Elasticsearch Documentation

Release 7.0.0

Honza Král

Apr 25, 2019

Contents

Official low-level client for Elasticsearch. Its goal is to provide common ground for all Elasticsearch-related code in Python; because of this it tries to be opinion-free and very extendable.

For a more high level client library with more limited scope, have a look at elasticsearch-dsl - it is a more pythonic library sitting on top of `elasticsearch-py`.

CHAPTER 1

Compatibility

The library is compatible with all Elasticsearch versions since `0.90.x` but you **have to use a matching major version**:

For **Elasticsearch 6.0** and later, use the major version 6 (`6.x.y`) of the library.

For **Elasticsearch 5.0** and later, use the major version 5 (`5.x.y`) of the library.

For **Elasticsearch 2.0** and later, use the major version 2 (`2.x.y`) of the library, and so on.

The recommended way to set your requirements in your *setup.py* or *requirements.txt* is:

```
# Elasticsearch 6.x
elasticsearch>=6.0.0,<7.0.0

# Elasticsearch 5.x
elasticsearch>=5.0.0,<6.0.0

# Elasticsearch 2.x
elasticsearch>=2.0.0,<3.0.0
```

If you have a need to have multiple versions installed at the same time older versions are also released as `elasticsearch2` and `elasticsearch5`.

CHAPTER 2

Installation

Install the `elasticsearch` package with pip:

```
pip install elasticsearch
```

CHAPTER 3

Example Usage

```
from datetime import datetime
from elasticsearch import Elasticsearch
es = Elasticsearch()

doc = {
    'author': 'kimchy',
    'text': 'Elasticsearch: cool. bonsai cool.',
    'timestamp': datetime.now(),
}
res = es.index(index="test-index", doc_type='tweet', id=1, body=doc)
print(res['result'])

res = es.get(index="test-index", doc_type='tweet', id=1)
print(res['_source'])

es.indices.refresh(index="test-index")

res = es.search(index="test-index", body={"query": {"match_all": {}}})
print("Got %d Hits:" % res['hits']['total'])
for hit in res['hits']['hits']:
    print("%(timestamp)s %(author)s: %(text)s" % hit["_source"])
```

CHAPTER 4

Features

This client was designed as very thin wrapper around Elasticsearch's REST API to allow for maximum flexibility. This means that there are no opinions in this client; it also means that some of the APIs are a little cumbersome to use from Python. We have created some *Helpers* to help with this issue as well as a more high level library (elasticsearch-dsl) on top of this one to provide a more convenient way of working with Elasticsearch.

4.1 Persistent Connections

`elasticsearch-py` uses persistent connections inside of individual connection pools (one per each configured or sniffed node). Out of the box you can choose between two `http` protocol implementations. See *Transport classes* for more information.

The transport layer will create an instance of the selected connection class per node and keep track of the health of individual nodes - if a node becomes unresponsive (throwing exceptions while connecting to it) it's put on a timeout by the *`ConnectionPool`* class and only returned to the circulation after the timeout is over (or when no live nodes are left). By default nodes are randomized before being passed into the pool and round-robin strategy is used for load balancing.

You can customize this behavior by passing parameters to the *Connection Layer API* (all keyword arguments to the *`Elasticsearch`* class will be passed through). If what you want to accomplish is not supported you should be able to create a subclass of the relevant component and pass it in as a parameter to be used instead of the default implementation.

4.2 Automatic Retries

If a connection to a node fails due to connection issues (raises *`ConnectionError`*) it is considered in faulty state. It will be placed on hold for `dead_timeout` seconds and the request will be retried on another node. If a connection fails multiple times in a row the timeout will get progressively larger to avoid hitting a node that's, by all indication, down. If no live connection is available, the connection that has the smallest timeout will be used.

By default retries are not triggered by a timeout (*ConnectionTimeout*), set `retry_on_timeout` to `True` to also retry on timeouts.

4.3 Sniffing

The client can be configured to inspect the cluster state to get a list of nodes upon startup, periodically and/or on failure. See *Transport* parameters for details.

Some example configurations:

```
from elasticsearch import Elasticsearch

# by default we don't sniff, ever
es = Elasticsearch()

# you can specify to sniff on startup to inspect the cluster and load
# balance across all nodes
es = Elasticsearch(["seed1", "seed2"], sniff_on_start=True)

# you can also sniff periodically and/or after failure:
es = Elasticsearch(["seed1", "seed2"],
          sniff_on_start=True,
          sniff_on_connection_fail=True,
          sniffer_timeout=60)
```

4.4 Thread safety

The client is thread safe and can be used in a multi threaded environment. Best practice is to create a single global instance of the client and use it throughout your application. If your application is long-running consider turning on *Sniffing* to make sure the client is up to date on the cluster location.

By default we allow `urllib3` to open up to 10 connections to each node, if your application calls for more parallelism, use the `maxsize` parameter to raise the limit:

```
# allow up to 25 connections to each node
es = Elasticsearch(["host1", "host2"], maxsize=25)
```

Note: Since we use persistent connections throughout the client it means that the client doesn't tolerate `fork` very well. If your application calls for multiple processes make sure you create a fresh client after call to `fork`. Note that Python's `multiprocessing` module uses `fork` to create new processes on POSIX systems.

4.5 SSL and Authentication

You can configure the client to use `SSL` for connecting to your elasticsearch cluster, including certificate verification and HTTP auth:

```
from elasticsearch import Elasticsearch

# you can use RFC-1738 to specify the url
```

(continues on next page)

(continued from previous page)

```
es = Elasticsearch(['https://user:secret@localhost:443'])

# ... or specify common parameters as kwargs

es = Elasticsearch(
    ['localhost', 'otherhost'],
    http_auth=('user', 'secret'),
    scheme="https",
    port=443,
)

# SSL client authentication using client_cert and client_key

from ssl import create_default_context

context = create_default_context(cafile="path/to/cert.pem")
es = Elasticsearch(
    ['localhost', 'otherhost'],
    http_auth=('user', 'secret'),
    scheme="https",
    port=443,
    ssl_context=context,
)
```

Warning: `elasticsearch-py` doesn't ship with default set of root certificates. To have working SSL certificate validation you need to either specify your own as `cafile` or `capath` or `cadata` or install certifi which will be picked up automatically.

See class *`Urllib3HttpConnection`* for detailed description of the options.

4.6 Logging

`elasticsearch-py` uses the standard logging library from python to define two loggers: `elasticsearch` and `elasticsearch.trace`. `elasticsearch` is used by the client to log standard activity, depending on the log level. `elasticsearch.trace` can be used to log requests to the server in the form of `curl` commands using pretty-printed json that can then be executed from command line. Because it is designed to be shared (for example to demonstrate an issue) it also just uses `localhost:9200` as the address instead of the actual address of the host. If the trace logger has not been configured already it is set to *propagate=False* so it needs to be activated separately.

CHAPTER 5

Environment considerations

When using the client there are several limitations of your environment that could come into play.

When using an HTTP load balancer you cannot use the *Sniffing* functionality - the cluster would supply the client with IP addresses to directly connect to the cluster, circumventing the load balancer. Depending on your configuration this might be something you don't want or break completely.

In some environments (notably on Google App Engine) your HTTP requests might be restricted so that `GET` requests won't accept body. In that case use the `send_get_body_as` parameter of *`Transport`* to send all bodies via post:

```
from elasticsearch import Elasticsearch
es = Elasticsearch(send_get_body_as='POST')
```

5.1 Compression

When using capacity-constrained networks (low throughput), it may be handy to enable compression. This is especially useful when doing bulk loads or inserting large documents. This will configure compression on the *request*.

```
from elasticsearch import Elasticsearch
es = Elasticsearch(hosts, http_compress=True)
```

5.2 Running on AWS with IAM

If you want to use this client with IAM based authentication on AWS you can use the requests-aws4auth package:

```
from elasticsearch import Elasticsearch, RequestsHttpConnection
from requests_aws4auth import AWS4Auth

host = 'YOURHOST.us-east-1.es.amazonaws.com'
awsauth = AWS4Auth(YOUR_ACCESS_KEY, YOUR_SECRET_KEY, REGION, 'es')
```

(continues on next page)

(continued from previous page)

```
es = Elasticsearch(
    hosts=[{'host': host, 'port': 443}],
    http_auth=awsauth,
    use_ssl=True,
    verify_certs=True,
    connection_class=RequestsHttpConnection
)
print(es.info())
```

CHAPTER 6

Customization

6.1 Custom serializers

By default, JSONSerializer is used to encode all outgoing requests. However, you can implement your own custom serializer:

```
from elasticsearch.serializer import JSONSerializer

class SetEncoder(JSONSerializer):
    def default(self, obj):
        if isinstance(obj, set):
            return list(obj)
        if isinstance(obj, Something):
            return 'CustomSomethingRepresentation'
        return JSONSerializer.default(self, obj)

es = Elasticsearch(serializer=SetEncoder())
```

CHAPTER 7

Contents

7.1 API Documentation

All the API calls map the raw REST api as closely as possible, including the distinction between required and optional arguments to the calls. This means that the code makes distinction between positional and keyword arguments; we, however, recommend that people **use keyword arguments for all calls for consistency and safety**.

Note: for compatibility with the Python ecosystem we use `from_` instead of `from` and `doc_type` instead of `type` as parameter names.

7.1.1 Global options

Some parameters are added by the client itself and can be used in all API calls.

Ignore

An API call is considered successful (and will return a response) if elasticsearch returns a 2XX response. Otherwise an instance of *`TransportError`* (or a more specific subclass) will be raised. You can see other exception and error states in *Exceptions*. If you do not wish an exception to be raised you can always pass in an `ignore` parameter with either a single status code that should be ignored or a list of them:

```
from elasticsearch import Elasticsearch
es = Elasticsearch()

# ignore 400 cause by IndexAlreadyExistsException when creating an index
es.indices.create(index='test-index', ignore=400)

# ignore 404 and 400
es.indices.delete(index='test-index', ignore=[400, 404])
```

Timeout

Global timeout can be set when constructing the client (see `Connection`'s `timeout` parameter) or on a per-request basis using `request_timeout` (float value in seconds) as part of any API call, this value will get passed to the `perform_request` method of the connection class:

```
# only wait for 1 second, regardless of the client's default
es.cluster.health(wait_for_status='yellow', request_timeout=1)
```

Note: Some API calls also accept a `timeout` parameter that is passed to Elasticsearch server. This timeout is internal and doesn't guarantee that the request will end in the specified time.

Response Filtering

The `filter_path` parameter is used to reduce the response returned by elasticsearch. For example, to only return `_id` and `_type`, do:

```
es.search(index='test-index', filter_path=['hits.hits._id', 'hits.hits._type'])
```

It also supports the `*` wildcard character to match any field or part of a field's name:

```
es.search(index='test-index', filter_path=['hits.hits._*'])
```

7.1.2 Elasticsearch

class elasticsearch.**Elasticsearch**(*hosts=None*, *transport_class=<class 'elasticsearch.transport.Transport'>*, ***kwargs*)

Elasticsearch low-level client. Provides a straightforward mapping from Python to ES REST endpoints.

The instance has attributes `cat`, `cluster`, `indices`, `ingest`, `nodes`, `snapshot` and `tasks` that provide access to instances of *CatClient*, *ClusterClient*, *IndicesClient*, *IngestClient*, *NodesClient*, *SnapshotClient* and *TasksClient* respectively. This is the preferred (and only supported) way to get access to those classes and their methods.

You can specify your own connection class which should be used by providing the `connection_class` parameter:

```
# create connection to localhost using the ThriftConnection
es = Elasticsearch(connection_class=ThriftConnection)
```

If you want to turn on *Sniffing* you have several options (described in *Transport*):

```
# create connection that will automatically inspect the cluster to get
# the list of active nodes. Start with nodes running on 'esnode1' and
# 'esnode2'
es = Elasticsearch(
    ['esnode1', 'esnode2'],
    # sniff before doing anything
    sniff_on_start=True,
    # refresh nodes after a node fails to respond
    sniff_on_connection_fail=True,
    # and also every 60 seconds
```

(continues on next page)

(continued from previous page)

```
    sniffer_timeout=60
)
```

Different hosts can have different parameters, use a dictionary per node to specify those:

```
# connect to localhost directly and another node using SSL on port 443
# and an url_prefix. Note that ``port`` needs to be an int.
es = Elasticsearch([
    {'host': 'localhost'},
    {'host': 'othernode', 'port': 443, 'url_prefix': 'es', 'use_ssl': True},
])
```

If using SSL, there are several parameters that control how we deal with certificates (see *Urllib3HttpConnection* for detailed description of the options):

```
es = Elasticsearch(
    ['localhost:443', 'other_host:443'],
    # turn on SSL
    use_ssl=True,
    # make sure we verify SSL certificates
    verify_certs=True,
    # provide a path to CA certs on disk
    ca_certs='/path/to/CA_certs'
)
```

If using SSL, but don't verify the certs, a warning message is showed optionally (see *Urllib3HttpConnection* for detailed description of the options):

```
es = Elasticsearch(
    ['localhost:443', 'other_host:443'],
    # turn on SSL
    use_ssl=True,
    # no verify SSL certificates
    verify_certs=False,
    # don't show warnings about ssl certs verification
    ssl_show_warn=False
)
```

SSL client authentication is supported (see *Urllib3HttpConnection* for detailed description of the options):

```
es = Elasticsearch(
    ['localhost:443', 'other_host:443'],
    # turn on SSL
    use_ssl=True,
    # make sure we verify SSL certificates
    verify_certs=True,
    # provide a path to CA certs on disk
    ca_certs='/path/to/CA_certs',
    # PEM formatted SSL client certificate
    client_cert='/path/to/clientcert.pem',
    # PEM formatted SSL client key
    client_key='/path/to/clientkey.pem'
)
```

Alternatively you can use RFC-1738 formatted URLs, as long as they are not in conflict with other options:

```
es = Elasticsearch(
    [
        'http://user:secret@localhost:9200/',
        'https://user:secret@other_host:443/production'
    ],
    verify_certs=True
)
```

By default, JSONSerializer is used to encode all outgoing requests. However, you can implement your own custom serializer:

```
from elasticsearch.serializer import JSONSerializer

class SetEncoder(JSONSerializer):
    def default(self, obj):
        if isinstance(obj, set):
            return list(obj)
        if isinstance(obj, Something):
            return 'CustomSomethingRepresentation'
        return JSONSerializer.default(self, obj)

es = Elasticsearch(serializer=SetEncoder())
```

Parameters

- **hosts** – list of nodes we should connect to. Node should be a dictionary ({“host”: “localhost”, “port”: 9200}), the entire dictionary will be passed to the `Connection` class as kwargs, or a string in the format of `host[:port]` which will be translated to a dictionary automatically. If no value is given the *`Urllib3HttpConnection`* class defaults will be used.
- **transport_class** – *`Transport`* subclass to use.
- **kwargs** – any additional arguments will be passed on to the *`Transport`* class and, subsequently, to the `Connection` instances.

bulk(***kwargs*)

Perform many index/delete operations in a single API call.

