# 詹芮线上任务

## Oozie定时调度任务

### Zaful推荐相关

#### zaful\_pc\_list\_log

崔剑需求：类目页PC/M列表页埋点数据清洗，提供给算法组使用

相关wiki：<http://wiki.hqygou.com:8090/pages/viewpage.action?pageId=121143722>

#### zf\_old\_user\_device\_SNB25641\_3

Zaful冷启动区分新老用户需求，写入**我们redi**s，供接口调用

相关wiki：<http://wiki.hqygou.com:8090/pages/viewpage.action?pageId=93620078>

1. zf\_old\_user\_device\_hive 区分新老用户hive脚本
2. old\_user\_device\_toredis 老用户数据写入redis（算法的和推荐的）

#### ~~zf\_old\_user\_device2redis\_SNB25641\_3~~

~~李龙需求：Zaful新老用户写入~~**~~算法Redis~~**~~，供算法组使用~~

1. ~~zf\_old\_user\_device2redis 老用户数据写入算法redis，代码一样，打包时配置文件替换为算法的redis集群~~

~~合并到了zf\_old\_user\_device\_SNB25641\_3这个任务~~

#### zaful\_pc\_list\_plat\_country\_SNB25641\_3

彭富需求：Zaful类目页算法推荐基础数据整理，导入算法Mongo供算法组使用

1. zaful\_pc\_list\_plat\_country\_hive 类目页推荐基础数据整理hive脚本
2. hive\_to\_mongodb 数据导入算法Mongo zaful\_pc\_list\_plat\_country
3. hive\_to\_mongodb2 数据导入算法Mongo zaful\_pc\_list\_plat\_global
4. doneflag\_toredis 数据同步完成标识导入算法redis

done\_flag：

/bigdata/log\_clean/job/done\_flag/

/bigdata/ods/sqoop/zf/ods\_zaful\_wf01/done\_flag/

#### zaful\_ods\_base\_SNB25641\_3

Zaful网站推荐基础数据，关联各种商品信息（pipline、lang等等），汇总成最终的商品信息表，写入Hbase。也包含网站的新品，热销等推荐后补数据写入Redis。

1. zaful\_ods\_base zaful商品基本信息整理hive脚本，重要任务，参考刘庆凡历史任务zaful\_base.q 做的stg迁移ods数据，增加pipline\_code维度。Zaful每增加一个国家站和语言后台会新增表，需要大数据配合抽数，同时脚本也需要修改。
2. apl\_result\_hotsell\_cat\_zaful\_fact\_to\_redis 推荐后补数据写入redis
3. apl\_result\_newgoods\_zaful\_fact\_to\_redis 推荐后补数据写入redis
4. apl\_result\_hotsell\_all\_zaful\_fact\_to\_redis 推荐后补数据写入redis
5. apl\_zaful\_goods\_info\_tohbase 商品信息写入Hbase，hive表是用ods数据做的带pipline维度的商品信息
6. apl\_zaful\_goods\_info\_stg\_tohbase 商品信息写入Hbase，hive表是用老的stg商品信息

#### zaful\_recommend\_base\_data\_SNB25641\_1

重要任务：Zaful推荐相关基础数据（网站的用户埋点行为数据，网站后台订单信息）整理，导入算法Mongo供算法组使用

1. base\_data\_hive 算法所需推荐基础数据准备，包括埋点行为数据，后台订单数据，商品在售信息（emp推荐中用到，过滤上下架无库存）
2. email\_zaful\_pc\_event\_info2mongo 埋点行为数据写入算法Mongo
3. email\_zaful\_goods\_event\_info2mongo 后台订单数据写入算法Mongo
4. email\_emp\_zaful\_onsale2mongo 商品在售信息写入算法Mongo
5. doneflag\_toredis 数据同步完成标识导入算法redis

### GB 推荐相关

#### recommend\_soa\_detail\_page\_sponsored\_SNB25641\_3

GB 商详页第四推荐位按规则推荐，随机推荐DA商品

1. recommend\_soa\_detail\_page\_sponsored\_hive 推荐数据准备hive脚本
2. recommend\_soa\_detail\_page\_sponsored\_redis 推荐结果写入redis

