

The SACC logo is rendered in a bold, white, sans-serif font with a blue glow effect. It is positioned in the upper right quadrant of the image, above the main conference title. The background features a blue wireframe architectural design with a perspective view of a city skyline and a large gear-like structure at the bottom left.

2021 中国系统架构师大会

SYSTEM ARCHITECT CONFERENCE CHINA 2021

数字转型 架构重塑

IT168.com

ChinaUnix

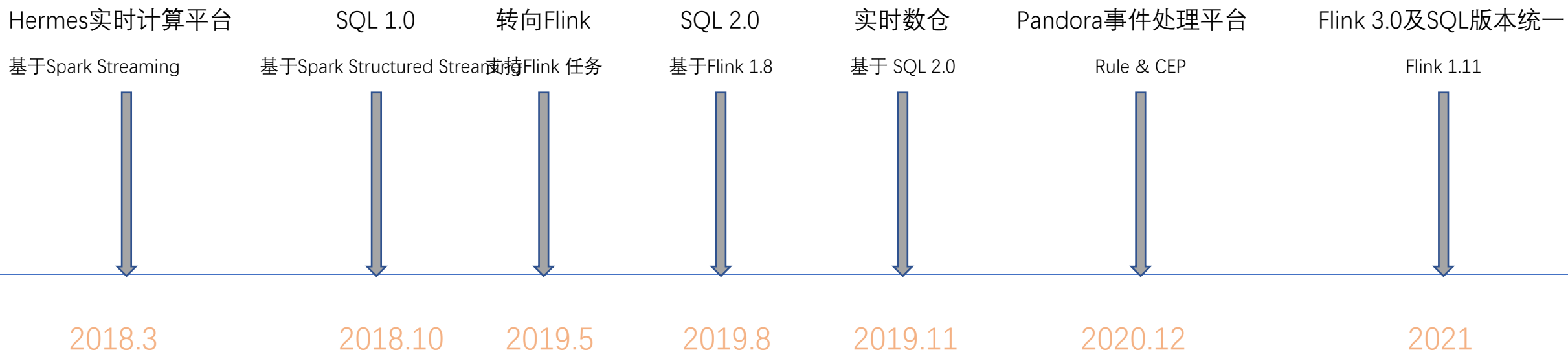
ITPUB

云上会议 网络直播 | 2021.5.20-2021.5.22

贝壳实时计算平台建设实践

- 发展历程
- Hermes实时计算平台介绍
- 实时数仓建设
- Pandora事件处理平台建设
- 未来规划

发展历程



Hermes实时计算平台介绍-平台概览

生命周期管理

监控报警

任务托管

Java

Scala

Python

多语言支持

自定义任务

模板任务

场景任务

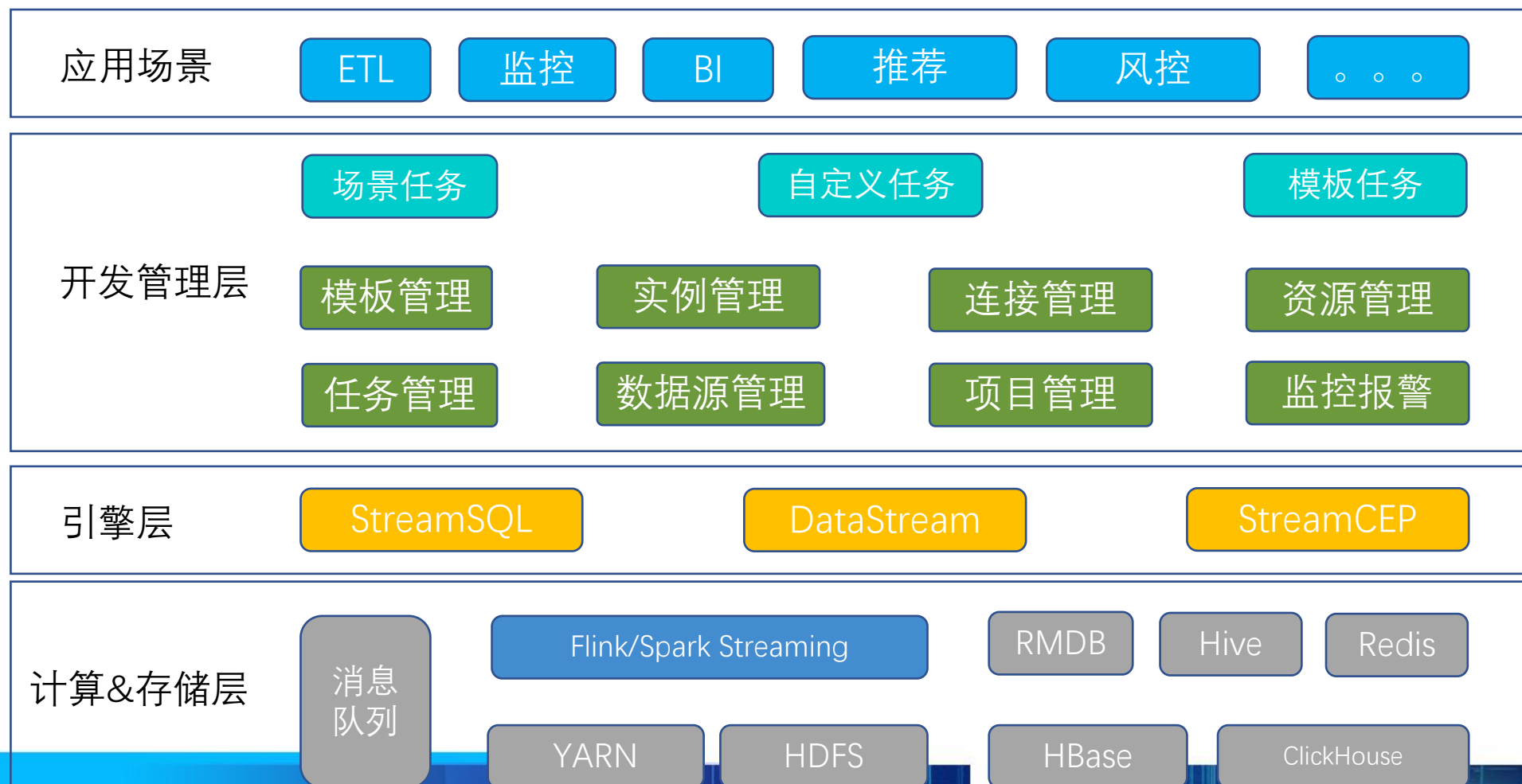
多任务类型支持

专有队列

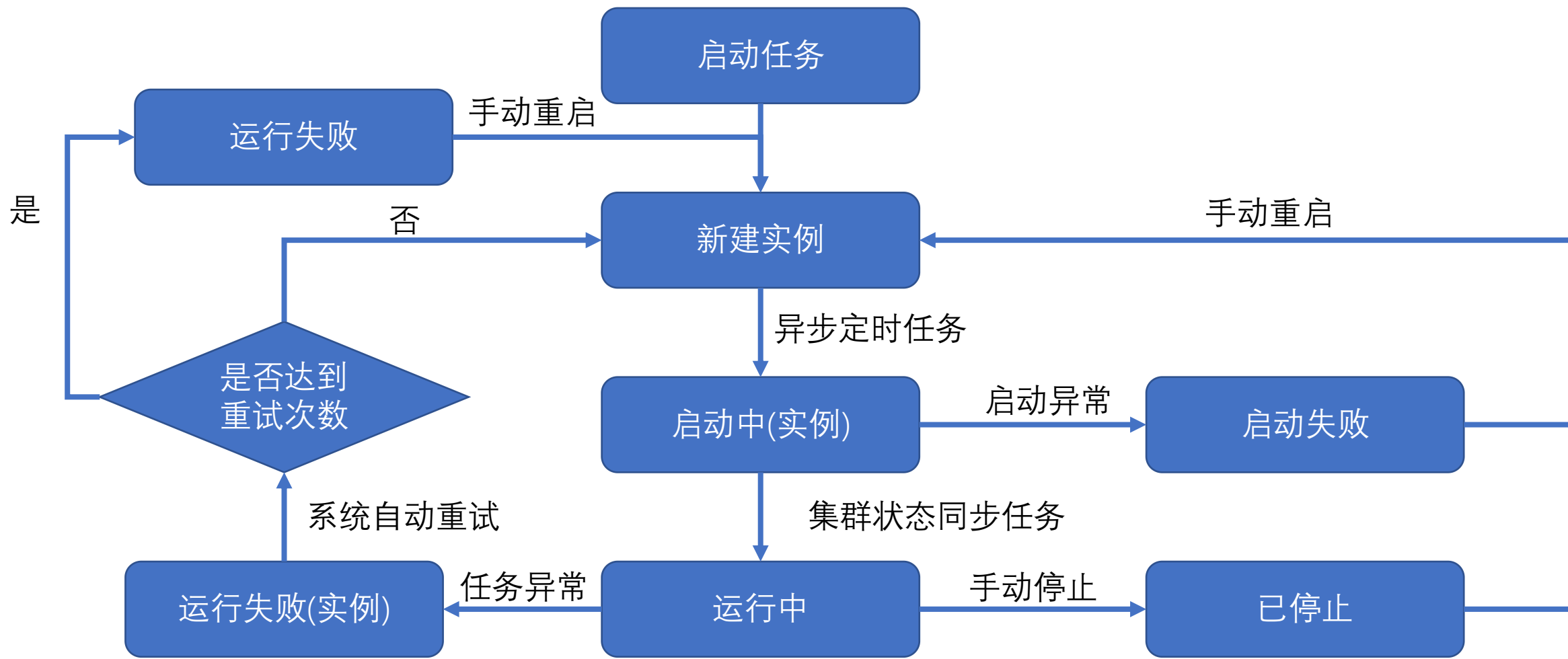
公共队列

资源隔离

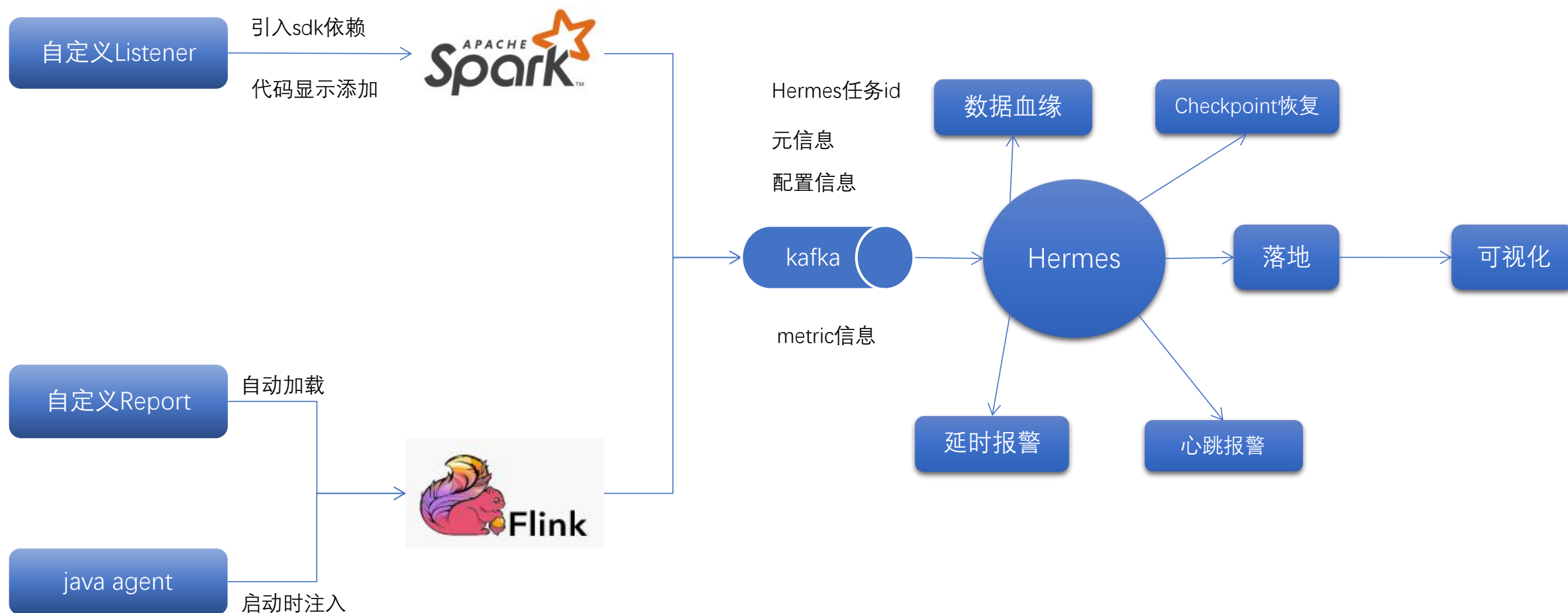
Hermes实时计算平台介绍- 平台架构



Hermes实时计算平台介绍- 任务生命周期



Hermes实时计算平台介绍-监控报警



Hermes实时计算平台介绍-监控报警



实时数仓建设 - 能力

元数据管理

连接管理
表管理

