Settings

The Scrapy settings allows you to customize the behaviour of all Scrapy components, including the core, extensions, pipelines and spiders themselves.

The infrastructure of the settings provides a global namespace of key-value mappings that the code can use to pull configuration values from. The settings can be populated through different mechanisms, which are described below.

The settings are also the mechanism for selecting the currently active Scrapy project (in case you have many).

For a list of available built-in settings see: :ref:`topics-settings-ref`.

> **System Message: ERROR/3 (**`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`**, line 17);** *backlink*
>
> Unknown interpreted text role "ref".

Designating the settings

When you use Scrapy, you have to tell it which settings you're using. You can do this by using an environment variable, `SCRAPY_SETTINGS_MODULE`.

The value of `SCRAPY_SETTINGS_MODULE` should be in Python path syntax, e.g. `myproject.settings`. Note that the settings module should be on the Python :ref:`import search path <tut-searchpath>`.

> **System Message: ERROR/3 (**`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`**, line 27);** *backlink*
>
> Unknown interpreted text role "ref".

Populating the settings

Settings can be populated using different mechanisms, each of which having a different precedence. Here is the list of them in decreasing order of precedence:

1. Command line options (most precedence)
2. Settings per-spider
3. Project settings module
4. Default settings per-command
5. Default global settings (less precedence)

The population of these settings sources is taken care of internally, but a manual handling is possible using API calls. See the :ref:`topics-api-settings` topic for reference.

> **System Message: ERROR/3 (**`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`**, line 46);** *backlink*
>
> Unknown interpreted text role "ref".

These mechanisms are described in more detail below.

1. Command line options

Arguments provided by the command line are the ones that take most precedence, overriding any other options. You can explicitly override one (or more) settings using the `-s` (or `--set`) command line option.

> **System Message: ERROR/3 (**`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`**, line 59)**
>
> Unknown directive type "highlight".
>
> ```
>     .. highlight:: sh
> ```

Example:

```
scrapy crawl myspider -s LOG_FILE=scrapy.log
```

2. Settings per-spider

Spiders (See the :ref:`topics-spiders` chapter for reference) can define their own settings that will take precedence and override the project ones. They can do so by setting their :attr:`~scrapy.Spider.custom_settings` attribute:

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 68); *backlink*
>
> Unknown interpreted text role "ref".

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 68); *backlink*
>
> Unknown interpreted text role "attr".

```
class MySpider(scrapy.Spider):
    name = 'myspider'

    custom_settings = {
        'SOME_SETTING': 'some value',
    }
```

3. Project settings module

The project settings module is the standard configuration file for your Scrapy project, it's where most of your custom settings will be populated. For a standard Scrapy project, this means you'll be adding or changing the settings in the `settings.py` file created for your project.

4. Default settings per-command

Each :doc:`Scrapy tool </topics/commands>` command can have its own default settings, which override the global default settings. Those custom command settings are specified in the `default_settings` attribute of the command class.

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 90); *backlink*
>
> Unknown interpreted text role "doc".

5. Default global settings

The global defaults are located in the `scrapy.settings.default_settings` module and documented in the :ref:`topics-settings-ref` section.

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 98); *backlink*
>
> Unknown interpreted text role "ref".

Import paths and classes

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 105)
>
> Unknown directive type "versionadded".
>
> ```
> .. versionadded:: 2.4.0
> ```

When a setting references a callable object to be imported by Scrapy, such as a class or a function, there are two different ways you can specify that object:

- As a string containing the import path of that object
- As the object itself

For example:

```
from mybot.pipelines.validate import ValidateMyItem
ITEM_PIPELINES = {
    # passing the classname...
    ValidateMyItem: 300,
    # ...equals passing the class path
```

```
    'mybot.pipelines.validate.ValidateMyItem': 300,
}
```

> **Note**
>
> Passing non-callable objects is not supported.

How to access settings

In a spider, the settings are available through `self.settings`:

```
class MySpider(scrapy.Spider):
    name = 'myspider'
    start_urls = ['http://example.com']

    def parse(self, response):
        print(f"Existing settings: {self.settings.attributes.keys()}")
```

> **Note**
>
> The `settings` attribute is set in the base Spider class after the spider is initialized. If you want to use the settings before the initialization (e.g., in your spider's `__init__()` method), you'll need to override the :meth:`~scrapy.Spider.from_crawler` method.
>
> >

Settings can be accessed through the :attr:`scrapy.crawler.Crawler.settings` attribute of the Crawler that is passed to `from_crawler` method in extensions, middlewares and item pipelines:

```
class MyExtension:
    def __init__(self, log_is_enabled=False):
        if log_is_enabled:
            print("log is enabled!")

    @classmethod
    def from_crawler(cls, crawler):
        settings = crawler.settings
        return cls(settings.getbool('LOG_ENABLED'))
```

The settings object can be used like a dict (e.g., `settings['LOG_ENABLED']`), but it's usually preferred to extract the setting in the format you need it to avoid type errors, using one of the methods provided by the :class:`~scrapy.settings.Settings` API.

Rationale for setting names

Setting names are usually prefixed with the component that they configure. For example, proper setting names for a fictional robots.txt extension would be `ROBOTSTXT_ENABLED`, `ROBOTSTXT_OBEY`, `ROBOTSTXT_CACHEDIR`, etc.

Built-in settings reference

Here's a list of all available Scrapy settings, in alphabetical order, along with their default values and the scope where they apply.

The scope, where available, shows where the setting is being used, if it's tied to any particular component. In that case the module of that component will be shown, typically an extension, middleware or pipeline. It also means that the component must be enabled in order for the setting to have any effect.

> **System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 187)**
>
> Unknown directive type "setting".
>
>   .. setting:: AWS_ACCESS_KEY_ID

AWS_ACCESS_KEY_ID

Default: `None`

The AWS access key used by code that requires access to Amazon Web services, such as the :ref:`S3 feed storage backend <topics-feed-storage-s3>`.

> **System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 194); *backlink***
>
> Unknown interpreted text role "ref".

> **System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 197)**
>
> Unknown directive type "setting".
>
>   .. setting:: AWS_SECRET_ACCESS_KEY

AWS_SECRET_ACCESS_KEY

Default: `None`

The AWS secret key used by code that requires access to Amazon Web services, such as the :ref:`S3 feed storage backend <topics-feed-storage-s3>`.

> **System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 204); *backlink***
>
> Unknown interpreted text role "ref".

> **System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 207)**
>
> Unknown directive type "setting".
>
>   .. setting:: AWS_SESSION_TOKEN

AWS_SESSION_TOKEN

Default: `None`

The AWS security token used by code that requires access to Amazon Web services, such as the :ref:`S3 feed storage backend <topics-feed-storage-s3>`, when using temporary security credentials.

