**PIP**

**By**

**Augustine, Arif, Sanjog, Masinde (Group 9)**

**Åbo Akademi University**

**Introduction**

pip is a recursive acronym that stands for Pip installs Packages. It is a package management system for installing and managing Python software packages. It is an open source software by the Python Packaging Authority (PyPA) initiated and led by Donald Stufft and was initially released on 4th April, 2011. The most recent stable release was on 6th November, 2016.

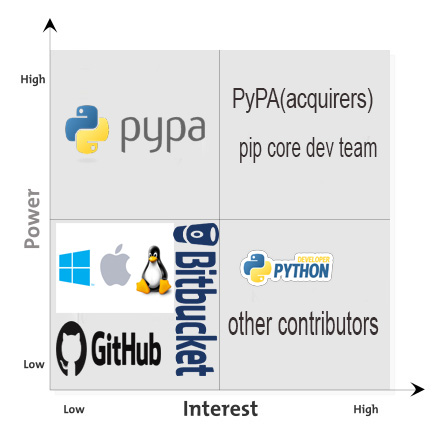
pip is the defacto package manager for software packages developed with Python and ships with most Python distributions. As such its user base is almost as large as the entire Python community of developers. It is notable for the ease of its command-line interface, which simplifies the installation and uninstallation of Python software packages.

**Stakeholder Analysis**

Stakeholders of project/product refer to individuals and/or organizations who have direct impact on the development and success of the product/project. They have varied levels of power and interest in the project/product. Listed below are some stakeholders we have identified in pip project and their respective roles.

* **Acquirers**: PyPA is the acquirer of pip. Pip was initiated by Donald Stufft who is a core PyPA developer.
* **Assessors:** The Python Packaging Authority is responsible for assessing and approving contributions from pip contributors to the source code. It takes care of legal matters and licensing issues regarding pip and other Python packages stored in the Python Packages Index.
* **Communicators**: The communicators maintain communication with other stakeholders. They analyze and make documentation of the system. Based on this definition, the PyPA team members are the communicators in pip project.
* **Developers**: PyPA developers and other contributors from Python Developers community.
* **Maintainers**: PyPA core developers are responsible for maintaining pip. Other Python developers also contribute to the maintenance of pip.
* **Production Engineers:** pip like most other Python packages are stored in a repository called Python Packages Index (PyPI) owned and managed by PyPA so PyPA engineers serve as the production engineers. pip source code is hosted on github, therefore engineers from github could also be seen as part of the production engineers
* **Suppliers:** PIP works on different operating system such as Unix/Linux, Mac, and Windows and it also uses github version control and issue tracking. Vendors of these software platforms could be seen as suppliers in pip project.
* **Support Staff:** PyPA is a working groupthat supports many of the relevant projects in Python packaging. Besides this support staff for pip could also be github and bitbucket staff and many other volunteers.
* **System Administrators:**  PyPA is responsible for the system administration.
* **Testers:** The pip development team is responsible for testing the software to ensure that it meets the requirements. All contributors are required to perform both unit testing and integration testing on the code they are contributing to the pip project. Python developers could also be seen as playing the role of testers.
* **Users:** Python developers

