



2019

05

08-10

北京新云南皇冠假日酒店

# 数据风云 十年变迁

DTCC

第十届中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2019



+

○

○

○

# 中国农业银行 数据库使用实践和发展规划

研发中心 高级专家 蔡仕志



数据库使用实践



数据管理体系建设



数据管理典型案例



数据库发展规划

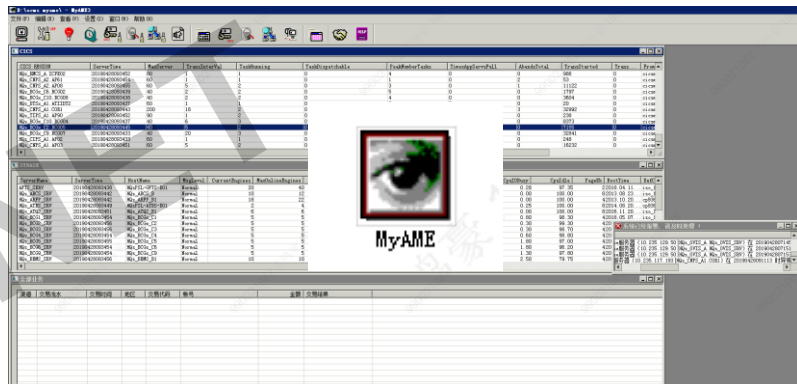
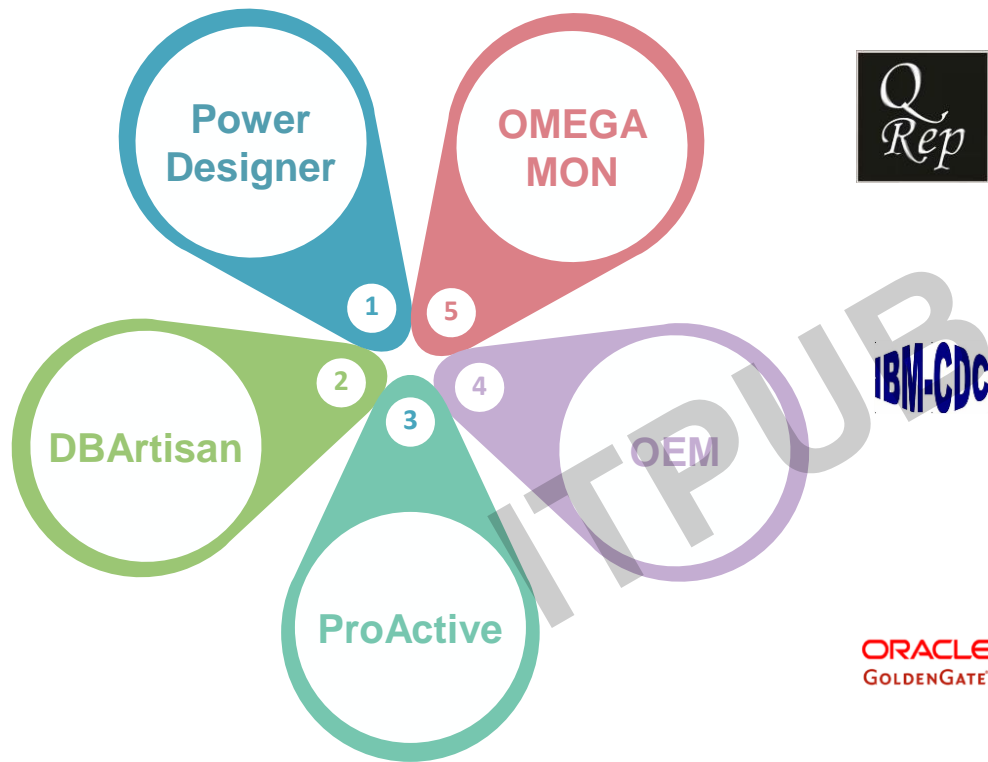
# 1.1 农行省域及数据大集中