#### gb\_black\_gray\_result\_SNB25641\_3

GB 网站商品黑灰标打标

相关wiki：<http://wiki.hqygou.com:8090/pages/viewpage.action?pageId=117638923>

1. gb\_black\_gray\_result\_hive 黑灰标打标hive脚本，其中关联的3a，2a商品是网站提供的csv文件将sku手动导入的表 不会更新
2. data\_delete 之前导给搜索组使用，全量删除搜索组后台表，目前已停用
3. data\_push 导给搜索组使用，采用sqoop增量更新，目前已停用。如果今后有增量更新的需求可以参考该配置，需注意sqoop的update-key对应的mysql字段应该为唯一键

#### gb\_recommend\_base\_data\_SNB25641\_1

葛腾青，张哲

重要任务：GB推荐相关基础数据（网站的用户埋点行为数据，网站后台订单信息）整理，导入算法Mongo供算法组使用

1. pipeline\_language\_import GB后台国家站语言对应关系表导入

2. base\_data\_hive 推荐基础数据准备

3. email\_gb\_pc\_event\_info2mongo 埋点行为数据导入算法Mongo email\_gb\_pc\_event\_info

4. email\_gb\_goods\_event\_info2mongo 后台订单数据导入算法Mongo email\_gb\_goods\_event\_info

done\_flag：

/bigdata/log\_clean/job/done\_flag/

/bigdata/ods/sqoop/gb/ods\_m\_gearbest\_gb\_order\_tables\_job02/done\_flag

/bigdata/ods/sqoop/gb/ods\_m\_gearbest\_gb\_order\_tables\_job04/done\_flag

#### ~~pc\_homepage\_flow\_day\_gb~~

~~GB PC首页瀑布流~~

~~相关wiki：<http://wiki.hqygou.com:8090/pages/viewpage.action?pageId=111577590>~~

~~按照~~

~~庆凡历史任务GB APP首页瀑布流：app\_homepage\_flow\_day\_gb 中的脚本即可，GB PC首页瀑布流之比APP瀑布流多了过滤敏感商品的需求，而且没有多馆区。~~

1. ~~gb\_flow\_pc 首页瀑布流规则hive脚本~~
2. ~~gb\_flow\_pc\_da首页瀑布流规则hive脚本，插入DA商品~~
3. ~~result\_toredis 推荐数据写入redis~~

~~已换成新的pc\_homepage\_waterfall\_flow~~

#### pc\_homepage\_waterfall\_flow

GB PC首页瀑布流

1. 瀑布流按规则计算脚本
2. 脚本写入redis

#### recommend\_gb\_backup-coord

GB 推荐后补数据

按pipeline\_lang\_categoryid分组，每组至多50个后补商品，后补数据集写入Redis

按pipeline\_lang分组，每组至多1000个后补商品，后补数据集写入Redis

1. recommend\_gb\_backup\_hive 推荐有分类后补数据hive脚本
2. recommend\_gb\_backup\_noctg\_hive 推荐无分类后补数据hive脚本
3. recommend\_gb\_backup\_redis 有分类后补数据导入redis
4. recommend\_gb\_recommend\_gb\_backup\_noctg\_redis\_redis 无分类后补数据导入redis

#### goods\_bf\_result-coord

GB 网采商品禁售国家和非禁售SKU，结果写入redis，推荐过滤网采使用

1. recommend\_gb\_bf\_hive 过滤网采商品hive脚本
2. bf\_pipline\_country\_toredis 禁售国家写入redis
3. bf\_good\_sn\_toredis 非禁售商品写入redis

#### recommend\_gb\_static\_backup-coord

双十一做的降级后补静态数据

1. recommend\_gb\_static\_backup\_hive 降级后补数据准备hive脚本
2. recommend\_gb\_static\_backup\_redis 降级后补数据写入redis

