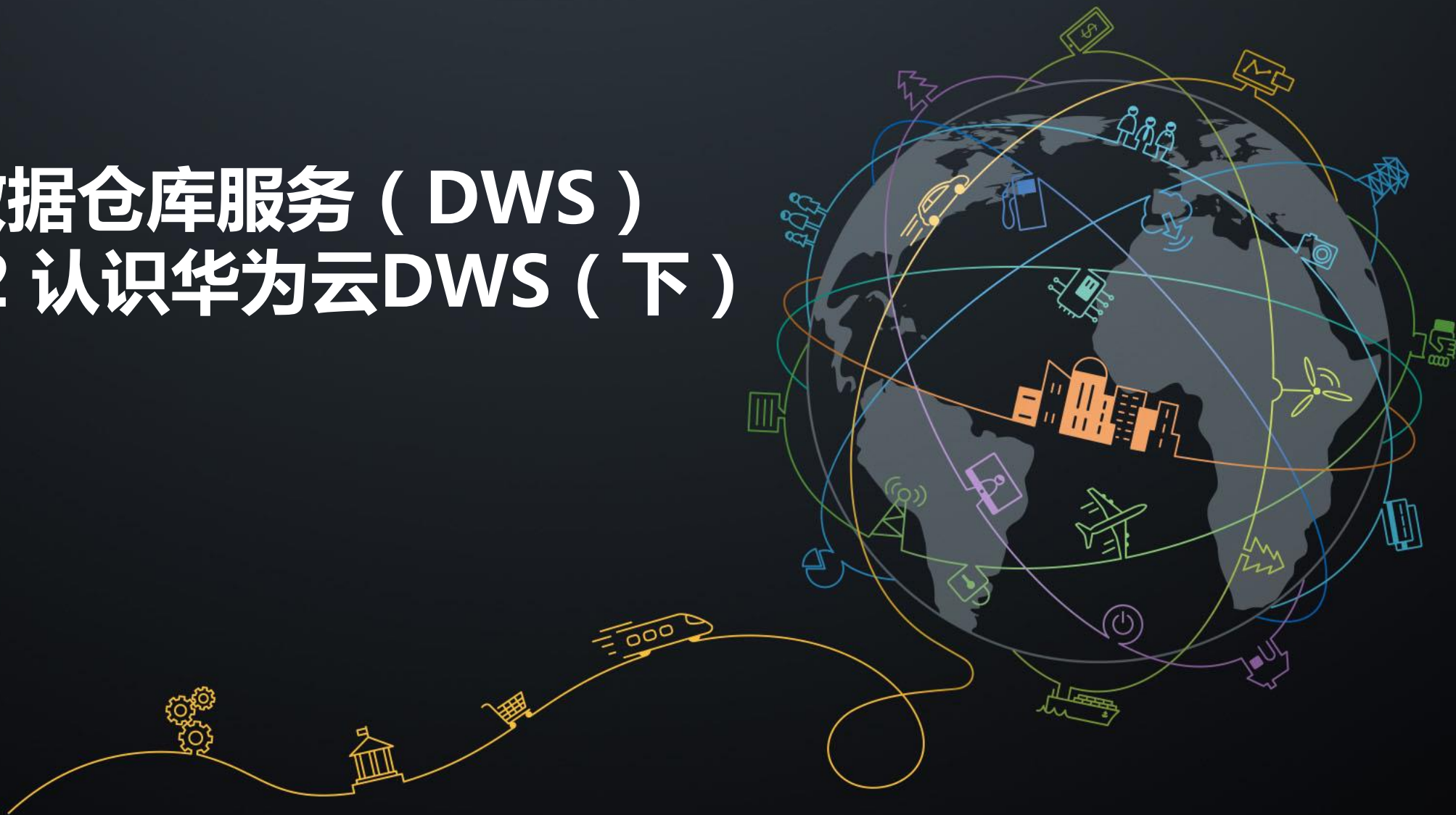


华为数据仓库服务 (DWS)

