**大数据平台部署与运维**

|  |  |  |  |
| --- | --- | --- | --- |
| 文档更新记录 | | | |
| 日期 | 作者 | 版本 | 备注说明 |
| 2016-05-20 | 吴钢奇 | 1.0 | Rsyslog+kafka+fleuntd+elastic+hadoop部署 |
|  |  |  |  |
|  |  |  |  |

目录

[一、大数据平台架构图 3](#_Toc451527197)

[二、rsyslog 3](#_Toc451527198)

[1、安装 3](#_Toc451527199)

[2、配置 4](#_Toc451527200)

[3、维护 5](#_Toc451527201)

[三、kafka 5](#_Toc451527202)

[1、安装 5](#_Toc451527203)

[2、配置 6](#_Toc451527204)

[3、维护 7](#_Toc451527205)

[四、fleuntd 7](#_Toc451527206)

[1、安装 7](#_Toc451527207)

[2、配置 8](#_Toc451527208)

[3、维护 13](#_Toc451527209)

[五、elastic 14](#_Toc451527210)

[1、安装 14](#_Toc451527211)

[2、配置 15](#_Toc451527212)

[3、维护 18](#_Toc451527213)

[六、hadoop 19](#_Toc451527214)

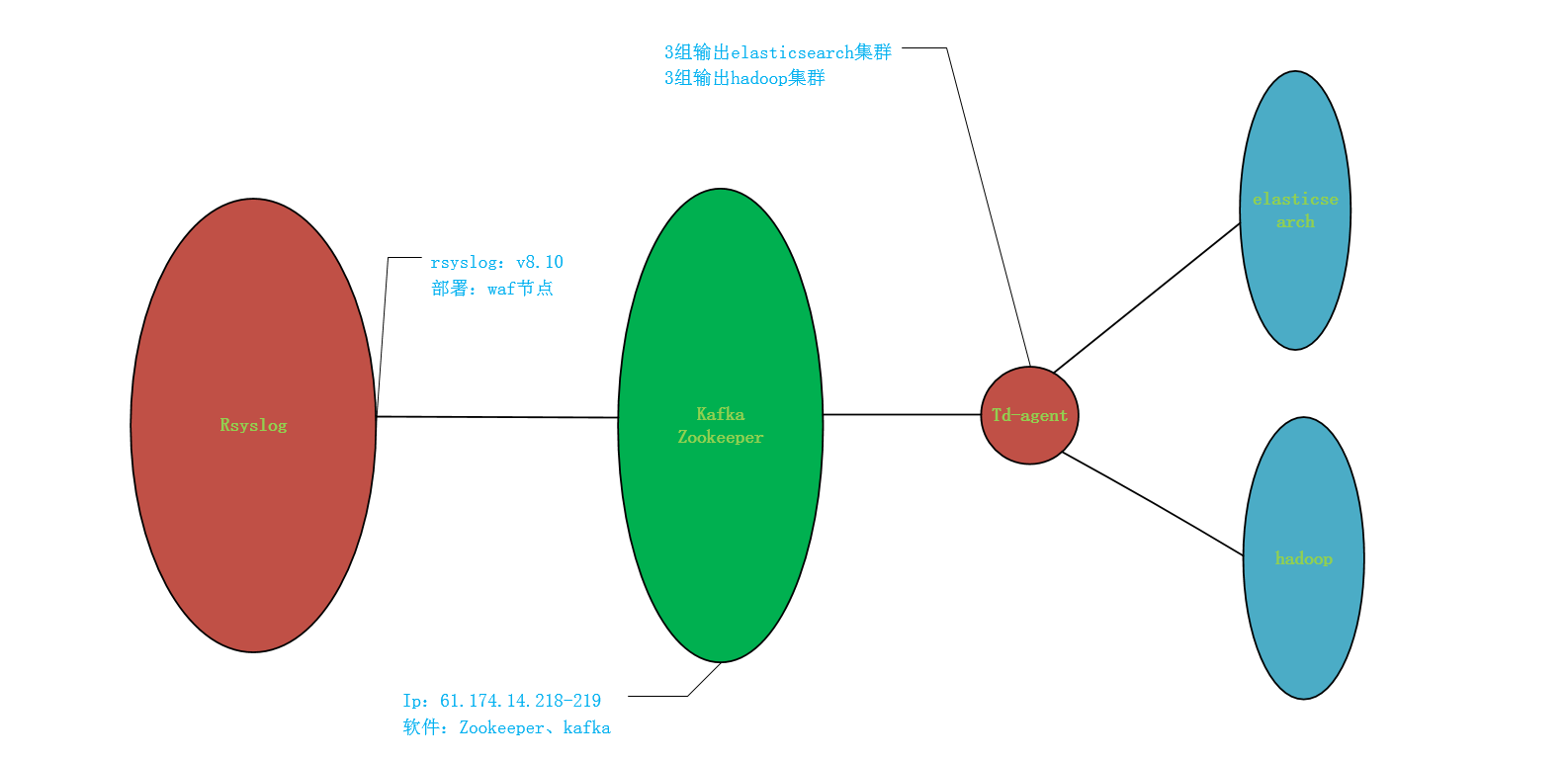
[1、安装 19](#_Toc451527215)

[2、配置 21](#_Toc451527216)

[3、维护 22](#_Toc451527217)

[4、数据输出 22](#_Toc451527218)

# 一、大数据平台架构图



# 二、rsyslog

## 1、安装

[root@CNC-TJ-005-045 ~]# cd /etc/yum.repos.d/

[root@CNC-TJ-005-045 yum.repos.d]# ls

CentOS-Base.repo epel.repo epel-testing.repo rsyslogall.repo

[root@CNC-TJ-005-045 yum.repos.d]# cat rsyslogall.repo

[rsyslog-v8-stable]

name=Adiscon Rsyslog v8-stable for CentOS-$releasever-$basearch

baseurl=http://rpms.adiscon.com/v8-stable/epel-$releasever/$basearch

enabled=1

gpgcheck=0

protect=1

[root@CNC-TJ-005-045 yum.repos.d]#

[root@CNC-TJ-005-045 yum.repos.d]# yum -y install rsyslog rsyslog-kafka

[root@CNC-TJ-005-045 yum.repos.d]# rpm -qa |grep rsyslog

rsyslog-8.18.0-1.el6.x86\_64

rsyslog-kafka-8.18.0-1.el6.x86\_64

## 2、配置

[root@CNC-TJ-005-045 yum.repos.d]# cat /etc/rsyslog.conf |grep -vE "^#|^$"

$MaxMessageSize 20k

$ModLoad imuxsock # provides support for local system logging (e.g. via logger command)

$ModLoad imklog # provides kernel logging support (previously done by rklogd)

module(load="omkafka")

$ActionFileDefaultTemplate RSYSLOG\_TraditionalFileFormat

$IncludeConfig /etc/rsyslog.d/\*.conf

kern.\* @61.174.9.68:61200

authpriv.\* /var/log/secure

mail.\* -/var/log/maillog

cron.\* /var/log/cron

uucp,news.crit /var/log/spooler

local7.\* /var/log/boot.log

$template strtpl,"%TIMESTAMP:::date-pgsql% %msg%\n"

local6.info /var/log/history;strtpl

local6.info @ops.cdn.syslog.secure.anquan.io:5140

authpriv.\* @ops.cdn.syslog.secure.anquan.io:5656

$WorkDirectory /var/spool/rsyslog

module(load="imfile" mode="inotify")

input(type="imfile"

file="/cache/logs/201\*\_user\_access.log"

Tag="nginx-access:"

Facility="local5"

Severity="err"

PersistStateInterval="200")

