HOWTO Use Python in the web

Author
Marek Kubica

Abstract

This document shows how Python fits into the web. It presents some ways to integrate Python with a web server, and general practices useful for developing web sites.

Programming for the Web has become a hot topic since the rise of “Web 2.0”, which focuses on user-generated content on web sites. It has always been possible to use Python for creating web sites, but it was a rather tedious task. Therefore, many frameworks and helper tools have been created to assist developers in creating faster and more robust sites. This HOWTO describes some of the methods used to combine Python with a web server to create dynamic content. It is not meant as a complete introduction, as this topic is far too broad to be covered in one single document. However, a short overview of the most popular libraries is provided.

See also: While this HOWTO tries to give an overview of Python in the web, it cannot always be as up to date as desired. Web development in Python is rapidly moving forward, so the wiki page on Web Programming may be more in sync with recent development.

The Low-Level View

When a user enters a web site, their browser makes a connection to the site’s web server (this is called the *request*). The server looks up the file in the file system and sends it back to the user’s browser, which displays it (this is the *response*). This is roughly how the underlying protocol, HTTP, works.

Dynamic web sites are not based on files in the file system, but rather on programs which are run by the web server when a request comes in, and which *generate* the content that is returned to the user. They can do all sorts of useful things, like display the postings of a bulletin board, show your email, configure software, or just display the current time. These programs can be written in any programming language the server supports. Since most servers support Python, it is easy to use Python to create dynamic web sites.

Most HTTP servers are written in C or C++, so they cannot execute Python code directly – a bridge is needed between the server and the program. These bridges, or rather interfaces, define how programs interact with the server. There have been numerous attempts to create the best possible interface, but there are only a few worth mentioning.

Not every web server supports every interface. Many web servers only support old, now-obsolete interfaces; however, they can often be extended using third-party modules to support newer ones.

Common Gateway Interface

This interface, most commonly referred to as “CGI”, is the oldest, and is supported by nearly every web server out of the box. Programs using CGI to communicate with their web server need to be started by the server for every request. So, every request starts a new Python interpreter – which takes some time to start up – thus making the whole interface only usable for low load situations.

The upside of CGI is that it is simple – writing a Python program which uses CGI is a matter of about three lines of code. This simplicity comes at a price: it does very few things to help the developer.

Writing CGI programs, while still possible, is no longer recommended. With WSGI, a topic covered later in this document, it is possible to write programs that emulate CGI, so they can be run as CGI if no better option is available.

See also: The Python standard library includes some modules that are helpful for creating plain CGI programs:

- `cgi` – Handling of user input in CGI scripts
- `cgitb` – Displays nice tracebacks when errors happen in CGI applications, instead of presenting a “500 Internal Server Error” message

The Python wiki features a page on CGI scripts with some additional information about CGI in Python.

Simple script for testing CGI

To test whether your web server works with CGI, you can use this short and simple CGI program:

```
#!/usr/bin/env python
# -*- coding: UTF-8 -*-

# enable debugging
import cgitb
cgitb.enable()

print "Content-Type: text/plain;charset=utf-8"
print

print "Hello World!"
```

Depending on your web server configuration, you may need to save this code with a `.py` or `.cgi` extension. Additionally, this file may also need to be in a `cgi-bin` folder, for security reasons.

You might wonder what the `cgitb` line is about. This line makes it possible to display a nice traceback instead of just crashing and displaying an “Internal Server Error” in the user’s browser. This is useful for debugging, but it might risk exposing some confidential data to the user. You should not use `cgitb` in production code for this reason. You should *always* catch exceptions, and display proper error pages – end-users don’t like to see nondescript “Internal Server Errors” in their browsers.

Setting up CGI on your own server

If you don’t have your own web server, this does not apply to you. You can check whether it works as-is, and if not you will need to talk to the administrator of your web server. If it is a big host, you can try filing a ticket asking for Python support.

