

Tendis在实时数仓中的应用

王新春@唯品会 数据平台

实时数仓架构

数据产品

数据服务（指标系统）

数据运营平台

供应商平台

其他产品

ADS

应用层(KV/MySQL/HDFS)

运营看板

品牌看板

预热看板...

实时特征

广告

智能监控..

DWS

轻度汇总层(Kafka/Kudu/HDFS)

订单实时宽表

用户实时宽表

流量实时宽表

DWD

事实数据(Kafka/HDFS)

泛订单数据

订单

商品

用户

流量数据

PV

点击

曝光

监控

ODS

原始数据(Kafka)

订单

商品

用户

APP/微信

PC/WAP

Binlog

MySQL

Log

APP/Wechat

Log

PC/WAP

开发平台

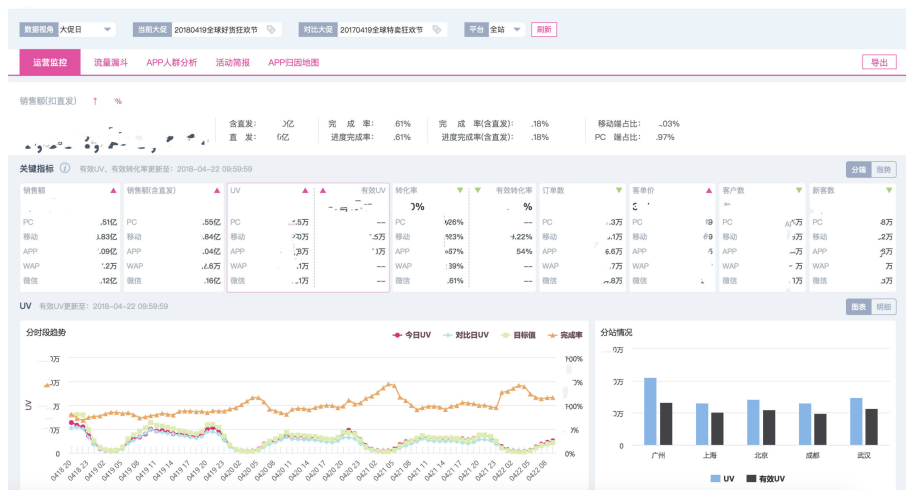
IDE

权限

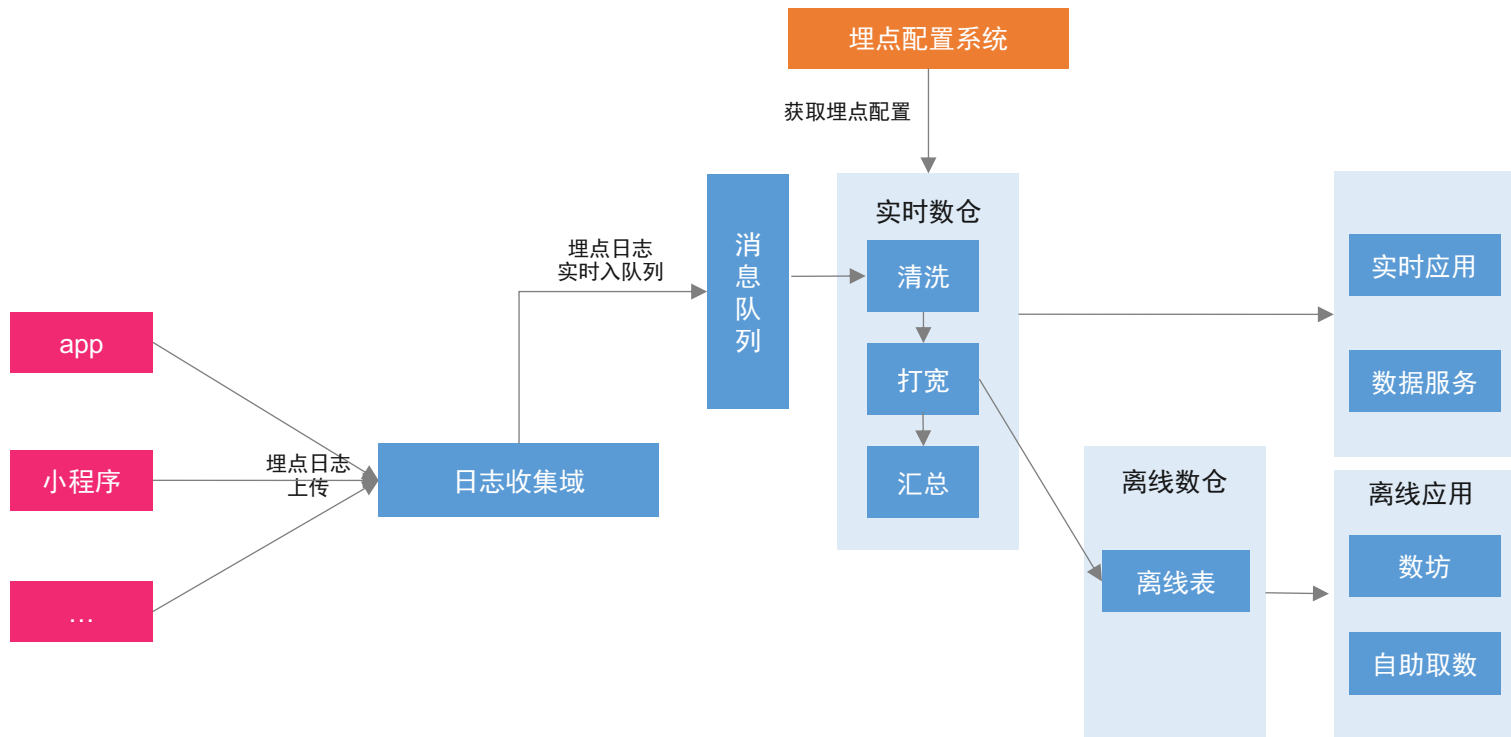
监控

质量

实时数仓——数据看板

