# 海狗2安装指南

## 海狗性能指标

* 总数据量500亿，100~400个维度.
* 单表最大存储500亿。
* 数据表总数小于10。
* 单表每天数据增量小于1亿。
* 单次扫描最大时间跨度为该时间跨度下数据量小于30亿行。
* 单次扫描的列的总数不超过12列。

## 推荐硬件

* 机器数量：10~12台
* CPU:E5-2630(6核)\*2
* 内存：>=48GB
* 硬盘规格数量：>=480GB\*12 （SSD最佳，普通磁盘也可）。
* RAID要求：最好将12块盘raid成统一的一块，raid 1即可

## 源码编译

在linux环境下使用maven编译

mvn clean && mvn package -DskipTests assembly:assembly

编译好的程序在/target/alipay-bluewhale.tar.gz 下

## 安装海狗依赖环境

* zookeeper安装

zookeeper:<http://zookeeper.apache.org/>

推荐使用zookeeper-3.4.5版本

请注意，zoo.cfg的如下几个参数按下面来配置

tickTime=2000

initLimit=10

syncLimit=5

maxClientCnxns=300

maxSessionTimeout=20000

另外zk运行久了，硬盘占用会越来越多，可以使用如下命令清理

cd /xxxx/zookeeper

java -cp

./zookeeper-3.4.5.jar:./lib/log4j-1.2.15.jar:./lib/slf4j-api-1.6.1.jar:./conf

org.apache.zookeeper.server.PurgeTxnLog /xxx/data/zook /xxx/data/zook -n 100

* 部署hadoop环境

hadoop:<http://hadoop.apache.org/>

* ZeroMQ 2.1.7

tar -xzf zeromq-2.1.7.tar.gz

cd zeromq-2.1.7

./configure

make

sudo make install

* JZMQ

./autogen.sh

./configure

make

sudo make install

* Java 6
* Python 2.6.6
* unzip

## 环境变量配置

* 每台机器配置好hostname
* shell环境变量，编辑~/.bashrc

export HADOOP\_CONF\_DIR=$HOME/config

export HADOOP\_HOME=/home/hadoop/hadoop-current

export HADOOP\_ROOT=/home/hadoop

export HADOOP\_HEAPSIZE=128

export PATH=$PATH:$HADOOP\_HOME/bin

## 蓝鲸配置

将编译好的/target/alipay-bluewhale.tar.gz解压

进入/xxxx/bluewhale/alipay/bluewhale-core/conf

编辑storm.yaml

一般我们在每台机器上配置6个shard，1个merger server,均分配5GB的内存

配置文件说明参考如下，可以根据实际情况修改

#zookeeper的连接

storm.zookeeper.servers:

- "adhoc2.kgb.cm6"

- "adhoc3.kgb.cm6"

- "adhoc4.kgb.cm6"

#无用已经作废

hadoop.merger.conf.dir: "/home/taobao/hadoop/conf"

#hadoop mapreduce，在创建海狗索引中可以使用的内存

hadoop.java.opts: "-Xmx1024m"

#蓝鲸任务的心跳超时时间

nimbus.task.timeout.secs: 240

#蓝鲸任务的心跳频率

task.heartbeat.frequency.secs: 12

#nimbus认为一个supervisor多久没有心跳，认为其超时

nimbus.supervisor.timeout.secs: 360

#supervisor心跳频率

supervisor.heartbeat.frequency.secs: 10

#nimbus多久check一次心跳

nimbus.monitor.freq.secs: 20

#supervisor认为一个worker多久没有心跳，认为其超时

supervisor.worker.timeout.secs: 240

#supervisor多久check一次心跳

supervisor.monitor.frequency.secs: 12

#worker多久给supervisor发送一次心跳

worker.heartbeat.frequency.secs: 4

#zookeeper的配置参数

storm.zookeeper.session.timeout: 60000

storm.zookeeper.retry.interval: 6000

storm.zookeeper.retry.times: 10

#蓝鲸本地用来存储索引和临时文件的目录列表

higo.workdir.list: "/disk1/taobao/bluewhile,/disk2/taobao/bluewhile,/disk3/taobao/bluewhile,/disk4/taobao/bluewhile,/disk5/taobao/bluewhile,/disk6/taobao/bluewhile,/disk7/taobao/bluewhile,/disk8/taobao/bluewhile,/disk9/taobao/bluewhile,/disk10/taobao/bluewhile"