If you are your own administrator or want to set up CGI for testing purposes on your own computers, you have to configure it by yourself. There is no single way to configure CGI, as there are many web servers with different configuration options. Currently the most widely used free web server is Apache HTTPd, or Apache for short. Apache can be easily installed on nearly every system using the system’s package management tool. lighttpd is another alternative and is said to have better performance. On many systems this server can also be installed using the package management tool, so manually compiling the web server may not be needed.

- On Apache you can take a look at the Dynamic Content with CGI tutorial, where everything is described. Most of the time it is enough just to set `+ExecCGI`. The tutorial also describes the most common gotchas that might arise.
- On lighttpd you need to use the CGI module, which can be configured in a straightforward way. It boils down to setting `cgi.assign` properly.

Common problems with CGI scripts

Using CGI sometimes leads to small annoyances while trying to get these scripts to run. Sometimes a seemingly correct script does not work as expected, the cause being some small hidden problem that's difficult to spot.

Some of these potential problems are:

- The Python script is not marked as executable. When CGI scripts are not executable most web servers will let the user download it, instead of running it and sending the output to the user. For CGI scripts to run properly on Unix-like operating systems, the `+x` bit needs to be set. Using `chmod a+x your_script.py` may solve this problem.

- On a Unix-like system, The line endings in the program file must be Unix style line endings. This is important because the web server checks the first line of the script (called shebang) and tries to run the program specified there. It gets easily confused by Windows line endings (Carriage Return & Line Feed, also called CRLF), so you have to convert the file to Unix line endings (only Line Feed, LF). This can be done automatically by uploading the file via FTP in text mode instead of binary mode, but the preferred way is just telling your editor to save the files with Unix line endings. Most editors support this.

- Your web server must be able to read the file, and you need to make sure the permissions are correct. On unix-like systems, the server often runs as user and group `www-data`, so it might be worth a try to change the file ownership, or making the file world readable by using `chmod a+r your_script.py`.

- The web server must know that the file you're trying to access is a CGI script. Check the configuration of your web server, as it may be configured to expect a specific file extension for CGI scripts.

- On Unix-like systems, the path to the interpreter in the shebang (`#!/usr/bin/env python`) must be correct. This line calls `/usr/bin/env` to find Python, but it will fail if there is no `/usr/bin/env`, or if Python is not in the web server's path. If you know where your Python is installed, you can also use that full path. The commands `whereis python` and `type -p python` could help you find where it is installed. Once you know the path, you can change the shebang accordingly: `#!/usr/bin/python`.

- The file must not contain a BOM (Byte Order Mark). The BOM is meant for determining the byte order of UTF-16 and UTF-32 encodings, but

some editors write this also into UTF-8 files. The BOM interferes with the shebang line, so be sure to tell your editor not to write the BOM.

- If the web server is using mod_python, `mod_python` may be having problems. `mod_python` is able to handle CGI scripts by itself, but it can also be a source of issues.

mod_python

People coming from PHP often find it hard to grasp how to use Python in the web. Their first thought is mostly mod_python, because they think that this is the equivalent to `mod_php`. Actually, there are many differences. What `mod_python` does is embed the interpreter into the Apache process, thus speeding up requests by not having to start a Python interpreter for each request. On the other hand, it is not “Python intermixed with HTML” in the way that PHP is often intermixed with HTML. The Python equivalent of that is a template engine. `mod_python` itself is much more powerful and provides more access to Apache internals. It can emulate CGI, work in a “Python Server Pages” mode (similar to JSP) which is “HTML intermingled with Python”, and it has a “Publisher” which designates one file to accept all requests and decide what to do with them.

