### 一、下载Elasticsearch安装包

<https://www.elastic.co/downloads/elasticsearch>

### 二、解压elasticsearch-2.3.5.tar.gz到/data

|  |  |  |
| --- | --- | --- |
| |  | | --- | | mkdir /data  cd /data  tar -zxvf elasticsearch-2.3.5.tar.gz | |  |

### 三、Elasticsearch配置

|  |  |  |
| --- | --- | --- |
| |  | | --- | | #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的配置    #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。  cluster.name: elasticsearch    #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.name: "node-1"    #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.master: true    #指定该节点是否存储索引数据，默认为true。  node.data: true    #设置默认索引分片个数，默认为5片。  index.number\_of\_shards: 5    #设置默认索引副本个数，默认为1个副本。  index.number\_of\_replicas: 1    #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.conf: /path/to/conf    #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/data    #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.data: /path/to/data1,/path/to/data2  path.work: /path/to/work    #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.logs: /path/to/logs    #设置插件的存放路径，默认是es根目录下的plugins文件夹  path.plugins: /path/to/plugins    #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。  bootstrap.mlockall: true    #设置绑定的ip地址，可以是ipv4或ipv6的，默认为0.0.0.0。  network.bind\_host: 192.168.0.1    #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.publish\_host: 192.168.0.1    #这个参数是用来同时设置bind\_host和publish\_host上面两个参数。  network.host: 192.168.0.1    #设置节点间交互的tcp端口，默认是9300。  transport.tcp.port: 9300    #设置是否压缩tcp传输时的数据，默认为false，不压缩。  transport.tcp.compress: true    #设置对外服务的http端口，默认为9200。  http.port: 9200    #设置内容的最大容量，默认100mb  http.max\_content\_length: 100mb    #是否使用http协议对外提供服务，默认为true，开启。  http.enabled: false    #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  gateway.type: local    #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\_nodes: 1    #设置初始化数据恢复进程的超时时间，默认是5分钟。  gateway.recover\_after\_time: 5m    #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。  gateway.expected\_nodes: 2    #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\_initial\_primaries\_recoveries: 4    #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。  cluster.routing.allocation.node\_concurrent\_recoveries: 2    #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.max\_size\_per\_sec: 0    #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。  indices.recovery.concurrent\_streams: 5    #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.minimum\_master\_nodes: 1    #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.timeout: 3s    #设置是否打开多播发现节点，默认是true。  discovery.zen.ping.multicast.enabled: false    #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。由于到了2.x版本之后，ES取消了默认的广播模式来发现master节点  discovery.zen.ping.unicast.hosts: ["host1", "host2:port", "host3[portX-portY]"]    #下面是一些查询时的慢日志参数设置  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    # ------------- Custom Analyze Begin-----------------------------  # <https://www.elastic.co/guide/en/elasticsearch/reference/2.3/analysis-custom-analyzer.html>  # 分词器也可以定义在elasticsearch.yml配置中，也可以在索引创建时定义，这样不会影响到其他索引  # 以下自定义分词器配置仅供参数    index :  analysis :  analyzer :  myAnalyzer2 :  type : custom  tokenizer : myTokenizer1  filter : [myTokenFilter1, myTokenFilter2]  char\_filter : [my\_html]  position\_increment\_gap: 256  user\_name\_analyzer :  type : custom  tokenizer : whitespace  filter : pinyin\_first\_letter\_and\_full\_pinyin\_filter  ik\_pinyin :  type : custom  tokenizer : ik  filter : [pinyin\_mcl, word\_delimiter]  tokenizer :  myTokenizer1 :  type : standard  max\_token\_length : 900  my\_pinyin :  type : pinyin  keep\_separate\_first\_letter : false  keep\_full\_pinyin : true  keep\_original : true  limit\_first\_letter\_length : 16  lowercase" : true  filter :  myTokenFilter1 :  type : stop  stopwords : [stop1, stop2, stop3, stop4]  myTokenFilter2 :  type : length  min : 0  max : 2000  pinyin\_first\_letter\_and\_full\_pinyin\_filter :  type : pinyin  keep\_first\_letter : true  keep\_full\_pinyin : false  keep\_none\_chinese : true  keep\_original : false  limit\_first\_letter\_length : 16  lowercase : true  trim\_whitespace : true  keep\_none\_chinese\_in\_first\_letter : true  pinyin\_mcl :  type: pinyin  first\_letter: none  keep\_full\_pinyin: false  keep\_joined\_full\_pinyin: true  lowercase : true  keep\_original : true  padding\_char: " "  char\_filter :  my\_html :  type : html\_strip  escaped\_tags : [xxx, yyy]  read\_ahead : 1024    # --------------- Custom Analyze End---------------- | |  |

