**雷神探测V1.1.0迁移方案**

一、组件描述：

* 组件逻辑结构：

从kafka中获取数据，通过开发程序实现数据的清洗加工放入kafka中，然后通过kylin消费，调用后台接口5分钟发送一次数据。

依赖图如下：

Kafka（topic:helios-processed-log）

forest数据清洗

Kafka（topic: thor-probe-log）

Kylin(whaley\_thor\_probe)

http post 发送

后台接口

* 组件清单

|  |  |  |
| --- | --- | --- |
| 层级 | 类别 | 文件清单 |
| 输入层 | Kafka | topic: helios-processed-log |
| 中间层 | Kafka(forest程序) | forest程序（2-22：spark用户，/home/spark/thor/），topic: thor-probe-log |
| kylin | whaley\_thor\_probe、whaley\_thor\_model、 whaley\_thor\_cube |
| 输出层 | Java程序 | WhaleyThorProbe程序（10.10.1.1 hadoop用户  /hadoopecosystem/thor） |
| 任务调度 | crontab 执行 | 10.10.1.1 hadoop用户  \*/5 \* \* \* \* sh /hadoopecosystem/thor/start.sh |

1. 迁移步骤：
2. 前置条件：
3. kafka topic helios-processed-log 中有数据
4. kafka topic thor-probe-log 存在
5. kylin 构建完成
6. 后台接口可用（telnet ip 可用）
7. 迁移步骤：
8. 创建topic thor-probe-log
9. 构建kylin cube（见confluence）
10. forest数据加工清洗程序迁移：

迁移2-22节点（spark用户）/home/spark/thor/ 下全量目录及文件

启动脚本：sh thorprobe.sh

1. 消息推送程序迁移：

迁移1-1节点（hadoop用户）/hadoopecosystem/thor下全量目录及文件

1. 任务调度（crontab）迁移（为消息推送服务）：

\*/5 \* \* \* \* sh /hadoopecosystem/thor/start.sh

三、检查清单：

|  |  |  |
| --- | --- | --- |
| 阶段 | 检测点 | 检测方法 |
| Kafka | topic: helios-processed-log | 通过kafka监控页面，查看helios-processed-log是否有数据 |
| topic: thor-probe-log | Forest程序启动后，通过kafka监控页面查看thor\_probe消费组有数据消费变化, thor-probe-log中数据是否增长 |
| Kylin cube构建 | whaley\_thor\_probe、whaley\_thor\_probe\_model、  whaley\_thor\_probe\_cube | 1. 三者存在 2. Kylin 页面可以查询， 3. 2-3分钟后数据有增长 |
| 数据推送 | probe.aginomoto.com | 是否可以ping通 |
| 任务调度 | crontab | 1. crontab是否配置 2. 每隔5分钟，日志是否有输出 |