See the *`bulk()`* helper function for a more friendly API. http://www.elastic.co/guide/en/elasticsearch/reference/current/docs-bulk.html

Parameters

- **body** – The operation definition and data (action-data pairs), separated by newlines
- **index** – Default index for items which don’t provide one
- **_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(***kwargs*)

Clear the scroll request created by specifying the scroll parameter to search. http://www.elastic.co/guide/en/elasticsearch/reference/current/search-request-scroll.html

Parameters

- **scroll_id** – A comma-separated list of scroll IDs to clear
- **body** – A comma-separated list of scroll IDs to clear if none was specified via the scroll_id parameter

count(***kwargs*)

Execute a query and get the number of matches for that query. http://www.elastic.co/guide/en/elasticsearch/reference/current/search-count.html

Parameters

- **index** – A list of index names or a string containing a comma-separated list of index names to restrict the results to
- **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
- **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

create(***kwargs*)
Adds a typed JSON document in a specific index, making it searchable. Behind the scenes this method calls index(..., op_type='create') http://www.elastic.co/guide/en/elasticsearch/reference/current/docs-index_.html

Parameters

- **index** – The name of the index
- **id** – Document ID
- **body** – The document
- **parent** – ID of the parent document
- **pipeline** – The pipeline id to preprocess incoming documents with
- **refresh** – If *true* then refresh the affected 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
- **timestamp** – Explicit timestamp for the document
- **ttl** – Expiration time for the document
- **version** – Explicit version number for concurrency control
- **version_type** – Specific version type, valid choices are: 'internal', 'external', 'external_gte', 'force'
- **wait_for_active_shards** – Sets the number of shard copies that must be active before proceeding with the index 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)

delete(***kwargs*)
Delete a typed JSON document from a specific index based on its id. http://www.elastic.co/guide/en/elasticsearch/reference/current/docs-delete.html

Parameters

- **index** – The name of the index
- **id** – The document ID
- **parent** – ID of parent document
- **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
- **version** – Explicit version number for concurrency control
- **version_type** – Specific version type, valid choices are: 'internal', 'external', 'external_gte', 'force'
- **wait_for_active_shards** – Sets the number of shard copies that must be active before proceeding with the delete 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)

delete_by_query(***kwargs*)

Delete all documents matching a query. https://www.elastic.co/guide/en/elasticsearch/reference/current/docs-delete-by-query.html

Parameters

- **index** – A list of index names to search, or a string containing a comma-separated list of index names to search; use *_all* or the empty string to perform the operation on all indices
- **body** – The search definition using the Query DSL
- **_source** – True or false to return the _source field or not, or a list of fields to return
- **_source_exclude** – A list of fields to exclude from the returned _source field
- **_source_include** – A list of fields to extract and return from the _source field
- **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
- **conflicts** – What to do when the delete-by-query hits version conflicts?, default 'abort', valid choices are: 'abort', 'proceed'
- **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'
- **from_** – Starting offset (default: 0)
- **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
- **preference** – Specify the node or shard the operation should be performed on (default: random)
- **q** – Query in the Lucene query string syntax
- **refresh** – Should the effected indexes be refreshed?
- **request_cache** – Specify if request cache should be used for this request or not, defaults to index level setting
- **requests_per_second** – The throttle for this request in sub-requests per second. -1 means no throttle., default 0
- **routing** – A comma-separated list of specific routing values
- **scroll** – Specify how long a consistent view of the index should be maintained for scrolled search
- **scroll_size** – Size on the scroll request powering the update_by_query

- **search_timeout** – Explicit timeout for each search request. Defaults to no timeout.
- **search_type** – Search operation type, valid choices are: 'query_then_fetch', 'dfs_query_then_fetch'
- **size** – Number of hits to return (default: 10)
- **slices** – The number of slices this task should be divided into. Defaults to 1 meaning the task isn't sliced into subtasks., default 1
- **sort** – A comma-separated list of <field>:<direction> pairs
- **stats** – Specific 'tag' of the request for logging and statistical purposes
- **terminate_after** – The maximum number of documents to collect for each shard, upon reaching which the query execution will terminate early.
- **timeout** – Time each individual bulk request should wait for shards that are unavailable., default '1m'
- **version** – Specify whether to return document version as part of a hit
- **wait_for_active_shards** – Sets the number of shard copies that must be active before proceeding with the delete by query 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)
- **wait_for_completion** – Should the request should block until the delete-by-query is complete., default True

delete_script(***kwargs*)

Remove a stored script from elasticsearch. http://www.elastic.co/guide/en/elasticsearch/reference/master/modules-scripting.html

Parameters **id** – Script ID

exists(***kwargs*)

Returns a boolean indicating whether or not given document exists in Elasticsearch. http://www.elastic.co/guide/en/elasticsearch/reference/current/docs-get.html

Parameters

- **index** – The name of the index
- **id** – The document ID
- **_source** – True or false to return the _source field or not, or a list of fields to return
- **_source_exclude** – A list of fields to exclude from the returned _source field
- **_source_include** – A list of fields to extract and return from the _source field
- **parent** – The ID of the parent document
- **preference** – Specify the node or shard the operation should be performed on (default: random)
- **realtime** – Specify whether to perform the operation in realtime or search mode
- **refresh** – Refresh the shard containing the document before performing the operation
- **routing** – Specific routing value
- **stored_fields** – A comma-separated list of stored fields to return in the response
- **version** – Explicit version number for concurrency control

- **version_type** – Specific version type, valid choices are: 'internal', 'external', 'external_gte', 'force'

exists_source(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/master/docs-get.html

Parameters

- **index** – The name of the index
- **id** – The document ID
- **_source** – True or false to return the _source field or not, or a list of fields to return
- **_source_exclude** – A list of fields to exclude from the returned _source field
- **_source_include** – A list of fields to extract and return from the _source field
- **parent** – The ID of the parent document
- **preference** – Specify the node or shard the operation should be performed on (default: random)
- **realtime** – Specify whether to perform the operation in realtime or search mode
- **refresh** – Refresh the shard containing the document before performing the operation
- **routing** – Specific routing value
- **version** – Explicit version number for concurrency control
- **version_type** – Specific version type, valid choices are: 'internal', 'external', 'external_gte', 'force'

explain(***kwargs*)
The explain api computes a score explanation for a query and a specific document. This can give useful feedback whether a document matches or didn't match a specific query. http://www.elastic.co/guide/en/elasticsearch/reference/current/search-explain.html

Parameters

- **index** – The name of the index
- **id** – The document ID
- **body** – The query definition using the Query DSL
- **_source** – True or false to return the _source field or not, or a list of fields to return
- **_source_exclude** – A list of fields to exclude from the returned _source field
- **_source_include** – A list of fields to extract and return from the _source field
- **analyze_wildcard** – Specify whether wildcards and prefix queries in the query string query should be analyzed (default: false)
- **analyzer** – The analyzer for the query string query
- **default_operator** – The default operator for query string query (AND or OR), default 'OR', valid choices are: 'AND', 'OR'
- **df** – The default field for query string query (default: _all)
- **lenient** – Specify whether format-based query failures (such as providing text to a numeric field) should be ignored
- **parent** – The ID of the parent document

- **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
- **stored_fields** – A comma-separated list of stored fields to return in the response

field_caps(***kwargs*)
The field capabilities API allows to retrieve the capabilities of fields among multiple indices. http://www.elastic.co/guide/en/elasticsearch/reference/current/search-field-caps.html

Parameters

- **index** – A list of index names, or a string containing a comma-separated list of index names; use *_all* or the empty string to perform the operation on all indices
- **body** – Field json objects containing an array of field names
- **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)
- **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'
- **fields** – A comma-separated list of field names
- **ignore_unavailable** – Whether specified concrete indices should be ignored when unavailable (missing or closed)

get(***kwargs*)
Get a typed JSON document from the index based on its id. http://www.elastic.co/guide/en/elasticsearch/reference/current/docs-get.html

Parameters

- **index** – The name of the index
- **id** – The document ID
- **_source** – True or false to return the _source field or not, or a list of fields to return
- **_source_exclude** – A list of fields to exclude from the returned _source field
- **_source_include** – A list of fields to extract and return from the _source field
- **parent** – The ID of the parent document
- **preference** – Specify the node or shard the operation should be performed on (default: random)
- **realtime** – Specify whether to perform the operation in realtime or search mode
- **refresh** – Refresh the shard containing the document before performing the operation
- **routing** – Specific routing value
- **stored_fields** – A comma-separated list of stored fields to return in the response
- **version** – Explicit version number for concurrency control
- **version_type** – Specific version type, valid choices are: 'internal', 'external', 'external_gte', 'force'

get_script (***kwargs*)
Retrieve a script from the API. http://www.elastic.co/guide/en/elasticsearch/reference/master/modules-scripting.html

Parameters **id** – Script ID

get_source (***kwargs*)
Get the source of a document by it's index, type and id. http://www.elastic.co/guide/en/elasticsearch/reference/current/docs-get.html

Parameters

- **index** – The name of the index
- **id** – The document ID
- **_source** – True or false to return the _source field or not, or a list of fields to return
- **_source_exclude** – A list of fields to exclude from the returned _source field
- **_source_include** – A list of fields to extract and return from the _source field
- **parent** – The ID of the parent document
- **preference** – Specify the node or shard the operation should be performed on (default: random)
- **realtime** – Specify whether to perform the operation in realtime or search mode
- **refresh** – Refresh the shard containing the document before performing the operation
- **routing** – Specific routing value
- **version** – Explicit version number for concurrency control
- **version_type** – Specific version type, valid choices are: 'internal', 'external', 'external_gte', 'force'

index (***kwargs*)
Adds or updates a typed JSON document in a specific index, making it searchable. http://www.elastic.co/guide/en/elasticsearch/reference/current/docs-index_.html

Parameters

- **index** – The name of the index
- **body** – The document
- **id** – Document ID
- **doc_type** – Document type, defaults to *_doc*. Not used on ES 7 clusters.
- **op_type** – Explicit operation type, default 'index', valid choices are: 'index', 'create'
- **parent** – ID of the parent document
- **pipeline** – The pipeline id to preprocess incoming documents with
- **refresh** – If *true* then refresh the affected 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
- **timestamp** – Explicit timestamp for the document

- **ttl** – Expiration time for the document
- **version** – Explicit version number for concurrency control
- **version_type** – Specific version type, valid choices are: 'internal', 'external', 'external_gte', 'force'
- **wait_for_active_shards** – Sets the number of shard copies that must be active before proceeding with the index 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)

info(***kwargs*)
Get the basic info from the current cluster. http://www.elastic.co/guide/

mget(***kwargs*)
Get multiple documents based on an index, type (optional) and ids. http://www.elastic.co/guide/en/elasticsearch/reference/current/docs-multi-get.html

Parameters

- **body** – Document identifiers; can be either *docs* (containing full document information) or *ids* (when index and type is provided in the URL.
- **index** – The name of the index
- **_source** – True or false to return the _source field or not, or a list of fields to return
- **_source_exclude** – A list of fields to exclude from the returned _source field
- **_source_include** – A list of fields to extract and return from the _source field
- **preference** – Specify the node or shard the operation should be performed on (default: random)
- **realtime** – Specify whether to perform the operation in realtime or search mode
- **refresh** – Refresh the shard containing the document before performing the operation
- **routing** – Specific routing value
- **stored_fields** – A comma-separated list of stored fields to return in the response

msearch(***kwargs*)
Execute several search requests within the same API. http://www.elastic.co/guide/en/elasticsearch/reference/current/search-multi-search.html

Parameters

- **body** – The request definitions (metadata-search request definition pairs), separated by newlines
- **index** – A list of index names, or a string containing a comma-separated list of index names, to use as the default
- **max_concurrent_searches** – Controls the maximum number of concurrent searches the multi search api will execute
- **pre_filter_shard_size** – A threshold that enforces a pre-filter roundtrip to pre-filter search shards based on query rewriting if the number of shards the search request expands to exceeds the threshold. This filter roundtrip can limit the number of shards significantly if for instance a shard can not match any documents based on it's rewrite method ie. if date filters are mandatory to match but the shard bounds and the query are disjoint., default 128

- **search_type** – Search operation type, valid choices are: 'query_then_fetch', 'query_and_fetch', 'dfs_query_then_fetch', 'dfs_query_and_fetch'
- **typed_keys** – Specify whether aggregation and suggester names should be prefixed by their respective types in the response

msearch_template(***kwargs*)

The /_search/template endpoint allows to use the mustache language to pre render search requests, before they are executed and fill existing templates with template parameters. http://www.elastic.co/guide/en/elasticsearch/reference/current/search-template.html

Parameters

- **body** – The request definitions (metadata-search request definition pairs), separated by newlines
- **index** – A list of index names, or a string containing a comma-separated list of index names, to use as the default
- **max_concurrent_searches** – Controls the maximum number of concurrent searches the multi search api will execute
- **search_type** – Search operation type, valid choices are: 'query_then_fetch', 'query_and_fetch', 'dfs_query_then_fetch', 'dfs_query_and_fetch'
- **typed_keys** – Specify whether aggregation and suggester names should be prefixed by their respective types in the response

mtermvectors(***kwargs*)

Multi termvectors API allows to get multiple termvectors based on an index, type and id. http://www.elastic.co/guide/en/elasticsearch/reference/current/docs-multi-termvectors.html

Parameters

- **index** – The index in which the document resides.
- **body** – Define ids, documents, parameters or a list of parameters per document here. You must at least provide a list of document ids. See documentation.
- **field_statistics** – Specifies if document count, sum of document frequencies and sum of total term frequencies should be returned. Applies to all returned documents unless otherwise specified in body "params" or "docs"., default True
- **fields** – A comma-separated list of fields to return. Applies to all returned documents unless otherwise specified in body "params" or "docs".
- **ids** – A comma-separated list of documents ids. You must define ids as parameter or set "ids" or "docs" in the request body
- **offsets** – Specifies if term offsets should be returned. Applies to all returned documents unless otherwise specified in body "params" or "docs"., default True
- **parent** – Parent id of documents. Applies to all returned documents unless otherwise specified in body "params" or "docs".
- **payloads** – Specifies if term payloads should be returned. Applies to all returned documents unless otherwise specified in body "params" or "docs"., default True
- **positions** – Specifies if term positions should be returned. Applies to all returned documents unless otherwise specified in body "params" or "docs"., default True
- **preference** – Specify the node or shard the operation should be performed on (default: random) .Applies to all returned documents unless otherwise specified in body "params" or "docs".

- **realtime** – Specifies if requests are real-time as opposed to near- real-time (default: true).
- **routing** – Specific routing value. Applies to all returned documents unless otherwise specified in body "params" or "docs".
- **term_statistics** – Specifies if total term frequency and document frequency should be returned. Applies to all returned documents unless otherwise specified in body "params" or "docs"., default False
- **version** – Explicit version number for concurrency control
- **version_type** – Specific version type, valid choices are: 'internal', 'external', 'external_gte', 'force'

ping(***kwargs*)
Returns True if the cluster is up, False otherwise. http://www.elastic.co/guide/

put_script(***kwargs*)
Create a script in given language with specified ID. http://www.elastic.co/guide/en/elasticsearch/reference/current/modules-scripting.html

Parameters

- **id** – Script ID
- **body** – The document

reindex(***kwargs*)
Reindex all documents from one index to another. https://www.elastic.co/guide/en/elasticsearch/reference/current/docs-reindex.html

Parameters

- **body** – The search definition using the Query DSL and the prototype for the index request.
- **refresh** – Should the effected indexes be refreshed?
- **requests_per_second** – The throttle to set on this request in sub- requests per second. -1 means no throttle., default 0
- **slices** – The number of slices this task should be divided into. Defaults to 1 meaning the task isn't sliced into subtasks., default 1
- **timeout** – Time each individual bulk request should wait for shards that are unavailable., default '1m'
- **wait_for_active_shards** – Sets the number of shard copies that must be active before proceeding with the reindex 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)
- **wait_for_completion** – Should the request should block until the reindex is complete., default True

reindex_rethrottle(***kwargs*)
Change the value of `requests_per_second` of a running `reindex` task. https://www.elastic.co/guide/en/elasticsearch/reference/current/docs-reindex.html

Parameters

- **task_id** – The task id to rethrottle
- **requests_per_second** – The throttle to set on this request in floating sub-requests per second. -1 means set no throttle.

render_search_template(***kwargs*)
http://www.elasticsearch.org/guide/en/elasticsearch/reference/current/search-template.html

Parameters

- **id** – The id of the stored search template
- **body** – The search definition template and its params

scroll(***kwargs*)
Scroll a search request created by specifying the scroll parameter. http://www.elastic.co/guide/en/elasticsearch/reference/current/search-request-scroll.html

Parameters

- **scroll_id** – The scroll ID
- **body** – The scroll ID if not passed by URL or query parameter.
- **scroll** – Specify how long a consistent view of the index should be maintained for scrolled search
- **rest_total_hits_as_int** – This parameter is used to restore the total hits as a number in the response. This param is added version 6.x to handle mixed cluster queries where nodes are in multiple versions (7.0 and 6.latest)

search(***kwargs*)
Execute a search query and get back search hits that match the query. http://www.elastic.co/guide/en/elasticsearch/reference/current/search-search.html

Parameters

- **index** – A list of index names to search, or a string containing a comma-separated list of index names to search; use *_all* or empty string to perform the operation on all indices
- **body** – The search definition using the Query DSL
- **_source** – True or false to return the _source field or not, or a list of fields to return
- **_source_exclude** – A list of fields to exclude from the returned _source field
- **_source_include** – A list of fields to extract and return from the _source field
- **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)
- **allow_partial_search_results** – Set to false to return an overall failure if the request would produce partial results. Defaults to True, which will allow partial results in the case of timeouts or partial failures
- **analyze_wildcard** – Specify whether wildcard and prefix queries should be analyzed (default: false)
- **analyzer** – The analyzer to use for the query string
- **batched_reduce_size** – The number of shard results that should be reduced at once on the coordinating node. This value should be used as a protection mechanism to reduce the memory overhead per search request if the potential number of shards in the request can be large., default 512
- **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

- **docvalue_fields** – A comma-separated list of fields to return as the docvalue representation of a field for each hit
- **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'
- **explain** – Specify whether to return detailed information about score computation as part of a hit
- **from_** – Starting offset (default: 0)
- **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
- **max_concurrent_shard_requests** – The number of concurrent shard requests this search executes concurrently. This value should be used to limit the impact of the search on the cluster in order to limit the number of concurrent shard requests, default 'The default grows with the number of nodes in the cluster but is at most 256.'
- **pre_filter_shard_size** – A threshold that enforces a pre-filter roundtrip to prefilter search shards based on query rewriting if the number of shards the search request expands to exceeds the threshold. This filter roundtrip can limit the number of shards significantly if for instance a shard can not match any documents based on it's rewrite method ie. if date filters are mandatory to match but the shard bounds and the query are disjoint., default 128
- **preference** – Specify the node or shard the operation should be performed on (default: random)
- **q** – Query in the Lucene query string syntax
- **rest_total_hits_as_int** – This parameter is used to restore the total hits as a number in the response. This param is added version 6.x to handle mixed cluster queries where nodes are in multiple versions (7.0 and 6.latest)
- **request_cache** – Specify if request cache should be used for this request or not, defaults to index level setting
- **routing** – A comma-separated list of specific routing values
- **scroll** – Specify how long a consistent view of the index should be maintained for scrolled search
- **search_type** – Search operation type, valid choices are: 'query_then_fetch', 'dfs_query_then_fetch'
- **size** – Number of hits to return (default: 10)
- **sort** – A comma-separated list of <field>:<direction> pairs
- **stats** – Specific 'tag' of the request for logging and statistical purposes
- **stored_fields** – A comma-separated list of stored fields to return as part of a hit
- **suggest_field** – Specify which field to use for suggestions
- **suggest_mode** – Specify suggest mode, default 'missing', valid choices are: 'missing', 'popular', 'always'
- **suggest_size** – How many suggestions to return in response
- **suggest_text** – The source text for which the suggestions should be returned

