

A Study on Open Source Software development - An Apache Web Server Case Study

Srinivas Nanduri

Botho College P.O.Box. 501564,

Gaborone, Botswana.

srinivas@bothocollege.ac.bw

Takwirira Tafireyi

Botho College P.O.Box. 501564,

Gaborone, Botswana.

takwirara.tafireyi@bothocollege.ac.bw

Abstract

Technologies have evolved and rapidly changing and in the current environment of business and competitive environment Open Source software comes as a relief for programmers. The developing world such as African countries have benefited from the use of Open Source Software. The ability to use resources from the internet without having to worry about the cost implications is what motivates many enthusiastic programmers to become open source software developers. The fact that sharing of information has been a mainstay in the academic environment and the development of science and technology without considering monetary gains has been the cause of knowledge and resource transfer and growth across the world. A simple case to open source learning in education is Wikipedia and the various internet resources which have revolutionized the learning processes that in fact many academicians and seekers of knowledge have become sometimes over reliant on these web resources.

The Apache web server serves as a case study for Open Source software which is currently being used by 60% of all web servers. It is interesting to observe that this software was developed by developers since 1995 and caters to the ever growing demand of its various products from the Apache Software foundation. This paper serves to first understand factors that contribute to the popularity of open source software among software developers followed by the understanding of the Apache web server and its development.

Keywords: Open Source Software, Apache Web servers, licensing procedures, Free and Open Source Software(FOSS).

2. INTRODUCTION

Open Source is defined as the software that meets the open source definition. It is for the software developer to access and modify the source code for their specific requirement and has been largely successful tool for ICT development in the developing world. It is interesting to note that in the African context many developments have been made in the open source environment with various organizations, formed to assist developers and the software engineering environment.

The technological community developed the open source software as an answer to the proprietary software developed by large corporations. The certification standard [1] used in this case is Open source Initiative (OSI) which uses the Open Source definition (OSD) and community-recognized body for reviewing and approving licenses as OSD-conformant. This is considered a relief for programmers from the proprietary software that are developed and marketed by large corporations, with a purpose to creating and distributing bug free software for everyone to use. It uses peer review mechanism to identify and eliminate bugs in the software, which proprietary based software do not employ. This process assists in improved performance of the software which is possible through removing bugs and developing the software through a shared resource community.

The milestones [2] that were achieved in the open source software are given as follows:

- 1983- The GNU project by the Richard Stallman
- 1985- The free software foundation creation.
- 1991- The Linux kernel created and developed by Linus Torvalds.
- 1998- Open source initiative created by Eric Raymond.

There is a misconception that open source does not innovate, which is based on the view of many proprietary vendors. However it is interesting to note that many browsers such as Firefox, Chrome and scripting languages such as Perl, Python and Ruby started out as open source projects. The open Source projects remain a source of innovation in not just the commercial sector but in universities, and many public institutions.

The reasons for choosing open source software:

- i. The price of commercially available software is high and many organizations such as libraries, schools cannot purchase the software.
- ii. The fact that open source software can be used anywhere and by anybody is an advantage in development of community as the distribution facilitates easy use and distribution of information.
- iii. The reasons that peer groups use the software adds to the advantage of using the software for the purpose of interoperability and custom designed approaches which adds value in using the software. It assists in creating systems such as digital libraries which can be used by communities at large.

3. Literature Review

3.1 Free and open source software:

It is important to understand free software for the purpose of distinguishing this from open source software.

FreeSoftware Definition: (According to Free software Foundation)

A program is free software if the program's users have the four essential freedoms:

- The freedom to run the program, for any purpose (freedom 0).

- The freedom to study how the program works, and change it so it does your computing as you wish (freedom 1). Access to the source code is a precondition for this.
- The freedom to redistribute copies so you can help your neighbor (freedom 2).
- The freedom to distribute copies of your modified versions to others (freedom 3). By doing this you can give the whole community a chance to benefit from your changes. Access to the source code is a precondition for this.

Free software [1] allows unrestricted access and usage of software, thus allowing it to be copied and modified and redistributed. Free word symbolizes liberty without taking into account the monetary aspects. However the Free software foundation does not encourage terms such as “for free” or give away”, as this linked to the aspect of monetary benefits rather than freedom to use the software.