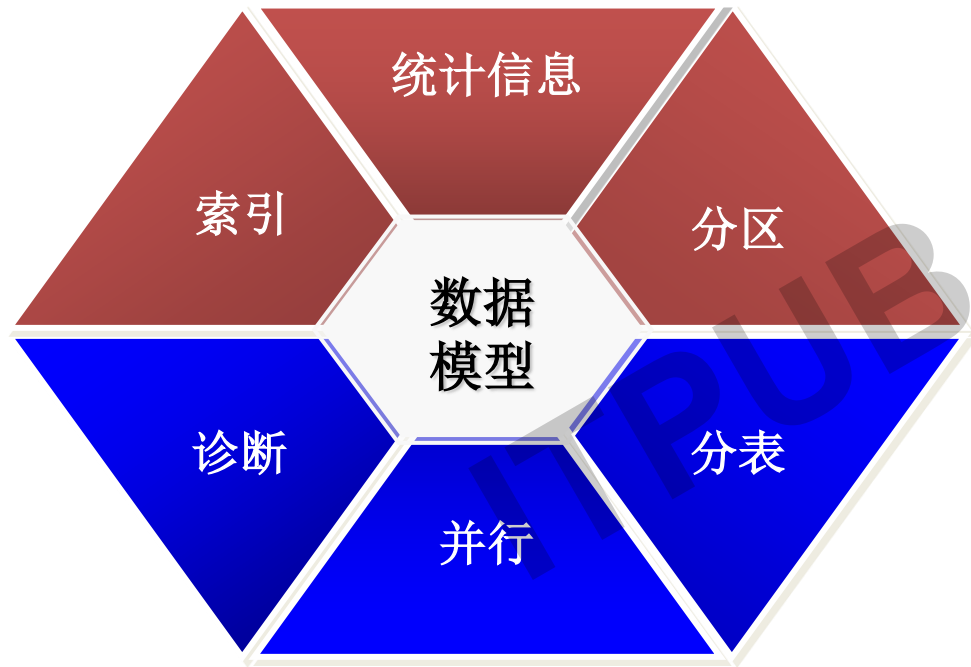
# 1.2 农行数据库产品总览



# 1.3 关系数据库工具箱



# 1.4 关系数据库经验谈——OLTP



SQL

- 事务颗粒度
- SQL组合

运维

- 配置规范
- 数据库整合



# 1.4 分析型数据库经验谈——OLAP

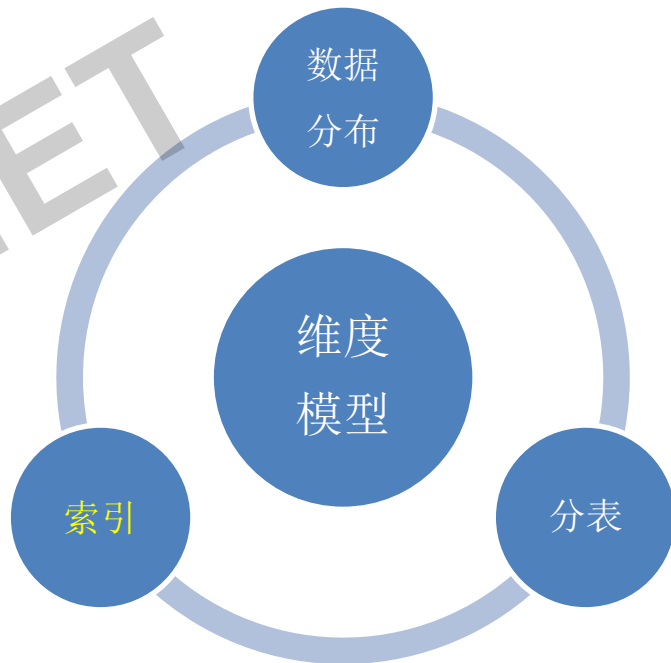
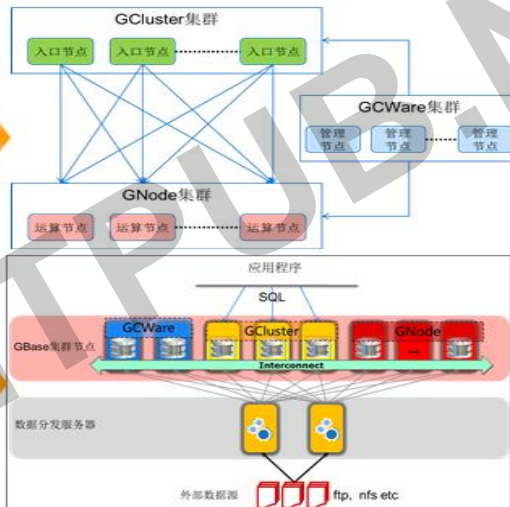
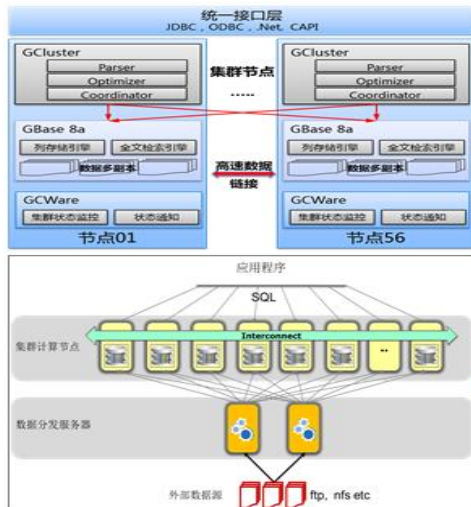
GBASE®

列存储

MPP架构

无Master架构：最多支持64节点

联邦架构：最多支持300节点







数据库使用实践



数据管理体系建设



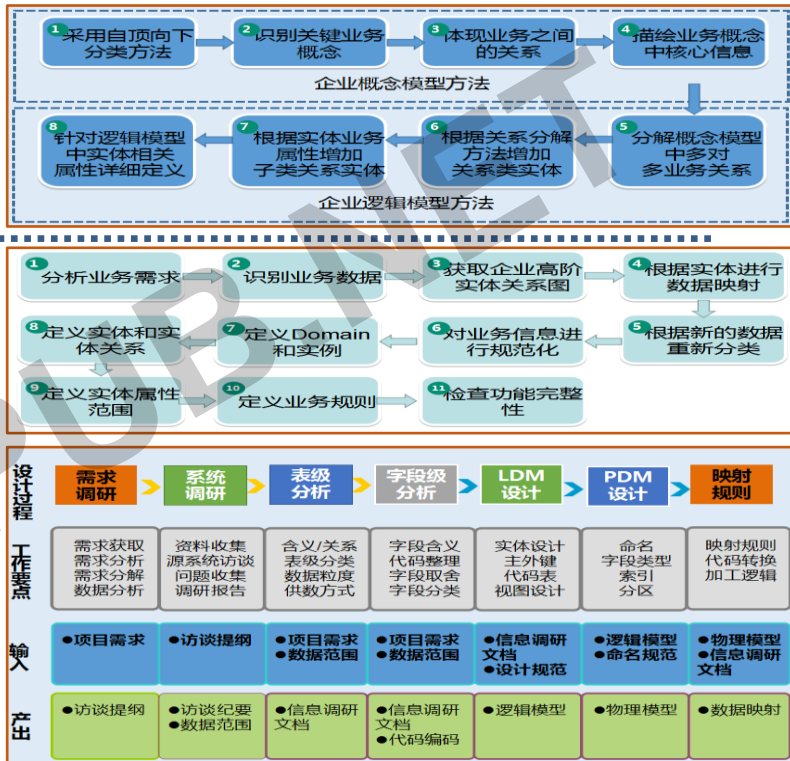
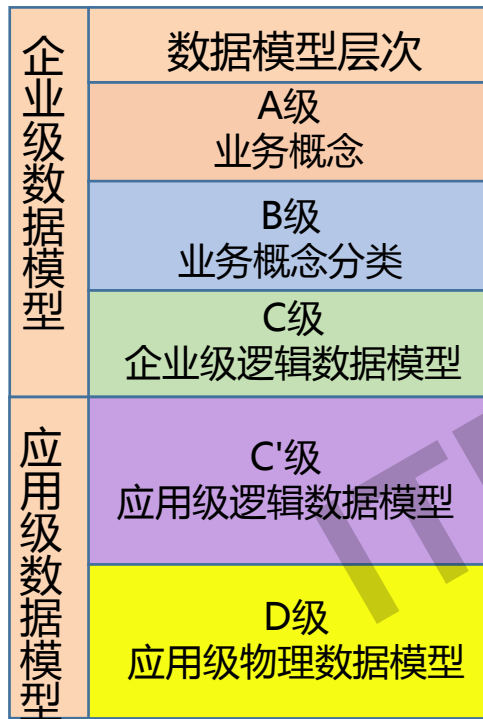
数据管理典型案例



