# 一、国内服务器

国内大数据集群服务器列表：

|  |  |  |  |
| --- | --- | --- | --- |
| Host | 服务器ip | Admin密码 | 主要角色 |
| gzcloud11 | 119.29.148.206 | CWoU4r7D | trino管理 |
| gzcloud12 | 43.139.244.142 | CWoU4r7D | Hadoop从管理 |
| gzcloud13 | 114.132.178.44 | CWoU4r7D | HDFS数据节点、YARN计算节点、TRINO计算节点 |
| gzcloud19 | 211.159.188.153 | CWoU4r7D | Hadoop主管理 |
| gzcloud20 | 114.132.169.71 | CWoU4r7D | HDFS数据节点、YARN计算节点、TRINO计算节点 |
| gzcloud21 | 43.138.183.154 | CWoU4r7D | HDFS数据节点、YARN计算节点、TRINO计算节点 |
| gzcloud22 | 139.199.180.66 | CWoU4r7D | HDFS数据节点、YARN计算节点、TRINO计算节点 |
| gzcloud23 | 43.138.196.222 | 2qyKH~%\*)(ZfL&c | HDFS数据节点、YARN计算节点、TRINO计算节点 |

”

Hadoop、Trino平台相关进程

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 大数据集群 | | 分配内存（G） | | | | | | | | | 关键组件计算内存 |
| 组件 | 组件实例 | gzcloud19 (主管理) | gzcloud12 (从管理) | gzcloud20 | gzcloud21 | gzcloud22 | gzcloud23 | gzcloud13 | gzcloud11 (trino管理) | gzcloud10 |
| 服务器带宽(M) |  | 1 | 1 | 10 | 1 | 1 | 5 | 1 | 5 | 1 |  |
| 汇总 | | 63 | 60 | 54 | 76 | 75 | 67 | 68 | 42 | 56 |  |
| 服务器规划内存 | | 64 | 64 | 64 | 64 | 64 | 64 | 64 | 64 | 64 |  |
| HDFS | NameNode | 15 | 15 |  |  |  |  |  |  |  |  |
| Failover Controller | 1 | 1 |  |  |  |  |  |  |  |  |
| JournalNode | 2 | 2 |  | 1 |  |  |  |  |  |  |
| DataNode |  |  | 12 | 12 | 12 | 12 | 12 | 12 |  | 72 |
| SPARK | History Server | 2 |  |  |  |  |  |  |  |  |  |
| gateway | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |  |  |
| HIVE | Metastore Server |  | 4 |  |  |  |  |  |  |  |  |
| HiveServer2 |  | 2 |  |  |  |  |  |  |  |  |
| gateway | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |  |  |
| HBASE | Master | 2 | 2 |  |  |  |  |  |  |  |  |
| RegionServer |  |  | 4 | 4 | 4 |  |  |  |  |  |
| CM | ClouderaManager | 10 |  |  |  |  |  |  |  |  |  |
| ZOOKEEPER | Server |  |  | 1 | 1 | 1 |  |  |  |  |  |
| KAFKA | Kafka Broker |  |  | 2 | 2 | 2 |  |  |  |  |  |
| YARN | JobHistory Server | 2 |  |  |  |  |  |  |  |  |  |
| ResourceManager | 2 | 2 |  |  |  |  |  |  |  |  |
| NodeManager |  |  | 20 | 20 | 20 | 20 | 20 |  |  | 100 |
| TRINO | coordinator |  |  |  |  |  |  |  | 30 |  |  |
| worker |  | 32 |  | 30 | 30 | 30 | 30 |  |  | 150 |
| MYSQL | Server | 12 |  |  |  |  |  |  |  |  |  |
| KUDU | master/tserver |  |  | 4 | 4 |  |  | 4 |  |  |  |

