

微软并行数据仓库

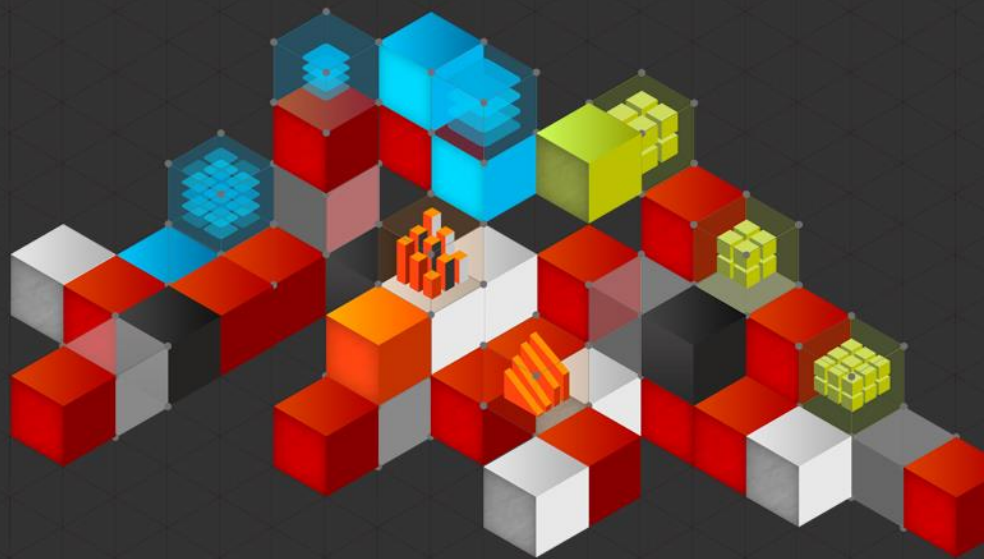
SQL SERVER 2012 PARALLEL DATA WAREHOUSE APPLIANCE

Insights on Any Data of Any Size

乔怡

高级项目经理

微软亚太研发集团 - 云计算创新中心



SQL SERVER 2012 PARALLEL DATA WAREHOUSE 软硬件整合优化的并行数据仓库一体机



- 从TB到PB 的高可扩展性平台
- 使用MPP架构实现卓越的可扩展性
- 硬件厂商的选择 
- 通过使用遵循行业标准的硬件降低成本
- 与Microsoft 商业智能工具的深度集成
- 充分利用最新软件技术
 - Windows Server 2012 Storage Spaces
 - Windows Server 2012 Hyper-V
 - SQL Server 2012 xVelocity 列存储

PDW多节点分布式无共享（MPP）并行处理架构

客户端请求

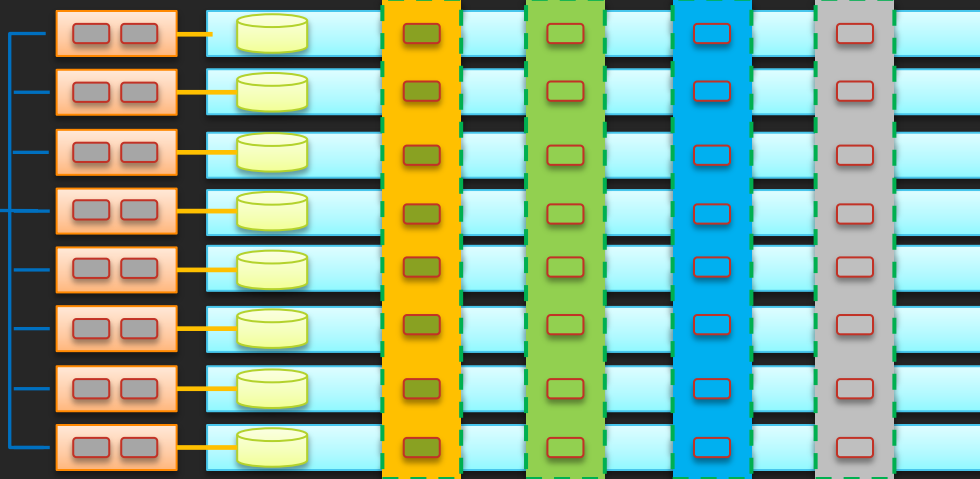


...

