger	Pat	555 Bailey Ave	San Jose	CA	95141

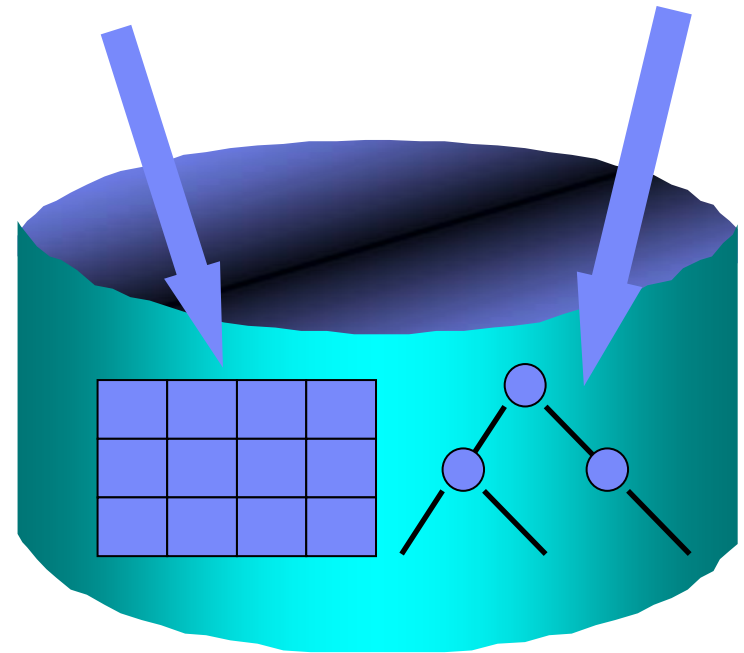
ItemID	Name
2	#6 wire nut
5	Small Walrus
4	Apollo moon rocket