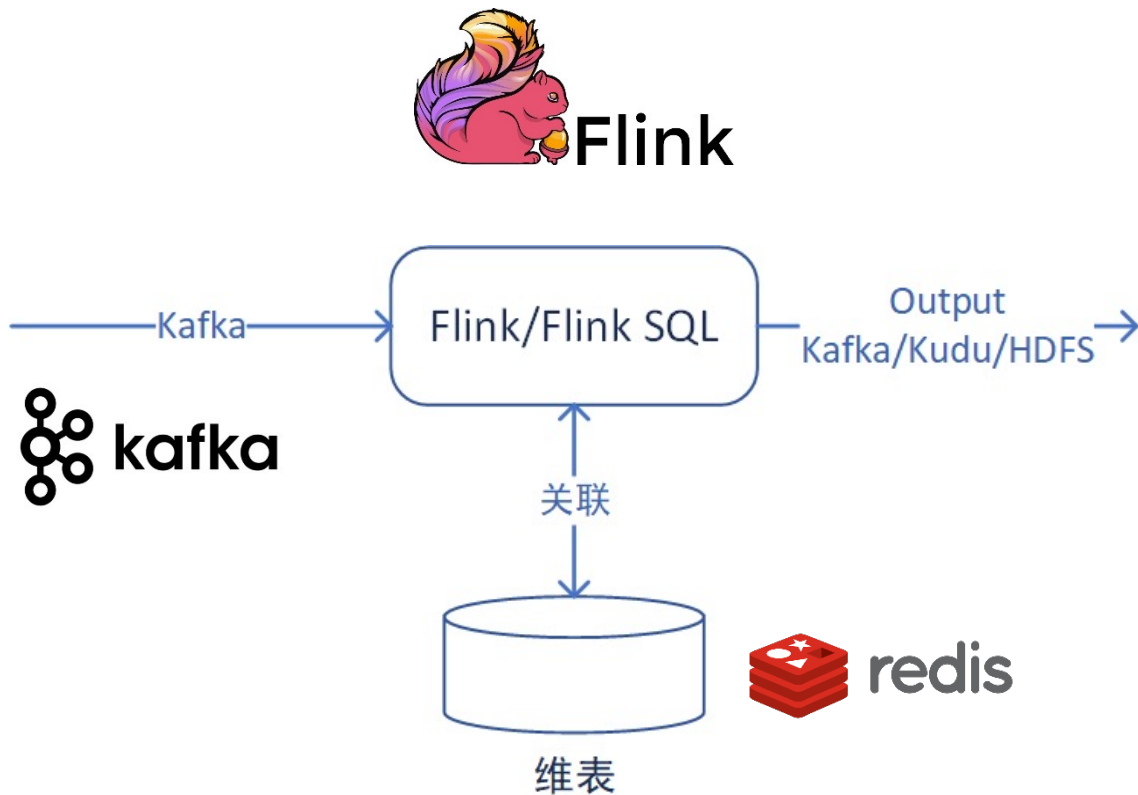


实时数仓——数据产品

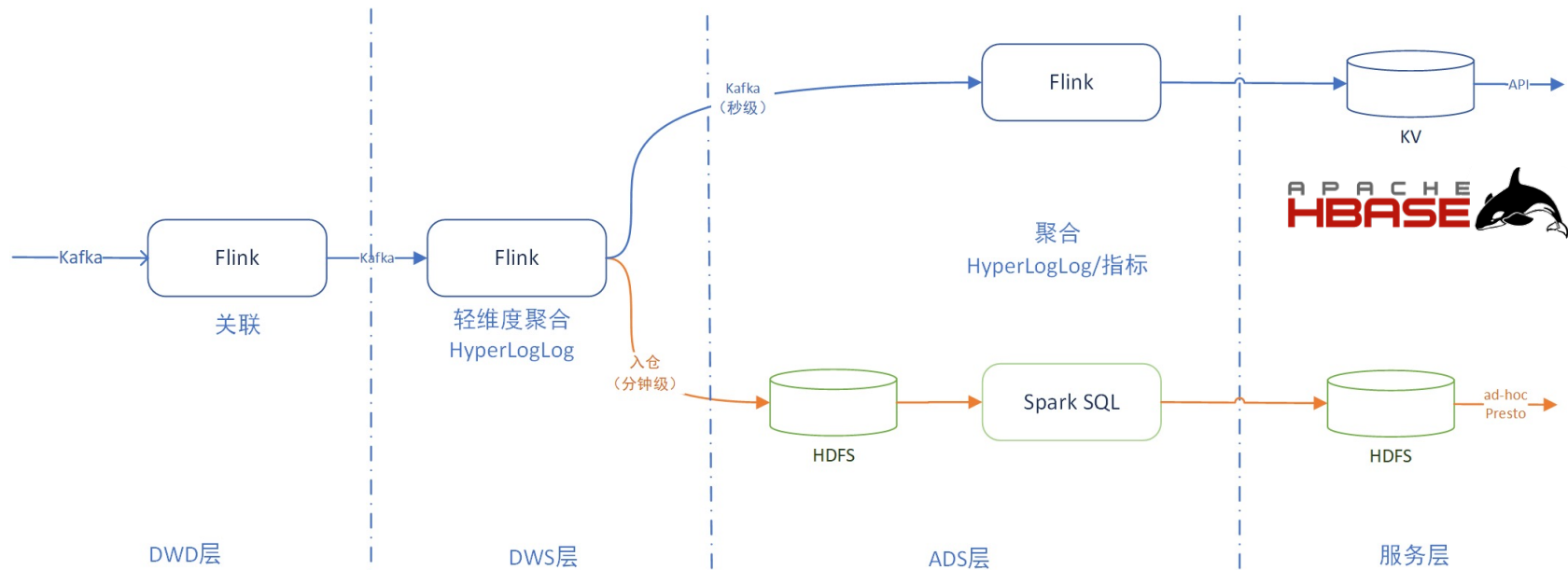


- 各端埋点统一登记在埋点配置系统，包含埋点详细字段、埋点注释含义等
- 实时数据清洗埋点日志时，只认已经在埋点配置系统登记过的埋点字段，未登记的则不识别

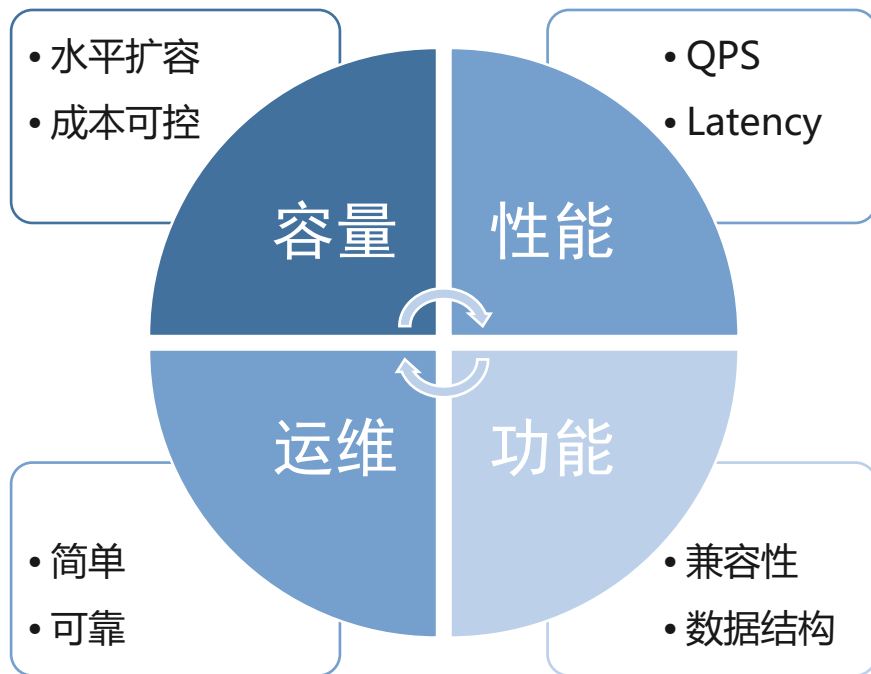
实时数仓——维表关联（打宽）



实时数仓——UV计算



实时数仓——KV强需求



选择Tendis的主要原因



自主可控

项目代码完全开源



快速迭代

社区版迭代节奏与云上基本一致



社区活跃

专职研发对接社区的issue



协议兼容

完全兼容Redis协议和命令



去中心化

保持Redis的使用和运维习惯



内核自治管理

故障探测，水平扩展，副本迁移



关注Tendis开源项目



关注Tendis官方社区交流群

Tendis的版本演进