**Stakeholders Power/Interest Grid**

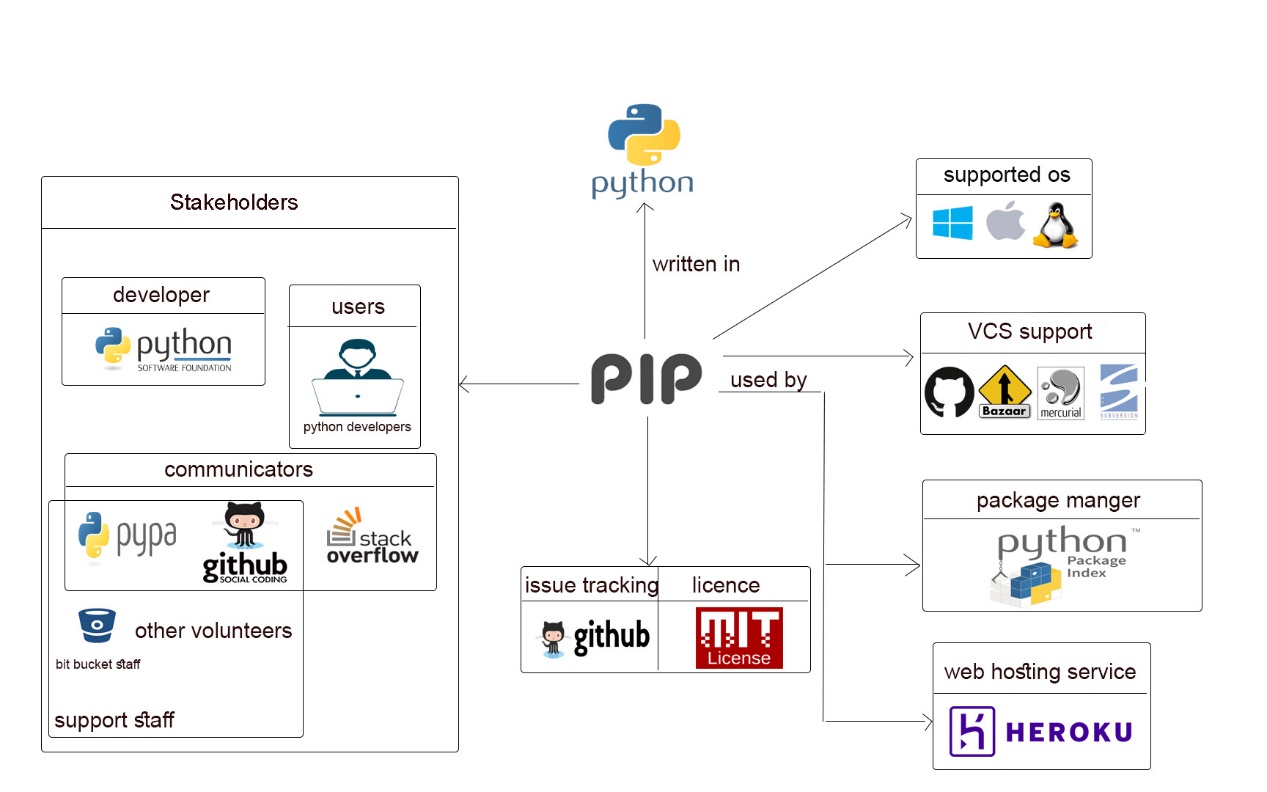
****

**Fig 1. Power Interest Grid**

Figure 1 above visualizes the level of interest and power of the different stakeholders in the pip project. From the figure, PyPA at the top right corner of the grid, as the acquirer and represented by Donald Stufft, has the highest interest and power in the project. The success of the project is of utmost importance to it and it has the final say on important decisions concerning the project. At the top left corner of the grid is PyPA again, this time as the accessor. In this role, it has very little or no interest in the project but has significant power regarding decision making in the direction of the project. At the bottom right corner of the grid are the developers and maintainers, which include PyPA core developers, led by Donald Stufft, and other contributors from the Python developers community, and the users, which include Python developers. These do not have much say in the decision making concerning the project but they have relatively high interest in the success of the project. Lastly, at the bottom left corner of the grid are the communicators, testers, support staff, etc. Although these play significant role in the project’s life cycle as stakeholders, they have relatively low power and interest in the project.

**The Context Viewpoint**

The Context viewpoint describes the relationships, communication interfaces and channels, interdependencies between a system and its environment, which includes users, developers, maintainers, other systems, data input source, etc. Its concern is to describe the nature of the contents of the system’s environment (context), how the system relates with the different players and systems within its context, the nature of the interfaces between the system and its environment, what the system needs from its dependencies and what it does and does not do (scope).