### 临时任务

#### ~~bts\_import-coord~~

~~帮数据分析师 谢海滨做的定时同步bts后台表的任务，方便他做bts相关数据分析~~

~~一直失败，mysql连不上，应该不用了，killed了~~

#### ~~gb\_sales\_list\_detail-coord~~

~~算法临时需求，GB在售SKU统计，是以前邮件用的，现在不用了~~

#### ~~gb\_sku\_info\_report-coord~~

~~算法临时需求，GB sku指标数据（曝光点击加购收藏等）统计，是以前邮件用的，现在不用了~~

### 报表相关

#### zaful\_search\_report-coord

殷铨均报表需求：Zaful PC 搜索页报表

#### zaful\_cat\_report-coord

殷铨均报表需求：Zaful PC 类目页报表

#### zf\_recommend\_position\_report-coord

殷铨均报表需求：Zaful PC 首页、商详页、购物车页报表

#### gb\_recommend\_position\_report-coord

按照Zaful推荐报表取数规则做的一套GB PC 推荐位报表

#### stg\_yoshop\_import\_SNB25641\_3

yoshop网站的抽数任务，统计yoshop在售商品，彭富做yoshop推荐使用

### EMP邮件推荐相关

#### gb\_goods\_country\_hotsell\_SNB25641\_3

GB 国家热销数据，写入Hbase

1. gb\_goods\_country\_hotsell\_hive 网站国家热销数据hive脚本
2. gb\_goods\_country\_plat\_hotsell\_hive 网站国家平台热销数据hive脚本
3. gb\_goods\_country\_hotsell\_tohbase 热销数据写入hbase
4. gb\_goods\_site\_multi\_lang\_category\_tohbase 分类信息写入hbase

#### result\_gb\_systm\_email\_fact\_SNB25641\_3

GB 系统邮件，将算法推荐结果写入Hbase

1. add\_partition 给算法推荐结果表添加分区，以便接下来从mongo导入hive
2. result\_gb\_systm\_email\_gtq\_tohive 从算法Mongo导入推荐结果导hive
3. result\_gb\_systm\_email\_fact\_hive 关联商品信息，生成推荐结果hive脚本
4. result\_gb\_systm\_email\_fact\_tohbase 将推荐结果写入hbase

#### gb\_email\_daily\_SNB25641\_3

GB 每日邮件，EMP提供的sql 逻辑，每日定时将数据写入Mongo供算法使用

1. gb\_review\_info\_hive 每日邮件基础数据准备hive脚本，逻辑由emp直接提供
2. hive\_to\_mongodb1 评论数据导入算法mongo
3. hive\_to\_mongodb2 订单数据导入算法mongo
4. doneflag\_toredis数据同步完成标识导入算法redis

#### gb\_furniture\_SNB25641\_3

GB 家具家电邮件，每两周发一次，定时将EMP提供的Hive表同步到Mongo供算法使用

1. hive\_to\_mongodb 将emp生成的hive表导入算法Mongo
2. finished2mongodb 以Post/Get方式通知算法数据已导入完成

#### gb\_computer\_SNB25641\_3

GB 电脑周边邮件，每周发一次，定时将EMP提供的Hive表同步到Mongo供算法使用

1. hive\_to\_mongodb 将emp生成的hive表导入算法Mongo
2. finished2mongodb 以Post/Get方式通知算法数据已导入完成

#### gb\_browse\_SNB25641\_3

GB 浏览记录邮件，每日定时将大数据生成的Hive表同步到Mongo供算法使用

1. hive\_to\_mongodb 将大数据生成的hive表导入算法Mongo，注意该任务配置了done\_flag依赖
2. finished2mongodb 以Post/Get方式通知算法数据已导入完成

#### gb\_cart\_SNB25641\_3

GB 购物车邮件，每日定时将大数据生成的Hive表同步到Mongo供算法使用

1. hive\_to\_mongodb 将大数据生成的hive表导入算法Mongo，注意该任务配置了done\_flag依赖
2. finished2mongodb以Post/Get方式通知算法数据已导入完成

#### zf\_goods\_country\_hotsell\_SNB25641\_3

Zaful 国家热销商品数据，写入Hbase

1. zf\_goods\_country\_hotsell\_hive Zaful 国家热销hive脚本
2. zaful\_country\_hotsell2Hbase Zaful国家热销数据写入hbase