数据分层

ODS
DWD
DWS
OLAP

SQL

标准SQL语法
维表关联
图形化开发环境
丰富的内置函数
UDF支持

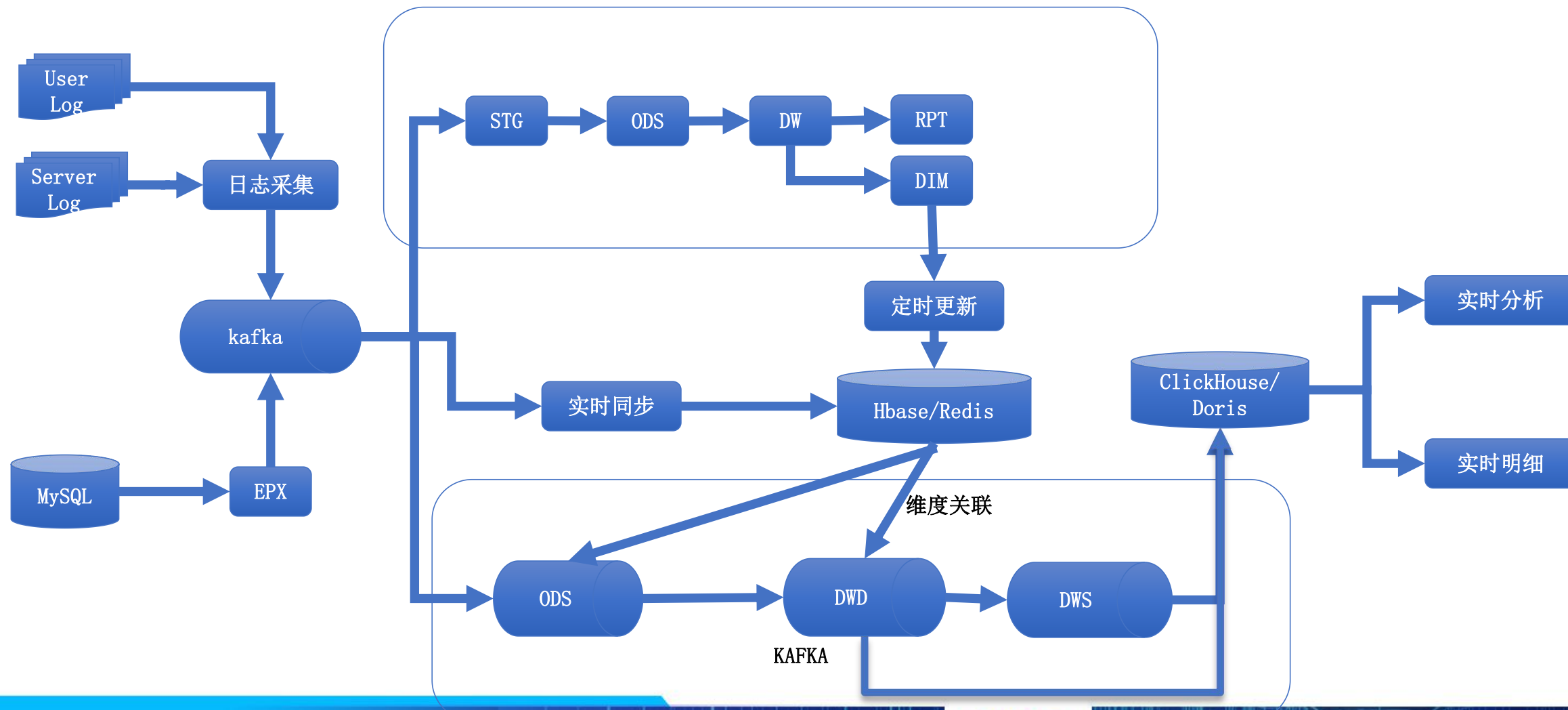
数据血缘

图形化展示
链路分析
实时数据流显示

多源支持

Kafka
Hbase
Redis
Doris
ClickHouse
MySQL

实时数仓建设 - 架构



实时数仓建设 – SQL IDE

dw_dwd_

查看Flink执行计划

任务调试查看DDLSQL校验折叠文本回格式化工屏

```
1 CREATE view dwd_ AS
2 SELECT t1.id AS activity_id
3 ,t1.activity_name AS activity_name
4 ,t1.city_id AS city_code
5 ,t2.city_name AS city_name
6 ,t1.start_time AS activity_start_time
7 ,t1.end_time AS activity_end_time
8 ,t1.data_status AS activity_status
9 ,t1.status AS is_valid
10 ,t1.operator_id AS operate_ucid
11 ,t1.operator_name AS operator_name
12 ,t1.ctime AS create_time
13 ,t1.mtime AS update_time
14 ,t1.focus_id AS focus_id
15 ,t2.focus_name AS focus_name
16 ,t2.start_time AS focus_start_time
17 ,t2.end_time AS focus_end_time
18 ,t2.sign_start_time AS focus_sign_start_time
19 ,t2.sign_end_time AS focus_sign_end_time
20 ,t1.src_db AS src_db
21 ,t1.record_type_code AS record_type_code