> **System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 214); *backlink***
>
> Unknown interpreted text role "ref".

> **System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-**

Unknown directive type "setting".

```
.. setting:: AWS_ENDPOINT_URL
```

AWS_ENDPOINT_URL

Default: `None`

Endpoint URL used for S3-like storage, for example Minio or s3.scality.

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst,` **line 229)**

Unknown directive type "setting".

```
.. setting:: AWS_USE_SSL
```

AWS_USE_SSL

Default: `None`

Use this option if you want to disable SSL connection for communication with S3 or S3-like storage. By default SSL will be used.

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst,` **line 239)**

Unknown directive type "setting".

```
.. setting:: AWS_VERIFY
```

AWS_VERIFY

Default: `None`

Verify SSL connection between Scrapy and S3 or S3-like storage. By default SSL verification will occur.

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst,` **line 249)**

Unknown directive type "setting".

```
.. setting:: AWS_REGION_NAME
```

AWS_REGION_NAME

Default: `None`

The name of the region associated with the AWS client.

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst,` **line 258)**

Unknown directive type "setting".

```
.. setting:: ASYNCIO_EVENT_LOOP
```

ASYNCIO_EVENT_LOOP

Default: `None`

Import path of a given `asyncio` event loop class.

If the asyncio reactor is enabled (see :setting:`TWISTED_REACTOR`) this setting can be used to specify the asyncio event loop to be used with it. Set the setting to the import path of the desired asyncio event loop class. If the setting is set to `None` the default asyncio event loop will be used.

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst,` **line 267);** *backlink*

Unknown interpreted text role "setting".

If you are installing the asyncio reactor manually using the :func:`~scrapy.utils.reactor.install_reactor` function, you can use the `event_loop_path` parameter to indicate the import path of the event loop class to be used.

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst, line 272`); *backlink*
>
> Unknown interpreted text role "func".

Note that the event loop class must inherit from :class:`asyncio.AbstractEventLoop`.

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst, line 276`); *backlink*
>
> Unknown interpreted text role "class".

> **Caution!**
>
> Please be aware that, when using a non-default event loop (either defined via :setting:`ASYNCIO_EVENT_LOOP` or installed with :func:`~scrapy.utils.reactor.install_reactor`), Scrapy will call :func:`asyncio.set_event_loop`, which will set the specified event loop as the current loop for the current OS thread.
>
> > **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst, line 278`); *backlink*
> >
> > Unknown interpreted text role "setting".
>
> > **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst, line 278`); *backlink*
> >
> > Unknown interpreted text role "func".
>
> > **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst, line 278`); *backlink*
> >
> > Unknown interpreted text role "func".

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst, line 284`)
>
> Unknown directive type "setting".
>
> ```
> .. setting:: BOT_NAME
> ```

BOT_NAME

Default: `'scrapybot'`

The name of the bot implemented by this Scrapy project (also known as the project name). This name will be used for the logging too.

It's automatically populated with your project name when you create your project with the :command:`startproject` command.

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst, line 294`); *backlink*
>
> Unknown interpreted text role "command".

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst, line 297`)

Unknown directive type "setting".

```
.. setting:: CONCURRENT_ITEMS
```

CONCURRENT_ITEMS

Default: `100`

Maximum number of concurrent items (per response) to process in parallel in :ref:`item pipelines <topics-item-pipeline>`.

CONCURRENT_REQUESTS

Default: `16`

The maximum number of concurrent (i.e. simultaneous) requests that will be performed by the Scrapy downloader.

CONCURRENT_REQUESTS_PER_DOMAIN

Default: `8`

The maximum number of concurrent (i.e. simultaneous) requests that will be performed to any single domain.

See also: :ref:`topics-autothrottle` and its :setting:`AUTOTHROTTLE_TARGET_CONCURRENCY` option.

CONCURRENT_REQUESTS_PER_IP

Default: `0`

The maximum number of concurrent (i.e. simultaneous) requests that will be performed to any single IP. If non-zero, the :setting:`CONCURRENT_REQUESTS_PER_DOMAIN` setting is ignored, and this one is used instead. In other words, concurrency limits will be applied per IP, not per domain.

This setting also affects :setting:`DOWNLOAD_DELAY` and :ref:`topics-autothrottle`: if :setting:`CONCURRENT_REQUESTS_PER_IP` is non-zero, download delay is enforced per IP, not per domain.

DEFAULT_ITEM_CLASS

Default: `'scrapy.Item'`

The default class that will be used for instantiating items in the :ref:`the Scrapy shell <topics-shell>`.

DEFAULT_REQUEST_HEADERS

Default:

```
{
    'Accept': 'text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8',
    'Accept-Language': 'en',
}
```

The default headers used for Scrapy HTTP Requests. They're populated in the :class:`~scrapy.downloadermiddlewares.defaultheaders.DefaultHeadersMiddleware`.

Caution!

Cookies set via the `Cookie` header are not considered by the :ref:`cookies-mw`. If you need to set cookies for a request, use the :class:`Request.cookies <scrapy.Request>` parameter. This is a known current limitation that is being worked on.

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 374);** *backlink*
>
> Unknown interpreted text role "ref".

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 374);** *backlink*
>
> Unknown interpreted text role "class".

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 379)**
>
> Unknown directive type "setting".
>
> ```
>     .. setting:: DEPTH_LIMIT
> ```

DEPTH_LIMIT

Default: `0`

Scope: `scrapy.spidermiddlewares.depth.DepthMiddleware`

The maximum depth that will be allowed to crawl for any site. If zero, no limit will be imposed.

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 391)**
>
> Unknown directive type "setting".
>
> ```
>     .. setting:: DEPTH_PRIORITY
> ```

DEPTH_PRIORITY

Default: `0`

Scope: `scrapy.spidermiddlewares.depth.DepthMiddleware`

An integer that is used to adjust the :attr:`~scrapy.Request.priority` of a :class:`~scrapy.Request` based on its depth.

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 400);** *backlink*
>
> Unknown interpreted text role "attr".

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 400);** *backlink*
>
> Unknown interpreted text role "class".

The priority of a request is adjusted as follows:

```
request.priority = request.priority - ( depth * DEPTH_PRIORITY )
```

As depth increases, positive values of `DEPTH_PRIORITY` decrease request priority (BFO), while negative values increase request priority (DFO). See also :ref:`faq-bfo-dfo`.

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 407);** *backlink*
>
> Unknown interpreted text role "ref".

Note

This setting adjusts priority **in the opposite way** compared to other priority settings
:setting:`REDIRECT_PRIORITY_ADJUST` and :setting:`RETRY_PRIORITY_ADJUST`.

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-`
> `resources\scrapy-master\docs\topics\[scrapy-master][docs]`
> `[topics]settings.rst`, **line 413**); *backlink*
>
> Unknown interpreted text role "setting".