PDW控制节点



PDW计算节点



...

微软并行数据仓库全球案例

#1 Unit Share for
DW All-up
(SQL Server)

SQL Server
Enterprise Revenue
Growing > 30%

PDW Volume Grew
6x Year Over Year

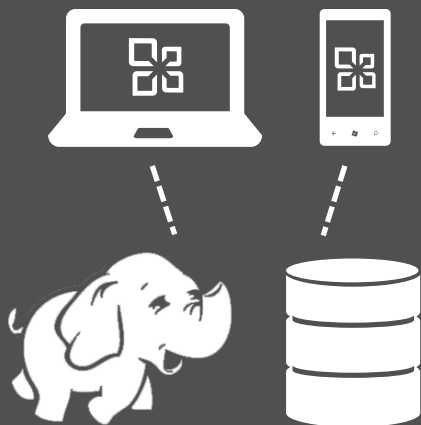
Strategic PDW
Customer
Deployments in
Fortune 1000
Companies



SQL SERVER 2012 PARALLEL DATA WAREHOUSE

Insights on Any Data of Any Size

为大数据而建



新一代的性能和规模

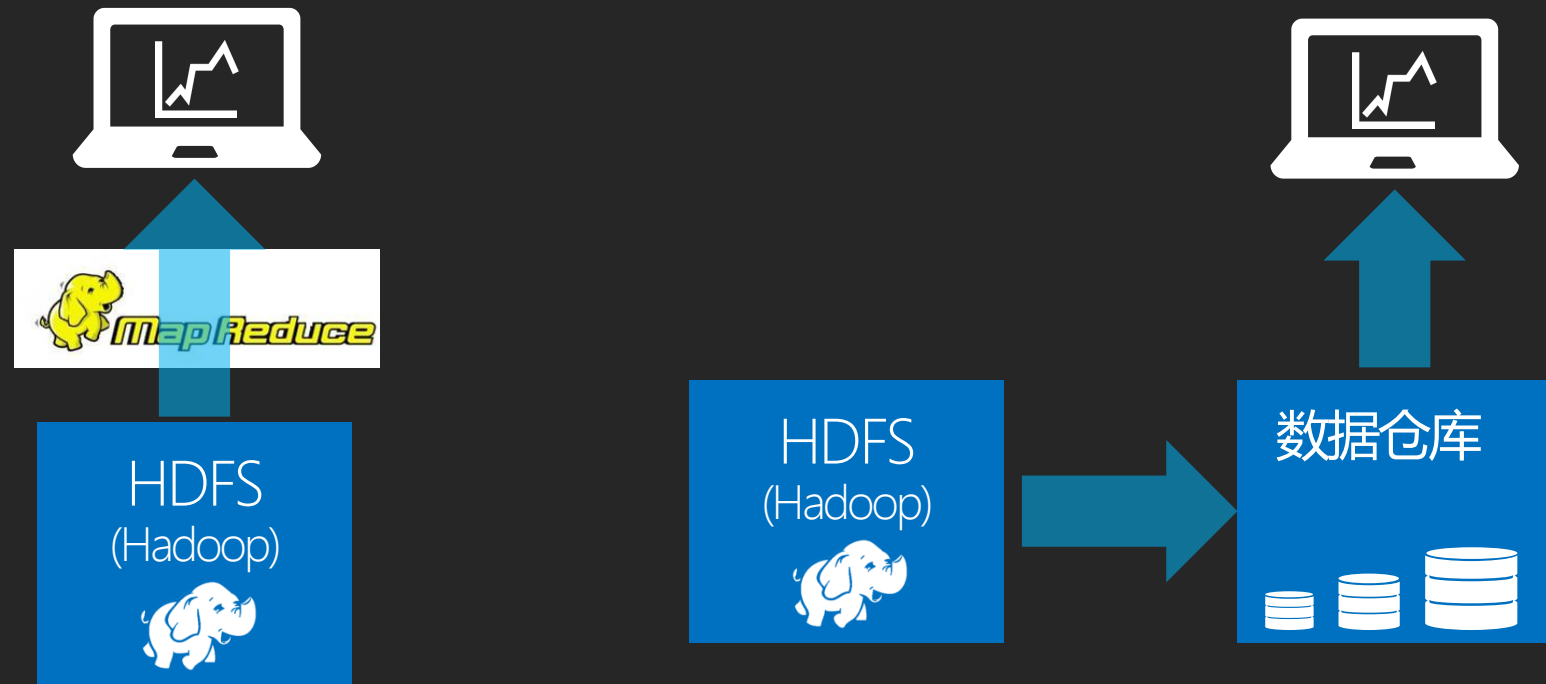


最优化的软硬件价值



为大数据而建

