Apache Flink在网易的实践

How Netease Uses Apache Flink

吴良波 Wu Liangbo
网易JAVA技术专家 Netease Java Technical Specialist

FLINK FORWARD # ASIA
实时即未来 # Real-time Is The Future

FLINK FORWARD

Contents
目录

FLINK FORWARD

业务与规模演进

The evolution of business and scale

01

网易流计算演进

The evolution of stream computing in Netease

Storm

Storm task

2012

Sloth0.x

Flink 1.5 Flink 1.3

2017.2

Sloth1.0

Flink1.7 Blink

2019.1

Sloth2.0

Flink1.5 Flink1.7
Flink1.9 Blink

2019.7

基于流计算业务规模

The business scale based on stream computing

1000+ 任务

1000+ tasks

2W+ vcores

2w+ vcores

80+ T 内存

80+ T mem

业务场景

Business scenarios

FLINK FORWARD

Flink平台化

Flink platformization

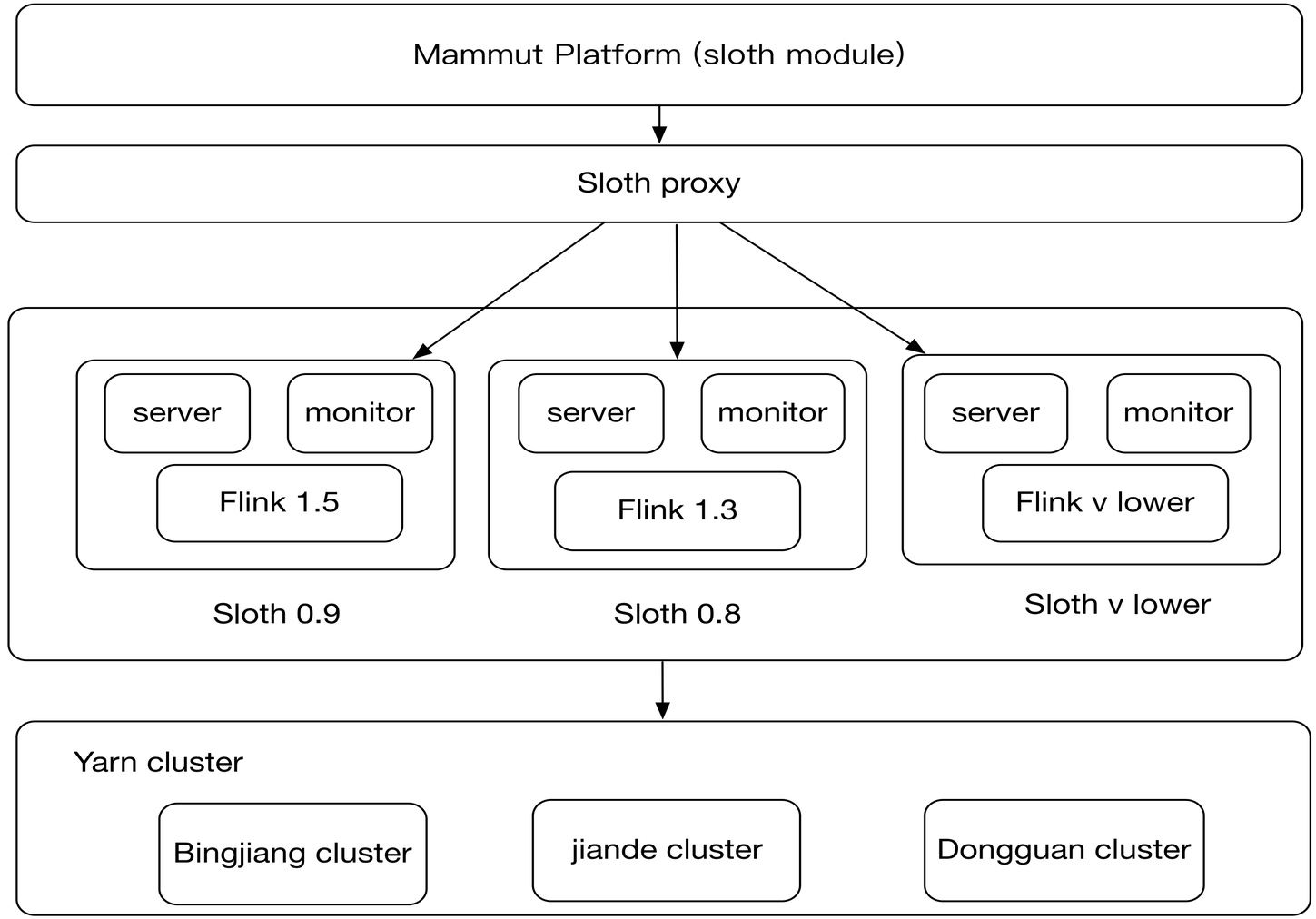
02

平台架构演进-Sloth0.x

The evolution of platform architecture- Sloth0.x

自研 codegen
Self-develop codegen

运维困难
Hard to operation and maintenance



平台架构演进-Sloth1.0

The evolution of platform architecture – Sloth1.0

Flink版本插件化
Pluggable of different Flink versions

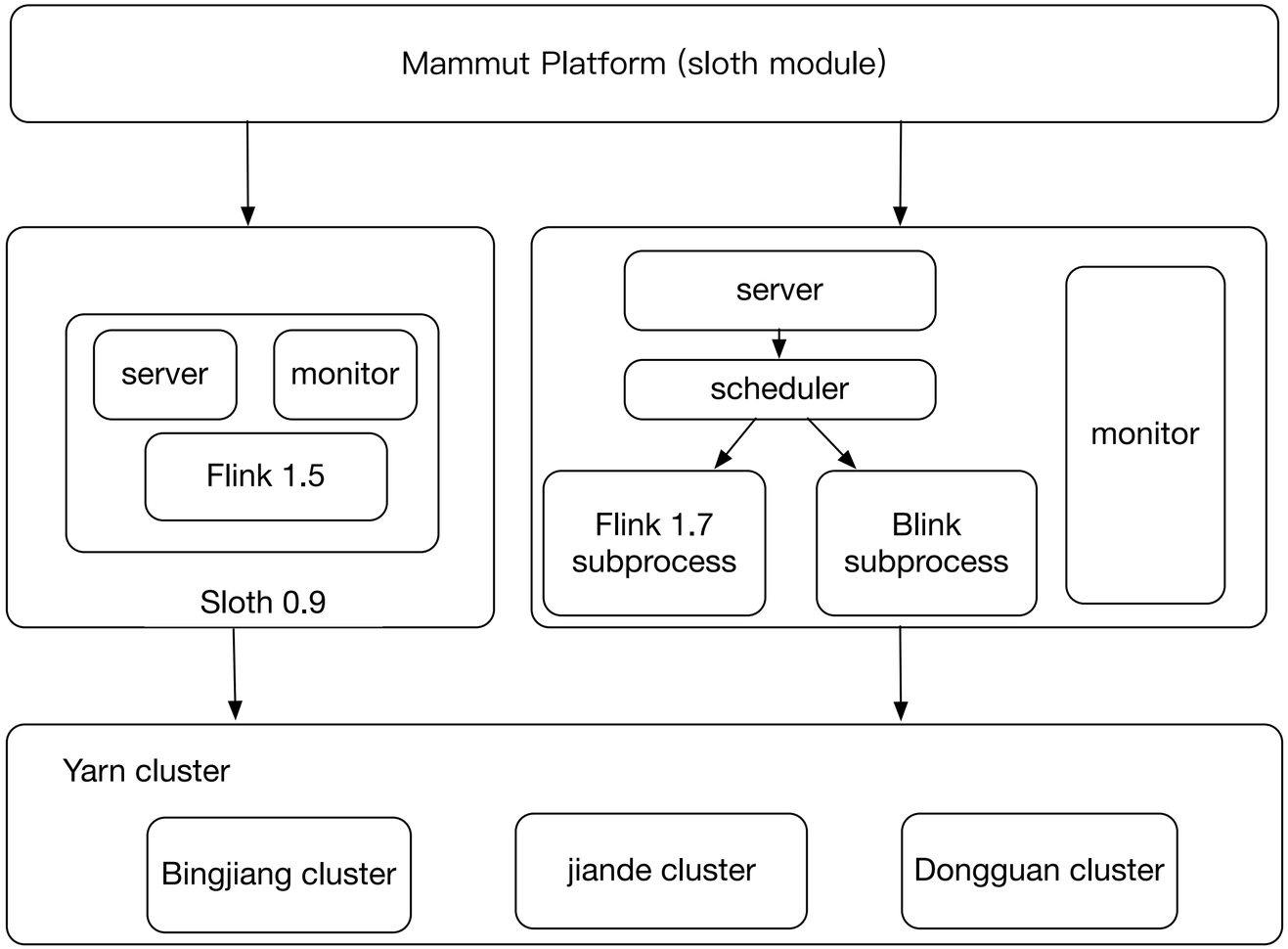
父子进程架构
Father and children process architecture