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-`
> `resources\scrapy-master\docs\topics\[scrapy-master][docs]`
> `[topics]settings.rst`, **line 413**); *backlink*
>
> Unknown interpreted text role "setting".

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-`
> `master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 417**)
>
> Unknown directive type "setting".
>
> ```
>     .. setting:: DEPTH_STATS_VERBOSE
> ```

DEPTH_STATS_VERBOSE

Default: `False`

Scope: `scrapy.spidermiddlewares.depth.DepthMiddleware`

Whether to collect verbose depth stats. If this is enabled, the number of requests for each depth is collected in the stats.

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-`
> `master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 429**)
>
> Unknown directive type "setting".
>
> ```
>     .. setting:: DNSCACHE_ENABLED
> ```

DNSCACHE_ENABLED

Default: `True`

Whether to enable DNS in-memory cache.

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-`
> `master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 438**)
>
> Unknown directive type "setting".
>
> ```
>     .. setting:: DNSCACHE_SIZE
> ```

DNSCACHE_SIZE

Default: `10000`

DNS in-memory cache size.

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-`
> `master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 447**)
>
> Unknown directive type "setting".
>
> ```
>     .. setting:: DNS_RESOLVER
> ```

DNS_RESOLVER

Default: `'scrapy.resolver.CachingThreadedResolver'`

The class to be used to resolve DNS names. The default `scrapy.resolver.CachingThreadedResolver` supports specifying a timeout for DNS requests via the :setting:`DNS_TIMEOUT` setting, but works only with IPv4 addresses. Scrapy provides an alternative resolver, `scrapy.resolver.CachingHostnameResolver`, which supports IPv4/IPv6 addresses but does not take the :setting:`DNS_TIMEOUT` setting into account.

DNS_TIMEOUT

Default: `60`

Timeout for processing of DNS queries in seconds. Float is supported.

DOWNLOADER

Default: `'scrapy.core.downloader.Downloader'`

The downloader to use for crawling.

DOWNLOADER_HTTPCLIENTFACTORY

Default: `'scrapy.core.downloader.webclient.ScrapyHTTPClientFactory'`

Defines a Twisted `protocol.ClientFactory` class to use for HTTP/1.0 connections (for `HTTP10DownloadHandler`).

> **Note**
>
> HTTP/1.0 is rarely used nowadays so you can safely ignore this setting, unless you really want to use HTTP/1.0 and override :setting:`DOWNLOAD_HANDLERS` for `http(s)` scheme accordingly, i.e. to `'scrapy.core.downloader.handlers.http.HTTP10DownloadHandler'`.

DOWNLOADER_CLIENTCONTEXTFACTORY

Default: `'scrapy.core.downloader.contextfactory.ScrapyClientContextFactory'`

Represents the classpath to the ContextFactory to use.

Here, "ContextFactory" is a Twisted term for SSL/TLS contexts, defining the TLS/SSL protocol version to use, whether to do certificate verification, or even enable client-side authentication (and various other things).

> **Note**
>
> Scrapy default context factory **does NOT perform remote server certificate verification**. This is usually fine for web scraping.
>
> If you do need remote server certificate verification enabled, Scrapy also has another context factory class that you can set, `'scrapy.core.downloader.contextfactory.BrowserLikeContextFactory'`, which uses the platform's certificates to validate remote endpoints.

If you do use a custom ContextFactory, make sure its `__init__` method accepts a `method` parameter (this is the `OpenSSL.SSL` method mapping :setting:`DOWNLOADER_CLIENT_TLS_METHOD`), a `tls_verbose_logging` parameter (`bool`) and a `tls_ciphers` parameter (see :setting:`DOWNLOADER_CLIENT_TLS_CIPHERS`).

DOWNLOADER_CLIENT_TLS_CIPHERS

Default: `'DEFAULT'`

Use this setting to customize the TLS/SSL ciphers used by the default HTTP/1.1 downloader.

The setting should contain a string in the [OpenSSL cipher list format](), these ciphers will be used as client ciphers. Changing this setting may be necessary to access certain HTTPS websites: for example, you may need to use `'DEFAULT:!DH'` for a website with weak DH parameters or enable a specific cipher that is not included in `DEFAULT` if a website requires it.

DOWNLOADER_CLIENT_TLS_METHOD

Default: `'TLS'`

Use this setting to customize the TLS/SSL method used by the default HTTP/1.1 downloader.

This setting must be one of these string values:

- `'TLS'`: maps to OpenSSL's `TLS_method()` (a.k.a `SSLv23_method()`), which allows protocol negotiation, starting from the highest supported by the platform; **default, recommended**
- `'TLSv1.0'`: this value forces HTTPS connections to use TLS version 1.0 ; set this if you want the behavior of Scrapy<1.1
- `'TLSv1.1'`: forces TLS version 1.1
- `'TLSv1.2'`: forces TLS version 1.2
- `'SSLv3'`: forces SSL version 3 (**not recommended**)

> **System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst, line 566)**
>
> Unknown directive type "setting".
>
> ```
>     .. setting:: DOWNLOADER_CLIENT_TLS_VERBOSE_LOGGING
> ```

DOWNLOADER_CLIENT_TLS_VERBOSE_LOGGING

Default: `False`

Setting this to `True` will enable DEBUG level messages about TLS connection parameters after establishing HTTPS connections. The kind of information logged depends on the versions of OpenSSL and pyOpenSSL.

This setting is only used for the default :setting:`DOWNLOADER_CLIENTCONTEXTFACTORY`.

> **System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst, line 577); *backlink***
>
> Unknown interpreted text role "setting".

> **System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst, line 580)**
>
> Unknown directive type "setting".
>
> ```
>     .. setting:: DOWNLOADER_MIDDLEWARES
> ```

DOWNLOADER_MIDDLEWARES

Default:: `{}`

A dict containing the downloader middlewares enabled in your project, and their orders. For more info see :ref:`topics-downloader-middleware-setting`.

> **System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst, line 587); *backlink***
>
> Unknown interpreted text role "ref".