22 ,t1.record_time AS record_time
23 FROM ods_l t1
24 LEFT JOIN dwd_comm_raw_nh_focus_info_rt_si t2
25 ON CAST(t1.focus_id AS varchar) = t2.rowKey;
26
27 INSERT INTO dwd_comm_raw_nh_focus_activity_info_rt_si( activity_id ,activity_name ,city_code ,city_name ,activity_start_time
28 ,activity_end_time ,activity_status ,is_valid ,operate_ucid ,operator_name ,create_time ,update_time ,focus_id ,focus_name
29 ,focus_start_time ,focus_end_time ,focus_sign_start_time ,focus_sign_end_time ,src_db ,record_type_code ,record_time )
30 SELECT activity_id
31 ,activity_name
32 ,cast(city_code AS int) AS city_code
33 ,city_name
34 ,activity_start_time
```

使用文档语法手册发布配置保存草稿返回详情

基本信息高级设置配置信息

* 所属项目实时数仓开发

* 任务名称dwd_co

任务描述新房

负责人

* 所属队列root

添加输入源

ods

添加维表

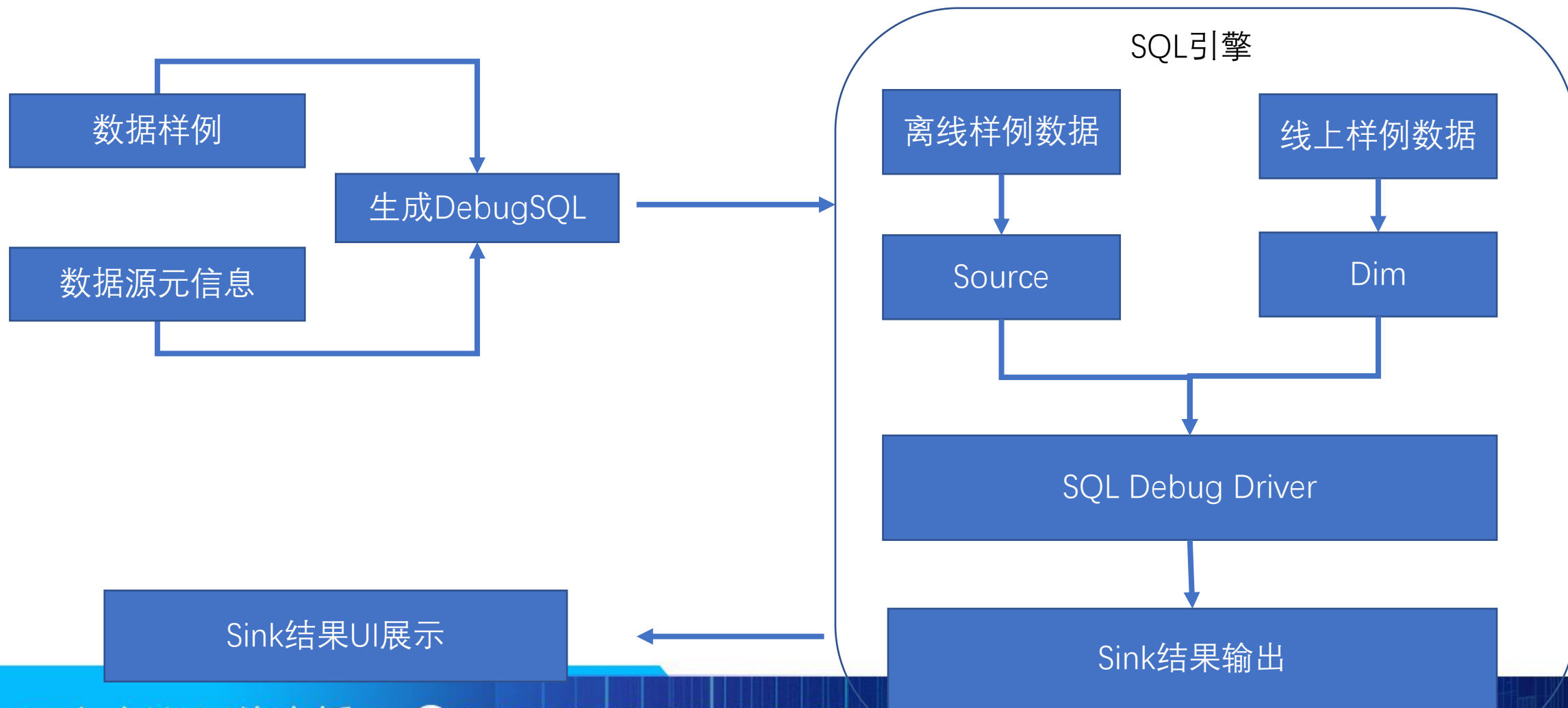
dwd_co

添加输出源

dwd_com

dwd_

实时数仓建设 – SQL任务调试



实时数仓建设 - SQL任务调试

任务管理 / SQL编辑器

dwd_com

使用文档

语法手册

发布配置

保存草稿

返回详情

任务调试

任务调试帮助文档

