Kettle-elasticsearch环境搭建

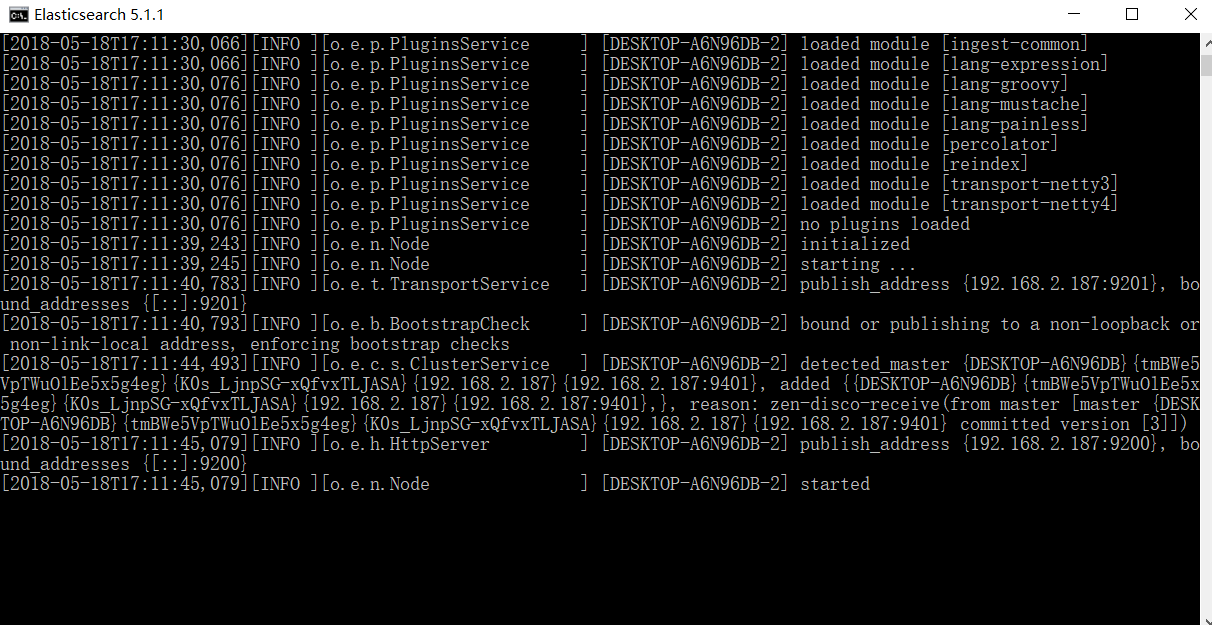
1. elasticsearch下载地址：（elasticsearch-5.1.1.zip）

[**https://www.elastic.co/downloads/past-releases**](https://www.elastic.co/downloads/past-releases) **(包含各种版本)**

**最好是** elasticsearch-5.1.1.zip 版本不然Kettle链接elasticsearch不成功

**把下载好的文件 例如：**elasticsearch-5.2.0.zip **解压到**  D:\software\elasticsearch-5.2.0**到目录下，无需安装。**

1. **启动 bin目录下找到 elasticeearch.bat双击即可启动 如下图启动成功**



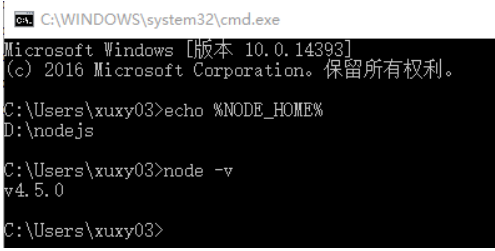
**3．windows环境下ElasticSearch5以上版本安装head插件**