> **System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst, line 590)**
>
> Unknown directive type "setting".
>
> ```
>     .. setting:: DOWNLOADER_MIDDLEWARES_BASE
> ```

DOWNLOADER_MIDDLEWARES_BASE

Default:

```
{
    'scrapy.downloadermiddlewares.robotstxt.RobotsTxtMiddleware': 100,
    'scrapy.downloadermiddlewares.httpauth.HttpAuthMiddleware': 300,
```

```
    'scrapy.downloadermiddlewares.downloadtimeout.DownloadTimeoutMiddleware': 350,
    'scrapy.downloadermiddlewares.defaultheaders.DefaultHeadersMiddleware': 400,
    'scrapy.downloadermiddlewares.useragent.UserAgentMiddleware': 500,
    'scrapy.downloadermiddlewares.retry.RetryMiddleware': 550,
    'scrapy.downloadermiddlewares.ajaxcrawl.AjaxCrawlMiddleware': 560,
    'scrapy.downloadermiddlewares.redirect.MetaRefreshMiddleware': 580,
    'scrapy.downloadermiddlewares.httpcompression.HttpCompressionMiddleware': 590,
    'scrapy.downloadermiddlewares.redirect.RedirectMiddleware': 600,
    'scrapy.downloadermiddlewares.cookies.CookiesMiddleware': 700,
    'scrapy.downloadermiddlewares.httpproxy.HttpProxyMiddleware': 750,
    'scrapy.downloadermiddlewares.stats.DownloaderStats': 850,
    'scrapy.downloadermiddlewares.httpcache.HttpCacheMiddleware': 900,
}
```

A dict containing the downloader middlewares enabled by default in Scrapy. Low orders are closer to the engine, high orders are closer to the downloader. You should never modify this setting in your project, modify :setting:`DOWNLOADER_MIDDLEWARES` instead. For more info see :ref:`topics-downloader-middleware-setting`.

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 614**); *backlink*

Unknown interpreted text role "setting".

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 614**); *backlink*

Unknown interpreted text role "ref".

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 620**)

Unknown directive type "setting".

```
.. setting:: DOWNLOADER_STATS
```

DOWNLOADER_STATS

Default: `True`

Whether to enable downloader stats collection.

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 629**)

Unknown directive type "setting".

```
.. setting:: DOWNLOAD_DELAY
```

DOWNLOAD_DELAY

Default: `0`

The amount of time (in secs) that the downloader should wait before downloading consecutive pages from the same website. This can be used to throttle the crawling speed to avoid hitting servers too hard. Decimal numbers are supported. Example:

```
DOWNLOAD_DELAY = 0.25    # 250 ms of delay
```

This setting is also affected by the :setting:`RANDOMIZE_DOWNLOAD_DELAY` setting (which is enabled by default). By default, Scrapy doesn't wait a fixed amount of time between requests, but uses a random interval between 0.5 * :setting:`DOWNLOAD_DELAY` and 1.5 * :setting:`DOWNLOAD_DELAY`.

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 643**); *backlink*

Unknown interpreted text role "setting".

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 643**); *backlink*

Unknown interpreted text role "setting".

When :setting:`CONCURRENT_REQUESTS_PER_IP` is non-zero, delays are enforced per ip address instead of per domain.

You can also change this setting per spider by setting `download_delay` spider attribute.

DOWNLOAD_HANDLERS

Default: `{}`

A dict containing the request downloader handlers enabled in your project. See :setting:`DOWNLOAD_HANDLERS_BASE` for example format.

DOWNLOAD_HANDLERS_BASE

Default:

```
{
    'data': 'scrapy.core.downloader.handlers.datauri.DataURIDownloadHandler',
    'file': 'scrapy.core.downloader.handlers.file.FileDownloadHandler',
    'http': 'scrapy.core.downloader.handlers.http.HTTPDownloadHandler',
    'https': 'scrapy.core.downloader.handlers.http.HTTPDownloadHandler',
    's3': 'scrapy.core.downloader.handlers.s3.S3DownloadHandler',
    'ftp': 'scrapy.core.downloader.handlers.ftp.FTPDownloadHandler',
}
```

A dict containing the request download handlers enabled by default in Scrapy. You should never modify this setting in your project, modify :setting:`DOWNLOAD_HANDLERS` instead.

You can disable any of these download handlers by assigning `None` to their URI scheme in :setting:`DOWNLOAD_HANDLERS`. E.g., to disable the built-in FTP handler (without replacement), place this in your `settings.py`:

```
DOWNLOAD_HANDLERS = {
    'ftp': None,
}
```

The default HTTPS handler uses HTTP/1.1. To use HTTP/2:

1. Install `Twisted[http2]>=17.9.0` to install the packages required to enable HTTP/2 support in Twisted.

2. Update :setting:`DOWNLOAD_HANDLERS` as follows:

```
DOWNLOAD_HANDLERS = {
    'https': 'scrapy.core.downloader.handlers.http2.H2DownloadHandler',
}
```

> **Warning**
>
> HTTP/2 support in Scrapy is experimental, and not yet recommended for production environments. Future Scrapy versions may introduce related changes without a deprecation period or warning.

> **Note**
>
> Known limitations of the current HTTP/2 implementation of Scrapy include:
>
> - No support for HTTP/2 Cleartext (h2c), since no major browser supports HTTP/2 unencrypted (refer http2 faq).
> - No setting to specify a maximum frame size larger than the default value, 16384. Connections to servers that send a larger frame will fail.
> - No support for server pushes, which are ignored.
> - No support for the :signal:`bytes_received` and :signal:`headers_received` signals.
>
>

DOWNLOAD_TIMEOUT

Default: `180`

The amount of time (in secs) that the downloader will wait before timing out.

> **Note**
>
> This timeout can be set per spider using :attr:`download_timeout` spider attribute and per-request using :reqmeta:`download_timeout` Request.meta key.
>

DOWNLOAD_MAXSIZE

Default: `1073741824` (1024MB)

The maximum response size (in bytes) that downloader will download.

If you want to disable it set to 0.

Note

This size can be set per spider using :attr:`download_maxsize` spider attribute and per-request using
:reqmeta:`download_maxsize` Request.meta key.

DOWNLOAD_WARNSIZE

Default: `33554432` (32MB)

The response size (in bytes) that downloader will start to warn.

If you want to disable it set to 0.

Note

This size can be set per spider using :attr:`download_warnsize` spider attribute and per-request using :reqmeta:`download_warnsize` Request.meta key.

DOWNLOAD_FAIL_ON_DATALOSS

Default: `True`

Whether or not to fail on broken responses, that is, declared `Content-Length` does not match content sent by the server or chunked response was not properly finish. If `True`, these responses raise a `ResponseFailed([_DataLoss])` error. If `False`, these responses are passed through and the flag `dataloss` is added to the response, i.e.: `'dataloss' in response.flags` is `True`.

Optionally, this can be set per-request basis by using the :reqmeta:`download_fail_on_dataloss` Request.meta key to `False`.

Note

A broken response, or data loss error, may happen under several circumstances, from server misconfiguration to network errors to data corruption. It is up to the user to decide if it makes sense to process broken responses considering they may contain partial or incomplete content. If :setting:`RETRY_ENABLED` is `True` and this setting is set to `True`, the `ResponseFailed([_DataLoss])` failure will be retried as usual.

Warning

This setting is ignored by the :class:`~scrapy.core.downloader.handlers.http2.H2DownloadHandler` download handler (see :setting:`DOWNLOAD_HANDLERS`). In case of a data loss error, the corresponding HTTP/2 connection may be corrupted, affecting other requests that use the same connection; hence, a `ResponseFailed([InvalidBodyLengthError])` failure is always raised for every request that was using that connection.