运维方便
Easy to operate and maintenance

支持jar包任务开发
Support Jar task development

Blink SQL开发适配
Support Blink SQL development

监控接入Grafana
Monitor using Grafana



平台架构演进-Sloth2.0

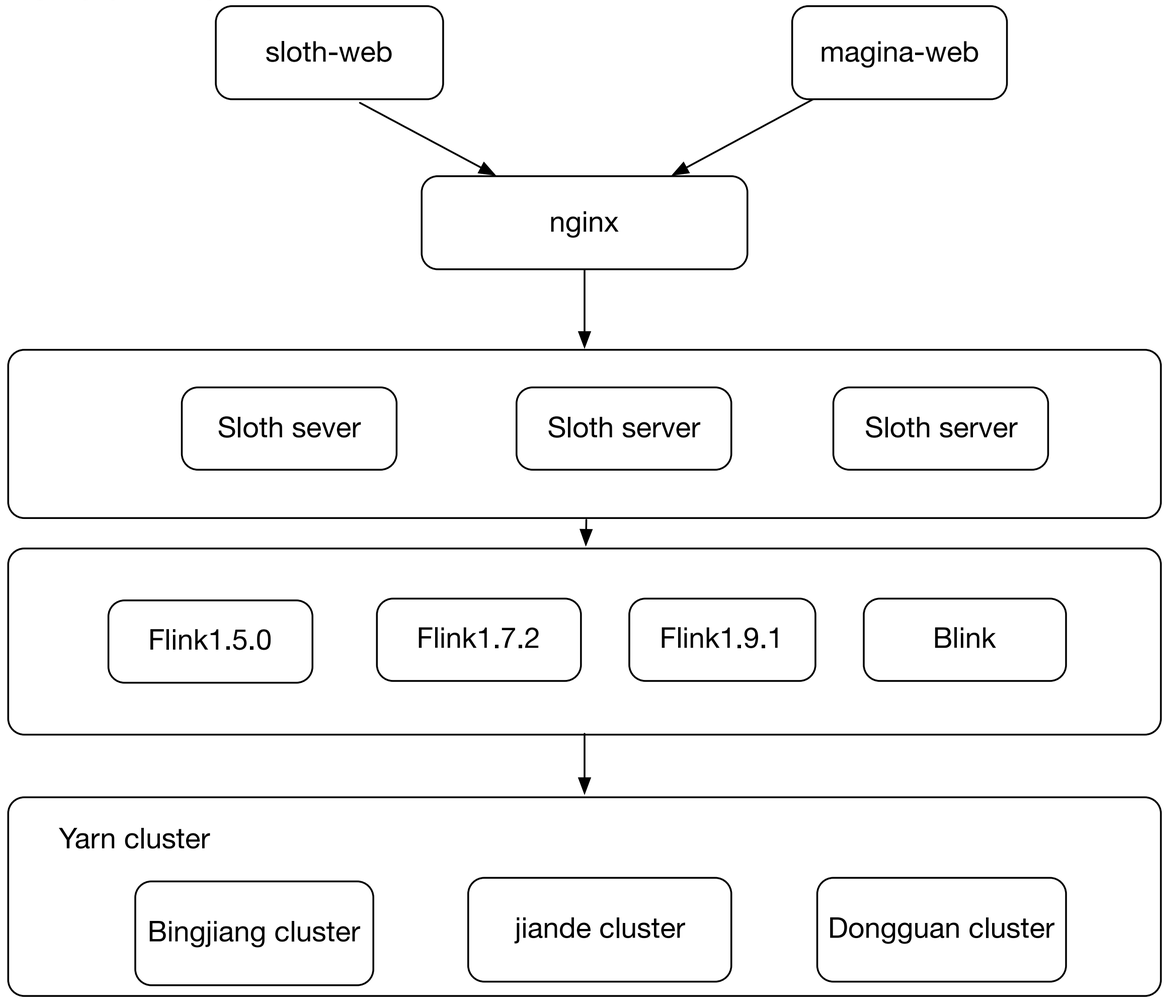
The evolution of platform architecture – Sloth2.0