- XML 数据模型:
 - 节点(元素、属性、备注等)
 - 节点间存在着相互的关系
 - 基于序列的操作
 - 灵活的Schema



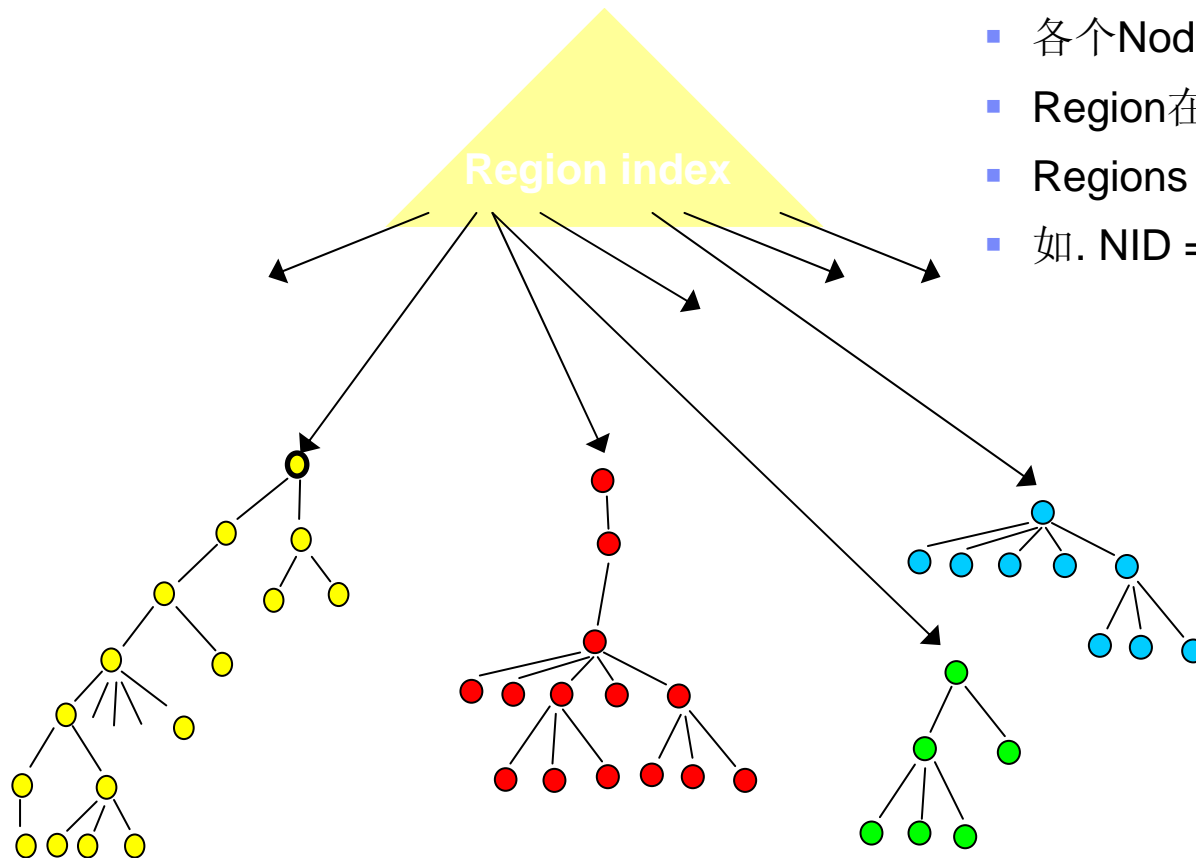
XML在DB2中的存储: Native XML Storage

- XML被解析为层次方式进行存储
`create table dept`
`(deptID char(8), ..., deptdoc xml);`
- 关系列仍以关系方式进行存储 (tables)
- XML列也以Native方式存储
- 以UTF8方式存储



XML Region Index

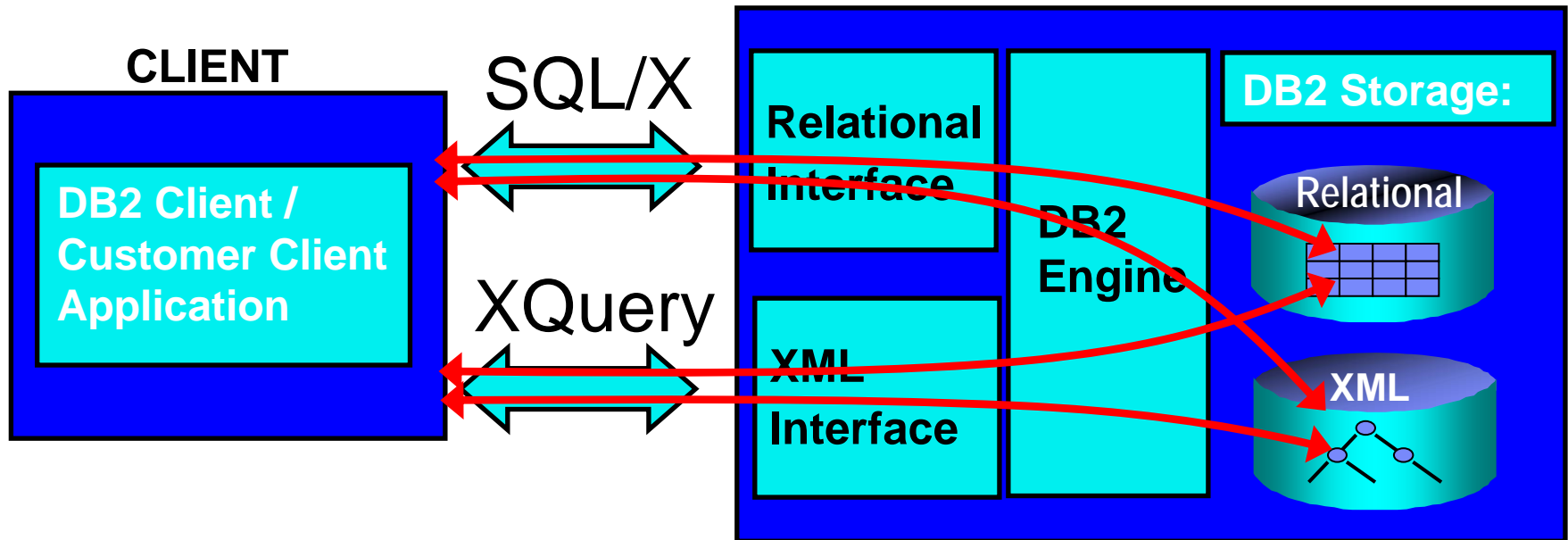
- 无需用户定义, XML存储的缺省部分



- 各个Node在物理上相互连接
- Region在逻辑上相互连接
- Regions index 的键值是 NodeID (NID)
- 如. NID = 1.2.4.2

DB2 中使用 XML

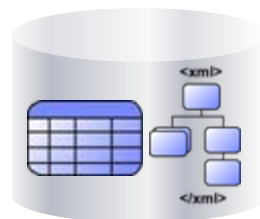
- 同时访问关系型数据 或 XML数据
- 可使用SQL风格 或 完全的XML风格的语句
- 可用SQL方式访问XML数据
- 也可用XML方式访问关系型数据



DB2 9 中的XML



SQL Person... "I see a world class RDBMS that also supports XML"



DB2 with XML Support



XML Person... "I see a world class XML repository that also supports SQL"

DB2 在所有特性/接口中都整合了对XML数据的支持

新的基于XML的应用程序使用DB2可以获得如下优势:

- 可以无缝地访问关系数据库资源
- 体系架构拥有企业级XML应用的访问能力



IBM Software Group

表压缩

节省大型数据库的存储设备投资



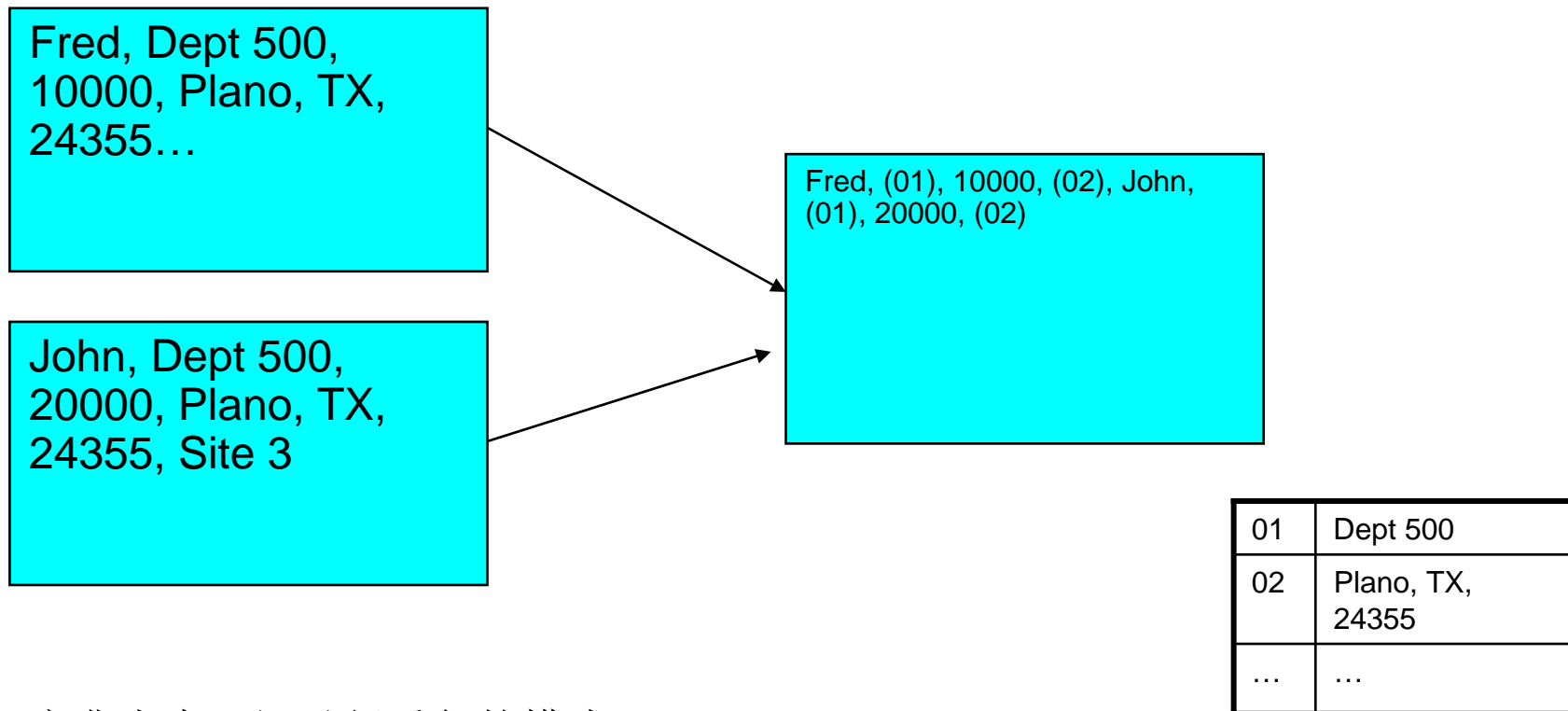
ON DEMAND BUSINESS™

行压缩

- 基于字典的压缩——为压缩/解压数据建立符号表
- 使用基于静态字典表的Lempel-Ziv (LZ) 算法
- 数据在内存/磁盘中均以方式驻留
 - ▶ 节省大量I/O
 - ▶ 显著节省内存占用
 - ▶ 占用CPU负载: 行数据处理前, 需要进行解压缩操作
- 日志数据也以压缩方式进行存储
- 如果有些行数据的压缩并不能显著节省空间, 那么DB2将自动选择不进行压缩
- 数据中拥有较多重复模式时, 压缩可获得较好的效率. 文本数据一般会获得较好的压缩效果。