- **terminate_after** – The maximum number of documents to collect for each shard, upon reaching which the query execution will terminate early.
- **timeout** – Explicit operation timeout
- **track_scores** – Whether to calculate and return scores even if they are not used for sorting
- **track_total_hits** – Indicate if the number of documents that match the query should be tracked
- **typed_keys** – Specify whether aggregation and suggester names should be prefixed by their respective types in the response
- **version** – Specify whether to return document version as part of a hit

search_shards(***kwargs*)

The search shards api returns the indices and shards that a search request would be executed against. This can give useful feedback for working out issues or planning optimizations with routing and shard preferences. http://www.elastic.co/guide/en/elasticsearch/reference/current/search-shards.html

Parameters

- **index** – A list of index names to search, or a string containing a comma-separated list of index names to search; use *_all* or the empty string to perform the operation on all indices
- **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)
- **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)
- **local** – Return local information, do not retrieve the state from master node (default: false)
- **preference** – Specify the node or shard the operation should be performed on (default: random)
- **routing** – Specific routing value

search_template(***kwargs*)

A query that accepts a query template and a map of key/value pairs to fill in template parameters. http://www.elastic.co/guide/en/elasticsearch/reference/current/search-template.html

Parameters

- **index** – A list of index names to search, or a string containing a comma-separated list of index names to search; use *_all* or the empty string to perform the operation on all indices
- **body** – The search definition template and its params
- **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)
- **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'
- **explain** – Specify whether to return detailed information about score computation as part of a hit
- **ignore_unavailable** – Whether specified concrete indices should be ignored when unavailable (missing or closed)

- **preference** – Specify the node or shard the operation should be performed on (default: random)
- **profile** – Specify whether to profile the query execution
- **routing** – A comma-separated list of specific routing values
- **scroll** – Specify how long a consistent view of the index should be maintained for scrolled search
- **search_type** – Search operation type, valid choices are: 'query_then_fetch', 'query_and_fetch', 'dfs_query_then_fetch', 'dfs_query_and_fetch'
- **typed_keys** – Specify whether aggregation and suggester names should be prefixed by their respective types in the response

termvectors(***kwargs*)

Returns information and statistics on terms in the fields of a particular document. The document could be stored in the index or artificially provided by the user (Added in 1.4). Note that for documents stored in the index, this is a near realtime API as the term vectors are not available until the next refresh. http://www.elastic.co/guide/en/elasticsearch/reference/current/docs-termvectors.html

Parameters

- **index** – The index in which the document resides.
- **id** – The id of the document, when not specified a doc param should be supplied.
- **body** – Define parameters and or supply a document to get termvectors for. See documentation.
- **field_statistics** – Specifies if document count, sum of document frequencies and sum of total term frequencies should be returned., default True
- **fields** – A comma-separated list of fields to return.
- **offsets** – Specifies if term offsets should be returned., default True
- **parent** – Parent id of documents.
- **payloads** – Specifies if term payloads should be returned., default True
- **positions** – Specifies if term positions should be returned., default True
- **preference** – Specify the node or shard the operation should be performed on (default: random).
- **realtime** – Specifies if request is real-time as opposed to near- real-time (default: true).
- **routing** – Specific routing value.
- **term_statistics** – Specifies if total term frequency and document frequency should be returned., default False
- **version** – Explicit version number for concurrency control
- **version_type** – Specific version type, valid choices are: 'internal', 'external', 'external_gte', 'force'

update(***kwargs*)

Update a document based on a script or partial data provided. http://www.elastic.co/guide/en/elasticsearch/reference/current/docs-update.html

Parameters

- **index** – The name of the index

- **id** – Document ID
- **body** – The request definition using either *script* or partial *doc*
- **_source** – True or false to return the _source field or not, or a list of fields to return
- **_source_exclude** – A list of fields to exclude from the returned _source field
- **_source_include** – A list of fields to extract and return from the _source field
- **fields** – A comma-separated list of fields to return in the response
- **if_seq_no** –
- **if_primary_term** –
- **lang** – The script language (default: painless)
- **parent** – ID of the parent document. Is is only used for routing and when for the upsert request
- **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_forarg retry_on_conflict: Specify how many times should the operation be retried when a conflict occurs (default: 0)
- **routing** – Specific routing value
- **timeout** – Explicit operation timeout
- **timestamp** – Explicit timestamp for the document
- **ttl** – Expiration time for the document
- **version** – Explicit version number for concurrency control
- **version_type** – Specific version type, valid choices are: 'internal', 'force'
- **wait_for_active_shards** – Sets the number of shard copies that must be active before proceeding with the update 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)

update_by_query(***kwargs*)

Perform an update on all documents matching a query. https://www.elastic.co/guide/en/elasticsearch/reference/current/docs-update-by-query.html

Parameters

- **index** – A list of index names to search, or a string containing a comma-separated list of index names to search; use *_all* or the empty string to perform the operation on all indices
- **body** – The search definition using the Query DSL
- **_source** – True or false to return the _source field or not, or a list of fields to return
- **_source_exclude** – A list of fields to exclude from the returned _source field
- **_source_include** – A list of fields to extract and return from the _source field
- **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
- **conflicts** – What to do when the update by query hits version conflicts?, default 'abort', valid choices are: 'abort', 'proceed'
- **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'
- **from_** – Starting offset (default: 0)
- **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
- **pipeline** – Ingest pipeline to set on index requests made by this action. (default: none)
- **preference** – Specify the node or shard the operation should be performed on (default: random)
- **q** – Query in the Lucene query string syntax
- **refresh** – Should the effected indexes be refreshed?
- **request_cache** – Specify if request cache should be used for this request or not, defaults to index level setting
- **requests_per_second** – The throttle to set on this request in sub- requests per second. -1 means no throttle., default 0
- **routing** – A comma-separated list of specific routing values
- **scroll** – Specify how long a consistent view of the index should be maintained for scrolled search
- **scroll_size** – Size on the scroll request powering the update_by_query
- **search_timeout** – Explicit timeout for each search request. Defaults to no timeout.
- **search_type** – Search operation type, valid choices are: 'query_then_fetch', 'dfs_query_then_fetch'
- **size** – Number of hits to return (default: 10)
- **slices** – The number of slices this task should be divided into. Defaults to 1 meaning the task isn't sliced into subtasks., default 1
- **sort** – A comma-separated list of <field>:<direction> pairs
- **stats** – Specific 'tag' of the request for logging and statistical purposes
- **terminate_after** – The maximum number of documents to collect for each shard, upon reaching which the query execution will terminate early.
- **timeout** – Time each individual bulk request should wait for shards that are unavailable., default '1m'
- **version** – Specify whether to return document version as part of a hit
- **version_type** – Should the document increment the version number (internal) on hit or not (reindex)

- **wait_for_active_shards** – Sets the number of shard copies that must be active before proceeding with the update by query 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)
- **wait_for_completion** – Should the request should block until the update by query operation is complete., default True

7.1.3 Indices

class elasticsearch.client.**IndicesClient**(*client*)

analyze(***kwargs*)
Perform the analysis process on a text and return the tokens breakdown of the text. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-analyze.html

Parameters

- **index** – The name of the index to scope the operation
- **body** – Define analyzer/tokenizer parameters and the text on which the analysis should be performed
- **format** – Format of the output, default 'detailed', valid choices are: 'detailed', 'text'
- **prefer_local** – With *true*, specify that a local shard should be used if available, with *false*, use a random shard (default: true)

clear_cache(***kwargs*)
Clear either all caches or specific cached associated with one ore more indices. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-clearcache.html

Parameters

- **index** – A comma-separated list of index name to limit the operation
- **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)
- **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'
- **field_data** – Clear field data
- **fielddata** – Clear field data
- **fields** – A comma-separated list of fields to clear when using the *field_data* parameter (default: all)
- **ignore_unavailable** – Whether specified concrete indices should be ignored when unavailable (missing or closed)
- **query** – Clear query caches
- **recycler** – Clear the recycler cache
- **request** – Clear request cache
- **request_cache** – Clear request cache

close(***kwargs*)
Close an index to remove it's overhead from the cluster. Closed index is blocked for read/write operations. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-open-close.html

Parameters

- **index** – The name of the index
- **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)
- **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)
- **master_timeout** – Specify timeout for connection to master
- **timeout** – Explicit operation timeout

create(***kwargs*)

Create an index in Elasticsearch. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-create-index.html

Parameters

- **index** – The name of the index
- **body** – The configuration for the index (*settings* and *mappings*)
- **master_timeout** – Specify timeout for connection to master
- **timeout** – Explicit operation timeout
- **wait_for_active_shards** – Set the number of active shards to wait for before the operation returns.
- **include_type_name** – Specify whether requests and responses should include a type name (default: depends on Elasticsearch version).

delete(***kwargs*)

Delete an index in Elasticsearch http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-delete-index.html

Parameters

- **index** – A comma-separated list of indices to delete; use *_all* or * string to delete all indices
- **allow_no_indices** – Ignore if a wildcard expression resolves to no concrete indices (default: false)
- **expand_wildcards** – Whether wildcard expressions should get expanded to open or closed indices (default: open), default 'open', valid choices are: 'open', 'closed', 'none', 'all'
- **ignore_unavailable** – Ignore unavailable indexes (default: false)
- **master_timeout** – Specify timeout for connection to master
- **timeout** – Explicit operation timeout

delete_alias(***kwargs*)

Delete specific alias. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-aliases.html

Parameters

- **index** – A comma-separated list of index names (supports wildcards); use *_all* for all indices

- **name** – A comma-separated list of aliases to delete (supports wildcards); use *_all* to delete all aliases for the specified indices.
- **master_timeout** – Specify timeout for connection to master
- **timeout** – Explicit timeout for the operation

delete_template(***kwargs*)
Delete an index template by its name. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-templates.html

Parameters

- **name** – The name of the template
- **master_timeout** – Specify timeout for connection to master
- **timeout** – Explicit operation timeout

exists(***kwargs*)
Return a boolean indicating whether given index exists. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-exists.html

Parameters

- **index** – A comma-separated list of index names
- **allow_no_indices** – Ignore if a wildcard expression resolves to no concrete indices (default: false)
- **expand_wildcards** – Whether wildcard expressions should get expanded to open or closed indices (default: open), default 'open', valid choices are: 'open', 'closed', 'none', 'all'
- **flat_settings** – Return settings in flat format (default: false)
- **ignore_unavailable** – Ignore unavailable indexes (default: false)
- **include_defaults** – Whether to return all default setting for each of the indices., default False
- **local** – Return local information, do not retrieve the state from master node (default: false)

exists_alias(***kwargs*)
Return a boolean indicating whether given alias exists. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-aliases.html

Parameters

- **index** – A comma-separated list of index names to filter aliases
- **name** – A comma-separated list of alias names to return
- **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)
- **expand_wildcards** – Whether to expand wildcard expression to concrete indices that are open, closed or both., default 'all', valid choices are: 'open', 'closed', 'none', 'all'
- **ignore_unavailable** – Whether specified concrete indices should be ignored when unavailable (missing or closed)
- **local** – Return local information, do not retrieve the state from master node (default: false)

exists_template(***kwargs*)
Return a boolean indicating whether given template exists. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-templates.html

Parameters

- **name** – The comma separated names of the index templates
- **flat_settings** – Return settings in flat format (default: false)
- **local** – Return local information, do not retrieve the state from master node (default: false)
- **master_timeout** – Explicit operation timeout for connection to master node

exists_type(***kwargs*)
Check if a type/types exists in an index/indices. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-types-exists.html

Parameters

- **index** – A comma-separated list of index names; use *_all* to check the types across all indices
- **doc_type** – A comma-separated list of document types to check
- **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)
- **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)
- **local** – Return local information, do not retrieve the state from master node (default: false)

flush(***kwargs*)
Explicitly flush one or more indices. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-flush.html

Parameters

- **index** – A comma-separated list of index names; use *_all* or empty string for all indices
- **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)
- **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'
- **force** – Whether a flush should be forced even if it is not necessarily needed ie. if no changes will be committed to the index. This is useful if transaction log IDs should be incremented even if no uncommitted changes are present. (This setting can be considered as internal)
- **ignore_unavailable** – Whether specified concrete indices should be ignored when unavailable (missing or closed)
- **wait_if_ongoing** – If set to true the flush operation will block until the flush can be executed if another flush operation is already executing. The default is true. If set to false the flush will be skipped iff if another flush operation is already running.

flush_synced(***kwargs*)
Perform a normal flush, then add a generated unique marker (sync_id) to all shards. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-synced-flush.html

Parameters

- **index** – A comma-separated list of index names; use *_all* or empty string for all indices
- **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)
- **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)

forcemerge(***kwargs*)
The force merge API allows to force merging of one or more indices through an API. The merge relates to the number of segments a Lucene index holds within each shard. The force merge operation allows to reduce the number of segments by merging them.

This call will block until the merge is complete. If the http connection is lost, the request will continue in the background, and any new requests will block until the previous force merge is complete. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-forcemerge.html

Parameters

- **index** – A comma-separated list of index names; use *_all* or empty string to perform the operation on all indices
- **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)
- **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'
- **flush** – Specify whether the index should be flushed after performing the operation (default: true)
- **ignore_unavailable** – Whether specified concrete indices should be ignored when unavailable (missing or closed)
- **max_num_segments** – The number of segments the index should be merged into (default: dynamic)
- **only_expunge_deletes** – Specify whether the operation should only expunge deleted documents
- **operation_threading** – TODO: ?

get(***kwargs*)
The get index API allows to retrieve information about one or more indexes. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-get-index.html

Parameters

- **index** – A comma-separated list of index names
- **allow_no_indices** – Ignore if a wildcard expression resolves to no concrete indices (default: false)

- **expand_wildcards** – Whether wildcard expressions should get expanded to open or closed indices (default: open), default 'open', valid choices are: 'open', 'closed', 'none', 'all'
- **flat_settings** – Return settings in flat format (default: false)
- **ignore_unavailable** – Ignore unavailable indexes (default: false)
- **include_defaults** – Whether to return all default setting for each of the indices., default False
- **local** – Return local information, do not retrieve the state from master node (default: false)
- **include_type_name** – Specify whether requests and responses should include a type name (default: depends on Elasticsearch version).

get_alias(***kwargs*)

Retrieve a specified alias. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-aliases.html

Parameters

- **index** – A comma-separated list of index names to filter aliases
- **name** – A comma-separated list of alias names to return
- **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)
- **expand_wildcards** – Whether to expand wildcard expression to concrete indices that are open, closed or both., default 'all', valid choices are: 'open', 'closed', 'none', 'all'
- **ignore_unavailable** – Whether specified concrete indices should be ignored when unavailable (missing or closed)
- **local** – Return local information, do not retrieve the state from master node (default: false)

get_field_mapping(***kwargs*)

Retrieve mapping definition of a specific field. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-get-field-mapping.html

Parameters

- **fields** – A comma-separated list of fields
- **index** – A comma-separated list of index names
- **doc_type** – A comma-separated list of document types
- **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)
- **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)
- **include_defaults** – Whether the default mapping values should be returned as well
- **local** – Return local information, do not retrieve the state from master node (default: false)

- **include_type_name** – Specify whether requests and responses should include a type name (default: depends on Elasticsearch version).

get_mapping(***kwargs*)
Retrieve mapping definition of index or index/type. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-get-mapping.html

Parameters

- **index** – A comma-separated list of index names
- **doc_type** – A comma-separated list of document types
- **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)
- **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)
- **local** – Return local information, do not retrieve the state from master node (default: false)
- **include_type_name** – Specify whether requests and responses should include a type name (default: depends on Elasticsearch version).

get_settings(***kwargs*)
Retrieve settings for one or more (or all) indices. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-get-settings.html

Parameters

- **index** – A comma-separated list of index names; use *_all* or empty string to perform the operation on all indices
- **name** – The name of the settings that should be included
- **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)
- **expand_wildcards** – Whether to expand wildcard expression to concrete indices that are open, closed or both., default ['open', 'closed'], valid choices are: 'open', 'closed', 'none', 'all'
- **flat_settings** – Return settings in flat format (default: false)
- **ignore_unavailable** – Whether specified concrete indices should be ignored when unavailable (missing or closed)
- **include_defaults** – Whether to return all default setting for each of the indices., default False
- **local** – Return local information, do not retrieve the state from master node (default: false)

get_template(***kwargs*)
Retrieve an index template by its name. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-templates.html

Parameters

- **name** – The name of the template

- **flat_settings** – Return settings in flat format (default: false)
- **local** – Return local information, do not retrieve the state from master node (default: false)
- **master_timeout** – Explicit operation timeout for connection to master node
- **include_type_name** – Specify whether requests and responses should include a type name (default: depends on Elasticsearch version).

get_upgrade(***kwargs*)

Monitor how much of one or more index is upgraded. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-upgrade.html

Parameters

- **index** – A comma-separated list of index names; use *_all* or empty string to perform the operation on all indices
- **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)
- **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)

open(***kwargs*)

Open a closed index to make it available for search. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-open-close.html

Parameters

- **index** – The name of the index
- **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)
- **expand_wildcards** – Whether to expand wildcard expression to concrete indices that are open, closed or both., default 'closed', valid choices are: 'open', 'closed', 'none', 'all'
- **ignore_unavailable** – Whether specified concrete indices should be ignored when unavailable (missing or closed)
- **master_timeout** – Specify timeout for connection to master
- **timeout** – Explicit operation timeout

put_alias(***kwargs*)