$template FormatNginx,"%msg%\n"

local5.err action(type="omkafka" topic="nginx-access" broker=["120.92.245.58:9092","120.92.245.59:9092","120.92.245.60:9092","120.92.245.59:9092","120.92.245.60:9092","120.92.245.58:9092","120.92.245.60:9092","120.92.245.59:9092","120.92.245.58:9092"] partitions.number="9" errorFile="/cache/logs/rsyslog/kafkaerr.log" template="FormatNginx" confParam=["compression.codec=snappy","socket.timeout.ms=5000","socket.keepalive.enable=true"] queue.spoolDirectory="/cache/logs/rsyslog" queue.type="LinkedList" queue.size="16777216" queue.filename="toKafka" queue.highwatermark="10066330" queue.lowwatermark="5033165" queue.maxdiskspace="20g" queue.maxfilesize="1g" queue.dequeuebatchsize="500" queue.saveonshutdown="on" queue.workerThreads="4" queue.discardseverity="8" queue.discardmark="16000000" queue.timeoutenqueue="5000")

if $programname == 'nginx-access' then stop

\*.info;mail.none;authpriv.none;cron.none;local6.none /var/log/messages

## 3、维护

[root@CNC-TJ-005-045 ~]# service rsyslog start

[root@CNC-TJ-005-045 ~]# service rsyslog stop

# 三、kafka

## 1、安装

[root@ctl-zj-61-174-14-218 ~]# wget -c http://download.oracle.com/otn-pub/java/jdk/8u65-b17/jdk-8u65-linux-x64.tar.gz?AuthParam=1446451514\_9a4acfee12c17886366e74da5d6cdef9

[root@ctl-zj-61-174-14-218 ~]# cat /etc/profile

export JAVA\_HOME=/usr/local/jdk1.8.0\_65

export CLASSPATH=.:$JAVA\_HOME/lib/dt.jar:$JAVA\_HOME/lib/tools.jar

export LIBJLI\_PATH=/usr/local/jdk1.8.0\_65/jre/lib/amd64/jli

export PATH=/usr/local/bin:$PATH:$JAVA\_HOME/bin:$HOME/bin:$LIBJLI\_PATH:/usr/local/kafka/bin:/usr/local/zookeeper/bin

[root@ctl-zj-61-174-14-218 ~]# ls /usr/local/zookeeper/

bin CHANGES.txt contrib docs ivy.xml LICENSE.txt README\_packaging.txt recipes zookeeper-3.4.6.jar zookeeper-3.4.6.jar.md5

build.xml conf dist-maven ivysettings.xml lib NOTICE.txt README.txt src zookeeper-3.4.6.jar.asc zookeeper-3.4.6.jar.sha1

[root@ctl-zj-61-174-14-218 ~]#

[root@ctl-zj-61-174-14-218 ~]# ls /usr/local/kafka/

bin config libs LICENSE logs NOTICE version

## 2、配置

[root@ctl-zj-61-174-14-218 ~]# cat /usr/local/zookeeper/conf/zoo.cfg

tickTime=2000

initLimit=10

syncLimit=5

clientPort=2181

server.0=61.174.14.218:2888:3888

server.1=61.174.14.219:2888:3888

server.2=61.174.14.220:2888:3888

dataLogDir=/data01/zookeeper/logs

dataDir=/data01/zookeeper/data

autopurge.purgeInterval=6

autopurge.snapRetainCount=72

[root@ctl-zj-61-174-14-218 ~]#

[root@ctl-zj-61-174-14-218 ~]# cat /usr/local/kafka/config/server.properties

broker.id=0

port=9092

advertised.host.name=61.174.14.218

num.network.threads=16

num.io.threads=32

socket.send.buffer.bytes=102400

socket.receive.buffer.bytes=102400

socket.request.max.bytes=1083741824

log.dirs=/data02/kafka,/data03/kafka,/data04/kafka,/data05/kafka,/data06/kafka,/data07/kafka,/data08/kafka,/data09/kafka,/data10/kafka,/data11/kafka,/data12/kafka,/data13/kafka

num.partitions=8

num.recovery.threads.per.data.dir=1

log.segment.bytes=1073741824

log.retention.check.interval.ms=300000

log.cleaner.enable=false

zookeeper.connect=61.174.14.218:2181,61.174.14.219:2181,61.174.14.220:2181

zookeeper.connection.timeout.ms=6000

delete.topic.enable=true

log.flush.interval.messages=100000

log.flush.interval.ms=1000

log.retention.hours=720

queued.max.requests=2000

[root@ctl-zj-61-174-14-218 ~]#

## 3、维护

/usr/local/zookeeper/bin/zkServer.sh start

/usr/local/zookeeper/bin/zkServer.sh stop

/usr/local/kafka/bin/kafka-server-start.sh /usr/local/kafka/config/server.properties &

/usr/local/kafka/bin/kafka-server-stop.sh

# 四、fleuntd

## 1、安装

[root@ctl-zj-61-174-14-218 ~]# cd /etc/yum.repos.d

[root@ctl-zj-61-174-14-218 yum.repos.d]# ll

total 32

-rw-r--r-- 1 root root 2572 May 15 2015 CentOS-Base.repo

-rw-r--r-- 1 root root 954 Jul 6 2015 epel.repo

-rw-r--r-- 1 root root 1056 Nov 5 2012 epel-testing.repo

-rw-r--r-- 1 root root 174 Nov 4 2015 td.repo

-rw-r--r-- 1 root root 401 Nov 12 2013 zabbix.repo

[root@ctl-zj-61-174-14-218 yum.repos.d]# cat td.repo

[treasuredata]