使用字典表实现的数据行压缩方法



字典表中可记录行重复的模式

- 重复信息可以是 跨列组合 或 在列内提取子元素

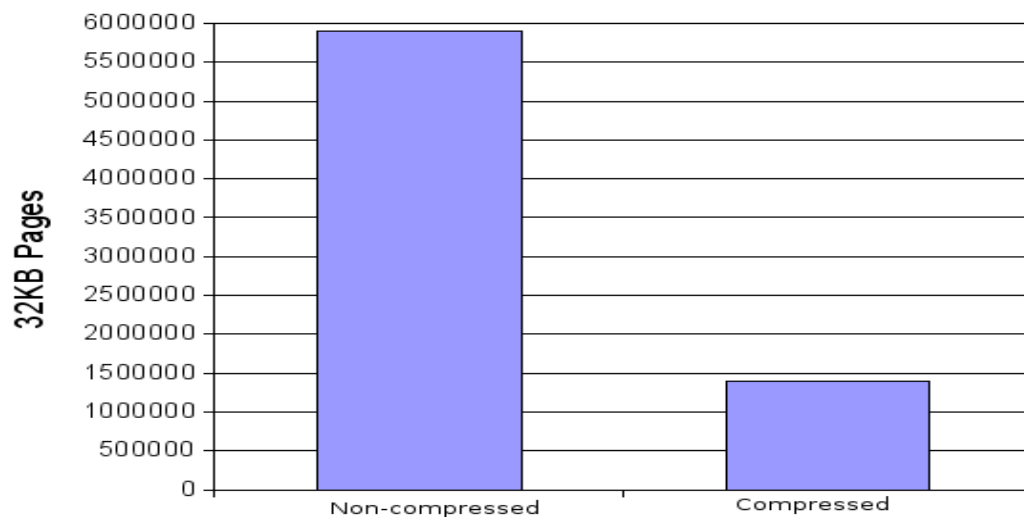
DB2 Viper 提供强大的数据压缩能力

大大减少存储成本，提高I/O效率

压缩类型	32KB 页数	磁盘空间需求
无压缩	5893888	179.9GB
记录压缩	1392446	42.5GB

% 节省存储: 76.4%

T1 Compression - 179.9GB Initial Size



国内测试结果:

- 金蝶软件: 数据存储减少60%以上, 同时交易性能提升10%
- 广东移动: 数据存储减少50%
- 招商银行: 数据存储减少50%



IBM Software Group

DB2 自主管理特性的增强

降低用户的总体数据库系统拥有成本

 ON DEMAND BUSINESS™

自动化 —— DB2做得更多！

- 大部分参数缺省设为Automatic。
- 例如：
 - ▶ 配置助手 (2 second tuning)
 - ▶ 自主化内存伸缩管理
 - ▶ 自动化统计信息收集机制
- 自动化页面刷新与磁盘预读进程管理
 - ▶ NUM_IOSERVERs 与 NUM_IOCLEANERS 缺省将设置为 AUTOMATIC
- 自动化存储管理，存储空间可以自动动态伸缩
 - ▶ CREATE DATABASE DB1 AUTOMATIC STORAGE YES
 - ▶ CREATE TABLESPACE TS2 MANAGED BY AUTOMATIC STORAGE
- 自动化数据库恢复、重定向脚本生成
- 更多的容错机制：Read Retry, Isolation Level Diagnose ,Scehma Copy



拥有自适应能力的内存自我调整机制

DB2 9 中引入了一个具有革命性意义的内存自我调整系统——the Self Tuning Memory Manager (STMM)

- ▶ 自动控制DB2主要的内存对象：
 - Sort, locklist, package cache, buffer pools, and total database memory
- ▶ 无需人工干预的内存自我在线调优
- ▶ 自我感知工作负载、按需调整内存大小
- ▶ 能够迅速适应工作负载的突然变化，自动重新划分内存区域
- ▶ 自动化适应工作负载的周期性变化