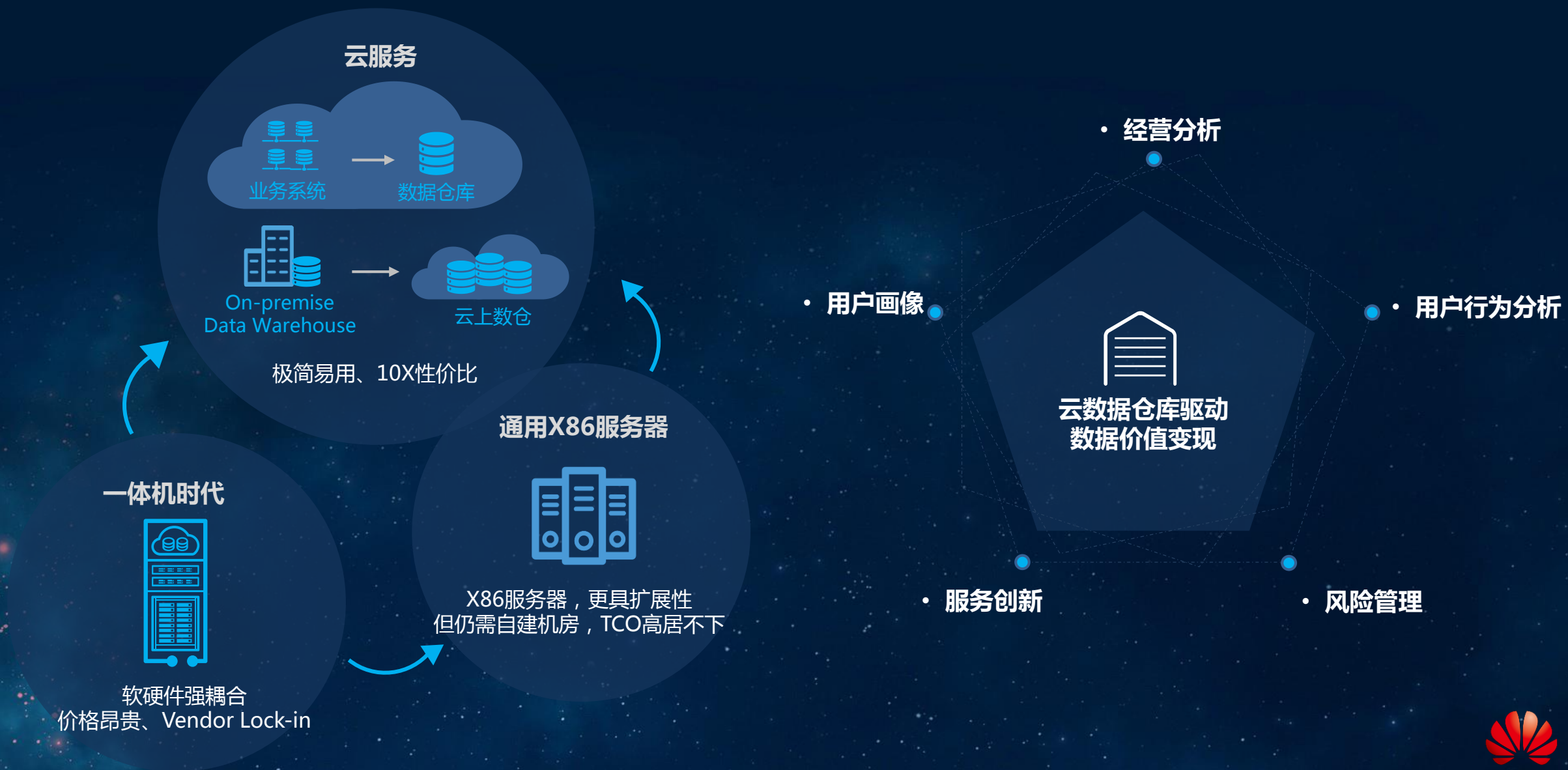
day02 认识华为云DWS (下)



目录

- 一 . 数据分析发展趋势与挑战
- 二 . 华为数据仓库解决方案介绍
- 三 . 华为数据仓库优势