Unknown interpreted text role "class".

DUPEFILTER_CLASS

Default: `'scrapy.dupefilters.RFPDupeFilter'`

The class used to detect and filter duplicate requests.

The default (`RFPDupeFilter`) filters based on request fingerprint using the `scrapy.utils.request.request_fingerprint` function. In order to change the way duplicates are checked you could subclass `RFPDupeFilter` and override its `request_fingerprint` method. This method should accept scrapy :class:`~scrapy.Request` object and return its fingerprint (a string).

You can disable filtering of duplicate requests by setting :setting:`DUPEFILTER_CLASS` to `'scrapy.dupefilters.BaseDupeFilter'`. Be very careful about this however, because you can get into crawling loops. It's usually a better idea to set the `dont_filter` parameter to `True` on the specific :class:`~scrapy.Request` that should not be filtered.

DUPEFILTER_DEBUG

Default: `False`

By default, `RFPDupeFilter` only logs the first duplicate request. Setting :setting:`DUPEFILTER_DEBUG` to `True` will make it log all duplicate requests.

EDITOR

Default: `vi` (on Unix systems) or the IDLE editor (on Windows)

The editor to use for editing spiders with the :command:\`edit\` command. Additionally, if the `EDITOR` environment variable is set, the :command:\`edit\` command will prefer it over the default setting.

EXTENSIONS

Default:: `{}`

A dict containing the extensions enabled in your project, and their orders.

EXTENSIONS_BASE

Default:

```
{
    'scrapy.extensions.corestats.CoreStats': 0,
    'scrapy.extensions.telnet.TelnetConsole': 0,
    'scrapy.extensions.memusage.MemoryUsage': 0,
    'scrapy.extensions.memdebug.MemoryDebugger': 0,
    'scrapy.extensions.closespider.CloseSpider': 0,
    'scrapy.extensions.feedexport.FeedExporter': 0,
    'scrapy.extensions.logstats.LogStats': 0,
    'scrapy.extensions.spiderstate.SpiderState': 0,
    'scrapy.extensions.throttle.AutoThrottle': 0,
}
```

A dict containing the extensions available by default in Scrapy, and their orders. This setting contains all stable built-in extensions. Keep in mind that some of them need to be enabled through a setting.

For more information See the :ref:\`extensions user guide <topics-extensions>\` and the :ref:\`list of available extensions <topics-extensions-ref>\`.

FEED_TEMPDIR

The Feed Temp dir allows you to set a custom folder to save crawler temporary files before uploading with :ref:`FTP feed storage <topics-feed-storage-ftp>` and :ref:`Amazon S3 <topics-feed-storage-s3>`.

FEED_STORAGE_GCS_ACL

The Access Control List (ACL) used when storing items to :ref:`Google Cloud Storage <topics-feed-storage-gcs>`. For more information on how to set this value, please refer to the column *JSON API* in Google Cloud documentation.

FTP_PASSIVE_MODE

Default: `True`

Whether or not to use passive mode when initiating FTP transfers.

master\docs\topics\[scrapy-master][docs][topics]settings.rst, line 926)

Unknown directive type "setting".

```
.. setting:: FTP_PASSWORD
```

FTP_PASSWORD

Default: `"guest"`

The password to use for FTP connections when there is no "`ftp_password`" in `Request` meta.

> **Note**
>
> Paraphrasing RFC 1635, although it is common to use either the password "guest" or one's e-mail address for anonymous FTP, some FTP servers explicitly ask for the user's e-mail address and will not allow login with the "guest" password.

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst, line 944)

Unknown directive type "reqmeta".

```
.. reqmeta:: ftp_user
```

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst, line 945)

Unknown directive type "setting".

```
.. setting:: FTP_USER
```

FTP_USER

Default: `"anonymous"`

The username to use for FTP connections when there is no "`ftp_user`" in `Request` meta.

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst, line 955)

Unknown directive type "setting".

```
.. setting:: GCS_PROJECT_ID
```

GCS_PROJECT_ID

Default: `None`

The Project ID that will be used when storing data on Google Cloud Storage.

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst, line 964)

Unknown directive type "setting".

```
.. setting:: ITEM_PIPELINES
```

ITEM_PIPELINES

Default: `{}`

A dict containing the item pipelines to use, and their orders. Order values are arbitrary, but it is customary to define them in the 0-1000 range. Lower orders process before higher orders.

Example:

```
ITEM_PIPELINES = {
    'mybot.pipelines.validate.ValidateMyItem': 300,
    'mybot.pipelines.validate.StoreMyItem': 800,
```

```
    }
```

<div style="border:1px solid">

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 982**)

Unknown directive type "setting".

```
    .. setting:: ITEM_PIPELINES_BASE
```

</div>

ITEM_PIPELINES_BASE

Default: `{}`

A dict containing the pipelines enabled by default in Scrapy. You should never modify this setting in your project, modify :setting:`ITEM_PIPELINES` instead.

<div style="border:1px solid">

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 989**); *backlink*

Unknown interpreted text role "setting".

</div>

<div style="border:1px solid">

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 992**)

Unknown directive type "setting".

```
    .. setting:: JOBDIR
```

</div>

JOBDIR

Default: `''`

A string indicating the directory for storing the state of a crawl when :ref:`pausing and resuming crawls <topics-jobs>`.

<div style="border:1px solid">

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 999**); *backlink*

Unknown interpreted text role "ref".

</div>

<div style="border:1px solid">

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 1002**)

Unknown directive type "setting".

```
    .. setting:: LOG_ENABLED
```

</div>

LOG_ENABLED

Default: `True`

Whether to enable logging.

<div style="border:1px solid">

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 1011**)

Unknown directive type "setting".

```
    .. setting:: LOG_ENCODING
```

</div>

LOG_ENCODING

Default: `'utf-8'`

The encoding to use for logging.

<div style="border:1px solid">

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 1020**)

Unknown directive type "setting".

</div>

```
       .. setting:: LOG_FILE
```

LOG_FILE

Default: `None`

File name to use for logging output. If `None`, standard error will be used.

LOG_FILE_APPEND

Default: `True`

If `False`, the log file specified with :setting:`LOG_FILE` will be overwritten (discarding the output from previous runs, if any).

LOG_FORMAT

Default: `'%(asctime)s [%(name)s] %(levelname)s: %(message)s'`

String for formatting log messages. Refer to the :ref:`Python logging documentation <logrecord-attributes>` for the qwhole list of available placeholders.

