**日志抽取程序**

**安装手册**



修 改 记 录

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 版本号 | 日 期 | 地点 | 修订 | 描 述 |
| V1.0 | 2014/05/27 | 杭州 | 黄建丽 | 文档修订与维护 |
|  |  |  |  |  |

目录

[㈠.说明 3](#_Toc390361970)

[㈡.解压安装 3](#_Toc390361971)

[2.1解压tar.gz包 3](#_Toc390361972)

[2.2将lib包放在snappyFilter-multisp目录下 3](#_Toc390361973)

[㈢.配置文件 3](#_Toc390361974)

[3.1 customer-ref.xml 3](#_Toc390361975)

[3.2 domain.properties 4](#_Toc390361976)

[3.3 snappy.properties 4](#_Toc390361977)

[3.4 hdfs-site.xml 5](#_Toc390361978)

[3.5 log4j.properties 5](#_Toc390361979)

[（四）.启动 5](#_Toc390361980)

[4.1注意要修改启动脚本run.sh标粗的配置 5](#_Toc390361981)

[4.2启动 5](#_Toc390361982)

[（五）. 定制服务 5](#_Toc390361983)

# ㈠.说明

安装包：snappyFilter\_multisp\_${tag}的tar.gz

依赖包：lib.tgz

下载路径：**ftp://viewtoo:release@releaseftp.sihuatech.com/webcache/4.1.1.5/20140527/snappy**

# ㈡.解压安装

## 2.1解压tar.gz包

生成snappyFilter-multisp目录，目录结构如下

[root@HTCDN-VM-TEST-10 tmp]# tar -zxvf snappyFilter-multisp-R4.1.1.5\_Alpha\_B04\_20140611.tar.gz

snappyFilter-multisp/snappyFilter-multisp.jar

snappyFilter-multisp/config/customer-ref.xml

snappyFilter-multisp/config/domain.properties

snappyFilter-multisp/config/source-conf.xml

snappyFilter-multisp/config/hdfs-site.xml

snappyFilter-multisp/config/log4j.properties

snappyFilter-multisp/config/snappy.properties

snappyFilter-multisp/run.sh

## 2.2将lib包放在snappyFilter-multisp目录下

# ㈢.配置文件

配置文档包括customer-ref.xml、domain.properties、snappy.properties、hdfs-site.xml、log4j.properties五个文件。

## 3.1 customer-ref.xml

该版本只支持两种策略default和天地祥云，后续版本更新为多种策略。

[root@HTCDN-VOPE-LOGCALC-02 config]# cat customer-ref.xml

<?xml version="1.0" encoding="utf-8"?>

<customers>

<!-- SP 淘宝图片-->

<customer>

<name>taobao</name> <outputdir>/data/logdownload/taobao\_video</outputdir> <filePrefix>taobao\_htcdn</filePrefix>

<strategy>default</strategy>

</customer>

<!-- SP 天地祥云-->

<customer>

<name>cloudvsp</name>

<outputdir>/data/logdownload/cloudvsp\_video</outputdir>

<filePrefix>cloudvsp\_htcdn</filePrefix>

<strategy>cloudvsp</strategy>

</customer>

<customer>

<!-- SP 其他-->

<name>other</name>

<outputdir>/data/logdownload/other\_video</outputdir>

<filePrefix>other\_htcdn</filePrefix>

<strategy>default</strategy>

</customer>

</customers>

## 3.2 domain.properties

[root@HTCDN-VOPE-LOGCALC-02 config]# cat domain.properties

tb-img=taobao

tdxy-tdxy=cloudvsp

## 3.3 snappy.properties

[root@HTCDN-VOPE-LOGCALC-02 config]# cat snappy.properties

logHome=hdfs\://hdfs1/flume/test/huatong/picture/

partitions=001,006

tempdir=/opt/snappyvid\_wasuott/snappyFilter-multisp/tmp

#######默认策略######

default.fields=3,1,2,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28

default.quote.fields=5,12,15,23,18

default.resp.fields=8,9

default.prefix=default

default.suffix=log

default.tempdir=/opt/snappyvid\_wasuott/snappyFilter-multisp/tmp

#######天地祥云#########

cloudvsp.fields=1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21

cloudvsp.quote.fields=5,12,15,23,18

cloudvsp.resp.fields=8,9

cloudvsp.prefix=cloudvsp

cloudvsp.suffix=log

cloudvsp.tempdir=/opt/snappyvid\_wasuott/snappyFilter-multisp/tmp\_cloud

################

workers=10

############

## 3.4 hdfs-site.xml

为hadoop集群的配置，可以去集群上copy过来，无需修改（略）

## 3.5 log4j.properties

略

# （四）.启动

## 4.1注意要修改启动脚本run.sh标粗的配置

[root@HTCDN-VOPE-LOGCALC-02 snappyFilter-multisp]# cat run.sh

#!/bin/bash

export LD\_LIBRARY\_PATH=**/opt/snappyvid\_wasuott/snappyFilter-multisp/lib/native**

nohup java -cp config:snappyFilter-multisp.jar:**/opt/snappyvid\_wasuott/snappyFilter-multisp/lib/\*** -server -Xmx4G -Xms4G com.sihuatech.snappy.SnappyHandlerMain >> snappyFilter.log 2>&1 &

## 4.2启动

[root@HTCDN-VOPE-LOGCALC-02 snappyFilter-multisp]# ./run.sh

# （五）. 定制服务

使用脚本推到对方FTP

vi /etc/crontab

\*/1 \* \* \* \* root /data/soft/ftp\_cp.sh

0 1 \* \* \* root /data/soft/ftp\_yes\_cp.sh

每一分钟运行ftp\_cp.sh

**[root@HTCDN-VOPE-LOGCALC-02 soft]# cat ftp\_cp.sh**

#/bin/sh

TODAY=`date +%Y%m%d`

YESTERDAY="`date -d"1 day ago" +%Y%m%d`"

LOG\_DIR=/data/logdownload/cloudvsp\_video

FTP\_USER=hjl

FTP\_PASSWORD=hjl

FTP\_IP=10.254.240.102

SPNAME=cloudvsp\_htcdn

FILE\_DATE=`date +%Y%m%d%H%M`

cd $LOG\_DIR/$TODAY

ftp -i -n $FTP\_IP <<EOF

user $FTP\_USER $FTP\_PASSWORD

bin

mput $SPNAME"\_\*.log.gz

quit

EOF

每天凌晨1点运行ftp\_yes\_cp.sh

**[root@HTCDN-VOPE-LOGCALC-02 soft]# cat ftp\_yes\_cp.sh**

#/bin/sh

TODAY=`date +%Y%m%d`

YESTERDAY="`date -d"1 day ago" +%Y%m%d`"

LOG\_DIR=/data/logdownload/cloudvsp\_video

FTP\_USER=hjl

FTP\_PASSWORD=hjl

FTP\_IP=10.254.240.102

SPNAME=cloudvsp\_htcdn

cd $LOG\_DIR/$YESTERDAY

#cd $LOG\_DIR/$TODAY

ftp -i -n $FTP\_IP <<EOF

user $FTP\_USER $FTP\_PASSWORD

bin

mput $SPNAME"\_\*.log.gz

quit

EOF