数据库发展规划

## 2.1 企业级数据模型

数据模型管理



企业级 OLTP OLAP

数据模型设计方法

## 2.1 数据模型规范

## 主题域

主题域	主题域中文名称规范	主题域英文名称规范
参与人	参与人	IP_业务含义英文缩写_ERD
合约	合约	AR_业务含义英文缩写_ERD
产品工厂	产品工厂	PD_业务含义英文缩写_ERD
结算	结算	ST_业务含义英文缩写_ERD
核算	核算	AT_业务含义英文缩写_ERD
存款	存款	DP_业务含义英文缩写_ERD
资产	资产	LP_业务含义英文缩写_ERD
服务	服务	SV_业务含义英文缩写_ERD
投资	投资	IF_业务含义英文缩写_ERD
公共服务	公共服务	PR_业务含义英文缩写_ERD
国际业务	国际业务	FB_业务含义英文缩写_ERD
市场营销	市场营销	MK_业务含义英文缩写_ERD
总控管理	总控管理	SC_业务含义英文缩写_ERD

## 实体特征

序号	实体特征名称	实体特征代码
1	基本信息实体	INFO
2	辅助实体	AUX
3	生命周期管理实体	LCS
4	计划实体	PLG
5	附加信息实体	AINF
6	衍生信息实体	DINF
7	历史信息实体	HIST
8	附加信息相关历史信息实体	AHIST
9	审计信息实体	AUDT
10	关系实体	RLN
11	计划相关附加信息	PLG_AINF
12	附加信息相关计划信息	AINF_PLG

## 属性分类

序号	分类词英文代码	分类词英文名称	分类词中文名称
1	COD	Code	代码类
2	ID	ID	ID标识类
3	IND	Identifier	状态标识类
4	NAM	Name	名称类
5	TXT	Text	描述类
6	VLU	Value	取值类
7	NUM	Number	编号类
8	DATE	Date	日期类
9	TIME	Time	时间类
10	AMT	Amount	金额类
11	BAL	Balance	余额类
12	CNT	Count	数量类

## 数据类型应用规范

序号	分类词英文代码	分类词英文名称	分类词中文名称	数据类型应用规范描述
1	COD	Code	代码类	Char, 原则上数据长度不得超过8位
2	ID	ID	ID标识类	Char, 原则上数据长度不得超过30位
3	IND	Identifier	状态标识类	Char, 原则上数据长度不得超过2位
4	NAM	Name	名称类	Char
5	TXT	Text	描述类	Char、Varchar
6	VLU	Value	取值类	Char、short integer、integer、decimal (含numeric、double、money)
7	NUM	Number	编号类	Char, 原则上数据长度不得超过30位
8	DATE	Date	日期类	Char (8), 对于YYYY、MM、DD、YYYYMM、MMDD等日期定义, 建议使用VLU分类词定义。
9	TIME	Time	时间类	Timestamp[26位]、Char (6) [HHMMSS]、Char (14) [YYYYMMDDHHMMSS]
10	AMT	Amount	金额类	Decimal (18, 2)
11	BAL	Balance	余额类	Decimal (18, 2)
12	CNT	Count	数量类	Short integer、Integer

## 2.2 数据管理制度和规范

### 制度规范

一阶：战略规划

《数据管理与  
服务规划》  
《数据集市建设规划》

二阶：规章制度

《数据管理基本制度》  
《数据质量管理办法》  
《数据信息分级分类指南》  
《数据安全实施实施细则》  
《数据模型设计与管理工作手册》  
《应用系统接口规范管理规程》

三阶：技术标准

《项目总体方案编制规范》  
《企业级数据标准》  
《应用系统数据模型设计技术手册》  
《数据库设计开发规范》  
《核心业务系统建表规范》  
《服务方接口规范》...