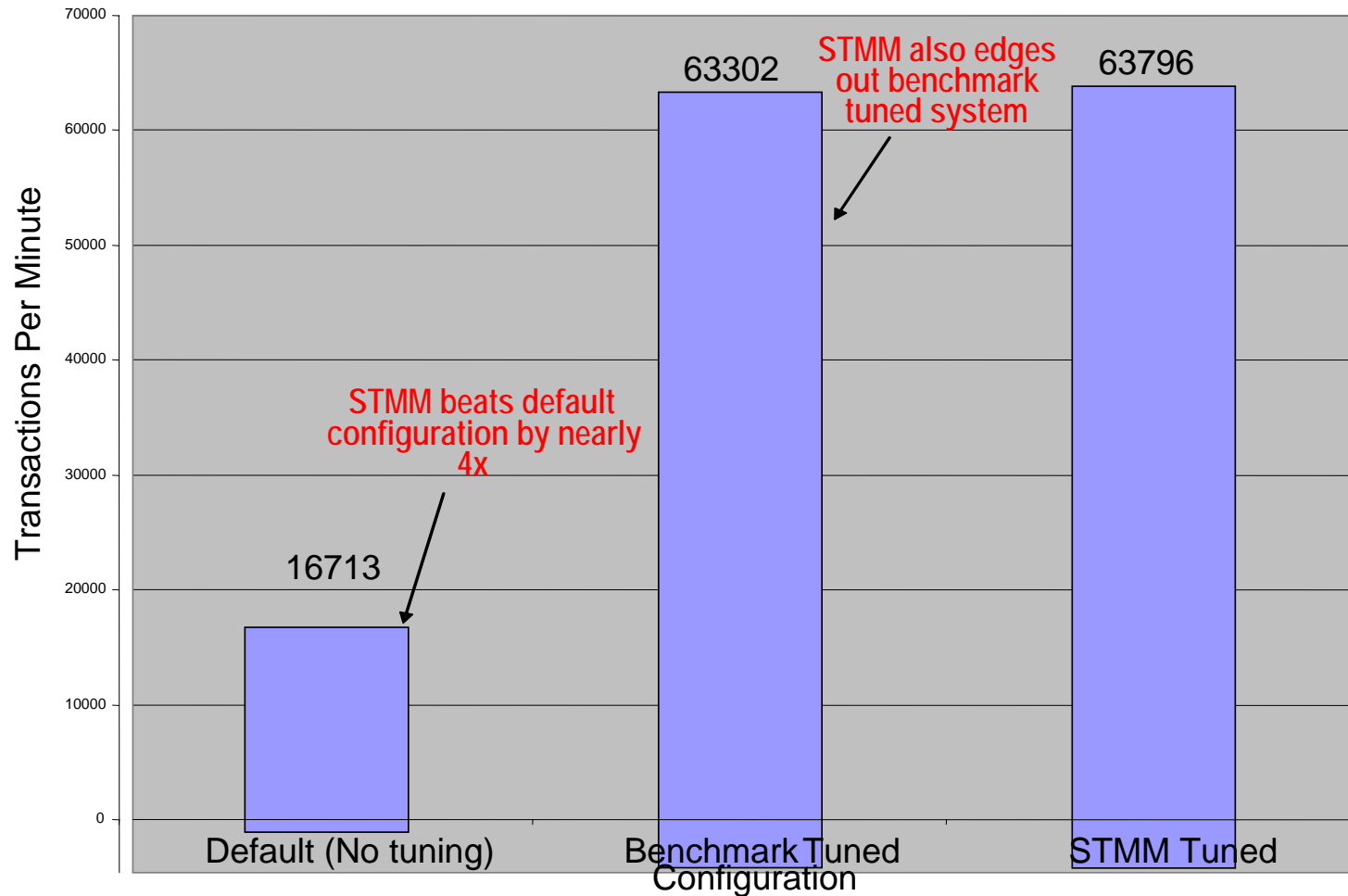


STMM 实际测试 – Dropping an Important Index

TPCH Query 21 - After drop index - Average times for the 10 streams



STMM 测试 – 不同配置方法的比较





IBM Software Group

精确的数据安全性

精确到表的行和列的安全控制



ON DEMAND BUSINESS™

安全性 - LBAC

- **Label Based Access Control (LBAC)**
 - ▶ “label” 与用户会话和数据的行列关联
 - ▶ 规则用于比较用户和数据的标记，提供记录级的访问控制
- 标记可包含多个部分
 - ▶ 层次型, 组或树类型
 - ▶ 行标记是被保护表额外增加的一个列
 - ▶ 用户标记由安全管理员授权
- 与DB2 for z/OS v8 中基于标记的安全性控制机制类似



LBAC 查询

```
SELECT * FROM EMP
WHERE
  SALARY >= 50000
```

No LBAC	SEC=254	SEC=100	SEC=50	ID	SALARY
Red				255	60000
Red	Green	Green		100	50000
Red	Green	Green	Green	50	70000
				50	45000
				60	30000
Red	Green			250	56000
Red	Green			102	82000
Red	Green	Green		100	54000
				75	33000
				253	46000
Red	Green	Green		90	83000
Red	Green			200	78000



