

实时计算在贝壳的实践

The Practices of Real-Time Computing in Beike

刘力云 Liu Liyun

资深工程师 Senior Engineer

FLINK FORWARD # ASIA

实时即未来 # Real-time Is The Future

**FLINK
FORWARD**

Contents

目录

01 业务规模与演进

Application Scale & Evolution

02 Hermes实时计算平台介绍

Introduction of Hermes Real-Time Computing Platform

03 未来发展与规划

Future Plan

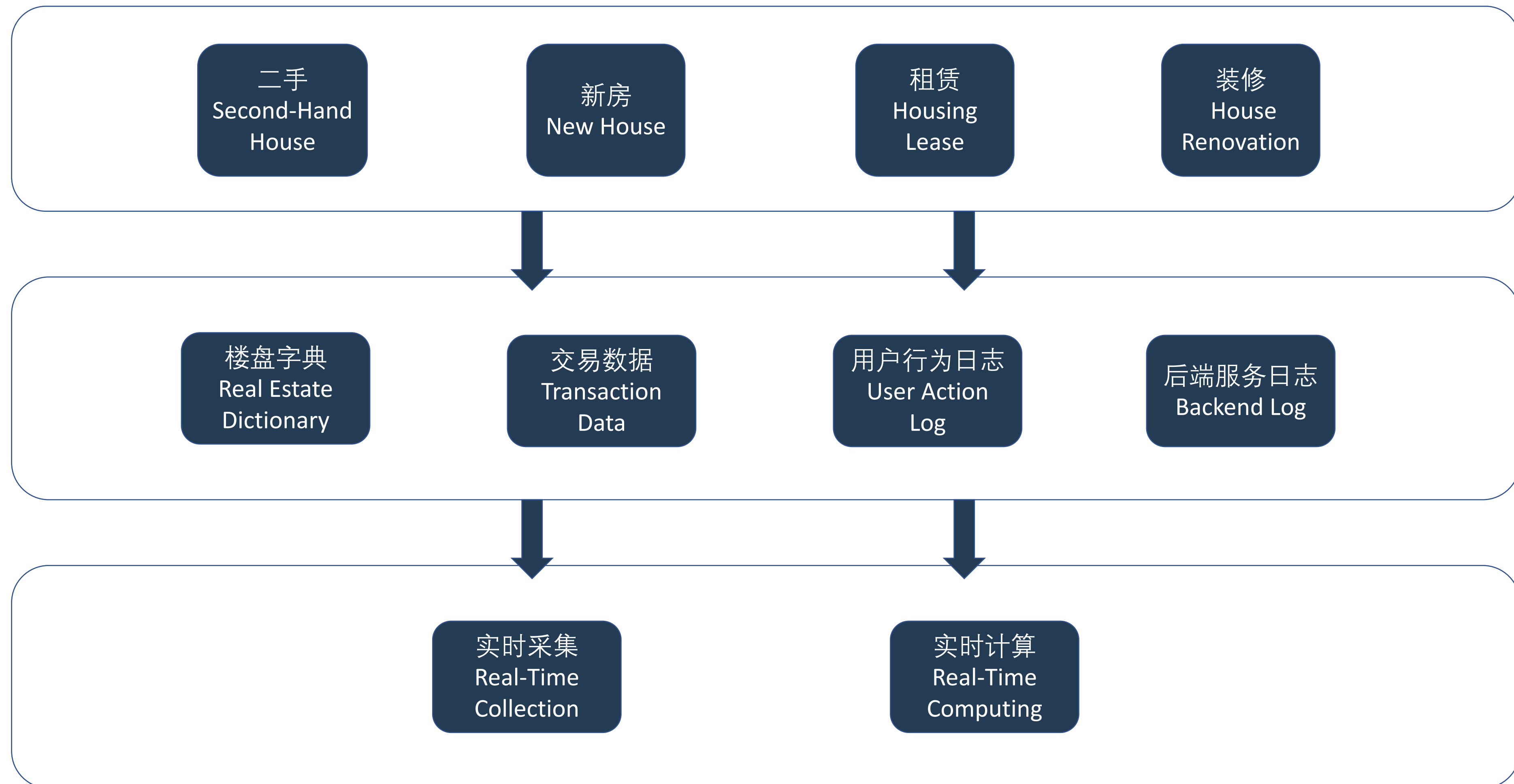
业务规模与演进

Application Scale & Evolution

01

业务场景

Business Scenarios



发展历程

Develop Steps

Hermes Platform

Based on Spark Streaming

2018.03

SQL V1

Based on Spark Structured Streaming

2018.10

Turn to Flink

Support Flink Jobs

2019.05

SQL V2

Based Flink SQL

2019.08

Real-Time Data Warehouse

Based on SQL V2

2019.11

应用规模

Application Scale

30

支持的项目
Supported
Projects

400

流计算任务数
Streaming Jobs

80 Billions

每日处理消息条数
Processed
messages per
day

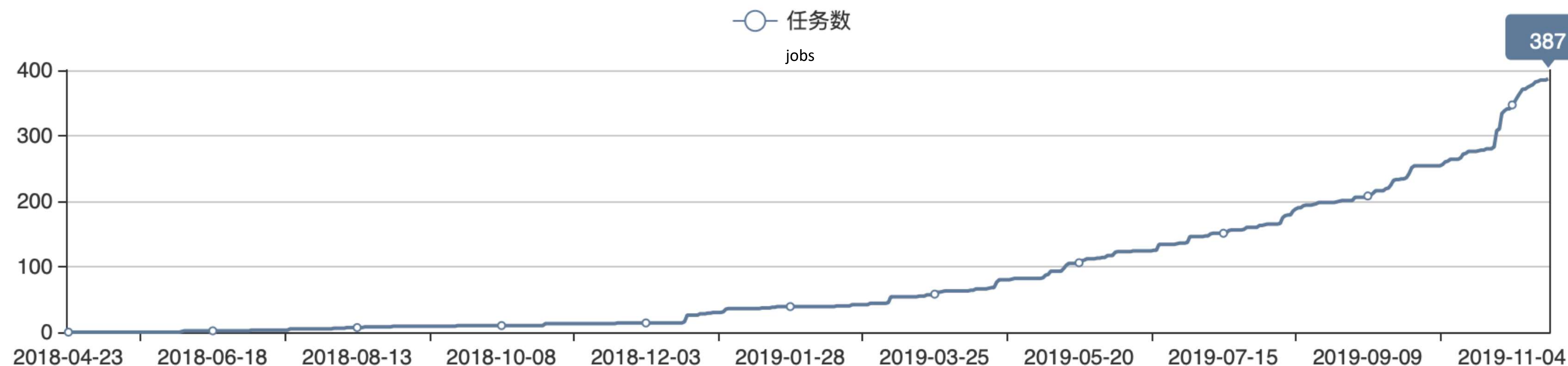
支持的项目

Supported Projects



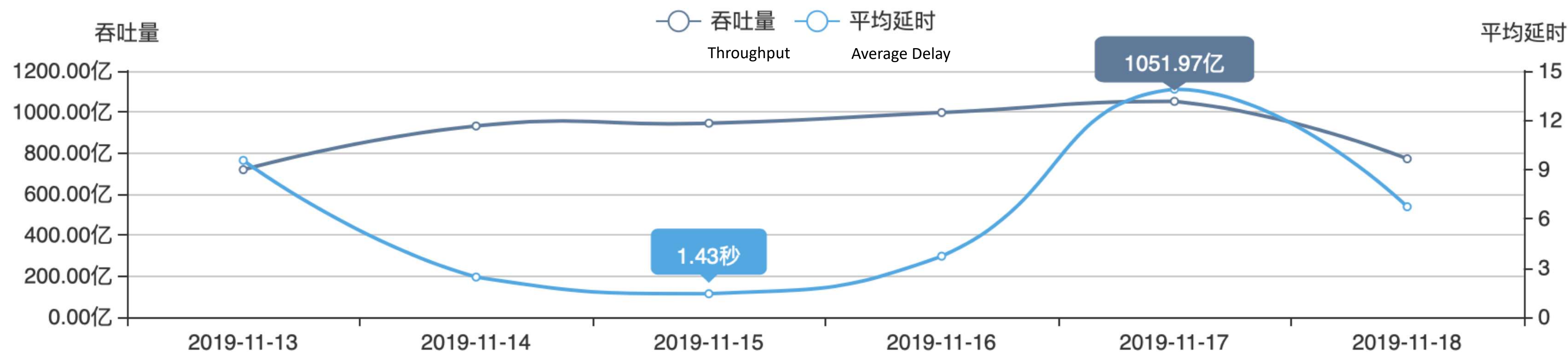
任务增长趋势

Jobs Continues Growing



平台数据统计

Data Statistics of the Real-Time Computing Platform



Hermes实时计算平台介绍

Introduction of Hermes Ream-Time Computing Platform

02

平台概览

Hermes Overview

编辑、发布
Edit、Deploy
启停
Start/Stop
版本管理
Version
Management
监控报警
Monitoring

任务托管
Job Management

Java
Scala
Python

多种开发语言
Programming
Languages

自定义任务
User Developed
Jobs
模板任务
Template Jobs
场景任务
SQL Jobs

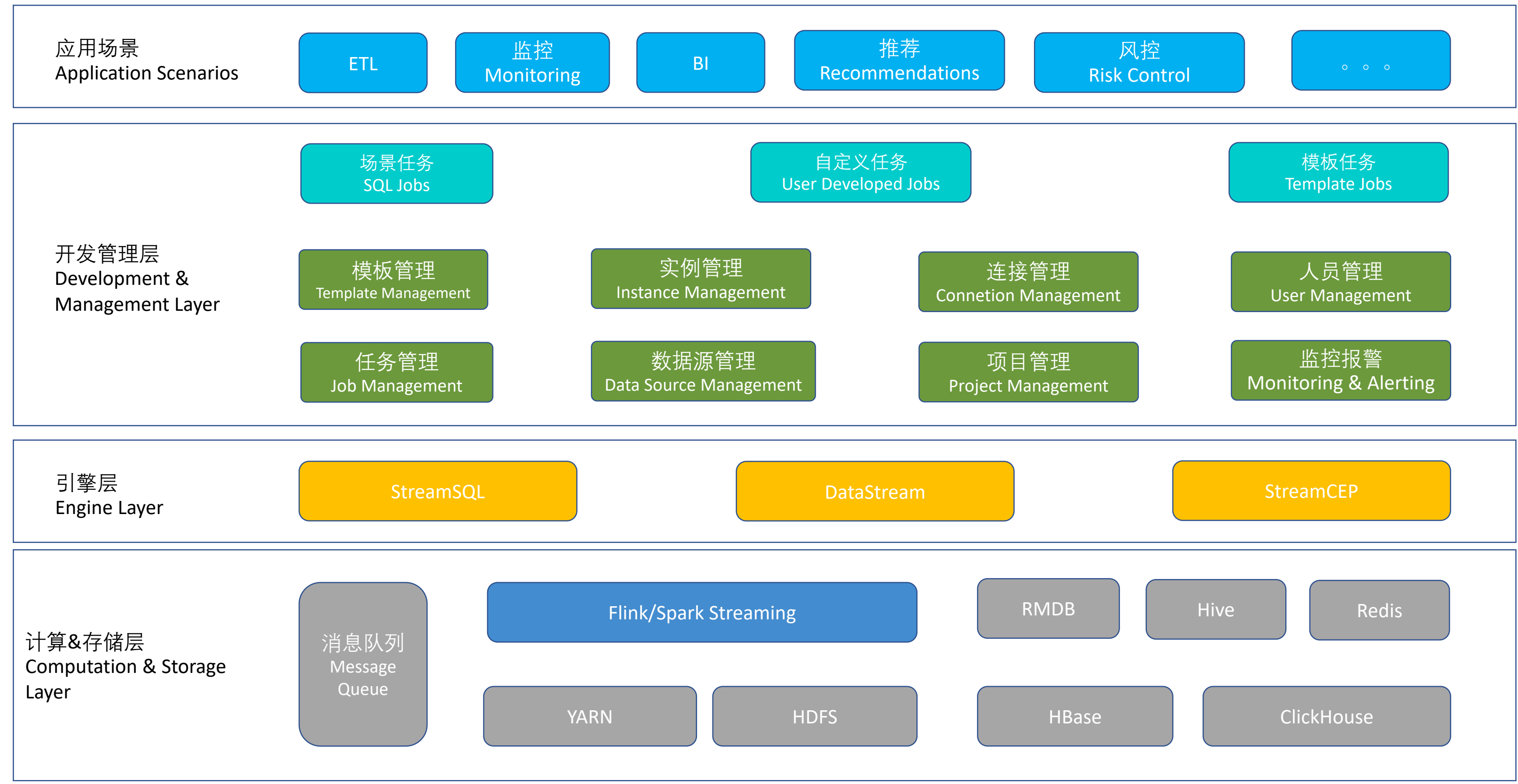
多种任务类型
Job Types

专有队列
Private Resource
Queue
公共队列
Public Resource
