

阿里云数智中台构建产业数字化转型

汤祯捷

阿里云计算平台产品解决方案总监

2023/03/03

目 录

01

阿里巴巴数据平台发展之路

02

大数据AI平台能力体系分享

03

平台最佳实践和沉淀的优势

数字化洞察

实现物理世界到数字世界的映射



数字化优化

通过数据分析优化业务提升绩效



数字化重塑

改变企业为客户创造价值的方式



数字化变革

创新数字化产品，产生商业模式变革



数字化是基础，数智化体系是引擎
业务价值是目的，技术应用是手段

“赢得竞争以生存”是本质

不断攻克技术难点

自研ODPS：突破性能瓶颈，构建多租户安全体系
平衡技术领先性和成本平衡

一体化自研大数据平台ODPS

一体化融合架构、可扩展技术引擎、
性价比全球领先、金融级安全稳定
提供MaxCompute、Hologres等自研引擎服务

持续降低数据应用门槛

统一调度、开发、治理的可视化界面
让集团1/3人员高质高效使用数据

统一多引擎开发界面

全链路数据治理与开发，全面提升数据质量与效率
通过统一可视化界面的DataWorks产品对外提供服务

阿里云大数据智能平台向4个一体化+1个平台架构演进

大数据AI行业场景应用

互联网 新金融 新零售 制造业 能源 汽车 公共安全 教育 医保卫健 检法司 运营商 ...

开放生态：全数据生命周期链路的专业服务和方

数字化咨询服务 >>> 行业方案设计 >>> 行业数据建模 >>> 数据应用开发 >>> 数字化运营服务

阿里云4+1大数据AI能力：数据智能新基石

DataWorks 数据开发与治理
开发 | 建模 | 治理 | 资产 | ...

+

AI开发与服务
DSW | Designer | OpenMind

大数据 AI 一体

业务系统

数据库

IoT

数据集成

MaxCompute

实时离线一体

流计算

Hologres

自研数仓产品

实时计算Flink版

湖仓一体

流批一体

E-MapReduce

DLF

生态产品

OSS

数据湖产品

机器学习PAI EAS

机器学习PAI

ES / OS / AIRec

AI产品

目 录

01

阿里巴巴数据平台发展之路

02

大数据AI平台能力体系分享

03

平台最佳实践和沉淀的优势

ODPS (Open Data Platform and Service)