关闭调试

调试配置

开始调试

调试模式

自动模式

手动模式

数据展示

原始数据

表格展示

样例获取

最新

最早

5

条

调试进度

调试成功

上游数据源样例

ods_li

```
{
  "activity_name": "西安-1125期聚焦",
  "focus_id": 21,
  "operator_id": 1000000026622126,
  "end_time": "2020-12-03 07:59:59",
  "create_time": "2020-12-03 07:59:59"
}, {
  "activity_name": "西安-1125期聚焦",
  "focus_id": 21,
  "operator_id": 1000000020130889,
  "end_time": "2020-12-03 07:59:59",
  "create_time": "2020-12-03 07:59:59"
}, {
  "activity_name": "西安-1125期聚焦",
  "focus_id": 21,
  "operator_id": 1000000020130889,
  "end_time": "2020-12-03 07:59:59",
  "create_time": "2020-12-03 07:59:59"
}, {
  "activity_name": "西安-1125期聚焦",
  "focus_id": 21,
  "operator_id": 1000000026722753,
  "end_time": "2020-12-03 07:59:59",
  "create_time": "2020-12-03 07:59:59"
}, {
  "activity_name": "西安-1125期聚焦",
  "focus_id": 21,
  "operator_id": 1000000026722753,
  "end_time": "2020-12-03 07:59:59",
  "create_time": "2020-12-03 07:59:59"
}
```