Queue

资源隔离
Resource Isolation

Hermes架构

Infrastructure of Hermes



SQL V1编辑器

SQL V1 IDE

任务管理 / sql编辑

线上环境

帮助文档

保存

返回

数据源(source)

kafka

操作符
(operator)

filter

union

目标源(sink)

kafka

console

druid

source_kafka_0
(IM消息)

operator_filter_0
(正常消息获取)

operator_filter_2

sink_kafka_0

operator_filter_1
(兼容客户端错误格式)

operator_filter_3

sink_kafka_1

基本信息

高级设置

配置信息

元素名称: operator_filter_1

兼容客户端错误格式

```
* SELECT      send_time,from_unixtime(cast(send_time/1000 as int),'yyyy-MM-dd\'T\'HH:mm:ss.SSS+08:00') as client_timestamp,from_ucid,to_ucid,get_json_object(get_json_object(msg_attr,'$biz_src'),'$app_data') as biz_src,from_unixtime(cast(send_time/1000 as int),'yyyy-MM-dd\'T\'HH:mm:ss.SSS+08:00') as client_timestamp,from_ucid,to_ucid,get_json_object(get_json_object(msg_attr,'$biz_src'),'$app_data') as biz_src
* FROM        source_kafka_0
WHERE         get_json_object(get_json_object(get_json_object(msg_attr,'$biz_src'),'$app_data'),'$add') is not null and get_json_object(get_json_object(get_json_object(msg_attr,'$biz_src'),'$app_data'),'$add')
```

SQL V2编辑器

SQL V2 IDE

查看Flink执行计划

语法校验

折叠sql

格式化

退出全屏

44

CREATE VIEW tmp_join_1 AS

45

SELECT

46

a.`type`,

47

a.`database`,

48

a.`table`,

49

a.`timestamp`,

50

a.`business_id`,

51

a.`process_id`,

52

a.`app_id`,

53

a.`job_id`,

54

a.`job_definition_key`,

55

a.`complete_time`,

56

b.`company_type`,

57

b.`sub_trade_product`,