01

通过一体化的制度设计，规范了系统研发与运维全流程中的数据生产与消费

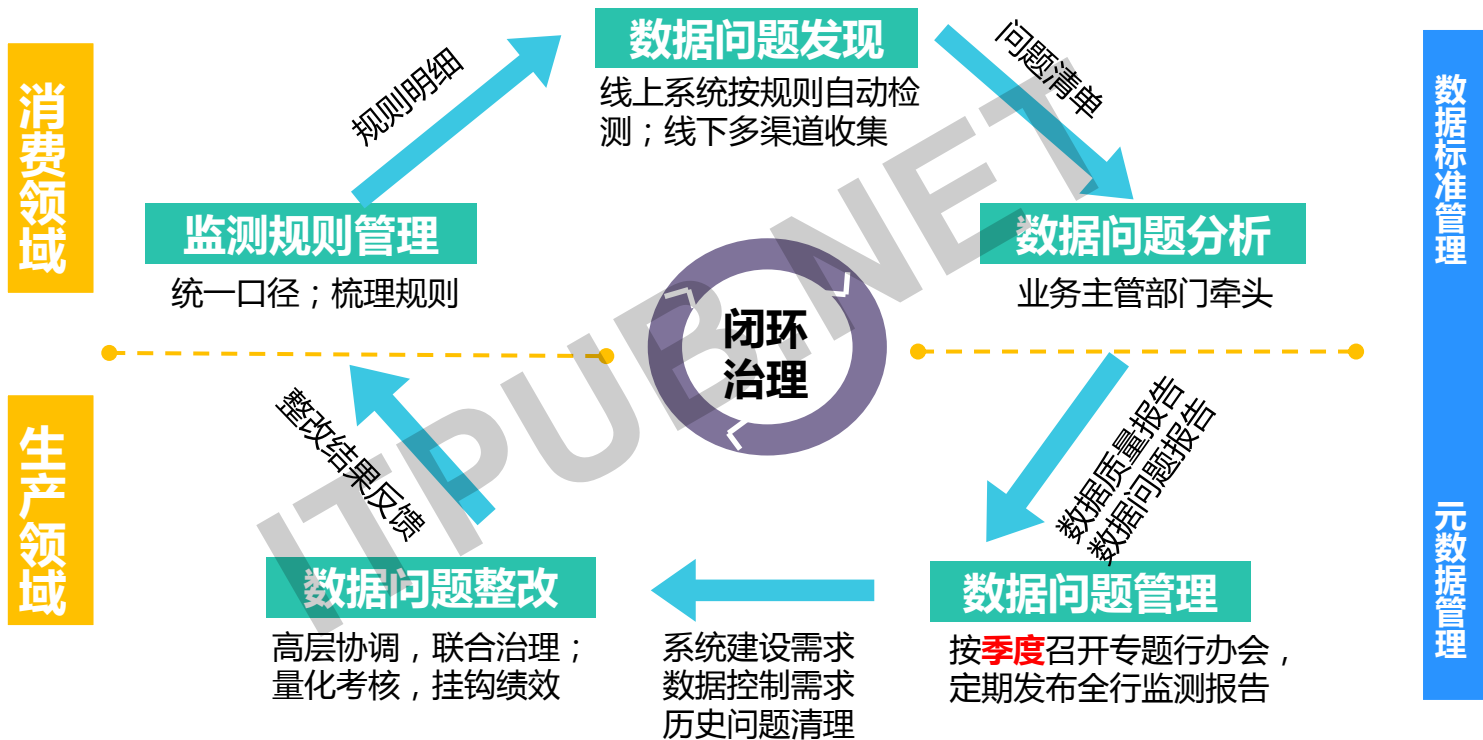
02

配以专业的数据管理团队、流程化的改进机制落实系统研发运维及数据全生命周期的架构管控

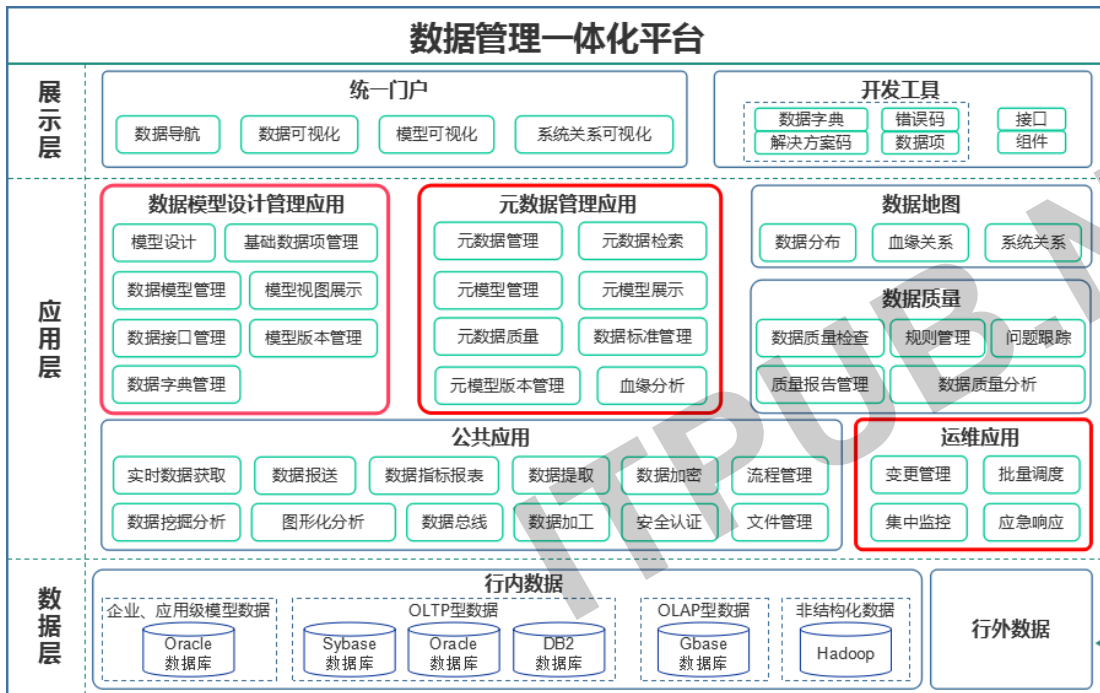
### 改进机制

- 数据架构管控机制
- 数据质量保障机制

## 2.3 数据质量保证机制



## 2.4 数据管理技术平台



### 流程线上化

数据逻辑设计管理：统一管理数据模型的资源 and 标准  
元数据管理：一致性分析、影响性分析  
运维应用：数据变更的流程化管理



### 数据可视化

数据地图：实现数据定位和多种形式展示  
数据质量：实现生产过程中问题数据的全流程管理

## 2.5 OLTP数据管理实践



### 数据模型方法研究

2008年 - 2011年10月  
研究FSDM、Inmon范式建模等，形成农业银行OLTP的11步建模方法和建模流程。



### BoEing数据模型实践

2009年 - 2015年10月  
成立BoEing数据模型组，统筹数据管理方面工作，完善数据管理各项规章制度，形成全行数据标准。打造数据管理平台，包括数据模型设计与管理系统、元数据管理系统、数据质量管理体系等。初步实现数据管理工作的线上化。



### 总结形成DOTA数据管理框架

2016年3月 - 2017年4月  
通过研究DAMA国际、IBM、同业实践等，总结形成DOTA数据管理框架，完善数据规范体系。整合数据管理各系统，打造研发运维数据管理一体化平台。在研发、运维、应用等方面均实现了数据的有效管理。



### 选取OLTP典型项目进行DOTA框架再实践

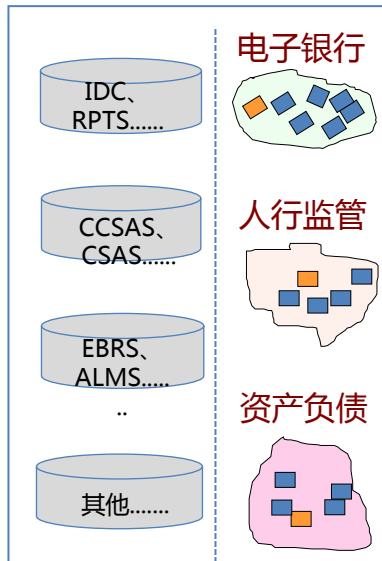
2017年2月 -  
选取OLTP典型项目，进行实践，从数据模型变更、数据模型设计、数据质量提升、数据标准落地等角度均进行了有效实践。结合数据规划，从数据源头着手，打通数据消费（OLAP）的通道。