LOG_DATEFORMAT

Default: `'%Y-%m-%d %H:%M:%S'`

String for formatting date/time, expansion of the `%(asctime)s` placeholder in :setting:`LOG_FORMAT`. Refer to the :ref:`Python datetime documentation <strftime-strptime-behavior>` for the whole list of available directives.

LOG_FORMATTER

Default: :class:`scrapy.logformatter.LogFormatter`

The class to use for :ref:`formatting log messages <custom-log-formats>` for different actions.

LOG_LEVEL

Default: `'DEBUG'`

Minimum level to log. Available levels are: CRITICAL, ERROR, WARNING, INFO, DEBUG. For more info see :ref:`topics-logging`.

LOG_STDOUT

Default: `False`

If `True`, all standard output (and error) of your process will be redirected to the log. For example if you `print('hello')` it will appear in the Scrapy log.

LOG_SHORT_NAMES

Default: `False`

If `True`, the logs will just contain the root path. If it is set to `False` then it displays the component responsible for the log output

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 1102**)
>
> Unknown directive type "setting".
>
> ```
>     .. setting:: LOGSTATS_INTERVAL
> ```

LOGSTATS_INTERVAL

Default: `60.0`

The interval (in seconds) between each logging printout of the stats by :class:`~scrapy.extensions.logstats.LogStats`.

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 1109**); *backlink*
>
> Unknown interpreted text role "class".

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 1112**)
>
> Unknown directive type "setting".
>
> ```
>     .. setting:: MEMDEBUG_ENABLED
> ```

MEMDEBUG_ENABLED

Default: `False`

Whether to enable memory debugging.

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 1121**)
>
> Unknown directive type "setting".
>
> ```
>     .. setting:: MEMDEBUG_NOTIFY
> ```

MEMDEBUG_NOTIFY

Default: `[]`

When memory debugging is enabled a memory report will be sent to the specified addresses if this setting is not empty, otherwise the report will be written to the log.

Example:

```
MEMDEBUG_NOTIFY = ['user@example.com']
```

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 1136**)
>
> Unknown directive type "setting".
>
> ```
>     .. setting:: MEMUSAGE_ENABLED
> ```

MEMUSAGE_ENABLED

Default: `True`

Scope: `scrapy.extensions.memusage`

Whether to enable the memory usage extension. This extension keeps track of a peak memory used by the process (it writes it to stats). It can also optionally shutdown the Scrapy process when it exceeds a memory limit (see :setting:`MEMUSAGE_LIMIT_MB`), and notify by email when that happened (see :setting:`MEMUSAGE_NOTIFY_MAIL`).

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 1145**); *backlink*

See :ref:`topics-extensions-ref-memusage`.

MEMUSAGE_LIMIT_MB

Default: `0`

Scope: `scrapy.extensions.memusage`

The maximum amount of memory to allow (in megabytes) before shutting down Scrapy (if MEMUSAGE_ENABLED is True). If zero, no check will be performed.

See :ref:`topics-extensions-ref-memusage`.

MEMUSAGE_CHECK_INTERVAL_SECONDS

Default: `60.0`

Scope: `scrapy.extensions.memusage`

The :ref:`Memory usage extension <topics-extensions-ref-memusage>` checks the current memory usage, versus the limits set by :setting:`MEMUSAGE_LIMIT_MB` and :setting:`MEMUSAGE_WARNING_MB`, at fixed time intervals.

This sets the length of these intervals, in seconds.

See :ref:`topics-extensions-ref-memusage`.

> **System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 1183); *backlink***
>
> Unknown interpreted text role "ref".

> **System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 1185)**
>
> Unknown directive type "setting".
>
> ```
>     .. setting:: MEMUSAGE_NOTIFY_MAIL
> ```

MEMUSAGE_NOTIFY_MAIL

Default: `False`

Scope: `scrapy.extensions.memusage`

A list of emails to notify if the memory limit has been reached.

Example:

```
MEMUSAGE_NOTIFY_MAIL = ['user@example.com']
```

See :ref:`topics-extensions-ref-memusage`.

> **System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 1200); *backlink***
>
> Unknown interpreted text role "ref".

> **System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 1202)**
>
> Unknown directive type "setting".
>
> ```
>     .. setting:: MEMUSAGE_WARNING_MB
> ```

MEMUSAGE_WARNING_MB

Default: `0`

Scope: `scrapy.extensions.memusage`

The maximum amount of memory to allow (in megabytes) before sending a warning email notifying about it. If zero, no warning will be produced.

> **System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 1214)**
>
> Unknown directive type "setting".
>
> ```
>     .. setting:: NEWSPIDER_MODULE
> ```

NEWSPIDER_MODULE

Default: `''`

Module where to create new spiders using the :command:`genspider` command.

> **System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 1221); *backlink***
>
> Unknown interpreted text role "command".

Example:

```
NEWSPIDER_MODULE = 'mybot.spiders_dev'
```

> **System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 1227)**
>
> Unknown directive type "setting".
>
> ```
>     .. setting:: RANDOMIZE_DOWNLOAD_DELAY
> ```

RANDOMIZE_DOWNLOAD_DELAY

Default: `True`

If enabled, Scrapy will wait a random amount of time (between 0.5 * :setting:`DOWNLOAD_DELAY` and 1.5 * :setting:`DOWNLOAD_DELAY`) while fetching requests from the same website.

> **System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 1234); *backlink***
>
> Unknown interpreted text role "setting".

> **System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 1234); *backlink***
>
> Unknown interpreted text role "setting".

This randomization decreases the chance of the crawler being detected (and subsequently blocked) by sites which analyze requests looking for statistically significant similarities in the time between their requests.

The randomization policy is the same used by wget `--random-wait` option.

If :setting:`DOWNLOAD_DELAY` is zero (default) this option has no effect.

> **System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 1243); *backlink***
>
> Unknown interpreted text role "setting".

> **System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 1247)**
>
> Unknown directive type "setting".
>
> ```
>     .. setting:: REACTOR_THREADPOOL_MAXSIZE
> ```

REACTOR_THREADPOOL_MAXSIZE

Default: `10`

The maximum limit for Twisted Reactor thread pool size. This is common multi-purpose thread pool used by various Scrapy components. Threaded DNS Resolver, BlockingFeedStorage, S3FilesStore just to name a few. Increase this value if you're experiencing problems with insufficient blocking IO.

> **System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 1259)**
>
> Unknown directive type "setting".
>
> ```
>     .. setting:: REDIRECT_PRIORITY_ADJUST
> ```

REDIRECT_PRIORITY_ADJUST

Default: `+2`

Scope: `scrapy.downloadermiddlewares.redirect.RedirectMiddleware`

Adjust redirect request priority relative to original request:

- **a positive priority adjust (default) means higher priority.**
- a negative priority adjust means lower priority.

ROBOTSTXT_OBEY

Default: `False`

Scope: `scrapy.downloadermiddlewares.robotstxt`

If enabled, Scrapy will respect robots.txt policies. For more information see :ref:`topics-dlmw-robots`.