58

b.`business_code`,

59

c.`ziduan_key` as `jine_key`,

60

c.`index_name`,

61

case when d.`unit` = '万元' then cast(cast(regexp_replace(d.`value`,` ','`) as double) * 10000 as varchar) else regexp_replace(d.`value`,` ','`) end as `jine`,

62

e.`ziduan_key` as `yinhang_key`,

63

e.`index_name` as `bank_key`,

64

g.`bank_name` as `bank`

65

FROM devide_1 a

66

LEFT JOIN businesses b

67

ON concat(a.`business_id`,`_`,a.`process_id`) = b.rowKey

68

LEFT JOIN bank_config c

69

ON concat(a.`app_id`,`_`,b.`sub_trade_product`,`_`,a.`job_definition_key`,`_`,`面签商贷金额`) = c.rowKey

70

LEFT JOIN job_vars d

71

ON concat(a.`business_id`,`_`,a.`job_id`,`_`,c.`ziduan_key`) = d.rowKey

72

LEFT JOIN bank_config e

73

ON concat(a.`app_id`,`_`,b.`sub_trade_product`,`_`,a.`job_definition_key`,`_`,`面签商贷总行`) = e.rowKey

74

LEFT JOIN job_vars f

75

ON concat(a.`business_id`,`_`,a.`job_id`,`_`,e.`ziduan_key`) = f.rowKey

76

LEFT JOIN bank_info g

77

ON regexp_replace(f.`value`,` ','`) = g.rowKey

78

WHERE b.`contract_source` <> '2' and c.`ziduan_key` <> '' and c.`ziduan_key` is not null

79

and concat(a.`app_id`,`_`,b.`sub_trade_product`) not in

80

('130100_10002012','210600_10002014','350600_10002001','420500_10002001','350600_10004005','420500_10004005','440100_10002016','441900_10002018','441900_10002019') and

81

concat(a.`app_id`,`_`,b.`company_type`) not in ('230100_29','360400_64','440800_15','500000_171');

