API文档

所有APP方法都尽可能向贴近原生REST API的方向设计，包括方法的必需参数与可选参数之间的区别。这也意味着不同位置参数（positional arguments）与关键词参数（keyword arguments）的区别会产生不同的影响。当然，我们推荐使用者优先使用关键词参数来保持代码的一致性和安全性。

Note

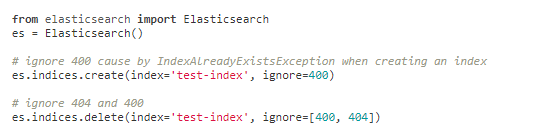
为了与Python语言生态系统相兼容，我们使用from\_与doc\_type而不是from与type（这两者是python的保留关键字）作为参数的参数名。

全局选项(Global options)

这里所讲述的是一些被添加到客户端本身并可以在任何API方法中使用的参数。

忽略（Ignore）

如果客户端的方法返回类似2XX这样的代码，则表示API调用是成功的。否则客户端会抛出一例TransportError异常或更加精准的子级异常。你可以在Exceptions相关的章节中查看各种异常（以及对应的情况）。如果你不希望客户端抛出异常，那么你在调用接口的时候可以传入ignore参数，这样对应你输入参数的返回状态码就会被忽略。



超时（Timeout）

全局超时（timeout）参数可以在两种情况设置：一、当你创建客户端的时候（查看 connection 的 timeout 参数），二、任何使用request\_timeout（以秒为单位的浮点数）并以请求为基础的API调用。超时这个数值会通过perform\_request这种connection类的方法进行传递。



Note

一些API函数调用也接受将被传递给Elasticsearch服务器的timeout参数。这个超时参数是内部使用的，并不能保证这个请求会在一个具体的时间点终止。

响应过滤（Response Filtering）

Filter\_path这个参数用于筛去一些由Elasticsearch返回的信息。举个例子，如果你使用es.search方法搜索某条记录，而只想得到\_id 和\_type两条信息，那么你可以这样操作：



当然使用 \* 这样的匹配的符号也是被支持的，比如我想查询任意分区下的数据。可以这样操作：



Elasticsearch



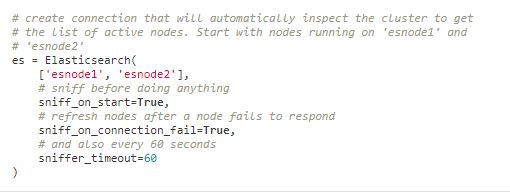
Elasticsearch是个低级客户端，查看、创建、搜索这些最基础和直接的操作都可以由这个客户端完成。

当你创建了一个Elasticsearch实例之后，它拥有cat,cluster,indices,ingest,nodes,snapshot,tasks这些属性，以及可以连接对应CatClient,ClusterClient,IndicesClient,IngestClient,NodesClient,SnapshotClient,TasksClient的实例客户端。这些xxxClient是推荐的，或说唯一支持的连接这些类方法的方式。

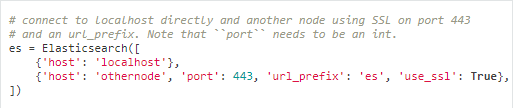
当然你可以通过提供connection\_class这个参数来指定你自己的连接类。



如果你希望打开 Sniffing，你还需填写几个额外的参数（在Transport相关章节中有描述）：



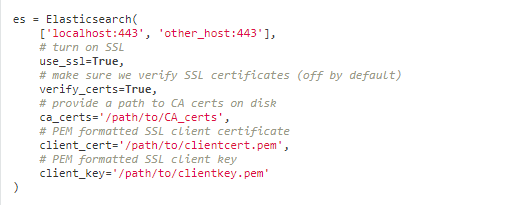
如果你要使用Elasticsearch来连接一个集群，而且不同的ES主机还有不同的连接配置，那你可以使用字典来对单独ES主机进行配置：