name=TreasureData

baseurl=http://packages.treasuredata.com/2/redhat/$releasever/$basearch

gpgcheck=1

gpgkey=https://packages.treasuredata.com/GPG-KEY-td-agent

[root@ctl-zj-61-174-14-218 yum.repos.d]#

[root@ctl-zj-61-174-14-218 yum.repos.d]# yum -y install td-agent

## 2、配置

[root@ctl-zj-61-174-14-218 ~]# cat /etc/td-agent/td-agent.conf

<source>

type kafka

topics nginx-access

hosts 61.174.14.218,61.174.14.219,61.174.14.220

port 9092

#format text

format json

client\_id elasticsearch-1

partition 1

max\_bytes 52428800

max\_wait\_ms 5000

min\_bytes 1024

socket\_timeout\_ms 50000

add\_prefix abc

</source>

<match abc.nginx-access>

type record\_modifier

char\_encoding gbk:utf-8

tag a.apache.access

</match>

<match a.apache.access>

type api

add\_prefix api

</match>

<match api.a.apache.access>

type copy

<store>

type forward

heartbeat\_type tcp

send\_timeout 60s

recover\_wait 10s

heartbeat\_interval 20s

phi\_threshold 16

hard\_timeout 60s

buffer\_type file

buffer\_path /data01/fluentd/buffer/myapp.\*.buffer

buffer\_queue\_limit 500

buffer\_chunk\_limit 2m

num\_threads 50

<server>

host 61.174.14.221

port 10086

</server>

<secondary>

type file

path /data01/fluentd/buffer/forward-failed

</secondary>

</store>

<store>

type rewrite\_tag\_filter

rewriterule1 blockid ^[-1-7]\d\*$ attack.log

rewriterule2 domain 81\.cn$ jfzb.log1

rewriterule3 domain chinamil\.com\.cn$ jfzb.log2

rewriterule4 domain 81rc\.mil\.cn$ jfzb.log3

rewriterule5 domain mod\.gov\.cn$ jfzb.log4

</store>

<store>

type elasticsearch

hosts 61.174.14.194:9200,61.174.14.195:9200,61.174.14.196:9200

logstash\_format true

flush\_interval 20s

logstash\_prefix access

index\_name access

utc\_index false

reload\_connections false

buffer\_type file

buffer\_path /data01/fluentd/buffer/nginx.access.buffer

buffer\_chunk\_limit 14M

buffer\_queue\_limit 2048

request\_timeout 30

</store>

</match>

<match attack.log>

type elasticsearch

hosts 61.174.14.194:9200,61.174.14.195:9200,61.174.14.196:9200

logstash\_format true

flush\_interval 20s

logstash\_prefix waf

index\_name waf

utc\_index false

reload\_connections false

buffer\_type file

buffer\_path /data01/fluentd/buffer/attack.buffer

buffer\_chunk\_limit 12M

buffer\_queue\_limit 1024

request\_timeout 30

</match>

<match jfzb.log\*>

type elasticsearch

hosts 61.174.14.194:9200,61.174.14.195:9200,61.174.14.196:9200

logstash\_format true

flush\_interval 20s

logstash\_prefix mod

index\_name mod

utc\_index false

reload\_connections false

buffer\_type file

buffer\_path /data01/fluentd/buffer/jfzb.log.buffer

buffer\_chunk\_limit 12M

buffer\_queue\_limit 2048

request\_timeout 30

</match>

[root@ctl-zj-61-174-14-218 ~]# cat /etc/td-agent/td-agenthadoop1.conf

<source>

type kafka

topics nginx-access

hosts 61.174.14.218,61.174.14.219,61.174.14.220

port 9092

#format text

format json

client\_id td-agent-input-hadoop1

partition 1

max\_bytes 52428800

max\_wait\_ms 5000

min\_bytes 1024

socket\_timeout\_ms 50000

add\_prefix abc

</source>

<match abc.nginx-access>

type record\_modifier

char\_encoding gbk:utf-8

tag a.apache.access

</match>

<match a.apache.access>

type api

add\_prefix api

</match>

<match api.a.apache.access>

type copy

<store>

type webhdfs

namenode 172.26.100.12:50070

standby\_namenode 172.26.100.11:50070

path /weblog/temp\_table/day=%Y%m%d/ht=%H/${hostname}\_%H.log

username hdfs

localtime true

retry\_known\_errors yes

field\_separator SOH

output\_include\_time true

output\_include\_tag false

time\_format %Y-%m-%d %H:%M:%S

output\_data\_type attr:clientip,method,domain,request,status,size,referer,agent,hit\_status,ruleid,prov,country,blockid,searchbot,p\_status,hackerid

flush\_interval 5s

buffer\_type file

buffer\_path /data01/fluentd/buffer/webhdfs2.buffer

buffer\_chunk\_limit 12M

buffer\_queue\_limit 2048

#retry\_times 10

#retry\_interval 3

</store>

<store>

type webhdfs

namenode 172.26.100.12:50070

standby\_namenode 172.26.100.11:50070

path /weblog/web\_source/day=%Y%m%d/${hostname}\_%H.log

username hdfs

localtime true

retry\_known\_errors yes

field\_separator SOH

output\_include\_time false

output\_include\_tag false

time\_format %Y-%m-%d %H:%M:%S

output\_data\_type attr:domain,referer,blockid

flush\_interval 5s

buffer\_type file

buffer\_path /data01/fluentd/buffer/webhdfs3.buffer

buffer\_chunk\_limit 12M

buffer\_queue\_limit 2048

#retry\_times 10

#retry\_interval 3

</store>

<store>

type webhdfs

namenode 172.26.100.12:50070

standby\_namenode 172.26.100.11:50070

path /weblog/top10\_url/day=%Y%m%d/${hostname}\_%H.log

username hdfs

localtime true

retry\_known\_errors yes

field\_separator SOH

output\_include\_time false

output\_include\_tag false

time\_format %Y-%m-%d %H:%M:%S

output\_data\_type attr:domain,request

flush\_interval 5s

buffer\_type file

buffer\_path /data01/fluentd/buffer/webhdfs4.buffer

buffer\_chunk\_limit 12M

buffer\_queue\_limit 2048

#retry\_times 10

#retry\_interval 3

</store>

<store>

type webhdfs

namenode 172.26.100.12:50070

standby\_namenode 172.26.100.11:50070

path /weblog/top5\_hacker/day=%Y%m%d/${hostname}\_%H.log

username hdfs

localtime true

retry\_known\_errors yes

field\_separator SOH

output\_include\_time false

output\_include\_tag false

time\_format %Y-%m-%d %H:%M:%S

output\_data\_type attr:domain,clientip,prov,hackerid

flush\_interval 5s

buffer\_type file

buffer\_path /data01/fluentd/buffer/webhdfs5.buffer

buffer\_chunk\_limit 12M

buffer\_queue\_limit 2048

#retry\_times 10

#retry\_interval 3

</store>

</match>

[root@ctl-zj-61-174-14-218 ~]#

## 3、维护

/etc/init.d/td-agent restart

kill `ps -ef |grep td-agent2 | grep -v grep |awk '{print $2}'`

kill `ps -ef |grep td-agent2 | grep -v grep |awk '{print $2}'`

kill `ps -ef |grep td-agent2 | grep -v grep |awk '{print $2}'`

sleep 30

/opt/td-agent/embedded/bin/ruby /usr/sbin/td-agent2 --group td-agent --log /var/log/td-agent/td-agent2.log --use-v1-config --daemon /var/run/td-agent/td-agent2.pid

kill `ps -ef |grep td-agent3 | grep -v grep |awk '{print $2}'`

kill `ps -ef |grep td-agent3 | grep -v grep |awk '{print $2}'`

kill `ps -ef |grep td-agent3 | grep -v grep |awk '{print $2}'`

sleep 30

/opt/td-agent/embedded/bin/ruby /usr/sbin/td-agent3 --group td-agent --log /var/log/td-agent/td-agent3.log --use-v1-config --daemon /var/run/td-agent/td-agent3.pid

kill `ps -ef |grep td-agenthadoop1 | grep -v grep |awk '{print $2}'`

kill `ps -ef |grep td-agenthadoop1 | grep -v grep |awk '{print $2}'`

kill `ps -ef |grep td-agenthadoop1 | grep -v grep |awk '{print $2}'`

sleep 30

/opt/td-agent/embedded/bin/ruby /usr/sbin/td-agenthadoop1 --group td-agent --log /var/log/td-agent/td-agenthadoop1.log --use-v1-config --daemon /var/run/td-agent/td-agenthadoop1.pid

kill `ps -ef |grep td-agenthadoop2 | grep -v grep |awk '{print $2}'`

kill `ps -ef |grep td-agenthadoop2 | grep -v grep |awk '{print $2}'`

kill `ps -ef |grep td-agenthadoop2 | grep -v grep |awk '{print $2}'`

sleep 30

/opt/td-agent/embedded/bin/ruby /usr/sbin/td-agenthadoop2 --group td-agent --log /var/log/td-agent/td-agenthadoop2.log --use-v1-config --daemon /var/run/td-agent/td-agenthadoop2.pid

kill `ps -ef |grep td-agenthadoop3 | grep -v grep |awk '{print $2}'`

kill `ps -ef |grep td-agenthadoop3 | grep -v grep |awk '{print $2}'`

kill `ps -ef |grep td-agenthadoop3 | grep -v grep |awk '{print $2}'`

sleep 30

/opt/td-agent/embedded/bin/ruby /usr/sbin/td-agenthadoop3 --group td-agent --log /var/log/td-agent/td-agenthadoop3.log --use-v1-config --daemon /var/run/td-agent/td-agenthadoop3.pid

# 五、elastic

## 1、安装

[root@ctl-zj-014-194 ~]# cd /etc/yum.repos.d/

[root@ctl-zj-014-194 yum.repos.d]# ls

abc CentOS-Base.repo elasticsearch.repo epel.repo epel-testing.repo kibana.repo nd.repo nodesource-el.repo