#### ~~email\_emp\_zaful\_users\_SNB25641\_3~~

~~Zaful 优惠券营销邮件，每周按规则将数据写入Mongo供算法使用~~

1. ~~zaful\_user\_hive 优惠券营销用户数据hive脚本~~
2. ~~email\_emp\_zaful\_users2mongo 优惠券营销数据写入Mongo~~
3. ~~doneflag\_toredis 数据同步完成标识写入算法redis，以前的交互是用Post/Get方式，后来已经切换为redis~~

~~邮件表示不再使用，邱倩文、董腾飞、葛腾青已确认~~

#### dresslily\_recommend\_base\_data\_SNB25641\_1

D网推荐基础数据

1. base\_data\_hive基础数据脚本
2. email\_d\_pc\_event\_info2mongo 埋点数据导入mongodb
3. email\_d\_goods\_event\_info2mongo 订单数据导入mongodb
4. email\_emp\_d\_onsale2mongo 上下架数据导入mongodb

#### dl\_goods\_country\_hotsell\_SNB25641\_3

D网国家热销

1. dl\_goods\_country\_hotsell\_hive 商品信息+国家热销 脚本
2. apl\_dresslily\_goods\_info 商品信息导hbase
3. apl\_dl\_goods\_country\_hotsell 国家信息导hbase

## Spark代码

### Git地址

<http://gitlab.egomsl.com/yanhan/ai-bigdata> 分支：zhanrui

<http://gitlab.egomsl.com/liuqingfan/glbg-userbehaviour-hadoop-1.0.0-product-jar-with-dependencies.git>

Mr代码，分支：zhanrui（添加了CookieMongo2HiveMapper，CookieMongo2HiveMR，EMPMongo2HiveMapper，EMPMongo2HiveMR）

分支：master（其他的用这个）

### 相关说明

**AlgorithmRedisFlagCheck**

算法数据同步标志位检查

参数1：算法 redis key

参数2：算法 redis field

参数3：推荐 redis key

参数4：推荐 redis field

#### BFResult2Redis

GB过滤网采数据写入redis，包括网采禁售国家 和 非禁售SKU

参数1：hive表名

参数2：redis key

#### CommonResult2Hbase

通用工具类，将Hive数据导入Hbase

参数1：hbase表名

参数2：hive表名

参数3：rowkey下标，逗号分隔。如0,4 代表rowkey为hive表第0列和第4列拼接

注意事项，rowkey中hive表字段类型必须为string型

#### CommonResult2Redis

通用工具类，将Hive数据导入Redis

参数1：hive表名

参数2：redis key

主要做新老用户数据导入redis，只用来导hive表只有一列的情况

#### CommonSortResult2Hbase

通用工具类，将Hive数据（商品推荐结果）按照score排序导入Hbase

参数1：hbase 表名

参数2：hive表名

参数3：rowkey下标，逗号分隔

注意事项，rowkey中hive表字段类型必须为string，hive表最后一个字段为score

#### Detail1Result2Hbase

GB 商详页推荐位推荐结果写入hbase，

用spark sql关联算法推荐结果和商品信息表，GB第一二推荐位使用

参数1：hbase表名

参数2：算法推荐结果时间分区

参数3：算法推荐结果hive表

#### Detail3Result2Hbase

GB 商详页推荐位推荐结果写入hbase，

用spark sql关联算法推荐结果和商品信息表，GB第三推荐位使用

参数1：hbase表名

参数2：算法推荐结果时间分区

参数3：算法推荐结果hive表

参数4：Hfile个数，可用默认值

#### DoneFlag2Redis

通用工具类，数据同步完成标识写入算法Redis，结果为hashmap。例如同步gb相关的表到Mongo之后，算法redis写入HSET gb\_recommend\_base\_data email\_gb\_pc\_event\_info 20190101

参数1：redis hashmap value

参数2：redis key

参数3：redis hashmap key

注意算法redis集群和推荐redis是两个集群，连接方式不一样

#### DoneFlagChecker

通用工具类，还未使用到，作为检查算法redis标识工具

#### EmailResult2Hbase

同步庆凡apl\_emp\_gb EMP三封邮件 hive表到hbase的任务

参数1：hbase表名

参数2：hive表名

#### RecommendResult2Hbase

通用工具类，算法推荐商品结果写入hbase

参数1：hbase 表名

参数2：hive表名

参数3：rowkey下标，逗号分隔

参数4：hive表时间分区

参数5：是否分区表

参数6：结果是否需要反排序

说明：以前的推荐流程是将算法Mongo的数据导入hive之后，关联上商品信息写入hbase，现在无需关联商品信息，直接将算法Mongo的数据导入hive后，将hive表原样导入hbase，接口查询的时候多查一次商品信息即可。由于历史任务有些hive表分区有些没有，所以做了兼容，还有个问题是hive分区表最终写入hbase后score排序会反，也做的特殊处理，没查出具体原因。