- 四 . 商业成功案例

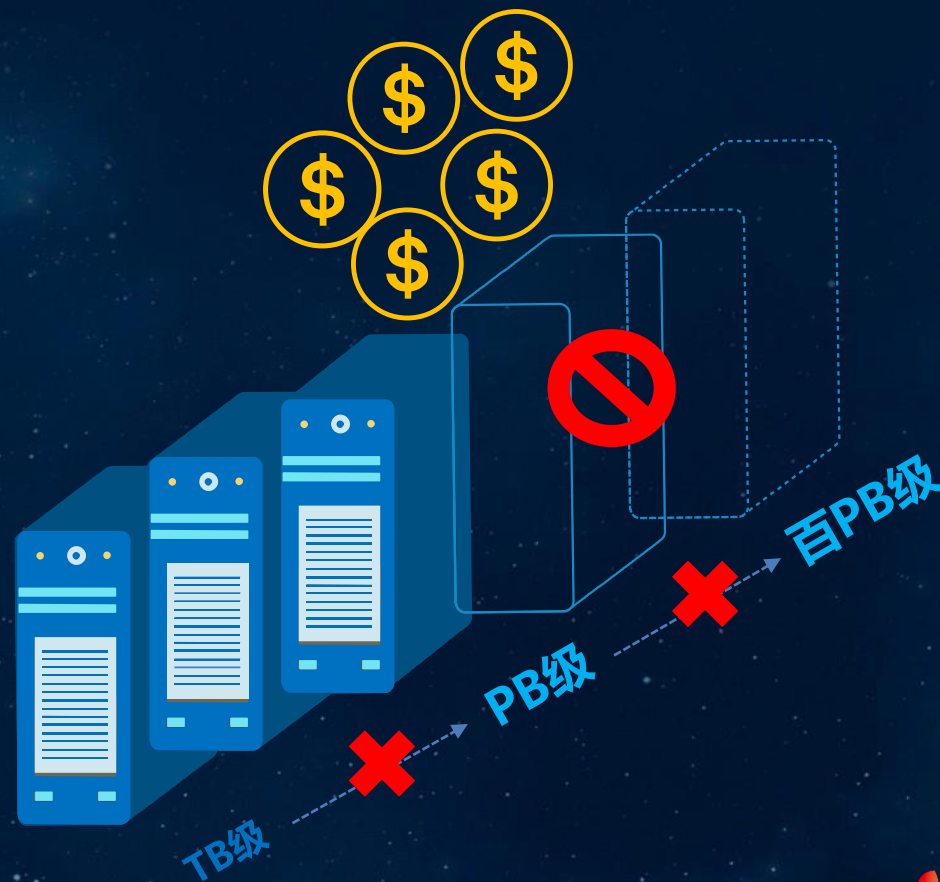
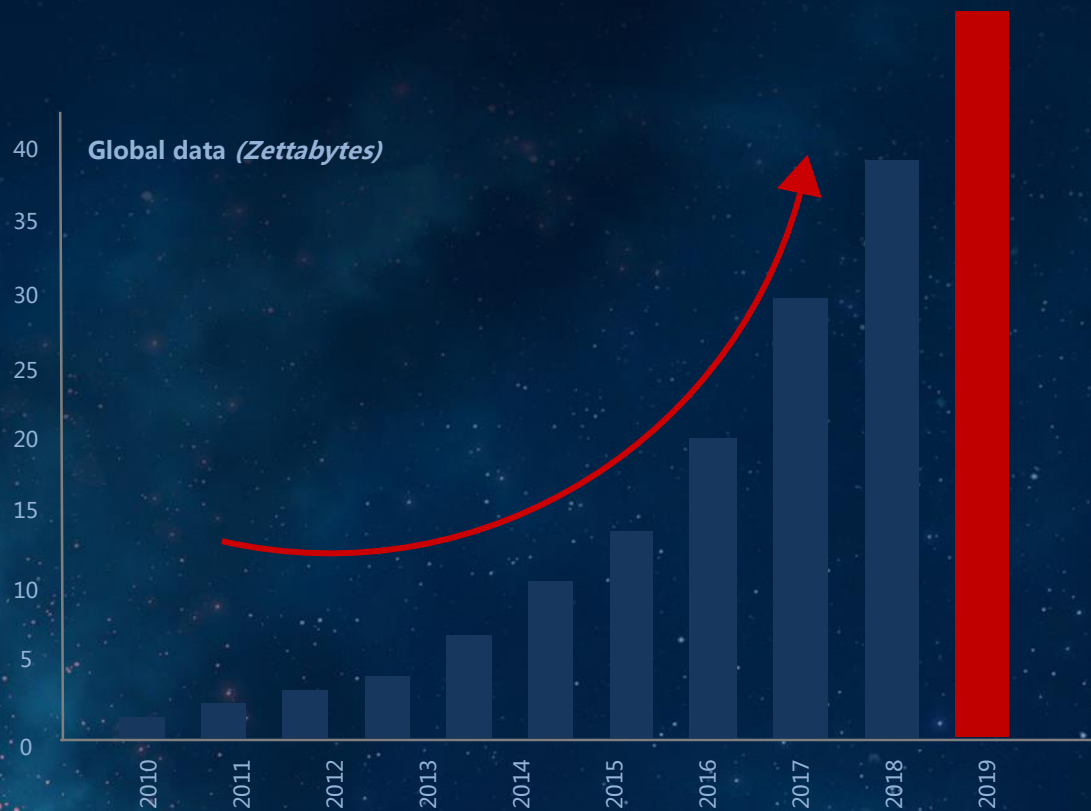
企业IT向云化快速推进，云数据仓库助力数据价值变现