# 步骤

下载node.js ,网址：<https://nodejs.org/en/>

安装node到D盘。如D:\nodejs。

把NODE\_HOME设置到环境变量里(安装包也可以自动加入PATH环境变量)。测试一下node是否生效

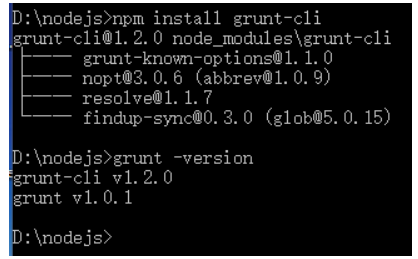


## 安装grunt

grunt是一个很方便的构建工具，可以进行打包压缩、测试、执行等等的工作，5.0里的head插件就是通过grunt启动的。因此需要安装grunt：

注意：路径切到D:\nodejs下。执行命令：nmp install -g grunt-cli

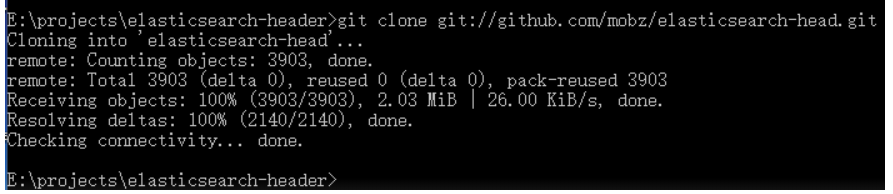
-g代表全局安装。安装路径为C:\Users\yourname\AppData\Roaming\npm，并且自动加入PATH变量。安装完成后检查一下：



把head插件的源码git clone下来：

git clone git://github.com/mobz/elasticsearch-head.git

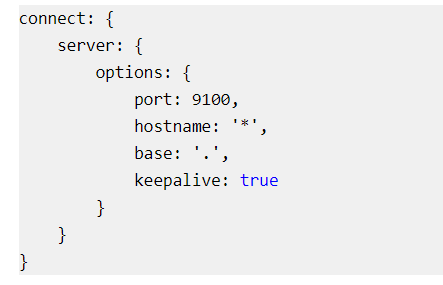
效果图：



## 修改head源码

由于head的代码还是2.6版本的，直接执行有很多限制，比如无法跨机器访问。因此需要用户修改两个地方：

目录：head/Gruntfile.js：



增加hostname属性，设置为\*

#### 修改连接地址：

目录：head/\_site/app.js

修改head的连接地址:

this.base\_uri = this.config.base\_uri || this.prefs.get("app-base\_uri") || "http://localhost:9200";

把localhost修改成你es的服务器地址，如：

this.base\_uri = this.config.base\_uri || this.prefs.get("app-base\_uri") || "http://10.10.10.10:9200";

