

Elasticsearch配置参数介绍

Elasticsearch的config文件夹里面有两个配置文件：elasticsearch.yml和logging.yml。第一个是es的基本配置文件，第二个是日志配置文件，es也是使用log4j来记录日志的，所以logging.yml里的设置按普通log4j配置文件来设置就行了。下面主要讲解下elasticsearch.yml这个文件中可配置的东西。



如果想及时了解Spark、Hadoop或者Hbase相关的文章，欢迎关注微信公共帐号：iteblog_hadoop

cluster.name:elasticsearch

配置es的集群名称，默认是elasticsearch，es会自动发现在同一网段下的es，如果在同一网段下有多个集群，就可以用这个属性来区分不同的集群。

node.name

节点名，默认随机指定一个name列表中名字，该列表在es的jar包中config文件夹里name.txt文件中，其中有很多作者添加的有趣名字。

node.master

指定该节点是否有资格被选举成为node，默认是true，es是默认集群中的第一台机器为master，如果这台机挂了就会重新选举master。

node.data

指定该节点是否存储索引数据，默认为true。

index.number_of_shards

设置默认索引分片个数，默认为5片。

index.number_of_replicas

设置默认索引副本个数，默认为1个副本。

path.conf

设置配置文件的存储路径，默认是es根目录下的config文件夹。

path.data

设置索引数据的存储路径，默认是es根目录下的data文件夹，可以设置多个存储路径，用逗号隔开，例：

path.data:/path/to/data1,/path/to/data2

path.work

设置临时文件的存储路径，默认是es根目录下的work文件夹。

path.logs

设置日志文件的存储路径，默认是es根目录下的logs文件夹

path.plugins

设置插件的存放路径，默认是es根目录下的plugins文件夹

bootstrap.mlockall

设置为true来锁住内存。因为当jvm开始swapping时es的效率会降低，所以要保证它不swap，可以把ES_MIN_MEM和ES_MAX_MEM两个环境变量设置成同一个值，并且保证机器有足够的内存分配给es。同时也要允许elasticsearch的进程可以锁住内存，linux下可以通过`ulimit-l unlimited`命令。

network.bind_host

设置绑定的ip地址，可以是ipv4或ipv6的，默认为0.0.0.0。

network.publish_host

设置其它节点和该节点交互的ip地址，如果不设置它会自动判断，值必须是个真实的ip地址。

network.host

这个参数是用来同时设置bind_host和publish_host上面两个参数。

transport.tcp.port

设置节点间交互的tcp端口，默认是9300。

transport.tcp.compress

设置是否压缩tcp传输时的数据，默认为false，不压缩。

http.port

设置对外服务的http端口，默认为9200。

http.max_content_length

设置内容的最大容量，默认100mb

http.enabled

是否使用http协议对外提供服务，默认为true，开启。

gateway.type

gateway的类型，默认为local即为本地文件系统，可以设置为本地文件系统，分布式文件系统，hadoop的HDFS，和amazon的s3服务器，其它文件系统的设置方法下次再详细说。

gateway.recover_after_nodes

设置集群中N个节点启动时进行数据恢复，默认为1。

gateway.recover_after_time

设置初始化数据恢复进程的超时时间，默认是5分钟。

gateway.expected_nodes

设置这个集群中节点的数量，默认为2，一旦这N个节点启动，就会立即进行数据恢复。

cluster.routing.allocation.node_initial_primaries_recoveries

初始化数据恢复时，并发恢复线程的个数，默认为4。

cluster.routing.allocation.node_concurrent_recoveries

添加删除节点或负载均衡时并发恢复线程的个数，默认为4。

indices.recovery.max_size_per_sec

设置数据恢复时限制的带宽，如入100mb，默认为0，即无限制。

indices.recovery.concurrent_streams

设置这个参数来限制从其它分片恢复数据时最大同时打开并发流的个数，默认为5。

discovery.zen.minimum_master_nodes

设置这个参数来保证集群中的节点可以知道其它N个有master资格的节点。默认为1，对于大的集群来说，可以设置大一点的值（2-4）

discovery.zen.ping.timeout

设置集群中自动发现其它节点时ping连接超时时间，默认为3秒，对于比较差的网络环境可以高点的值来防止自动发现时出错。

discovery.zen.ping.multicast.enabled

设置是否打开多播发现节点，默认是true。

discovery.zen.ping.unicast.hosts

设置集群中master节点的初始列表，可以通过这些节点来自动发现新加入集群的节点。

更多的参数介绍请参见Elasticsearch的官方文档。

本博客文章除特别声明，全部都是原创！
原创文章版权归过往记忆大数据（[过往记忆](#)）所有，未经许可不得转载。
本文链接: [【】（）](#)