82

— 批贷

83

CREATE VIEW tmp_oin_2 AS

sql编辑器

根据配置自动生成DDL

基本信息

高级设置

配置信息

添加输入源

source_base_1

source_circle

添加维表

businesses

bank_config

job_vars

bank_info

添加输出源

sink_clickhouse

sink_circle

表名: source_base_1

查看元数据

类型: kafka

* 表名: source_base_1

* 连接: common-log-02

* TOPIC: tradeData-flink-base

* GROUP: hermes_prod_sqlv2_857

消费位置: 最新 最早

水印:

并行度: 1

描述: 表的描述信息

语法校验提示窗口

错误提示区域

SQL V2架构

SQL V2 IDE

前端 Front End

sql编辑器
SQL Editor

语法/语义校验
Grammar/Semantic
Validation

执行计划
Execution Plan

自动生成DDL
DDL Generation

任务调试
Job Debugging

后台 Back End

连接管理
Connection
Management

数据源管理
Data Source
Management

UDF管理
UDF Management

sql引擎
SQL Engine

antlr4

calcite

Flink Sql

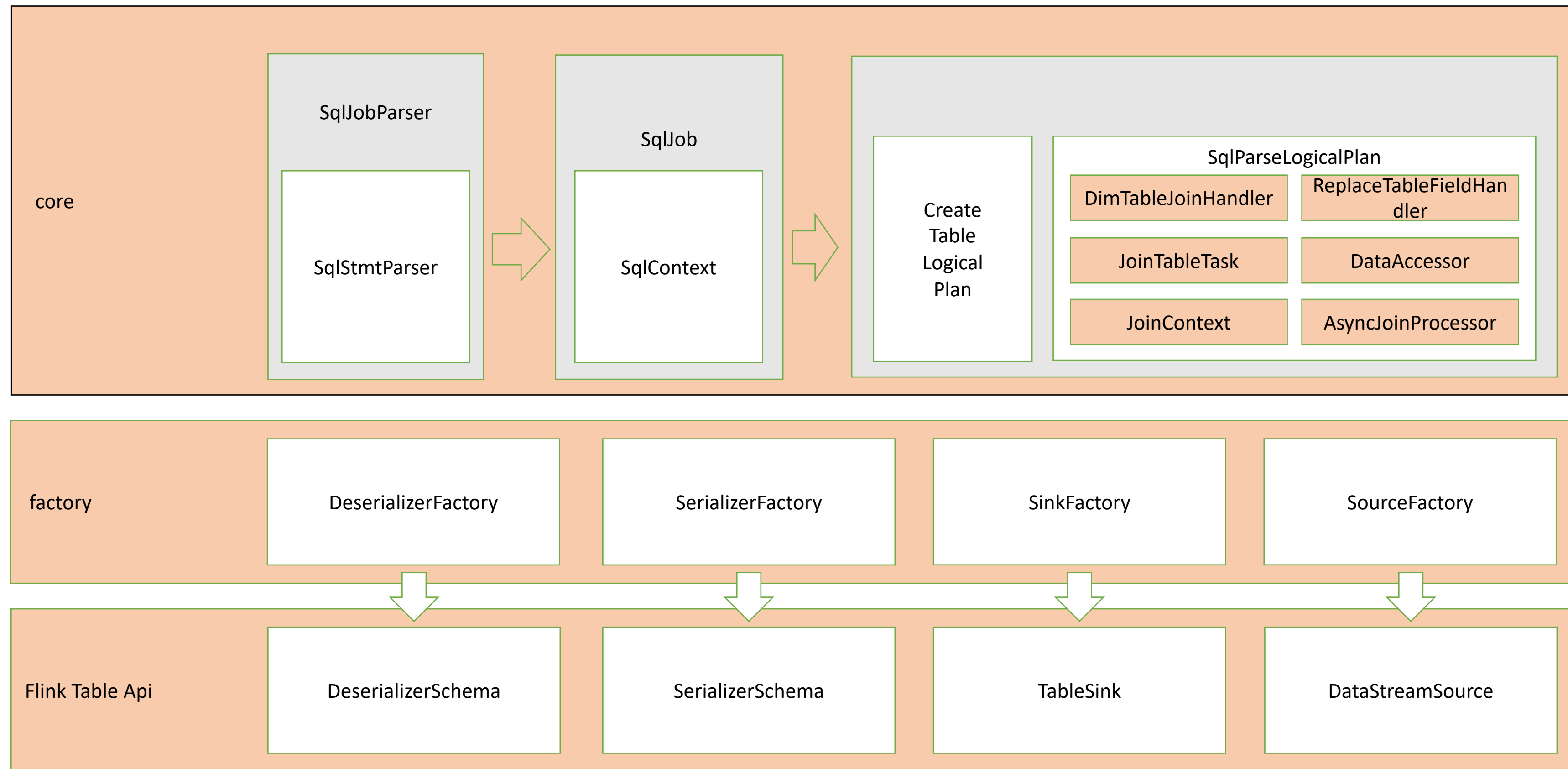
sql

Flink Jar

sql引擎
SQL Engine