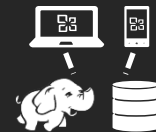
当前大数据分析的挑战： 大量新技术需要学习掌握、周期长、见效缓慢



学习MapReduce

需要IT人员提前手动将数据从HDFS导入数据仓库或集市，才可进行分析

POLYBASE: 数据处理的突破性技术



SQL



SQL Server 2012
PDW Powered
by PolyBase

统一查询，结构化和非结构化数据

- 查询关联Hadoop表和关系型数据库表
- 采用标准SQL语言
 - Select, From Where

Database



HDFS
(Hadoop)



沿用现有
SQL 技能



无需IT人员
介入



节省时间和
成本

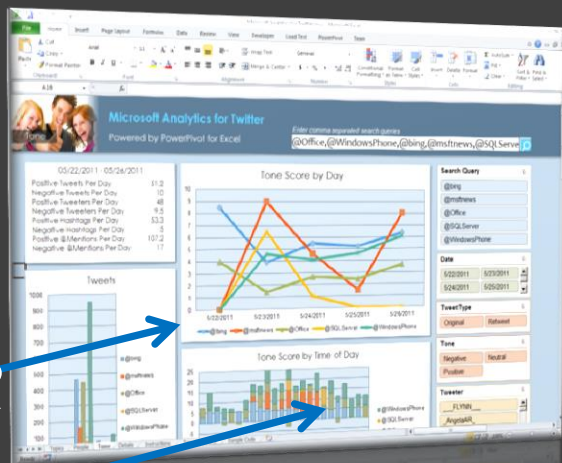
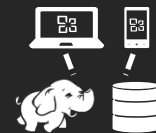