[root@ctl-zj-014-194 yum.repos.d]# cat elasticsearch.repo

[elasticsearch-2.x]

name=Elasticsearch repository for 2.x packages

baseurl=https://packages.elastic.co/elasticsearch/2.x/centos

gpgcheck=1

gpgkey=https://packages.elastic.co/GPG-KEY-elasticsearch

enabled=1

[root@ctl-zj-014-194 yum.repos.d]# cat kibana.repo

[kibana-4.5]

name=Kibana repository for 4.5.x packages

baseurl=http://packages.elastic.co/kibana/4.5/centos

gpgcheck=1

gpgkey=http://packages.elastic.co/GPG-KEY-elasticsearch

enabled=1

[root@ctl-zj-014-194 yum.repos.d]#

[root@ctl-zj-014-194 yum.repos.d]# yum install -y kibana

[root@ctl-zj-014-194 yum.repos.d]# yum install elasticsearch

[root@ctl-zj-014-194 yum.repos.d]# rpm -qa |grep elasticsearch

elasticsearch-2.3.2-1.noarch

[root@ctl-zj-014-194 yum.repos.d]# rpm -qa |grep kibana

kibana-4.5.0-1.x86\_64

[root@ctl-zj-014-194 yum.repos.d]#

## 2、配置

[root@ctl-zj-014-194 ~]# cat /etc/elasticsearch/elasticsearch.yml

# ---------------------------------- Cluster -----------------------------------

cluster.name: "2dafb4bf52d4e8e0"

cluster.routing.allocation.node\_initial\_primaries\_recoveries: 8

cluster.routing.allocation.node\_concurrent\_recoveries: 8

# ------------------------------------ Node ------------------------------------

#node.name: ${HOSTNAME}

node.name: "niudun\_node194"

node.master: true

node.data: true

# ----------------------------------- Paths ------------------------------------

path.data: /log/02,/log/03,/log/04,/log/05,/log/06,/log/07,/log/08,/log/09,/log/10,/log/11,/log/12

path.repo: ["/backup"]

# ----------------------------------- Memory -----------------------------------

bootstrap.mlockall: true

# ---------------------------------- Network -----------------------------------

network.bind\_host: 172.26.99.1

network.publish\_host: 172.26.99.1

http.host: 61.174.14.194

network.tcp.keep\_alive: true

# ---------------------------------- Gateway -----------------------------------

#gateway.type: local

gateway.recover\_after\_nodes: 6

gateway.recover\_data\_nodes: 9

gateway.recover\_after\_time: 5m

# ---------------------------------- Transport module -----------------------------------

transport.tcp.compress: true

# ---------------------------------- Http module -----------------------------------

http.enabled: true

http.tcp\_keep\_alive: true

http.max\_content\_length: 500mb

http.cors.allow-origin: "/.\*/"

http.cors.enabled: true

# --------------------------------- Discovery ----------------------------------

discovery.zen.minimum\_master\_nodes: 1

discovery.zen.fd.ping\_timeout: 120s

discovery.zen.fd.ping\_retries: 6

discovery.zen.fd.ping\_interval: 30s

discovery.zen.ping.timeout: 3s

discovery.zen.ping.unicast.hosts: ["172.26.99.1", "172.26.99.2", "172.26.99.3", "172.26.99.4", "172.26.99.5", "172.26.99.6", "172.26.99.7", "172.26.99.8", "172.26.99.9", "172.26.99.10"]

# ---------------------------------- Various -----------------------------------

node.max\_local\_storage\_nodes: 1

action.destructive\_requires\_name: true

# ---------------------------------- Indices module -----------------------------------

# indices recovery

indices.recovery.max\_bytes\_per\_sec: 100mb

indices.recovery.max\_size\_per\_sec: 400mb

indices.recovery.concurrent\_streams: 8

indices.recovery.compress: true

indices.recovery.translog\_ops: 2000

# circuit breaker

indices.breaker.total.limit: 70%

indices.breaker.fielddata.limit: 60%

indices.breaker.request.limit: 40%

# fielddata

indices.fielddata.cache.size: 30%

# node query cache

indices.queries.cache.size: 10%

# indexing buffer

indices.memory.index\_buffer\_size: 15%

#indices.memory.max\_index\_buffer\_size: 28gb

# shard request cache

indices.requests.cache.size: 2%

# basic

index.number\_of\_shards: 5

index.number\_of\_replicas: 1

index.refresh\_interval: 5s

# translog settings

#index.translog.flush\_threshold\_ops: 50000

index.translog.flush\_threshold\_size: 4g

index.translog.flush\_threshold\_period: 30m

index.translog.interval: 5s

index.translog.sync\_interval: 5s

# thread pool

threadpool.search.type: fixed

threadpool.search.size: 20

threadpool.search.queue\_size: 400

threadpool.bulk.type: fixed

threadpool.bulk.size: 60

threadpool.bulk.queue\_size: 300

threadpool.index.type: fixed

threadpool.index.size: 20

threadpool.index.queue\_size: 100

# ---------------------------------- Slowlog module -----------------------------------

#index.search.slowlog.level: TRACE

#index.search.slowlog.threshold.query.warn: 10s

#index.search.slowlog.threshold.query.info: 5s

#index.search.slowlog.threshold.query.debug: 2s

#index.search.slowlog.threshold.query.trace: 500ms

#

#index.search.slowlog.threshold.fetch.warn: 1s

#index.search.slowlog.threshold.fetch.info: 800ms

#index.search.slowlog.threshold.fetch.debug: 500ms

#index.search.slowlog.threshold.fetch.trace: 200ms

#

#index.indexing.slowlog.threshold.index.warn: 10s

#index.indexing.slowlog.threshold.index.info: 5s

#index.indexing.slowlog.threshold.index.debug: 2s

#index.indexing.slowlog.threshold.index.trace: 500ms

## mavel

marvel.agent.exporters:

id1:

type: http

host: ["61.174.14.194:9200","61.174.14.195:9200","61.174.14.196:9200"]

[root@ctl-zj-014-194 ~]#

[root@ctl-zj-014-194 ~]# grep -vE "^#|^$" /opt/kibana/config/kibana.yml

server.port: 56123

server.host: "127.0.0.1"

server.basePath: ""

server.maxPayloadBytes: 1048576

elasticsearch.url: "http://61.174.14.194:9200"

elasticsearch.preserveHost: true

kibana.defaultAppId: "discover"