IBM Software Group

DB2 数据库规模的扩展

更大的数据容量，更少的约束

——混合分区的强大支持



ON DEMAND BUSINESS™

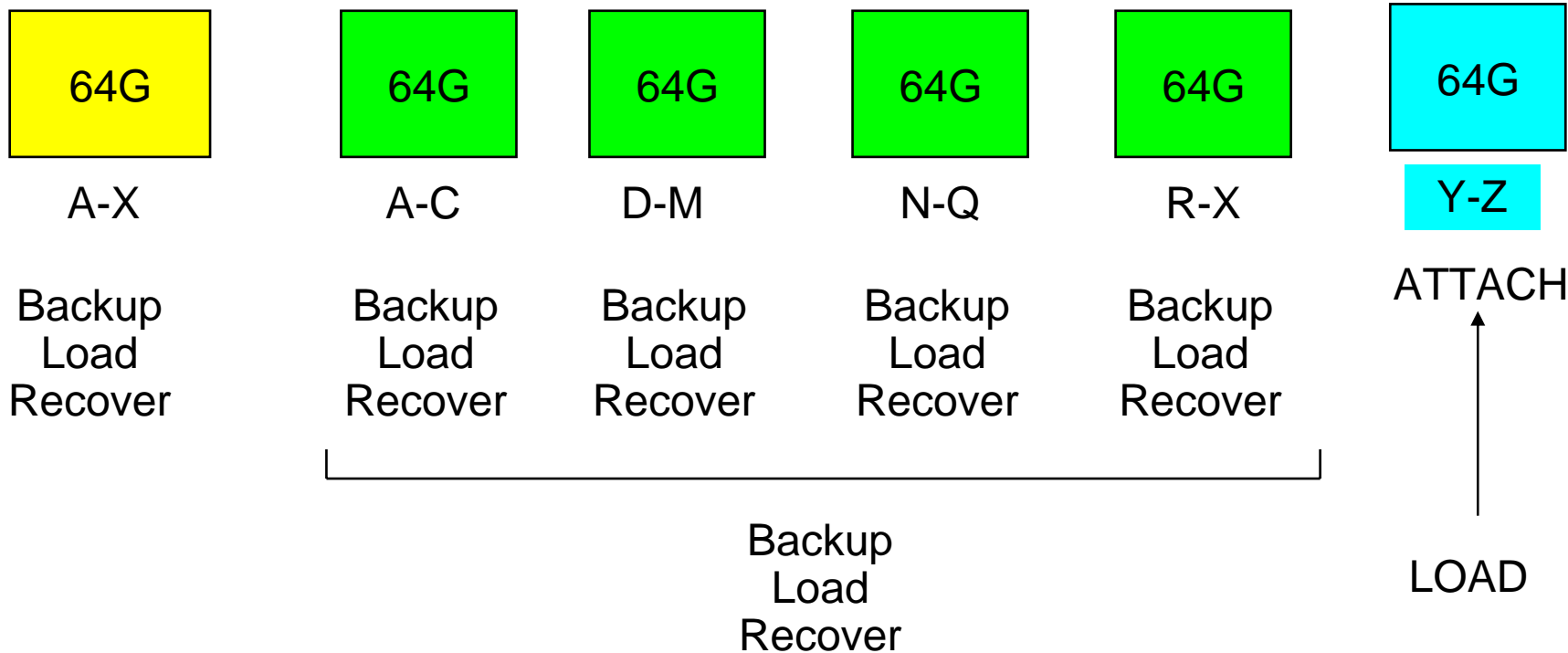
新的分区方式（Table Partitioning）

- 为什么要引入表分区？
 - ▶ 在多个物理设备上存储数据、跨多个表空间存储数据
 - ▶ 每个表空间中仅存储固定范围的数据，能够更有效地访问
- 表分区能带来什么？
 - ▶ 提高单表容量限制
 - ▶ 提高大数据量表的可管理性
 - ▶ 通过分区提高SQL执行效率
 - ▶ 提供快速并且在线的数据载入和卸除
 - ▶ 与Informix的更紧密融合

