



第十一届中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2020

架构革新 高效可控



北京国际会议中心 | 2020/12/21-12/23

混布数据库Hubble

深植于场景需求





架构革新 ◎ 高效可控
第十一届中国数据库技术大会
DATABASE TECHNOLOGY CONFERENCE CHINA 2020



变化趋势

DTCC
2020



北京国际会议中心

🕒 2020/12/21-12/23



万物互联，数据海量，非结构化，实时分析需求高



Transaction data



Engagement data



Observing data

一、生产方式变化	应用系统生产数据 IBM Oracle EMC	人的行为生产数据 Google Facebook BAT	机器生产数据 IoT 7*24小时产生数据
二、数据量级变化	TB级 千万级记录/天	PB级 10亿级记录/天	ZB级 千亿级记录/天
三、数据类别变化	结构化	结构化、半结构化	结构化、半结构化、 非结构化 (占比高达90%)
四、数据价值变化	单一业务应用 只需离线分析	流程业务应用 离线分析为主，实时分析为辅	混合业务应用 需要实时分析避免事故

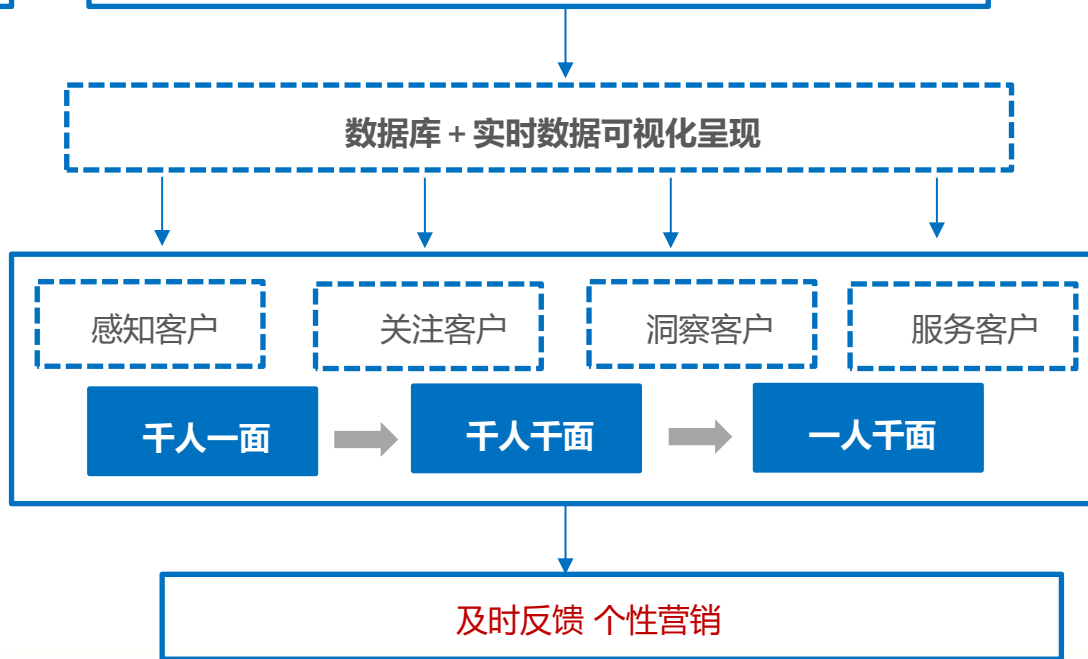
从决策层到服务层，一对一服务精准营销

Before: 管理驾驶仓----面向决策

领导决策



Now: 个性化驾驶仓----面向服务



混布数据库



对混布HTAP数据库的完美诠释

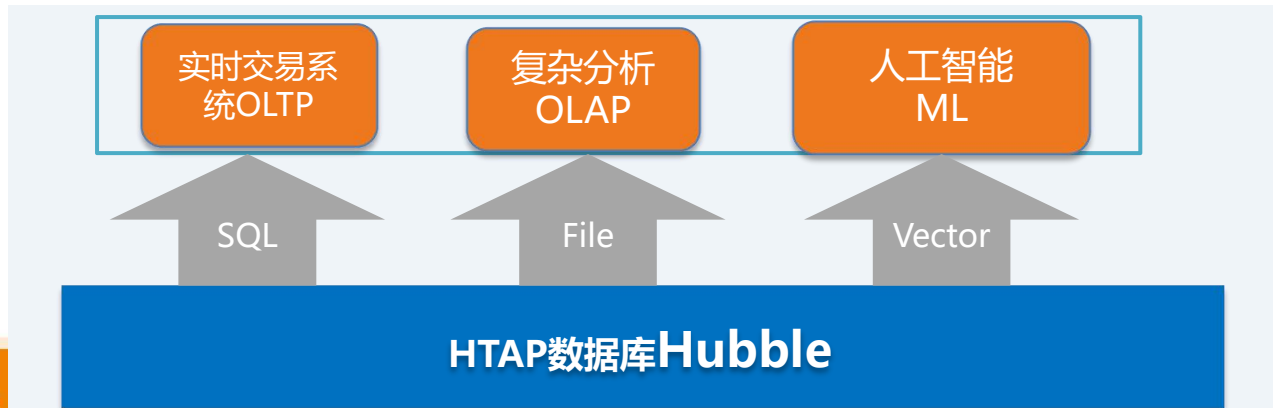
什么是HTAP?