下游调试结果输出

dwd_c

si

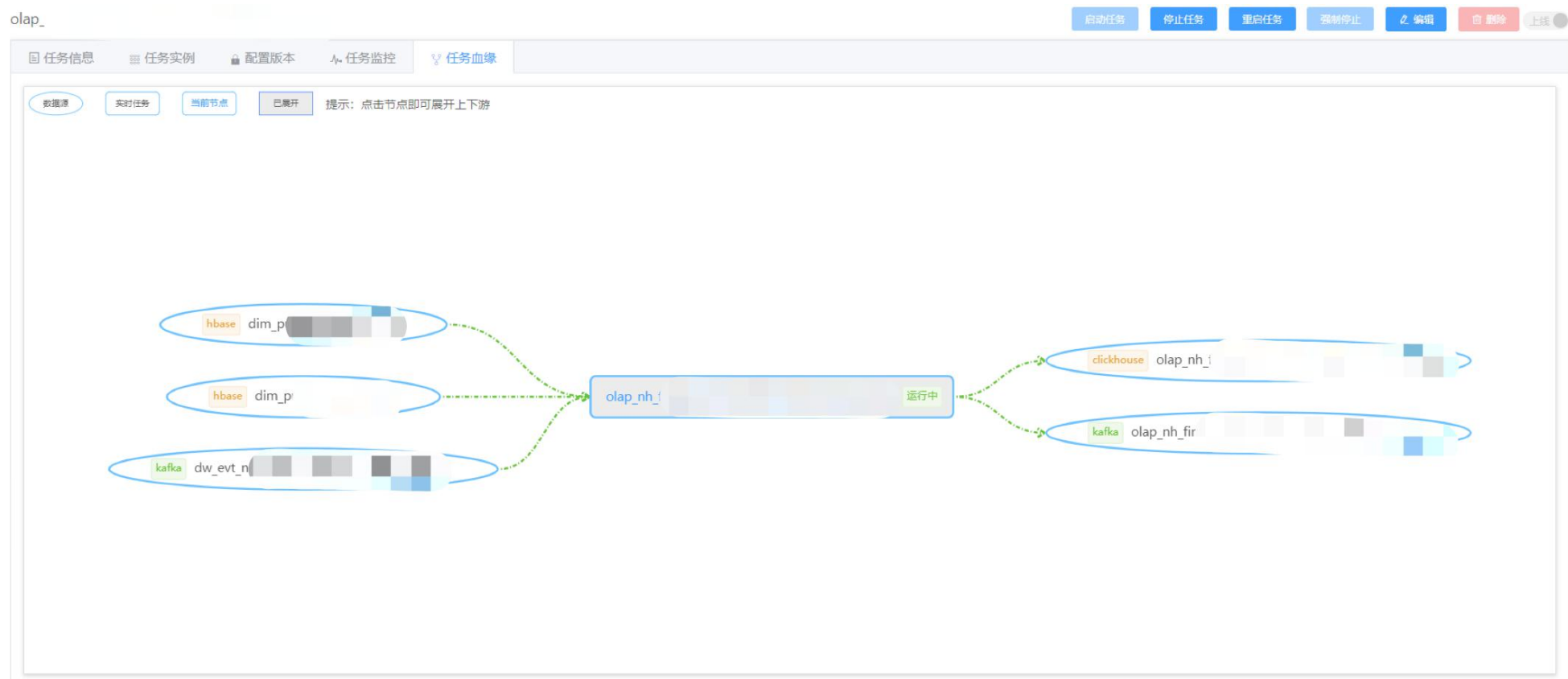
```
{
  "activity_name": "西安-1125期聚焦",
  "activity_end_time": "2020-12-03 07:59:59",
  "focus_id": 21,
  "create_time": "2020-12-03 07:59:59"
}, {
  "activity_name": "西安-1125期聚焦",
  "activity_end_time": "2020-12-03 07:59:59",
  "focus_id": 21,
  "create_time": "2020-12-03 07:59:59"
}, {
  "activity_name": "西安-1125期聚焦",
  "activity_end_time": "2020-12-03 07:59:59",
  "focus_id": 21,
  "create_time": "2020-12-03 07:59:59"
}, {
  "activity_name": "西安-1125期聚焦",
  "activity_end_time": "2020-12-03 07:59:59",
  "focus_id": 21,
  "create_time": "2020-12-03 07:59:59"
}, {
  "activity_name": "西安-1125期聚焦",
  "activity_end_time": "2020-12-03 07:59:59",
  "focus_id": 21,
  "create_time": "2020-12-03 07:59:59"
}
```

dwd_c

```
{
  "activity_name": "西安-1125期聚焦",
  "activity_end_time": "2020-12-03 07:59:59",
  "focus_id": 21,
  "create_time": "2020-12-03 07:59:59"
}, {
  "activity_name": "西安-1125期聚焦",
  "activity_end_time": "2020-12-03 07:59:59",
  "focus_id": 21,
  "create_time": "2020-12-03 07:59:59"
}, {
  "activity_name": "西安-1125期聚焦",
  "activity_end_time": "2020-12-03 07:59:59",
  "focus_id": 21,
  "create_time": "2020-12-03 07:59:59"
}, {
  "activity_name": "西安-1125期聚焦",
  "activity_end_time": "2020-12-03 07:59:59",
  "focus_id": 21,
  "create_time": "2020-12-03 07:59:59"
}, {
  "activity_name": "西安-1125期聚焦",
  "activity_end_time": "2020-12-03 07:59:59",
  "focus_id": 21,
  "create_time": "2020-12-03 07:59:59"
}
```