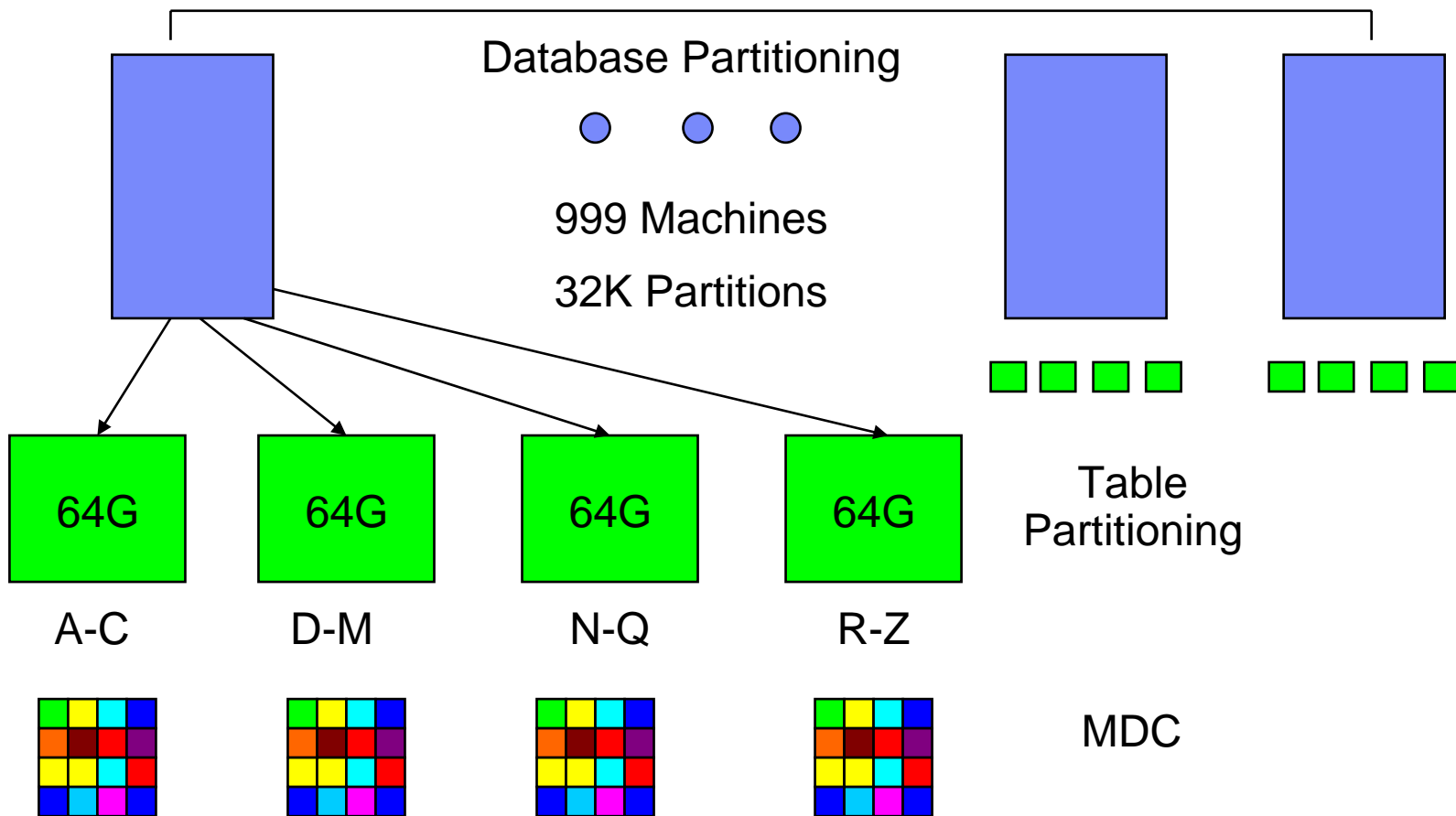


Table Partitioning

32K Partitions



DB2 9 混合分区

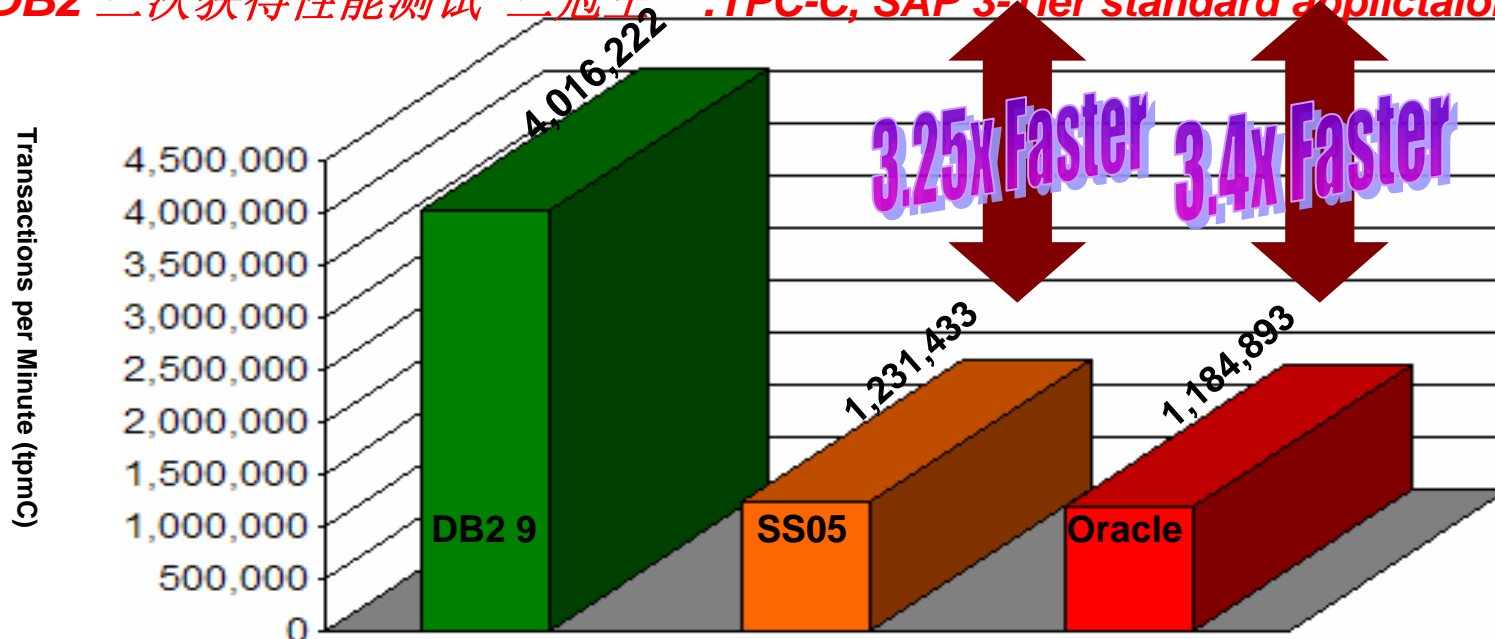


DB2 9的高性能

- 模拟320万用户在1590亿行数据的数据库上
- 数据库初始数据量32.7TB, 用2TB RAM,系统以每小时230G的数据增长 (15亿的交易量在2.5个小时内完成)
- 从16-way 到 64-way ,DB2 线性增长达 98%, \$/tpmC性价比提高 33%!!

Top TPC-C Results on 64 Core

DB2 三次获得性能测试“三冠王” :TPC-C, SAP 3-Tier standard applictaion, 10G TPC-H



Note: Oracle has a higher result on a 32-way, but even if it scaled linearly, DB2 would still be 70% faster

Results current as of 07/26/06

国内用户选择IBM DB2 UDB

DB2 UDB在中国目前最大的OLTP系统—中国银联系统中运行。

- DB2 UDB在中国未来最大的OLTP系统—中国人民银行国库信息处理系统中运行。
- 鞍钢新一代生产管理系统
- 华为开始基于DB2 UDB开发其新产品作为核心数据支持平台
- 中国烟草专卖体系及烟草生产企业
- 中国电信，中国移动等
- 北京市政府选用了DB2 on Linux作为其政府信息管理系统

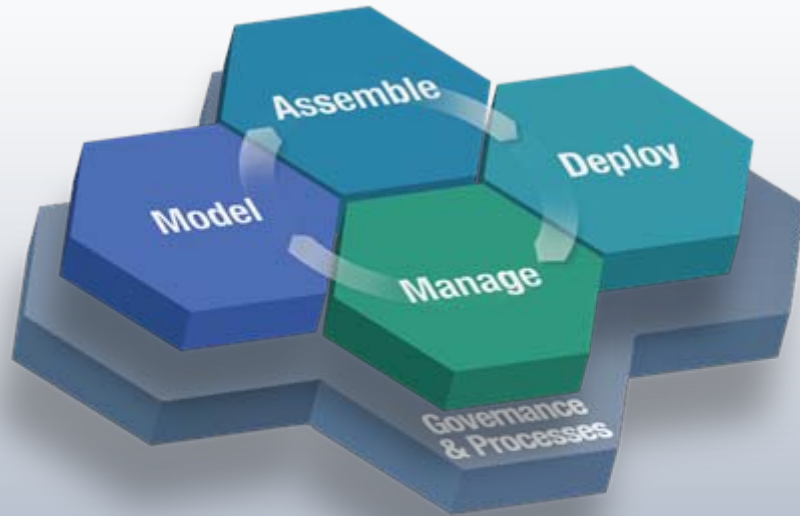