Create an alias for a specific index/indices. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-aliases.html

Parameters

- **index** – A comma-separated list of index names the alias should point to (supports wildcards); use *_all* to perform the operation on all indices.
- **name** – The name of the alias to be created or updated
- **body** – The settings for the alias, such as *routing* or *filter*
- **master_timeout** – Specify timeout for connection to master
- **timeout** – Explicit timeout for the operation

put_mapping(***kwargs*)
Register specific mapping definition for a specific type. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-put-mapping.html

Parameters

- **doc_type** – The name of the document type
- **body** – The mapping definition
- **index** – A comma-separated list of index names the mapping should be added to (supports wildcards); use *_all* or omit to add the mapping on all indices.
- **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)
- **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)
- **master_timeout** – Specify timeout for connection to master
- **timeout** – Explicit operation timeout
- **include_type_name** – Specify whether requests and responses should include a type name (default: depends on Elasticsearch version).

put_settings(***kwargs*)
Change specific index level settings in real time. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-update-settings.html

Parameters

- **body** – The index settings to be updated
- **index** – A comma-separated list of index names; use *_all* or empty string to perform the operation on all indices
- **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)
- **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'
- **flat_settings** – Return settings in flat format (default: false)
- **ignore_unavailable** – Whether specified concrete indices should be ignored when unavailable (missing or closed)
- **master_timeout** – Specify timeout for connection to master
- **preserve_existing** – Whether to update existing settings. If set to *true* existing settings on an index remain unchanged, the default is *false*

put_template(***kwargs*)
Create an index template that will automatically be applied to new indices created. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-templates.html

Parameters

- **name** – The name of the template
- **body** – The template definition

- **create** – Whether the index template should only be added if new or can also replace an existing one, default False
- **flat_settings** – Return settings in flat format (default: false)
- **master_timeout** – Specify timeout for connection to master
- **order** – The order for this template when merging multiple matching ones (higher numbers are merged later, overriding the lower numbers)
- **timeout** – Explicit operation timeout
- **include_type_name** – Specify whether requests and responses should include a type name (default: depends on Elasticsearch version).

recovery(***kwargs*)

The indices recovery API provides insight into on-going shard recoveries. Recovery status may be reported for specific indices, or cluster-wide. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-recovery.html

Parameters

- **index** – A comma-separated list of index names; use *_all* or empty string to perform the operation on all indices
- **active_only** – Display only those recoveries that are currently on- going, default False
- **detailed** – Whether to display detailed information about shard recovery, default False

refresh(***kwargs*)

Explicitly refresh one or more index, making all operations performed since the last refresh available for search. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-refresh.html

Parameters

- **index** – A comma-separated list of index names; use *_all* or empty string to perform the operation on all indices
- **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)
- **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)

rollover(***kwargs*)

The rollover index API rolls an alias over to a new index when the existing index is considered to be too large or too old.

The API accepts a single alias name and a list of conditions. The alias must point to a single index only. If the index satisfies the specified conditions then a new index is created and the alias is switched to point to the new alias. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-rollover-index.html

Parameters

- **alias** – The name of the alias to rollover
- **new_index** – The name of the rollover index
- **body** – The conditions that needs to be met for executing rollover
- **dry_run** – If set to true the rollover action will only be validated but not actually performed even if a condition matches. The default is false

- **master_timeout** – Specify timeout for connection to master
- **timeout** – Explicit operation timeout
- **wait_for_active_shards** – Set the number of active shards to wait for on the newly created rollover index before the operation returns.
- **include_type_name** – Specify whether requests and responses should include a type name (default: depends on Elasticsearch version).

segments(***kwargs*)

Provide low level segments information that a Lucene index (shard level) is built with. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-segments.html

Parameters

- **index** – A comma-separated list of index names; use *_all* or empty string to perform the operation on all indices
- **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)
- **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)
- **operation_threading** – TODO: ?
- **verbose** – Includes detailed memory usage by Lucene., default False

shard_stores(***kwargs*)

Provides store information for shard copies of indices. Store information reports on which nodes shard copies exist, the shard copy version, indicating how recent they are, and any exceptions encountered while opening the shard index or from earlier engine failure. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-shards-stores.html

Parameters

- **index** – A comma-separated list of index names; use *_all* or empty string to perform the operation on all indices
- **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)
- **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)
- **operation_threading** – TODO: ?
- **status** – A comma-separated list of statuses used to filter on shards to get store information for, valid choices are: 'green', 'yellow', 'red', 'all'

shrink(***kwargs*)

The shrink index API allows you to shrink an existing index into a new index with fewer primary shards. The number of primary shards in the target index must be a factor of the shards in the source index. For example an index with 8 primary shards can be shrunk into 4, 2 or 1 primary shards or an index with 15 primary shards can be shrunk into 5, 3 or 1. If the number of shards in the index is a prime number it can only be shrunk into a single primary shard. Before shrinking, a (primary or replica) copy of every shard

in the index must be present on the same node. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-shrink-index.html

Parameters

- **index** – The name of the source index to shrink
- **target** – The name of the target index to shrink into
- **body** – The configuration for the target index (*settings* and *aliases*)
- **master_timeout** – Specify timeout for connection to master
- **timeout** – Explicit operation timeout
- **wait_for_active_shards** – Set the number of active shards to wait for on the shrunken index before the operation returns.

stats(***kwargs*)

Retrieve statistics on different operations happening on an index. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-stats.html

Parameters

- **index** – A comma-separated list of index names; use *_all* or empty string to perform the operation on all indices
- **metric** – Limit the information returned the specific metrics.
- **completion_fields** – A comma-separated list of fields for *fielddata* and *suggest* index metric (supports wildcards)
- **fielddata_fields** – A comma-separated list of fields for *fielddata* index metric (supports wildcards)
- **fields** – A comma-separated list of fields for *fielddata* and *completion* index metric (supports wildcards)
- **groups** – A comma-separated list of search groups for *search* index metric
- **include_segment_file_sizes** – Whether to report the aggregated disk usage of each one of the Lucene index files (only applies if segment stats are requested), default False
- **level** – Return stats aggregated at cluster, index or shard level, default 'indices', valid choices are: 'cluster', 'indices', 'shards'
- **types** – A comma-separated list of document types for the *indexing* index metric

update_aliases(***kwargs*)

Update specified aliases. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-aliases.html

Parameters

- **body** – The definition of *actions* to perform
- **master_timeout** – Specify timeout for connection to master
- **timeout** – Request timeout

upgrade(***kwargs*)

Upgrade one or more indices to the latest format through an API. http://www.elastic.co/guide/en/elasticsearch/reference/current/indices-upgrade.html

Parameters

- **index** – A comma-separated list of index names; use *_all* or empty string to perform the operation on all indices
- **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)
- **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)
- **only_ancient_segments** – If true, only ancient (an older Lucene major release) segments will be upgraded
- **wait_for_completion** – Specify whether the request should block until the all segments are upgraded (default: false)

validate_query(***kwargs*)

Validate a potentially expensive query without executing it. http://www.elastic.co/guide/en/elasticsearch/reference/current/search-validate.html

Parameters

- **index** – A comma-separated list of index names to restrict the operation; use *_all* or empty string to perform the operation on all indices
- **doc_type** – A comma-separated list of document types to restrict the operation; leave empty to perform the operation on all types
- **body** – The query definition specified with the Query DSL
- **all_shards** – Execute validation on all shards instead of one random shard per index
- **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'
- **explain** – Return detailed information about the error
- **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
- **operation_threading** – TODO: ?
- **q** – Query in the Lucene query string syntax
- **rewrite** – Provide a more detailed explanation showing the actual Lucene query that will be executed.

7.1.4 Ingest

class elasticsearch.client.**IngestClient**(*client*)

delete_pipeline(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/plugins/current/ingest.html

Parameters

- **id** – Pipeline ID
- **master_timeout** – Explicit operation timeout for connection to master node
- **timeout** – Explicit operation timeout

get_pipeline(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/plugins/current/ingest.html

Parameters

- **id** – Comma separated list of pipeline ids. Wildcards supported
- **master_timeout** – Explicit operation timeout for connection to master node

put_pipeline(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/plugins/current/ingest.html

Parameters

- **id** – Pipeline ID
- **body** – The ingest definition
- **master_timeout** – Explicit operation timeout for connection to master node
- **timeout** – Explicit operation timeout

simulate(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/plugins/current/ingest.html

Parameters

- **body** – The simulate definition
- **id** – Pipeline ID
- **verbose** – Verbose mode. Display data output for each processor in executed pipeline, default False

7.1.5 Cluster

class elasticsearch.client.**ClusterClient**(*client*)

allocation_explain(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/cluster-allocation-explain.html

Parameters

- **body** – The index, shard, and primary flag to explain. Empty means 'explain the first unassigned shard'
- **include_disk_info** – Return information about disk usage and shard sizes (default: false)

- **include_yes_decisions** – Return 'YES' decisions in explanation (default: false)

get_settings(***kwargs*)
Get cluster settings. http://www.elastic.co/guide/en/elasticsearch/reference/current/cluster-update-settings.html

Parameters

- **flat_settings** – Return settings in flat format (default: false)
- **include_defaults** – Whether to return all default clusters setting., default False
- **master_timeout** – Explicit operation timeout for connection to master node
- **timeout** – Explicit operation timeout

health(***kwargs*)
Get a very simple status on the health of the cluster. http://www.elastic.co/guide/en/elasticsearch/reference/current/cluster-health.html

Parameters

- **index** – Limit the information returned to a specific index
- **level** – Specify the level of detail for returned information, default 'cluster', valid choices are: 'cluster', 'indices', 'shards'
- **local** – Return local information, do not retrieve the state from master node (default: false)
- **master_timeout** – Explicit operation timeout for connection to master node
- **timeout** – Explicit operation timeout
- **wait_for_active_shards** – Wait until the specified number of shards is active
- **wait_for_events** – Wait until all currently queued events with the given priority are processed, valid choices are: 'immediate', 'urgent', 'high', 'normal', 'low', 'languid'
- **wait_for_no_relocating_shards** – Whether to wait until there are no relocating shards in the cluster
- **wait_for_nodes** – Wait until the specified number of nodes is available
- **wait_for_status** – Wait until cluster is in a specific state, default None, valid choices are: 'green', 'yellow', 'red'

pending_tasks(***kwargs*)
The pending cluster tasks API returns a list of any cluster-level changes (e.g. create index, update mapping, allocate or fail shard) which have not yet been executed. http://www.elastic.co/guide/en/elasticsearch/reference/current/cluster-pending.html

Parameters

- **local** – Return local information, do not retrieve the state from master node (default: false)
- **master_timeout** – Specify timeout for connection to master

put_settings(***kwargs*)
Update cluster wide specific settings. http://www.elastic.co/guide/en/elasticsearch/reference/current/cluster-update-settings.html

Parameters

- **body** – The settings to be updated. Can be either *transient* or *persistent* (survives cluster restart).
- **flat_settings** – Return settings in flat format (default: false)
- **master_timeout** – Explicit operation timeout for connection to master node
- **timeout** – Explicit operation timeout

reroute(***kwargs*)
Explicitly execute a cluster reroute allocation command including specific commands. http://www.elastic.co/guide/en/elasticsearch/reference/current/cluster-reroute.html

Parameters

- **body** – The definition of *commands* to perform (*move*, *cancel*, *allocate*)
- **dry_run** – Simulate the operation only and return the resulting state
- **explain** – Return an explanation of why the commands can or cannot be executed
- **master_timeout** – Explicit operation timeout for connection to master node
- **metric** – Limit the information returned to the specified metrics. Defaults to all but metadata, valid choices are: '_all', 'blocks', 'metadata', 'nodes', 'routing_table', 'master_node', 'version'
- **retry_failed** – Retries allocation of shards that are blocked due to too many subsequent allocation failures
- **timeout** – Explicit operation timeout

state(***kwargs*)
Get a comprehensive state information of the whole cluster. http://www.elastic.co/guide/en/elasticsearch/reference/current/cluster-state.html

Parameters

- **metric** – Limit the information returned to the specified metrics
- **index** – A comma-separated list of index names; use *_all* or empty string to perform the operation on all indices
- **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)
- **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'
- **flat_settings** – Return settings in flat format (default: false)
- **ignore_unavailable** – Whether specified concrete indices should be ignored when unavailable (missing or closed)
- **local** – Return local information, do not retrieve the state from master node (default: false)
- **master_timeout** – Specify timeout for connection to master

stats(***kwargs*)
The Cluster Stats API allows to retrieve statistics from a cluster wide perspective. The API returns basic index metrics and information about the current nodes that form the cluster. http://www.elastic.co/guide/en/elasticsearch/reference/current/cluster-stats.html

Parameters

- **node_id** – A comma-separated list of node IDs or names to limit the returned information; use *_local* to return information from the node you're connecting to, leave empty to get information from all nodes
- **flat_settings** – Return settings in flat format (default: false)
- **timeout** – Explicit operation timeout

7.1.6 Nodes

class elasticsearch.client.**NodesClient**(*client*)

hot_threads(***kwargs*)
An API allowing to get the current hot threads on each node in the cluster. https://www.elastic.co/guide/en/elasticsearch/reference/current/cluster-nodes-hot-threads.html

Parameters

- **node_id** – A comma-separated list of node IDs or names to limit the returned information; use *_local* to return information from the node you're connecting to, leave empty to get information from all nodes
- **type** – The type to sample (default: cpu), valid choices are: 'cpu', 'wait', 'block'
- **ignore_idle_threads** – Don't show threads that are in known-idle places, such as waiting on a socket select or pulling from an empty task queue (default: true)
- **interval** – The interval for the second sampling of threads
- **snapshots** – Number of samples of thread stacktrace (default: 10)
- **threads** – Specify the number of threads to provide information for (default: 3)
- **timeout** – Explicit operation timeout

info(***kwargs*)
The cluster nodes info API allows to retrieve one or more (or all) of the cluster nodes information. https://www.elastic.co/guide/en/elasticsearch/reference/current/cluster-nodes-info.html

Parameters

- **node_id** – A comma-separated list of node IDs or names to limit the returned information; use *_local* to return information from the node you're connecting to, leave empty to get information from all nodes
- **metric** – A comma-separated list of metrics you wish returned. Leave empty to return all.
- **flat_settings** – Return settings in flat format (default: false)
- **timeout** – Explicit operation timeout

reload_secure_settings(***kwargs*)
Reload any settings that have been marked as "reloadable" https://www.elastic.co/guide/en/elasticsearch/reference/current/secure-settings.html#reloadable-secure-settings

stats(***kwargs*)
The cluster nodes stats API allows to retrieve one or more (or all) of the cluster nodes statistics. https://www.elastic.co/guide/en/elasticsearch/reference/current/cluster-nodes-stats.html

Parameters

- **node_id** – A comma-separated list of node IDs or names to limit the returned information; use *_local* to return information from the node you're connecting to, leave empty to get information from all nodes
- **metric** – Limit the information returned to the specified metrics
- **index_metric** – Limit the information returned for *indices* metric to the specific index metrics. Isn't used if *indices* (or *all*) metric isn't specified.
- **completion_fields** – A comma-separated list of fields for *fielddata* and *suggest* index metric (supports wildcards)
- **fielddata_fields** – A comma-separated list of fields for *fielddata* index metric (supports wildcards)
- **fields** – A comma-separated list of fields for *fielddata* and *completion* index metric (supports wildcards)
- **groups** – A comma-separated list of search groups for *search* index metric
- **include_segment_file_sizes** – Whether to report the aggregated disk usage of each one of the Lucene index files (only applies if segment stats are requested), default False
- **level** – Return indices stats aggregated at index, node or shard level, default 'node', valid choices are: 'indices', 'node', 'shards'
- **timeout** – Explicit operation timeout
- **types** – A comma-separated list of document types for the *indexing* index metric

usage(***kwargs*)
The cluster nodes usage API allows to retrieve information on the usage of features for each node. http://www.elastic.co/guide/en/elasticsearch/reference/master/cluster-nodes-usage.html

Parameters

- **node_id** – A comma-separated list of node IDs or names to limit the returned information; use *_local* to return information from the node you're connecting to, leave empty to get information from all nodes
- **metric** – Limit the information returned to the specified metrics
- **human** – Whether to return time and byte values in human-readable format., default False
- **timeout** – Explicit operation timeout

7.1.7 Cat

class elasticsearch.client.**CatClient**(*client*)

aliases(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/cat-alias.html

Parameters

- **name** – A comma-separated list of alias names to return
- **format** – a short version of the Accept header, e.g. json, yaml
- **h** – Comma-separated list of column names to display
- **help** – Return help information, default False

- **local** – Return local information, do not retrieve the state from master node (default: false)
- **master_timeout** – Explicit operation timeout for connection to master node
- **s** – Comma-separated list of column names or column aliases to sort by
- **v** – Verbose mode. Display column headers, default False

allocation(***kwargs*)

Allocation provides a snapshot of how shards have located around the cluster and the state of disk usage. https://www.elastic.co/guide/en/elasticsearch/reference/current/cat-allocation.html