The different categories of the free software are given as follows:

- i. Public domain software: This software does not incorporate copyright which allows the use for commercial or no commercial purposes.
- ii. Permissive licenses: This allows the author to retain copyrights at the same time ensuring that distribution without restrictions. The author however contributes to any work modified allows the redistribution of the software.
- iii. Copy left licenses: In this case the author keeps the copyright and allows redistribution with this process being possible under the same license. Any further modification will then part of the copy left licenses when they are distributed under the original license. E.g. GPL licence

A distinction must be made between open source software and free software. While the former, is restrictive in terms of licenses that are given the later gives freedom to use and distributed software. It is important to understand that open source software emphasizes the use of software which can be freely used but with some licencing agreements, such as the use of the operating system Linux. Some of the aspects of Open source software [3] can be given as follows:

- i. Free redistribution: The open source license does not restrict the selling or to give away the combined software which was formed as a result of several programs. There is no requirement for a royalty or fee for the license.
- ii. Source Code: The program must facilitate the distribution of the source code and its compiled form. If there is a product that is not available with the source code, there must be mechanism of the publicizing it for a reasonable cost, and downloading this from the internet free of cost. There must be clarity in the source code without confusing it and not allowing any other intermediate forms such as translator.
- iii. Derived works: The license should facilitate derived works and modification, and distribution should be allowed in terms applicable to the original software.
- iv. Integrity of the Authors' Source code: Restrictions are applicable to the distribution of the source code only if license considers patch files distribution with the source code at the build time. The license should ensure distribution of software which was built from the source code which was modified. There should be a different name or version applicable to the license for the derived works.

- v. No discrimination against persons or groups: There should be no discrimination in the licensing agreement against any individual or groups.
- vi. No discrimination against fields of endeavor: There should be no restriction in giving licensing to different fields such as business or genetic research.
- vii. Distribution of license: There is no need for any additional licenses for parties that have received programs which were redistributed.
- viii. License must not be specific to a product: The rights which are given for a program does not depend upon the distribution of particular software, and the process of redistribution as point number vii.
- ix. License must not restrict other software: There should be no restriction placed on other software while distributing the licensed software.
- x. License must be Technology neutral: The license should not predicate styles of interface or any technology.

3.2 Open Source software as driver of innovation:

There is a general opinion [4] that proprietary software hampers innovation and given proper organization, open source software can prove to be useful tool for innovation. The reason for this is the philosophy attached with the open source movement which gives way to transparency and partnerships which assists in solving the problems of organization such as government, health care. The idea here is according to Don Topscott when many developers contribute together, everyone benefits. In open sourcing unlike crowdsourcing where crowds work together for a common project, allows the project to be given to the developer. An example of open source platform is the Wikipedia which allows users to access to important and relevant information. The open source platform empowers people to change their lives and the societies they live in.

Examples of user innovation networks are Free and Open source software. This is unlike innovation by manufactured centered systems which uses was a mainstay of commercial enterprises. The advantages of having a user centered approach are that innovations are available to be shared between the developers to develop debug and use the software. Research conducted on the aspect of free revealing [5] have shown interesting patterns in which users in number of fields reveal their innovations even to commercial manufacturers. This is surprising to economists as this is a direct contradiction to the principle of innovation. This is because an innovation from the classical view is meant to be a secret which should follow patent laws. Spillovers of innovation and the related information which are not compensated should be avoided according to them.

There is an answer to this puzzle which can be described as follows:

- i. There are aspects of the code which is privy to the developer although the code is freely accessible to the public at large. The free riders that use the software benefit less from the software compared to the developers. The developer stands to gain as the users who use the system will debug and integrate the code with changes which in turn the leads to the advantages of “network effect”.
- ii. There is misunderstanding that creation and development of software with a profit motive is an easy exercise. On the contrary free revealing is beneficially compared to profit motive by commercial setups as in the latter there are licensing processes and selling the product involved.
- iii. Since the cost of free revealing is low, the competitive nature of using the software by many developers is less risky as the users will only modify the code free. The economic benefits attached with free revealing are more for multiple and users as well as the developers thus,

and hiding the software will have little benefits as there will be other developers who will freely revealing this information.

As shown in figure 1 there is difference in selling and distribution of information and physical products. Physical products need a manufacturer with a profit motive to sell the products while with information the developer can easily upload the information into the web. Software being information rather than a physical product this becomes even more possible to be easily distributed.