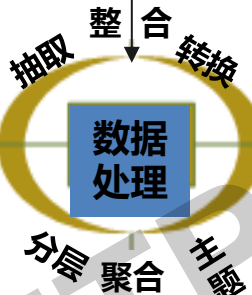
# 2.6 OLAP数据管理实践

孤岛 竖井式 多标准 共享差的OLAP系统

原OLAP系统



建模方法

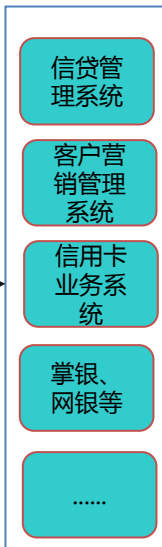


共享多层次OLAP系统



主题化 共享可复用 多层次OLAP系统

应用于OLTP



2012年1月 - 2012年10月

研究维度、范式建模理论，形成农行OLAP数据建模方法

2012年8月 - 2016年8月

在OLAP领域大数据平台工程中实践维度与范式结合的数据建模方法

2016年3月 - 2017年4月

总结融合OLTP和OLAP实践，贯穿研发与运维，研究形成DOTA框架

2017年2月 -

DOTA框架进一步在金融大脑工程再实践



数据库使用实践



数据管理体系建设



数据管理典型案例



数据库发展规划

## 3.1 核心系统

ITPUB.NET



## 3.1 核心系统下移成效



数据范围

按交易码分布，下移100多支交易  
按渠道分布，开放核心渠道接入50个  
按应用分布，核心查询下移应用12个

联机交易延迟0.75s，最大延迟4.6s  
批量时段延迟20s之内



数据复制



数据总量

主机50T，下移到开放12T  
总量4000多张表，下移400多张表  
下移约100亿条记录，最大表16亿

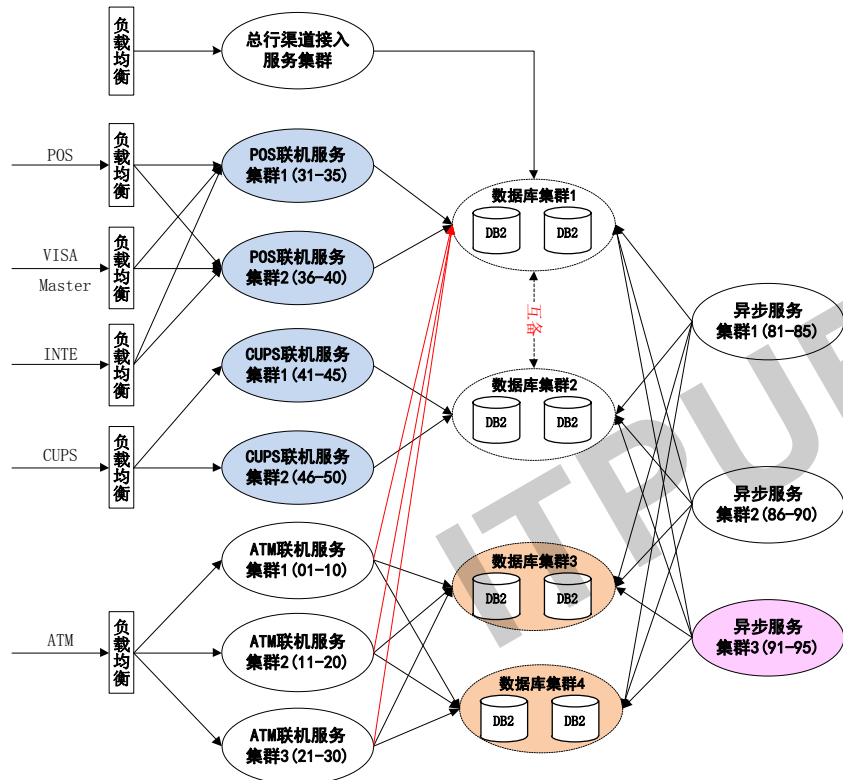
日均交易：主机3.6亿笔，开放平台  
2.3亿笔  
峰值交易：主机4.0亿笔，开放平台  
2.6亿笔