Gartner对HTAP数据库给出了明确的定义：
HTAP数据库需要同时支持OLTP和OLAP场景。

为什么是HTAP?

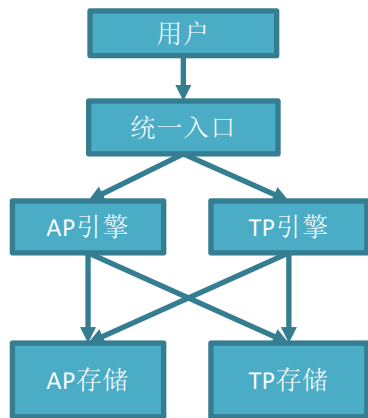
基于创新的计算存储框架，在**同一份数据**上
保证**事务**的同时支持**实时分析**，省去了费时的ETL过程。

Hubble完美诠释HTAP!

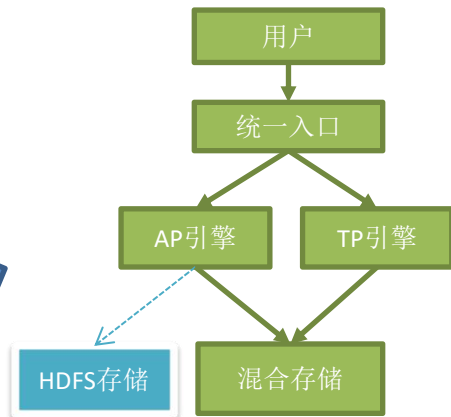


演进路线

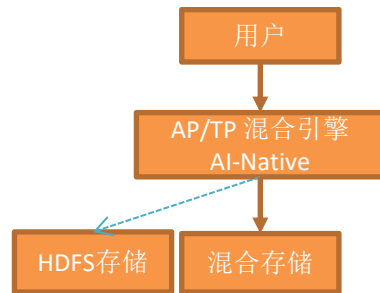
三个阶段



阶段一AP、TP混布



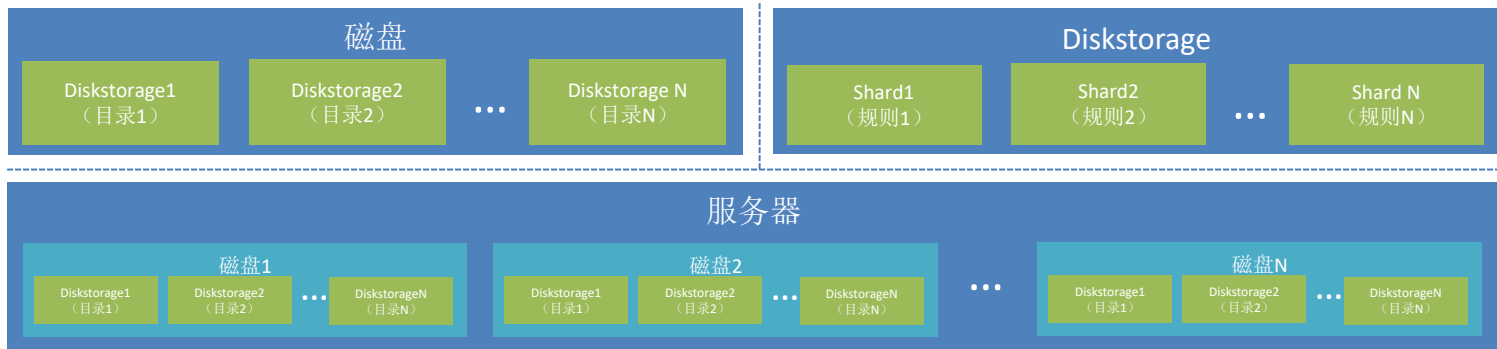
阶段二融合存储



阶段三融合引擎
& AI-Native

技术实现与优化

名称	中文	描述
Dataspace	数据空间	是一个分布式空间，通常一张表对应一个dataspace，一个dataspace对应多个Shard，dataspace可以是独占的存储空间也可以是非独占空间。
Exclusivespace	独占空间	是dataspace的一种，可以通过dataspace中的一个属性来描述，通常这个空间会在磁盘上有独占的存储目录，与其他空间是物理隔离的，独占空间包含多个shard，所有的shard属于同一个dataspace，属于同一张表。
Non-exclusivespace	非独占空间	是dataspace的一种，可以通过dataspace中的一个属性来描述，会与其他非独占空间的dataspace共用磁盘目录，通过逻辑来隔离，非独占空间会包含多个shard，多个shard可能属于多个dataspace，属于多张表。
Diskstorage	磁盘存储	1个diskstorage指的就是一个磁盘目录，也就是一个存储的实例。