2015.4

Tendis SSD正式立项

Based redis-2.8.17 + rocksdb

2018.9

Tendis存储版

多线程，集群自治，去中心化架构

2020

上云+开源

2020.9 Tendis冷热混合存储上线
2020.12 Tendis存储版开源
2021.6 Tendis存储版上线

2016.4

Tendis SSD v1.2发布

数据结构持久化，物理备份，增量复制

2019.5

Tendis冷热混合存储

热数据在内存，全量数据在Tendis存储版
缓存层+存储层支持分别扩缩容

Tendis存储版 vs Redis Cluster



可靠性

数据落地
Binlog落地
秒级备份

复制

镜像
+ binlog
断点续传
并行复制

并发机制

key级锁
多线程
线程池动态调整

集群

独立的gossip线程
压缩的gossip消息

数据搬迁

slot搬迁，独立线程
搬迁过程可观察，可
控制，对服务无影响



可靠性

内存
aof rewrite
bgsave

复制

rdb+aof
全量复制
串行复制

并发机制

单线程
redis 6.0引入网
络多线程

集群

Gossip共用主线程
gossip消息

数据搬迁

key搬迁，共用线程
依赖工具，大key会
阻塞服务

Tendis存储版 vs Pika



集群

独立的gossip线程
压缩的gossip消息



集群

Sentinel / Codis
初步的集群模式3.4

Pika v3.4.0 VS Tendis 2.12 单节点

	Tendis-24T	Pika-20T	Tendis-36T	Pika-36T
GET 100w	6.93s	6.00s	6.15s	5.83s
HGET 100w	7.87s	6.07s	5.81s	6.38s
HGETALL 100w	40.60s	21.80s	23.96s	24.97s
HSET 100w	14.10s	10.19s	9.51s	9.73s
SET 100w	10.79s	11.81s	7.30s	9.59s