大数据集群的服务器中的进程：

所有trino用户的进程为TRINO集群的主进程

所有用户为HDFS、HBASE、YARN、HIVE、SPARK的进程为相应大数据组件的主进程

大数据组件的主进程相应功能可以通过Cdh管理平台中查看

Cdh管理平台：http:// 114.132.169.71:7180/

用户名：admin

密码：admin!@#$%^

所有自己开发部署的程序的进程都在bigdata用户下，主要分为下面几类：

Flume数据发送和接收（flume）：ps –ef|grep flume

大数据任务定时调度（task-server）：ps –ef|grep V2Server

Spark Streaming定时入库任务（stream）：ps –ef|grep stream

## gzcloud11

### 1.flume：cos上传状态监控

部署位置：/data/bigdata/flume

脚本：37088-quality-flume.sh

配置文件：flume\_quality\_37088.conf

描述：flume接收日志上传cos状态信息，把信息存到mysql中，在页面可以查看每天的cos日志是否有正常上传。

## gzcloud12

无

## gzcloud13

无

## gzcloud19

### 1.flume：关键项目订单监控

部署位置：/data/bigdata/flume

脚本：9499\_check\_pay.sh

配置文件：flume\_pay\_check.conf

描述：flume接收项目组订单库监控信息，为关键项目添加订单监控，把每小时每个区的订单条数信息存到mysql中，监控程序会检查如果连续6个小时无订单数据就告警。

### 2.flume：hbase日志接收

部署位置：/data/bigdata/hbaseflume

脚本：9083-avro-flume.sh

配置文件：flume\_avro\_9083.conf

描述：flume接收服务端日志，把日志存入hbase中，供gmt查询日志数据，之前使用的项目已经下线，暂时没有项目用到。

### 3.stream：trino实时导入

部署位置：/data/bigdata/mutilstream

脚本：load\_huntress\_jp.sh

配置文件：gzcloud21平台->流程->实时->huntress\_japan\_load\_realtime

描述：spark streaming实时入trino进程，处理项目：huntress\_jp、huntress\_tw、huntress\_vn、tank\_jp的kafka服务端日志

### 4.stream：trino实时导入

部署位置：/data/bigdata/qkadjstream

脚本：load\_qkslg\_adjevent.sh

配置文件：gzcloud21平台->流程->实时->qkslg\_adjevent\_realtime\_load

描述：spark streaming实时入trino进程，处理项目：骑砍SLG的adjust日志里面新手引导等前端日志

### 5.stream：trino实时导入

部署位置：/data/bigdata/streamadj

脚本：load\_adjust\_old.sh

配置文件：gzcloud21平台->流程->实时->dow\_adjust\_stream\_load\_new

描述：spark streaming实时入trino进程，处理老项目的adjust数据：dow\_jp、huntress\_jp、huntress\_om、huntress\_tw、warship\_jp、tank\_jp

## gzcloud20

### 1.flume：dow日志接收

部署位置：/data/bigdata/flume

脚本：start\_9450.sh

配置文件：flume\_trino\_dow\_jp\_9450.conf

描述：接受从香港中转的dow\_jp的实时日志，写入kafka中。

### 3.task-server：Adjust数据入库

部署位置：/data/bigdata/bigdata-scheduling-adj

脚本：bin/start.sh

配置文件：conf/server.conf

调度标识：77

描述：定时调用平台配置的调度标识为77的任务，主要是老项目adjust日志按天定时入trino。

### 4.task-server：Adjust数据入库

部署位置：/data/bigdata/bigdata-scheduling-nadj

脚本：bin/start.sh

配置文件：conf/server.conf

调度标识：78

描述：定时调用平台配置的调度标识为78的任务，主要是新项目的adjust日志按天定时入trino，因为adjust表的结构有变化，需要重新起一个调度（目前就一个一菇当关的adjust日志使用此调度器）。

### 5.task-server：Adjust数据入库

部署位置：/data/bigdata/bigdata-scheduling-fadj

脚本：bin/start.sh

配置文件：conf/server.conf

调度标识：79

描述：定时调用平台配置的调度标识为79的任务，主要是新项目的adjust日志按天定时入trino，最新版对timezone的定义做了调整（骑砍SLG的adjust使用此调度）。