存储分配

假定有3台服务器，每台服务器有两块磁盘，总共有3张表：

表1对应 Dataspace1

表2对应 Dataspace2

表3对应 Dataspace3

表1 为**非独占空间**，有2个shard分别是shard1和shard4

表2 为**非独占空间**，有4个shard分别是shard2、shard3、shard5、shard6

表3 为**独占空间**，有6个shard分别是 shard7、shard8、shard9、shard10、shard11、shard12

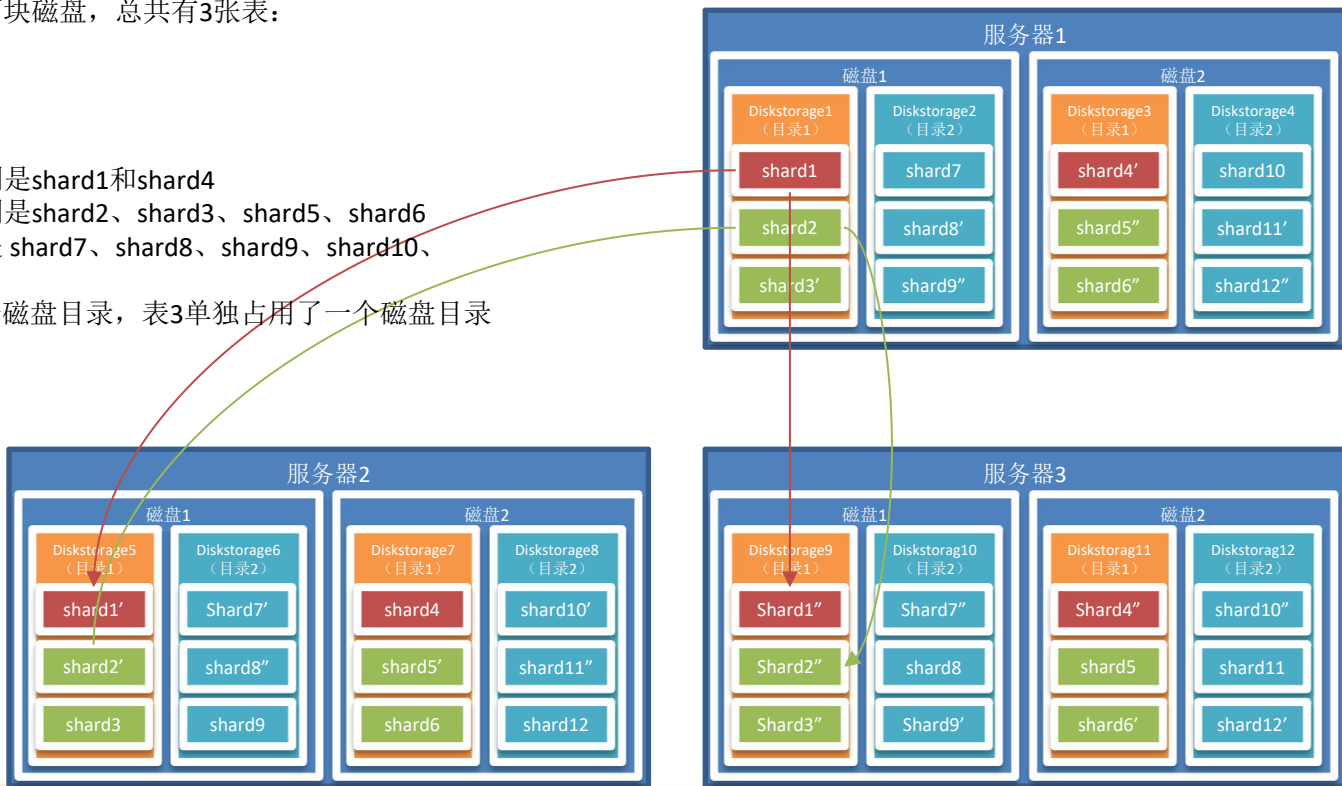
从图中可以看到表1 和 表2共用一个磁盘目录，表3单独占用了—个磁盘目录

图例：

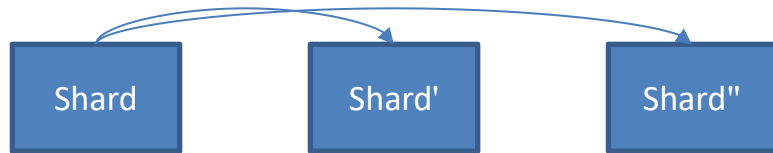
Dataspace1
非独占空间

Dataspace2
非独占空间

Dataspace3
独占空间



混合存储例子



样例说明(示例):

➢ 建表语句指定不同的存储格式

```
Create table sky(  
  id int,  
  address varchar(50)  
)  
with(  
  replication=3,  
  rep1-store='KV',      -- 副本1KV存储  
  rep2-store='COLUMN',  -- 副本2列存  
  rep3-store='KV',      -- 副本3 KV存储  
  shard-count=12  
);
```

说明:

- 可以建表时自定义shard各个副本的存储格式，格式选择有：列存、KV、MIX等，意思是Shard的数据存储格式可以与Shard'、Shard"是不同的。
- Shard与其副本之间通过Raft协议做数据复制
- 每个副本的数据格式可以做变更
- 默认表的副本数为3，默认表的3个副本存储格式都是KV

➢ 更新shard副本存储格式

alter table sky change store set rep3-store= 'MIX' ; -- 更新副本3存储格式为MIX

另外还有对shard的操作，如shard的下线、上线、切分、合并、检查、修复等先关维护功能。

潜在优势:

1. 逻辑上的一份数据支持了客户多样化的需求。
2. 真正的减少了数据的搬家。

数据分片



流程说明:

- 客户端通过不同的接口访问形式，直接访问主服务节点服务
- 主服务节点收到服务请求进行分析处理，分配到不同的分配服务节点执行
- 分片服务节点收到执行请求，进行sql解析处理并执行SQL计划
- SQL执行服务底层存储数据进行处理访问，并返回处理结果



数据索引

存储和索引在同一个
Dataspace中

KV索引

倒排索引

复合索引、索引选择
评分机制使用最
匹配的索引映射查
询

索引可以删除、修改等操作

索引index_id为自增

Key

Index_id, NULL-byte, 二级索引
列, 主键列

Value

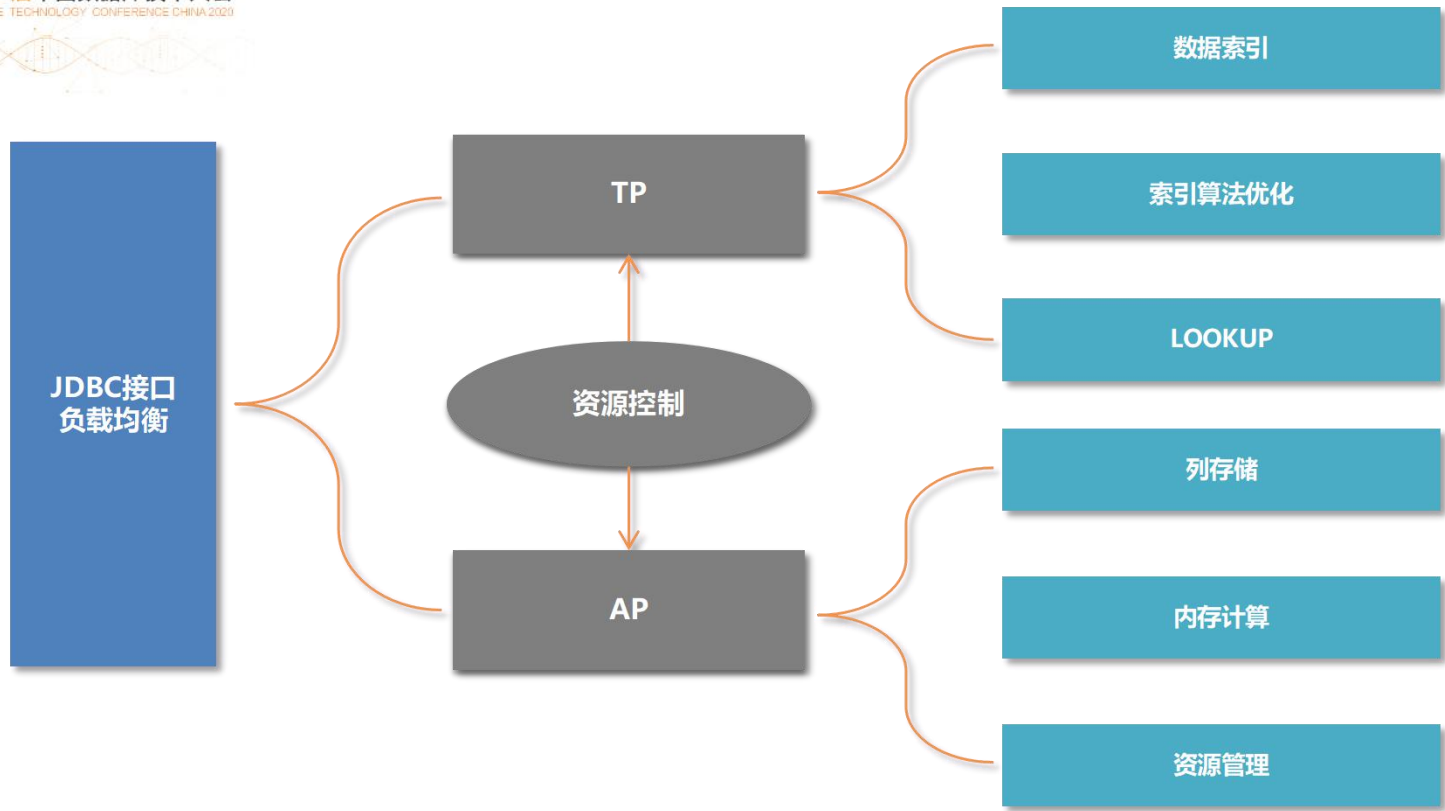
主键在key的起始位置



存储和索引在同一个dataspace中

索引字段为索引的ID

混合负载



资源管理





产品特性

改变传统架构、优化大数据架构

架构体系

IOE架构, 数据采用 oracle、mysql,
TD 和 mpp 数据库等

存在问题

- ✓ 节点无法水平动态扩展
- ✓ 对于大数据量的支撑比较受限
- ✓ 采用分库分表实现, 实施人员业务技术能力要求高
- ✓ 技术能力限制, TP, AP分离, 数据冗余存储, 数据应用迁移搬家, 耗时 效率低