实时数仓建设 – 数据血缘

- 溯源分析
- 问题排查
- 差异分析
- 提升用户体验
- 变动/异常预警



实时数仓建设 - DDL

表类型	支持的底层数据存储类型
Source	Kafka
Dim	HBase Redis MySQL
Sink	Redis HBase HTTP ClickHouse Kafka Doris MySQL

```
CREATE (SOURCE|SINK|DIM) TABLE tableName (
  (columnName dataType EXTRACTION '\\' extractPath '\\')
  [, (columnName dataType EXTRACTION '\\' extractPath '\\')]*
  [, WATERMARK KEY columnName DELAY FOR interval]
  [, PAIMARY KEY columnName]
) WITH (
  propertyName = propertyValue [, propertyName = propertyValue]*
)
```

```
--定义数据源表
create source table kafka_source_tb (
  system_type string extraction '$.system_type',
  client_os_type string extraction '$.client_os_type',
  ucid string extraction '$.ucid',
  ts bigint extraction '$.timestamp',
  watermark key ts delay for 5s
) with (
  type = 'kafka',
  dataType = 'json',
  brokers = 'off03-bigdata.mars.ljnode.com:9092',
  topics = 'data-pipeline-common-dev',
  groupId = 'test-group'
);

--定义维表
create dim table redis_dim_tb (
  ucid string,
  first_login_time string,
  device_id string,
  primary key ucid
) with (
  type = 'redis',
  server = '127.0.0.1',
  port = '6379',
  cache = 'all'
);

--定义输出表
create sink table console_sink_tb (
  ucid string,
  first_login_time string,
  device_id string,
  client_os_type string,
  system_type string,
  ts bigint
) with (
  type = 'console',
  dataType = 'json'
);
```

实时数仓建设 – OneSQL(建设中)

当前面临的问题:

1. 任务稳定性: 减少断流时长
 - 任务运行失败后的迅速恢复,
 - 任务的状态管理及状态恢复
 - 任务运行出错时的问题排查、迅速定位
2. 计算性能提升: 高吞吐情况下保持较低的延迟
 - 引擎执行优化
 - 任务资源合理调配
3. SQL能力提升, 满足日益增多的业务场景
 - 聚合能力, 支持下流写入数据的更新
 - 关联能力的提升, 解决相关表数据的乱序问题

The screenshot displays the OneSQL web interface. On the left, a sidebar contains '数据表' (Data Tables) and '函数列表' (Function List). The main area is titled '函数列表' (Function List) and includes a search bar and filters for '字符串函数' (String Functions), '用户函数(UDF)' (User Functions), and '标量函数' (Scalar Functions). Two functions are listed: 'GET_JSON_OBJECT' and 'LAST_VALUE'. The 'GET_JSON_OBJECT' function is selected, showing its details: '负责人: xxx', '描述: json解析函数', and '标量' (Scalar) type. Below this, the function's syntax is shown: 'STRING GET_JSON_OBJECT(STRING json, STRING jsonPath)'. A table lists the parameters: 'json' (STRING) and 'jsonPath' (STRING). The function description states: '通过jsonPath解析json字符串的数据'. The test sentence is 'select get_json_object('{"a": "1"}', '\$.a')' and the test result is '1'. On the right, there is a 'sql编辑器' (SQL Editor) with line numbers 1 to 25 and a '控制输出' (Control Output) section.