**Fig 2. Context Model**

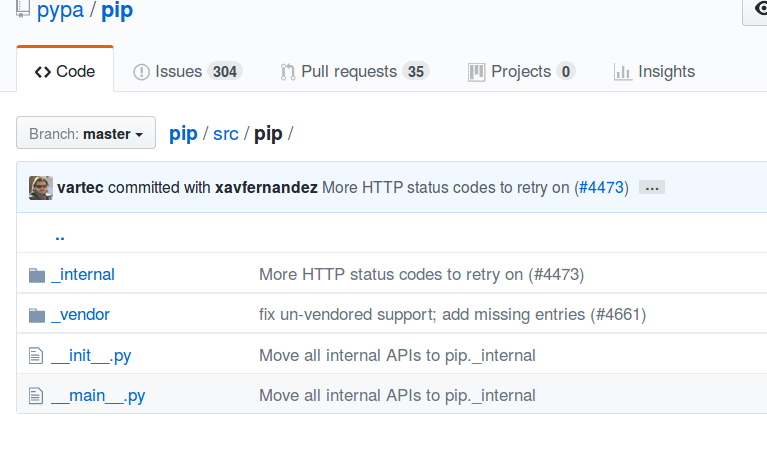
The figure 2 above is a pictorial overview of pip’s context. Pip is written in Python and has support for the three major operating system platforms, Windows, Mac and Linux. It depends on github for issue tracking and version control and is licensed under the MIT license. Pip installs python packages from a repository of software for the Python programming language called the Python Package Index (PyPI). It was developed, and is being maintained by PyPa core developers and contributors from the Python developers community. PyPa is responsible for moderating and supervising inputs from contributors. As a package manager for the Python programming language, pip is used by Python developers to install and manage Python packages in their Python-based projects.

**The Development Viewpoint**

The development viewpoint refers to a description of the architecture that supports a system’s development process. It is concerned with module structure and organization, module dependencies, standardization of design and testing, and codeline organization and consistency for technical integrity. This directly addresses the concerns of the stakeholders within the development and testing groups.

Pip is basically a command line based utility software for installing, uninstalling, and managing software packages developed with the Python programming language. Pip was originally developed to be self-contained, implementing all the functionalities it needed to eliminate the need for dependencies and its attendant challenges and problems. However, with version 1.5, pip introduced the idea of dependencies replacing some codes implemented internally with reusable libraries from PyPi in order to leverage the benefits of reusing libraries other than re-inventing the wheel.

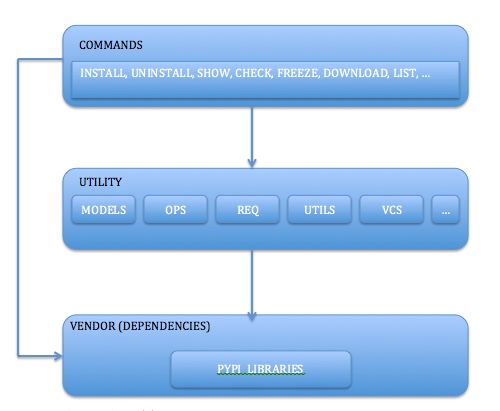
The pip source code folder on github shows the pip modules/components, which are group of files with implementations of similar/related functionalities. It shows the development structure when developing pip. The modules are easier to follow. In the modules folder there is internal module and vendor module.

**Fig. 3 pip src folder organization**

The approach to dependency in pip is different from the traditional way (that is via *install\_requires*). In pip, to avoid dependency-related issues, dependencies are bundled with the source. So pip’s source directory contains two main directories: **\_internal** and **\_vendor.**

**\_internal** : This contains pip’s internal API and helper functions. All the implementations of pip’s functionalities as a package manager are contained in this directory. It contains several directories and files one of which is the commands directory, which contains implementations of all pip’s commands.