传统业务应用

架构体系

大数据 Hadoop 平台架构, 数据采用 Hive、
Hbase, ES 和 solr 数据库等

存在问题

- ✓ 原生API开发实施难度大。
- ✓ 大数据体系SQL工具Hive性能缓慢
- ✓ 复杂场景多组件组合应用, 数据冗余存储多副本, 运行效率低

大数据业务应用

多源异构、超高实时并发、全部SQL标准



多源异构

统一服务入口，接入各类数据库源系统，自由编写SQL，实现数据访问服务，**无需将数据完全搬迁，即可以现有数据即席分析探查。**



超高实时并发

利用分布式技术架构，**支持上千万用户在线实时高并发修改与查询，**服务个性化数据服务应用



全部SQL标准

支持SQL99标准，封装多种查询调用接口，方便系统对接服务应用，**实现传统场景无缝迁移，降低生态合作伙伴参与大数据开发的门槛。**

在企业级数据
处理领域

支持OLTP 高并发事务的服务应用；

支持OLAP的实时数据仓库服务；

支持海量数据分析处理，即席探查的服务应用。



场景和案例

事务交易、数据仓库、大数据领域

事务交易领域

- 传统关系数据库的升级
- 核心交易系统

数据仓库领域

- 升级、优化传统数据仓库，构建**实时数仓**。
- 满足时效性、并发性需求

大数据领域

- 支持海量（PB级）数据的存储、处理及分析，升级替换Hadoop生态组件
- 满足**易用性**，**低门槛**需求

最终 用户 收益

大规模+实时+高并发：

打破“不可能”，实现业务梦想

实时高并发+事务一致性：

突破国产化性能低的现状，有力支持国产+安可

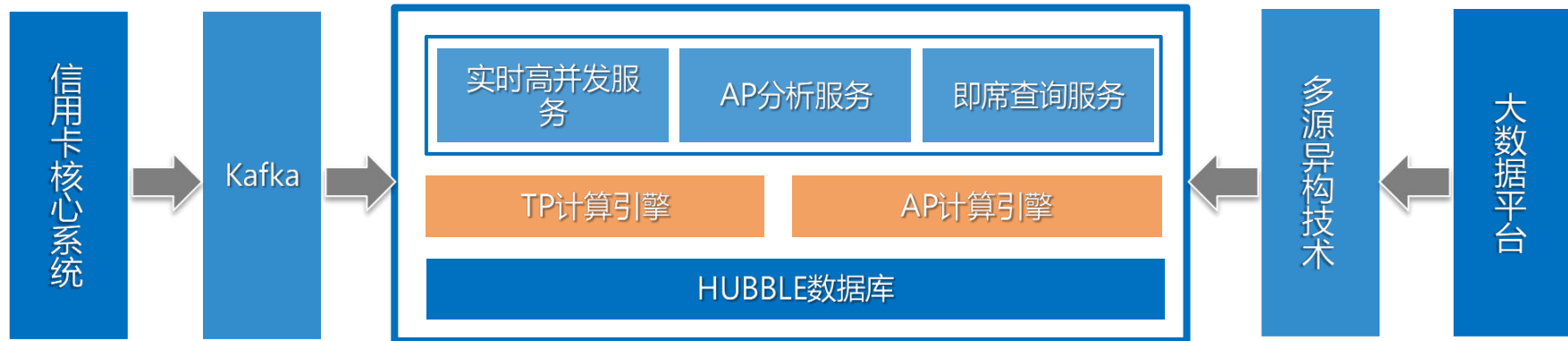
降低人力资源（业务实现、运维）成本：

仅SQL要求



系统快速上线整体逻辑

某大型股份制银行基础大数据平台



六大类业务应用场景支撑

贷中实时反欺诈业务应用

贷中实时交易欺诈规则判断应用

实时营销业务应用

SAS营销业务整体迁移支持

分行实时统计业务应用

信用卡业务即席查询和数据统计分析业务

进件审批反欺诈业务应用

进件审批多维规则实时判定应用

TP业务场景

促发卡、促激活和渠道类应用实时分析

AP业务场景

客服运营数据处理和分析

系统上线业务效果

在实时反欺诈场景中：

实现了**每天千万级**交易数据数据**实时并发**入库；
接近**100个**复杂欺诈规则分析
毫秒返回分析结果。

在实时营销场景中：

实现**5000多个**标签灵活**组合查询**分析，精准命中数据
毫秒级输出查询结果。
支持**100多个**节点的大数据平台秒级数据同步。

在实时审批场景中：

支撑**50多个**维度数据复杂规则分析，**毫秒级**返回分析结果，
支撑**实时**进件审批业务，
同时支撑**1万多家**分支行，**上千家**营业厅的**实时汇总统计**分析的业务应用。