Intel(R) Xeon(R) CPU E5-2630 v2 @ 2.60GHz 24/3.2T FIO/256G内存

注：Pika先采用20线程而非24线程测试是因为和线上标配一致，故尝试先观察下和当前线上的差别

数据出仓的业务视角对比数据导入量：515510873

* redis_batch_size 8000

* redis_key_expire_seconds 86400

* redis_commit_sleep_time 5

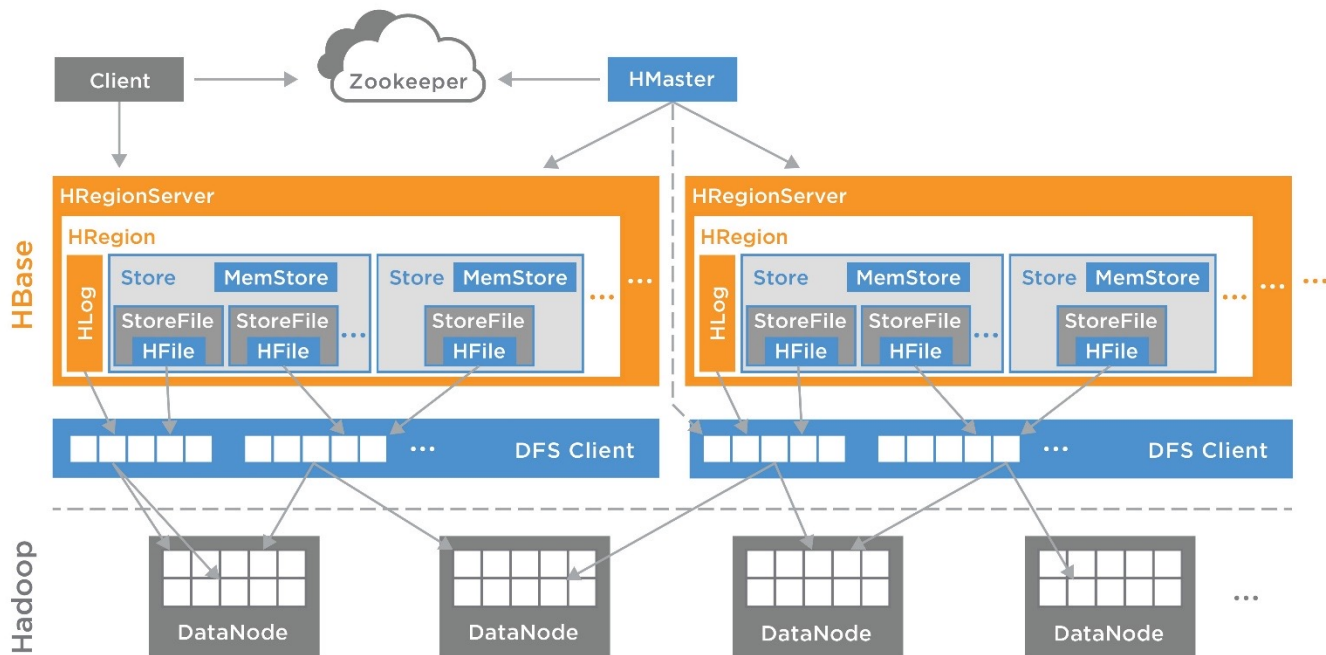
Pika : 04:33:05 Tendis : 04:34:43

Tendis存储版 vs HBase

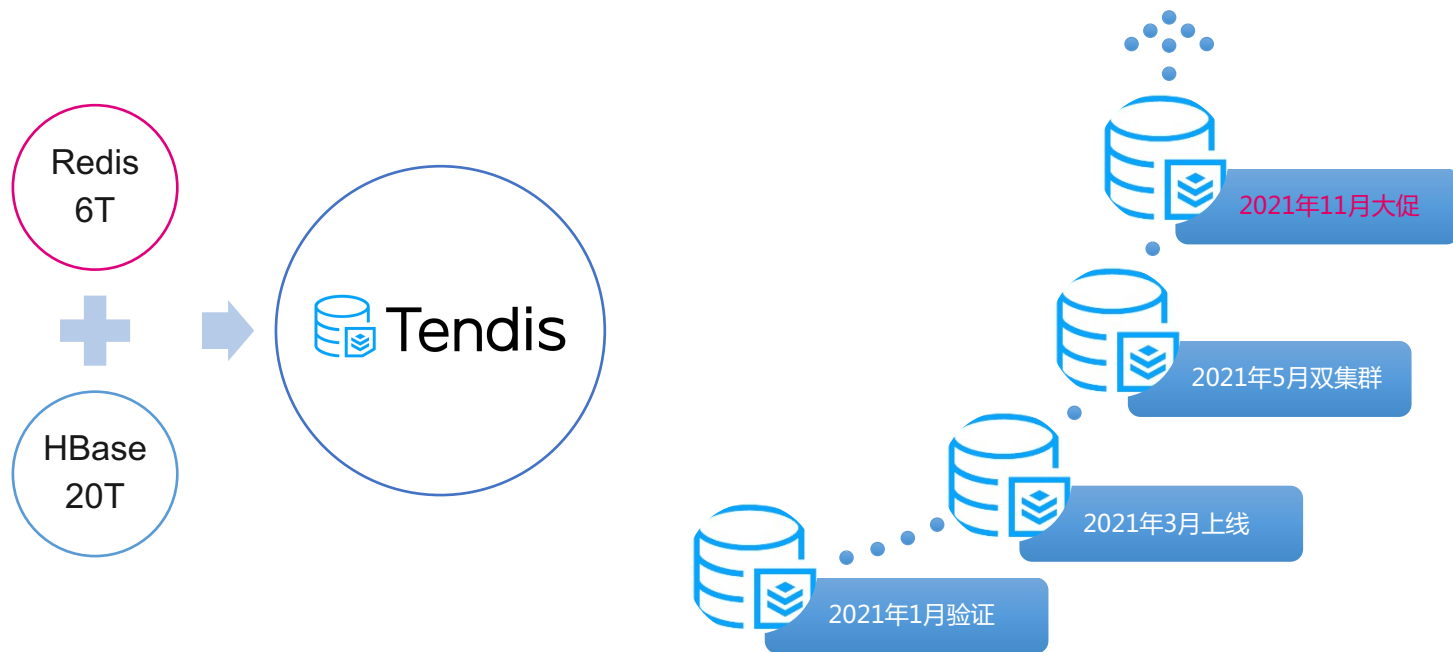
优点：极强的可扩展性 (PB)

缺点：complex (部署、运维、性能调优) hot region、cache；数据结构单一

示例 rowkey：获取一级部类A01在活动 202102211567 中app的2021年2月21日19时的指标数据A01FC765112201202HOURappSP2021022119



Tendis存储版@VIP——维表关联和指标存储



实时数仓Flink + Tendis

AI平台

任务

模型服务

模型训练

工作区

自定义发布

Kafka监控

即将下线

SQL

Flink

Storm

Spark

元数据管理
新建Flink元数据,查看DDL

2 SQL 管理
Flink SQL实现业务逻辑处理

3 上线管理
配置任务级别、机房、资源等

SQL管理

历史版本对比

保存

语法校验

执行计划

Tr 格式化

```
1 CREATE TEMPORARY TABLE tendis(  
2   `key` STRING,  
3   `field` STRING,  
4   `value` STRING  
5 ) WITH (  
6   'connector' = 'redis',  
7   'redis-mode' = 'cluster',  
8   'cluster-nodes' = 'gd16-xxx-xxx-m01.vip.vip.com:6379,gd16-xxx-xxx-m02.vip.vip.com:6380',  
9   'command' = 'HSET',  
10  'key.ttl' = '47304000'  
11 );  
12  
13 insert into tendis  
14 select  
15   `key`,  
16   col as field,  
17   case right(col,12)  
18     when '_last_update' then FROM_UNIXTIME(cast(val as bigint)/1000)  
19     else val  
20   end as `value`  
21 from hive.temp_vipflink.kafka_source_dataeye_uv2 /*+ OPTIONS('properties.group.id'='topi_uv_write_tendis', 'scan.startup.mode'='group-offsets') */  
22 ,lateral table(universal_json_array_parse(`value`, 'column', `value`)) as T1(col, val)  
23 ,lateral table(string_split(sinkType, ',')) as T2(sink)  
24 where sink='dataeye-tendis';
```