# 刘庆凡线上任务

## Oozie定时调度任务

### GB调度任务

#### recommend\_soa\_common\_gb

后补数据任务调度，需确认哪些在用？

1、common HIVE离线计算任务脚本

2、apl\_sale\_top100\_pipeline\_fact apl\_topsale\_category50\_fact apl\_top10\_category\_fact 数据写入redis

3、apl\_goods\_info\_result\_rec\_gb\_fact\_tohbase 商品基本信息更新至HBASE表apl\_goods\_info\_result\_rec\_gb\_fact

#### recommend\_soa\_base\_gb

基础数据任务调度

1. ~~action\_goods\_category\_cl~~

~~action\_sqoop\_goods\_category~~

~~action\_goods\_category\_level\_cl~~

~~action\_goods\_category\_level~~

~~商品分类信息通过oracle递归处理 ORACLE服务器不能被回收——用ods.ods\_o\_gearbest\_bigdata\_goods\_category\_level替换~~

1. soa\_base 商品基本信息离线计算脚本

产生8个中间表：

商品信息中间表1（分类、上下架、仓库、状态等信息）goods\_info\_mid1

商品信息中间表2（网站展示用信息（图片、title））goods\_info\_mid2

商品信息中间表3（库存）goods\_info\_mid3

商品信息中间表3（评论）goods\_info\_mid4

SPU/SKU一一对应表goods\_info\_mid5

商品属性信息表goods\_info\_result

商品属性信息表sku\_info\_data

可被推荐商品属性信息表goods\_info\_result\_rec

1. soa\_hot\_sell\_step\_1\_1

soa\_hot\_sell\_step\_1\_2

soa\_hot\_sell\_step\_1\_3

soa\_hot\_sell\_step\_2\_1

soa\_hot\_sell\_step\_2\_2

算法后补结果 脚本算法（崔剑）提供

#### recommend\_soa\_unbrand\_gb

非品牌商品推荐任务调度

1. apl\_result\_rpfy\_unbrand\_gb\_fact\_to\_hbase Spark任务包含离线计算及写入Hbase

#### recommend\_soa\_pay\_gb

支付完成页任务调度

1. soa\_pay\_spark Spark任务包含离线计算及写入Hbase

#### recommend\_soa\_homepage\_gb

视频页推荐任务调度(首页)

1、soa\_homepage 根据规则离线计算脚本

2、apl\_result\_rtiyv\_gb\_cookie\_fact\_tohbase apl\_result\_rtiyv\_gb\_uncookie\_fact 规则结果写入Redis或Hbase

#### recommend\_soa\_user\_center\_gb

用户中心页推荐任务调度

1、female\_recommend 女性推荐规则离线计算

2、apl\_female\_ymal\_fact 女性推荐结果写入hbase

3、apl\_cart\_10\_fact 分类top10写入hbase

#### app\_homepage\_flow\_hour\_bf\_gb

APP/M端置顶商品任务调度（每半小时一次）

1、top\_goods\_settings 同步置顶信息（来自OBS）

2、webapp\_category\_library\_result\_bf 置顶商品信息关联（商品顺序同OBS）

3、webapp\_category\_library\_result\_bf\_toredis 置顶信息更新至Redis

4、gb\_app\_homepage\_trending\_rec\_bf\_toredis trending排序商品半小时查询价格为价格筛选准备

#### recommend\_gb\_detail\_abtest1

GB商详页AB测试任务调度

1. wait\_mongodb1

wait\_mongodb2

~~wait\_mongodb3~~ 算法任务调度（算法需求数据由其他任务提供）

1. ~~gb\_result\_detail\_3\_page\_gtq\_to\_hive 这个表不再更新了~~

gb\_result\_detail\_page\_gtq\_to\_hive

gb\_result\_detail\_1\_page\_gtq\_to\_hive 算法结果同步至Hive表

1. ~~apl\_gb\_result\_detail\_page\_3\_fact\_tohbase~~

apl\_gb\_result\_detail\_page\_2\_abtest1\_fact\_tohbase apl\_gb\_result\_detail\_page\_1\_abtest1\_fact\_tohbase 算法结果写入Hbase