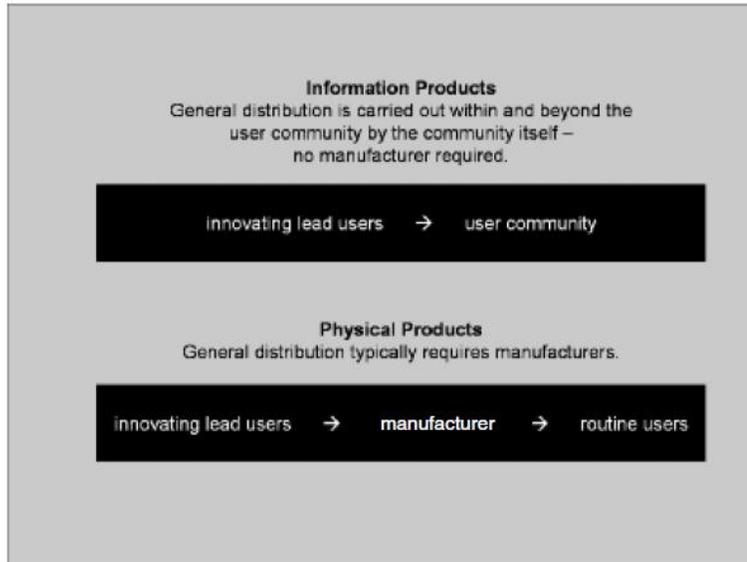


Figure 1: How lead user innovation are distributed (courtesy: Open Source Software Projects as User Innovation Networks by Eric Von Hippel(2012))

3.3 Motivation of providing open source assistance

Some of the major motives that developers [6] distribute software through open source are as follows:

- i. The need for the software and improvements developed by the users.
- ii. The joy experiencing by the entire process of creating and working software.
- iii. The reputation that follows in creating and distributing information in open source.

But contrary to the above arguments is the fact that open sourcing involves not just mundane tasks but essential work. There are several factors that contribute to providing which include the fun and learning that the developer benefits from sharing information., The immediate benefit for the user is their ability to use and implement the code developed, while cost of this is the time and effort spent in developing the software without any monetary gain. Other factors also involved are ego gratification and peer recognition.

Empirical studies conducted by Constant et.al (1996) revealed interesting facts the sample was taken out of 55 information seekers and 295 information providers. The system of open source assistance revealed that overall the exercise was useful, with 49% having been able to get the assistance they desired.

Reasons for participating	Points assigned (mean)
Personal benefits	
I enjoy helping others	16
I enjoy solving problems	9.5
I enjoy earning respect	4.8
The company rewards information sharing	0.9
Total	31.2
Organizational motivation	
Being a good company citizen	17.8
The problem is important to the company	14.0
It is part of my job to answer questions like this one	12.6
I expect others to help me, so it is only fair to help them	11.8
Total	56.2

^a Source of data: Constant et al. (1996), Table 5, p. 129.

Table 1: Survey feedback by Constant et.al (1996), (Courtesy: *How open source software works: “free” user-to-user assistance*, Lakhani.K.R. & Von Hippel.E(2002))

The information providers were also allocated 100 points for eight reasons and the results are shown in Table 1. Interestingly as can be observed in the Table “Being a good citizen” and “The problem is important to the company” were higher in the priority for the information providers in open source networks. Less strongly the information providers felt for reasons such as for were however for question such as “part of my job” and “I enjoy earning respect”

According to Kollock(1999) there are four possible motivation of contributing to public good online. They are given as follows:

- i. Since there is focus on information that was also already created reputation does not play a role in contribution to online sites.
- ii. There is a principle of reciprocity where individual see this as platform help each other although they do not know each other. This then becomes a generalized exchange (Ekeh, 1974).
- iii. The contributors feel that by contributing they have an effect over the environment. (Bandura, 1995)
- iv. There is also an attachment to a certain open source project by the contributor, and there is particular group dynamics that plays a role in being part of team.(ibid, pp 228-289)

3.4 Open Source Software in the context of Africa

Sub-Saharan Africa [7] has made rapid developments in the area of open source learning and software developments. The deployment of FOSS technologies has played a role in the development of ICT infrastructure in Africa. There are also many efforts currently ongoing in the direction of developing and improving FOSS’s potential in the African context.

However the usage of FOSS is disproportionately favoring the developed world compared to the African countries. It must be noted that FOSS technologies have been used by the developed world and the technological, educational and socio-economic developments of FOSS in Europe and North America is documented well. Currently however there are many initiatives in Africa in the direction of capacity building, and expansion in usage of these technologies.

Some of the examples of impact in African context include organizations such as

- FOSSFA is a nonprofit organization which represents Free and Open Source Software for Africa.
- ict@innovation-is a programme which involves FOSSFA and GIZ (a German owned organization) which looks at developing small and medium IT companies and the ICT sector in Africa.
- OSISA- A Southern African foundation which aims to promote and maintain values and ideals of an open society.
- AVOIR-stands for African Virtual Open initiatives and resources working with 11 institutions in developing software engineering in Africa.
- OERAFrica- This is a portal and an initiative of the South African Institute of distance education for individuals who want to benefit from OER(Open educational resources).
- AITI-KACE—which is the Ghana – India Kofi Annan Centre of Excellence is involves both the nations mentioned who goals is to stimulate and the growth of the ICT sector which includes FOSS, in the economic sector in West African Countries.