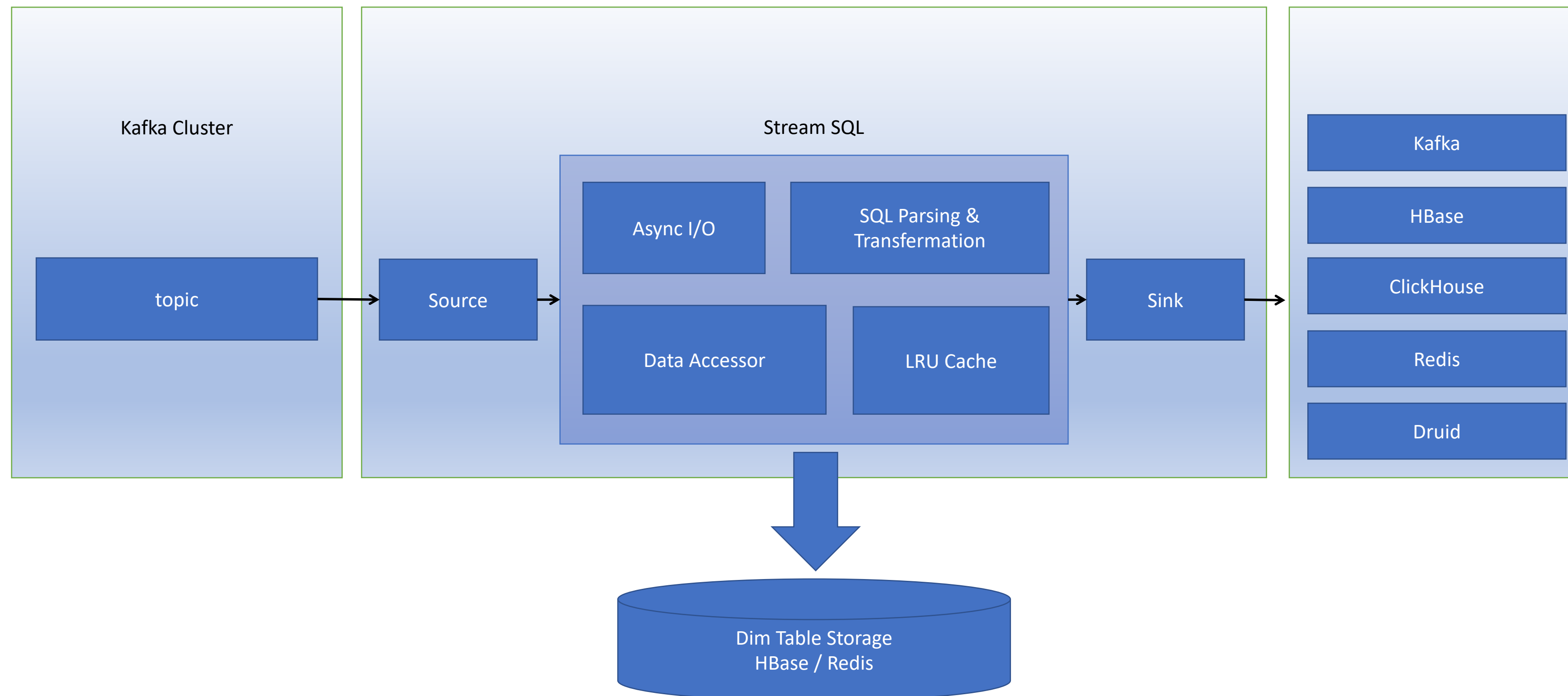
SQL引擎整体架构

Infrastructure of SQL Engine



维表关联

Stream Table Joins Dim Table



丰富的内置函数

Plentiful Functions Embedded

时间函数
Time Functions

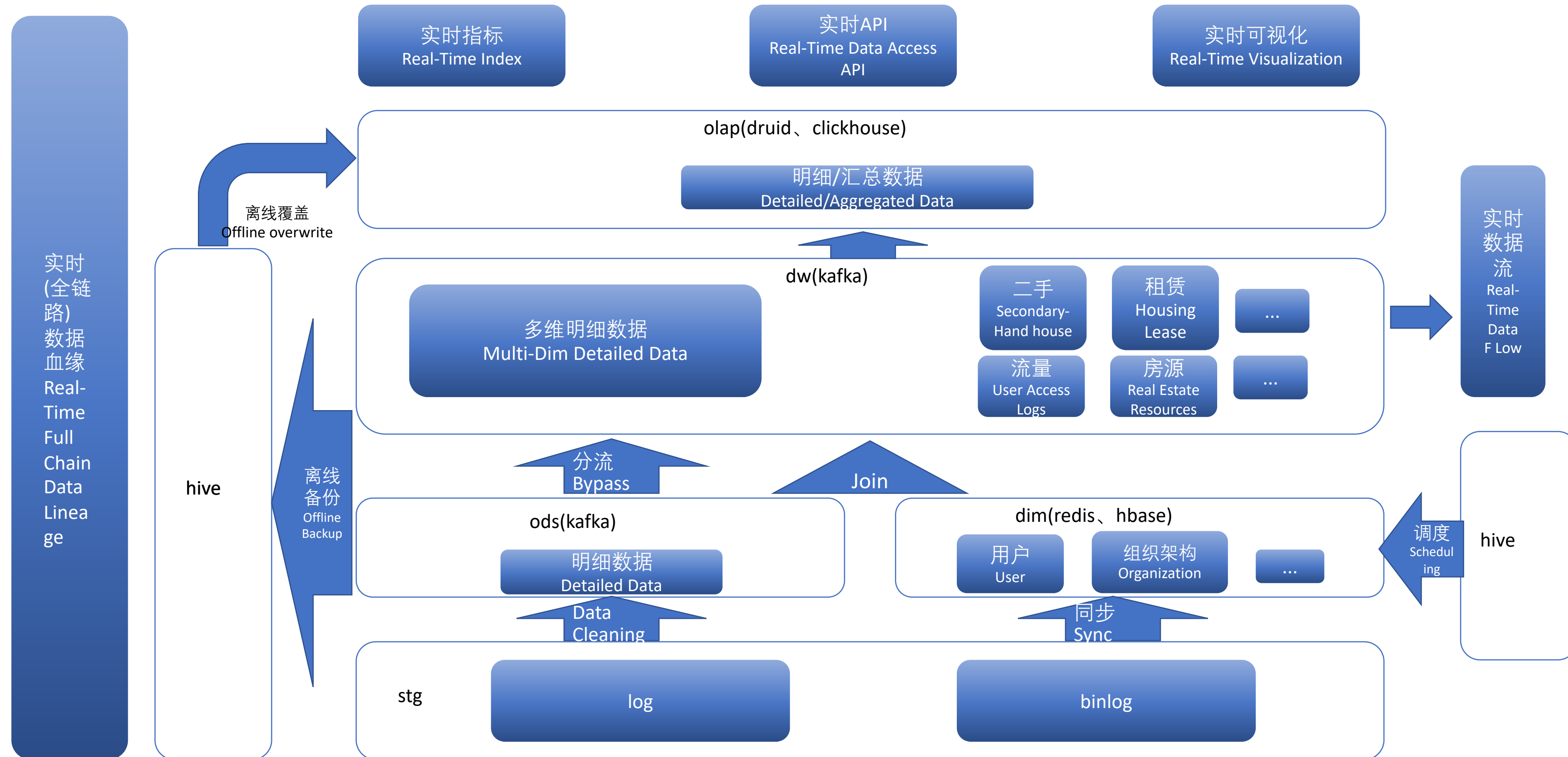
集合函数
Set Functions

Json处理
Json Functions

字符串函数
String Functions

实时数仓整体架构

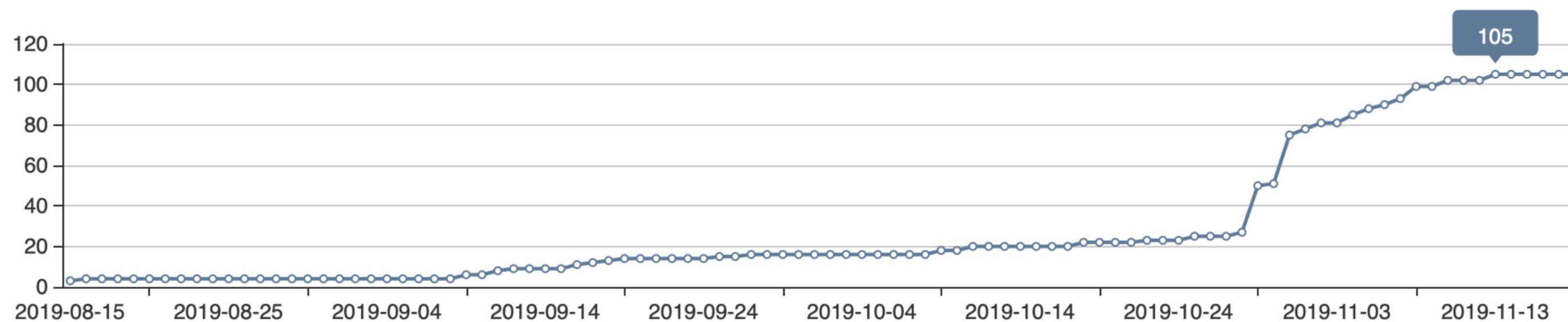
Infrastructure of Real-Time Data Warehouse