Parameters

- **node_id** – A comma-separated list of node IDs or names to limit the returned information
- **bytes** – The unit in which to display byte values, valid choices are: 'b', 'k', 'kb', 'm', 'mb', 'g', 'gb', 't', 'tb', 'p', 'pb'
- **format** – a short version of the Accept header, e.g. json, yaml
- **h** – Comma-separated list of column names to display
- **help** – Return help information, default False
- **local** – Return local information, do not retrieve the state from master node (default: false)
- **master_timeout** – Explicit operation timeout for connection to master node
- **s** – Comma-separated list of column names or column aliases to sort by
- **v** – Verbose mode. Display column headers, default False

count(***kwargs*)

Count provides quick access to the document count of the entire cluster, or individual indices. https://www.elastic.co/guide/en/elasticsearch/reference/current/cat-count.html

Parameters

- **index** – A comma-separated list of index names to limit the returned information
- **format** – a short version of the Accept header, e.g. json, yaml
- **h** – Comma-separated list of column names to display
- **help** – Return help information, default False
- **local** – Return local information, do not retrieve the state from master node (default: false)
- **master_timeout** – Explicit operation timeout for connection to master node
- **s** – Comma-separated list of column names or column aliases to sort by
- **v** – Verbose mode. Display column headers, default False

fielddata(***kwargs*)

Shows information about currently loaded fielddata on a per-node basis. https://www.elastic.co/guide/en/elasticsearch/reference/current/cat-fielddata.html

Parameters

- **fields** – A comma-separated list of fields to return the fielddata size

- **bytes** – The unit in which to display byte values, valid choices are: 'b', 'k', 'kb', 'm', 'mb', 'g', 'gb', 't', 'tb', 'p', 'pb'
- **format** – a short version of the Accept header, e.g. json, yaml
- **h** – Comma-separated list of column names to display
- **help** – Return help information, default False
- **local** – Return local information, do not retrieve the state from master node (default: false)
- **master_timeout** – Explicit operation timeout for connection to master node
- **s** – Comma-separated list of column names or column aliases to sort by
- **v** – Verbose mode. Display column headers, default False

health(***kwargs*)

health is a terse, one-line representation of the same information from `health()` API https://www.elastic.co/guide/en/elasticsearch/reference/current/cat-health.html

Parameters

- **format** – a short version of the Accept header, e.g. json, yaml
- **h** – Comma-separated list of column names to display
- **help** – Return help information, default False
- **local** – Return local information, do not retrieve the state from master node (default: false)
- **master_timeout** – Explicit operation timeout for connection to master node
- **s** – Comma-separated list of column names or column aliases to sort by
- **ts** – Set to false to disable timestamping, default True
- **v** – Verbose mode. Display column headers, default False

help(***kwargs*)

A simple help for the cat api. https://www.elastic.co/guide/en/elasticsearch/reference/current/cat.html

Parameters

- **help** – Return help information, default False
- **s** – Comma-separated list of column names or column aliases to sort by

indices(***kwargs*)

The indices command provides a cross-section of each index. https://www.elastic.co/guide/en/elasticsearch/reference/current/cat-indices.html

Parameters

- **index** – A comma-separated list of index names to limit the returned information
- **bytes** – The unit in which to display byte values, valid choices are: 'b', 'k', 'm', 'g'
- **format** – a short version of the Accept header, e.g. json, yaml
- **h** – Comma-separated list of column names to display
- **health** – A health status ("green", "yellow", or "red" to filter only indices matching the specified health status, default None, valid choices are: 'green', 'yellow', 'red'
- **help** – Return help information, default False

- **local** – Return local information, do not retrieve the state from master node (default: false)
- **master_timeout** – Explicit operation timeout for connection to master node
- **pri** – Set to true to return stats only for primary shards, default False
- **s** – Comma-separated list of column names or column aliases to sort by
- **v** – Verbose mode. Display column headers, default False

master(***kwargs*)
Displays the master's node ID, bound IP address, and node name. https://www.elastic.co/guide/en/elasticsearch/reference/current/cat-master.html

Parameters

- **format** – a short version of the Accept header, e.g. json, yaml
- **h** – Comma-separated list of column names to display
- **help** – Return help information, default False
- **local** – Return local information, do not retrieve the state from master node (default: false)
- **master_timeout** – Explicit operation timeout for connection to master node
- **s** – Comma-separated list of column names or column aliases to sort by
- **v** – Verbose mode. Display column headers, default False

nodeattrs(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/cat-nodeattrs.html

Parameters

- **format** – a short version of the Accept header, e.g. json, yaml
- **h** – Comma-separated list of column names to display
- **help** – Return help information, default False
- **local** – Return local information, do not retrieve the state from master node (default: false)
- **master_timeout** – Explicit operation timeout for connection to master node
- **s** – Comma-separated list of column names or column aliases to sort by
- **v** – Verbose mode. Display column headers, default False

nodes(***kwargs*)
The nodes command shows the cluster topology. https://www.elastic.co/guide/en/elasticsearch/reference/current/cat-nodes.html

Parameters

- **format** – a short version of the Accept header, e.g. json, yaml
- **full_id** – Return the full node ID instead of the shortened version (default: false)
- **h** – Comma-separated list of column names to display
- **help** – Return help information, default False
- **local** – Return local information, do not retrieve the state from master node (default: false)

- **master_timeout** – Explicit operation timeout for connection to master node
- **s** – Comma-separated list of column names or column aliases to sort by
- **v** – Verbose mode. Display column headers, default False

pending_tasks(***kwargs*)
pending_tasks provides the same information as the `pending_tasks()` API in a convenient tabular format. https://www.elastic.co/guide/en/elasticsearch/reference/current/cat-pending-tasks.html

Parameters

- **format** – a short version of the Accept header, e.g. json, yaml
- **h** – Comma-separated list of column names to display
- **help** – Return help information, default False
- **local** – Return local information, do not retrieve the state from master node (default: false)
- **master_timeout** – Explicit operation timeout for connection to master node
- **s** – Comma-separated list of column names or column aliases to sort by
- **v** – Verbose mode. Display column headers, default False

plugins(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/cat-plugins.html

Parameters

- **format** – a short version of the Accept header, e.g. json, yaml
- **h** – Comma-separated list of column names to display
- **help** – Return help information, default False
- **local** – Return local information, do not retrieve the state from master node (default: false)
- **master_timeout** – Explicit operation timeout for connection to master node
- **s** – Comma-separated list of column names or column aliases to sort by
- **v** – Verbose mode. Display column headers, default False

recovery(***kwargs*)
recovery is a view of shard replication. https://www.elastic.co/guide/en/elasticsearch/reference/current/cat-recovery.html

Parameters

- **index** – A comma-separated list of index names to limit the returned information
- **bytes** – The unit in which to display byte values, valid choices are: 'b', 'k', 'kb', 'm', 'mb', 'g', 'gb', 't', 'tb', 'p', 'pb'
- **format** – a short version of the Accept header, e.g. json, yaml
- **h** – Comma-separated list of column names to display
- **help** – Return help information, default False
- **master_timeout** – Explicit operation timeout for connection to master node
- **s** – Comma-separated list of column names or column aliases to sort by
- **v** – Verbose mode. Display column headers, default False

repositories(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/cat-repositories.html

Parameters

- **format** – a short version of the Accept header, e.g. json, yaml
- **h** – Comma-separated list of column names to display
- **help** – Return help information, default False
- **local** – Return local information, do not retrieve the state from master node, default False
- **master_timeout** – Explicit operation timeout for connection to master node
- **s** – Comma-separated list of column names or column aliases to sort by
- **v** – Verbose mode. Display column headers, default False

segments(***kwargs*)
The segments command is the detailed view of Lucene segments per index. https://www.elastic.co/guide/en/elasticsearch/reference/current/cat-segments.html

Parameters

- **index** – A comma-separated list of index names to limit the returned information
- **bytes** – The unit in which to display byte values, valid choices are: 'b', 'k', 'kb', 'm', 'mb', 'g', 'gb', 't', 'tb', 'p', 'pb'
- **format** – a short version of the Accept header, e.g. json, yaml
- **h** – Comma-separated list of column names to display
- **help** – Return help information, default False
- **s** – Comma-separated list of column names or column aliases to sort by
- **v** – Verbose mode. Display column headers, default False

shards(***kwargs*)
The shards command is the detailed view of what nodes contain which shards. https://www.elastic.co/guide/en/elasticsearch/reference/current/cat-shards.html

Parameters

- **index** – A comma-separated list of index names to limit the returned information
- **bytes** – The unit in which to display byte values, valid choices are: 'b', 'k', 'kb', 'm', 'mb', 'g', 'gb', 't', 'tb', 'p', 'pb'
- **format** – a short version of the Accept header, e.g. json, yaml
- **h** – Comma-separated list of column names to display
- **help** – Return help information, default False
- **local** – Return local information, do not retrieve the state from master node (default: false)
- **master_timeout** – Explicit operation timeout for connection to master node
- **s** – Comma-separated list of column names or column aliases to sort by
- **v** – Verbose mode. Display column headers, default False

snapshots (***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/cat-snapshots.html

Parameters

- **repository** – Name of repository from which to fetch the snapshot information
- **format** – a short version of the Accept header, e.g. json, yaml
- **h** – Comma-separated list of column names to display
- **help** – Return help information, default False
- **ignore_unavailable** – Set to true to ignore unavailable snapshots, default False
- **master_timeout** – Explicit operation timeout for connection to master node
- **s** – Comma-separated list of column names or column aliases to sort by
- **v** – Verbose mode. Display column headers, default False

tasks (***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/tasks.html

Parameters

- **actions** – A comma-separated list of actions that should be returned. Leave empty to return all.
- **detailed** – Return detailed task information (default: false)
- **format** – a short version of the Accept header, e.g. json, yaml
- **h** – Comma-separated list of column names to display
- **help** – Return help information, default False
- **nodes** – A comma-separated list of node IDs or names to limit the returned information; use *_local* to return information from the node you're connecting to, leave empty to get information from all nodes
- **parent_task_id** – Return tasks with specified parent task id. Set to -1 to return all.
- **s** – Comma-separated list of column names or column aliases to sort by
- **v** – Verbose mode. Display column headers, default False

templates (***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/cat-templates.html

Parameters

- **name** – A pattern that returned template names must match
- **format** – a short version of the Accept header, e.g. json, yaml
- **h** – Comma-separated list of column names to display
- **help** – Return help information, default False
- **local** – Return local information, do not retrieve the state from master node (default: false)
- **master_timeout** – Explicit operation timeout for connection to master node
- **s** – Comma-separated list of column names or column aliases to sort by
- **v** – Verbose mode. Display column headers, default False

thread_pool(***kwargs*)
Get information about thread pools. https://www.elastic.co/guide/en/elasticsearch/reference/current/cat-thread-pool.html

Parameters

- **thread_pool_patterns** – A comma-separated list of regular-expressions to filter the thread pools in the output
- **format** – a short version of the Accept header, e.g. json, yaml
- **h** – Comma-separated list of column names to display
- **help** – Return help information, default False
- **local** – Return local information, do not retrieve the state from master node (default: false)
- **master_timeout** – Explicit operation timeout for connection to master node
- **s** – Comma-separated list of column names or column aliases to sort by
- **size** – The multiplier in which to display values, valid choices are: '', 'k', 'm', 'g', 't', 'p'
- **v** – Verbose mode. Display column headers, default False

7.1.8 Snapshot

class elasticsearch.client.**SnapshotClient**(*client*)

create(***kwargs*)
Create a snapshot in repository http://www.elastic.co/guide/en/elasticsearch/reference/current/modules-snapshots.html

Parameters

- **repository** – A repository name
- **snapshot** – A snapshot name
- **body** – The snapshot definition
- **master_timeout** – Explicit operation timeout for connection to master node
- **wait_for_completion** – Should this request wait until the operation has completed before returning, default False

create_repository(***kwargs*)
Registers a shared file system repository. http://www.elastic.co/guide/en/elasticsearch/reference/current/modules-snapshots.html

Parameters

- **repository** – A repository name
- **body** – The repository definition
- **master_timeout** – Explicit operation timeout for connection to master node
- **timeout** – Explicit operation timeout
- **verify** – Whether to verify the repository after creation

delete(***kwargs*)
Deletes a snapshot from a repository. http://www.elastic.co/guide/en/elasticsearch/reference/current/modules-snapshots.html

Parameters

- **repository** – A repository name
- **snapshot** – A snapshot name
- **master_timeout** – Explicit operation timeout for connection to master node

delete_repository(***kwargs*)
Removes a shared file system repository. http://www.elastic.co/guide/en/elasticsearch/reference/current/modules-snapshots.html

Parameters

- **repository** – A comma-separated list of repository names
- **master_timeout** – Explicit operation timeout for connection to master node
- **timeout** – Explicit operation timeout

get(***kwargs*)
Retrieve information about a snapshot. http://www.elastic.co/guide/en/elasticsearch/reference/current/modules-snapshots.html

Parameters

- **repository** – A repository name
- **snapshot** – A comma-separated list of snapshot names
- **ignore_unavailable** – Whether to ignore unavailable snapshots, defaults to false which means a NotFoundError *snapshot_missing_exception* is thrown
- **master_timeout** – Explicit operation timeout for connection to master node
- **verbose** – Whether to show verbose snapshot info or only show the basic info found in the repository index blob

get_repository(***kwargs*)
Return information about registered repositories. http://www.elastic.co/guide/en/elasticsearch/reference/current/modules-snapshots.html

Parameters

- **repository** – A comma-separated list of repository names
- **local** – Return local information, do not retrieve the state from master node (default: false)
- **master_timeout** – Explicit operation timeout for connection to master node

restore(***kwargs*)
Restore a snapshot. http://www.elastic.co/guide/en/elasticsearch/reference/current/modules-snapshots.html

Parameters

- **repository** – A repository name
- **snapshot** – A snapshot name
- **body** – Details of what to restore

- **master_timeout** – Explicit operation timeout for connection to master node
- **wait_for_completion** – Should this request wait until the operation has completed before returning, default False

status(***kwargs*)
Return information about all currently running snapshots. By specifying a repository name, it's possible to limit the results to a particular repository. http://www.elastic.co/guide/en/elasticsearch/reference/current/modules-snapshots.html

Parameters

- **repository** – A repository name
- **snapshot** – A comma-separated list of snapshot names
- **ignore_unavailable** – Whether to ignore unavailable snapshots, defaults to false which means a NotFoundError *snapshot_missing_exception* is thrown
- **master_timeout** – Explicit operation timeout for connection to master node

verify_repository(***kwargs*)
Returns a list of nodes where repository was successfully verified or an error message if verification process failed. http://www.elastic.co/guide/en/elasticsearch/reference/current/modules-snapshots.html

Parameters

- **repository** – A repository name
- **master_timeout** – Explicit operation timeout for connection to master node
- **timeout** – Explicit operation timeout

7.1.9 Tasks

class elasticsearch.client.**TasksClient**(*client*)

cancel(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/tasks.html

Parameters

- **task_id** – Cancel the task with specified task id (node_id:task_number)
- **actions** – A comma-separated list of actions that should be cancelled. Leave empty to cancel all.
- **nodes** – A comma-separated list of node IDs or names to limit the returned information; use *_local* to return information from the node you're connecting to, leave empty to get information from all nodes
- **parent_task_id** – Cancel tasks with specified parent task id (node_id:task_number). Set to -1 to cancel all.

get(***kwargs*)
Retrieve information for a particular task. http://www.elastic.co/guide/en/elasticsearch/reference/current/tasks.html

Parameters

- **task_id** – Return the task with specified id (node_id:task_number)
- **wait_for_completion** – Wait for the matching tasks to complete (default: false)

- **timeout** – Maximum waiting time for *wait_for_completion*

list(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/tasks.html

Parameters

- **actions** – A comma-separated list of actions that should be returned. Leave empty to return all.
- **detailed** – Return detailed task information (default: false)
- **group_by** – Group tasks by nodes or parent/child relationships, default 'nodes', valid choices are: 'nodes', 'parents'
- **nodes** – A comma-separated list of node IDs or names to limit the returned information; use *_local* to return information from the node you're connecting to, leave empty to get information from all nodes
- **parent_task_id** – Return tasks with specified parent task id (node_id:task_number). Set to -1 to return all.
- **wait_for_completion** – Wait for the matching tasks to complete (default: false)
- **timeout** – Maximum waiting time for *wait_for_completion*

7.2 X-Pack APIs

X-Pack is an Elastic Stack extension that bundles security, alerting, monitoring, reporting, and graph capabilities into one easy-to-install package. While the X-Pack components are designed to work together seamlessly, you can easily enable or disable the features you want to use.

