

### 第十三届中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2022

### 数据智能 价值创新











OceanBase

数据来源:数据库产品上市商用时间

openGauss

RASESQL



# Hubble Ai-Natve数据库混合存储之列存

乔旺龙+天云数据+分布式平台负责人









### Hubble数据库架构







混	资源管理			资源管理			资源管理		
合资	资源收集	资源池	资源告警	资源收集	资源池	资源告警	资源收集	资源池	资源告警
合资源管	资源分配			资源分配			资源分配		
理	AP/TP任务分配			AP/TP任务分配			AP/TP任务分配		
混	执行器(Execution)				执行器(Executior	1)	执行器(Execution)		
合计算框架	任务管理			任务管理			任务管理		
	内存计算			内存计算		内存计算			
架	数据扫描(条件下压)				数据扫描(条件下压	)	数据扫描(条件下压)		

混	存储层(Storage)						
混合左		存储/索引引擎(Storage Engine)					
仔储	分片 (Shard)	分片 (Shard)	分片 (Shard)		分片 (Shard)		

存储层(Storage)							
	存储/索引引擎(Storage Engine)						
分片 (Shard)	分片 (Shard)	分片 (Shard)		分片 (Shard)			



IT<sub>PUB</sub>

### 逻辑计划融合-图逻辑计划融合实现

数据结构映射

- 图的存储结构映射到Hubble的存储层
- 图的索引结构映射Hubble的索引
- 通过SQL可以查看图的底层数据

多源异构图结 构

- 图的节点、边、属性等映射为数据库的表和字段等
- 图的数据可以供给Hubble的计算层
- 通过SQL可以方便的访问图的数据,且可以把结构化数据转换为图的数据结构

支撑图算法

- 用Hubble的计算引擎支撑图的算法
- 在统一的组件中实现图的AP、TP操作
- 图算法计算结果可以直接在图库中表达

兼容**Gremli**n

- 延续用户使用图数据库习惯
- Gremlin中扩展SQL表达,可以同时操作结构化数据和图结构数据

数据复用

- 在实际应用中,没有必要把全部的数据加载到图库中
- 因为图结构和数据库结构为一体的所以图的数据直接调用数据库中的数据
- 图结构更适合高维的表达,数据库结构更适合表结构数据实现互补

用户 **ThinkerPop** 嵌入 SQL接口 管理API **API- Gremlin** 接口服 供数 数据转换 混合计算引擎 数据映射 读取数据 存取数据 混合存储



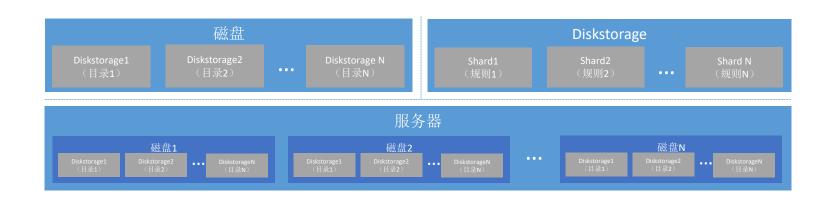








名称	中文	描述
Dataspace	数据空间	是一个分布式空间,通常一张表对应一个dataspace,一个dataspace对应多个Shard,dataspace可以是独占的存储空间也可以是非独占空间。
Exclusivespace	独占空间	是dataspace的一种,可以通过dataspace中的一个属性来描述,通常这个空间会在磁盘上有独占的存储目录,与其他空间是物理隔离的,独占空间包含多个shard,所有的shard属于同一个dataspace,属于同一张表。
Non- exclusivespace	非独占空间	是dataspace的一种,可以通过dataspace中的一个属性来描述,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。











### 存储样例

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

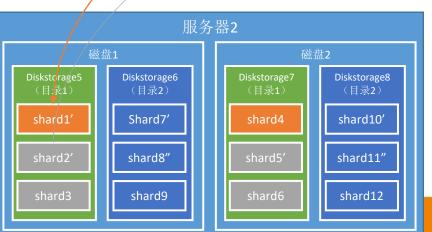
表2 为非独占空间,有4个shard分别是shard2、shard3、shard5、shard6