### 6.task-server：GS日志数据入库

部署位置：/data/bigdata/bigdata-scheduling-load

脚本：bin/start.sh

配置文件：conf/server.conf

调度标识：111

描述：定时调用平台配置的调度标识为111的任务，项目GS日志按天定时导入trino。

### 7.task-server：GS日志数据入库

部署位置：/data/bigdata/bigdata-scheduling-kload

脚本：bin/start.sh

配置文件：conf/server.conf

调度标识：211

描述：定时调用平台配置的调度标识为211的任务，项目GS日志按天定时导入trino，和上一个调度相比，更改了对于空字符串的处理逻辑，此版会把空字符串转为NULL存到trino中（骑砍SLG使用此调度）。

### 8.stream：trino实时导入

部署位置：/data/bigdata/mutilstream

脚本：load\_tankws2\_pc.sh

配置文件：gzcloud21平台->流程->实时->ws2\_warship2\_load

描述：spark streaming实时入trino进程，处理项目：warship2\_asia、warship2\_cn、modernship\_cn、warship\_jp的kafka服务端日志

## gzcloud21

### 1.tomcat：大数据任务平台

部署位置：/data/bigdata/apache-tomcat-8.5.70

脚本：bin/startup.sh

部署文件：/data/bigdata/apache-tomcat-8.5.70/bigdata/ROOT

描述：大数据任务平台，数据导入trino以及新版通用统计使用。

访问地址：http://43.138.183.154:9090/main

admin

shufen@#$%

### 2.stream：trino实时导入

部署位置：/data/bigdata/streamdow

脚本：load\_dow\_new.sh

配置文件：gzcloud21平台->流程->实时->dow\_jpnew\_realtime\_load

描述：spark streaming实时入trino进程，处理项目：huntress\_kr、huntress\_rus、redwar\_pc、dow\_jp的kafka服务端日志

### 3.task-server：日志数据入KUDU库

部署位置：/data/bigdata/bigdata-scheduling

脚本：bin/start.sh

配置文件：conf/server.conf

调度标识：33

描述：定时调用平台配置的调度标识为33的任务，主要处理Adjust cost等需要更新的数据入kudu库。

### 4.task-server：日志数据入库

部署位置：/data/bigdata/bigdata-scheduling-his

脚本：bin/start.sh

配置文件：conf/server.conf

调度标识：66

描述：定时调用平台配置的调度标识为66的任务，主要处理早期项目支付数据入trino。

### 5.task-server：日志数据入库

部署位置：/data/bigdata/bigdata-scheduling-json

脚本：bin/start.sh

配置文件：conf/server.conf

调度标识：37

描述：定时调用平台配置的调度标识为37的任务，为将来上线的骑砍新开发，解析处理json格式日志并入trino。

## gzcloud22

### 1.tomcat：大数据任务平台

部署位置：/data/bigdata/apache-tomcat-8.5.70

脚本：bin/startup.sh

部署文件：/data/bigdata/apache-tomcat-8.5.70/bigdata/ROOT

描述：大数据任务平台，由老集群迁移过来的老项目统计任务。

访问地址：

http://139.199.180.66:9090/main

admin

admin!@#$%^

### 2.flume：redwar日志接收

部署位置：/data/bigdata/trinoflume

脚本：start-flume-9451.sh

配置文件：flume\_trino\_redwarws2\_9451.conf

描述：集群flume接收端，接收实时日志并写入kafka：redwar\_pc、warship2\_asia、warship2\_cn、modernship，因为数数所需的日志要从此处取，所有 redwar\_pc除了入kafka外还要落到本地磁盘。