数据驱动而生的超大规模多场景融合的大数据智能计算平台，开放架构，面向未来可拓展



10万台

服务器并行计算

2.79EB

单日最大数据处理规模

6年

蝉联TPCx-BB世界冠军

新纪录

TPC-H 2022年全球性能

300+项

海内外专利

20+项

软件著作权



基于Pangu存储

- 计算存储分离架构
- 多副本高可靠底座
- 分层存储，性价比高



计算统一调度

- 基于容器技术
- 资源按需弹性
- 资源合池与混部



存储引擎直读

- 一份存储，多个引擎
- 直读IO提速10+倍
- 联邦查询，离线实时一体



统一元数据管理

- Meta集中管理
- 外表自动映射
- 血缘跨引擎打通



统一开发界面

- 基于DataWorks最佳实践
- 持全链路数据开发、治理

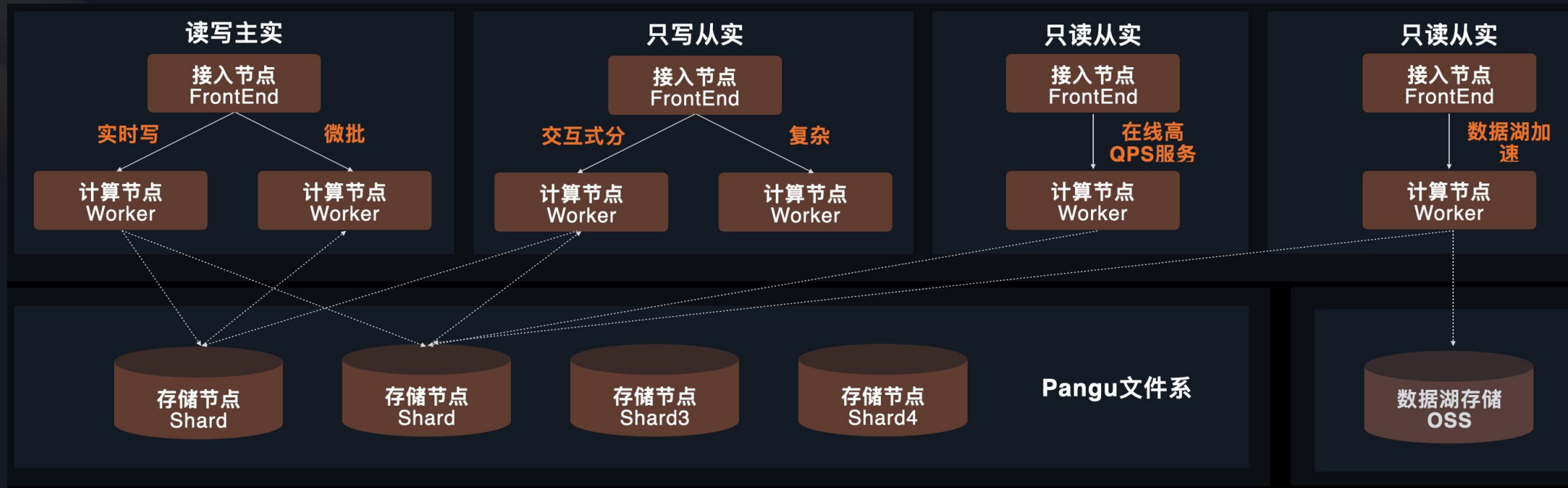


统一安全管理

- 企业级金融级安全能力
- 跨引擎加密解密
- 引擎间细粒度授权

Hologres统一数据服务层，分析服务一体化

一站式实时数据仓库引擎，支持海量数据实时写入、实时更新、实时分析，支持标准SQL（兼容PostgreSQL协议），支持PB级数据**多维分析**（OLAP）与**即席分析**（Ad Hoc），支持高并发低延迟的**在线数据服务**（Serving）



Rank	Company	System	QphH	Price/kQphH	Watts/KQphH	System Availability	Database
1	Alibaba	Alibaba Cloud Hologres	27,865,885	1.591.87 CNY	NR	10/31/22	Hologres V1.3
	Alibaba	Alibaba Cloud Hologres	22,664,825	84.95 USD	NR	06/01/21	EXASOL 7.1
			15,265,305	4,542.13 CNY	NR	07/31/21	OceanBase V3.2
4							
5							
6							