某商业银行全量数据交易系统

痛点

用户核心系统采用oracle架构体系，目前只能提供13个月内的数据服务应用；

13个月之前的数据存储在带库系统和数据仓库系统；

带库系统无法提供高并发数据服务应用；

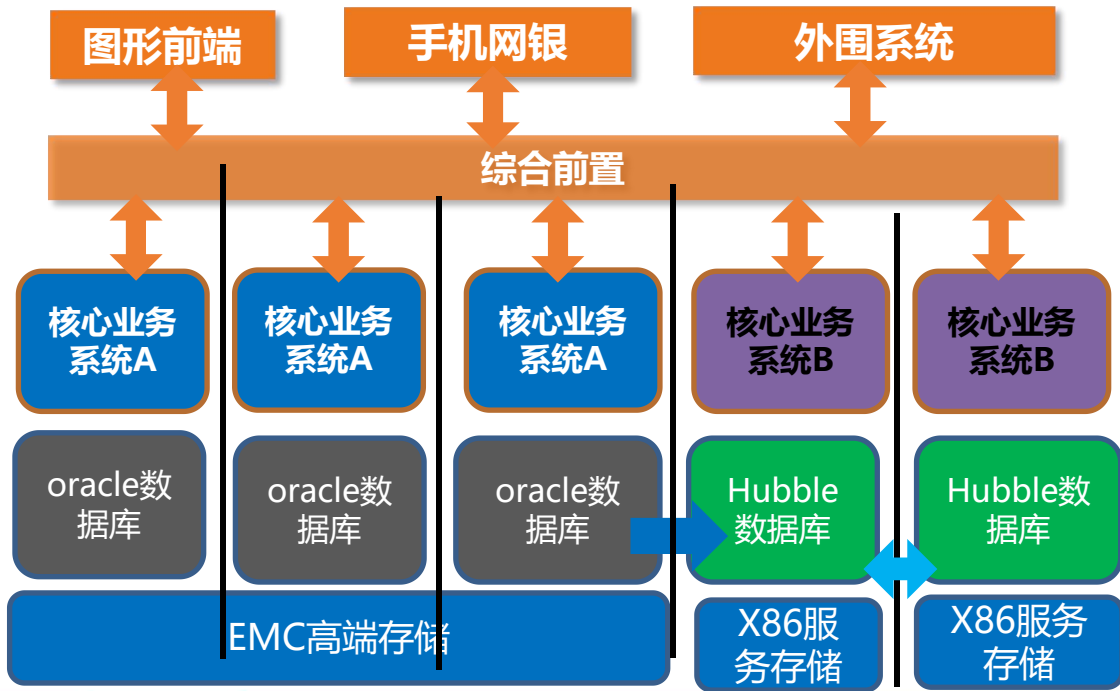
数据仓库系统数据已经进行了加工，无法提供原始数据。

终端用户无法获取全量数据服务。

AI分析建模需要用户全量历史数据也无法提供。

现状

项目实施架构



- “四双架构”实现系统高可用性
- 为了满足金融行业A类核心系统灾备要求，本系统在生产和灾备两中心采用“双中心 双集群 双活 双服务应用”的四双架构，构建高可用应用系统。两集群间服务、数据独立，互相同步，保持高可用。

海量数据服务



每天海量的感知数据如何
存储与使用



需求 & 痛点

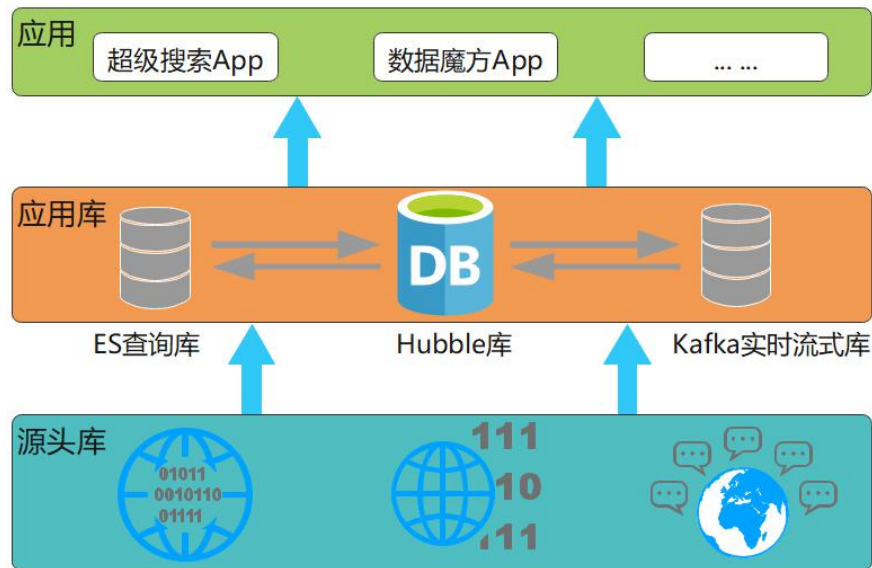
➤ 数据入库:

- 1、源头库种类多，包括Oracle、MySQL、Gbase甚至还有FTP；
- 2、数据量大，日数据量超过5亿条；

➤ 查询响应:

- 1、同时服务于省厅的统计分析和地市一线干警；
- 2、从批处理向即席查询转变，查询结果时间要求 < 5秒。

项目实施架构



数据形式多样，数据来源较多



分别为上层统计分析应用与一线干警提供数据服务，并对接到前端应用；



通过将原先的业务库数据统一汇聚到Hubble数据库，构建用户的实时数仓，在只改变数据库接口指向的前提下使用Hubble支撑原有业务；



凭借数据库超高性能，将新的感知网数据通过实时方式入库到Hubble数据库（每天超过5亿条）；

实施效果

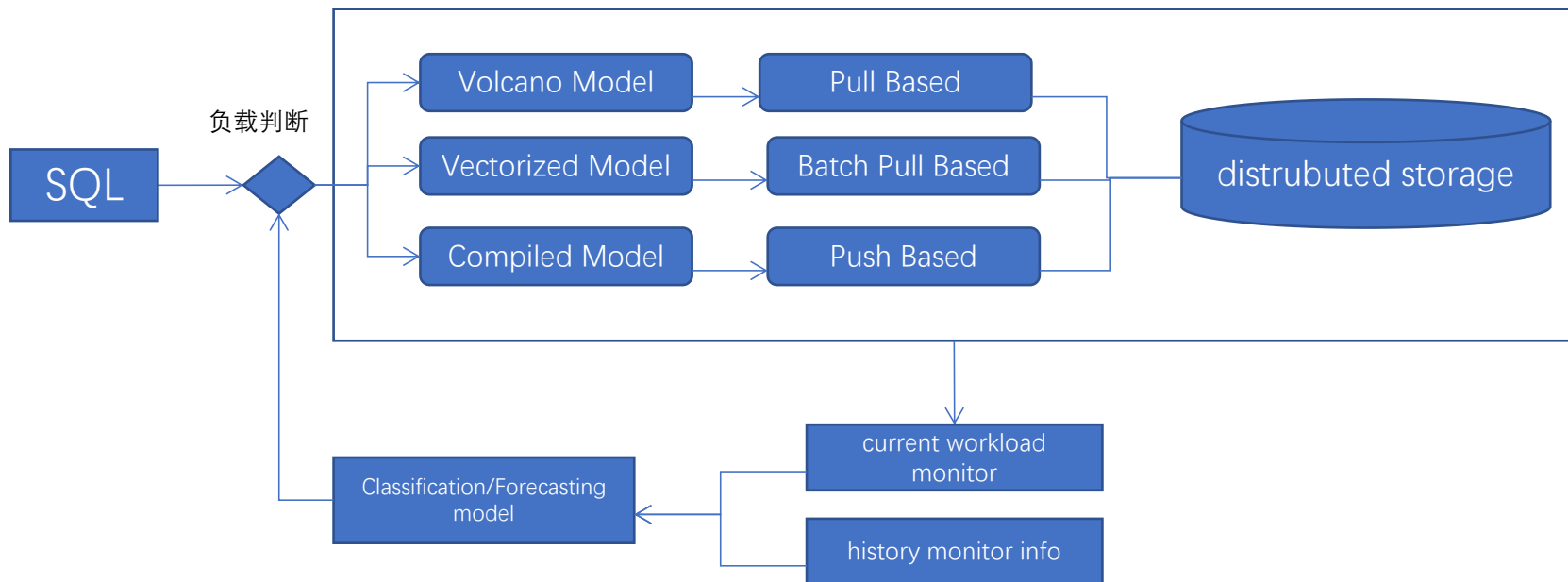
通过灵活的扩展能力、高效的读写/查询性能等优势顺利解决用户对于存储效率、高并发查询效率，同时还要具备便于扩容，并做到同时兼顾业务读写与统计分析的问题；帮助用户将感知网数据融入到日常办案工作中，在提高案件侦破率的同时也减少了办案的时间。