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

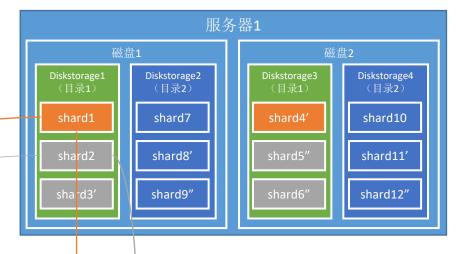
从图中可以看到表1和表2共用一个磁盘目录,表3单独占用了一个磁盘目录

注:因为对应的线过多,图中仅画shard1和shard2的复制路径

独占空间的进一步细化,针对多个目录下如何分片数据。







Dataspace1

非独占空间

图例:





### 图逻辑计划-数据结构映射



- ▶ Hubble混合存储提供分布式KV接口
- ▶ 数据结构映射需要把图的点、边、属性、标签 等内容映射到KV的数据结构中
- ▶ 图数据库ID,复用数据库分布式序列管理
- ▶ 图索引,复用数据库的索引机制

存储结构	描述
图ID	图的ID数据管理,负责图ID自动生成,复用序列管理  为 节点序列  为 边序列  为 属性  为 标签序列等
数据存储	用于存储点、边、属性、标签等数据
图索引	图的索引管理,复用Hubble的索引
系统参数	系统参数表

- 1、通过多源异构的方式,实现了图的SQL表达:通过SQL来访问图的点、边、属性等数据
- 2、图的批量数据加载,只需一条SQL语句( create table as 或者 insert into <图表> select \* from ...)
- 3、对分析类操作直接在Hubble数据库的计算层和调度层处理,真正实现图数据库的一体化





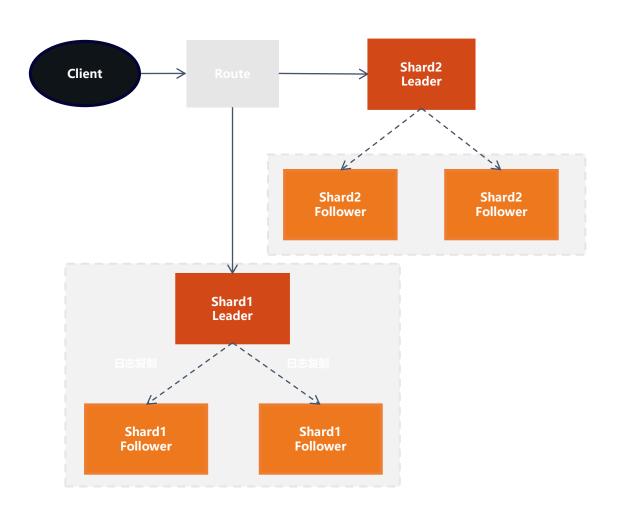


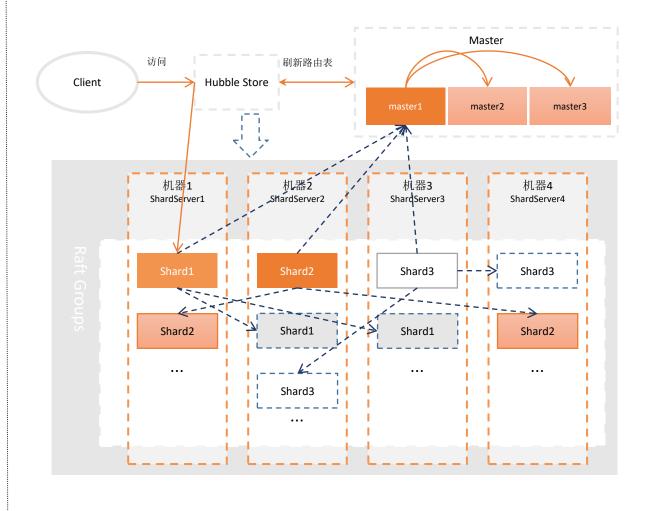




# Hubble关键技术点 — 数据一致性

















# Hubble关键技术点-高并发存储



#### 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sq1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,并反回 处理结果
- Zookeeper保证相关服务应用的高可用
- HDFS持久化底层存储数据,并利用三副本技术保证 数据不丢失