#### 修改elasticsearch的参数

修改一下es使用的参数。编辑config/elasticsearch.yml：

# 增加新的参数，这样head插件可以访问es

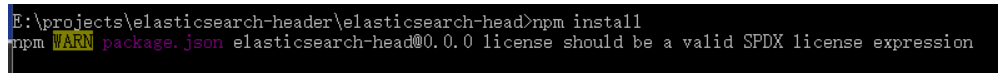
http.cors.enabled: true

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

注意，设置参数的时候 : 后面要有空格！参数前面也要有空格！

然后在head源码目录中，执行npm install 下载的包：npm install

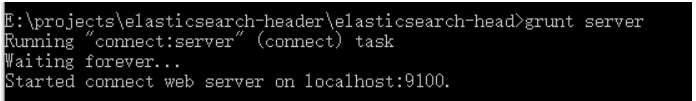
**效果如图：**



初次运行安装可能会报警告或错误。可以重新运行一次npm install。

最后，在head源代码目录下启动nodejs：grunt server

效果如图：



**启动成功后访问：**[**http://127.0.0.1:9200**](http://127.0.0.1:9200)**访问效果如下：**



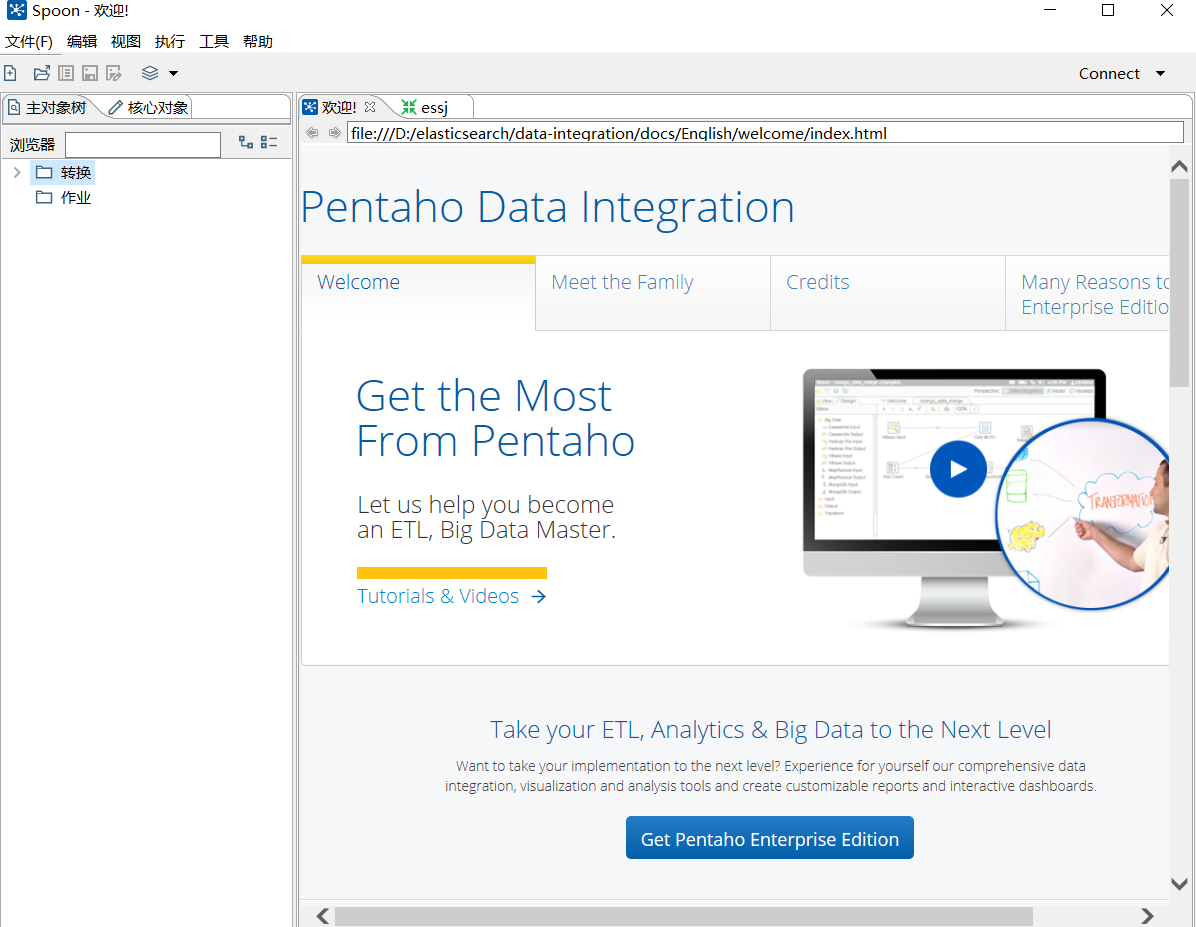
**至此elasticsearch 与 elasticsearch-head 环境搭建完成。**

Kettle 下载地址：（pdi-ce-7.1.0.0-12.zip）

[**https://sourceforge.net/projects/pentaho/files/Data%20Integration/**](https://sourceforge.net/projects/pentaho/files/Data%20Integration/) **（各个版本都有自己挑选）**

**下载解压后找到：Spoon.bat 双击启动**

**启动效果如下：**



从oracle数据库读出数据写入到es

1. 流程



1. 设置

Kettle中设置ElasticSearch需要配置如下几个参数

（1）服务地址：包括地址和[端口号](https://www.ggdoc.com/56uv5Y_j5Y_30)

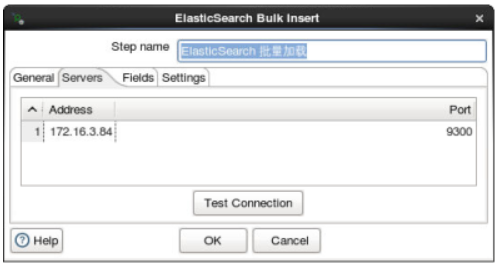
（2）索引名

（3）索引中的type表名

（4）集群名：如果更换了集群的名字（即不是默认的），需要修改集群名

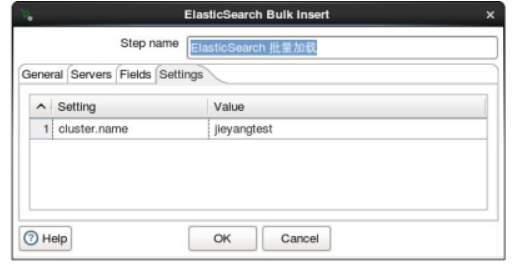
（5）需要插入的数据字段

Servers选项卡设置（此处的prot的端口为：ElasticSearch中elasticsearch.yml中的transport.tcp.port: 9401端口 不能错）