分析多样的
数据类型



利用现有的EXCEL技能来分析大数据

使用熟悉工具来分析结构化和非结构化数据



Hadoop
数据

结构化
数据

用熟悉的工具来分析大数据

- PDW与微软BI工具的天然集成
- 在同一表格中分析结构化和非结构化数据
- 广泛使用的熟悉工具

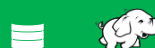
广泛使用的
Excel



无需IT
人员介入



分析多样的
数据类型



新一代的性能和规模

当今性能和规模的挑战

纵向扩展



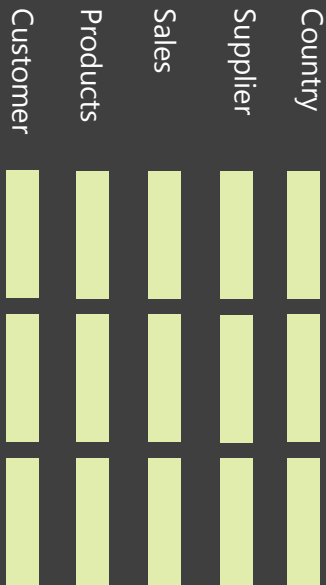
纵向扩展规模有限，无法满足需求的增长

行存储



数据仓库分析查询性能欠佳

XVELOCITY 技术提供新一代性能



列存储技术提供强大性能

- 可更新的聚集列存储
- 数据以列形式存储
- 内存优化技术提供新一代性能
- 灵活可更新，同时满足批量导入或小量数据的加载更新需求

高至**50X**
性能提升



高至**15x**
压缩比例



节省时间和
成本



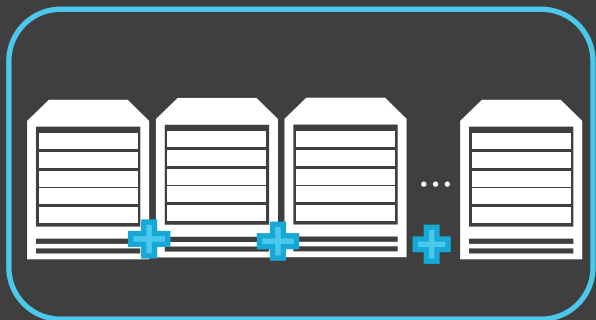
实时
数据仓库



大规模并行处理引擎 MASSIVELY PARALLEL PROCESSING (MPP) PDW作为SQL SERVER的横向扩展解决方案



横向扩展



MPP 提供线性横向扩展能力

- Massively Parallel Processing (MPP)
大规模并行处理架构
- 横向扩展：按需添加硬件实现线性扩展
- 无共享架构

比SMP DW
快10倍



复杂作业计算



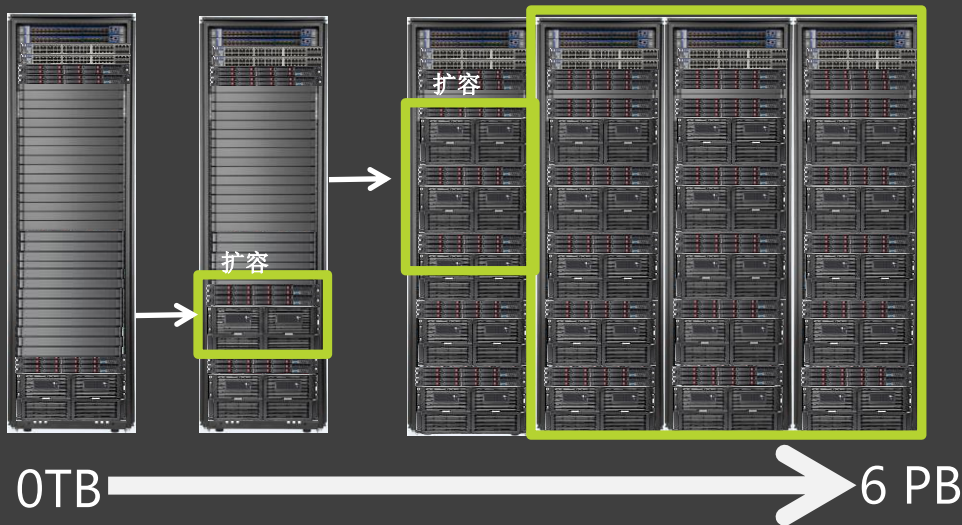
线性扩展



扩展灵活方便
(No forklift)



灵活可扩展 – 从几个TB到6PB的线性扩展



最小 (0TB) 到最大 (6PB)