### Hubble关键技术点-高并发索引



存储和索引在同一 个shard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



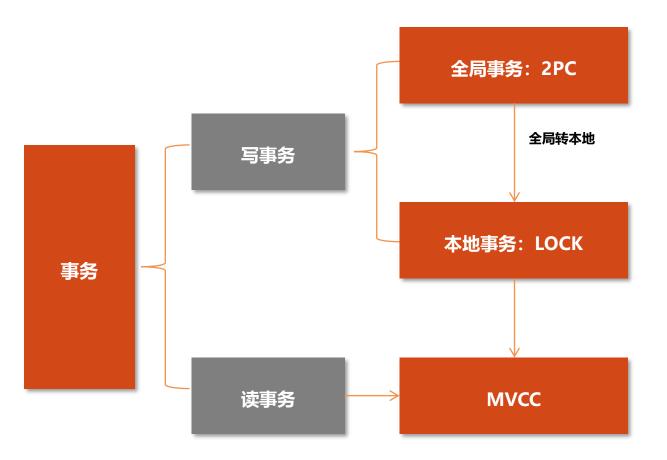








# Hubble关键技术点-高并发事务



事务开始,记录事务唯一ID,执行操作,记录修改的shard,执行预提交动作,提交或回滚

写入时当前采用锁机制

读取使用快照读取,存储层每次写入都是追加写入,通过覆盖机制 进行数据变更









## Hubble关键技术点-高并发事务



1、开始事务,产生事务唯一ID (UUID)

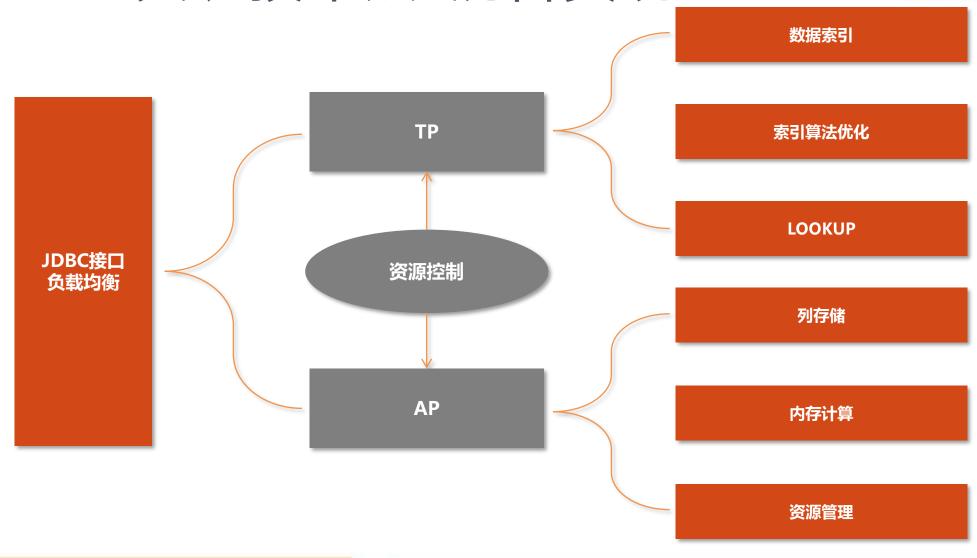




1 168.com

# Hubble关键技术点-混合实现















# Hubble关键技术点 一资源管理

#### 资源管理

#### 调度模式

- FIFO Scheduler 先进先出
- Fair Scheduler 公平调度 每个单独的任务可以指定优先级

#### CPU管理

- 实时获取每个shardserver cpu的使用情况
- 以及每个任务cpu资源占用情况

#### 内存管理

- 实时获取每个shardserver 内存的使用情况
- 以及每个任务cpu资源占用情况









# Hubble关键技术点一性能保证





据的落地

















### 基于传统技术的银行数据建设现状



#### 银行科技人员



怎么跑这么慢, 今天要是不能按时开门就惨了。

领导要的报表还没跑出来,要扣奖金了。

#### 总行分行领导



今天月末了,不知道指标都完成的怎么样了?

市场有新的动向,不知行内客户的反应怎么样?