挑战一：海量数据查询分析性能低，难以支撑业务敏捷分析决策

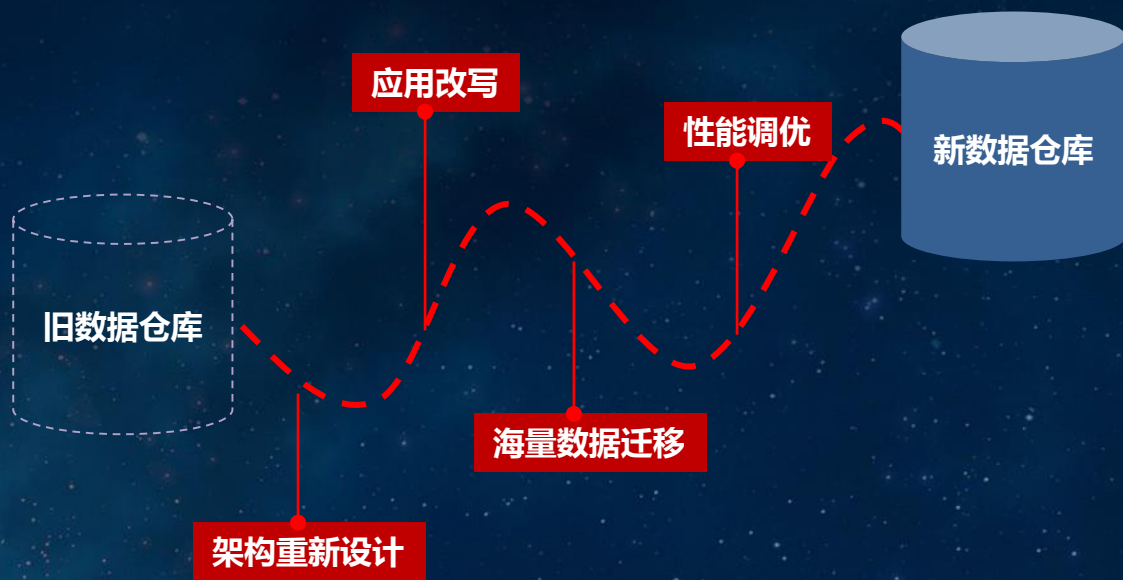
互联网时代，数据爆炸式增长
根据摩尔定理，数据总量大约每24个月**翻一倍**
企业数据，从GB到TB，从TB到**PB~百PB**

原有数据分析平台的**低性能、低扩展性、高成本**，成为挖掘数据金矿的阻力



挑战二：替换原有数仓“伤筋动骨”，新建数仓耗时数月“投入巨大”

旧数据仓库**替换，工作繁琐**



新建数仓耗时数月，花费巨大



业务数据上云成为必然趋势

【自建数据仓库】

- 购买并安装服务器、系统、数据仓库等软硬件
- 租用机房，费用高昂
- 招聘专业DBA,运维人员

应用优化
数据仓库弹性伸缩
数据仓库高可用
数据备份与恢复
数据仓库软件版本升级或打补丁
数据仓库软件安装
操作系统版本升级或打补丁
操作系统安装
服务器部署及维护
机架堆叠
机房、电力、空调、网络等基础设施

v.s.