平台PaaS化

Supply platform interface to different web

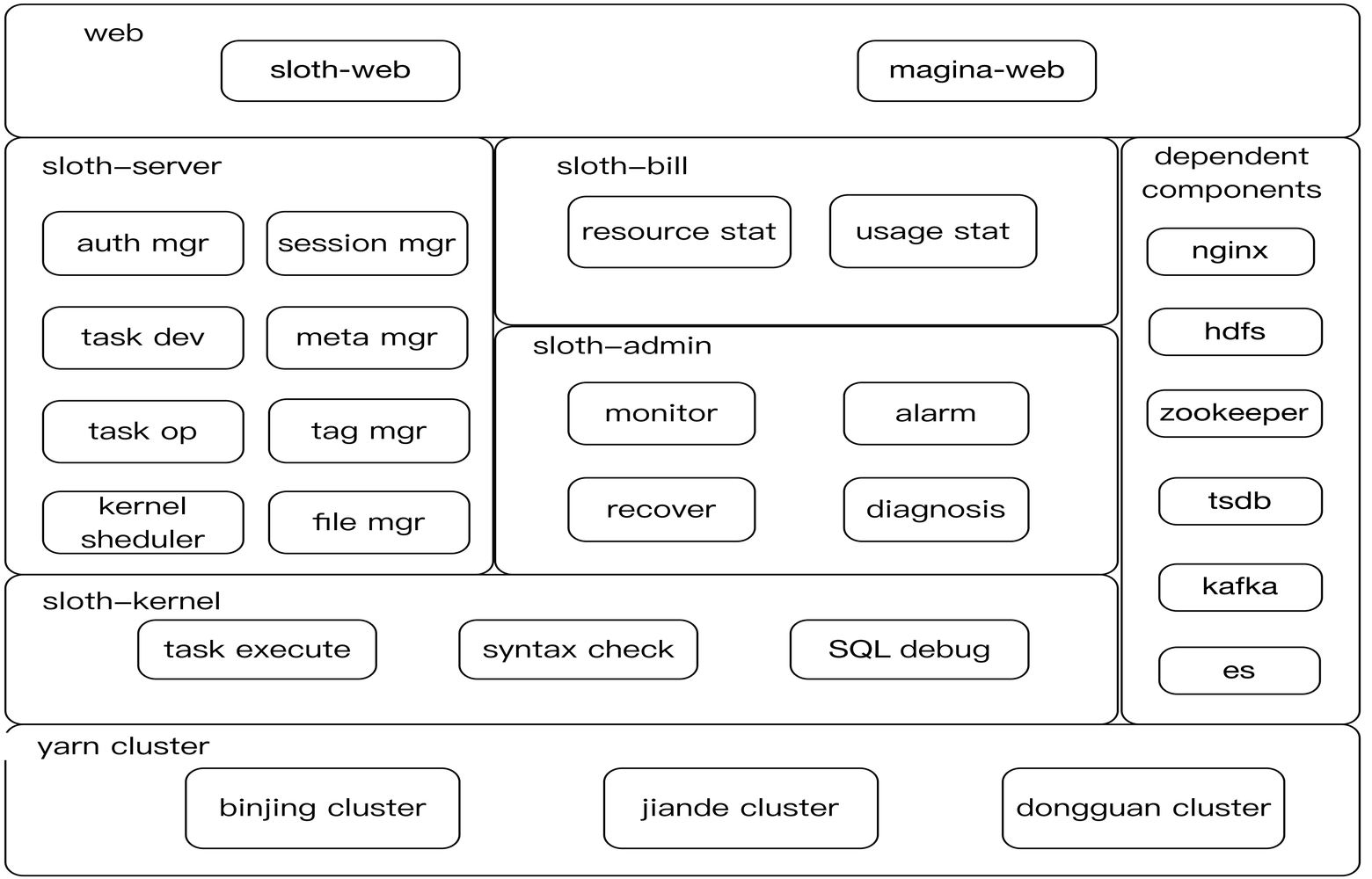
支持server分布式部署，实现高可用

Support distribute deployment of server, service high available



平台模块图

The modules of platform



事件管理

Event management

事件包括任务的启动和停止两个操作，由Server, Kernel,Admin三个模块共同完成

The event includes two operations of task start and stop, which are completed by three modules: Server, Kernel and Admin

Server: 事件执行的发起者，接受事件的请求，进行数据校验，拼装，将事件发送给kernel执行

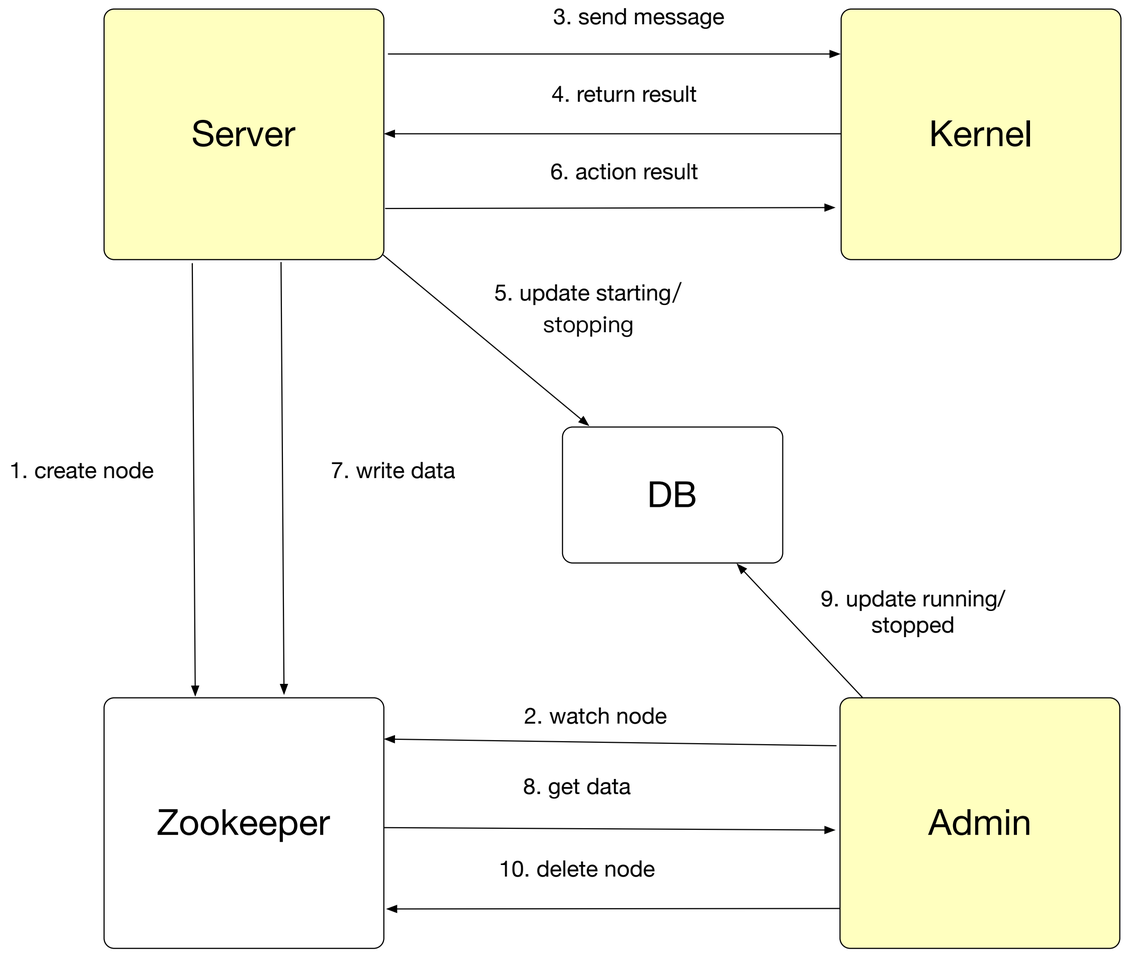
Server:The initiator of event execution, accepts the request of the event, performs data verification and assembly, and sends the event to the kernel for execution

Kernel: 事件具体逻辑的执行者，根据请求向集群发送指令（shell脚本方式）

Kernel:The executor of the event specific logic, sends instructions to the cluster according to the request (shell script mode)