elasticsearch.pingTimeout: 1500

elasticsearch.requestTimeout: 300000

elasticsearch.shardTimeout: 0

elasticsearch.startupTimeout: 5000

pid.file: /var/run/kibana.pid

logging.dest: stdout

logging.silent: false

logging.quiet: false

logging.verbose: false

[root@ctl-zj-014-194 ~]#

## 3、维护

service elasticsearch start

/etc/init.d/kibana start

/usr/local/openresty/nginx/sbin/nginx

service elasticsearch stop

/etc/init.d/kibana stop

/usr/local/openresty/nginx/sbin/nginx –s stop

# 六、hadoop

## 1、安装

[root@master211 ~]# cat /etc/hosts

127.0.0.1 localhost localhost.localdomain localhost4 localhost4.localdomain4

::1 localhost localhost.localdomain localhost6 localhost6.localdomain6

172.26.100.1 slave201

172.26.100.2 slave202

172.26.100.3 slave203

172.26.100.4 slave204

172.26.100.5 slave205

172.26.100.6 slave206

172.26.100.7 slave207

172.26.100.8 slave208

172.26.100.9 slave209

172.26.100.10 slave210

172.26.100.13 slave211

172.26.100.14 slave212

172.26.100.15 slave213

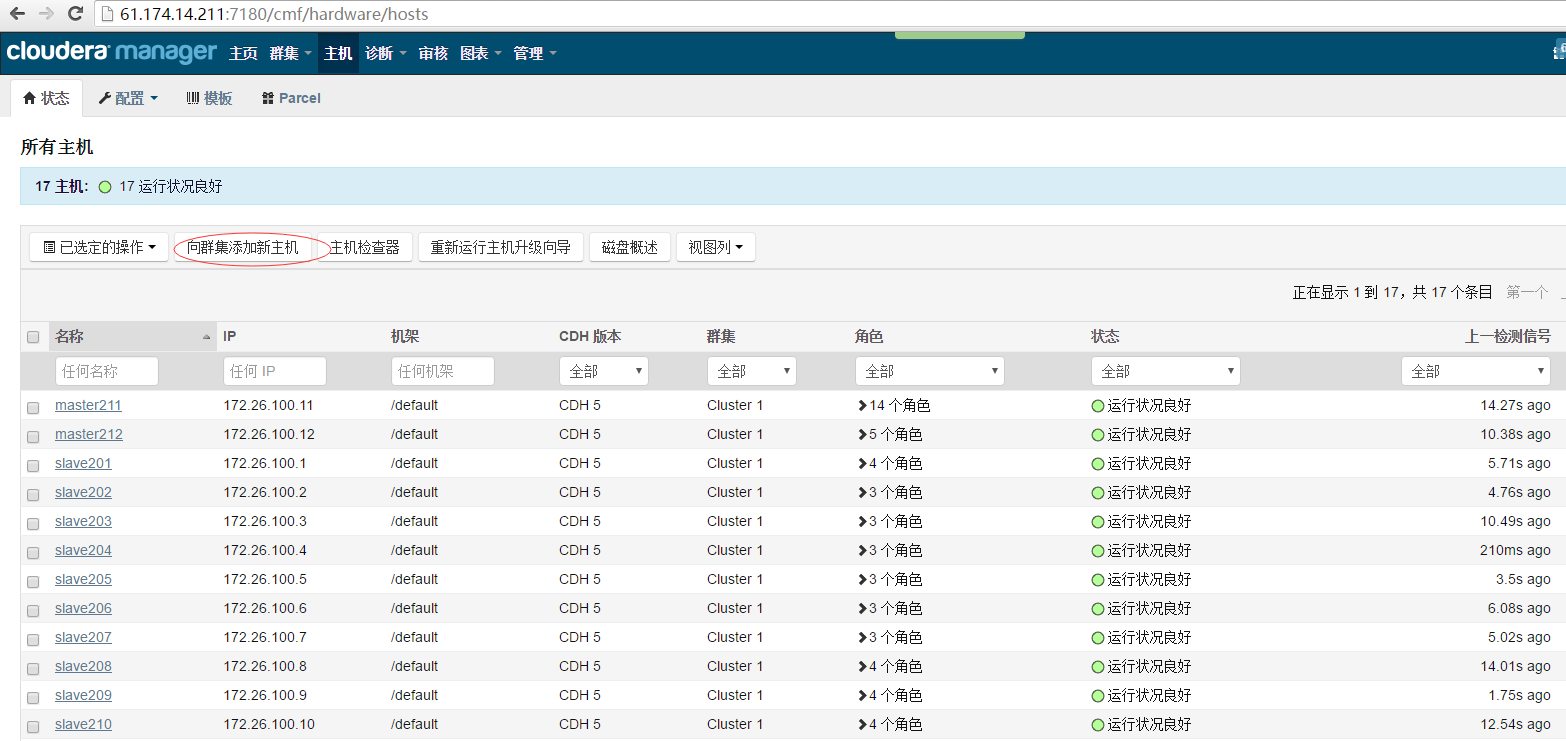
172.26.100.16 slave214

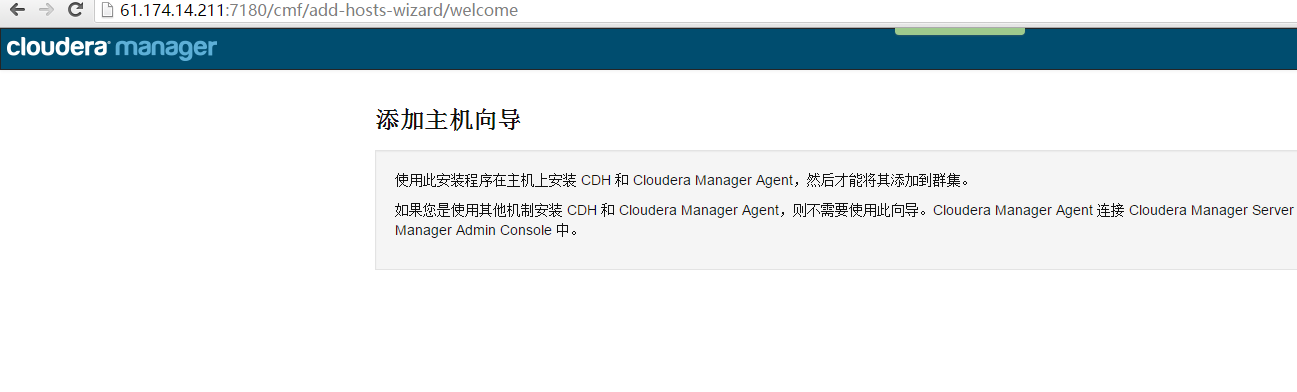
172.26.100.17 slave215

172.26.100.11 master211

172.26.100.12 master212

[root@master211 ~]#

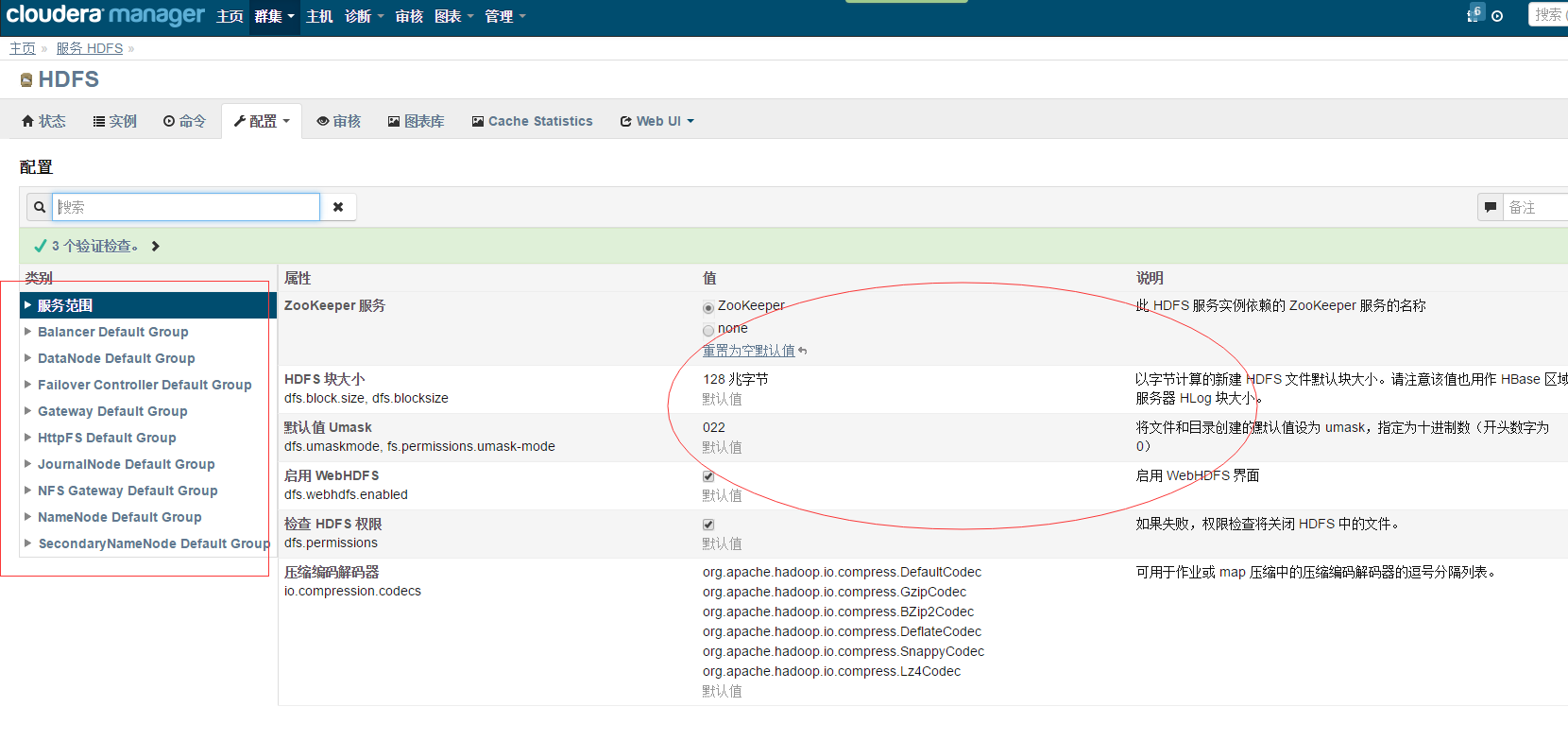




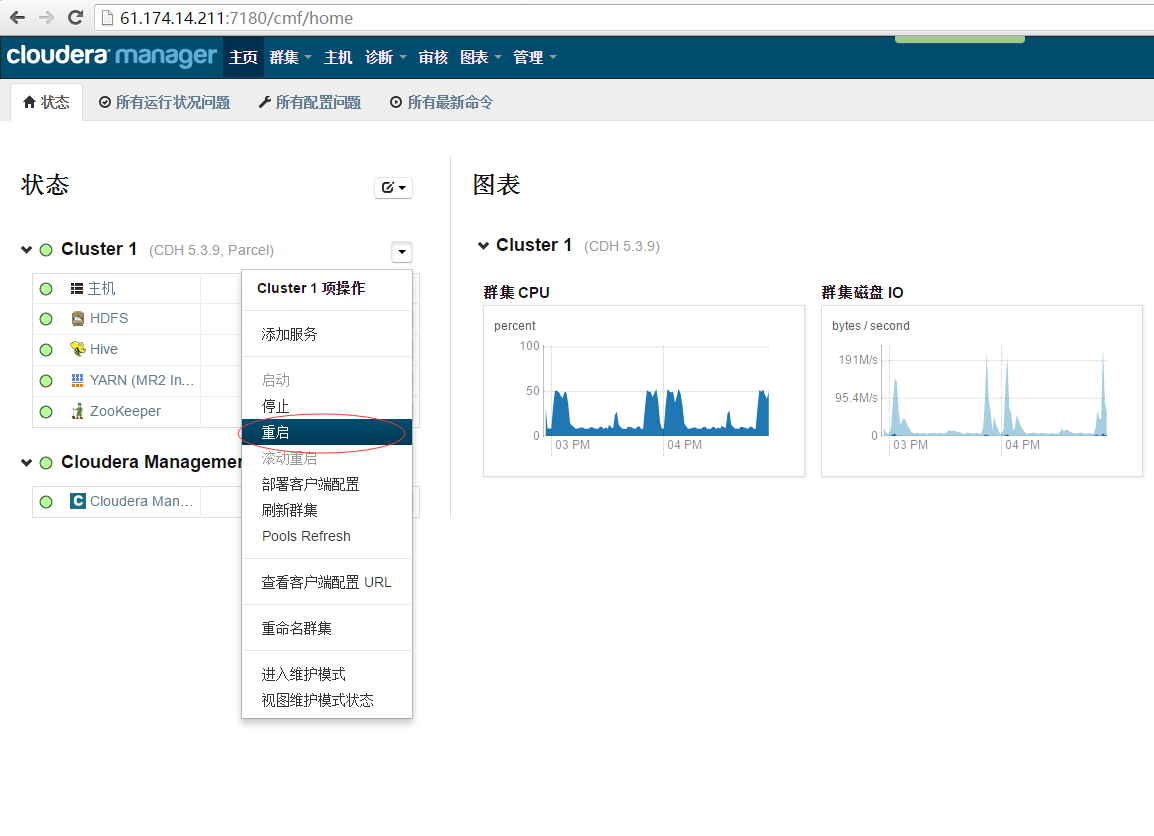


## 2、配置





## 3、维护



## 4、数据输出

[root@CTL\_ZJ\_014\_200 ~]# /root/wugq 61.174.9.94

Last login: Thu May 19 15:23:19 2016 from 61.174.14.200

[root@ctl-zj-061-174-009-094 ~]# crontab -l

0 \* \* \* \* /usr/sbin/ntpdate ntp.api.bz;/usr/sbin/hwclock -w > /dev/null 2>&1

#15 7 \* \* \* /var/lib/hadoop-hdfs/move\_data.sh> /var/result/logs/move\_data.sh.log 2>&1

#\*/5 \* \* \* \* /bin/bash /opt/bin/restart.sh >> /var/result/logs/td\_agent\_restart.log 2>&1

[root@ctl-zj-061-174-009-094 ~]# su - hdfs

-bash: ulimit: open files: cannot modify limit: Operation not permitted

-bash-4.3$

-bash-4.3$ crontab -l

#\*/15 \* \* \* \* /var/lib/hadoop-hdfs/jobs/total.sh >> /var/result/logs/total.log 2>&1

#\*/6 \* \* \* \* /var/lib/hadoop-hdfs/jobs/hourliuliang.sh >> /var/result/logs/hourliuliang.log 2>&1

\*/30 \* \* \* \* /var/lib/hadoop-hdfs/jobs/total.sh >> /var/result/logs/total.log 2>&1

\*/7 \* \* \* \* /var/lib/hadoop-hdfs/jobs/hourliuliang.sh >> /var/result/logs/hourliuliang.log 2>&1

\*/5 \* \* \* \* /var/lib/hadoop-hdfs/jobs/mod\_new.sh >> /var/result/logs/mod.log 2>&1

\*/49 \* \* \* \* /var/lib/hadoop-hdfs/jobs/domaintoday.sh >> /var/result/logs/domaintoday.log 2>&1

#10 0 \* \* \* /var/lib/hadoop-hdfs/p.sh >> /var/result/logs/p.log 2>&1