#蓝鲸nimbus的地址

nimbus.host: "tiansuan1.kgb.cm4"

#蓝鲸zookeeper的root

storm.zookeeper.root: "/bluewhale\_higo"

#shell连接的海狗jdbc的地址

higo.jdbc.conn: "jdbc:higo://localhost:1107"

#无用保留不变即可

higo.download.offline.conn: "jdbc:mysql://tiansuan1.kgb.cm4:3306/adhoc\_download?useUnicode=true&characterEncoding=utf8"

#无用保留不变即可

higo.download.offline.username: "adhoc"

#无用保留不变即可

higo.download.offline.passwd: "adhoc"

#无用保留不变即可

higo.download.offline.store: "/group/taobao/external/p4p/p4padhoc/download/offline"

#配置那几个端口作为merger 使用

higo.merge.ports: "6601,6602,6603,6604,6605,6606,6607,6608,6609,6610"

#配置那几个端口作为实时数据源使用（目前该功能在开发中）

higo.realtime.ports: "6501,6502,6503,6504,6505,6506,6507,6508,6509,6510"

#使用的merger server的数量（请根据机器数量调整）

higo.mergeServer.count: 8

#使用的实时数据源的数量（目前该功能还在开发中）

higo.realtime.count: 0

#允许海狗cache住的打开的分区个数，越多消耗的内存越大，请不要随意更改

higo.cache.partions: 10

#默认每个进程的jvm配置

worker.childopts: "-Xms5g -Xmx5g -Xmn2g -XX:SurvivorRatio=3 -XX:PermSize=96m -XX:MaxPermSize=256m -XX:+UseParallelGC -XX:ParallelGCThreads=16 -XX:+UseAdaptiveSizePolicy -XX:+PrintGCDetails -XX:+PrintGCTimeStamps -Xloggc:%storm.home%/logs/gc-%port%.log "

#特定端口的jvm配置

worker.childopts.6601: "-Xms5g -Xmx5g -Xmn2g -XX:SurvivorRatio=3 -XX:PermSize=96m -XX:MaxPermSize=256m -XX:+UseParallelGC -XX:ParallelGCThreads=16 -XX:+UseAdaptiveSizePolicy -XX:+PrintGCDetails -XX:+PrintGCTimeStamps -Xloggc:%storm.home%/logs/gc-%port%.log "

#每台机器上启动的端口

supervisor.slots.ports:

- 6701

- 6702

- 6703

- 6704

- 6705

- 6706

- 6707

- 6708

- 6709

- 6710

- 6711

- 6712

- 6713

- 6714

- 6601 #注意看，这个端口是用来做merger server的

## 蓝鲸的启动

1. 将程序和配置分别copy到每台机器
2. 启动nimbus

nimbus是真个蓝鲸任务的总调度，有点类似hadoop的jobtracker

只需要在其中一台机器启动即可，启动命令如下

cd alipay/bluewhale-core/bin/

chmod 777 ./bluewhale

nohup ./bluewhale nimbus >nimbus.log &

启动完毕后可以看logs目录查看nimbus的运行日志，一般排错都去看nimbus.log来定位原因

1. 启动supervisor

supervisor用来管理其所在机器的work进程，其角色有点类似hadoop的tasktracker

需要在每台机器上启动，启动命令如下

cd alipay/bluewhale-core/bin/

chmod 777 ./bluewhale

nohup ./bluewhale supervisor >supervisor.log &

1. 启动ui和jdbc接口

ui里面暂时了海狗执行状态，内存使用，任务的分配等

jdbc接口是海狗对外的统一接口，应用层通过jdbc的方式连接海狗来进行查询

启动命令如下

cd alipay/bluewhale-core/bin/

chmod 777 ./bluewhale

mkdir ./ui

nohup ./bluewhale higoui 1107 ../lib/bluewhale-web-1.0.0-SNAPSHOT.war ./ui >ui.log &

启动后，可以通过浏览器打开对应ip的1107端口，看是否能正常打开即可。

## 数据的准备

* 由于历史原因，阿里一般都使用云梯，数据基本都存储在hive上，故海狗目前只能处理sequencefile格式的数据，对key没有要求，value则为对应数据表中一条记录，记录的列与列之间的分隔符为\001,有兴趣的可以通过简单的修改源码来改变
* 海狗要求数据按照天进行分目录