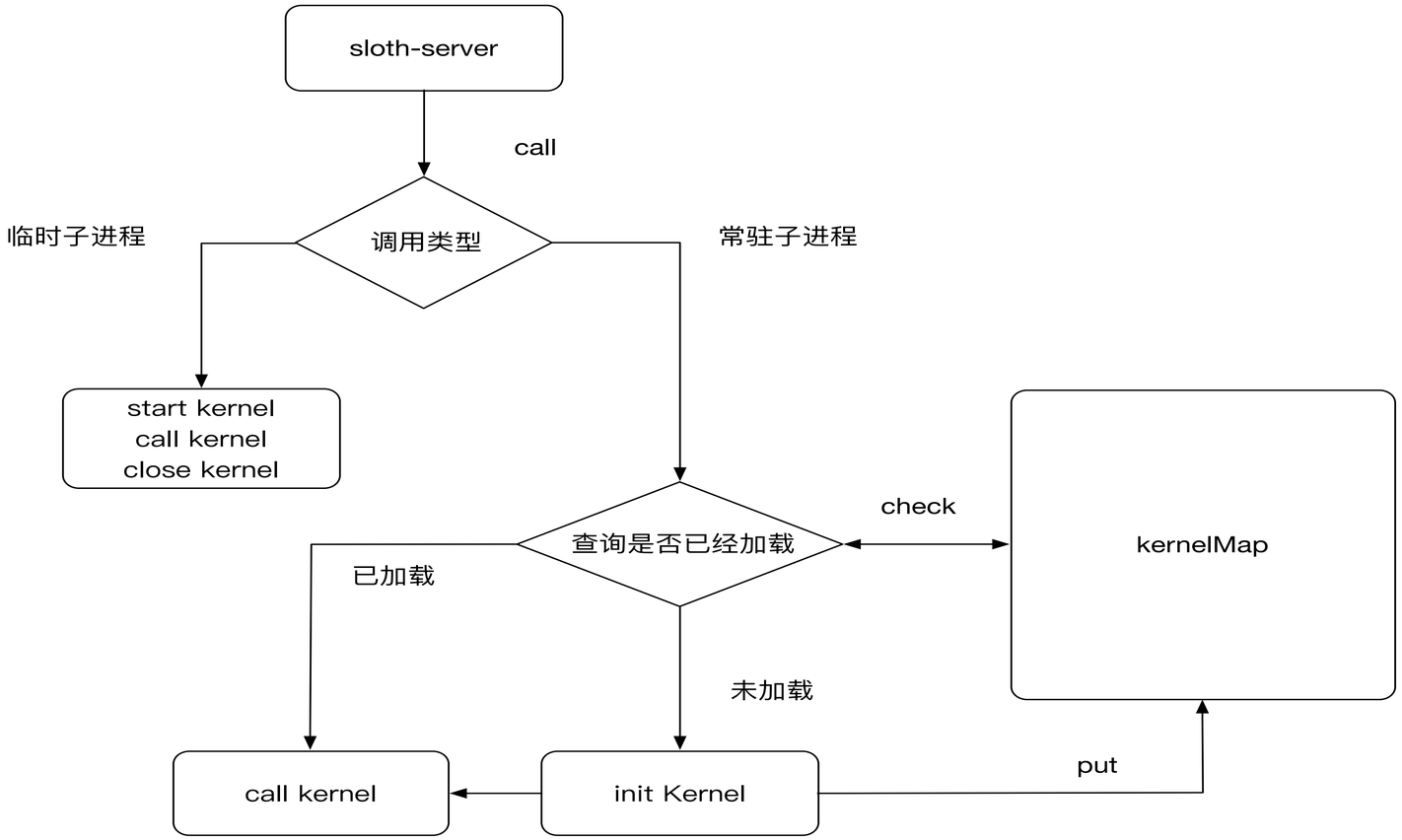
Admin: 事件执行结果的确认者，根据事件类型，获取事件的最终结果，保证结果的正确性

Admin: The confirmer of the event execution result, obtains the final result of the event according to the event type to ensure the correctness of the result