【数据仓库 on 云主机】

- 购买并安装数据仓库软件
- 租用云主机
- 招聘专业DBA,运维人员

应用优化
数据仓库弹性伸缩
数据仓库高可用
数据备份与恢复
数据仓库软件版本升级或打补丁
数据仓库软件安装
操作系统版本升级或打补丁
操作系统安装
服务器部署及维护
机架堆叠
机房、电力、空调、网络等基础设施

v.s.

【DWS】

- 无需购买和安装任何软硬件
- 按需随时租用DDS
- 无需招聘DBA,运维人员

应用优化
数据仓库弹性伸缩
数据仓库高可用
数据备份与恢复
数据仓库软件版本升级或打补丁
数据仓库软件安装
操作系统版本升级或打补丁
操作系统安装
服务器部署及维护
机架堆叠
机房、电力、空调、网络等基础设施

■ 您自行解决

■ 云服务提供商解决

目录

- 一 . 数据分析发展趋势与挑战
- 二 . 华为数据仓库解决方案介绍
- 三 . 华为数据仓库优势
- 四 . 商业成功案例

DWS：实时、简单、安全可信的企业级数据仓库



数据实时洞见

- 支持流式数据**实时入库**、业务数据准**实时同步**
- 数据**入库即可查**，零等待
- 万亿数据查询分析**毫秒级**响应

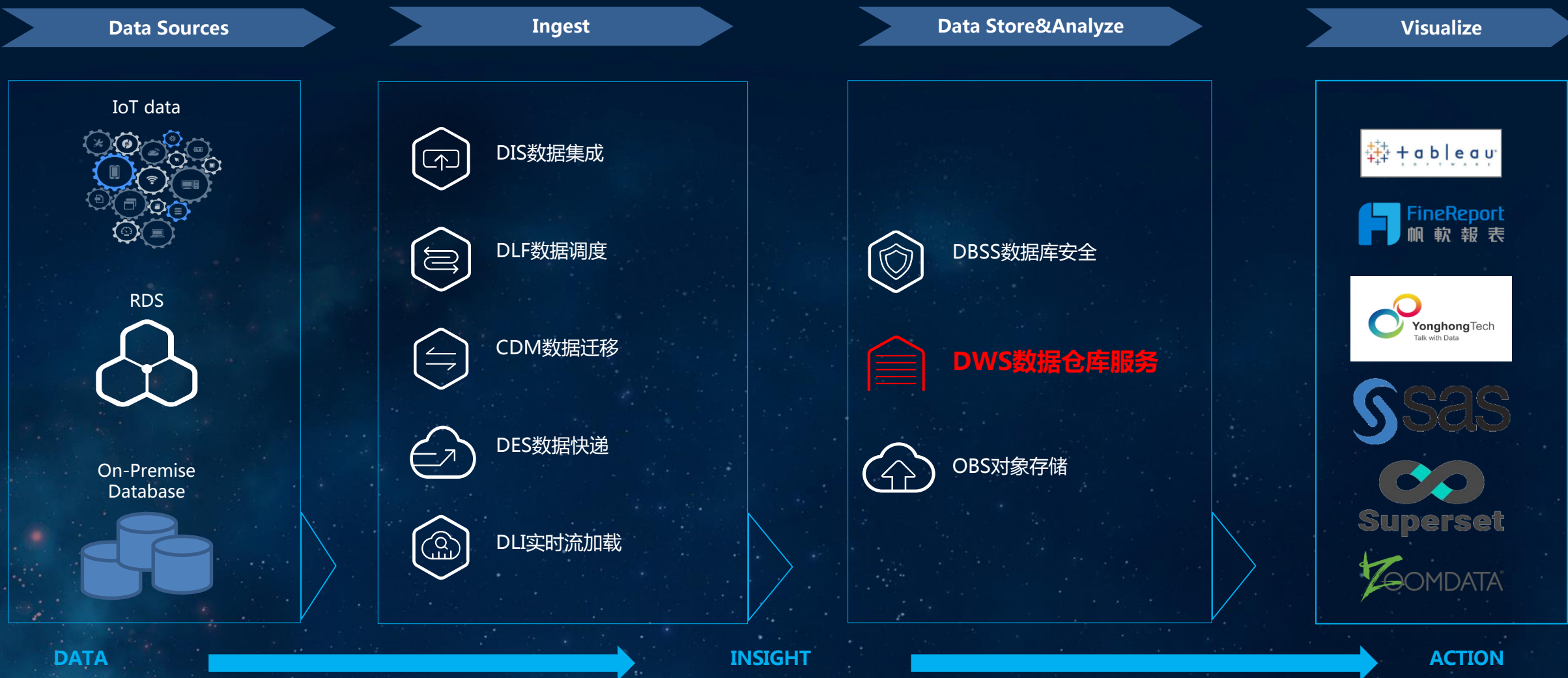
极简易用

- 数据迁移工具，最大化降低**TTM**
- 兼容标准**SQL 2003**，内置丰富OLAP函数
- TPC-H、TPC-DS真正**100%**支持

企业级、安全可信

- 支持分布式事务**ACID**，数据强一致保证
- 满足史上最严安全合规要求**GDPR**
- **业界唯一**数据库防火墙服务

华为云围绕DWS提供数据全生命周期解决方案



华为数仓聚焦两类水平解决方案

行业



金融



教育



零售电商



医药健康



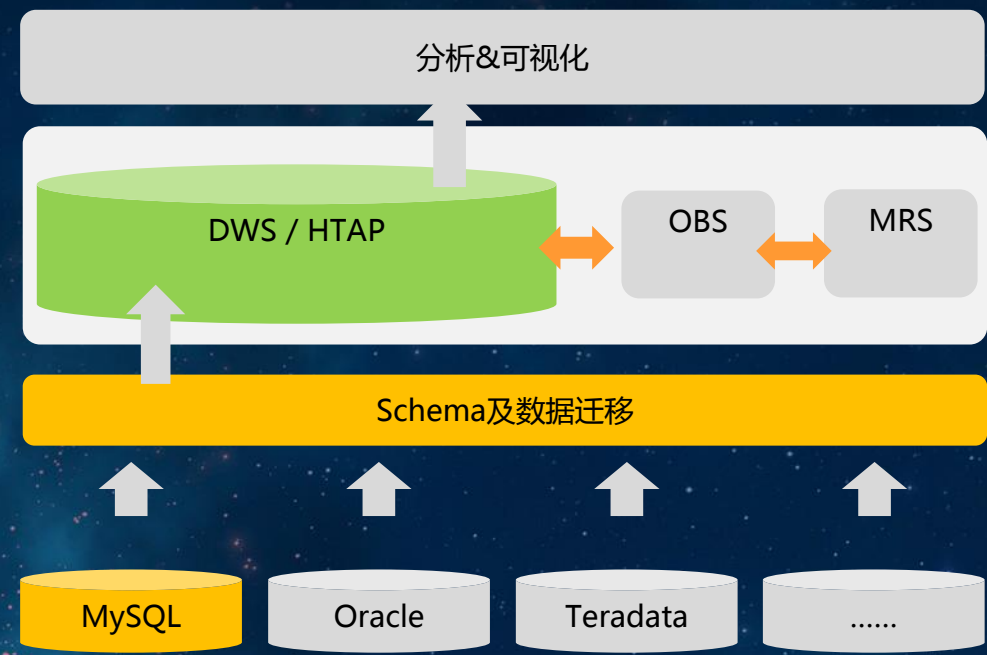
IoT与车辆网



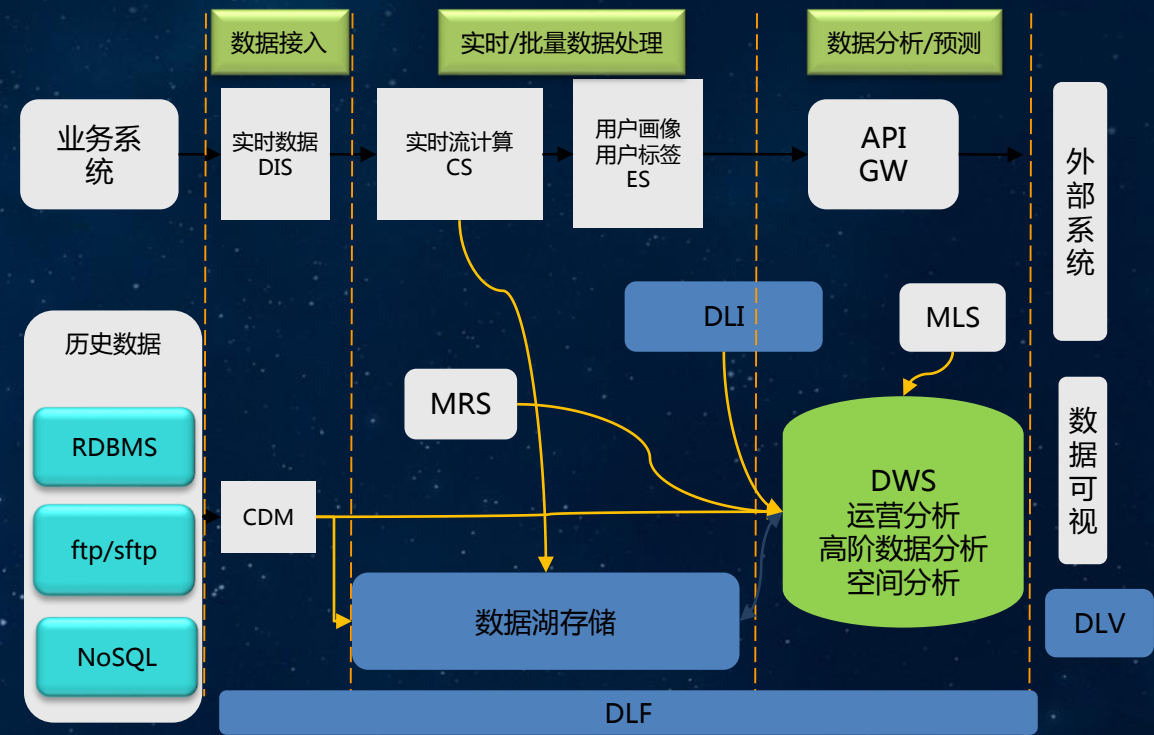
园区智能体

方案

传统数据库搬迁（替代Oracle等国外数据库厂商）



大数据平台解决方案（数据湖生态）



DWS具备强大的生态系统

数据集成工具



普元信息



保障信息安全 促进信息共享



REPLICATE



商业智能工具



帆軟報表



Talk with Data



开发工具

Data Studio

SQL Workbench/J



pgAdmin3



Navicat™



目录

- 一 . 数据分析发展趋势与挑战
- 二 . 华为数据仓库解决方案介绍
- 三 . 华为数据仓库优势
- 四 . 商业成功案例

连续两年进入Gartner魔力四象限，国产数据仓库第一



2017



2018

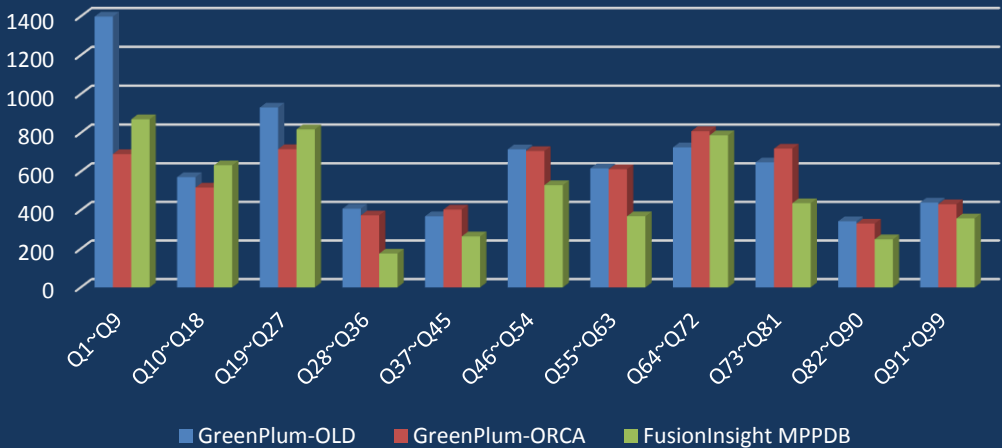


极致的性能与开放的SQL兼容性