#### rrecommend\_gb\_detail\_realtime

GB商详页AB测试任务调度

新的通信方式，算法计算完写入redis，每一个小时检测一次

1. 监测redis是否有新的时间写入

check\_mongodb\_gb\_result\_detail\_2\_page\_gtq\_abtest\_v1

check\_mongodb\_gb\_detail2\_result\_fpg

1. 设置表建分区

gb\_result\_detail\_2\_page\_gtq\_abtest\_v1\_add\_partition

gb\_detail2\_result\_fpg\_add\_partition

1. Mongodb导入hive

gb\_result\_detail\_2\_page\_gtq\_abtest\_v1\_mongo2hive

gb\_detail2\_result\_fpg\_mongo2hive

1. Hive数据通过spark聚合后写入hbase

gb\_result\_detail\_2\_page\_gtq\_abtest\_v1\_tohbase

gb\_detail2\_result\_fpg\_tohbase

#### app\_homepage\_flow\_day\_gb

APP/M 瀑布流每日端任务调度

1、site\_multi\_lang\_brand interest app\_index\_tab 商品信息同步（来自OBS及SOA）

2、trending\_data trending计算

3、gb\_app\_homepage\_trending\_rec\_da DA商品插入

4、gb\_app\_homepage\_trending\_thirty 每页后补（需求已经废弃）

5、gb\_app\_homepage\_trending\_toredis trending规则结果写入redis

6、gb\_app\_homepage\_cat\_name gb\_app\_homepage\_brand\_name 每个馆区的商品分类名称及品牌名称汇总

注意：调用HiveToRedisPriceMR会查一遍搜索的接口查soa的价格，更新价格

#### apl\_lable\_app\_fact\_gb

APP专享价任务调度

1、APP专享价任务（已停止），线上还在，严涵还在用

#### apl\_recommend\_lable\_gb

新品盈利任务调度

1. 新品盈利标签任务

### Zaful调度任务

#### stg\_zaful\_wf03

STG层ZAFUL后台任务调度，sqoop导数

1. eload\_goods\_id

eload\_goods\_ar

eload\_goods\_de ......（多语言数据表，每增加一个语言，增加一张表）

2、eload\_goods\_attr\_multi\_lang eload\_goods\_attr 多语言商品属性规则

ZAFUL数据同步至STG库，后续使用ODS库数据

zaful\_eload\_goods\_attr在用（recommend\_zaful\_list\_page，recommend\_zaful\_baseinfo），其他表全没用了

#### recommend\_homepage\_v2\_zaful

多巴胺首页瀑布流任务调度 (失败了一定要重跑，不然第二天跑不了)

1、ip\_country\_region 每日日志IP转换成country region

2、add\_partition\_result\_detail\_uncookie\_user 表加分区

3、click\_count\_region\_data click\_country\_data lr\_base\_info4\_hive 算法需求 离线脚本

4、country\_to\_mongo\_m country\_to\_mongo\_pc hive\_to\_mongodb 数据写入mongodb

5、wait\_mongodb 算法任务调度

6、mongodb\_to\_hive 算法结果同步至Hive表

7、apl\_result\_detail\_uncookie\_zaful\_fact 算法结果商品信息关联

8、hive\_to\_redis 算法结果写入redis

**Hive数据写入Mongodb表：**

lr\_base\_info\_country\_m 先停掉

lr\_base\_info\_country\_pc 先停掉

lr\_base\_info\_3 没用了

lr\_base\_info\_4 还在用，不过之前也是stg的数据，建议安排计划改到ods，优先级可以往后排

**结果数据写入redis:**

result\_zaful\_detail\_uncookie\_sort\_regionid\_platform\_lang 多巴胺后补数据，严涵说之前有用，暂时没用，以后不知道用不用