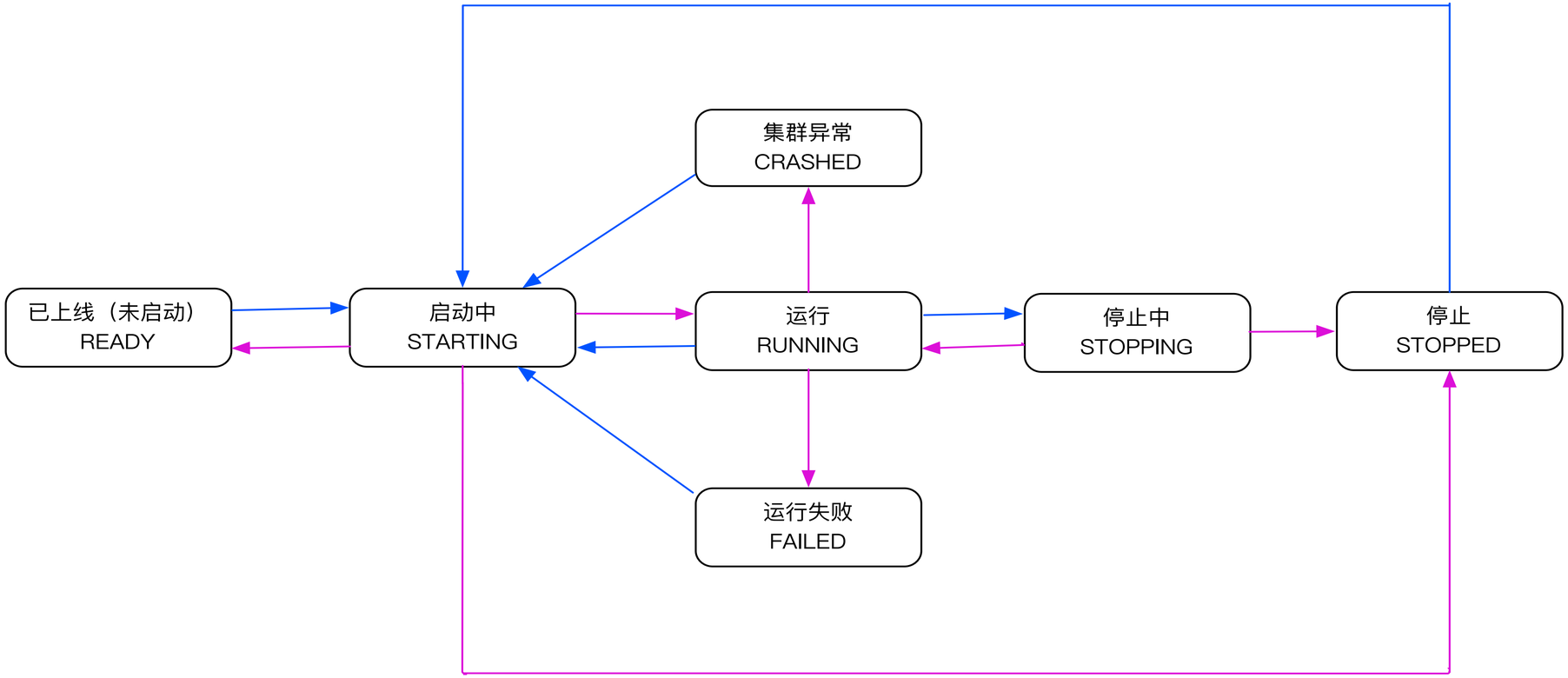
内核调度

Kernel scheduler



平台任务状态图

The statechart of platform



任务开发
Task development

支持任务调试
Support task debug

支持任务Tab页
Support tab page for tasks

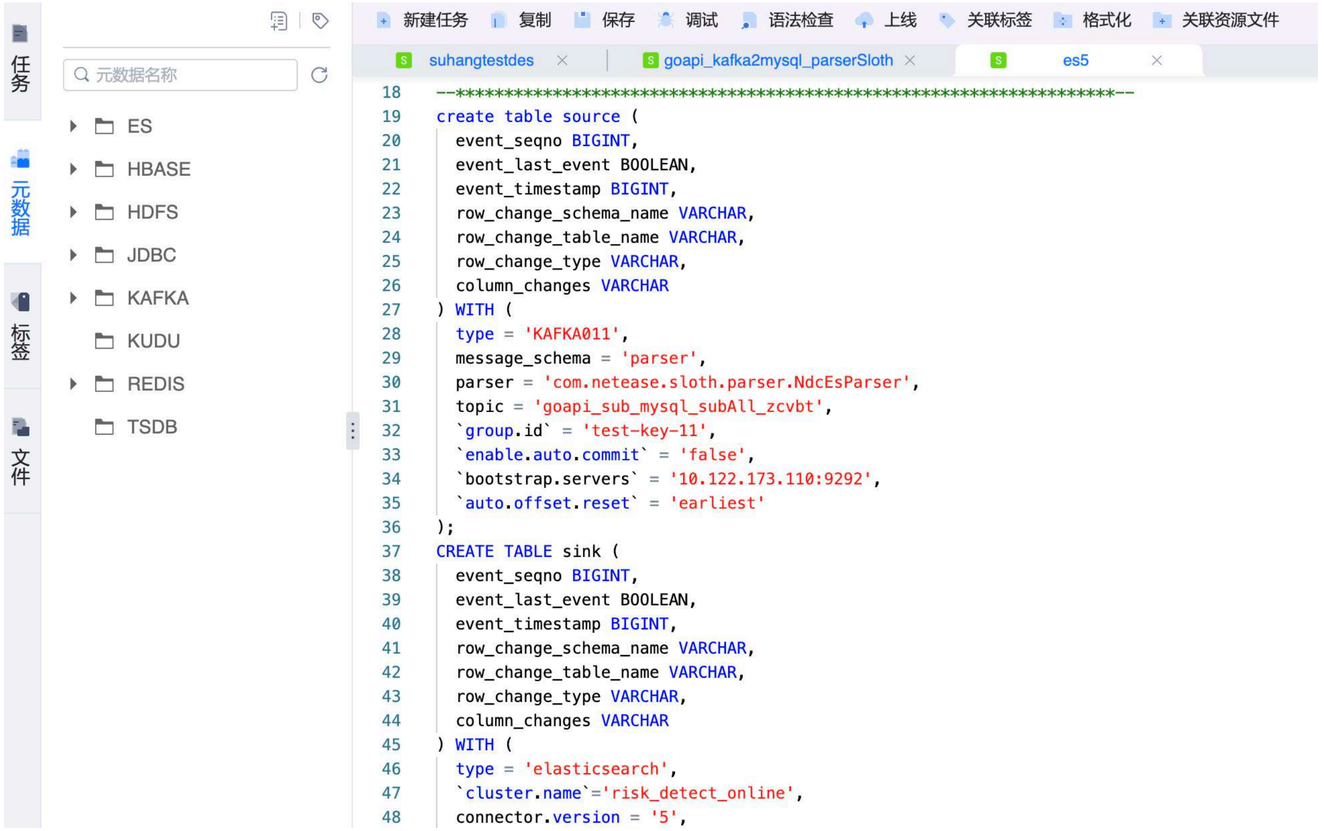
支持语法检查
Support syntax check

支持任务标签
Support tags for tasks

支持元数据管理
Support metadata management

支持用户资源文件管理
Support user resource file management

支持任务复制
Support task replication

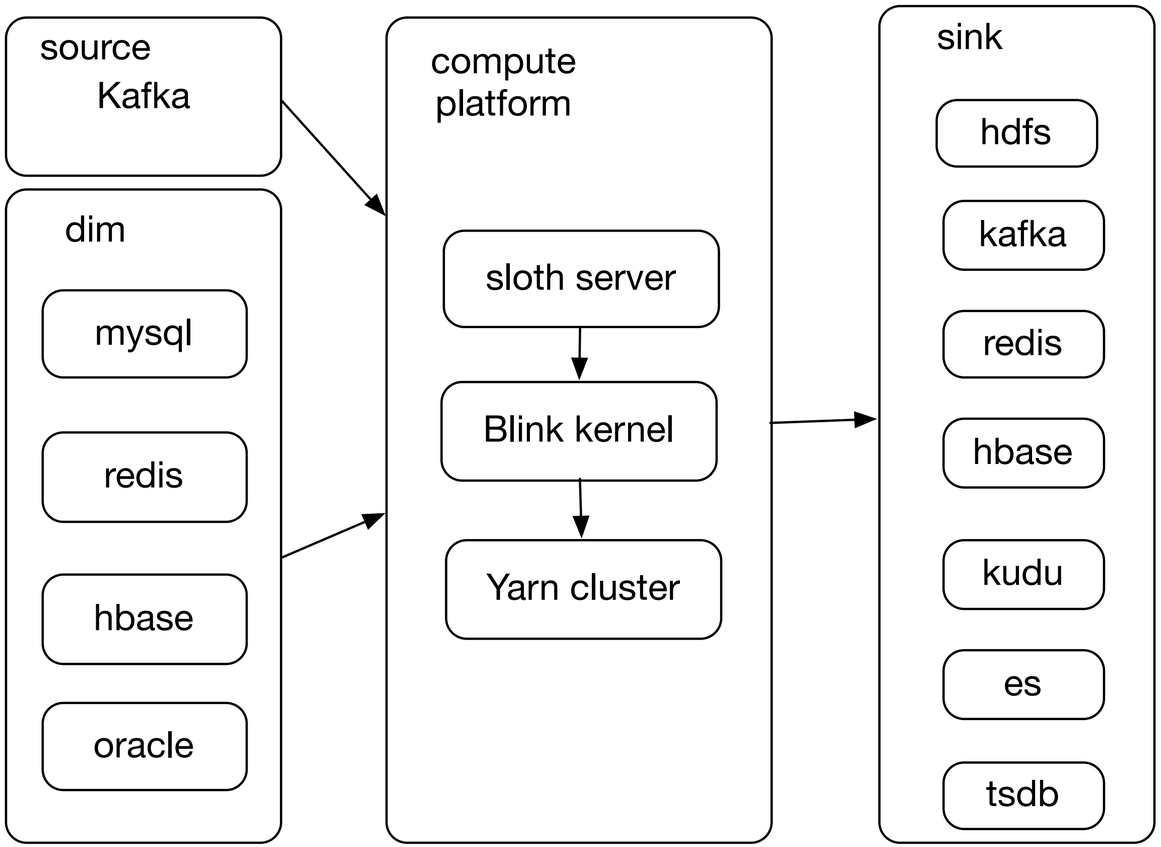


Blink SQL

Blink SQL

扩充完善了开源Blink对维表JOIN支持，以及SINK端的支持。