**\_vendor** : This contains basically pip’s bundled dependencies (implementations of the libraries from PyPi on which pip depends). Each of these dependencies with its version is listed in the file vendor.txt also contained in this directory. This list is used for automatic dependency update.



**Fig. 4 pip module structure**

The figure 4 above is a simplified model of pip’s module structure describing the different layers of dependencies. Pip commands implementations are dependent on several internal utility modules and helper functions, which make use of third party libraries from PyPI. These commands also make use of the external libraries from PyPI directly.

**Contributing to pip**

The open source project pip is maintained by The Python Packaging Authority (PyPA), and is developed by Python core developer and the community of Python developers registered on GitHub. Any contributor can contribute to the project through GitHub.

To contribute in pip development, the contributor needs to submit a Pull Request against the master branch on GitHub. The contributor must provide an explicit description of the tasks his/her contribution performs and they are tested locally first.

**Automated testing:** all kinds of pull requests and push to ‘master’ are tested in Travis, and Travis is based on .travis.yml file. Jenkins CI is used for certain versions of Python on centos and windows.

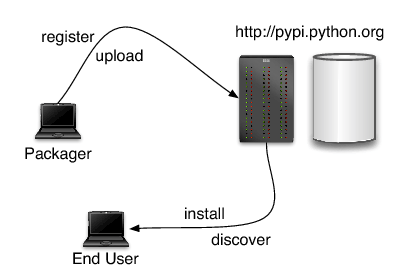
**Tests running:** to run the test an operating system needs to have subversion, git, bazaar, and mercurial. And Python requirements are a virtualenv, tox or pytest, scripttest, and mock.



**Fig 5. Snapshot of running test locally**

**Pip usage**

The Python Package Index (PyPI) is a repository of software packages written in Python. It is the host of different Python packages and their dependencies. At present there are about 119685 packages on PyPI. Donald Stufft is a core developer of pip, and PyPI is solely developed and maintained by him.



**Figure: Package management process in PyPI using pip**

Developers can register/upload their own Python software packages in PyPI using pip command and other users can browse existing projects/packages. Source can be uploaded to an existing package. Pip uses PyPI as a default host for the packages and their dependencies.

|  |  |
| --- | --- |
| Command | Action |
| Pip install | Install packages |
| Pip uninstall | Uninstall packages |
| Pip freeze | Output installed packages in requirements format |
| Pip list | List installed packages |
| Pip show | Show information about installed packages |
| Pip search | Search PyPI for packages |
| Pip zip | Zip individual packages |
| Pip unzip | Unzip individual packages |
| Pip bundle | Create pybundles |
| Pip help | Show help for commands |

**Table 1 Some important pip commands**

**Quality Attribute Analysis**

Pip is known for its ease of use. Pip ships with most Python distributions so it requires zero installation time and effort. Its ease of use is mostly shown in the simplification of installation and uninstallation of Python software packages which can be done simply with the command ‘pip install’ <package name> and ‘pip uninstall <package name> respectively. It is also shown in the fact that pip can be used to update pip itself since it is a PyPI package like all other Python packages. The command ‘pip install –upgrade pip’ upgrades pip.

Another very useful feature of pip is the possibility of managing a list of several packages and corresponding version numbers through a ‘requirements’ file. With this feature, an entire group of packages can be efficiently created in multiple environments, physical or virtual. This can be achieved with the following command:

‘pip install –r requirements.txt’ where requirements.txt is a well formatted list of packages and their corresponding version numbers to be installed/managed.

**Conclusion**

The Python packaging management tool is easy to install and use. In python 3.4 and later versions it is included. In the earlier versions of python it has to be installed in the particular OS. It is the recommended tool for installing python packages.

References:

1. Python pip

<https://github.com/pypa/pip>

1. pypa

https://github.com/pypa