测试项	TPCH-1000X 性能对比测试结果	
	DWS	阿里云HybridDB
	分区单并发	分区单并发
Q1	109.768	952.75
Q2	15.195	83.85
Q3	110.497	188.29
Q4	124.056	140.82
Q5	72.568	869.96
Q6	24.423	30.01
Q7	90.509	782.55
Q8	55.02	273.26
Q9	324.853	1022.43
Q10	75.277	195.45
Q11	10.759	50.19
Q12	87.064	81.23
Q13	53.739	189.7
Q14	29.26	30.51
Q15	69.479	79.07
Q16	16.948	115.37
Q17	320.243	2202.23
Q18	182.895	787.94
Q19	62.225	165.35
Q20	69.622	978.34
Q21	283.528	891.34
Q22	51.422	124.69
SUM	2239.35	10235.33

针对GreenPlum原生的优化器，以及改进后的ORCA优化器进行了对比测试。
DWS可以不修改SQL一次性跑通TPC-DS，且整体性能优于GreenPlum。

TPC-DS 1000X, DWS vs GreenPlum



机器	规格	CPU	内存	网络	磁盘	操作系统
4台本地盘规格VM	D1.8xlarge	2*10*Intel(R) Xeon(R) CPU E5-2680 v2 @ 2.80GHz	256GB	10GE	26*600G SAS 系统盘：2块600GB，RAID1 数据盘：24块1.8TB SAS本地盘	SLE 11 SP2

SQL无需修改，性能优于阿里云HybridDB 2-5倍，优于开源Greenplum



DWS核心技术优势

01

性能

- 全并行极致性能的MPPDB
- 行列混存及向量化计算
- 极速并行Bulk Load工具-GDS

02

可靠性

- 多层次冗余实现系统无单点故障
- CN多活，更高的并发度、可靠性、扩展性

03

扩展性

- 开放架构，按需水平扩展，容量和性能线性增长；
- 表级别在线扩容技术，保障扩容期间业务不中断、无感知；

04

易用性

- 一站式可视化集群管理，运维便捷；
- SQL On OBS，实现互联互通，且兼容标准SQL
- 应用开发和迁移工具列表

05

易管理

- 全面集群监控与告警、自动增量备份
- 快速升级，降低版本升级对业务的影响度；

DWS支持数据实时入库，T+0业务敏捷分析决策

数据实时入库，入库即可查，零等待！

业界首创端到端全并行引擎，万亿数据毫秒级响应



以某市交通套牌车分析为例，6000w条记录，全市1000智能卡口，查询分析计算量大： $6000w \times 6000w \times 1000 \times 1000$



过车记录表



车辆信息表

...



