**filebeat1.3组件配置文件(filebeat.yml)中文翻译**

################### Filebeat Configuration Example #########################

############################# Filebeat ######################################

filebeat:

# List of prospectors to fetch data.

prospectors:

# Each - is a prospector. Below are the prospector specific configurations

-

# Paths that should be crawled and fetched. Glob based paths.

# To fetch all ".log" files from a specific level of subdirectories

# /var/log/\*/\*.log can be used.

# For each file found under this path, a harvester is started.

# Make sure not file is defined twice as this can lead to unexpected behaviour.

# 指定要监控的日志，可以指定具体得文件或者目录

paths:

- /var/log/\*.log

#- c:\programdata\elasticsearch\logs\\*

# Configure the file encoding for reading files with international characters

# following the W3C recommendation for HTML5 (http://www.w3.org/TR/encoding).

# Some sample encodings:

# plain, utf-8, utf-16be-bom, utf-16be, utf-16le, big5, gb18030, gbk,

# hz-gb-2312, euc-kr, euc-jp, iso-2022-jp, shift-jis, ...

# 指定被监控的文件的编码类型，使用plain和utf-8都是可以处理中文日志的

#encoding: plain

# Type of the files. Based on this the way the file is read is decided.

# The different types cannot be mixed in one prospector

#

# Possible options are:

# \* log: Reads every line of the log file (default)

# \* stdin: Reads the standard in

# 指定文件的输入类型log(默认)或者stdin

input\_type: log

# Exclude lines. A list of regular expressions to match. It drops the lines that are

# matching any regular expression from the list. The include\_lines is called before

# 在输入中排除符合正则表达式列表的那些行。

# exclude\_lines. By default, no lines are dropped.

# exclude\_lines: ["^DBG"]

# Include lines. A list of regular expressions to match. It exports the lines that are

# matching any regular expression from the list. The include\_lines is called before

# exclude\_lines. By default, all the lines are exported.

# 包含输入中符合正则表达式列表的那些行（默认包含所有行），include\_lines执行完毕之后会执行exclude\_lines

# include\_lines: ["^ERR", "^WARN"]

# Exclude files. A list of regular expressions to match. Filebeat drops the files that

# are matching any regular expression from the list. By default, no files are dropped.

# 忽略掉符合正则表达式列表的文件

# exclude\_files: [".gz$"]

# Optional additional fields. These field can be freely picked

# to add additional information to the crawled log files for filtering

# 向输出的每一条日志添加额外的信息，比如“level:debug”，方便后续对日志进行分组统计。

# 默认情况下，会在输出信息的fields子目录下以指定的新增fields建立子目录，例如fields.level

# 这个得意思就是会在es中多添加一个字段，格式为 "filelds":{"level":"debug"}

#fields:

# level: debug

# review: 1

# Set to true to store the additional fields as top level fields instead

# of under the "fields" sub-dictionary. In case of name conflicts with the

# fields added by Filebeat itself, the custom fields overwrite the default

# fields.

# 如果该选项设置为true，则新增fields成为顶级目录，而不是将其放在fields目录下。

# 自定义的field会覆盖filebeat默认的field

# 如果设置为true，则在es中新增的字段格式为："level":"debug"

#fields\_under\_root: false

# Ignore files which were modified more then the defined timespan in the past.

# In case all files on your system must be read you can set this value very large.

# Time strings like 2h (2 hours), 5m (5 minutes) can be used.

# 可以指定Filebeat忽略指定时间段以外修改的日志内容，比如2h（两个小时）或者5m(5分钟)。

#ignore\_older: 0

# Close older closes the file handler for which were not modified

# for longer then close\_older

# Time strings like 2h (2 hours), 5m (5 minutes) can be used.

# 如果一个文件在某个时间段内没有发生过更新，则关闭监控的文件handle。默认1h

#close\_older: 1h

# Type to be published in the 'type' field. For Elasticsearch output,

# the type defines the document type these entries should be stored

# in. Default: log

# 设定Elasticsearch输出时的document的type字段 可以用来给日志进行分类。Default: log

#document\_type: log

# Scan frequency in seconds.

# How often these files should be checked for changes. In case it is set

# to 0s, it is done as often as possible. Default: 10s

# Filebeat以多快的频率去prospector指定的目录下面检测文件更新（比如是否有新增文件）

# 如果设置为0s，则Filebeat会尽可能快地感知更新（占用的CPU会变高）。默认是10s

#scan\_frequency: 10s

# Defines the buffer size every harvester uses when fetching the file

# 每个harvester监控文件时，使用的buffer的大小

#harvester\_buffer\_size: 16384

# Maximum number of bytes a single log event can have

# All bytes after max\_bytes are discarded and not sent. The default is 10MB.

# This is especially useful for multiline log messages which can get large.

# 日志文件中增加一行算一个日志事件，max\_bytes限制在一次日志事件中最多上传的字节数，多出的字节会被丢弃

#max\_bytes: 10485760

# Mutiline can be used for log messages spanning multiple lines. This is common

# for Java Stack Traces or C-Line Continuation

# 适用于日志中每一条日志占据多行的情况，比如各种语言的报错信息调用栈

#multiline:

# The regexp Pattern that has to be matched. The example pattern matches all lines starting with [

# 多行日志开始的那一行匹配的pattern

#pattern: ^\[

# Defines if the pattern set under pattern should be negated or not. Default is false.

# 是否需要对pattern条件转置使用，不翻转设为true，反转设置为false。 【建议设置为true】

#negate: false

# Match can be set to "after" or "before". It is used to define if lines should be append to a pattern

# that was (not) matched before or after or as long as a pattern is not matched based on negate.

# Note: After is the equivalent to previous and before is the equivalent to to next in Logstash

# 匹配pattern后，与前面（before）还是后面（after）的内容合并为一条日志

#match: after

# The maximum number of lines that are combined to one event.

# In case there are more the max\_lines the additional lines are discarded.

# Default is 500

# 合并的最多行数（包含匹配pattern的那一行）

#max\_lines: 500

# After the defined timeout, an multiline event is sent even if no new pattern was found to start a new event

# Default is 5s.

# 到了timeout之后，即使没有匹配一个新的pattern（发生一个新的事件），也把已经匹配的日志事件发送出去

#timeout: 5s

# Setting tail\_files to true means filebeat starts readding new files at the end

# instead of the beginning. If this is used in combination with log rotation

# this can mean that the first entries of a new file are skipped.

# 如果设置为true，Filebeat从文件尾开始监控文件新增内容，把新增的每一行文件作为一个事件依次发送，

# 而不是从文件开始处重新发送所有内容

#tail\_files: false

# Backoff values define how agressively filebeat crawls new files for updates

# The default values can be used in most cases. Backoff defines how long it is waited

# to check a file again after EOF is reached. Default is 1s which means the file

# is checked every second if new lines were added. This leads to a near real time crawling.

# Every time a new line appears, backoff is reset to the initial value.

# Filebeat检测到某个文件到了EOF（文件结尾）之后，每次等待多久再去检测文件是否有更新，默认为1s

#backoff: 1s

# Max backoff defines what the maximum backoff time is. After having backed off multiple times

# from checking the files, the waiting time will never exceed max\_backoff idenependent of the

# backoff factor. Having it set to 10s means in the worst case a new line can be added to a log

# file after having backed off multiple times, it takes a maximum of 10s to read the new line

# Filebeat检测到某个文件到了EOF之后，等待检测文件更新的最大时间，默认是10秒

#max\_backoff: 10s

# The backoff factor defines how fast the algorithm backs off. The bigger the backoff factor,

# the faster the max\_backoff value is reached. If this value is set to 1, no backoff will happen.

# The backoff value will be multiplied each time with the backoff\_factor until max\_backoff is reached

# 定义到达max\_backoff的速度，默认因子是2，到达max\_backoff后，变成每次等待max\_backoff那么长的时间才backoff一次，

# 直到文件有更新才会重置为backoff

# 根据现在的默认配置是这样的，每隔1s检测一下文件变化，如果连续检测两次之后文件还没有变化，下一次检测间隔时间变为10s

#backoff\_factor: 2

# This option closes a file, as soon as the file name changes.

# This config option is recommended on windows only. Filebeat keeps the files it's reading open. This can cause

# issues when the file is removed, as the file will not be fully removed until also Filebeat closes

# the reading. Filebeat closes the file handler after ignore\_older. During this time no new file with the

# same name can be created. Turning this feature on the other hand can lead to loss of data

# on rotate files. It can happen that after file rotation the beginning of the new

# file is skipped, as the reading starts at the end. We recommend to leave this option on false

# but lower the ignore\_older value to release files faster.

# 这个选项关闭一个文件,当文件名称的变化。#该配置选项建议只在windows

#force\_close\_files: false

# Additional prospector

#-

# Configuration to use stdin input

#input\_type: stdin

# General filebeat configuration options

#

# Event count spool threshold - forces network flush if exceeded

# spooler的大小，spooler中的事件数量超过这个阈值的时候会清空发送出去（不论是否到达超时时间）

#spool\_size: 2048

# Enable async publisher pipeline in filebeat (Experimental!)

# 是否采用异步发送模式（实验功能）

#publish\_async: false

# Defines how often the spooler is flushed. After idle\_timeout the spooler is

# Flush even though spool\_size is not reached.

# spooler的超时时间，如果到了超时时间，spooler也会清空发送出去（不论是否到达容量的阈值）

#idle\_timeout: 5s