### 3.flume：数数redwar日志发送

部署位置：/data/bigdata/ssflume

脚本：start-flume-send19.sh

配置文件：flume\_ss\_redwar\_send19\_9485.conf

描述：读取前面落地的redwar实时数据，发送给数数入库的服务器

### 4.stream：trino实时导入

部署位置：/data/bigdata/streamadj

脚本：load\_adj\_qkslg\_om.sh

配置文件：gzcloud21平台->流程->实时->blockwar\_adjust\_realtime\_load

描述：spark streaming实时入adjust数据到trino进程，处理项目：骑砍SLG的kafka Adjust日志。

### 5.stream：trino实时导入

部署位置：/data/bigdata/streamqkslg

脚本：load\_qkslg\_cn.sh

配置文件：gzcloud21平台->流程->实时->qkslg\_om\_realtime\_load

描述：spark streaming实时入GS数据到trino进程，处理项目：骑砍SLG的kafka服务端日志。

### 6.task-server：老统计

部署位置：/data/bigdata/bigdata-scheduling

脚本：bin/start.sh

配置文件：conf/server.conf

调度标识：22

描述：定时调用平台配置的调度标识为22的任务，主要处理老统计平台（http://139.199.180.66:9090/main）统计任务的定时调度，从老集群中迁移过来。

### 7.task-server：老统计2

部署位置：/data/bigdata/bigdata-scheduling-new

脚本：bin/start.sh

配置文件：conf/server.conf

调度标识：223

描述：定时调用平台配置的调度标识为223的任务，主要处理老统计平台（http://139.199.180.66:9090/main）统计任务的定时调度，和前面相比，修改了部分原始计算逻辑，dow统计中用到。

## gzcloud23

### 1.flume：adjust日志接收

部署位置：/data/bigdata/trinoflume

脚本：start-flume-9451.sh

配置文件：flume\_trino\_adjust\_9451.conf

描述：集群flume接收端，接收实时香港服务器发送的adjust日志并写入kafka：huntress\_jp、huntress\_tw、huntress\_om、tank\_jp、warship\_jp、warship2\_asia、qkslg\_om。

### 2.flume：GS日志接收

部署位置：/data/bigdata/trinoflume

脚本：start-flume-9452.sh

配置文件：flume\_trino\_wstank\_9452.conf

描述：集群flume接收端，接收实时香港服务器发送的gs日志并写入kafka： warship\_jp、tank\_jp、qkslg\_om。

### 3.flume：GS日志接收

部署位置：/data/bigdata/trinoflume

脚本：start-flume-9453.sh

配置文件：flume\_trino\_huntress\_new\_9453.conf

描述：集群flume接收端，接收实时香港服务器发送的gs日志并写入kafka： huntress\_jp、huntress\_om、huntress\_tw、huntress\_kr、huntress\_rus、huntress\_vn。

# 二、香港服务器

|  |  |  |  |
| --- | --- | --- | --- |
| cloud1 | 119.28.83.11 | HKCDH!@34%^spark | 日志中转 |
| cloud2 | 119.28.84.97 | HKCDH!@34%^spark | 日志中转 |
| cloud3 | 119.28.41.240 | HKCDH!@34%^spark | 日志中转 |

香港服务器主要部署flume，通用规则：

/data/bigdata/ssflume 下为数据接收flume，接收项目组日志服务器的数据，然后落地本地。

本地路径为 /data1/sslog，这下面日志为trino的flume和数数logbus共用

/data/bigdata/trinoflume 下面为trino日志采集flume，读取上面文件夹内日志，传回国内入到kafka里面

Adjust日志接收程序生成文件也在香港服务器，flume也会读取这些文件传回国内入kafka

## cloud1

### 1.flume：日志接收

路径/data/bigdata/ssflume

|  |  |  |
| --- | --- | --- |
| 配置文件 | 启动脚本 | 说明 |
| flume\_ss\_huntress\_jptw\_9461.conf | start-flume-ss-9461.sh | 龙女日本、龙女台湾 |
| flume\_ss\_huntress\_sea\_rus\_9462.conf | start-flume-ss-9462.sh | 龙女东南亚、俄罗斯 |
| flume\_ss\_huntress\_kr\_vn\_9463.conf | start-flume-ss-9463.sh | 龙女韩国、越南 |
| flume\_ss\_warship\_tank\_9464.conf | start-flume-ss-9464.sh | 海战日本、坦克日本 |
| flume\_ss\_qkslg\_9465.conf | start-flume-ss-9465.sh | 骑砍SLG日志 |
| flume\_resend\_paycheck\_9499.conf | start\_flume\_paycheck\_9499.sh | 转发关键项目订单监控日志 |