实时数仓Flink + Tendis

5 --大促日大数据销售计算

6 CREATE temporary view hive.temp_dw.view_trd_order_act_big_index as

7 select

8 order_id,

9 order_goods_id,

10 order_sn,

11 parent_sn,

12 user_id,

Overview

Exceptions

TimeLine

Checkpoints

Configuration

13 --more columns

14 cast('cgu' as float)

15 is_haitao

16 FROM hive.vipfli

17 /*+ OPTIONS('pr

18 where

19 (dim_source is

20 -- more

21 and normal_a

23 create temporan

24 create temporan

25 create temporan

26 create temporan

182 CREATE temporary TABLE hive.temp_vipflink.redis_realtime_topi_act_big_index (

183 rowkey STRING,

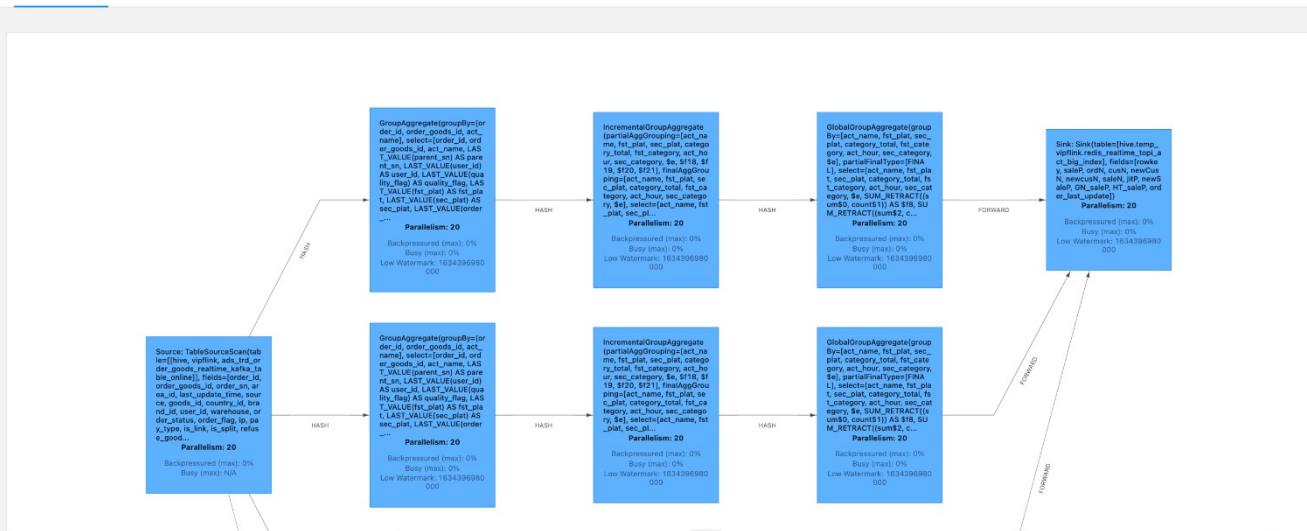
184 saleP STRING,

185 ordN STRING,

186 cusN STRING,

187 --more columns

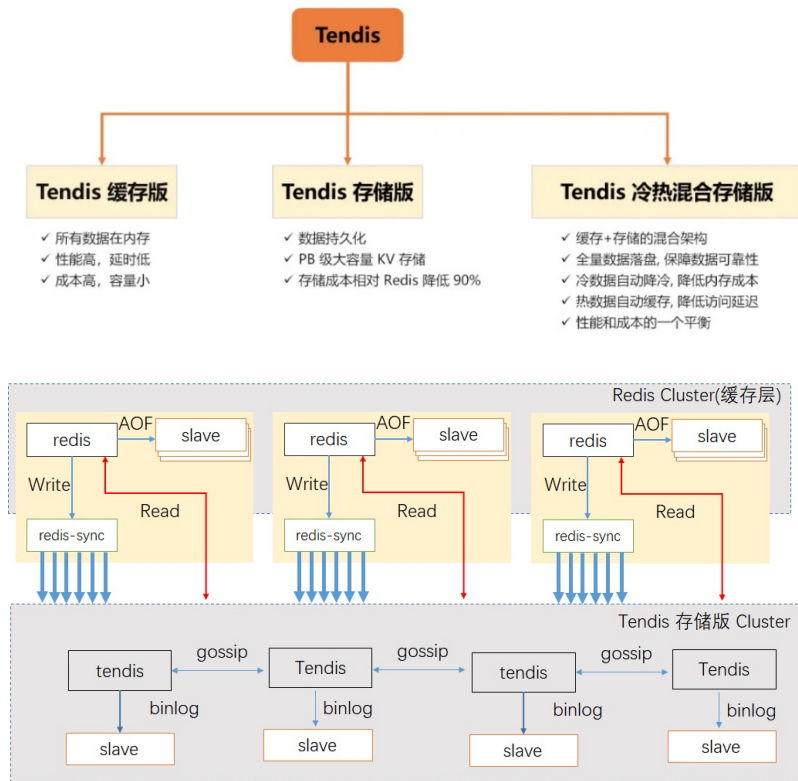
xx-m02.vip.vip.com:6380',