日均交易

时间	节约主机MIPS	节约财务支出
2018年底	3.2万MIPS	6.25亿
2019年底	1万MIPS	1.95亿

## 3.2 银行卡受理中心系统



高并发业务持续服务

1

数据分库

2

数据分表

3

短事务

4

多种切换方式

## 3.3 分布式缓存云



## 3.3 大数据实践

### 首家全面自主可控大数据体系

自主研发预处理平台、分析挖掘平台、流计算平台及ETL、监控、调度等大数据工具集

### MPP数据库+HADOOP技术

高可用、高性能混搭

全天候、全模式的数据服务

海量结构化和非结构化数据处理和存储

### “科研驱动+创新引擎”模式

专利（21个）+科技论文（8篇）

实现异构、异源、异库访问

人民银行科技发展奖一等奖

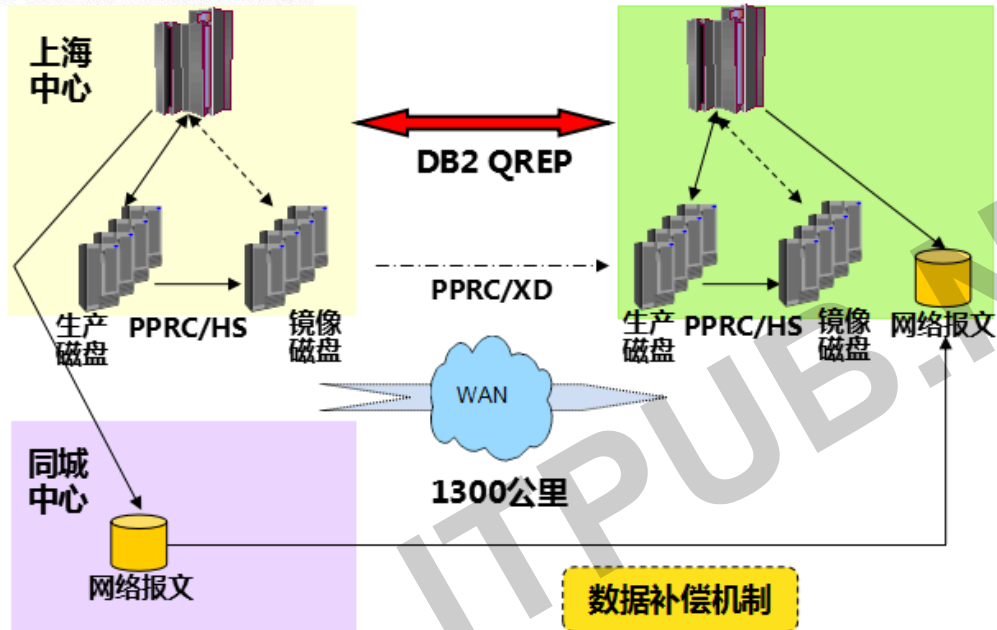
### 大数据+AI智能化服务框架

构建智能感知+智能服务的金融大脑

提供人脸识别、智能投顾、量化交易、智能推荐等数据服务



## 3.4 两地三中心建设—主机系统



技术架构  
GDPS/Active-Active sites

数据复制  
DB2 Qrep异步复制

运行模式  
Active / Query

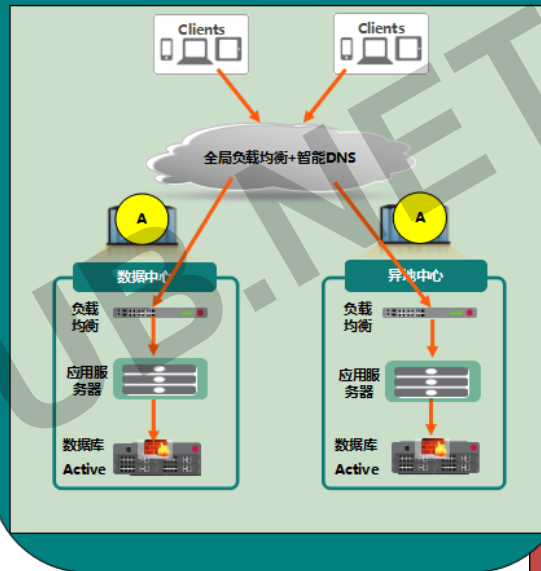
园区级灾难：RT0<5分钟（一键切换），RPO≈0分钟（数据补偿保障数据完整）

区域级灾难：RT0<5分钟（一键切换），RPO<5分钟（大量应用改造降低时延）

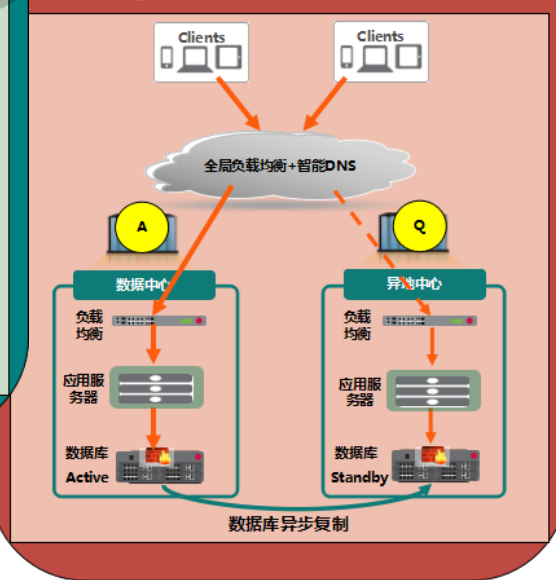
## 3.4 两地三中心建设—开放系统

- A-A架构：两中心同时提供全功能服务
- A-Q架构：主中心提供全功能服务，备中心提供查询服务，同城中心提供数据
- A-S架构：主中心提供全功能服务，备中心仅作应用和数据备份，不提供查询服务，同城中心提供数据

### A-A架构：无数据依赖的应用



### A-Q架构：严格数据依赖的应用



# 3.5 快捷支付系统

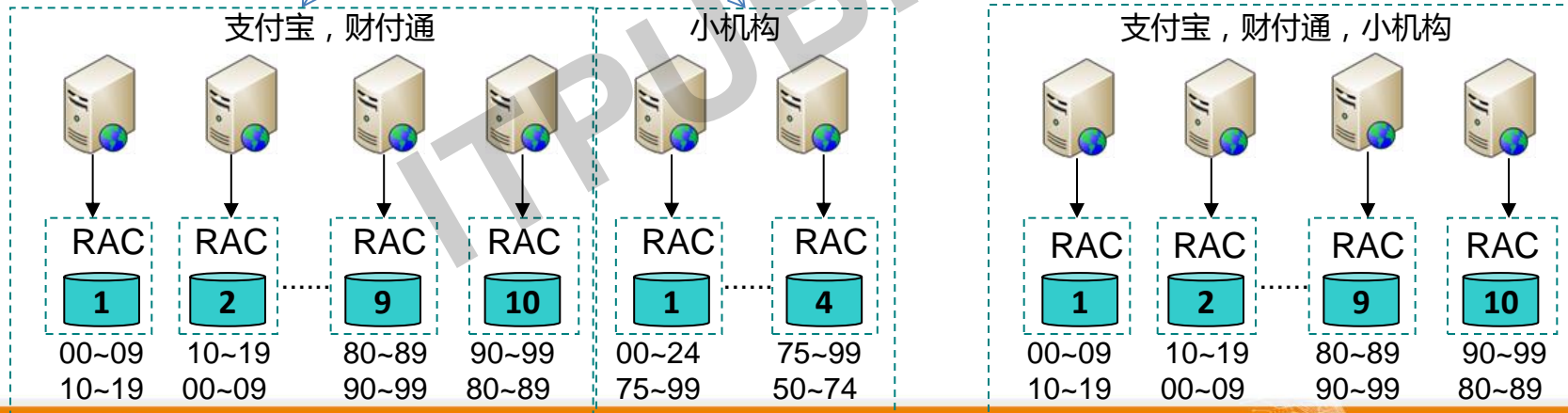
➤ A-S架构  
(北京中心忽略)