### 2.flume：日志发送

路径/data/bigdata/trinoflume

|  |  |  |
| --- | --- | --- |
| 配置文件 | 启动脚本 | 说明 |
| warship\_tank\_jp\_send23\_9452.conf | start-flume-gzcloud23-9452.sh | 海战日本、坦克日本日志和支付 |
| huntress\_log\_send23\_9453.conf | start-logflume-gz23-9453.sh | 龙女日本、韩国、台湾、越南、俄罗斯日志 |
| huntress\_pay\_send23\_9453.conf | start-payflume-gz23-9453.sh | 龙女日本、韩国、台湾、越南、俄罗斯支付 |
| qkslg\_om\_send23\_9452.conf | start-qkslg-gz23-9452.sh | 骑砍SLG日志和支付 |
| flume\_trino\_adjust\_new\_send23\_9451.conf | adjust-flume-send23-9451.sh | 骑砍SLG Adjust日志 |

### 3.flume：支付入库

路径/data/bigdata/statflume

|  |  |  |
| --- | --- | --- |
| 配置文件 | 启动脚本 | 说明 |
| flume\_warship\_jp\_pay.conf | start-flume-warshipjp-paydb.sh | 海战日本支付实时入mysql，统计top500支付页面用 |
| flume\_tank\_jp\_pay.conf | start-flume-tankjp-paydb.sh | 坦克日本支付实时入mysql，统计top500支付页面用 |

## cloud2

### 1.flume：日志接收

路径/data/bigdata/ssflume

|  |  |  |
| --- | --- | --- |
| 配置文件 | 启动脚本 | 说明 |
| flume\_ss\_huntressom\_9461.conf | start\_flume\_huntressom\_9461.sh | 龙女欧美 |
| flume\_resend\_paycheck\_9499.conf | start\_flume\_paycheck\_9499.sh | 转发关键项目订单监控日志 |

### 2.flume：日志发送

路径/data/bigdata/trinoflume

|  |  |  |
| --- | --- | --- |
| 配置文件 | 启动脚本 | 说明 |
| flume\_trino\_huntress\_om\_9451.conf | trino-flume-9451.sh | 龙女欧美日志和支付 |
| flume\_trino\_adjust\_send23\_9451.conf | adj-flume-9451.sh | 龙女日本、坦克日本Adjust日志 |

## cloud3

### 1.flume：日志接收

路径/data/bigdata/ssflume

|  |  |  |
| --- | --- | --- |
| 配置文件 | 启动脚本 | 说明 |
| flume\_ss\_warship2\_9461.conf | start-flume-ss-9462.sh | 海战2亚服日志和支付接收 |

路径/data/bigdata/dowflume

|  |  |  |
| --- | --- | --- |
| 配置文件 | 启动脚本 | 说明 |
| flume\_dow\_ss\_9487.conf | start\_flume\_dow\_ssnew.sh | 战国日本日志接收 |
| flume\_dow\_ss\_9485.conf | start\_flume\_dow\_sspay.sh | 战国日本支付接收 |

### 2.flume：日志发送

路径/data/bigdata/trinoflume

|  |  |  |
| --- | --- | --- |
| 配置文件 | 启动脚本 | 说明 |
| flume\_trino\_downew\_send20\_9450.conf | trino-flume-send20-9450.sh | 战国日本日志和支付 |
| flume\_trino\_ws2\_send22\_9451.conf | trino-flume-send22-9451.sh | 海战2日志和支付 |
| flume\_trino\_huntress\_adj\_9451.conf | trino-adj-flume-9451.sh | 龙女欧美、龙女台湾、海战日本 Adjust日志 |