AI-Native支持

负载识别-与AI结合

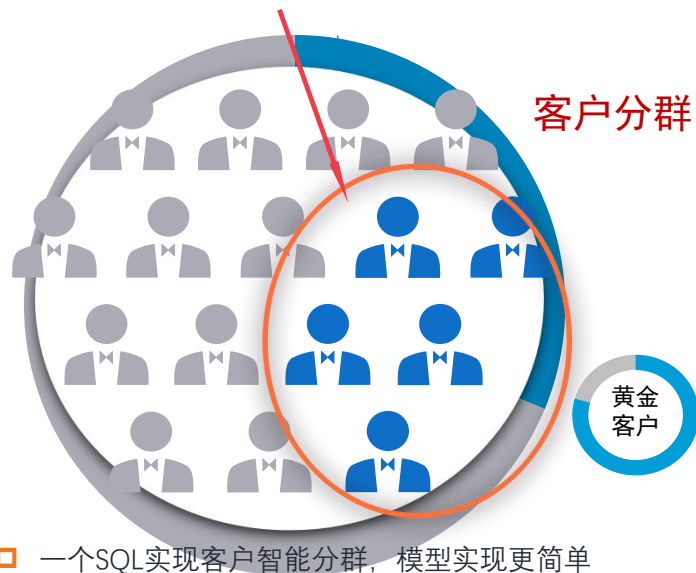


利用AI算法的优势，优化实现Hubble数据库分层设计技术难题，让Hubble数据库设计**更全面智能**

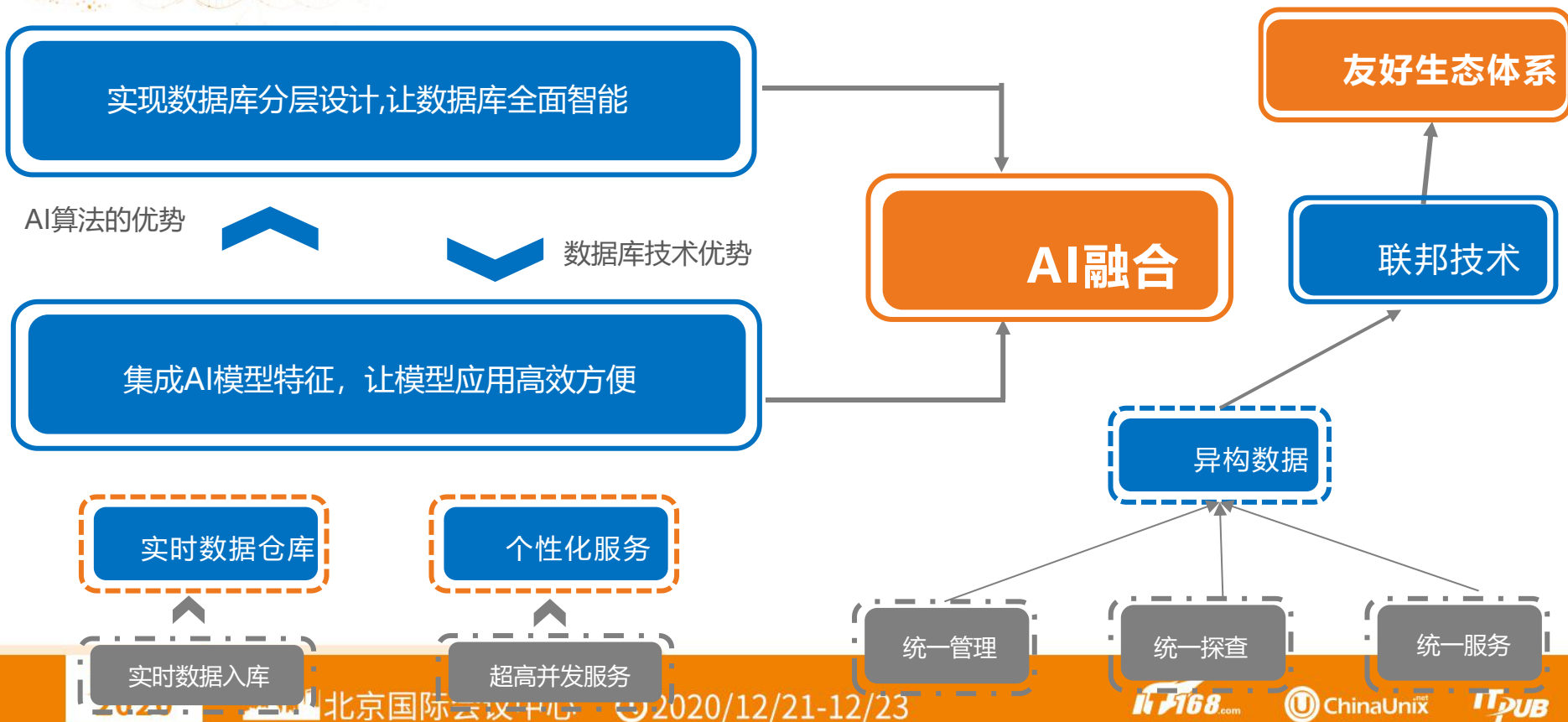
- ABO-基于人工智能SQL解析，运行AI模型，让SQL解析优化更智能！
- 智能存储-基于人工智能数据分片，运行AI模型，让数据三副本分片更智能！

利用Hubble数据库技术优势，集成AI模型特征SQL代码实现，让AI模型应用**更高效方便**

```
insert into db1.user_info  
select analyze_udf(user_code, user_name) as user_classification, serial_number,  
info,  
From db1.user_info;
```



生态友好 AI融合



THANKS