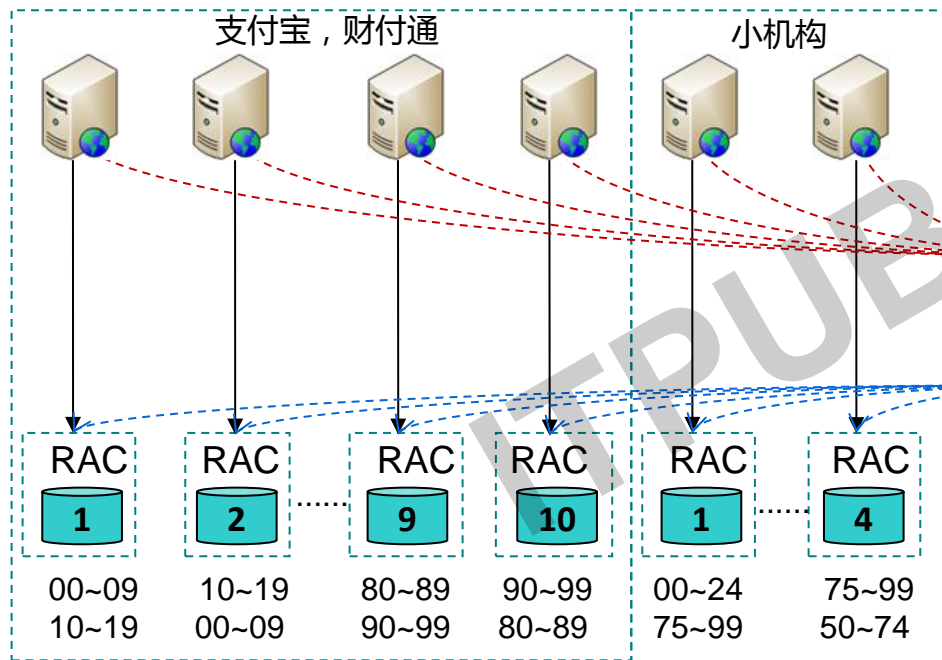
网联应用



银联应用



## 3.5 快捷支付系统



**1** 为了减少数据库访问压力，通过把静态数据及变化频率低的数据缓存到Redis中。对客户限额也进行缓存，应用对于限额的访问和回写都只访问Redis，缩短交易耗时。

**2** 为进一步提升系统高可用性，计划新增Redis同步功能，日终将月限额写回至数据库，实现限额数据持久化。当Redis出现故障时，通过访问数据库实现限额控制。



数据库使用实践



数据管理体系建设

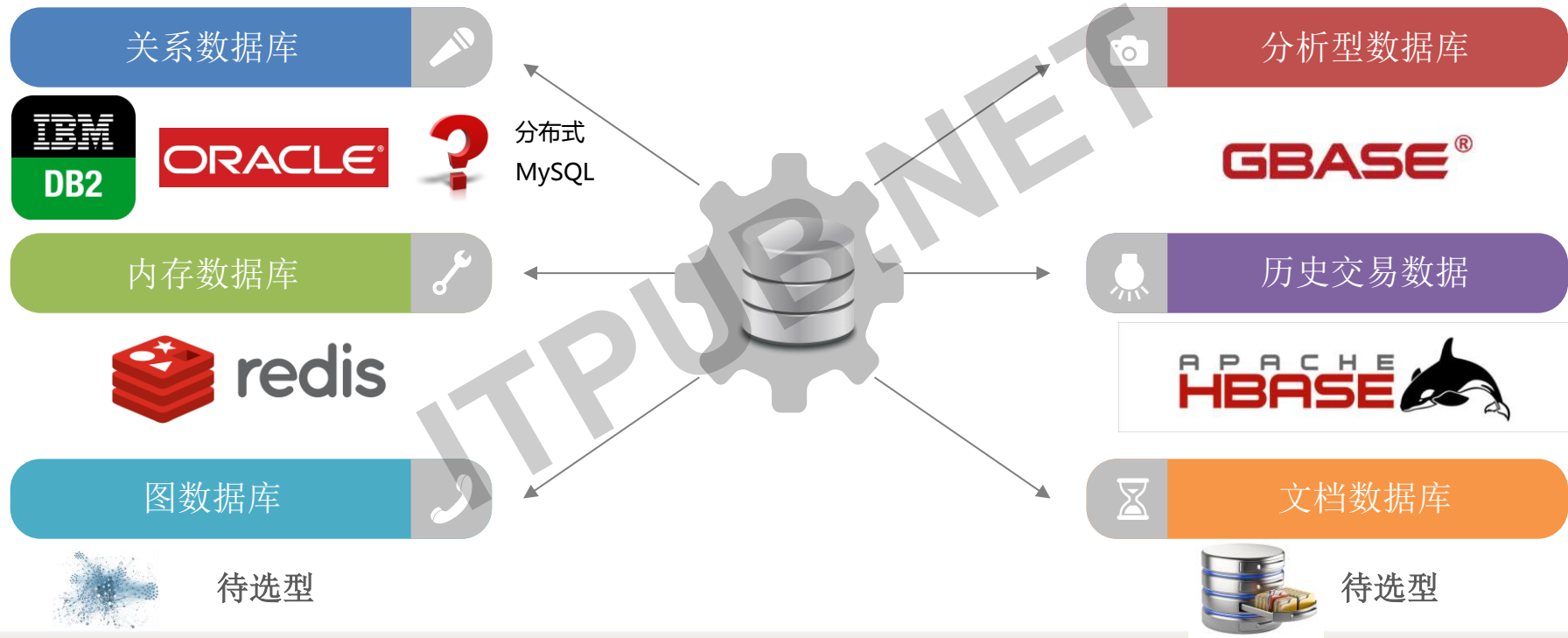


数据管理典型案例