TPCH 30TB Top1 Performance

- TPCH 30TB性能评测中，超过原有世界记录23%+
- 向量化、异步化、智能化、云原生、行列共存，计算存储协同
- 用更少的资源，推动更高的性能

实时计算 Flink让实时处理更加智能化

实时计算Flink版 提供端到端亚秒级实时数据分析能力，并通过标准SQL降低业务开发门槛，助力企业向实时化、智能化大数据计算升级转型



从云原生走向云深处

DataWorks 一站式数据开发和治理工具



致力于构建企业数据仓库/数据湖/湖仓一体的数据平台架构，全流程工具提效，减少开发人员低效重复工作，保持螺旋式上升，加速企业数字化转型



目 录

01

阿里巴巴数据平台发展之路

02

大数据AI平台能力体系分享

03

平台最佳实践和沉淀的优势

在国家电网

总部及13个省份实现**一云多Region**，65%的电力经过阿里云计算

上云应用

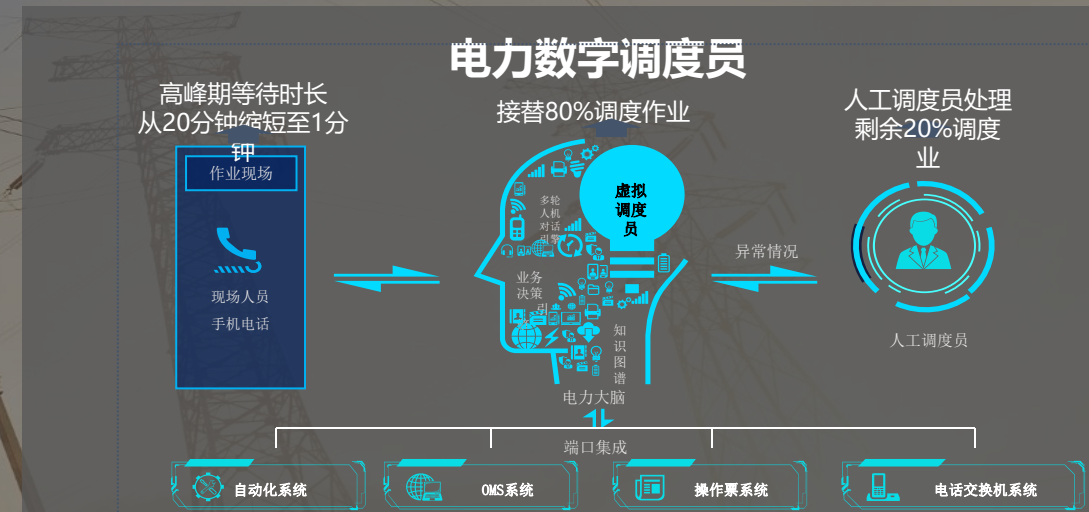
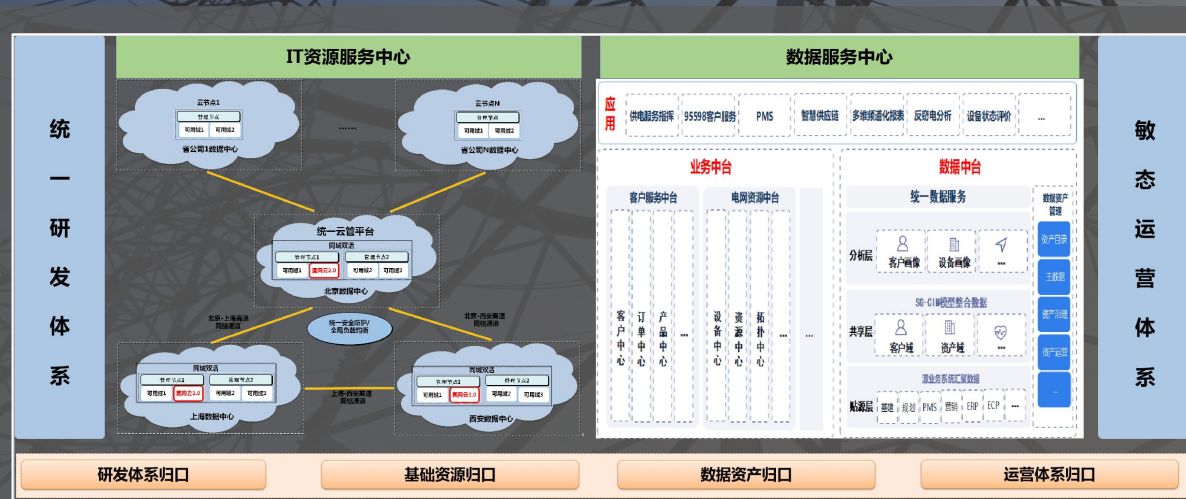
217个

数据中台

110万张表

AI调度提效

2000万/年



在长城汽车

通过搭建企业**数据中台**，整合53个领域的业务数据，实现秒级数据分析决策能力

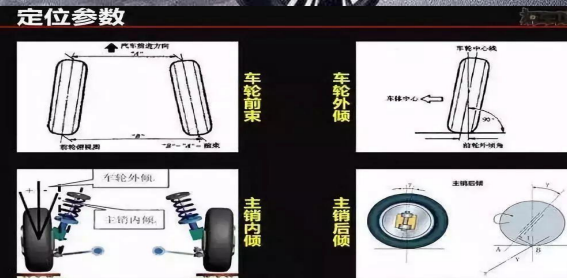
构建企业全域数据架构



基于数字化闭环的生产制造流程改进



- 基于产线生产数据的统计过程分析
- 底盘参数设定推荐最佳工艺参数
- 供应商评价及甄选



统一整合拉通各环节数据

总结：阿里云大数据智能平台与行业数字化改革共发展，业界认可



最大规模实践

- 单集群规模离线数仓》**3万节点**；实时数仓》**8192节点**
- 集团内案例**10+万台物理服务器节点**(跨域计算)
- IDC：大数据平台**领导者**；中国数据治理平台**市场分额第一**



最先进技术引领

- **一个平台+四个一体化**：一站式数据开发和治理平台；实时离线一体、流批一体、湖仓一体、大数据AI一体化
- 信通院：数据平台整体解决方案评测达到**先进级(3级)**



最受国际化认可

- 中国唯一进入 Forrester **强劲表现者**象限的大数据产品
- 中国唯一进入Gartner**领导者**象限的大数据产品
- AI能力进入Forrester**领导者**象限
- TPCx-Bigbench测试，100T和30T连续**五年全球双料冠军**



最全安全策略

- 全栈数据安全体系
- 数据全生命周期的安全保障
- **多租户**隔离、沙箱机制
- 国家等保**2.0四级**
- 领先的**脱敏和隐私计算技术**