Expansion and improvement of the support for dim table join and sink



SQL任务调试

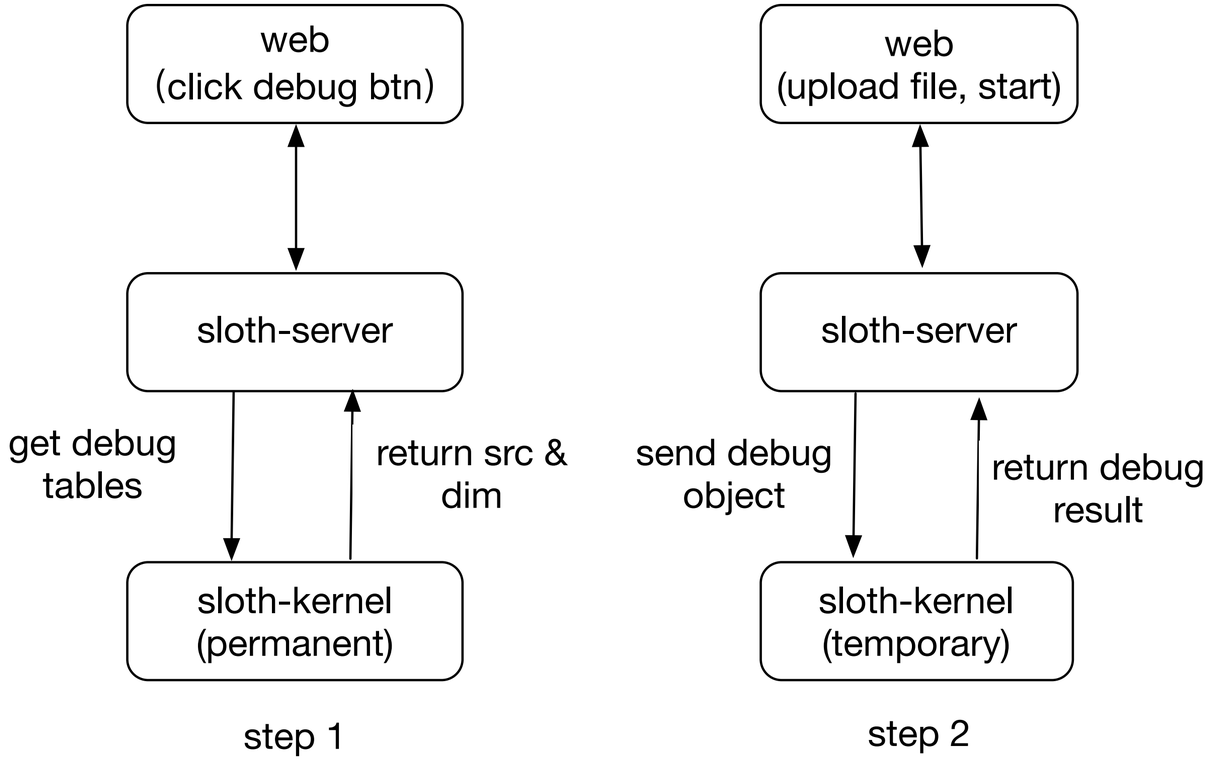
SQL task debug

SQL类型的任务支持调试功能，用户可以根据不同的source表和dim表，上传不同的csv文件作为输入数据，进行调试

SQL task support debug by uploading different csv files as source according to different source table and dim table

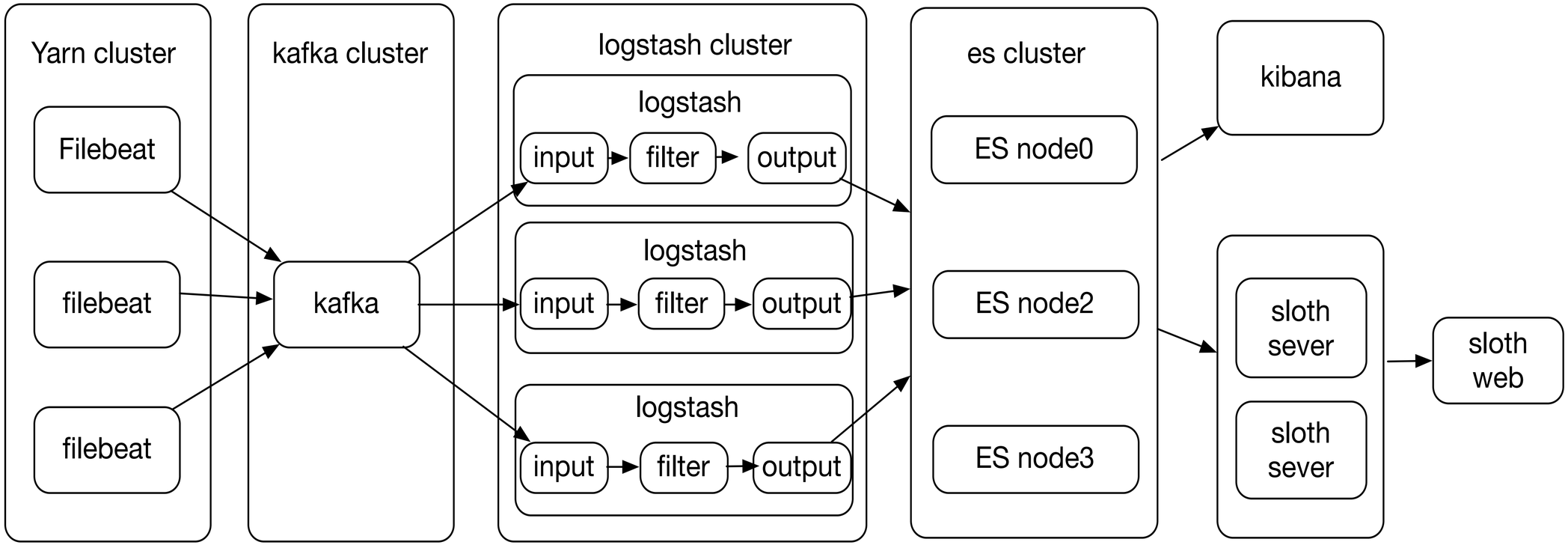
调试执行由指定的kernel来完成，sloth-server负责组装请求，调用kernel，返回结果，搜集日志

The execution of debug use assigned kernel, sloth server response for packaging request, invoking kernel, returning results, searching logs