设置地址和端口

（2）配置集群名字

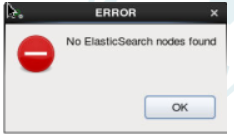


Setting中的值固定，为cluster.name

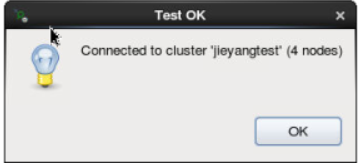
Value为集群名

设置完这两项可以点击test connection查看是否正确连接

这里一般会出现如下问题：

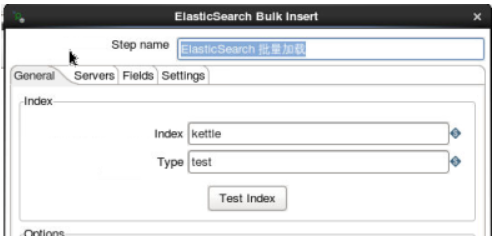


解决办法：下载elasticsearch-bulk-insert-plugin插件替换掉原来的插件

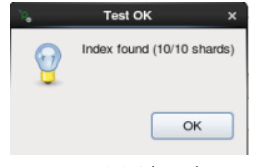


连接正确！

（3）General选项卡（index必须先创建，否则失败）



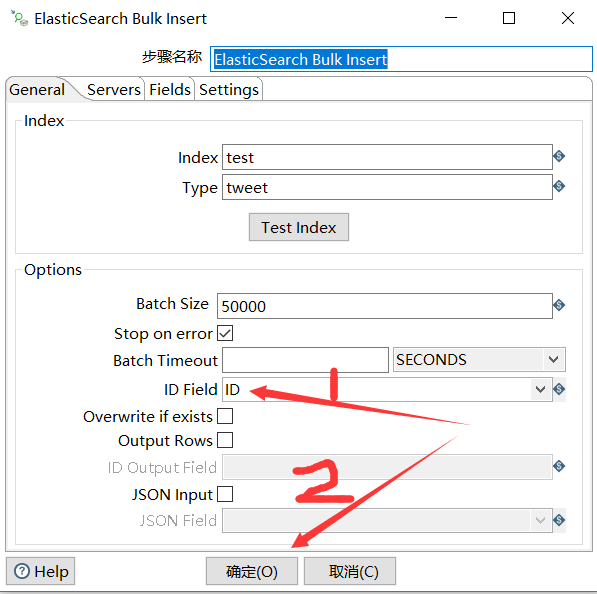
设置索引名和表名，设置完可以测试是否有这个索引（表名即使没有也可以，插入时可以自动增加）



4）Fields选项卡

需要插入的[字段名](https://www.ggdoc.com/5a2X5q615ZCN0)，从上一步骤获取（链接数据库后才能获取）

至此配置完毕（id Field必须写）

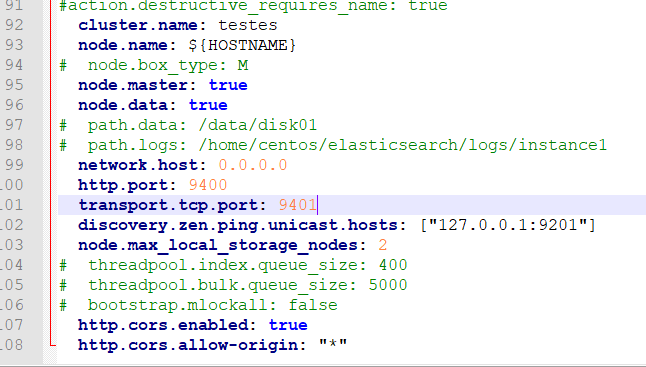


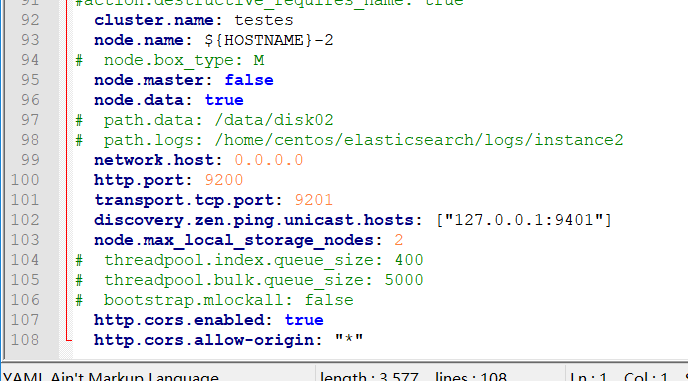
Windows 下elasticsearch 单机 集群配置

弄两份 elasticsearch 解压出来的文件



两份elasticsearch.yml文件的配置如下：





**注意两个节点的启动顺序：先启动master节点**