卡口配置信息表

在数据量达到千万级别时，传统数据库的查询响应时间极速增大



极简易用，助力企业聚焦业务

新建数仓

- 10min部署，即开即用
- 标准SQL 2003，TPC-DS全面兼容
- 数据透明切分，无需应用改写

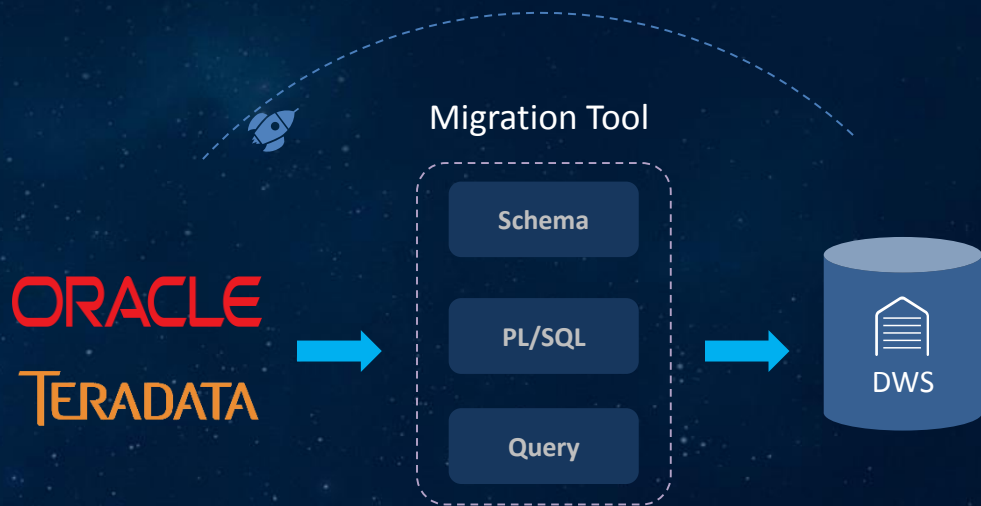


支持：

- 标准ANSI SQL92、SQL2003
- 标准开发接口JDBC、ODBC
- 支持事务和存储过程
- TPC、TPC-DS真正100%支持

旧数仓替换

- 提供数据迁移工具，海量数据快速迁移
- 独有Oracle、Teradata兼容模式



“宇宙第一大行” 2万+ EDW作业，人工投入数百人天，使用数据迁移工具仅需xx人天，效率提升**十倍**。

安全可信，满足“史上最严”安全合规要求

DWS具备全方位安全策略

配合华为独有数据库防火墙服务，构建超级安全壁垒



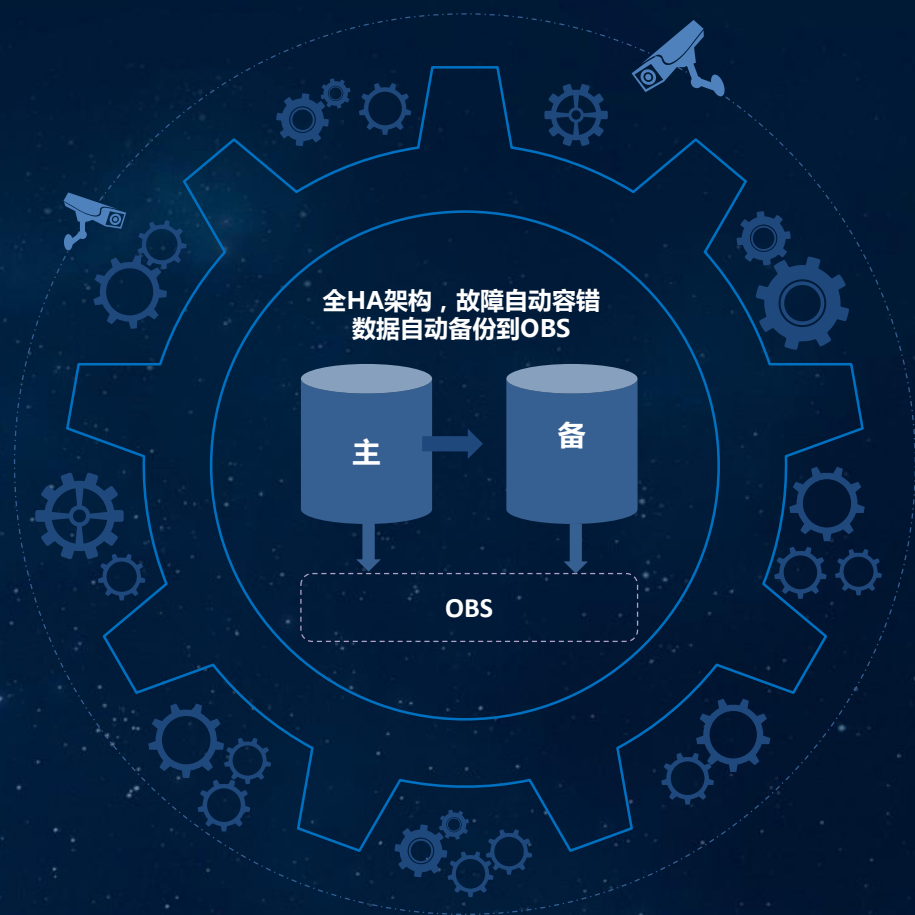
- 2018年5月25日，“史上最严数据保护条例”欧盟GDPR正式生效；
- DWS在欧洲DT、FT等公有云提供数仓服务，完全满足GDPR安全合规要求！