#### 一线营销经理



查到客户存了一笔钱进来,想针对客户推个新产品,结果不知道客户已经把钱取走了。

#### 一线客户经理



一个贷款的审批流程加工好再审批都过了一天了,客户不能及时得到批复,客户反映体验不好。









# 业务驱动数据分析成为流行



### 由IT驱动变为业务驱动





### 高效的OLAP分析





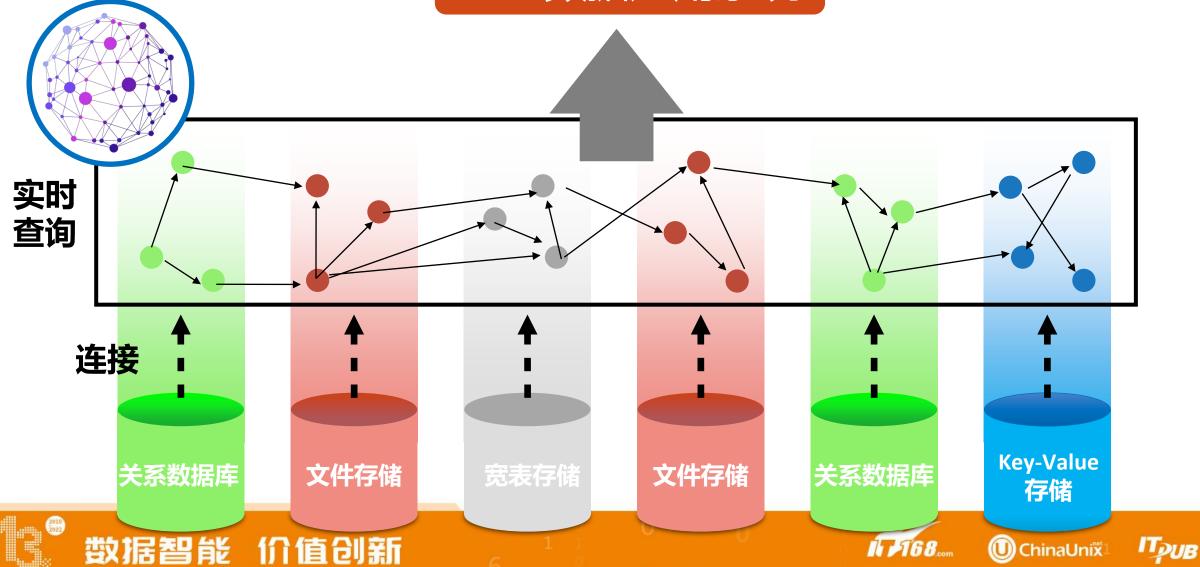




### Hubble的使用价值



### 企业数据应用系统





### Hubble数据库适用场景

BI分析的小屏化 AP场景高并发访问 数仓消费化 数据下沉 数据在线化 数据即席服务 历史数据在线服务 实时数据在线服务 数据迁移问题 Hive加工数据性能问题 Spark加工数据并发访问问题 ES 开发复杂度问题 加工性能问题 Hbase数据库查询不灵活问题 Hbase开发复杂度问题 开发难度问题 Phenix组件稳定性问题 Kylin数据加工时效性和数据灵活访问问题 ... ...









## 场景1 - 行业数据中台



**************************************		海鳥物根区	物伊西台	计每亚令	数据产品	门户
数据通道	数据标准化区	海量数据区	数据中台	计算平台	女X打占/一口口	
内部结构 化数据		实时流数据		图计算 平台	人物画像	统
内部非结构化数据	统 一 数 据			数据魔方	<b>一</b> 数	
外部结构 化数据	数 据 交 换	非结构化数据平台		人工智能	实时预警	据
外部非结 构化数据		行外数据平台		平台	模型探索	门 户
数据管控	数据治理平台					

数据管控

数据治理平台





## 场景1 - 行业数据中台



数据通道	数据标准化区	海量数据区	数据中台	计算平台	数据产品	门户
内部结构 化数据		Kafka + Flink	(S) (S) Hubbl∈	Hilbert	人物画像	统
内部非结构化数据	统 一 数 据	MPP			数据魔方	<b>—</b> 数
外部结构 化数据	数 据 交 换	Hadoop Hbase	ー HTAP数据库 -	MaximAl	实时预警	据
外部非结 构化数据		Elastic Search			模型探索	门 户
数据答控	※ <b>↑</b>					

数据管控

数据治理平台



14 168.com



### 场景1 - 数据中台