# Name of the registry file. Per default it is put in the current working

# directory. In case the working directory is changed after when running

# filebeat again, indexing starts from the beginning again.

# 记录filebeat处理日志文件的位置的文件，默认是在启动的根目录下

#registry\_file: .filebeat

# Full Path to directory with additional prospector configuration files. Each file must end with .yml

# These config files must have the full filebeat config part inside, but only

# the prospector part is processed. All global options like spool\_size are ignored.

# The config\_dir MUST point to a different directory then where the main filebeat config file is in.

# 如果要在本配置文件中引入其他位置的配置文件，可以写在这里（需要写完整路径），但是只处理prospector的部分

#config\_dir:

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

############################# Libbeat Config ##################################

# Base config file used by all other beats for using libbeat features

############################# Output ##########################################

# Configure what outputs to use when sending the data collected by the beat.

# Multiple outputs may be used.

output:

### Elasticsearch as output

elasticsearch:

# Array of hosts to connect to.

# Scheme and port can be left out and will be set to the default (http and 9200)

# In case you specify and additional path, the scheme is required: http://localhost:9200/path

# IPv6 addresses should always be defined as: https://[2001:db8::1]:9200

hosts: ["localhost:9200"]

# Optional protocol and basic auth credentials.

#protocol: "https"

#username: "admin"

#password: "s3cr3t"

# Number of workers per Elasticsearch host.

#worker: 1

# Optional index name. The default is "filebeat" and generates

# [filebeat-]YYYY.MM.DD keys.

#index: "filebeat"

# A template is used to set the mapping in Elasticsearch

# By default template loading is disabled and no template is loaded.

# These settings can be adjusted to load your own template or overwrite existing ones

#template:

# Template name. By default the template name is filebeat.

#name: "filebeat"

# Path to template file

#path: "filebeat.template.json"

# Overwrite existing template

#overwrite: false

# Optional HTTP Path

#path: "/elasticsearch"

# Proxy server url

#proxy\_url: http://proxy:3128

# The number of times a particular Elasticsearch index operation is attempted. If

# the indexing operation doesn't succeed after this many retries, the events are

# dropped. The default is 3.

#max\_retries: 3

# The maximum number of events to bulk in a single Elasticsearch bulk API index request.

# The default is 50.

#bulk\_max\_size: 50

# Configure http request timeout before failing an request to Elasticsearch.

#timeout: 90

# The number of seconds to wait for new events between two bulk API index requests.

# If `bulk\_max\_size` is reached before this interval expires, addition bulk index

# requests are made.

#flush\_interval: 1

# Boolean that sets if the topology is kept in Elasticsearch. The default is

# false. This option makes sense only for Packetbeat.

#save\_topology: false

# The time to live in seconds for the topology information that is stored in

# Elasticsearch. The default is 15 seconds.

#topology\_expire: 15

# tls configuration. By default is off.

#tls:

# List of root certificates for HTTPS server verifications

#certificate\_authorities: ["/etc/pki/root/ca.pem"]

# Certificate for TLS client authentication

#certificate: "/etc/pki/client/cert.pem"

# Client Certificate Key

#certificate\_key: "/etc/pki/client/cert.key"

# Controls whether the client verifies server certificates and host name.

# If insecure is set to true, all server host names and certificates will be

# accepted. In this mode TLS based connections are susceptible to

# man-in-the-middle attacks. Use only for testing.

#insecure: true

# Configure cipher suites to be used for TLS connections

#cipher\_suites: []

# Configure curve types for ECDHE based cipher suites

#curve\_types: []

# Configure minimum TLS version allowed for connection to logstash

#min\_version: 1.0

# Configure maximum TLS version allowed for connection to logstash

#max\_version: 1.2

### Logstash as output

#logstash:

# The Logstash hosts

#hosts: ["localhost:5044"]