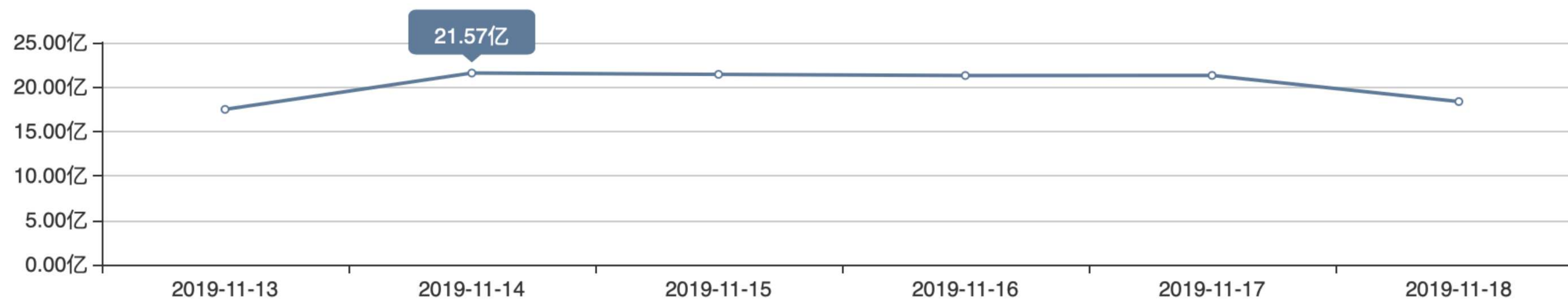
实时数仓数据统计

Data Statistics of Real-Time data warehouse

任务数



吞吐量



实时数仓案例

Applications Implemented Real-Time data warehouse

01

交易平台

Deal Platform

02

经纪人行程量

Agent Performance Monitoring

03

实时用户画像

Real-Time User Profile Computing



监控报警

Monitoring & altering

flink_metrics

启动任务 修复实例 停止任务 编辑 删除

任务信息

任务实例

任务监控

应用监控

kafka offset

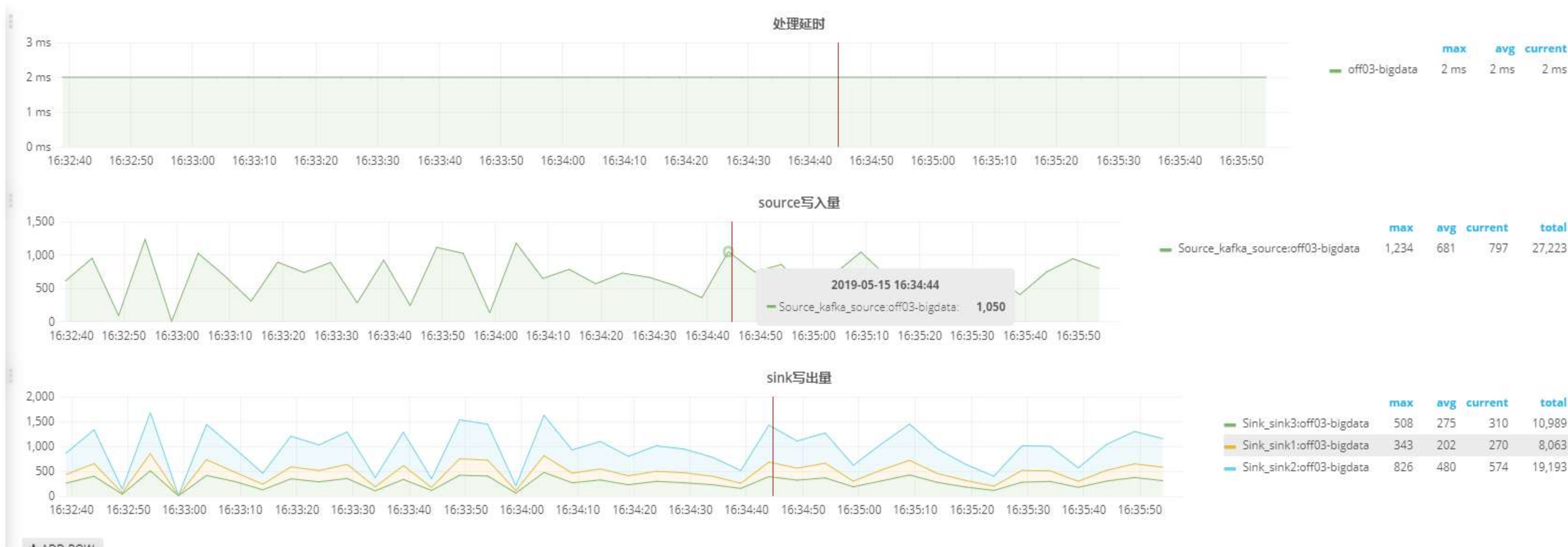


hermes.flink.task.metrics



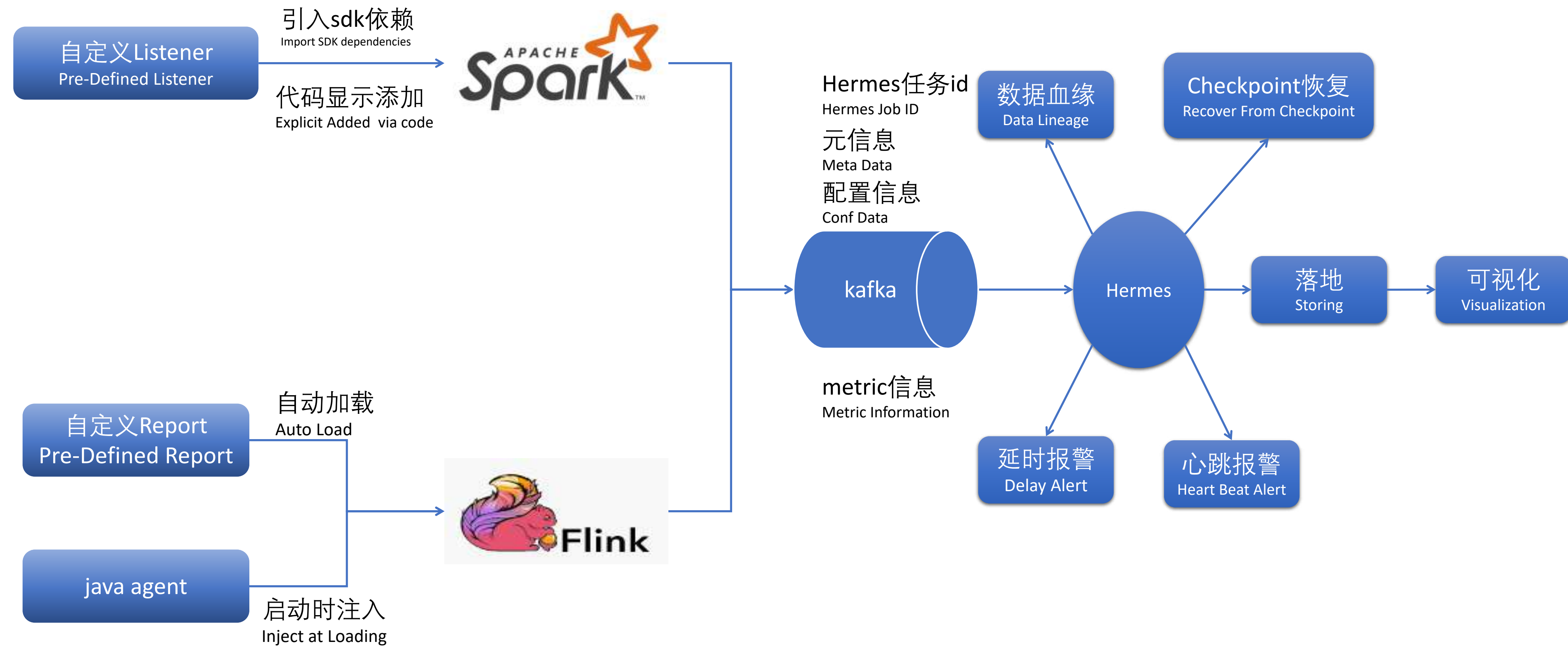
Zoom Out 5月 15, 2019 16:32:37 to 5月 15, 2019 16:35:57