> **Note**
>
> While the default value is `False` for historical reasons, this option is enabled by default in settings.py file generated by `scrapy startproject` command.

ROBOTSTXT_PARSER

Default: `'scrapy.robotstxt.ProtegoRobotParser'`

The parser backend to use for parsing `robots.txt` files. For more information see :ref:`topics-dlmw-robots`.

ROBOTSTXT_USER_AGENT

Default: `None`

The user agent string to use for matching in the robots.txt file. If `None`, the User-Agent header you are sending with the request or the :setting:`USER_AGENT` setting (in that order) will be used for determining the user agent to use in the robots.txt file.

SCHEDULER

Default: `'scrapy.core.scheduler.Scheduler'`

The scheduler class to be used for crawling. See the :ref:`topics-scheduler` topic for details.

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 1320**); *backlink*
>
> Unknown interpreted text role "ref".

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 1323**)
>
> Unknown directive type "setting".
>
> ```
>     .. setting:: SCHEDULER_DEBUG
> ```

SCHEDULER_DEBUG

Default: `False`

Setting to `True` will log debug information about the requests scheduler. This currently logs (only once) if the requests cannot be serialized to disk. Stats counter (`scheduler/unserializable`) tracks the number of times this happens.

Example entry in logs:

```
1956-01-31 00:00:00+0800 [scrapy.core.scheduler] ERROR: Unable to serialize request:
<GET http://example.com> - reason: cannot serialize <Request at 0x9a7c7ec>
(type Request)> - no more unserializable requests will be logged
(see 'scheduler/unserializable' stats counter)
```

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 1342**)
>
> Unknown directive type "setting".
>
> ```
>     .. setting:: SCHEDULER_DISK_QUEUE
> ```

SCHEDULER_DISK_QUEUE

Default: `'scrapy.squeues.PickleLifoDiskQueue'`

Type of disk queue that will be used by scheduler. Other available types are `scrapy.squeues.PickleFifoDiskQueue`, `scrapy.squeues.MarshalFifoDiskQueue`, `scrapy.squeues.MarshalLifoDiskQueue`.

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 1353**)
>
> Unknown directive type "setting".
>
> ```
>     .. setting:: SCHEDULER_MEMORY_QUEUE
> ```

SCHEDULER_MEMORY_QUEUE

Default: `'scrapy.squeues.LifoMemoryQueue'`

Type of in-memory queue used by scheduler. Other available type is: `scrapy.squeues.FifoMemoryQueue`.

> **System Message: ERROR/3** (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 1362**)
>
> Unknown directive type "setting".
>
> ```
>     .. setting:: SCHEDULER_PRIORITY_QUEUE
> ```

SCHEDULER_PRIORITY_QUEUE

Default: `'scrapy.pqueues.ScrapyPriorityQueue'`

Type of priority queue used by the scheduler. Another available type is `scrapy.pqueues.DownloaderAwarePriorityQueue`. `scrapy.pqueues.DownloaderAwarePriorityQueue` works better than `scrapy.pqueues.ScrapyPriorityQueue` when you crawl many different domains in parallel. But currently `scrapy.pqueues.DownloaderAwarePriorityQueue` does not work together with :setting:`CONCURRENT_REQUESTS_PER_IP`.

> **System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 1368); *backlink***
>
> Unknown interpreted text role "setting".

> **System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 1375)**
>
> Unknown directive type "setting".
>
> ```
>     .. setting:: SCRAPER_SLOT_MAX_ACTIVE_SIZE
> ```

SCRAPER_SLOT_MAX_ACTIVE_SIZE

> **System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 1380)**
>
> Unknown directive type "versionadded".
>
> ```
>     .. versionadded:: 2.0
> ```

Default: `5_000_000`

Soft limit (in bytes) for response data being processed.

While the sum of the sizes of all responses being processed is above this value, Scrapy does not process new requests.

> **System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 1389)**
>
> Unknown directive type "setting".
>
> ```
>     .. setting:: SPIDER_CONTRACTS
> ```

SPIDER_CONTRACTS

Default:: `{}`

A dict containing the spider contracts enabled in your project, used for testing spiders. For more info see :ref:`topics-contracts`.

> **System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 1396); *backlink***
>
> Unknown interpreted text role "ref".

> **System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 1399)**
>
> Unknown directive type "setting".
>
> ```
>     .. setting:: SPIDER_CONTRACTS_BASE
> ```

SPIDER_CONTRACTS_BASE

Default:

```
{
    'scrapy.contracts.default.UrlContract' : 1,
    'scrapy.contracts.default.ReturnsContract': 2,
    'scrapy.contracts.default.ScrapesContract': 3,
}
```

A dict containing the Scrapy contracts enabled by default in Scrapy. You should never modify this setting in your project, modify

:setting:`SPIDER_CONTRACTS` instead. For more info see :ref:`topics-contracts`.

You can disable any of these contracts by assigning `None` to their class path in :setting:`SPIDER_CONTRACTS`. E.g., to disable the built-in `ScrapesContract`, place this in your `settings.py`:

```
SPIDER_CONTRACTS = {
    'scrapy.contracts.default.ScrapesContract': None,
}
```

SPIDER_LOADER_CLASS

Default: `'scrapy.spiderloader.SpiderLoader'`

The class that will be used for loading spiders, which must implement the :ref:`topics-api-spiderloader`.

SPIDER_LOADER_WARN_ONLY

Default: `False`

By default, when Scrapy tries to import spider classes from :setting:`SPIDER_MODULES`, it will fail loudly if there is any `ImportError` exception. But you can choose to silence this exception and turn it into a simple warning by setting `SPIDER_LOADER_WARN_ONLY = True`.

> **Note**
>
> Some :ref:`scrapy commands <topics-commands>` run with this setting to `True` already (i.e. they will only issue a warning and will not fail) since they do not actually need to load spider classes to work: :command:`scrapy runspider <runspider>`, :command:`scrapy settings <settings>`, :command:`scrapy startproject <startproject>`, :command:`scrapy version <version>`.

SPIDER_MIDDLEWARES

Default:: `{}`

A dict containing the spider middlewares enabled in your project, and their orders. For more info see :ref:`topics-spider-middleware-setting`.

SPIDER_MIDDLEWARES_BASE

Default:

```
{
    'scrapy.spidermiddlewares.httperror.HttpErrorMiddleware': 50,
    'scrapy.spidermiddlewares.offsite.OffsiteMiddleware': 500,
    'scrapy.spidermiddlewares.referer.RefererMiddleware': 700,
```

```
    'scrapy.spidermiddlewares.urllength.UrlLengthMiddleware': 800,
    'scrapy.spidermiddlewares.depth.DepthMiddleware': 900,
}
```

A dict containing the spider middlewares enabled by default in Scrapy, and their orders. Low orders are closer to the engine, high orders are closer to the spider. For more info see :ref:`topics-spider-middleware-setting`.