如果要使用SSL，那么就要填写几个控制处理认证相关的参数。（查看 Urllib3HttpConnection 的具体使用说明）



SSL客户端认证也是支持的。



另外地，你也可以使用RFC-1738格式的URL，前提是他们不和其他选项产生冲突。

参数： \* hosts 填写我们要连接的ES节点的列表。节点的配置信息应该是一个字典，类似 ({“host”: “localhost”, “port”: 9200}),整个字典中的内容都会作为参数被传递给Connections 类。或者你可以用host[:port] 这种字符串的形式传递这个信息。如果没有数值传入，那么Urllib3httpConnection类就会被设置为默认。

transport\_class 作为Transport的次级类。

kwargs 额外参数

bulk(\*args,\*\*kwargs)

在单独的API函数调用中承担大量index和delete操作。

查看bulk()的帮助函数。

<http://www.elastic.co/guide/en/elasticsearch/reference/current/docs-bulk.html>

参数：

body 包含操作的定义以及数据内容（以动作-数据对的形式），通过换行符来分隔

index 数据内容(item)的默认索引

doc\_type 数据内容的默认文档类型

\_source True or false to return the \_source field or not, or default list of fields to return, can be overridden on each sub- request

\_source\_exclude – Default list of fields to exclude from the returned \_source field, can be overridden on each sub-request

\_source\_include – Default list of fields to extract and return from the \_source field, can be overridden on each sub-request

fields – Default comma-separated list of fields to return in the response for updates, can be overridden on each sub-request

pipeline – The pipeline id to preprocess incoming documents with

refresh – If true then refresh the effected shards to make this operation visible to search, if wait\_for then wait for a refresh to make this operation visible to search, if false (the default) then do nothing with refreshes., valid choices are: ‘true’, ‘false’, ‘wait\_for’

routing – Specific routing value

timeout – Explicit operation timeout

wait\_for\_active\_shards – Sets the number of shard copies that must be active before proceeding with the bulk operation. Defaults to 1, meaning the primary shard only. Set to all for all shard copies, otherwise set to any non-negative value less than or equal to the total number of copies for the shard (number of replicas + 1)

clear\_scroll(\*args,\*\*kwargs)

清除由指定的滚动参数(scroll parameter)进行搜索所创建的滚动请求(scroll request)。

<http://www.elastic.co/guide/en/elasticsearch/reference/current/search-request-scroll.html>

参数:

scroll\_id 要清除的滚动编号(scroll IDS)，如果有多个值，则要用逗号进行分隔。

body A comma-separated list of scroll IDs to clear if none was specified via the scroll\_id parameter

count(\*args,\*\*kwargs)

执行一次查询，并获取所有满足查询条件的内容。

<http://www.elastic.co/guide/en/elasticsearch/reference/current/search-count.html>

参数:

index 数据所在索引，若多个值则使用逗号分隔。

doc\_type 文档类型，同样多值逗号分隔。

body – A query to restrict the results specified with the Query DSL (optional)

allow\_no\_indices – Whether to ignore if a wildcard indices expression resolves into no concrete indices. (This includes \_all string or when no indices have been specified)

analyze\_wildcard – Specify whether wildcard and prefix queries should be analyzed (default: false)

analyzer – The analyzer to use for the query string

default\_operator – The default operator for query string query (AND or OR), default ‘OR’, valid choices are: ‘AND’, ‘OR’

df – The field to use as default where no field prefix is given in the query string

expand\_wildcards – Whether to expand wildcard expression to concrete indices that are open, closed or both., default ‘open’, valid choices are: ‘open’, ‘closed’, ‘none’, ‘all’

ignore\_unavailable – Whether specified concrete indices should be ignored when unavailable (missing or closed)

lenient – Specify whether format-based query failures (such as providing text to a numeric field) should be ignored

lowercase\_expanded\_terms – Specify whether query terms should be lowercased

min\_score – Include only documents with a specific \_score value in the result

preference – Specify the node or shard the operation should be performed on (default: random)

q – Query in the Lucene query string syntax

routing – Specific routing value

count\_percolate(\*args,\*\*kwargs)