7.2.1 Info

X-Pack info provides general info about the installed X-Pack.

class elasticsearch.client.xpack.**XPackClient**(**args*, ***kwargs*)

info(***kwargs*)
Retrieve information about xpack, including build number/timestamp and license status https://www.elastic.co/guide/en/elasticsearch/reference/current/info-api.html

Parameters

- **categories** – Comma-separated list of info categories. Can be any of: build, license, features
- **human** – Presents additional info for humans (feature descriptions and X-Pack tagline)

usage(***kwargs*)
Retrieve information about xpack features usage

Parameters **master_timeout** – Specify timeout for watch write operation

7.2.2 Graph Explore

X-Pack Graph Explore enables you to extract and summarize information about the documents and terms in your Elasticsearch index.

class elasticsearch.client.xpack.graph.**GraphClient**(*client*)

explore(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/graph-explore-api.html

Parameters

- **index** – A comma-separated list of index names to search; use *_all* or empty string to perform the operation on all indices
- **doc_type** – A comma-separated list of document types to search; leave empty to perform the operation on all types
- **body** – Graph Query DSL
- **routing** – Specific routing value
- **timeout** – Explicit operation timeout

7.2.3 Licensing API

Licensing API can be used to manage your licences.

class elasticsearch.client.xpack.license.**LicenseClient**(*client*)

delete(***kwargs*)
https://www.elastic.co/guide/en/x-pack/current/license-management.html

get(***kwargs*)
https://www.elastic.co/guide/en/x-pack/current/license-management.html

Parameters **local** – Return local information, do not retrieve the state from master node (default: false)

get_basic_status(***kwargs*)
https://www.elastic.co/guide/en/x-pack/current/license-management.html

get_trial_status(***kwargs*)
https://www.elastic.co/guide/en/x-pack/current/license-management.html

post(***kwargs*)
https://www.elastic.co/guide/en/x-pack/current/license-management.html

Parameters

- **body** – licenses to be installed
- **acknowledge** – whether the user has acknowledged acknowledge messages (default: false)

post_start_basic(***kwargs*)
https://www.elastic.co/guide/en/x-pack/current/license-management.html

Parameters **acknowledge** – whether the user has acknowledged acknowledge messages (default: false)

post_start_trial(***kwargs*)
https://www.elastic.co/guide/en/x-pack/current/license-management.html

Parameters

- **acknowledge** – whether the user has acknowledged acknowledge messages (default: false)
- **doc_type** – The type of trial license to generate (default: "trial")

7.2.4 Machine Learning APIs

Machine Learning can be useful for discovering new patterns about your data. For a more detailed explanation about X-Pack's machine learning please refer to the official documentation.

class elasticsearch.client.xpack.ml.**MlClient**(*client*)

close_job(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-close-job.html

Parameters

- **job_id** – The name of the job to close
- **body** – The URL params optionally sent in the body
- **allow_no_jobs** – Whether to ignore if a wildcard expression matches no jobs. (This includes *_all* string or when no jobs have been specified)
- **force** – True if the job should be forcefully closed
- **timeout** – Controls the time to wait until a job has closed. Default to 30 minutes

delete_calendar(***kwargs*)
'<>'_

Parameters **calendar_id** – The ID of the calendar to delete

delete_calendar_event(***kwargs*)
'<>'_

Parameters

- **calendar_id** – The ID of the calendar to modify
- **event_id** – The ID of the event to remove from the calendar

delete_calendar_job(***kwargs*)
'<>'_

Parameters

- **calendar_id** – The ID of the calendar to modify
- **job_id** – The ID of the job to remove from the calendar

delete_datafeed(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-delete-datafeed.html

Parameters

- **datafeed_id** – The ID of the datafeed to delete
- **force** – True if the datafeed should be forcefully deleted

delete_expired_data(***kwargs*)
‘<>‘_

delete_filter(***kwargs*)
‘<>‘_

Parameters **filter_id** – The ID of the filter to delete

delete_forecast(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-delete-forecast.html

Parameters

- **job_id** – The ID of the job from which to delete forecasts
- **forecast_id** – The ID of the forecast to delete, can be comma delimited list. Leaving blank implies *_all*
- **allow_no_forecasts** – Whether to ignore if *_all* matches no forecasts
- **timeout** – Controls the time to wait until the forecast(s) are deleted. Default to 30 seconds

delete_job(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-delete-job.html

Parameters

- **job_id** – The ID of the job to delete
- **force** – True if the job should be forcefully deleted, default False
- **wait_for_completion** – Should this request wait until the operation has completed before returning, default True

delete_model_snapshot(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-delete-snapshot.html

Parameters

- **job_id** – The ID of the job to fetch
- **snapshot_id** – The ID of the snapshot to delete

find_file_structure(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-file-structure.html

Parameters

- **body** – The contents of the file to be analyzed
- **charset** – Optional parameter to specify the character set of the file
- **column_names** – Optional parameter containing a comma separated list of the column names for a delimited file
- **delimiter** – Optional parameter to specify the delimiter character for a delimited file - must be a single character
- **explain** – Whether to include a commentary on how the structure was derived, default False
- **format** – Optional parameter to specify the high level file format, valid choices are: ‘ndjson’, ‘xml’, ‘delimited’, ‘semi_structured_text’
- **grok_pattern** – Optional parameter to specify the Grok pattern that should be used to extract fields from messages in a semi-structured text file

- **has_header_row** – Optional parameter to specify whether a delimited file includes the column names in its first row
- **lines_to_sample** – How many lines of the file should be included in the analysis, default 1000
- **quote** – Optional parameter to specify the quote character for a delimited file - must be a single character
- **should_trim_fields** – Optional parameter to specify whether the values between delimiters in a delimited file should have whitespace trimmed from them
- **timeout** – Timeout after which the analysis will be aborted, default '25s'
- **timestamp_field** – Optional parameter to specify the timestamp field in the file
- **timestamp_format** – Optional parameter to specify the timestamp format in the file - may be either a Joda or Java time format

flush_job(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-flush-job.html

Parameters

- **job_id** – The name of the job to flush
- **body** – Flush parameters
- **advance_time** – Advances time to the given value generating results and updating the model for the advanced interval
- **calc_interim** – Calculates interim results for the most recent bucket or all buckets within the latency period
- **end** – When used in conjunction with calc_interim, specifies the range of buckets on which to calculate interim results
- **skip_time** – Skips time to the given value without generating results or updating the model for the skipped interval
- **start** – When used in conjunction with calc_interim, specifies the range of buckets on which to calculate interim results

forecast(***kwargs*)
'<>'_

Parameters

- **job_id** – The ID of the job to forecast for
- **duration** – The duration of the forecast
- **expires_in** – The time interval after which the forecast expires. Expired forecasts will be deleted at the first opportunity.

get_buckets(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-get-bucket.html

Parameters

- **job_id** – ID of the job to get bucket results from
- **timestamp** – The timestamp of the desired single bucket result
- **body** – Bucket selection details if not provided in URI
- **anomaly_score** – Filter for the most anomalous buckets

- **desc** – Set the sort direction
- **end** – End time filter for buckets
- **exclude_interim** – Exclude interim results
- **expand** – Include anomaly records
- **from** – skips a number of buckets
- **size** – specifies a max number of buckets to get
- **sort** – Sort buckets by a particular field
- **start** – Start time filter for buckets

get_calendar_events (***kwargs*)
‘<>‘_

Parameters

- **calendar_id** – The ID of the calendar containing the events
- **end** – Get events before this time
- **from** – Skips a number of events
- **job_id** – Get events for the job. When this option is used calendar_id must be ‘_all’
- **size** – Specifies a max number of events to get
- **start** – Get events after this time

get_calendars (***kwargs*)
‘<>‘_

Parameters

- **calendar_id** – The ID of the calendar to fetch
- **body** – The from and size parameters optionally sent in the body
- **from** – skips a number of calendars
- **size** – specifies a max number of calendars to get

get_categories (***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-get-category.html

Parameters

- **job_id** – The name of the job
- **category_id** – The identifier of the category definition of interest
- **body** – Category selection details if not provided in URI
- **from** – skips a number of categories
- **size** – specifies a max number of categories to get

get_datafeed_stats (***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-get-datafeed-stats.html

Parameters

- **datafeed_id** – The ID of the datafeeds stats to fetch
- **allow_no_datafeeds** – Whether to ignore if a wildcard expression matches no datafeeds. (This includes *_all* string or when no datafeeds have been specified)

get_datafeeds(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-get-datafeed.html

Parameters

- **datafeed_id** – The ID of the datafeeds to fetch
- **allow_no_datafeeds** – Whether to ignore if a wildcard expression matches no datafeeds. (This includes *_all* string or when no datafeeds have been specified)

get_filters(***kwargs*)
'<>'_

Parameters

- **filter_id** – The ID of the filter to fetch
- **from** – skips a number of filters
- **size** – specifies a max number of filters to get

get_influencers(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-get-influencer.html

Parameters

- **job_id** – None
- **body** – Influencer selection criteria
- **desc** – whether the results should be sorted in decending order
- **end** – end timestamp for the requested influencers
- **exclude_interim** – Exclude interim results
- **from** – skips a number of influencers
- **influencer_score** – influencer score threshold for the requested influencers
- **size** – specifies a max number of influencers to get
- **sort** – sort field for the requested influencers
- **start** – start timestamp for the requested influencers

get_job_stats(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-get-job-stats.html

Parameters

- **job_id** – The ID of the jobs stats to fetch
- **allow_no_jobs** – Whether to ignore if a wildcard expression matches no jobs. (This includes *_all* string or when no jobs have been specified)

get_jobs(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-get-job.html

Parameters

- **job_id** – The ID of the jobs to fetch
- **allow_no_jobs** – Whether to ignore if a wildcard expression matches no jobs. (This includes *_all* string or when no jobs have been specified)

get_model_snapshots(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-get-snapshot.html

Parameters

- **job_id** – The ID of the job to fetch
- **snapshot_id** – The ID of the snapshot to fetch
- **body** – Model snapshot selection criteria
- **desc** – True if the results should be sorted in descending order
- **end** – The filter 'end' query parameter
- **from** – Skips a number of documents
- **size** – The default number of documents returned in queries as a string.
- **sort** – Name of the field to sort on
- **start** – The filter 'start' query parameter

get_overall_buckets(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-get-overall-buckets.html

Parameters

- **job_id** – The job IDs for which to calculate overall bucket results
- **body** – Overall bucket selection details if not provided in URI
- **allow_no_jobs** – Whether to ignore if a wildcard expression matches no jobs. (This includes *_all* string or when no jobs have been specified)
- **bucket_span** – The span of the overall buckets. Defaults to the longest job bucket_span
- **end** – Returns overall buckets with timestamps earlier than this time
- **exclude_interim** – If true overall buckets that include interim buckets will be excluded
- **overall_score** – Returns overall buckets with overall scores higher than this value
- **start** – Returns overall buckets with timestamps after this time
- **top_n** – The number of top job bucket scores to be used in the overall_score calculation

get_records(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-get-record.html

Parameters

- **job_id** – None
- **body** – Record selection criteria
- **desc** – Set the sort direction
- **end** – End time filter for records
- **exclude_interim** – Exclude interim results
- **from** – skips a number of records
- **record_score** –
- **size** – specifies a max number of records to get
- **sort** – Sort records by a particular field
- **start** – Start time filter for records

info (***kwargs*)
‘<>‘_

open_job (***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-open-job.html

Parameters **job_id** – The ID of the job to open

post_calendar_events (***kwargs*)
‘<>‘_

Parameters

- **calendar_id** – The ID of the calendar to modify
- **body** – A list of events

post_data (***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-post-data.html

Parameters

- **job_id** – The name of the job receiving the data
- **body** – The data to process
- **reset_end** – Optional parameter to specify the end of the bucket resetting range
- **reset_start** – Optional parameter to specify the start of the bucket resetting range

preview_datafeed (***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-preview-datafeed.html

Parameters **datafeed_id** – The ID of the datafeed to preview

put_calendar (***kwargs*)
‘<>‘_

Parameters

- **calendar_id** – The ID of the calendar to create
- **body** – The calendar details

put_calendar_job (***kwargs*)
‘<>‘_

Parameters

- **calendar_id** – The ID of the calendar to modify
- **job_id** – The ID of the job to add to the calendar

put_datafeed (***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-put-datafeed.html

Parameters

- **datafeed_id** – The ID of the datafeed to create
- **body** – The datafeed config

put_filter (***kwargs*)
‘<>‘_

Parameters

- **filter_id** – The ID of the filter to create

- **body** – The filter details

put_job(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-put-job.html

Parameters

- **job_id** – The ID of the job to create
- **body** – The job

revert_model_snapshot(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-revert-snapshot.html

Parameters

- **job_id** – The ID of the job to fetch
- **snapshot_id** – The ID of the snapshot to revert to
- **body** – Reversion options
- **delete_intervening_results** – Should we reset the results back to the time of the snapshot?

set_upgrade_mode(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-set-upgrade-mode.html

Parameters

- **enabled** – Whether to enable upgrade_mode ML setting or not. Defaults to false.
- **timeout** – Controls the time to wait before action times out. Defaults to 30 seconds

start_datafeed(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-start-datafeed.html

Parameters

- **datafeed_id** – The ID of the datafeed to start
- **body** – The start datafeed parameters
- **end** – The end time when the datafeed should stop. When not set, the datafeed continues in real time
- **start** – The start time from where the datafeed should begin
- **timeout** – Controls the time to wait until a datafeed has started. Default to 20 seconds

stop_datafeed(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-stop-datafeed.html

Parameters

- **datafeed_id** – The ID of the datafeed to stop
- **allow_no_datafeeds** – Whether to ignore if a wildcard expression matches no datafeeds. (This includes *_all* string or when no datafeeds have been specified)
- **force** – True if the datafeed should be forcefully stopped.
- **timeout** – Controls the time to wait until a datafeed has stopped. Default to 20 seconds

update_datafeed(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-update-datafeed.html

Parameters

- **datafeed_id** – The ID of the datafeed to update
- **body** – The datafeed update settings

update_filter(***kwargs*)
'<>'_

Parameters

- **filter_id** – The ID of the filter to update
- **body** – The filter update

update_job(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-update-job.html

Parameters

- **job_id** – The ID of the job to create
- **body** – The job update settings

update_model_snapshot(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/ml-update-snapshot.html

Parameters

- **job_id** – The ID of the job to fetch
- **snapshot_id** – The ID of the snapshot to update
- **body** – The model snapshot properties to update

validate(***kwargs*)
'<>'_

Parameters **body** – The job config

validate_detector(***kwargs*)
'<>'_

Parameters **body** – The detector

7.2.5 Security APIs

Security API can be used to help secure your Elasticsearch cluster. Integrating with LDAP and Active Directory.

class elasticsearch.client.xpack.security.**SecurityClient**(*client*)

authenticate(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/security-api-authenticate.html

change_password(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/security-api-change-password.html

Parameters

- **body** – the new password for the user
- **username** – The username of the user to change the password for
- **refresh** – If *true* (the default) then refresh the affected 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* then do nothing with refreshes., valid choices are: 'true', 'false', 'wait_for'

clear_cached_realms(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/security-api-clear-cache.html

Parameters

- **realms** – Comma-separated list of realms to clear
- **usernames** – Comma-separated list of usernames to clear from the cache

clear_cached_roles(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/security-api-clear-role-cache.html

Parameters **name** – Role name

create_api_key(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/security-api-create-api-key.html

Parameters

- **body** – The api key request to create an API key
- **refresh** – If *true* (the default) then refresh the affected 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* then do nothing with refreshes., valid choices are: 'true', 'false', 'wait_for'

delete_privileges(***kwargs*)
TODO

Parameters

- **application** – Application name
- **name** – Privilege name
- **refresh** – If *true* (the default) then refresh the affected 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* then do nothing with refreshes., valid choices are: 'true', 'false', 'wait_for'

delete_role(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/security-api-delete-role.html

Parameters

- **name** – Role name
- **refresh** – If *true* (the default) then refresh the affected 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* then do nothing with refreshes., valid choices are: 'true', 'false', 'wait_for'

delete_role_mapping(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/security-api-delete-role-mapping.html

Parameters

- **name** – Role-mapping name
- **refresh** – If *true* (the default) then refresh the affected 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* then do nothing with refreshes., valid choices are: 'true', 'false', 'wait_for'

delete_user(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/security-api-delete-user.html

Parameters

- **username** – username

- **refresh** – If *true* (the default) then refresh the affected 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* then do nothing with refreshes., valid choices are: 'true', 'false', 'wait_for'

disable_user(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/security-api-disable-user.html

Parameters

- **username** – The username of the user to disable
- **refresh** – If *true* (the default) then refresh the affected 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* then do nothing with refreshes., valid choices are: 'true', 'false', 'wait_for'

enable_user(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/security-api-enable-user.html

Parameters

- **username** – The username of the user to enable
- **refresh** – If *true* (the default) then refresh the affected 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* then do nothing with refreshes., valid choices are: 'true', 'false', 'wait_for'

get_api_key(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/security-api-get-api-key.html

Parameters

- **id** – API key id of the API key to be retrieved
- **name** – API key name of the API key to be retrieved
- **realm_name** – realm name of the user who created this API key to be retrieved
- **username** – user name of the user who created this API key to be retrieved

get_privileges(***kwargs*)
TODO

Parameters

- **application** – Application name
- **name** – Privilege name

get_role(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/security-api-get-role.html

Parameters **name** – Role name

get_role_mapping(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/security-api-get-role-mapping.html

Parameters **name** – Role-Mapping name

get_token(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/security-api-get-token.html

Parameters **body** – The token request to get

get_user(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/security-api-get-user.html

Parameters **username** – A comma-separated list of usernames

get_user_privileges(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/security-api-get-user-privileges.html

has_privileges(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/security-api-has-privileges.html

Parameters

- **body** – The privileges to test
- **user** – Username

invalidate_api_key(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/security-api-invalidate-api-key.html

Parameters **body** – The api key request to invalidate API key(s)

invalidate_token(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/security-api-invalidate-token.html

Parameters **body** – The token to invalidate

put_privileges(***kwargs*)
TODO

Parameters

- **body** – The privilege(s) to add
- **refresh** – If *true* (the default) then refresh the affected 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* then do nothing with refreshes., valid choices are: 'true', 'false', 'wait_for'

put_role(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/security-api-put-role.html

Parameters

- **name** – Role name
- **body** – The role to add
- **refresh** – If *true* (the default) then refresh the affected 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* then do nothing with refreshes., valid choices are: 'true', 'false', 'wait_for'

put_role_mapping(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/security-api-put-role-mapping.html