`mod_python` does have some problems. Unlike the PHP interpreter, the Python interpreter uses caching when executing files, so changes to a file will require the web server to be restarted. Another problem is the basic concept – Apache starts child processes to handle the requests, and unfortunately every child process needs to load the whole Python interpreter even if it does not use it. This makes the whole web server slower. Another problem is that, because `mod_python` is linked against a specific version of `libpython`, it is not possible to switch from an older version to a newer (e.g. 2.4 to 2.5) without recompiling `mod_python`. `mod_python` is also bound to the Apache web server, so programs written for `mod_python` cannot easily run on other web servers.

These are the reasons why `mod_python` should be avoided when writing new programs. In some circumstances it still might be a good idea to use `mod_python` for deployment, but WSGI makes it possible to run WSGI programs under `mod_python` as well.

FastCGI and SCGI

FastCGI and SCGI try to solve the performance problem of CGI in another way. Instead of embedding the interpreter into the web server, they create long-running background processes. There is still a module in the web

server which makes it possible for the web server to “speak” with the background process. As the background process is independent of the server, it can be written in any language, including Python. The language just needs to have a library which handles the communication with the webserver.

The difference between FastCGI and SCGI is very small, as SCGI is essentially just a “simpler FastCGI”. As the web server support for SCGI is limited, most people use FastCGI instead, which works the same way. Almost everything that applies to SCGI also applies to FastCGI as well, so we’ll only cover the latter.

These days, FastCGI is never used directly. Just like `mod_python`, it is only used for the deployment of WSGI applications.

Setting up FastCGI

Each web server requires a specific module.

- Apache has both mod_fastcgi and mod_fcgid. `mod_fastcgi` is the original one, but it has some licensing issues, which is why it is sometimes considered non-free. `mod_fcgid` is a smaller, compatible alternative. One of these modules needs to be loaded by Apache.
- lighttpd ships its own FastCGI module as well as an SCGI module.
- nginx also supports FastCGI.

Once you have installed and configured the module, you can test it with the following WSGI-application:

```
#!/usr/bin/env python
# -*- coding: UTF-8 -*-

from cgi import escape
import sys, os
from flup.server.fcgi import WSGIServer

def app(environ, start_response):
    start_response('200 OK', [('Content-Type', 'text/html')])

    yield '<h1>FastCGI Environment</h1>'
    yield '<table>'
    for k, v in sorted(environ.items()):
        yield '<tr><th>%s</th><td>%s</td></tr>' % (escape(k), escape(v))
    yield '</table>'

WSGIServer(app).run()
```

This is a simple WSGI application, but you need to install flup first, as flup handles the low level FastCGI access.

See also: There is some documentation on setting up Django with WSGI, most of which can be reused for other WSGI-compliant frameworks and libraries. Only the `manage.py` part has to be changed, the example used here can be used instead. Django does more or less the exact same thing.

mod_wsgi

mod_wsgi is an attempt to get rid of the low level gateways. Given that FastCGI, SCGI, and mod_python are mostly used to deploy WSGI applications, mod_wsgi was started to directly embed WSGI applications into the Apache web server. mod_wsgi is specifically designed to host WSGI applications. It makes the deployment of WSGI applications much easier than deployment using other low level methods, which need glue code. The downside is that mod_wsgi is limited to the Apache web server; other servers would need their own implementations of mod_wsgi.

mod_wsgi supports two modes: embedded mode, in which it integrates with the Apache process, and daemon mode, which is more FastCGI-like. Unlike FastCGI, mod_wsgi handles the worker-processes by itself, which makes administration easier.

Step back: WSGI

WSGI has already been mentioned several times, so it has to be something important. In fact it really is, and now it is time to explain it.

The *Web Server Gateway Interface*, or WSGI for short, is defined in **PEP 333** and is currently the best way to do Python web programming. While it is great for programmers writing frameworks, a normal web developer does not need to get in direct contact with it. When choosing a framework for web development it is a good idea to choose one which supports WSGI.