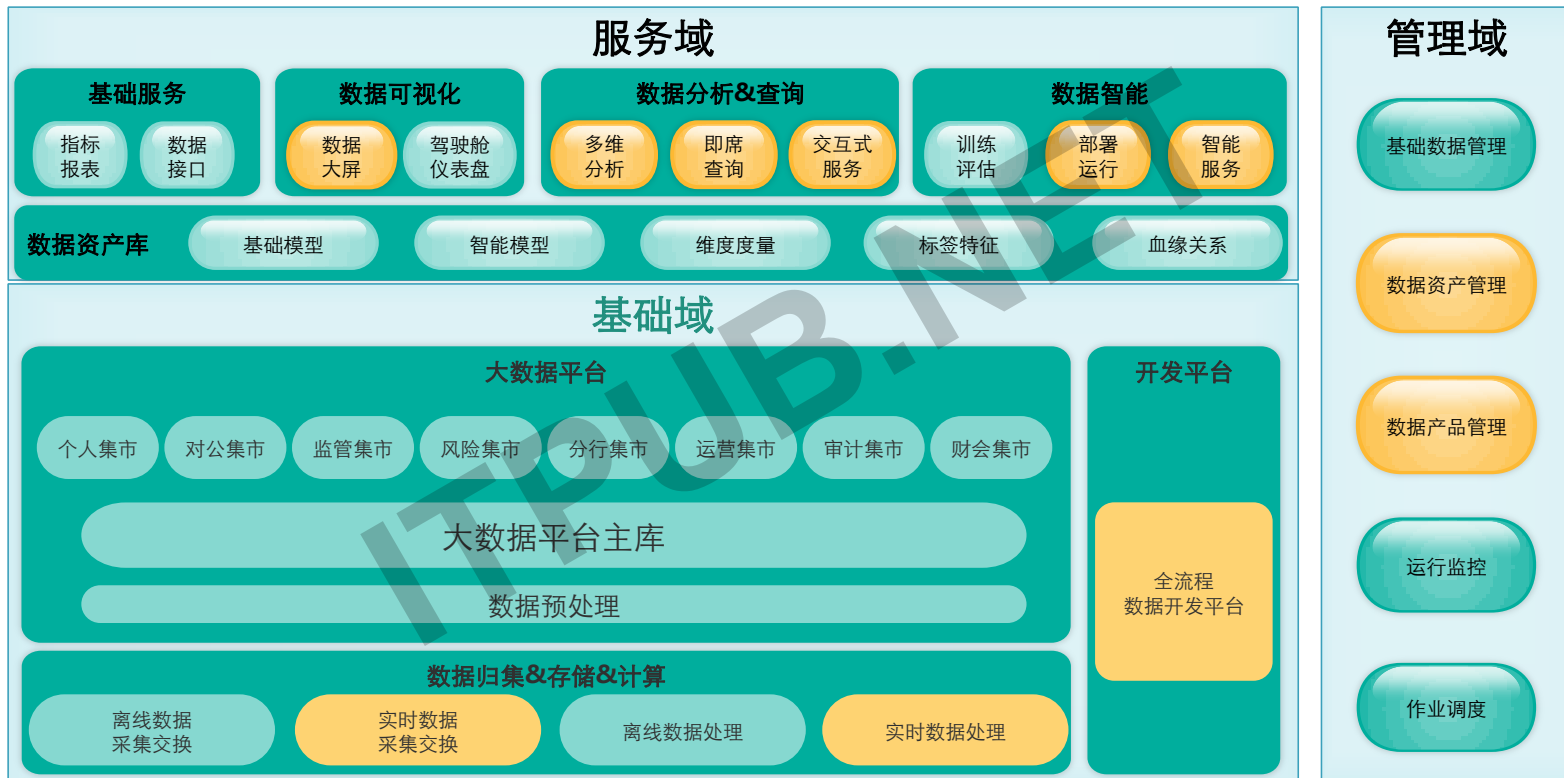


数据库发展规划

## 3.1 数据库产品战略

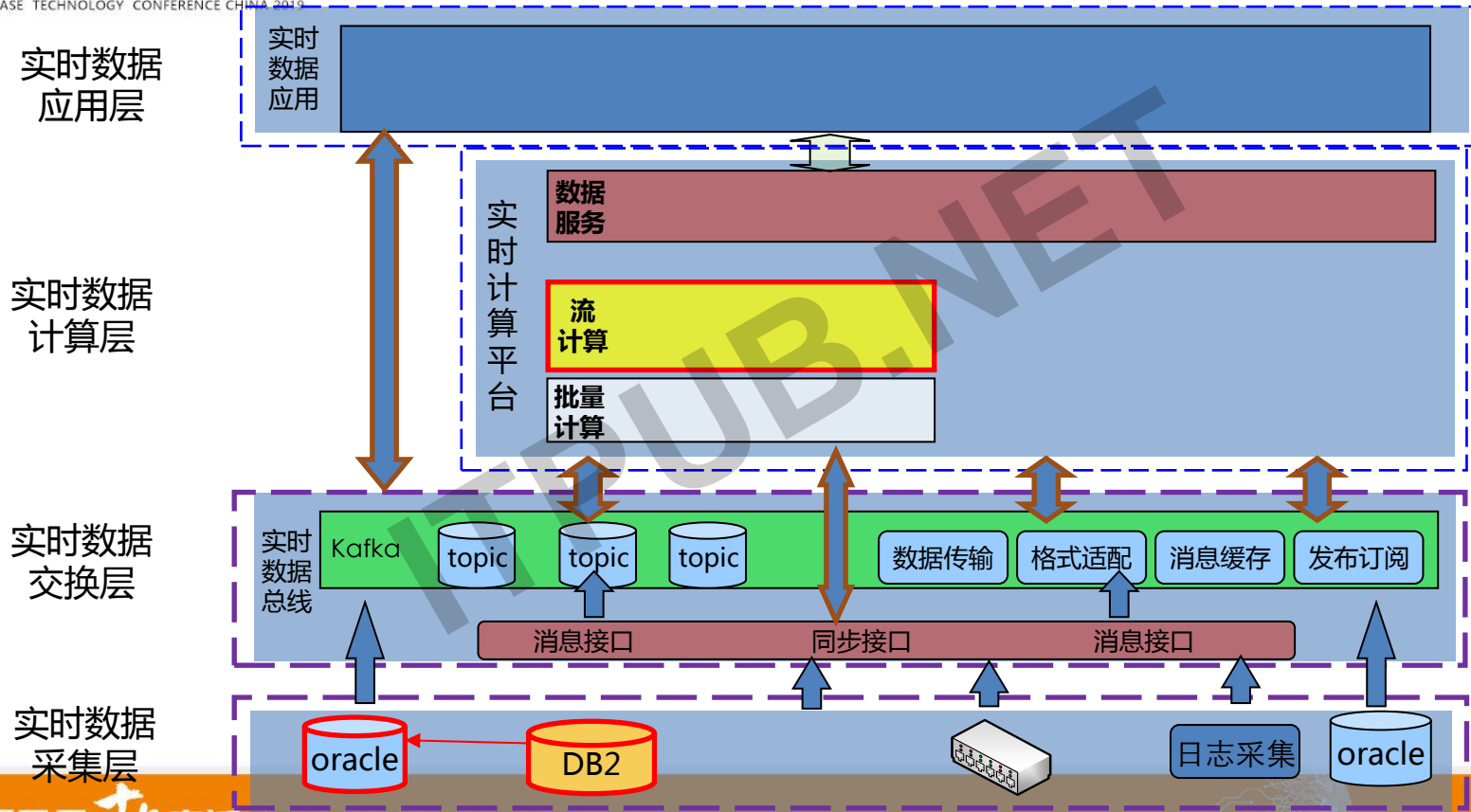


## 3.2 批量数据分布架构





## 3.3 实时数据分布架构



# 中国农业银行 稳健行远



2019

05

08-10

北京新云南皇冠假日酒店

## 数据风云 十年变迁

DTCC

第十届中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2019



+

○

○

○



THANKS

ITPUB.NET