#### recommend\_zaful\_list\_page

原列表页算法数据任务调度

任务已经被（ODS数据）取代

zaful\_pc\_list\_plat\_country\_SNB25641\_3詹芮已重写

**导入mongodb表：**

zaful\_list\_plat 已替换

zaful\_list\_plat\_country 已替换

zaful\_list\_source\_sku 在用

zaful\_list\_source\_sku\_1 已不用

apl\_sku\_lang\_fact 待彭富确认

#### recommend\_zaful\_baseinfo

基础数据任务调度

1、zaful\_base 商品基本信息离线脚本

2、zaful\_hot\_sell\_step\_1\_1

zaful\_hot\_sell\_step\_1\_2

zaful\_hot\_sell\_step\_2（后补数据 脚本由崔剑提供）

1. apl\_result\_hotsell\_cat\_zaful\_fact\_to\_redis

apl\_result\_newgoods\_zaful\_fact\_to\_redis

apl\_result\_hotsell\_all\_zaful\_fact\_to\_redis

apl\_result\_detail\_cookie\_zaful\_abtest\_3\_fact apl\_result\_detail\_page\_abtest2\_fact\_tohbase\_fact apl\_result\_detail\_page\_abtest2\_fact\_tohbase\_new （AB测试基础算法及新品热销数据写入redis或hbase）

#### recommend\_detailpage\_al2rec\_zaful-coord

商详页算法数据数据同步到 hbase

1. wait\_mongodb\_w2v\_detail  
    wait\_mongodb\_abtestv5\_detail
2. zaful\_detail\_w2v 最优算法+人工干预 pc/m的数据 mongodb 到 hive  
    result\_detail\_page\_abtest\_v5 最优算法+人工干预 ios/android的数据 mongodb 到 hive
3. apl\_result\_detailpage\_zaful\_abtest\_w2v\_tohbase\_new 算法数据 hive 到 hbase  
    apl\_result\_detailpage\_zaful\_abtest\_v5\_tohbase\_new 算法数据 hive 到 hbase

#### recommend\_detailpage\_zaful

商详页任务调度，包含购物车原始算法

1、add\_partition\_result\_goodsdetail\_lr 增加分区

~~2、zf\_pc\_event\_info\_detail\_exp\_to\_mongodb (李龙使用ods推的数据)~~

~~zaful\_eload\_order\_detail\_exp\_to\_mongodb (李龙使用ods推的数据)~~

~~zaful\_eload\_goods\_to\_mongodb 算法需求数据（已经被ODS数据取代，数据来自于其他任务）~~

1. wait\_mongodb

wait\_mongodb\_cart

~~wait\_mongodb4\_detail （改为新的同步服务）~~

~~wait\_mongodb5\_detail AB测试算法调度（李龙，崔剑）（改为新的同步服务）~~

1. result\_datail\_page\_to\_hive

result\_cart\_page\_to\_hive

result\_detail\_page\_abtest\_v4

result\_detail\_page\_abtest\_v5 算法结果同步至Hive表

1. apl\_result\_cartpage\_zaful\_fact

apl\_result\_detailpage\_zaful\_fact

apl\_result\_detail\_page\_abtest3\_fact

apl\_result\_detailpage\_zaful\_abtest\_v5 算法结果关联商品信息

1. apl\_result\_detailpage\_tohbase\_zaful\_fact

apl\_result\_detailpage\_tohbase\_zaful\_fact\_old

apl\_result\_cartpage\_tohbase\_zaful\_fact apl\_result\_detail\_page\_abtest1\_fact\_tohbase\_fact apl\_result\_detail\_page\_abtest3\_fact\_tohbase\_fact apl\_result\_detail\_page\_abtest3\_fact\_tohbase\_new apl\_result\_detailpage\_zaful\_abtest\_v5\_tohbase\_fact 算法结果写入Hbase或redis

### EMP 邮件推荐相关

#### apl\_emp\_gb

EMP三封邮件调度，下单成功发一封，过一个小时没付款发一封......