企业级保障，无后顾之忧

数据强一致性保证，确保数据分析无毫厘偏差

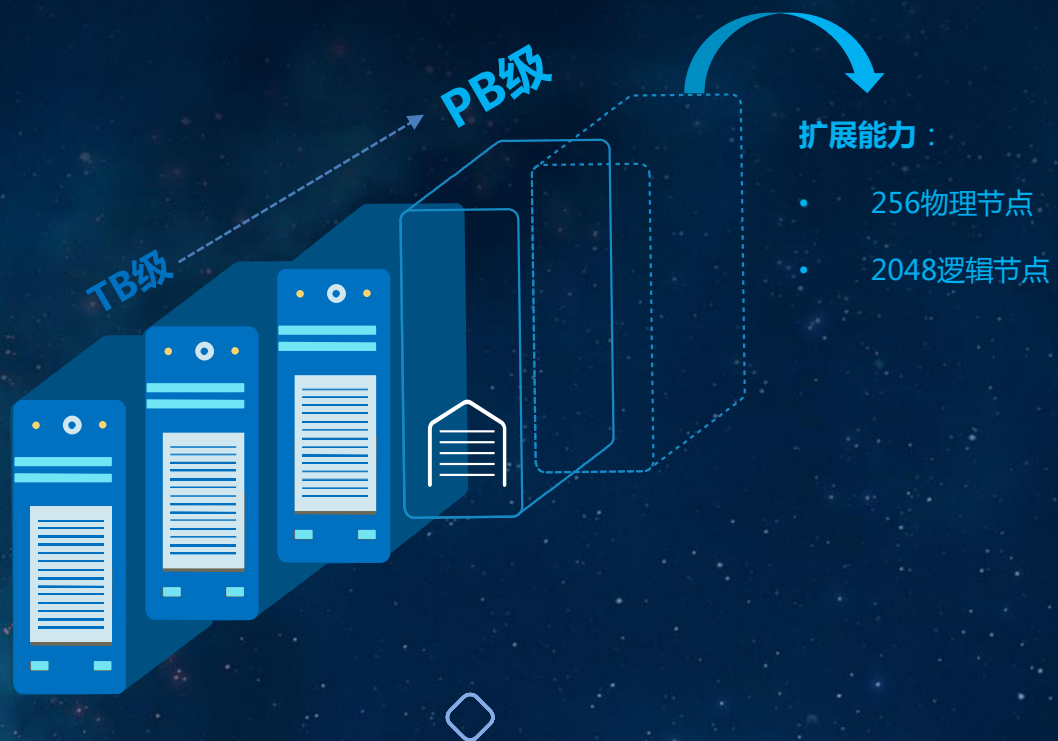


全自动健康监控，故障自动替换

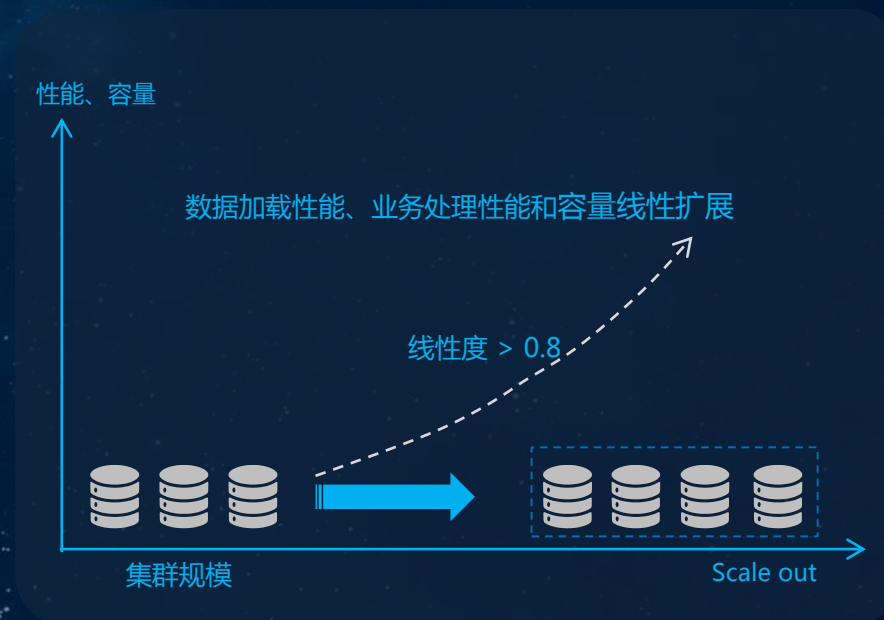


面对数据爆炸式增长，按需扩展，从容面对

横向、纵向按需扩展，业务0宕机



性能线性提升，投资回报成正比



目录

- 一 . 数据分析发展趋势与挑战
- 二 . 华为数据仓库解决方案介绍
- 三 . 华为数据仓库优势
- 四 . 商业成功案例

众多行业大客户的选择

金融



中国工商银行



深圳证券交易所



上海证券交易所



太平洋保险



深圳农村商业银行



中原银行



Hong Leong Bank



上海农商银行

政府

16+省级：

北京、重庆、上海、江苏、广东、新疆、贵州、阿根廷Salta省...

100+地级市：

深圳、广州、厦门、西安、无锡、苏州、青岛、南京、中山、宜昌、天津、石家庄、扬州、连云港...

电信



移动通信专家



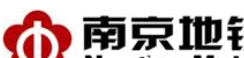
企业



国家电网公司
STATE GRID
CORPORATION OF CHINA



中国石油



南京地铁
Nanjing Metro



中国铁道科学研究院
CHINA ACADEMY OF RAILWAY SCIENCES
铁道建筑研究所



HUAWEI

打造“宇宙第一人行” PB级金融数据仓库，性价比提升xx%

- 依托华为分布式数据仓库，完成传统一体机Teradata全面替代，G行金融数仓正式迈入云化新时代
- 华为数据仓库集群规模总计330节点，总数据量~1PB



解决方案特点

- 高性能：列存储+向量化+全并行执行，为SQL交互查询提供极致性能，万亿级数据关联分析秒级响应。
- 融合数仓：统一SQL查询访问数据仓库和HDFS数据。
- 大规模，单集群120+节点

客户价值

- 战略意义：助力宇宙第一人行完成云化分布式改造
- 性价比提升xx倍
- 业务快速迁移：依托TD、Oracle迁移工具，迁移代价由人工xx人月缩减至xx人天，TTM效率提升xx倍

DWS Express : 助力华为终端云冷热数据关联分析，性能提升10倍



业务痛点：

- ◆ 探索查询HDFS 10PB级历史数据，耗时平均约1小时，全量扫描耗资源
- ◆ 业务系统存储3个月热数据，3个月至2年历史数据存储于HDFS，现有系统对热数据和历史数据无法进行关联分析

解决方案：

- ◆ DWS支持直接访问HDFS数据
- ◆ DWS提供标准SQL能力与Tableau无缝兼容，助力商业洞察
- ◆ 资源负载管理，即时业务性能保障

客户价值：

- ◆ OBS海量历史数据在线查询，免数据搬迁
- ◆ 冷热数据关联分析，小时级→分钟级，分析效率提升10倍
- ◆ 节省2/3报表开发人力，开发周期从每张1~2周减少至0.5小时

JOIN US IN
BUILDING A BETTER CONNECTED WORLD

THANK YOU

Copyright©2014 Huawei Technologies Co., Ltd. All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.