taskManager host All source All sink All



监控报警架构

Infrastructure of Monitor & altering



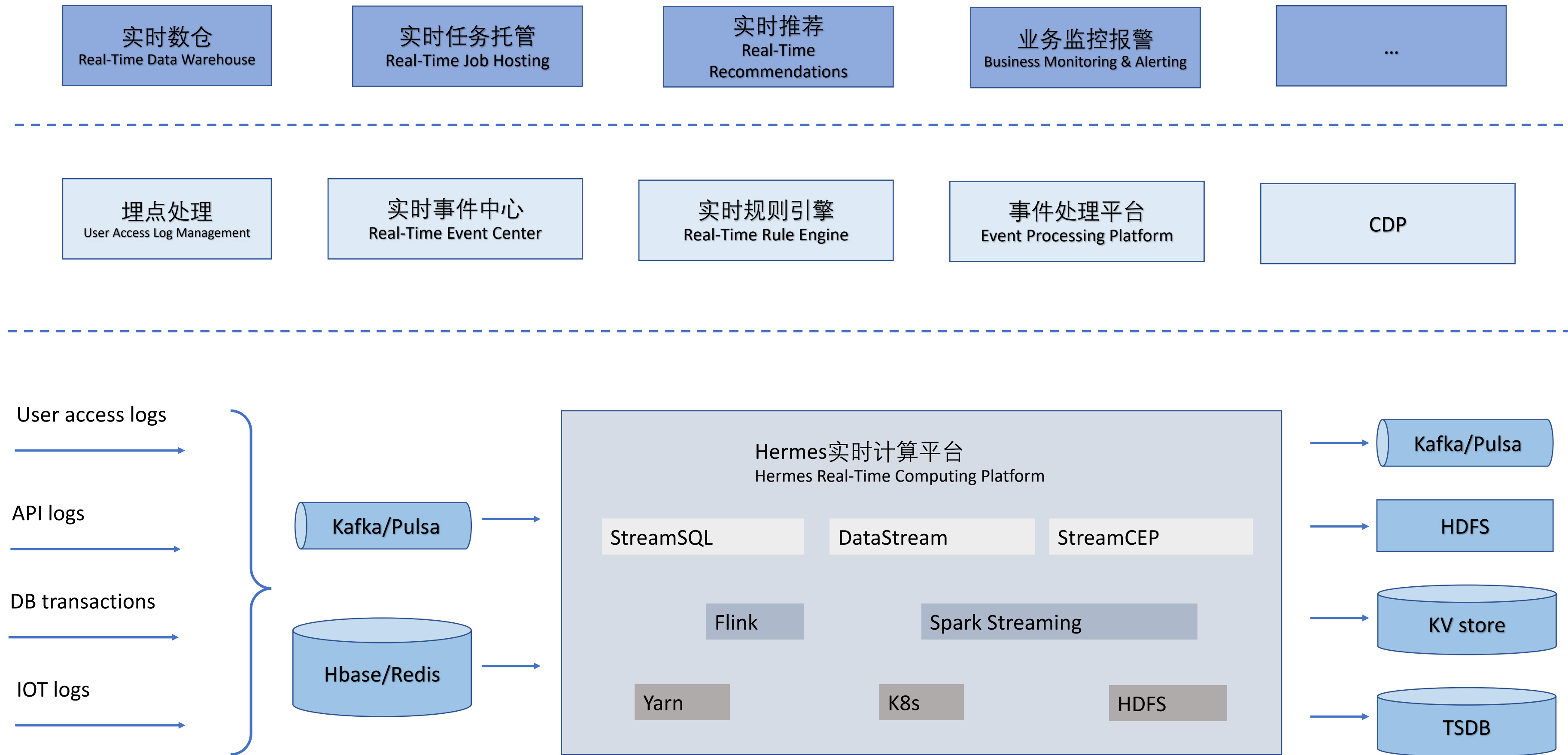
未来发展与规划

Future Plan

03

整体架构

Infrastructure of the Platform



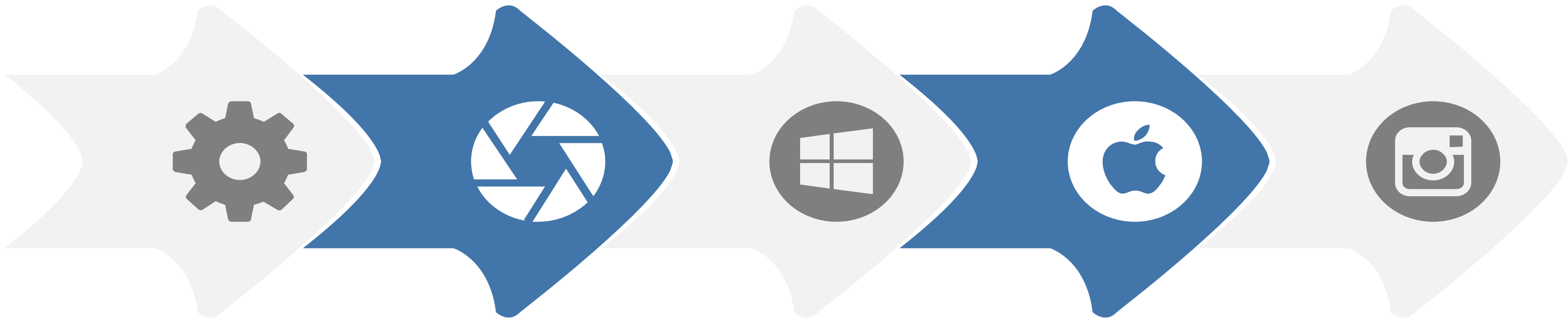
未来发展

Future Plan

资源动态分配
Auto Scaling

事件处理平台
Event Processing Platform

KAPPA模式探索
KAPPA Implementation



实时规则引擎
Real-Time rule engine

用户数据平台
Customer Data Platform

THANKS