- 从几个TB的小型数据仓库开始
- 扩展到 6 PB

聚沙成塔
成长无限



容量最大
按需扩容

PB

宕机为零
高枕无忧



最优化的软硬件价值

当今的挑战：硬件和软件的集成



用户必须搭建硬件环境、
配置并调试软件环境

\$\$\$\$

一体机具备预装的硬件和软件，
提供性价比优势。

硬件和软件整合优化：一体机的体验



开箱即用的一体机：预装好的硬件 + 软件

- 微软联合HP和Dell共同研发推出
- 预装硬件
- 预装软件



插电即用



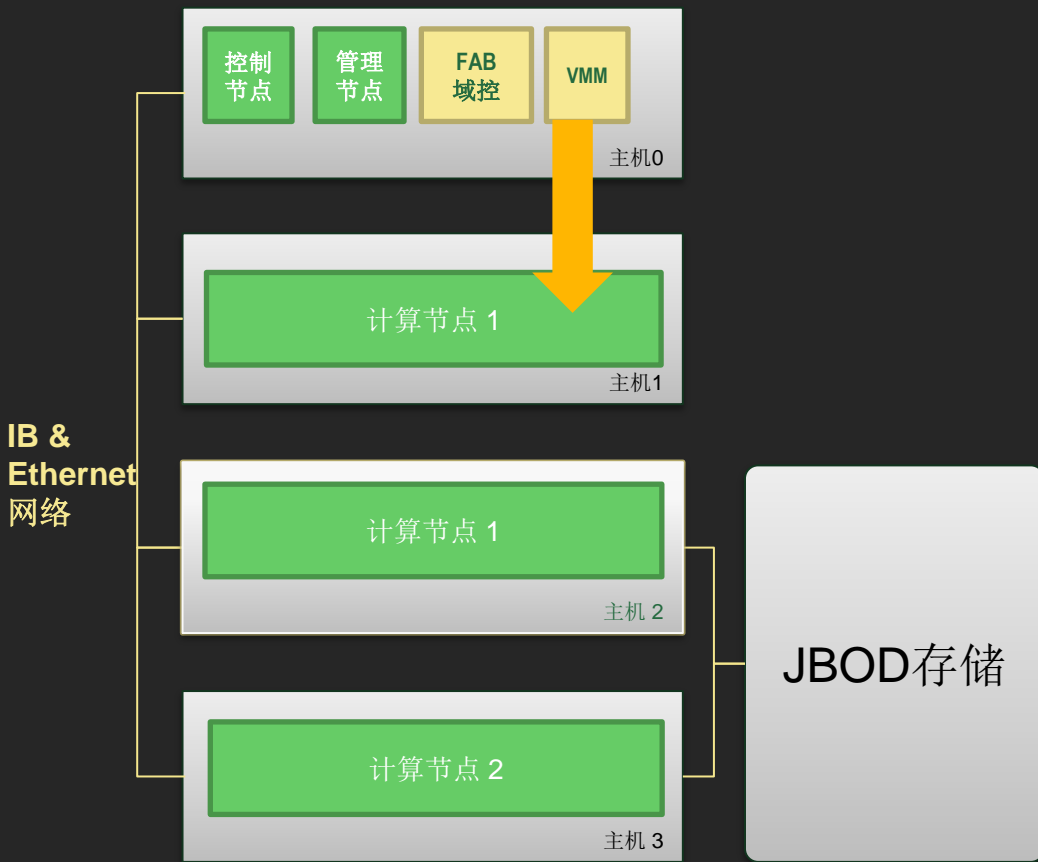
内置
最佳实践



节省时间

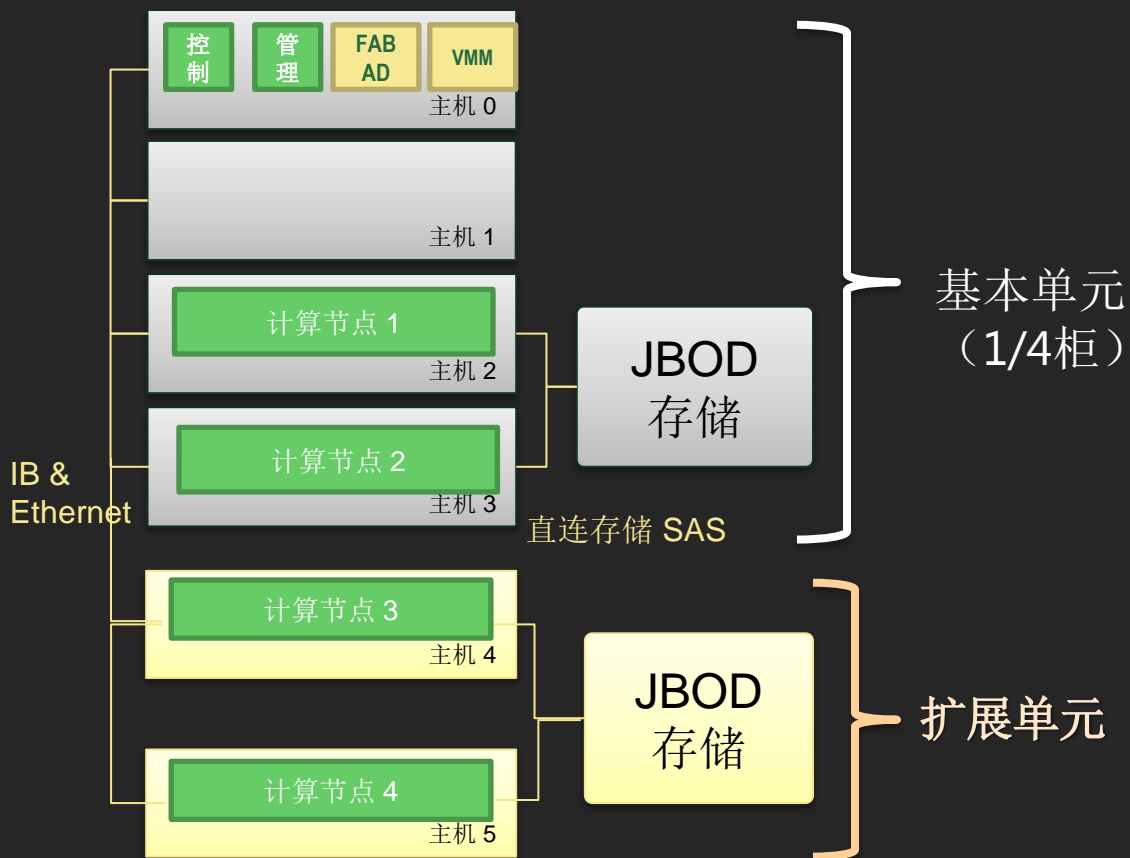


WINDOWS SERVER 2012云计算技术带来更加精简高效的硬件资源设计



- 所有管理、计算负载均在虚拟机当中进行
- 系统自动管理JBOD上的物理磁盘使用
 - ✓ 逻辑镜像磁盘
 - ✓ 热备盘
- 只要有一个连到某个存储的节点是活动的，就可以通过集群共享存储机制（Clustered Shared Volumes）让所有节点都能访问该存储阵列
- 虚机自动进行故障迁移
- 采用Windows 故障迁移管理

按需增加扩展单元 – 从几个TB到6PB按需扩容



- 从小规模开始
 - 1/4柜,
 - < 50TB
- 最大可扩展至
 - 56个计算节点
 - 高达 6 PB 容量
- 系统内置高可用设计, 自动故障转移

硬件配置

服务器	CPU	内存	存储	硬盘	网络1 数据	网络2 管理
1u Intel 2 PCIe G3	E5-2690 2 Socket 16 Cores	256 GB	JBOD (60-70 HDD) 6Gb SAS	1,2 or 3 TB 7200 RPM NL-SAS	56Gb FDR Infiniband	1Gb Ethernet

CPU
E5-2690 2 Socket 16 Cores

存储	硬盘
JBOD (60-70 HDD) 6Gb SAS	1,2 or 3 TB 7200 RPM NL-SAS

网络1 数据	网络2 管理
56Gb FDR Infiniband	1Gb Ethernet

- 总共32 CPU逻辑核
- 较前一代Intel处理器的SQL 查询吞吐量提升达20%

- 更加节约成本的 Direct Attached Storage直连存储，以及 3.5” NL-SAS Terabyte Drives
- Average of 2x higher spindle density over V1

- 升级至 56Gb FDR Infiniband
- 4x10GbE external interfaces
- 不再使用Fiber Channel Storage Network

SQL 2012 PDW为你带来无限的选择



Microsoft®
SQL Server® 2012
Parallel Data Warehouse

灵活可扩展



新一代
性能和规模

低成本



低成本实现
最优化的软硬件价值

SQL



采用标准SQL
实现大数据集成查询

简单易用



与广泛使用的BI工具
天然集成

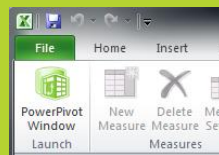
微软端到端的大数据平台实现数据与信息消费化

洞察分析



自助
运维
预测

移动
实时
协作



数据集市
外部数据源
及服务



发现和采集

数据丰富



清洗和转换



分享和监控

数据管理



关系型数据



非关系型数据



多维数据



流数据





© 2012 Microsoft Corporation. All rights reserved. Microsoft, Windows, Windows Vista and other product names are or may be registered trademarks and/or trademarks in the U.S. and/or other countries. The information herein is for informational purposes only and represents the current view of Microsoft Corporation as of the date of this presentation. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information provided after the date of this presentation.

MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS PRESENTATION.