IBM DB2数据库：商业价值

- **DB2 UDB** 领先的技术为业界应用提供极具竞争力的支持.
- **DB2 UDB**集成业界领先的开发环境, **ISV** 应用,以节约成本,加速新应用的上线
- **DB2 UDB** 通过一组自主管理特性极大地降低了系统管理的复杂度,减少系统管理时间,降低对**DBA** 的技能要求. 与管理其他企业级数据库相比,为用户节省巨大开支.
- **DB2 UDB** 在业界良好的服务与支持声誉,为保护用户投资和数据资产提供强有利保障.
- **DB2 UDB** 加速价值体现,为客户提供最优的总体拥有成本.
- **IBM** 帮助合作伙伴提高利润,减低风险,为客户提供最好的解决方案.

IBM 信息管理软件

——提供超越传统知识库的价值

信息随需应变的面向服务架构



除了IBM,没有任何一个厂家能提供如此深度和广度的能力

■ 数据服务

- DB2, Informatica, Cloudscape, U2
- Tools...

市场份额 #1, #2
450,000+ 用户

■ 内容和发现服务

- Content Management, Information Services
- Discovery, ...

市场份额 #1
13,000+ 用户

■ 信息集成服务

- Quality Services
- Transformation
- Federated
- Metadata Services...

市场份额 #1
5,000+ 用户

■ 信息加速

- Master Data
- Entity
- Information Warehousing
- Customer Dashboards
- Industry Data Models...

领导地位
1,000+ 用户

谢谢!