5 0 \* \* \* /var/lib/hadoop-hdfs/q.sh >> /var/result/logs/q.log 2>&1

7 \* \* \* \* /var/lib/hadoop-hdfs/analyze\_table.sh >> /var/result/logs/analyze\_table.log 2>&1

20 0 \* \* \* /var/lib/hadoop-hdfs/analyze\_table\_day.sh >> /var/result/logs/analyze\_table\_day.log 2>&1

50 0 \* \* \* /var/lib/hadoop-hdfs/jobs/domainyester.sh >> /var/result/logs/domainyester.log 2>&1

13 2 \* \* \* /var/lib/hadoop-hdfs/jobs/domain7.sh >> /var/result/logs/domain7.log 2>&1

33 3 \* \* \* /var/lib/hadoop-hdfs/jobs/domain30.sh >> /var/result/logs/domain30.log 2>&1

15 0 \* \* \* /var/lib/hadoop-hdfs/jobs/liuliang.sh> /var/result/logs/liuliang.log 2>&1

#15 6 \* \* \* /var/lib/hadoop-hdfs/move\_data.sh> /var/result/logs/move\_data.sh.log 2>&1

#15 6 \* \* \* /var/lib/hadoop-hdfs/dk/dp.sh> /var/lib/hadoop-hdfs/dk/dump\_error.log 2>&1

### migrate table from temp\_table to new\_accesslog2 ##

3 0 \* \* \* /var/lib/hadoop-hdfs/jobs/migrate\_table.sh >> /var/result/logs/migrate\_table.log 2>&1

-bash-4.3$

#数据移动

-bash-4.3$ cat /var/lib/hadoop-hdfs/jobs/migrate\_table.sh

#!/bin/bash

export PATH=/usr/local/bin:/usr/local/bin:/usr/lib64/qt-3.3/bin:/usr/local/bin:/bin:/usr/bin:/usr/local/sbin:/usr/sbin:/sbin:/usr/local/jdk1.8.0\_66/bin:/var/lib/hadoop-hdfs/bin:/usr/local/jdk1.8.0\_66/bin:/var/lib/hadoop-hdfs/bin

## 1. create hadoop dir ###

#hdfs dfs -mkdir /weblog/temp\_table

## 2. create temp table ###

#hive -e "CREATE EXTERNAL TABLE `temp\_table`(

# `time` string,

# `clientip` string,

# `method` string,

# `domain` string,

# `request` string,

# `status` string,

# `size` string,

# `referer` string,

# `agent` string,

# `hit\_status` string,

# `ruleid` string,

# `prov` string,

# `country` string,

# `blockid` string,

# `searchbot` string,

# `p\_status` string,

# `hackerid` int)

#PARTITIONED BY (

# `day` int,

# `ht` int)

#CLUSTERED BY (

# domain)

#INTO 10 BUCKETS

#ROW FORMAT SERDE

# 'org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe'

#STORED AS INPUTFORMAT

# 'org.apache.hadoop.mapred.TextInputFormat'

#OUTPUTFORMAT

# 'org.apache.hadoop.hive.ql.io.HiveIgnoreKeyTextOutputFormat'

#LOCATION

# 'hdfs://nameservice1/weblog/temp\_table'; "

## 3. hdfs move data ###

echo "============================ `date '+%Y%m%d %H%M%S'` Start =========================================="

day=`date -d "2 day ago" +%Y%m%d`

hdfs dfs -mv /weblog/temp\_table/day=$day /weblog/new\_access2

echo "============================ `date '+%Y%m%d %H%M%S'` hdfs mv complete =========================================="