参考：<http://www.linuxidc.com/Linux/2015-02/114244.htm>

### 四、创建elasticsearch用户

|  |
| --- |
| groupadd elasticsearch  useradd elasticsearch -g elasticsearch -s /sbin/nologin  ##修改elasticsearch文件的用户属性  chown -R elasticsearch:elasticsearch /data/elasticsearch-2.3.5 |

### 五、开启端口

|  |
| --- |
| iptables -I INPUT -p tcp --dport 9200 -j ACCEPT  iptables -I INPUT -p tcp --dport 9300 -j ACCEPT    #保存配置到/etc/sysconfig/iptables  service iptables save |

### 六、安装ES监控管理插件es-head

插件地址：<https://github.com/mobz/elasticsearch-head>

|  |
| --- |
| ./plugin install mobz/elasticsearch-head  ##离线安装  bin/plugin install file:/home/××/elasticsearch-head.zip |

测试：<http://192.168.xxx.xxx:9200/_plugin/head/>

### 七、将elasticsearch添加到服务

注：ES的服务插件elasticsearch-servicewrapper只适应2.x以下版本，不包2.x

可通过rpm方式安装得到启动脚本

|  |
| --- |
| wget https://artifacts.elastic.co/downloads/elasticsearch/elasticsearch-2.3.5.rpm  sha1sum elasticsearch-2.3.5.rpm  sudo rpm --install elasticsearch-2.3.5.rpm  #启动服务的脚本存放于/etc/init.d/elasticsearch |

参考文件：<elasticsearch服务启动脚本.zip>

### 八、添加各种分词器插件

ik分词器：<https://github.com/medcl/elasticsearch-analysis-ik>

pinyin分词器：<https://github.com/medcl/elasticsearch-analysis-pinyin>

注：1.7.5版本配置补全，需使用1.8.2版本

|  |
| --- |
| #拼音分词的配置参数说明  remove\_duplicated\_term：  默认：false  当开启此选项，将删除重复项节约指标，如：de的 -> de  注：位置相关查询可能受到影响    keep\_first\_letter：  默认true  当开启此选项，将保留分词的每个汉字的第一个字母，例如：刘德华 -> ldh    keep\_separate\_first\_letter：  默认:false  时启用此选项，将第一个字母分开，例如：刘德华 -> l，d，h  注：查询结果可能因分词颗粒过细导致准确率降低    limit\_first\_letter\_length：  长度默认值：16  设置最大的first\_letter结果    keep\_full\_pinyin：  默认：true  每个汉字都进行拼音分词，例如：刘德华 -> [liu,de,hua]    keep\_joined\_full\_pinyin：  默认：flase  分词的全拼音，例如：刘德华 -> [liudehua]    keep\_none\_chinese  默认：true  保留非中文字母或数字的结果    keep\_none\_chinese\_together：  默认：true  不拆分非中文字符  例如：DJ音乐家 -> DJ,yin,yue,jia，当设置为false，例如：DJ音乐家-> D,J,yin,yue,jia，  注：keep\_none\_chinese 建议开启    keep\_none\_chinese\_in\_first\_letter：  默认：true  当存在中文和非中文字符时，将中文保留首字母，例如：刘德华at2016 -> ldhat2016，    none\_chinese\_pinyin\_tokenize：  默认：true  非中文的拼音分词，，例如：liudehuaalibaba13zhuanghan ->liu,de,hua,a,li,ba,ba,13,zhuang,han  注：keep\_none\_chinese 和 keep\_none\_chinese\_together 建议开启    keep\_original：  默认：false  保留原来字符    lowercase：  默认为：真  非中文字符转小写    trim\_whitespace:  默认：真 |

注意，分布式系统整个集群节点个数N要为奇数个！！

### Q&A

[root@w03 elasticsearch-2.2.0]# ./bin/elasticsearchException in thread "main" java.lang.RuntimeException: don't run elasticsearch as root.

       at org.elasticsearch.bootstrap.Bootstrap.initializeNatives(Bootstrap.java:93)

       at org.elasticsearch.bootstrap.Bootstrap.setup(Bootstrap.java:144)

       at org.elasticsearch.bootstrap.Bootstrap.init(Bootstrap.java:285)

       at org.elasticsearch.bootstrap.Elasticsearch.main(Elasticsearch.java:35)

Refer to the log for complete error details.

|  |  |
| --- | --- |
| 1 | bin/elasticsearch -Des.insecure.allow.root=true |