Parameters

- **name** – Role-mapping name
- **body** – The role to add
- **refresh** – If *true* (the default) then refresh the affected 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* then do nothing with refreshes., valid choices are: 'true', 'false', 'wait_for'

put_user(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/security-api-put-user.html

Parameters

- **username** – The username of the User
- **body** – The user to add

- **refresh** – If *true* (the default) then refresh the affected 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* then do nothing with refreshes., valid choices are: 'true', 'false', 'wait_for'

7.2.6 Watcher APIs

Watcher API can be used to notify you when certain pre-defined thresholds have happened.

class elasticsearch.client.xpack.watcher.**WatcherClient**(*client*)

ack_watch(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/watcher-api-ack-watch.html

Parameters

- **watch_id** – Watch ID
- **action_id** – A comma-separated list of the action ids to be acked

activate_watch(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/watcher-api-activate-watch.html

Parameters **watch_id** – Watch ID

deactivate_watch(***kwargs*)
https://www.elastic.co/guide/en/elasticsearch/reference/current/watcher-api-deactivate-watch.html

Parameters **watch_id** – Watch ID

delete_watch(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/watcher-api-delete-watch.html

Parameters **id** – Watch ID

execute_watch(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/watcher-api-execute-watch.html

Parameters

- **id** – Watch ID
- **body** – Execution control
- **debug** – indicates whether the watch should execute in debug mode

get_watch(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/watcher-api-get-watch.html

Parameters **id** – Watch ID

put_watch(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/watcher-api-put-watch.html

Parameters

- **id** – Watch ID
- **body** – The watch
- **active** – Specify whether the watch is in/active by default
- **if_primary_term** – only update the watch if the last operation that has changed the watch has the specified primary term

- **if_seq_no** – only update the watch if the last operation that has changed the watch has the specified sequence number
- **version** – Explicit version number for concurrency control

start(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/watcher-api-start.html

stats(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/watcher-api-stats.html

Parameters

- **metric** – Controls what additional stat metrics should be include in the response
- **emit_stacktraces** – Emits stack traces of currently running watches

stop(***kwargs*)
http://www.elastic.co/guide/en/elasticsearch/reference/current/watcher-api-stop.html

7.2.7 Migration APIs

Migration API helps simplify upgrading X-Pack indices from one version to another.

class elasticsearch.client.xpack.migration.**MigrationClient**(*client*)

deprecations(***kwargs*)
http://www.elastic.co/guide/en/migration/current/migration-api-deprecation.html

Parameters **index** – Index pattern

7.3 Exceptions

class elasticsearch.**ImproperlyConfigured**
Exception raised when the config passed to the client is inconsistent or invalid.

class elasticsearch.**ElasticsearchException**
Base class for all exceptions raised by this package's operations (doesn't apply to *ImproperlyConfigured*).

class elasticsearch.**SerializationError**(*ElasticsearchException*)
Data passed in failed to serialize properly in the `Serializer` being used.

class elasticsearch.**TransportError**(*ElasticsearchException*)
Exception raised when ES returns a non-OK (>=400) HTTP status code. Or when an actual connection error happens; in that case the `status_code` will be set to `'N/A'`.

error
A string error message.

info
Dict of returned error info from ES, where available, underlying exception when not.

status_code
The HTTP status code of the response that precipitated the error or `'N/A'` if not applicable.

class elasticsearch.**ConnectionError**(*TransportError*)
Error raised when there was an exception while talking to ES. Original exception from the underlying `Connection` implementation is available as `.info`.

class elasticsearch.**ConnectionTimeout**(*ConnectionError*)
A network timeout. Doesn't cause a node retry by default.

class elasticsearch.**SSLError**(*ConnectionError*)
Error raised when encountering SSL errors.

class elasticsearch.**NotFoundError**(*TransportError*)
Exception representing a 404 status code.

class elasticsearch.**ConflictError**(*TransportError*)
Exception representing a 409 status code.

class elasticsearch.**RequestError**(*TransportError*)
Exception representing a 400 status code.

7.4 Connection Layer API

All of the classes responsible for handling the connection to the Elasticsearch cluster. The default subclasses used can be overriden by passing parameters to the *Elasticsearch* class. All of the arguments to the client will be passed on to *Transport*, *ConnectionPool* and Connection.

For example if you wanted to use your own implementation of the *ConnectionSelector* class you can just pass in the selector_class parameter.

Note: *ConnectionPool* and related options (like selector_class) will only be used if more than one connection is defined. Either directly or via the *Sniffing* mechanism.

7.4.1 Transport

class elasticsearch.**Transport**(*hosts*, *connection_class=Urllib3HttpConnection*, *connection_pool_class=ConnectionPool*, *host_info_callback=construct_hosts_list*, *sniff_on_start=False*, *sniffer_timeout=None*, *sniff_on_connection_fail=False*, *serializer=JSONSerializer()*, *max_retries=3*, *** kwargs*)
Encapsulation of transport-related to logic. Handles instantiation of the individual connections as well as creating a connection pool to hold them.

Main interface is the *perform_request* method.

Parameters

- **hosts** – list of dictionaries, each containing keyword arguments to create a *connection_class* instance
- **connection_class** – subclass of Connection to use
- **connection_pool_class** – subclass of *ConnectionPool* to use
- **host_info_callback** – callback responsible for taking the node information from */_cluser/nodes*, along with already extracted information, and producing a list of arguments (same as *hosts* parameter)
- **sniff_on_start** – flag indicating whether to obtain a list of nodes from the cluser at startup time
- **sniffer_timeout** – number of seconds between automatic sniffs

- **sniff_on_connection_fail** – flag controlling if connection failure triggers a sniff
- **sniff_timeout** – timeout used for the sniff request - it should be a fast api call and we are talking potentially to more nodes so we want to fail quickly. Not used during initial sniffing (if `sniff_on_start` is on) when the connection still isn't initialized.
- **serializer** – serializer instance
- **serializers** – optional dict of serializer instances that will be used for deserializing data coming from the server. (key is the mimetype)
- **default_mimetype** – when no mimetype is specified by the server response assume this mimetype, defaults to *'application/json'*
- **max_retries** – maximum number of retries before an exception is propagated
- **retry_on_status** – set of HTTP status codes on which we should retry on a different node. defaults to `(502, 503, 504)`
- **retry_on_timeout** – should timeout trigger a retry on different node? (default *False*)
- **send_get_body_as** – for GET requests with body this option allows you to specify an alternate way of execution for environments that don't support passing bodies with GET requests. If you set this to 'POST' a POST method will be used instead, if to 'source' then the body will be serialized and passed as a query parameter *source*.

Any extra keyword arguments will be passed to the *connection_class* when creating and instance unless overridden by that connection's options provided as part of the hosts parameter.

add_connection(*host*)
Create a new `Connection` instance and add it to the pool.

Parameters **host** – kwargs that will be used to create the instance

close()
Explicitly closes connections

get_connection()
Retreive a `Connection` instance from the *ConnectionPool* instance.

mark_dead(*connection*)
Mark a connection as dead (failed) in the connection pool. If sniffing on failure is enabled this will initiate the sniffing process.

Parameters **connection** – instance of `Connection` that failed

perform_request(*method*, *url*, *headers=None*, *params=None*, *body=None*)
Perform the actual request. Retrieve a connection from the connection pool, pass all the information to it's perform_request method and return the data.

If an exception was raised, mark the connection as failed and retry (up to *max_retries* times).

If the operation was succesful and the connection used was previously marked as dead, mark it as live, resetting it's failure count.

Parameters

- **method** – HTTP method to use
- **url** – absolute url (without host) to target
- **headers** – dictionary of headers, will be handed over to the underlying `Connection` class

- **params** – dictionary of query parameters, will be handed over to the underlying `Connection` class for serialization
- **body** – body of the request, will be serializes using serializer and passed to the connection

set_connections(*hosts*)
Instantiate all the connections and create new connection pool to hold them. Tries to identify unchanged hosts and re-use existing `Connection` instances.

Parameters **hosts** – same as *__init__*

sniff_hosts(*initial=False*)
Obtain a list of nodes from the cluster and create a new connection pool using the information retrieved.

To extract the node connection parameters use the `nodes_to_host_callback`.

Parameters **initial** – flag indicating if this is during startup (`sniff_on_start`), ignore the `sniff_timeout` if `True`

7.4.2 Connection Pool

class elasticsearch.**ConnectionPool**(*connections*, *dead_timeout=60*, *selector_class=RoundRobinSelector*, *randomize_hosts=True*, *** kwargs*)
Container holding the `Connection` instances, managing the selection process (via a *ConnectionSelector*) and dead connections.

It's only interactions are with the *Transport* class that drives all the actions within *ConnectionPool*.

Initially connections are stored on the class as a list and, along with the connection options, get passed to the *ConnectionSelector* instance for future reference.

Upon each request the *Transport* will ask for a *Connection* via the *get_connection* method. If the connection fails (it's *perform_request* raises a *ConnectionError*) it will be marked as dead (via *mark_dead*) and put on a timeout (if it fails N times in a row the timeout is exponentially longer - the formula is *default_timeout * 2 ** (fail_count - 1)*). When the timeout is over the connection will be resurrected and returned to the live pool. A connection that has been previously marked as dead and succeeds will be marked as live (its fail count will be deleted).

Parameters

- **connections** – list of tuples containing the `Connection` instance and it's options
- **dead_timeout** – number of seconds a connection should be retired for after a failure, increases on consecutive failures
- **timeout_cutoff** – number of consecutive failures after which the timeout doesn't increase
- **selector_class** – *ConnectionSelector* subclass to use if more than one connection is live
- **randomize_hosts** – shuffle the list of connections upon arrival to avoid dog piling effect across processes

close()
Explicitly closes connections

get_connection()
Return a connection from the pool using the *ConnectionSelector* instance.

It tries to resurrect eligible connections, forces a resurrection when no connections are availible and passes the list of live connections to the selector instance to choose from.

Returns a connection instance and it's current fail count.

mark_dead(*connection*, *now=None*)
Mark the connection as dead (failed). Remove it from the live pool and put it on a timeout.

Parameters **connection** – the failed instance

mark_live(*connection*)
Mark connection as healthy after a resurrection. Resets the fail counter for the connection.

Parameters **connection** – the connection to redeem

resurrect(*force=False*)
Attempt to resurrect a connection from the dead pool. It will try to locate one (not all) eligible (it's timeout is over) connection to return to the live pool. Any resurrected connection is also returned.

Parameters **force** – resurrect a connection even if there is none eligible (used when we have no live connections). If force is specified resurrect always returns a connection.

7.4.3 Connection Selector

class elasticsearch.**ConnectionSelector**(*opts*)
Simple class used to select a connection from a list of currently live connection instances. In init time it is passed a dictionary containing all the connections' options which it can then use during the selection process. When the *select* method is called it is given a list of *currently* live connections to choose from.

The options dictionary is the one that has been passed to *Transport* as *hosts* param and the same that is used to construct the Connection object itself. When the Connection was created from information retrieved from the cluster via the sniffing process it will be the dictionary returned by the *host_info_callback*.

Example of where this would be useful is a zone-aware selector that would only select connections from it's own zones and only fall back to other connections where there would be none in it's zones.

Parameters **opts** – dictionary of connection instances and their options

select(*connections*)
Select a connection from the given list.

Parameters **connections** – list of live connections to choose from

7.4.4 Urllib3HttpConnection (default connection_class)

If you have complex SSL logic for connecting to Elasticsearch using an *SSLContext* object might be more helpful. You can create one natively using the python SSL library with the *create_default_context* (https://docs.python.org/3/library/ssl.html#ssl.create_default_context) method.

To create an *SSLContext* object you only need to use one of cafile, capath or cadata:

```
>>> from ssl import create_default_context
>>> context = create_default_context(cafile=None, capath=None, cadata=None)
```

- *cafile* is the path to your CA File
- *capath* is the directory of a collection of CA's
- *cadata* is either an ASCII string of one or more PEM-encoded certificates or a bytes-like object of DER-encoded certificates.

Please note that the use of SSLContext is only available for Urllib3.

class elasticsearch.**Urllib3HttpConnection**(*host='localhost'*, *port=9200*, *http_auth=None*, *use_ssl=False*, *verify_certs=None*, *ssl_show_warn=True*, *ca_certs=None*, *client_cert=None*, *client_key=None*, *ssl_version=None*, *ssl_assert_hostname=None*, *ssl_assert_fingerprint=None*, *maxsize=10*, *headers=None*, *ssl_context=None*, *http_compress=False*, ***kwargs*)

Default connection class using the *urllib3* library and the http protocol.

Parameters

- **host** – hostname of the node (default: localhost)
- **port** – port to use (integer, default: 9200)
- **url_prefix** – optional url prefix for elasticsearch
- **timeout** – default timeout in seconds (float, default: 10)
- **http_auth** – optional http auth information as either ':' separated string or a tuple
- **use_ssl** – use ssl for the connection if *True*
- **verify_certs** – whether to verify SSL certificates
- **ssl_show_warn** – show warning when verify certs is disabled
- **ca_certs** – optional path to CA bundle. See https://urllib3.readthedocs.io/en/latest/security.html#using-certifi-with-urllib3 for instructions how to get default set
- **client_cert** – path to the file containing the private key and the certificate, or cert only if using client_key
- **client_key** – path to the file containing the private key if using separate cert and key files (client_cert will contain only the cert)
- **ssl_version** – version of the SSL protocol to use. Choices are: SSLv23 (default) SSLv2 SSLv3 TLSv1 (see `PROTOCOL_*` constants in the `ssl` module for exact options for your environment).
- **ssl_assert_hostname** – use hostname verification if not *False*
- **ssl_assert_fingerprint** – verify the supplied certificate fingerprint if not *None*
- **maxsize** – the number of connections which will be kept open to this host. See https://urllib3.readthedocs.io/en/1.4/pools.html#api for more information.
- **headers** – any custom http headers to be add to requests
- **http_compress** – Use gzip compression

close()

Explicitly closes connection

7.5 Transport classes

List of transport classes that can be used, simply import your choice and pass it to the constructor of *Elasticsearch* as *connection_class*. Note that the *RequestsHttpConnection* requires `requests` to be installed.

For example to use the `requests`-based connection just import it and use it:

```
from elasticsearch import Elasticsearch, RequestsHttpConnection
es = Elasticsearch(connection_class=RequestsHttpConnection)
```

The default connection class is based on `urllib3` which is more performant and lightweight than the optional `requests`-based class. Only use `RequestsHttpConnection` if you have need of any of `requests` advanced features like custom auth plugins etc.

7.5.1 Connection

class elasticsearch.connection.**Connection**(*host='localhost'*, *port=9200*, *use_ssl=False*, *url_prefix=''*, *timeout=10*, ***kwargs*)

Class responsible for maintaining a connection to an Elasticsearch node. It holds persistent connection pool to it and it's main interface (*perform_request*) is thread-safe.

Also responsible for logging.

Parameters

- **host** – hostname of the node (default: localhost)
- **port** – port to use (integer, default: 9200)
- **url_prefix** – optional url prefix for elasticsearch
- **timeout** – default timeout in seconds (float, default: 10)

7.5.2 Urllib3HttpConnection

class elasticsearch.connection.**Urllib3HttpConnection**(*host='localhost'*, *port=9200*, *http_auth=None*, *use_ssl=False*, *verify_certs=None*, *ssl_show_warn=True*, *ca_certs=None*, *client_cert=None*, *client_key=None*, *ssl_version=None*, *ssl_assert_hostname=None*, *ssl_assert_fingerprint=None*, *maxsize=10*, *headers=None*, *ssl_context=None*, *http_compress=False*, ***kwargs*)

Default connection class using the *urllib3* library and the http protocol.

Parameters

- **host** – hostname of the node (default: localhost)
- **port** – port to use (integer, default: 9200)
- **url_prefix** – optional url prefix for elasticsearch
- **timeout** – default timeout in seconds (float, default: 10)
- **http_auth** – optional http auth information as either ':' separated string or a tuple
- **use_ssl** – use ssl for the connection if *True*

- **verify_certs** – whether to verify SSL certificates
- **ssl_show_warn** – show warning when verify certs is disabled
- **ca_certs** – optional path to CA bundle. See https://urllib3.readthedocs.io/en/latest/security.html#using-certifi-with-urllib3 for instructions how to get default set
- **client_cert** – path to the file containing the private key and the certificate, or cert only if using client_key
- **client_key** – path to the file containing the private key if using separate cert and key files (client_cert will contain only the cert)
- **ssl_version** – version of the SSL protocol to use. Choices are: SSLv23 (default) SSLv2 SSLv3 TLSv1 (see `PROTOCOL_*` constants in the `ssl` module for exact options for your environment).
- **ssl_assert_hostname** – use hostname verification if not *False*
- **ssl_assert_fingerprint** – verify the supplied certificate fingerprint if not *None*
- **maxsize** – the number of connections which will be kept open to this host. See https://urllib3.readthedocs.io/en/1.4/pools.html#api for more information.
- **headers** – any custom http headers to be add to requests
- **http_compress** – Use gzip compression

7.5.3 RequestsHttpConnection

class elasticsearch.connection.**RequestsHttpConnection**(*host='localhost'*, *port=9200*, *http_auth=None*, *use_ssl=False*, *verify_certs=True*, *ssl_show_warn=True*, *ca_certs=None*, *client_cert=None*, *client_key=None*, *headers=None*, ***kwargs*)

Connection using the *requests* library.