SPIDER_MODULES

Default: `[]`

A list of modules where Scrapy will look for spiders.

Example:

```
SPIDER_MODULES = ['mybot.spiders_prod', 'mybot.spiders_dev']
```

STATS_CLASS

Default: `'scrapy.statscollectors.MemoryStatsCollector'`

The class to use for collecting stats, who must implement the :ref:`topics-api-stats`.

STATS_DUMP

Default: `True`

Dump the :ref:`Scrapy stats <topics-stats>` (to the Scrapy log) once the spider finishes.

For more info see: :ref:`topics-stats`.

Unknown interpreted text role "ref".

STATSMAILER_RCPTS

Default: `[]` (empty list)

Send Scrapy stats after spiders finish scraping. See :class:`~scrapy.extensions.statsmailer.StatsMailer` for more info.

TELNETCONSOLE_ENABLED

Default: `True`

A boolean which specifies if the :ref:`telnet console <topics-telnetconsole>` will be enabled (provided its extension is also enabled).

TEMPLATES_DIR

Default: `templates` dir inside scrapy module

The directory where to look for templates when creating new projects with :command:`startproject` command and new spiders with :command:`genspider` command.

The project name must not conflict with the name of custom files or directories in the `project` subdirectory.

Unknown directive type "setting".

```
.. setting:: TWISTED_REACTOR
```

TWISTED_REACTOR

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 1558**)

Unknown directive type "versionadded".

```
.. versionadded:: 2.0
```

Default: `None`

Import path of a given :mod:`~twisted.internet.reactor`.

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 1562**); *backlink*

Unknown interpreted text role "mod".

Scrapy will install this reactor if no other reactor is installed yet, such as when the `scrapy` CLI program is invoked or when using the :class:`~scrapy.crawler.CrawlerProcess` class.

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 1564**); *backlink*

Unknown interpreted text role "class".

If you are using the :class:`~scrapy.crawler.CrawlerRunner` class, you also need to install the correct reactor manually. You can do that using :func:`~scrapy.utils.reactor.install_reactor`:

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 1568**); *backlink*

Unknown interpreted text role "class".

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 1568**); *backlink*

Unknown interpreted text role "func".

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 1572**)

Unknown directive type "autofunction".

```
.. autofunction:: scrapy.utils.reactor.install_reactor
```

If a reactor is already installed, :func:`~scrapy.utils.reactor.install_reactor` has no effect.

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 1574**); *backlink*

Unknown interpreted text role "func".

:meth:`CrawlerRunner.__init__ <scrapy.crawler.CrawlerRunner.__init__>` raises :exc:`Exception` if the installed reactor does not match the :setting:`TWISTED_REACTOR` setting; therefore, having top-level :mod:`~twisted.internet.reactor` imports in project files and imported third-party libraries will make Scrapy raise :exc:`Exception` when it checks which reactor is installed.

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, **line 1577**); *backlink*

Unknown interpreted text role "meth".

In order to use the reactor installed by Scrapy:

```python
import scrapy
from twisted.internet import reactor


class QuotesSpider(scrapy.Spider):
    name = 'quotes'

    def __init__(self, *args, **kwargs):
        self.timeout = int(kwargs.pop('timeout', '60'))
        super(QuotesSpider, self).__init__(*args, **kwargs)

    def start_requests(self):
        reactor.callLater(self.timeout, self.stop)

        urls = ['https://quotes.toscrape.com/page/1']
        for url in urls:
            yield scrapy.Request(url=url, callback=self.parse)

    def parse(self, response):
        for quote in response.css('div.quote'):
            yield {'text': quote.css('span.text::text').get()}

    def stop(self):
        self.crawler.engine.close_spider(self, 'timeout')
```

which raises :exc:`Exception`, becomes:

```python
import scrapy


class QuotesSpider(scrapy.Spider):
    name = 'quotes'

    def __init__(self, *args, **kwargs):
        self.timeout = int(kwargs.pop('timeout', '60'))
        super(QuotesSpider, self).__init__(*args, **kwargs)

    def start_requests(self):
        from twisted.internet import reactor
        reactor.callLater(self.timeout, self.stop)

        urls = ['https://quotes.toscrape.com/page/1']
        for url in urls:
            yield scrapy.Request(url=url, callback=self.parse)

    def parse(self, response):
        for quote in response.css('div.quote'):
```

```
        yield {'text': quote.css('span.text::text').get()}

    def stop(self):
        self.crawler.engine.close_spider(self, 'timeout')
```

The default value of the :setting:`TWISTED_REACTOR` setting is `None`, which means that Scrapy will install the default reactor defined by Twisted for the current platform. This is to maintain backward compatibility and avoid possible problems caused by using a non-default reactor.

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 1640); *backlink*

Unknown interpreted text role "setting".

For additional information, see :doc:`core/howto/choosing-reactor`.

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 1645); *backlink*

Unknown interpreted text role "doc".

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 1648)

Unknown directive type "setting".

```
.. setting:: URLLENGTH_LIMIT
```

URLLENGTH_LIMIT

Default: `2083`

Scope: `spidermiddlewares.urllength`

The maximum URL length to allow for crawled URLs.

This setting can act as a stopping condition in case of URLs of ever-increasing length, which may be caused for example by a programming error either in the target server or in your code. See also :setting:`REDIRECT_MAX_TIMES` and :setting:`DEPTH_LIMIT`.

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 1659); *backlink*

Unknown interpreted text role "setting".

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 1659); *backlink*

Unknown interpreted text role "setting".

Use `0` to allow URLs of any length.

The default value is copied from the Microsoft Internet Explorer maximum URL length, even though this setting exists for different reasons.

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 1671)

Unknown directive type "setting".

```
.. setting:: USER_AGENT
```

USER_AGENT

Default: `"Scrapy/VERSION (+https://scrapy.org)"`

The default User-Agent to use when crawling, unless overridden. This user agent is also used by :class:`~scrapy.downloadermiddlewares.robotstxt.RobotsTxtMiddleware` if :setting:`ROBOTSTXT_USER_AGENT` setting is `None` and there is no overriding User-Agent header specified for the request.

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 1678); *backlink*

Unknown interpreted text role "class".

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 1678); *backlink*

Unknown interpreted text role "setting".

Settings documented elsewhere:

The following settings are documented elsewhere, please check each specific case to see how to enable and use them.

System Message: ERROR/3 (`D:\onboarding-resources\sample-onboarding-resources\scrapy-master\docs\topics\[scrapy-master][docs][topics]settings.rst`, line 1690)

Unknown directive type "settingslist".

```
.. settingslist::
```