**接口日志分析 设计文档**

****

**北京航天数据股份有限公司**

**2018 年 4月19日**

**接口日志分析 设计文档**

**修改历史**

|  |  |  |  |
| --- | --- | --- | --- |
| **日期** | **版本** | **作者** | **修改内容** |
| 2018-04-19 | 1.0 | 陈 思 | 创建文档 |

目录

[**接口日志分析 设计文档** 2](file:///C:\Users\anhaoyan\Desktop\ht\spark\文档\spark接口日志分析设计文档.docx#_Toc511916118)

[**一．概述** 3](#_Toc511916119)

[**二．模块组件介绍** 3](#_Toc511916120)

[2.1启动类 4](#_Toc511916122)

[2.2数据统计 4](#_Toc511916123)

[**三．三方组件和库** 4](#_Toc511916124)

[3.1 Zookeeper 4](#_Toc511916125)

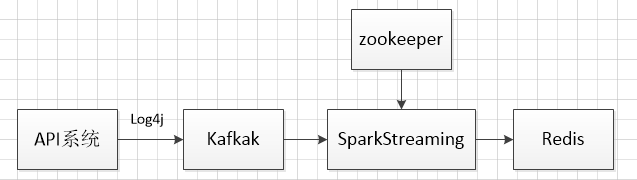
[3.2 Kafka 4](#_Toc511916126)

[3.3 Redis 5](#_Toc511916127)

1. **概述**

该模块主要负责处理api系统实时产生的日志文件。对日志文件进行不同维度的实时计算分析，对分析的结果实时存入redis数据库中供其他模块的调用。

1. **模块组件介绍**



## 2.1启动类

项目采用gradle搭建应用环境

入口: com.iiot.stream.HTLogAnalysisContext

## 2.2数据统计

> \* 各个接口累计访问次数

> \* 每个接口每分钟最大访问次数

> \* 每个接口响应总时长

1. **三方组件和库**
   1. Zookeeper

|  |  |
| --- | --- |
| 作用： | 统一配置管理。保存kafka ip,redis相关信息; |
| 通信地址： | Zookeeper地址由外部参数传入 |

* 1. Kafka

|  |  |
| --- | --- |
| 作用： | 作为消息中间件，存储由“API系统”生产的日志。 |
| Topic消息格式： | 2018-04-19 14:46:39,367 {"allowed":true,"date":1524120399358,"method":"DELETE","parameterMap":{},"reqUrl":"/1.1/knowledge/61088a","requestTime":1524120399358,"responseTime":1524120399367,"srcIP":"192.168.15.118","srcPort":7967,"useTime":9} |

* 1. Redis

|  |  |
| --- | --- |
| 作用： | 存储实时计算的结果 |
| 存储格式： | 采用redis的hash结构进行数据存储  key: api\_log\_filter\_times\_yyyy-MM-dd (按天计算各个接口累计访问次数) field:GET/people(提交方式+访问接口地址)value:值  key: api\_log\_filter\_times\_perminute\_yyyy-MM-dd (按天计算每个接口每分钟访问最大次数) field:GET/people(提交方式+访问接口地址)value:值  key: api\_log\_filter\_usetime\_yyyy-MM-dd (按天计算每个接口响应总时长) field:GET/people(提交方式+访问接口地址)value:值 |
| 通信地址： | 从zookeeper中读取**/conf\_htiiot/redis**节点 |

四.启动方式

|  |
| --- |
| spark-submit --master yarn --deploy-mode cluster --class com.iiot.stream.HTLogAnalysisContext ./logAnalysis.jar 192.168.0.83:2181 |