Name	Status	Bytes Received	Records Received	Bytes Sent	Records Sent	Parallelism	Start Time	Tasks
Source: TableSourceScan(table=[[hive, vipflink, ads_trd_order_g...	RUNNING	0 B	0	155 GB	544,944,600	20	2021-09-	20
GroupAggregate(groupBy=[order_id, order_goods_id, act_name]...	RUNNING	38.9 GB	136,236,150	694 GB	1,372,157,488	20	2021-09-	20
IncrementalGroupAggregate(partialAggGrouping=[act_name, fst...	RUNNING	694 GB	1,372,157,488	61.8 GB	201,209,808	20	2021-09-	20
GlobalGroupAggregate(groupBy=[act_name, fst_plat, sec_plat, c...	RUNNING	61.9 GB	201,209,808	4.62 GB	21,499,681	20	2021-09-	20

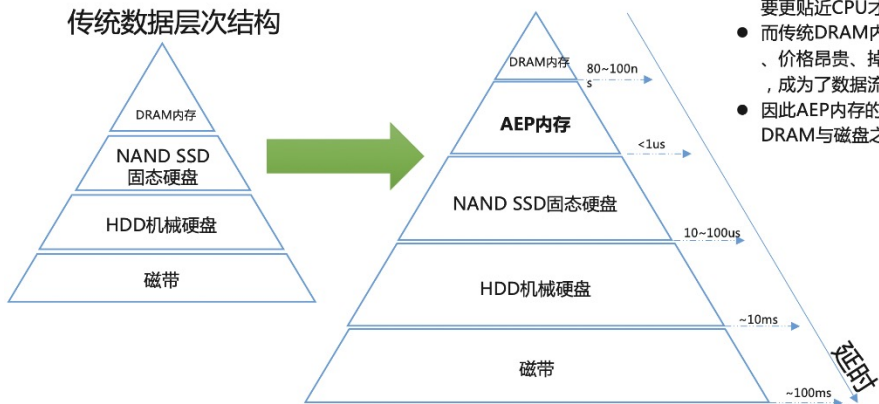
期待Tendis

- 更多特性的开源，如Tendis冷热混合存储



- AEP支持；Redis on AEP (memKeyDB)

为什么会有AEP内存



- 随着算力不断提高，大量数据需要更贴近CPU才能完成工作
- 而传统DRAM内存受限于容量小、价格昂贵、掉电丢数据等特征，成为了数据流转瓶颈。
- 因此AEP内存的出现补齐了DRAM与磁盘之间的需求缺口。



感谢聆听

THANKS!