The big benefit of WSGI is the unification of the application programming interface. When your program is compatible with WSGI – which at the outer level means that the framework you are using has support for WSGI – your program can be deployed via any web server interface for which there are WSGI wrappers. You do not need to care about whether the application user uses mod_python or FastCGI or mod_wsgi – with WSGI your application will work on any gateway interface. The Python standard library contains its own WSGI server, `wsgiref`, which is a small web server that can be used for testing.

A really great WSGI feature is middleware. Middleware is a layer around

your program which can add various functionality to it. There is quite a bit of middleware already available. For example, instead of writing your own session management (HTTP is a stateless protocol, so to associate multiple HTTP requests with a single user your application must create and manage such state via a session), you can just download middleware which does that, plug it in, and get on with coding the unique parts of your application. The same thing with compression – there is existing middleware which handles compressing your HTML using gzip to save on your server's bandwidth. Authentication is another problem that is easily solved using existing middleware.

Although WSGI may seem complex, the initial phase of learning can be very rewarding because WSGI and the associated middleware already have solutions to many problems that might arise while developing web sites.

WSGI Servers

The code that is used to connect to various low level gateways like CGI or mod_python is called a *WSGI server*. One of these servers is `flup`, which supports FastCGI and SCGI, as well as AJP. Some of these servers are written in Python, as `flup` is, but there also exist others which are written in C and can be used as drop-in replacements.

There are many servers already available, so a Python web application can be deployed nearly anywhere. This is one big advantage that Python has compared with other web technologies.

See also: A good overview of WSGI-related code can be found in the WSGI homepage, which contains an extensive list of WSGI servers which can be used by *any* application supporting WSGI.

You might be interested in some WSGI-supporting modules already contained in the standard library, namely:

- `wsgiref` – some tiny utilities and servers for WSGI

Case study: MoinMoin

What does WSGI give the web application developer? Let's take a look at an application that's been around for a while, which was written in Python without using WSGI.

One of the most widely used wiki software packages is MoinMoin. It was created in 2000, so it predates WSGI by about three years. Older versions needed separate code to run on CGI, mod_python, FastCGI and standalone.

It now includes support for WSGI. Using WSGI, it is possible to deploy MoinMoin on any WSGI compliant server, with no additional glue code. Unlike the pre-WSGI versions, this could include WSGI servers that the authors of MoinMoin know nothing about.

Model-View-Controller

The term *MVC* is often encountered in statements such as “framework *foo* supports MVC”. MVC is more about the overall organization of code, rather than any particular API. Many web frameworks use this model to help the developer bring structure to their program. Bigger web applications can have lots of code, so it is a good idea to have an effective structure right from the beginning. That way, even users of other frameworks (or even other languages, since MVC is not Python-specific) can easily understand the code, given that they are already familiar with the MVC structure.

MVC stands for three components:

- The *model*. This is the data that will be displayed and modified. In Python frameworks, this component is often represented by the classes used by an object-relational mapper.
- The *view*. This component’s job is to display the data of the model to the user. Typically this component is implemented via templates.
- The *controller*. This is the layer between the user and the model. The controller reacts to user actions (like opening some specific URL), tells the model to modify the data if necessary, and tells the view code what to display,

While one might think that MVC is a complex design pattern, in fact it is not. It is used in Python because it has turned out to be useful for creating clean, maintainable web sites.

Note: While not all Python frameworks explicitly support MVC, it is often trivial to create a web site which uses the MVC pattern by separating the data logic (the model) from the user interaction logic (the controller) and the templates (the view). That’s why it is important not to write unnecessary Python code in the templates – it works against the MVC model and creates chaos in the code base, making it harder to understand and modify.

See also: The English Wikipedia has an article about the Model-View-Controller pattern. It includes a long list of web frameworks for various programming languages.

Ingredients for Websites

Websites are complex constructs, so tools have been created to help web developers make their code easier to write and more maintainable. Tools like these exist for all web frameworks in all languages. Developers are not forced to use these tools, and often there is no “best” tool. It is worth learning about the available tools because they can greatly simplify the process of developing a web site.