日志检索

Log retrieval



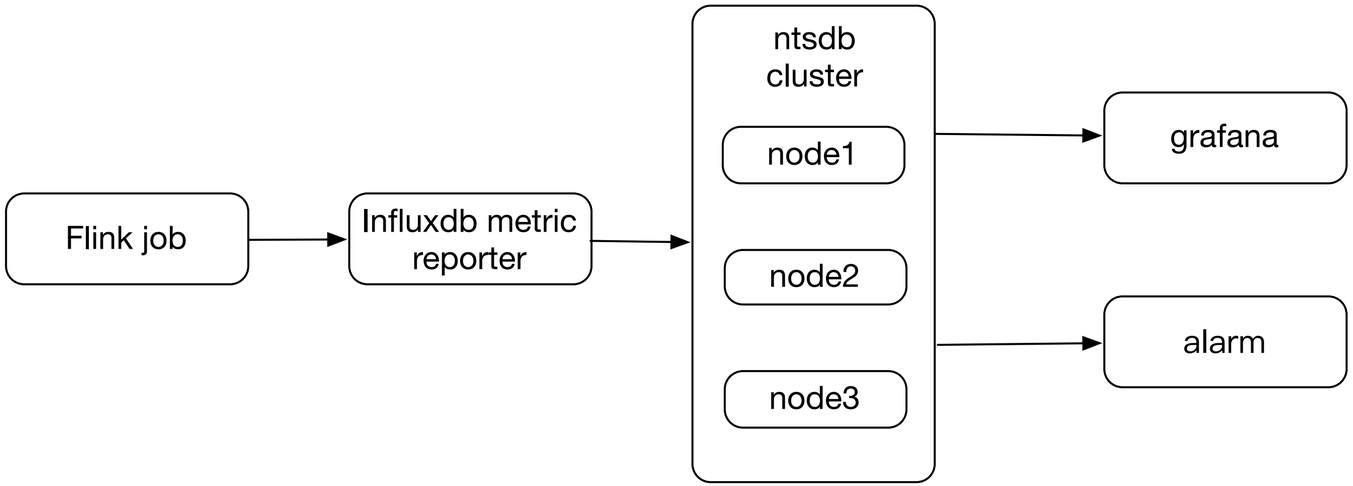
监控
Monitor

使用influxdb metirc reporter对metrics进行监控

Monitor metrics using influxdb metric reporter

使用网易自研的ntsdb时序数据库

Using self-develop time series database ntsdb



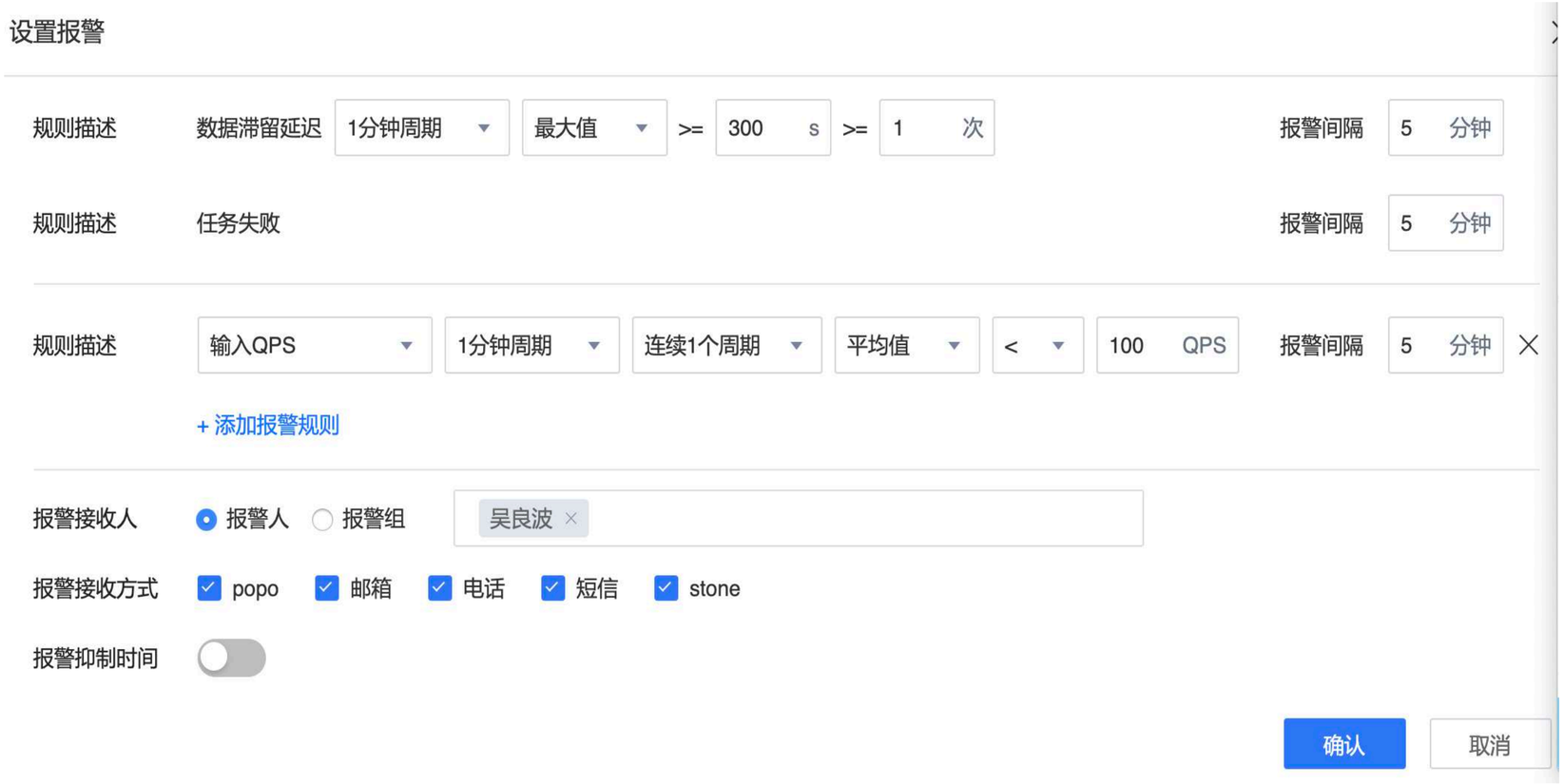
监控
Monitor


输入流量（IN_QPS）
20 K
18 K
15 K
13 K
10 K
8 K
5 K
13:52
13:54
13:56
13:58
14:00
Source: Custom Source Current:
输出流量（OUT_QPS）
200
175
150
125
100
75
13:52
13:54
13:56
13:58
14:00
Sink: Unnamed Current:
kafka每秒消费的消息数
20 K
18 K
15 K
13 K
10 K
8 K
13:52
13:54
13:56
13:58
14:00
Source: Custom Source Current:
kafka中Consumer落后于producer的消息数
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0
13:52
13:54
13:56
13:58
14:00
Source: Custom Source Current:

报警
Alarm

支持用户自定义报警规则
User can self define alarm rules

支持多种报警方式
Support multiple alarm styles



FLINK FORWARD

案例分析

Case study

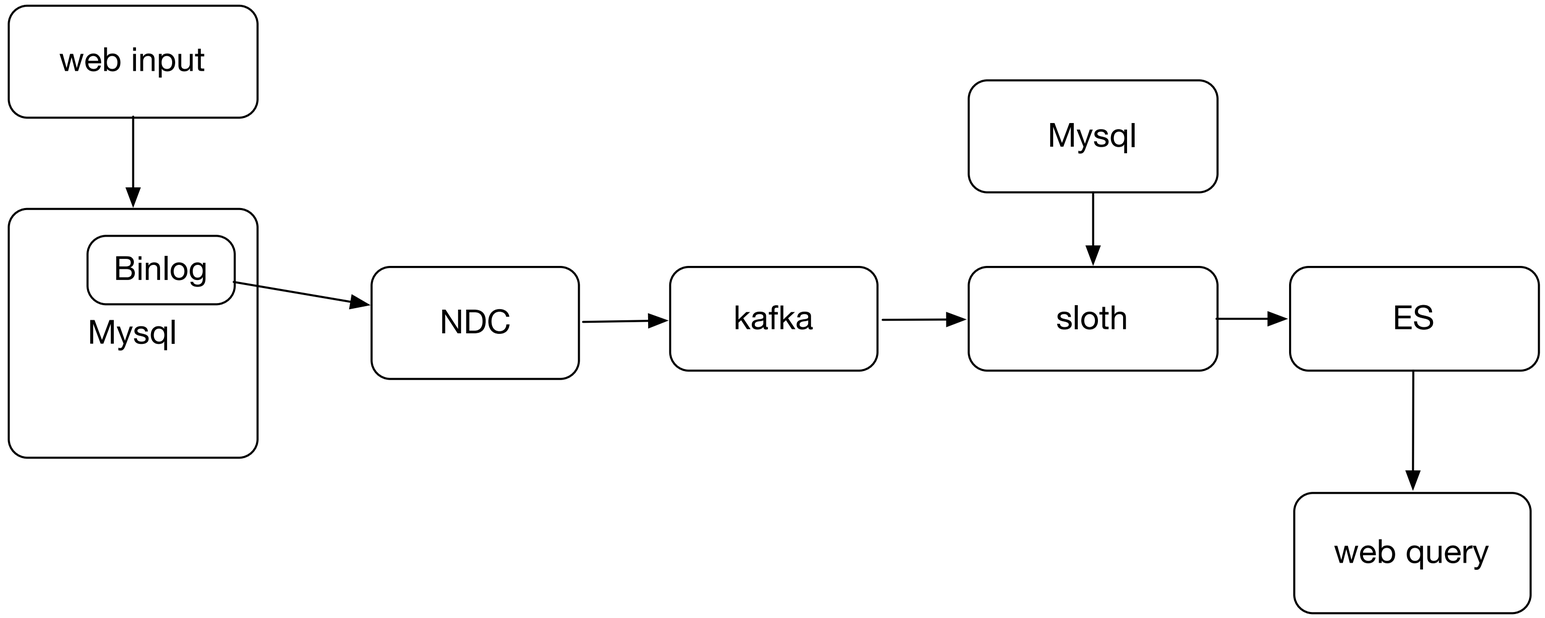
03

数据实时同步

Realtime data synchronization

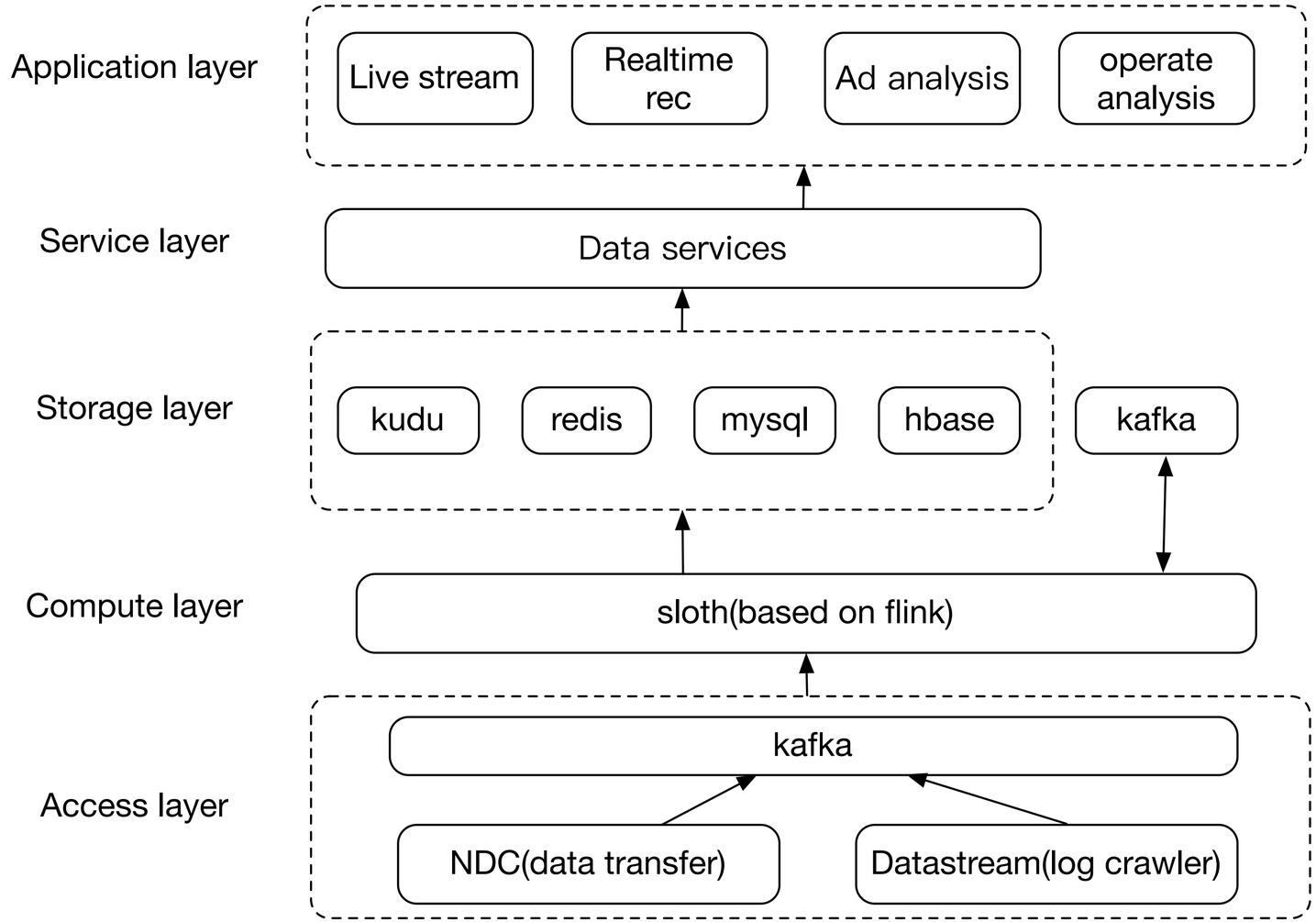
AI智能对话服务：客户在前端配置知识库数据,通过Sloth实时处理后，写入到ES中供查询端使用

AI smart talk services: client input the knowledge data in frontend, then real processed by Sloth, finally write into ES for user query



实时数仓

Realtime warehouse



电商应用-数据分析

E-commerce application – data analysis

业务场景：实时活动分析，首页资源分析，流量漏斗，实时毛利计算

Business scenario: real time activity analysis, home page resource analysis, flow funnel, real-time gross profit calculation

简要逻辑：从Hubble收集用户的访问日志推送到Kafka，使用Sloth清洗出明细层，写入Kafka,再用Sloth任务，关联维度，实时写入Kudu。落入Kudu表的数据，一方面给业务方，分析师开发实时查询，另一方面可以在这个实时的Kudu表上面，查询汇总数据，提供给数据应用

Brief logic: collect the user's access log from Hubble and push it to Kafka, use Sloth to clean out the details layer, write it to Kafka, then use sloth task to associate dimensions and write it to Kudu in real time. The data falling into the Kudu table, on the one hand, can be used to develop real-time queries for business parties and analysts; on the other hand, it can be used to query and summarize data on the real-time Kudu table and provide it to data applications

电商应用-搜索推荐

E-commerce application - search and recommendation

业务场景：用户实时足迹，用户实时特征，商品实时特征，实时CTR CVR样本组建，首页A区轮播，B区活动精选等UV,PV实时统计

Business scenario: user real-time footprints, user real-time features, real-time product features, real-time CTR CVR sample formation, home page a area rotation, B area activity selection and other UV, PV real-time statistics

简要逻辑： 使用Sloth读取应用日志（曝光，点击，加购物车，购买），进行数据清洗和维度拆分，写入Kafka，再使用Sloth读取Kafka数据，实时统计多维特征5min,30min,1h的PV和UV，写入Redis，供线上工程计算CTR,CVR,以及优化搜索和推荐结果

Brief logic: use Sloth to read application logs (exposure, click, add shopping cart, purchase), clean data and split dimensions, write into Kafka, then use sloth to read Kafka data, make real-time statistics of PV and UV of multi-dimensional features for 5min, 30min, 1H, write into Redis, for online engineering calculation CTR, CVR, and optimize search and recommendation results

FLINK FORWARD

未来发展与思考

Future development and thinking

04

未来发展与思考

Future development and thinking

1. 实时计算平台支持 Flink on K8s的任务
Support running K8s tasks in real time stream computing platform

2. 任务的自动配置功能，平台能根据业务类型，流量自动配置内存，并发度等，既保证业务SLA
,也能提升计算集群资源利用率
Task can auto config, the platform can auto config memory and parallelism according to business type and traffic. It can not only promise the SLA of business, also improve the resource utilization rate of the computing cluster

3. 智能诊断，对UDF以及代码构建的流计算任务，调试成本高，运行出错让业务和平台方疲于奔命，智能诊断是流计算平台根据任务的各种metrics信息，直指问题所在，减少业务和平台定位问题的时间，对于存在风险的任务，可以提前给出预警，并对调优给出建议
Smart diagnosis: It's hard to debug stream task with udf and building by code. The failure of task make the business and platform side exhausted. Smart diagnosis directly point out the problem according to the various task metrics, reduces the time of business and platform positioning problems. It should give warning in advance and give advice on tuning

4. 关注Flink 1.9后续对SQL的支持，以及Flink批流统一
Take a look at the support of SQL in Flink 1.9 later series and the unification of batch and streaming

5. 更多的参与到社区中去
More participant in Flink community


FLINK
FORWARD
THANKS