实时数仓建设 - 交易大屏

交易实时指标

A+签约量

D+1报单率

全流程时效

可视化渗透率

在途量

完结量

批贷时效

批贷量

批贷金额

放款时效

放款量

放款金额

综合五星好评率

过户时效

面签量

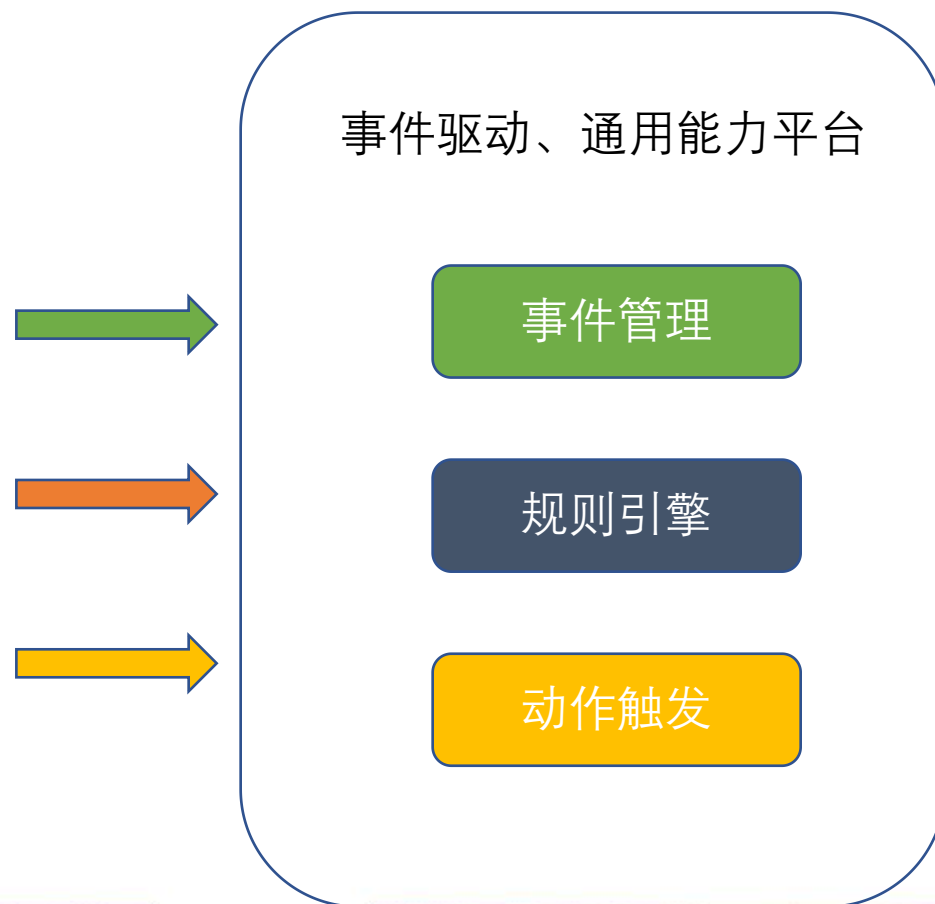
面签金额

黄绿灯率

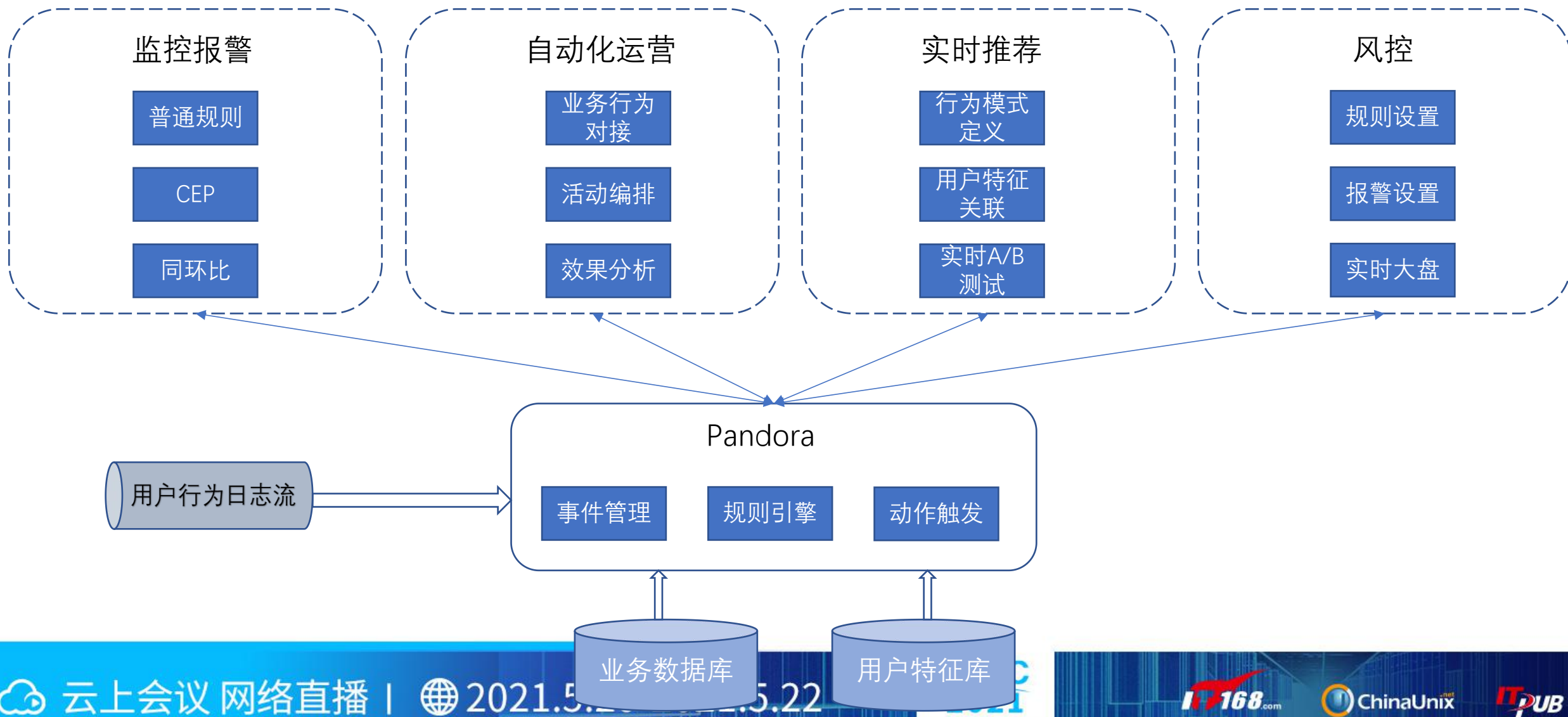


Pandora事件处理平台建设 – 需求背景

1. 用户行为事件缺乏统一抽象和管理；开发效率低、周期长，存在重复建设问题
2. 规则处理逻辑与业务系统耦合，难以灵活应对规则变化
3. 缺乏触发下游动作的统一灵活的管理和配置，以及监控报警