## 4. create partitions table on new\_accesslog2 ###

echo "============================ `date '+%Y%m%d %H%M%S'` create partitions new\_accesslog2 start =========================================="

for i in {00..23}

do

hive -e "ALTER TABLE new\_accesslog2 ADD PARTITION(day=$day, ht=\"$i\");";

done

echo "============================ `date '+%Y%m%d %H%M%S'` create partitions new\_accesslog2 complete =========================================="

## 5. rebuild index and analyze table on new\_accesslog2

echo "============================ `date '+%Y%m%d %H%M%S'` rebuild index new\_accesslog2 start =========================================="

for h in {00..23}

do

hive -e "alter index new\_accesslog2\_index on new\_accesslog2 partition(day=$day, ht=\"$h\") rebuild;"

hive -e "analyze table new\_accesslog2 partition(day=$day, ht=\"$h\") compute statistics;"

done

echo "============================ `date '+%Y%m%d %H%M%S'` rebuild index new\_accesslog2 complete =========================================="

## 6. delete partition table and rebuild index on temp\_table ###

echo "============================ `date '+%Y%m%d %H%M%S'` delete temp\_table partitions start =========================================="

for i in {00..23}

do

hive -e "ALTER TABLE temp\_table DROP IF EXISTS PARTITION(day=$day, ht=\"$i\");"

done

echo "============================ `date '+%Y%m%d %H%M%S'` delete temp\_table partitions complete =========================================="

## 7. delete indexes and create indexes on temp\_table ###

echo "============================ `date '+%Y%m%d %H%M%S'` delete temp\_table index and create temp\_table index start =========================================="

hive -e "drop index if exists temp\_table\_index on temp\_table;"

hive -e "create index temp\_table\_index on table temp\_table(domain,hackerid,blockid,hit\_status,referer) as 'org.apache.hadoop.hive.ql.index.compact.CompactIndexHandler' with deferred rebuild;"

echo "============================ `date '+%Y%m%d %H%M%S'` delete temp\_table index and create temp\_table index complete =========================================="

## 8. rebuild index on temp\_table

echo "============================ `date '+%Y%m%d %H%M%S'` rebuild temp\_table index start =========================================="

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

for i in {00..23}

do

hive -e "alter index temp\_table\_index on temp\_table partition(day=$day\_ago, ht=\"$i\") rebuild;"

hive -e "analyze table temp\_table partition(day=$day\_ago, ht=\"$i\") compute statistics;"

done

echo "============================ `date '+%Y%m%d %H%M%S'` rebuild temp\_table index complete =========================================="

echo "============================ \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\* Done!!! \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\* =========================================="

-bash-4.3$

-bash-4.3$ ps -ef |grep td-agent

td-agent 11088 1 0 May13 ? 00:00:00 /opt/td-agent/embedded/bin/ruby /usr/sbin/td-agent --log /var/log/td-agent/td-agent.log --use-v1-config --group td-agent --daemon /var/run/td-agent/td-agent.pid

td-agent 11091 11088 10 May13 ? 16:54:53 /opt/td-agent/embedded/bin/ruby /usr/sbin/td-agent --log /var/log/td-agent/td-agent.log --use-v1-config --group td-agent --daemon /var/run/td-agent/td-agent.pid

root 11148 1 0 May13 ? 00:00:00 /opt/td-agent/embedded/bin/ruby /usr/sbin/td-agent2 --log /var/log/td-agent/td-agent2.log --use-v1-config --group td-agent --daemon /var/run/td-agent/td-agent2.pid

root 11151 11148 0 May13 ? 01:03:40 /opt/td-agent/embedded/bin/ruby /usr/sbin/td-agent2 --log /var/log/td-agent/td-agent2.log --use-v1-config --group td-agent --daemon /var/run/td-agent/td-agent2.pid

hdfs 30680 20578 0 17:03 pts/0 00:00:00 grep td-agent

-bash-4.3$

-bash-4.3$ cat /etc/td-agent/td-agent.conf

<source>

type tail

path /var/result/today/total

#pos\_file /var/log/pos/today/total.pos

format /(?<count>[^ ]\*)\t(?<type>[^ ]\*)\t(?<date>[^ ]\*)\t(?<flag>[^ ]\*)/

tag total

</source>

<match total>

type bufferize

buffer\_type file

buffer\_path /var/result/fluentd/buffer/total.\*.buffer

<config>

type http

endpoint\_url http://www.anquan.io/api/totalData

http\_method put

serializer json

</config>

</match>

<source>

type tail

path /var/result/domaintoday/mod

#pos\_file /var/log/pos/today/total.pos

format /(?<count>[^ ]\*)\t(?<liuliang>[^ ]\*)\t(?<total\_count>[^ ]\*)\t(?<total\_liuliang>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<flag>[^ ]\*)/

tag mod

</source>

<match mod>

type bufferize

buffer\_type file

buffer\_path /var/result/fluentd/buffer/mod.\*.buffer

<config>

type http

endpoint\_url http://www.anquan.io/api/mod

http\_method put

serializer json

</config>

</match>

<source>

type tail

path /var/result/yesterday/liuliang

#pos\_file /var/log/pos/yesterday/liuliang.pos

format /(?<count>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<type>[^ ]\*)\t(?<date>[^ ]\*)\t(?<flag>[^ ]\*)/

tag yt.liuliang

</source>

<match yt.liuliang>

type bufferize

buffer\_type file

buffer\_path /var/result/fluentd/buffer/yt.liuliang.\*.buffer

<config>

type http

endpoint\_url http://www.anquan.io/api/ytdTotalFlow

http\_method put

serializer json

</config>

</match>

<source>

type tail

path /var/result/domaintoday/line

#pos\_file /var/log/pos/domaintoday/line.pos

format /(?<count>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<hour>[^ ]\*)\t(?<type>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag domainline

</source>

<match domainline>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/domainline.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 3s

</match>

<source>

type tail

path /var/result/domaintoday/attack

#pos\_file /var/log/pos/domaintoday/attack.pos

format /(?<count>[^ ]\*)\t(?<type>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag domainattack

</source>

<match domainattack>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/domainattack.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 3s

</match>

<source>

type tail

path /var/result/domaintoday/pie

#pos\_file /var/log/pos/domaintoday/pie.pos

format /(?<count>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<type>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag domainpie

</source>

<match domainpie>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/domainpie.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 3s

</match>

<source>

type tail

path /var/result/domaintoday/area

#pos\_file /var/log/pos/domaintoday/area.pos

format /(?<count>[^ ]\*)\t(?<prov>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<type>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag domainarea

</source>

<match domainarea>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/domainarea.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 3s

</match>

<source>

type tail

path /var/result/domaintoday/sum

#pos\_file /var/log/pos/domaintoday/sum.pos

format /(?<pv>[^ ]\*)\t(?<ip>[^ ]\*)\t(?<size>[^ ]\*)\t(?<attack>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag domainsum

</source>

<match domainsum>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/domainsum.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 3s

</match>

<source>

type tail

path /var/result/domaintoday/ref

#pos\_file /var/log/pos/domaintoday/ref.pos

format /(?<count>[^ ]\*)\t(?<ref>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag domainref

</source>

<match domainref>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/domainref.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 3s

</match>

<source>

type tail

path /var/result/domaintoday/searchbot

#pos\_file /var/log/pos/domaintoday/searchbot.pos

format /(?<day>[^ ]\*)\t(?<searchbot>[^ ]\*)\t(?<count>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag domainsearchbot

</source>

<match domainsearchbot>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/domainsearchbot.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 3s

</match>

<source>

type tail

path /var/result/domaintoday/4xx\_5xx

#pos\_file /var/log/pos/domaintoday/4xx\_5xx.pos

format /(?<count>[^ ]\*)\t(?<type>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag domain4xx\_5xx