See also: There are far more components than can be presented here. The Python wiki has a page about these components, called Web Components.

Templates

Mixing of HTML and Python code is made possible by a few libraries. While convenient at first, it leads to horribly unmaintainable code. That’s why templates exist. Templates are, in the simplest case, just HTML files with placeholders. The HTML is sent to the user’s browser after filling in the placeholders.

Python already includes two ways to build simple templates:

```
>>> template = "<html><body><h1>Hello %s!</h1></body></html>"
>>> print template % "Reader"
<html><body><h1>Hello Reader!</h1></body></html>

>>> from string import Template
>>> template = Template("<html><body><h1>Hello ${name}</h1></body></html>")
>>> print template.substitute(dict(name='Dinsdale'))
<html><body><h1>Hello Dinsdale!</h1></body></html>
```

To generate complex HTML based on non-trivial model data, conditional and looping constructs like Python’s *for* and *if* are generally needed. *Template engines* support templates of this complexity.

There are a lot of template engines available for Python which can be used with or without a framework. Some of these define a plain-text programming language which is easy to learn, partly because it is limited in scope. Others use XML, and the template output is guaranteed to be always be valid XML. There are many other variations.

Some frameworks ship their own template engine or recommend one in particular. In the absence of a reason to use a different template engine, using the one provided by or recommended by the framework is a good idea.

Popular template engines include:

- Mako
- Genshi
- Jinja

See also: There are many template engines competing for attention, because it is pretty easy to create them in Python. The page Templating in the wiki lists a big, ever-growing number of these. The three listed above are considered “second generation” template engines and are a good place to start.

Data persistence

Data persistence, while sounding very complicated, is just about storing data. This data might be the text of blog entries, the postings on a bulletin board or the text of a wiki page. There are, of course, a number of different ways to store information on a web server.

Often, relational database engines like MySQL or PostgreSQL are used because of their good performance when handling very large databases consisting of millions of entries. There is also a small database engine called SQLite, which is bundled with Python in the `sqlite3` module, and which uses only one file. It has no other dependencies. For smaller sites SQLite is just enough.

Relational databases are *queried* using a language called SQL. Python programmers in general do not like SQL too much, as they prefer to work with objects. It is possible to save Python objects into a database using a technology called ORM (Object Relational Mapping). ORM translates all object-oriented access into SQL code under the hood, so the developer does not need to think about it. Most frameworks use ORMs, and it works quite well.

A second possibility is storing data in normal, plain text files (some times called “flat files”). This is very easy for simple sites, but can be difficult to get right if the web site is performing many updates to the stored data.

A third possibility are object oriented databases (also called “object databases”). These databases store the object data in a form that closely parallels the way the objects are structured in memory during program execution. (By contrast, ORMs store the object data as rows of data in tables and relations between those rows.) Storing the objects directly has the advantage that nearly all objects can be saved in a straightforward

way, unlike in relational databases where some objects are very hard to represent.

Frameworks often give hints on which data storage method to choose. It is usually a good idea to stick to the data store recommended by the framework unless the application has special requirements better satisfied by an alternate storage mechanism.

See also:

- Persistence Tools lists possibilities on how to save data in the file system. Some of these modules are part of the standard library
- Database Programming helps with choosing a method for saving data
- SQLAlchemy, the most powerful OR-Mapper for Python, and Elixir, which makes SQLAlchemy easier to use
- SQLObject, another popular OR-Mapper
- ZODB and Durus, two object oriented databases

Frameworks

The process of creating code to run web sites involves writing code to provide various services. The code to provide a particular service often works the same way regardless of the complexity or purpose of the web site in question. Abstracting these common solutions into reusable code produces what are called “frameworks” for web development. Perhaps the most well-known framework for web development is Ruby on Rails, but Python has its own frameworks. Some of these were partly inspired by Rails, or borrowed ideas from Rails, but many existed a long time before Rails.