目录命名必须为dt=yyyyMMdd这种格式（你懂的），参考如下



* 数据的列中必须含有一列叫thedate,其值与其所在的目录dt=yyyyMMdd中的yyyyMMdd相等，thedate实际上为海狗的分区字段，任何的查询都必须指定thedate分区
* 列的数据类型目前只支持4种

string :字符串类型

tlong :long整形

tdouble :浮点型

tdate :日期类型

## 配置数据表

* 首先要确定数据表在hdfs上个根目录，比如说

/group/taobao/external/p4p/p4padhoc/tablelist

海狗jdbc会使用该目录，**请在jdbc的机器上配置环境变量**

export higo.table.path=/group/taobao/external/p4p/p4padhoc/tablelist

* 在根目录下可以有一个至多个数据表

参考如下



每个数据表目录下会有三个子目录

solr ：表的配置目录

index：海狗的索引在hdfs中的存储路径

tmp:临时目录

* 配置表

1. 将源码的bluewhale-higo/solr目录copy出来，并编辑conf/schema.xml进行表的列配置
2. 编辑注意事项

一个field一行，不能有换行，严格按照示例写，否则创建索引过程中有可能解析错误（创建索引使用的是正则匹配，非xml方式的解析）

数据存储顺序要与schema.xml中配置的顺序一致

保留字段不能删除

thedate:是分区的字段，格式为yyyyMMdd格式的字符串,将来与对应的分区目录要对应

higo\_uuid:是有系统生成的用来标记唯一性的字段，不能删除

dynamicField里的所有列都不能删除，将来扩充临时表的时候有可能会用到

为了节省存储空间，除了higo\_uuid外，其他字段都可以设置为不存储，也就是store=”false”;

1. 参考示例



* 上传至hdfs的表空间

假设hdfs的表目录空间跟目录为

/group/taobao/external/p4p/p4padhoc/tablelist/{tablename}/

我们需要将第一步修改好的配置文件上传到表空间中的solr目录里去

上传示例

hadoop fs -mkdir /group/taobao/external/p4p/p4padhoc/tablelist/{tablename}/solr

(注意这里多了个solr子目录)

hadoop fs -put ./bluewhale-higo/solr/\* /group/taobao/external/p4p/p4padhoc/tablelist/{tablename}/solr

## 创建索引

使用命令如下  
./bluewhale higo index {shards数量} {hdfs表空间目录} {hdfs源数据地址} 1 {清洗多少天的数据} {起始日期}   
./bluewhale higo index 60 /group/taobao/external/p4p/p4padhoc/tablelist/product/ /group/tbads/p4pdata/hive\_data/rpt/rpt\_p4padhoc\_product 1 100 20121001  
首次构建由于需要构建全部的索引，故会慢些，之后会采用按照每个月的上中下旬进行增量构建  
**关于参数的说明**

* shards的数量：在我们这个硬件条件下请设置为60
* hdfs表空间目录:注意这个目录下要有solr子目录，solr子目录下存储的这个表的字段配置等信息
* hdfs源数据地址

## 启动表

1. ./bluewhale higo table {表1,表2,表3} {shards数量} {hdfs的solr目录} {初始端口号} {topology名称}
2. 示例如下：启动7个表，每个表使用60个shard

./bluewhale higo table rpt\_hitfake\_auctionall\_d,rpt\_p4padhoc\_product,rpt\_p4padhoc\_cust,rpt\_b2bad\_hoc\_memb\_sum\_d,r\_rpt\_cps\_luna\_item,r\_rpt\_cps\_adhoc\_pid,r\_rpt\_cps\_adhoc\_seller 60 /group/taobao/external/p4p/p4padhoc/tablelist 51110 adhoc

1. 启动后大家可以在每台机器的logs目录去看系统日志，以及通过启动的监控ui查看表的一些监控信息

## 停止表

./bluewhale higo drop {表1,表2,表3} {topology名称}

./bluewhale higo drop rpt\_hitfake\_auctionall\_d,rpt\_p4padhoc\_product,rpt\_p4padhoc\_cust adhoc

## 如何重启服务

* 简单的通过停止表->启动表 重启服务
* 如果是因为蓝鲸本身的bug，需要重启蓝鲸，则可以这样做

1. 先停止表
2. 在每台机器上杀死所有的蓝鲸任务

ps -x|grep bluewhale|grep server|awk '{printf("%s\n",$1)}'|xargs kill

1. 清理蓝鲸的临时目录

rm xxx/bluewhale/stormwork/\* -rf

1. 重新按照原先的步骤 启动蓝鲸
2. 启动表