</source>

<match domain4xx\_5xx>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/domain4xx\_5xx.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 3s

</match>

<source>

type tail

path /var/result/domaintoday/top5\_attack

#pos\_file /var/log/pos/domaintoday/top5\_attack.pos

format /(?<ip>[^ ]\*)\t(?<prov>[^ ]\*)\t(?<count>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag domaintop5attack

</source>

<match domaintop5attack>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/domaintop5attack.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 3s

</match>

<source>

type tail

path /var/result/domaintoday/cc

#pos\_file /var/log/pos/domaintoday/cc.pos

format /(?<count>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<hour>[^ ]\*)\t(?<type>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag domaincc

</source>

<match domaincc>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/domaincc.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 3s

</match>

<source>

type tail

path /var/result/domaintoday/jiasu

#pos\_file /var/log/pos/domaintoday/jiasu.pos

format /(?<pv>[^ ]\*)\t(?<liuliang>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag domainjiasu

</source>

<match domainjiasu>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/domainjiasu.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 3s

</match>

<source>

type tail

path /var/result/domaintoday/top10\_url

#pos\_file /var/log/pos/domaintoday/top10\_url.pos

format /(?<request>[^ ]\*)\t(?<count>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<rank>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag domaintop10\_url

</source>

<match domaintop10\_url>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/domaintop10\_url.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 3s

</match>

<source>

type tail

path /var/result/day/7top5\_attack

#pos\_file /var/log/pos/day/7top5\_attack.pos

format /(?<ip>[^ ]\*)\t(?<prov>[^ ]\*)\t(?<count>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag 7top5\_attack

</source>

<match 7top5\_attack>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/7top5\_attack.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 3s

</match>

<source>

type tail

path /var/result/day/30top5\_attack

#pos\_file /var/log/pos/day/30top5\_attack.pos

format /(?<ip>[^ ]\*)\t(?<prov>[^ ]\*)\t(?<count>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag 30top5\_attack

</source>

<match 30top5\_attack>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/30top5\_attack.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 3s

</match>

<source>

type tail

path /var/result/day/7ref

#pos\_file /var/log/pos/day/7ref.pos

format /(?<count>[^ ]\*)\t(?<ref>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag 7ref

</source>

<match 7ref>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/7ref.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 3s

</match>

<source>

type tail

path /var/result/day/30ref

#pos\_file /var/log/pos/day/30ref.pos

format /(?<count>[^ ]\*)\t(?<ref>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag 30ref

</source>

<match 30ref>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/30ref.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 3s

</match>

<source>

type tail

path /var/result/day/7top10\_url

#pos\_file /var/log/pos/day/7top10\_url.pos

format /(?<request>[^ ]\*)\t(?<count>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<rank>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag 7top10\_url

</source>

<match 7top10\_url>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/7top10\_url.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 3s

</match>

<source>

type tail

path /var/result/day/30top10\_url

#pos\_file /var/log/pos/day/30top10\_url.pos

format /(?<request>[^ ]\*)\t(?<count>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<rank>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag 30top10\_url

</source>

<match 30top10\_url>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/30top10\_url.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 3s

</match>

-bash-4.3$ cat /etc/td-agent/td-agent2.conf

<source>

type tail

path /var/result/domainday/line

#pos\_file /var/log/pos/domainday/line.pos

format /(?<count>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<hour>[^ ]\*)\t(?<type>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag domainline\_day

</source>

<match domainline\_day>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/domainline\_day.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 5s

</match>

<source>

type tail

path /var/result/domainday/attack

#pos\_file /var/log/pos/domainday/attack.pos

format /(?<count>[^ ]\*)\t(?<type>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag domainattack\_day

</source>

<match domainattack\_day>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/domainattack\_day.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 5s

</match>

<source>

type tail

path /var/result/domainday/pie

#pos\_file /var/log/pos/domainday/pie.pos

format /(?<count>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<type>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag domainpie\_day

</source>

<match domainpie\_day>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/domainpie\_day.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 5s

</match>

<source>

type tail

path /var/result/domainday/area

#pos\_file /var/log/pos/domainday/area.pos

format /(?<count>[^ ]\*)\t(?<prov>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<type>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag domainarea\_day

</source>

<match domainarea\_day>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/domainarea\_day.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 5s

</match>

<source>

type tail

path /var/result/domainday/sum

#pos\_file /var/log/pos/domainday/sum.pos

format /(?<pv>[^ ]\*)\t(?<ip>[^ ]\*)\t(?<size>[^ ]\*)\t(?<attack>[^ ]\*)\t(?<cc>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag domainsum\_day

</source>

<match domainsum\_day>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/domainsum\_day.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 5s

</match>

<source>

type tail

path /var/result/domainday/ref

#pos\_file /var/log/pos/domainday/ref.pos

format /(?<count>[^ ]\*)\t(?<ref>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag domainref\_day

</source>

<match domainref\_day>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/domainref\_day.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 5s

</match>

<source>

type tail

path /var/result/domainday/searchbot

#pos\_file /var/log/pos/domainday/searchbotbot.pos

format /(?<day>[^ ]\*)\t(?<searchbot>[^ ]\*)\t(?<count>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag domainsearchbot\_day

</source>

<match domainsearchbot\_day>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/domainsearchbot\_day.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 5s

</match>

<source>

type tail

path /var/result/domainday/4xx\_5xx

#pos\_file /var/log/pos/domainday/4xx\_5xx.pos

format /(?<count>[^ ]\*)\t(?<type>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag domain4xx\_5xx\_day

</source>

<match domain4xx\_5xx\_day>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/domain4xx\_5xx\_day.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 5s

</match>

<source>

type tail

path /var/result/domainday/top5\_attack

#pos\_file /var/log/pos/domainday/top5\_attack.pos

format /(?<ip>[^ ]\*)\t(?<prov>[^ ]\*)\t(?<count>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag domaintop5attack\_day

</source>

<match domaintop5attack\_day>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/domaintop5attack\_day.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 5s

</match>

<source>

type tail

path /var/result/domainday/cc

#pos\_file /var/log/pos/domainday/cc.pos

format /(?<count>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<hour>[^ ]\*)\t(?<type>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag domaincc\_day

</source>

<match domaincc\_day>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/domaincc\_day.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 5s

</match>

<source>

type tail

path /var/result/domainday/jiasu

#pos\_file /var/log/pos/domainday/jiasu.pos

format /(?<pv>[^ ]\*)\t(?<liuliang>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag domainjiasu\_day

</source>

<match domainjiasu\_day>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/domainjiasu\_day.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 5s

</match>

<source>

type tail

path /var/result/domainday/top10\_url

#pos\_file /var/log/pos/domainday/top10\_url.pos

format /(?<request>[^ ]\*)\t(?<count>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<rank>[^ ]\*)\t(?<flag>[^ ]\*)\t(?<apiname>[^ ]\*)/

tag domaintop10\_url\_day

</source>

<match domaintop10\_url\_day>

type redis\_publish

host 61.174.14.199

port 6380

auth Qwe5sj32BSy9jpYC

db 0

buffer\_type file

buffer\_path /var/result/fluentd/buffer/domaintop10\_url\_day.buffer

buffer\_chunk\_limit 512m

buffer\_queue\_limit 500

flush\_interval 5s

</match>

<source>

type tail

path /var/result/domaintoday/hourliuliang

pos\_file /var/log/pos/domaintoday/hourliuliang.pos

format /(?<num>[^ ]\*)\t(?<domain>[^ ]\*)\t(?<hour>[^ ]\*)\t(?<flag>[^ ]\*)/

tag hourliuliang

</source>

<match hourliuliang>

type bufferize

buffer\_type file

buffer\_path /var/result/fluentd/buffer/hourliuliang.\*.buffer

<config>

type http

endpoint\_url http://www.anquan.io/api/hourliuliang

http\_method put

serializer json

</config>

</match>

include conf.d/\*.conf

-bash-4.3$