Parameters

- **http_auth** – optional http auth information as either ':' separated string or a tuple. Any value will be passed into requests as *auth*.
- **use_ssl** – use ssl for the connection if *True*
- **verify_certs** – whether to verify SSL certificates
- **ssl_show_warn** – show warning when verify certs is disabled
- **ca_certs** – optional path to CA bundle. By default standard requests' bundle will be used.
- **client_cert** – path to the file containing the private key and the certificate, or cert only if using client_key
- **client_key** – path to the file containing the private key if using separate cert and key files (client_cert will contain only the cert)
- **headers** – any custom http headers to be add to requests

7.6 Helpers

Collection of simple helper functions that abstract some specifics or the raw API.

7.6.1 Bulk helpers

There are several helpers for the `bulk` API since its requirement for specific formatting and other considerations can make it cumbersome if used directly.

All bulk helpers accept an instance of `Elasticsearch` class and an iterable `actions` (any iterable, can also be a generator, which is ideal in most cases since it will allow you to index large datasets without the need of loading them into memory).

The items in the `action` iterable should be the documents we wish to index in several formats. The most common one is the same as returned by *`search()`*, for example:

```
{
    '_index': 'index-name',
    '_type': 'document',
    '_id': 42,
    '_routing': 5,
    'pipeline': 'my-ingest-pipeline',
    '_source': {
        "title": "Hello World!",
        "body": "..."
    }
}
```

Alternatively, if *_source* is not present, it will pop all metadata fields from the doc and use the rest as the document data:

```
{
    "_id": 42,
    "_routing": 5,
    "title": "Hello World!",
    "body": "..."
}
```

The *`bulk()`* api accepts `index`, `create`, `delete`, and `update` actions. Use the `_op_type` field to specify an action (`_op_type` defaults to `index`):

```
{
    '_op_type': 'delete',
    '_index': 'index-name',
    '_type': 'document',
    '_id': 42,
}
{
    '_op_type': 'update',
    '_index': 'index-name',
    '_type': 'document',
    '_id': 42,
    'doc': {'question': 'The life, universe and everything.'}
}
```

Example:

Lets say we have an iterable of data. Lets say a list of words called `mywords` and we want to index those words into individual documents where the structure of the document is like `{"word": "<myword>"}`.

```
def gendata():
    mywords = ['foo', 'bar', 'baz']
    for word in mywords:
        yield {
            "_index": "mywords",
            "_type": "document",
            "doc": {"word": word},
        }

bulk(es, gendata())
```

For a more complete and complex example please take a look at https://github.com/elastic/elasticsearch-py/blob/master/example/load.py#L76-L130

Note: When reading raw json strings from a file, you can also pass them in directly (without decoding to dicts first). In that case, however, you lose the ability to specify anything (index, type, even id) on a per-record basis, all documents will just be sent to elasticsearch to be indexed as-is.

---

elasticsearch.helpers.**streaming_bulk**(*client*, *actions*, *chunk_size=500*, *max_chunk_bytes=104857600*, *raise_on_error=True*, *expand_action_callback=<function expand_action>*, *raise_on_exception=True*, *max_retries=0*, *initial_backoff=2*, *max_backoff=600*, *yield_ok=True*, **args*, ***kwargs*)

Streaming bulk consumes actions from the iterable passed in and yields results per action. For non-streaming usecases use *bulk()* which is a wrapper around streaming bulk that returns summary information about the bulk operation once the entire input is consumed and sent.

If you specify `max_retries` it will also retry any documents that were rejected with a `429` status code. To do this it will wait (**by calling time.sleep which will block**) for `initial_backoff` seconds and then, every subsequent rejection for the same chunk, for double the time every time up to `max_backoff` seconds.

Parameters

- **client** – instance of *Elasticsearch* to use
- **actions** – iterable containing the actions to be executed
- **chunk_size** – number of docs in one chunk sent to es (default: 500)
- **max_chunk_bytes** – the maximum size of the request in bytes (default: 100MB)
- **raise_on_error** – raise `BulkIndexError` containing errors (as *.errors*) from the execution of the last chunk when some occur. By default we raise.
- **raise_on_exception** – if `False` then don't propagate exceptions from call to `bulk` and just report the items that failed as failed.
- **expand_action_callback** – callback executed on each action passed in, should return a tuple containing the action line and the data line (*None* if data line should be omitted).
- **max_retries** – maximum number of times a document will be retried when `429` is received, set to 0 (default) for no retries on `429`

- **initial_backoff** – number of seconds we should wait before the first retry. Any subsequent retries will be powers of `initial_backoff * 2**retry_number`
- **max_backoff** – maximum number of seconds a retry will wait
- **yield_ok** – if set to False will skip successful documents in the output

elasticsearch.helpers.**parallel_bulk**(*client*, *actions*, *thread_count=4*, *chunk_size=500*, *max_chunk_bytes=104857600*, *queue_size=4*, *expand_action_callback=<function expand_action>*, **args*, ***kwargs*)

Parallel version of the bulk helper run in multiple threads at once.

Parameters

- **client** – instance of *Elasticsearch* to use
- **actions** – iterator containing the actions
- **thread_count** – size of the threadpool to use for the bulk requests
- **chunk_size** – number of docs in one chunk sent to es (default: 500)
- **max_chunk_bytes** – the maximum size of the request in bytes (default: 100MB)
- **raise_on_error** – raise `BulkIndexError` containing errors (as *.errors*) from the execution of the last chunk when some occur. By default we raise.
- **raise_on_exception** – if `False` then don't propagate exceptions from call to `bulk` and just report the items that failed as failed.
- **expand_action_callback** – callback executed on each action passed in, should return a tuple containing the action line and the data line (*None* if data line should be omitted).
- **queue_size** – size of the task queue between the main thread (producing chunks to send) and the processing threads.

elasticsearch.helpers.**bulk**(*client*, *actions*, *stats_only=False*, **args*, ***kwargs*)

Helper for the *bulk()* api that provides a more human friendly interface - it consumes an iterator of actions and sends them to elasticsearch in chunks. It returns a tuple with summary information - number of successfully executed actions and either list of errors or number of errors if `stats_only` is set to `True`. Note that by default we raise a `BulkIndexError` when we encounter an error so options like `stats_only` only apply when `raise_on_error` is set to `False`.

When errors are being collected original document data is included in the error dictionary which can lead to an extra high memory usage. If you need to process a lot of data and want to ignore/collect errors please consider using the *streaming_bulk()* helper which will just return the errors and not store them in memory.

Parameters

- **client** – instance of *Elasticsearch* to use
- **actions** – iterator containing the actions
- **stats_only** – if *True* only report number of successful/failed operations instead of just number of successful and a list of error responses

Any additional keyword arguments will be passed to *streaming_bulk()* which is used to execute the operation, see *streaming_bulk()* for more accepted parameters.

7.6.2 Scan

elasticsearch.helpers.**scan**(*client*, *query=None*, *scroll='5m'*, *raise_on_error=True*, *preserve_order=False*, *size=1000*, *request_timeout=None*, *clear_scroll=True*, *scroll_kwargs=None*, ***kwargs*)

Simple abstraction on top of the *scroll()* api - a simple iterator that yields all hits as returned by underlining scroll requests.

By default scan does not return results in any pre-determined order. To have a standard order in the returned documents (either by score or explicit sort definition) when scrolling, use `preserve_order=True`. This may be an expensive operation and will negate the performance benefits of using `scan`.

Parameters

- **client** – instance of *Elasticsearch* to use
- **query** – body for the *search()* api
- **scroll** – Specify how long a consistent view of the index should be maintained for scrolled search
- **raise_on_error** – raises an exception (`ScanError`) if an error is encountered (some shards fail to execute). By default we raise.
- **preserve_order** – don't set the `search_type` to `scan` - this will cause the scroll to paginate with preserving the order. Note that this can be an extremely expensive operation and can easily lead to unpredictable results, use with caution.
- **size** – size (per shard) of the batch send at each iteration.
- **request_timeout** – explicit timeout for each call to `scan`
- **clear_scroll** – explicitly calls delete on the scroll id via the clear scroll API at the end of the method on completion or error, defaults to true.
- **scroll_kwargs** – additional kwargs to be passed to *scroll()*

Any additional keyword arguments will be passed to the initial *search()* call:

```
scan(es,
    query={"query": {"match": {"title": "python"}}},
    index="orders-*",
    doc_type="books"
)
```

7.6.3 Reindex

elasticsearch.helpers.**reindex**(*client*, *source_index*, *target_index*, *query=None*, *target_client=None*, *chunk_size=500*, *scroll='5m'*, *scan_kwargs={}*, *bulk_kwargs={}*)

Reindex all documents from one index that satisfy a given query to another, potentially (if *target_client* is specified) on a different cluster. If you don't specify the query you will reindex all the documents.

Since `2.3` a *reindex()* api is available as part of elasticsearch itself. It is recommended to use the api instead of this helper wherever possible. The helper is here mostly for backwards compatibility and for situations where more flexibility is needed.

Note: This helper doesn't transfer mappings, just the data.

Parameters

- **client** – instance of *Elasticsearch* to use (for read if *target_client* is specified as well)
- **source_index** – index (or list of indices) to read documents from
- **target_index** – name of the index in the target cluster to populate
- **query** – body for the *search()* api
- **target_client** – optional, is specified will be used for writing (thus enabling reindex between clusters)
- **chunk_size** – number of docs in one chunk sent to es (default: 500)
- **scroll** – Specify how long a consistent view of the index should be maintained for scrolled search
- **scan_kwargs** – additional kwargs to be passed to *scan()*
- **bulk_kwargs** – additional kwargs to be passed to *bulk()*

7.7 Changelog

7.7.1 7.1.0 (dev)

7.7.2 7.0.0 (2019-04-11)

- Removed deprecated option `update_all_types`.
- Using insecure SSL configuration (`verify_cert=False`) raises a warning, this can be not showed with `ssl_show_warn=False`
- Add support for 7.x api's in Elasticsearch both xpack and oss flavors
- relative import for x-pack.
- include_type_name to pertinent indices operations
- fix python 3.x string decode exception
- better unicode chunking

6.3.0 (2018-06-20)

- Add an exponential wait on delays
- Fix issues with dependencies
- Adding X-pack Docs
- Adding forecast to x-pack ML client

7.7.3 6.2.0 (2018-03-20)

- cleanup for SSL Context
- Add X-Pack clients to -py

- Adding Gzip support for capacity constrained networks
- `_routing` in bulk action has been deprecated in ES. Introduces a breaking change if you use `routing` as a field in your documents.

7.7.4 6.1.1 (2018-01-05)

- Updates to SSLContext logic to make it easier to use and have saner defaults.
- Doc updates

7.7.5 6.1.0 (2018-01-05)

- bad release

7.7.6 6.0.0 (2017-11-14)

- compatibility with Elasticsearch 6.0.0

7.7.7 5.5.0 (2017-11-10)

- `streaming_bulk` helper now supports retries with incremental backoff
- `scan` helper properly checks for successful shards instead of just checking `failed`
- compatible release with elasticsearch `5.6.4`
- fix handling of UTF-8 surrogates

7.7.8 5.4.0 (2017-05-18)

- `bulk` helpers now extract `pipeline` parameter from the action dictionary.

7.7.9 5.3.0 (2017-03-30)

Compatibility with elasticsearch 5.3

7.7.10 5.2.0 (2017-02-12)

The client now automatically sends `Content-Type` http header set to `application/json`. If you are explicitly passing in other encoding than `json` you need to set the header manually.

7.7.11 5.1.0 (2017-01-11)

- Fixed sniffing

7.7.12 5.0.1 (2016-11-02)

Fixed performance regression in `scan` helper

7.7.13 5.0.0 (2016-10-19)

Version compatible with elasticsearch 5.0

- when using SSL certificate validation is now on by default. Install `certifi` or supply root certificate bundle.
- `elasticsearch.trace` logger now also logs failed requests, signature of internal logging method `log_request_fail` has changed, all custom connection classes need to be updated
- added `headers` arg to connections to support custom http headers
- passing in a keyword parameter with `None` as value will cause that param to be ignored

7.7.14 2.4.0 (2016-08-17)

- `ping` now ignores all `TransportError` exceptions and just returns `False`
- expose `scroll_id` on `ScanError`
- increase default size for `scan` helper to 1000

Internal:

- changed `Transport.perform_request` to just return the body, not status as well.

7.7.15 2.3.0 (2016-02-29)

- added `client_key` argument to configure client certificates
- debug logging now includes response body even for failed requests

7.7.16 2.2.0 (2016-01-05)

Due to change in json encoding the client will no longer mask issues with encoding - if you work with non-ascii data in python 2 you must use the `unicode` type or have proper encoding set in your environment.

- adding additional options for ssh - `ssl_assert_hostname` and `ssl_assert_fingerprint` to the default connection class
- fix sniffing

7.7.17 2.1.0 (2015-10-19)

- move multiprocessing import inside parallel bulk for Google App Engine

7.7.18 2.0.0 (2015-10-14)

- Elasticsearch 2.0 compatibility release

7.7.19 1.8.0 (2015-10-14)

- removed thrift and memcached connections, if you wish to continue using those, extract the classes and use them separately.
- added a new, parallel version of the bulk helper using thread pools
- In helpers, removed `bulk_index` as an alias for `bulk`. Use `bulk` instead.

7.7.20 1.7.0 (2015-09-21)

- elasticsearch 2.0 compatibility
- thrift now deprecated, to be removed in future version
- make sure urllib3 always uses keep-alive

7.7.21 1.6.0 (2015-06-10)

- Add `indices.flush_synced` API
- `helpers.reindex` now supports reindexing parent/child documents

7.7.22 1.5.0 (2015-05-18)

- Add support for `query_cache` parameter when searching
- helpers have been made more secure by changing defaults to raise an exception on errors
- removed deprecated options `replication` and the deprecated benchmark api.
- Added `AddonClient` class to allow for extending the client from outside

7.7.23 1.4.0 (2015-02-11)

- Using insecure SSL configuration (`verify_cert=False`) raises a warning
- `reindex` accepts a `query` parameter
- enable `reindex` helper to accept any kwargs for underlying `bulk` and `scan` calls
- when doing an initial sniff (via `sniff_on_start`) ignore special sniff timeout
- option to treat `TransportError` as normal failure in `bulk` helpers
- fixed an issue with sniffing when only a single host was passed in

7.7.24 1.3.0 (2014-12-31)

- Timeout now doesn't trigger a retry by default (can be overriden by setting `retry_on_timeout=True`)
- Introduced new parameter `retry_on_status` (defaulting to `(503, 504, )`) controls which http status code should lead to a retry.
- Implemented url parsing according to RFC-1738
- Added support for proper SSL certificate handling

- Required parameters are now checked for non-empty values
- ConnectionPool now checks if any connections were defined
- DummyConnectionPool introduced when no load balancing is needed (only one connection defined)
- Fixed a race condition in ConnectionPool

7.7.25 1.2.0 (2014-08-03)

Compatibility with newest (1.3) Elasticsearch APIs.

- Filter out master-only nodes when sniffing
- Improved docs and error messages

7.7.26 1.1.1 (2014-07-04)

Bugfix release fixing escaping issues with `request_timeout`.

7.7.27 1.1.0 (2014-07-02)

Compatibility with newest Elasticsearch APIs.

- Test helpers - `ElasticsearchTestCase` and `get_test_client` for use in your tests
- Python 3.2 compatibility
- Use `simplejson` if installed instead of stdlib json library
- Introducing a global `request_timeout` parameter for per-call timeout
- Bug fixes

7.7.28 1.0.0 (2014-02-11)

Elasticsearch 1.0 compatibility. See 0.4.X releases (and 0.4 branch) for code compatible with 0.90 elasticsearch.

- major breaking change - compatible with 1.0 elasticsearch releases only!
- Add an option to change the timeout used for sniff requests (`sniff_timeout`).
- empty responses from the server are now returned as empty strings instead of None
- `get_alias` now has `name` as another optional parameter due to issue #4539 in es repo. Note that the order of params have changed so if you are not using keyword arguments this is a breaking change.

7.7.29 0.4.4 (2013-12-23)

- `helpers.bulk_index` renamed to `helpers.bulk` (alias put in place for backwards compatibility, to be removed in future versions)
- Added `helpers.streaming_bulk` to consume an iterator and yield results per operation
- `helpers.bulk` and `helpers.streaming_bulk` are no longer limited to just index operations.
- unicode body (for `incices.analyze` for example) is now handled correctly

- changed `perform_request` on `Connection` classes to return headers as well. This is a backwards incompatible change for people who have developed their own connection class.
- changed deserialization mechanics. Users who provided their own serializer that didn't extend `JSONSerializer` need to specify a `mimetype` class attribute.
- minor bug fixes

7.7.30 0.4.3 (2013-10-22)

- Fixes to `helpers.bulk_index`, better error handling
- More benevolent `hosts` argument parsing for `Elasticsearch`
- `requests` no longer required (nor recommended) for install

7.7.31 0.4.2 (2013-10-08)

- `ignore` param accepted by all APIs
- Fixes to `helpers.bulk_index`

7.7.32 0.4.1 (2013-09-24)

Initial release.

CHAPTER 8

License

CHAPTER 9

Indices and tables

- genindex
- modindex
- search

Python Module Index

e

Index

E

F

G

H

I

L

M

N

O

P

R

S

T

U

V

W

X