1、add\_partition\_result\_emp\_data 分区表加分区

2、emp\_data 算法（张哲）提供脚本

3、gb\_goods\_info\_result\_detail\_exp\_to\_mongodb 数据写入mongodb

4、doneflag\_toredis 写入完成标志

5、wait\_mongodb 等待算法返回结果

6、result\_gb\_email\_create\_to\_hive

result\_gb\_email\_payed\_to\_hive

result\_gb\_email\_promptpay\_to\_hive 算法结果同步至Hive表

7、apl\_result\_gb\_email\_create\_fact

apl\_result\_gb\_email\_payed\_fact

apl\_result\_gb\_email\_promptpay\_fact 结果关联商品信息

8、apl\_result\_gb\_email\_create\_fact\_tohbase

apl\_result\_gb\_email\_payed\_fact\_tohbase

apl\_result\_gb\_email\_promptpay\_fact\_tohbase 算法结果写入Hbase

9、apl\_sale\_top100\_pipeline\_fact 后补数据写入redis

# 吴超线上任务

## Oozie定时调度任务

### Za**ful推荐相关**

注：（zaful推荐相关都是导入mongodb线上直接调用，需要保证每天更新）

#### zaful\_app\_sku\_cat\_id\_mongodb\_report\_SNB25641\_1

彭富需求：zf的推荐系统的列表页的商品信息分国家

#### zaful\_app\_sku\_cat\_id\_no\_country\_mongodb\_report\_SNB25641\_1

彭富需求：zf的推荐系统的列表页的商品信息不分国家

#### zaful\_app\_device\_id\_sku\_time\_report\_SNB25641\_1

崔剑需求：zf的推荐系统的订单信息

#### zaful\_app\_statistics\_result-coord\_SNB25641\_1

张哲需求：zf的推荐系统的商品统计信息

#### zaful\_app\_statistics\_homepage\_result\_coord

张哲需求：算法首页推荐位商品信息统计，专针对首页的

#### Media\_source\_zaful\_app\_device\_id\_sku\_to\_mongodb\_report

张哲需求：APP点击埋点数据

#### ~~af\_inner\_zaful\_app\_device\_id\_sku\_to\_mongodb\_report~~

~~崔剑需求：算法zaful用户订单信息，已没用~~

#### zaful\_app\_category\_planid\_report

彭富需求：算法列表页曝光，点击，加购，订单等统计信息

### RG**邮件相关**

#### ~~email\_emp\_rosegal\_users-coord\_SNB25641\_1~~

~~rg的emp邮件用户信息，写mongodb表email\_emp\_rg\_users~~

~~已停掉，邱倩文、葛腾青、董腾飞已确认~~

#### rosegal\_recommend\_base\_data-coord\_SNB25641\_1

rg的emp邮件基础信息，写mongodb表email\_rg\_pc\_event\_info,email\_rg\_goods\_event\_info

#### rg\_goods\_country\_hotsell-coord\_SNB25641\_1

rg的emp邮件热销商品信息，写hbase表apl\_rg\_goods\_country\_hotsell，apl\_rosegal\_goods\_info

#### email\_emp\_rosegal\_goods\_onsale-coord\_SNB25641\_1

rg的emp邮件在售商品信息，写mongodb表email\_emp\_rg\_onsale

### 数据分析相关

#### zaful\_catogory\_ios\_count

夏雪数据分析跑数任务

### BTS实验相关

#### ai\_bts\_prod\_planid\_416

bts实验

#### ai\_bts\_prod\_planid\_391

bts实验

# 张玉超线上任务

## Oozie定时调度任务

### Za**ful推荐相关**

### GB推荐相关

#### gb\_app\_old\_user\_device\_id\_ninety\_day-wf\_SNB25641\_3

新用户：90天内，未访问过网站的非注册用户

老用户：已注册用户和90天内访问过网站的非注册用户

#### gb\_new\_product\_SNB25641\_3

新品、A+商品数据处理

label\_sku\_to\_hive\_spark 标签系统里的商品同步到hive

new\_product\_recommend\_hive 新品计算

label\_sku\_hive\_to\_hbase\_spark 新品导入到hbase

label\_sku\_hbase\_cleaner 清除历史数据

label\_a\_plus\_recommend\_hive A+商品计算

label\_a\_plus\_toredis A+商品导入redis

### BTS实验相关