# Number of workers per Logstash host.

#worker: 1

# The maximum number of events to bulk into a single batch window. The

# default is 2048.

#bulk\_max\_size: 2048

# Set gzip compression level.

#compression\_level: 3

# Optional load balance the events between the Logstash hosts

#loadbalance: true

# Optional index name. The default index name depends on the each beat.

# For Packetbeat, the default is set to packetbeat, for Topbeat

# top topbeat and for Filebeat to filebeat.

#index: filebeat

# Optional TLS. By default is off.

#tls:

# List of root certificates for HTTPS server verifications

#certificate\_authorities: ["/etc/pki/root/ca.pem"]

# Certificate for TLS client authentication

#certificate: "/etc/pki/client/cert.pem"

# Client Certificate Key

#certificate\_key: "/etc/pki/client/cert.key"

# Controls whether the client verifies server certificates and host name.

# If insecure is set to true, all server host names and certificates will be

# accepted. In this mode TLS based connections are susceptible to

# man-in-the-middle attacks. Use only for testing.

#insecure: true

# Configure cipher suites to be used for TLS connections

#cipher\_suites: []

# Configure curve types for ECDHE based cipher suites

#curve\_types: []

### File as output

#file:

# Path to the directory where to save the generated files. The option is mandatory.

#path: "/tmp/filebeat"

# Name of the generated files. The default is `filebeat` and it generates files: `filebeat`, `filebeat.1`, `filebeat.2`, etc.

#filename: filebeat

# Maximum size in kilobytes of each file. When this size is reached, the files are

# rotated. The default value is 10 MB.

#rotate\_every\_kb: 10000

# Maximum number of files under path. When this number of files is reached, the

# oldest file is deleted and the rest are shifted from last to first. The default

# is 7 files.

#number\_of\_files: 7

### Console output

# console:

# Pretty print json event

#pretty: false

############################# Shipper #########################################

shipper:

# The name of the shipper that publishes the network data. It can be used to group

# all the transactions sent by a single shipper in the web interface.

# If this options is not defined, the hostname is used.

#name:

# The tags of the shipper are included in their own field with each

# transaction published. Tags make it easy to group servers by different

# logical properties.

#tags: ["service-X", "web-tier"]

# Uncomment the following if you want to ignore transactions created

# by the server on which the shipper is installed. This option is useful

# to remove duplicates if shippers are installed on multiple servers.

#ignore\_outgoing: true

# How often (in seconds) shippers are publishing their IPs to the topology map.

# The default is 10 seconds.

#refresh\_topology\_freq: 10

# Expiration time (in seconds) of the IPs published by a shipper to the topology map.

# All the IPs will be deleted afterwards. Note, that the value must be higher than

# refresh\_topology\_freq. The default is 15 seconds.

#topology\_expire: 15

# Internal queue size for single events in processing pipeline

#queue\_size: 1000

# Configure local GeoIP database support.

# If no paths are not configured geoip is disabled.

#geoip:

#paths:

# - "/usr/share/GeoIP/GeoLiteCity.dat"

# - "/usr/local/var/GeoIP/GeoLiteCity.dat"

############################# Logging #########################################

# There are three options for the log ouput: syslog, file, stderr.

# Under Windos systems, the log files are per default sent to the file output,

# under all other system per default to syslog.

# 建议在开发时期开启日志并把日志调整为debug或者info级别，在生产环境下调整为error级别

# 开启日志 必须设置to\_files 属性为true

logging:

# Send all logging output to syslog. On Windows default is false, otherwise

# default is true.

# 配置beats日志。日志可以写入到syslog也可以是轮滚日志文件。默认是syslog

# tail -f /var/log/messages

#to\_syslog: true

# Write all logging output to files. Beats automatically rotate files if rotateeverybytes

# limit is reached.

# 日志发送到轮滚文件

#to\_files: false

# To enable logging to files, to\_files option has to be set to true

# to\_files设置为true才可以开启轮滚日志记录

files:

# The directory where the log files will written to.

# 指定日志路径

#path: /var/log/mybeat

# The name of the files where the logs are written to.

# 指定日志名称

#name: mybeat

# Configure log file size limit. If limit is reached, log file will be

# automatically rotated

# 默认文件达到10M就会滚动生成新文件

rotateeverybytes: 10485760 # = 10MB

# Number of rotated log files to keep. Oldest files will be deleted first.

# 保留日志文件周期。 默认 7天。值范围为2 到 1024

#keepfiles: 7

# Enable debug output for selected components. To enable all selectors use ["\*"]

# Other available selectors are beat, publish, service

# Multiple selectors can be chained.

#selectors: [ ]

# Sets log level. The default log level is error.

# Available log levels are: critical, error, warning, info, debug

# 日志级别，默认是error

#level: error