Originally Python web frameworks tended to incorporate all of the services needed to develop web sites as a giant, integrated set of tools. No two web frameworks were interoperable: a program developed for one could not be deployed on a different one without considerable re-engineering work. This led to the development of “minimalist” web frameworks that provided just the tools to communicate between the Python code and the http protocol, with all other services to be added on top via separate components. Some ad hoc standards were developed that allowed for limited interoperability between frameworks, such as a standard that allowed different template engines to be used interchangeably.

Since the advent of WSGI, the Python web framework world has been evolving toward interoperability based on the WSGI standard. Now many web frameworks, whether “full stack” (providing all the tools one needs to deploy the most complex web sites) or minimalist, or anything in between, are built from collections of reusable components that can be used with more than one framework.

The majority of users will probably want to select a “full stack” framework that has an active community. These frameworks tend to be well documented, and provide the easiest path to producing a fully functional web site in minimal time.

Some notable frameworks

There are an incredible number of frameworks, so they cannot all be covered here. Instead we will briefly touch on some of the most popular.

Django

Django is a framework consisting of several tightly coupled elements which were written from scratch and work together very well. It includes an ORM which is quite powerful while being simple to use, and has a great online administration interface which makes it possible to edit the data in the database with a browser. The template engine is text-based and is designed to be usable for page designers who cannot write Python. It supports template inheritance and filters (which work like Unix pipes). Django has many handy features bundled, such as creation of RSS feeds or generic views, which make it possible to create web sites almost without writing any Python code.

It has a big, international community, the members of which have created many web sites. There are also a lot of add-on projects which extend Django’s normal functionality. This is partly due to Django’s well written online documentation and the Django book.

> **Note:** Although Django is an MVC-style framework, it names the elements differently, which is described in the Django FAQ.

TurboGears

Another popular web framework for Python is TurboGears. TurboGears takes the approach of using already existing components and combining them with glue code to create a seamless experience. TurboGears gives the user flexibility in choosing components. For example the ORM and template

engine can be changed to use packages different from those used by default.

The documentation can be found in the TurboGears documentation, where links to screencasts can be found. TurboGears has also an active user community which can respond to most related questions. There is also a TurboGears book published, which is a good starting point.

The newest version of TurboGears, version 2.0, moves even further in direction of WSGI support and a component-based architecture. TurboGears 2 is based on the WSGI stack of another popular component-based web framework, Pylons.

Zope

The Zope framework is one of the “old original” frameworks. Its current incarnation in Zope2 is a tightly integrated full-stack framework. One of its most interesting feature is its tight integration with a powerful object database called the ZODB (Zope Object Database). Because of its highly integrated nature, Zope wound up in a somewhat isolated ecosystem: code written for Zope wasn’t very usable outside of Zope, and vice-versa. To solve this problem the Zope 3 effort was started. Zope 3 re-engineers Zope as a set of more cleanly isolated components. This effort was started before the advent of the WSGI standard, but there is WSGI support for Zope 3 from the Repoze project. Zope components have many years of production use behind them, and the Zope 3 project gives access to these components to the wider Python community. There is even a separate framework based on the Zope components: Grok.

Zope is also the infrastructure used by the Plone content management system, one of the most powerful and popular content management systems available.

Other notable frameworks

Of course these are not the only frameworks that are available. There are many other frameworks worth mentioning.

Another framework that’s already been mentioned is Pylons. Pylons is much like TurboGears, but with an even stronger emphasis on flexibility, which comes at the cost of being more difficult to use. Nearly every component can be exchanged, which makes it necessary to use the documentation of every single component, of which there are many. Pylons builds upon Paste, an extensive set of tools which are handy for WSGI.

And that’s still not everything. The most up-to-date information can always

be found in the Python wiki.

See also: The Python wiki contains an extensive list of web frameworks.

Most frameworks also have their own mailing lists and IRC channels, look out for these on the projects' web sites.