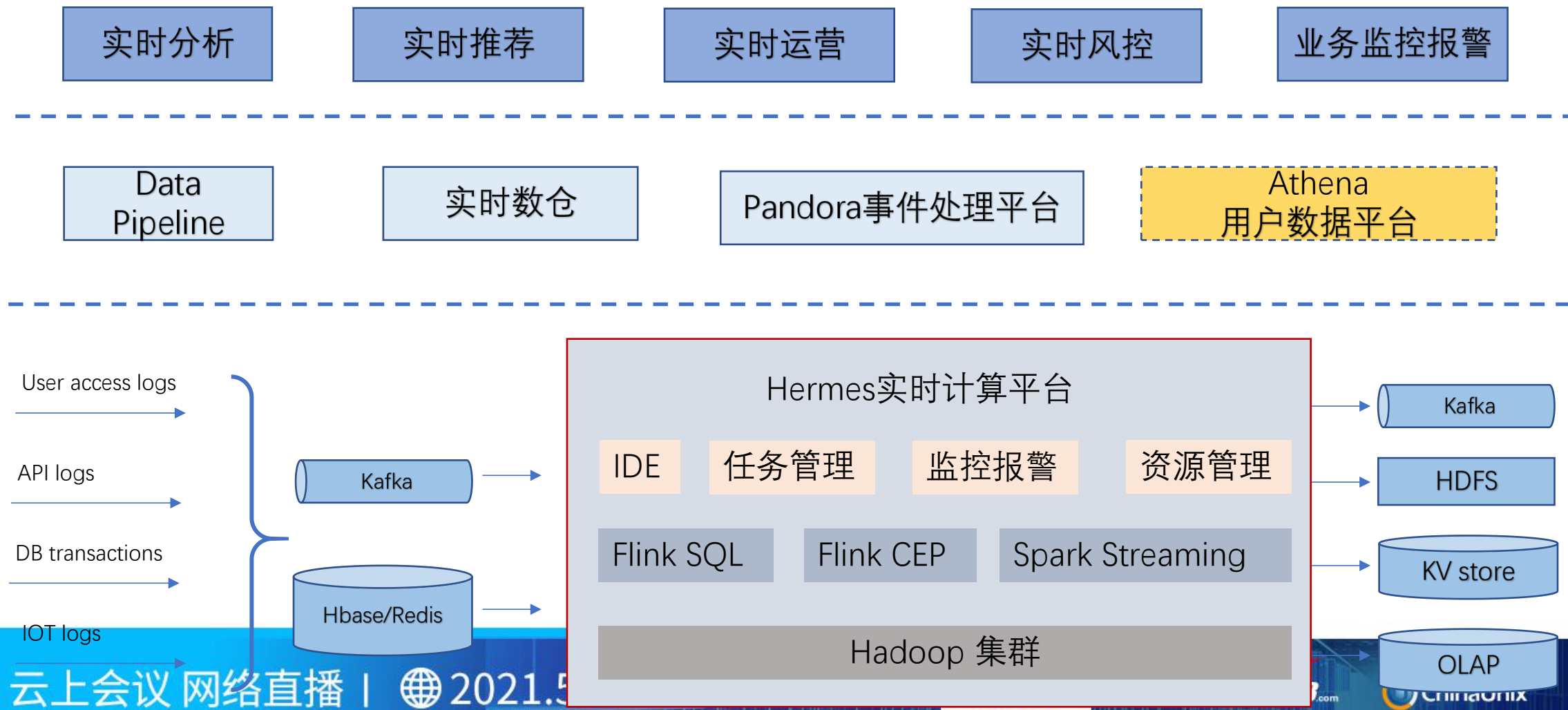
Pandora事件处理平台建设 – 应用场景



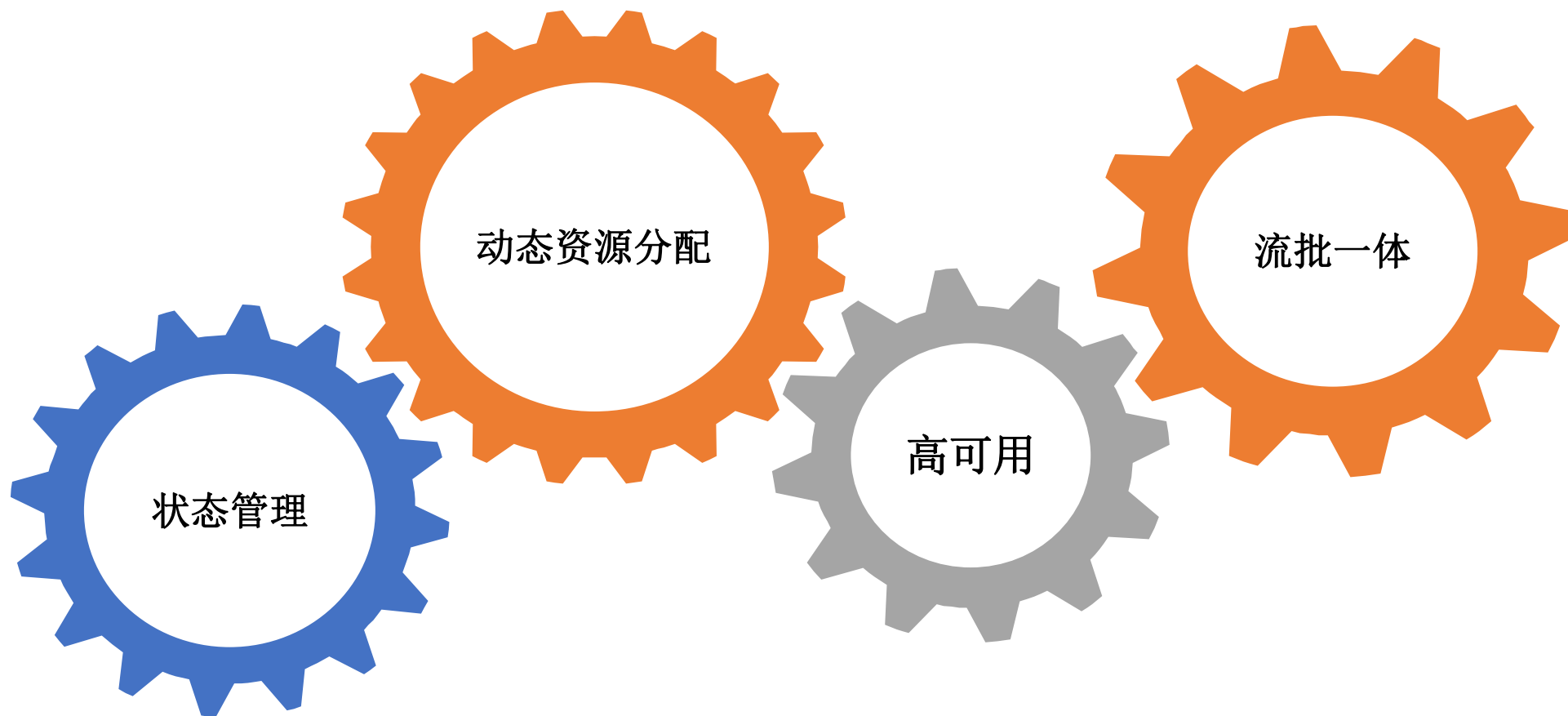
Pandora事件处理平台建设 - 架构



实时计算架构



未来规划



The background is a deep blue with a complex, abstract pattern of glowing blue wireframe cubes and rectangular prisms. These shapes are arranged in a way that creates a sense of depth and perspective, with some appearing to recede into the distance. A bright, horizontal lens flare or light streak cuts across the center of the image, passing behind the word 'THANKS'. In the upper left, there are some faint, stylized geometric shapes resembling flags or banners. The overall aesthetic is futuristic and digital.

THANKS