3.5 Interesting Survey facts about Open Sourcing

It is interesting to note that according to a survey [7] conducted by Optaros, IT consulting firm and InformationWeek magazine companies and governmental organizations in America are saving significant amount of money using the open source software. The survey was conducted in September, 2005 with 500 companies and governmental organizations and it was found that about 87% of these organizations use open source software. The motivation behind using open source software is cost saving measures, which according to Optaros in 2004, companies which generate revenues more than US\$1 billion save about \$3.3 million because of using the open source software. Companies with annual revenue between \$50 million and \$ 1billion save up to \$ 1.1 million. The survey also shows that the most popular open source technologies are the Linux, Apache and other assorted web browsers. It is interesting to note that open source content management systems is used by 42% of the participants of the survey while 16% of the respondents use customer relation management technologies. There are however worries about licensing, policies of software cost allocation, difficulties and choosing the software which act as impediments according to the survey.

There is however a survey which was conducted in Europe by IDC which showed contrasting results to the above. This was based on a survey conducted for 600 companies which showed that quality and flexibility played a role over cost saving measures for using the open source software. The IDC survey showed that 33% of people surveyed use open source databases while 25% use Linux operating system. This was a contradiction to the claim by proprietary software distributors users do have an interest in customizing software.

The survey conducted by Optaros is also different from the results by OSDL's desktop Linux survey which showed that more than cost saving and customization it is important to consider employee demand and competitors deployments.

The conclusion one can make is that open source deployments are becoming popular around the world, with American companies concentrating on the savings while European organization interested in the flexibility that these software can give. It is also important to note from the surveys that the demand for customizable software exists, contrary to the claims by proprietary software distributors.

4. Apache Web Server- Case Study

The Apache Web Server is also referred to as Httpd[8], is an open source software which is available under the apache licence version 2. It was developed by a group of programmers , and the first version of the Apache was developed in 1995, using the NCSA httpd Web Server as the basis. This software is used by 60% of the web domains. Twenty volunteer programmers called themselves the Apache group was involved in the core development of this Web Server. The source code of the software is available for free and users can adapt the software to their specific needs, this also includes Apache add-ons from a large public library. Development of the Apache software followed the same path as the Linux Operating System and was first written for the Unix operating system and the newer versions run under the Windows, OS/s and other platforms. It is also called the patchy server as it was developed from the NCSA code along with a variety of patches.

In further understanding the Open Source software development of the Apache Web Server it first important to explain the formation of the Apache Foundation. This is important phase for the Apache group where regulatory frameworks were set for further developments for their products.

4.1 Apache Software foundation (ASF)

The next natural step of the Apache group was the formation of the Apache Software foundation[10], which is non-profit corporation in June 1999. The foundation has management which includes Board of Directors and the corporation bylaws are used for the yearly election by the members. There are a set of officers appointed by the board to manage the day-to day operations and administer the ASF projects. There is a team of experts who contribute to the projects, and this process of collaboration involves certain guidelines.

The purpose of the foundation is as follows:

1. Ability to provide for software development which will include aspects such as open , collaborative approach by the supply of hardware, infrastructure for business, and communication.
2. Provision of legal entity which is independent where individuals or companies can provide their resources for the benefit of public.
3. Provide a sanctuary for individuals from legal suits who are working on the Foundation's Projects.
4. Protection of the Brand 'Apache' from abuse by other organizations.

The Apache Software foundation can be regarded as an organic entity and benefits extends to various areas such as fixing bugs, enhancing the features, and support system for other developers in the news groups and public lists. Although an individual contribution is light, the products are very strong.

There is also a factor that with commercial software, the developer will not be willing to fix bugs for free. The strength of the Apache software comes from the fact that it is free, which is why it has easy accessibility and acceptability.

4.2 Apache Web server

The Apache server supports the languages[9] such as Perl, Python, PHP. Additional functionality can be added which includes support for authentication processes, language support and other features. Popular examples of modules which support authentication are mod_digest, mod_auth_digest, mod_access, and mod_auth. SSL/TLS support can be provides modules such as URL rewriting (mod_rewrite), custom

logging (mod_log_config), filtering support (mod_include and mod_ext_filter), mod_ssl, and proxying (mod_proxy).

The Apache server supports CGI(common gateway Interface) , SSI(Server side Includes) , user authentication, anonymous user access, URL redirection, and automatic directory listings. After the launch of the web server in April 1996 several versions were released which made changes to the configuration of the software. For example the latest version Apache 2.0 provides the users with simple to use functionality, IPv6 and Unix support, multi-language error messaging and better compilation.

Apache 2.2 provides for proxy caching modules [11], SQL support and user friendly authentication systems. For the web administrator there is comprehensive documentation provided. It is interesting to note that commercial software use their own HTTP server development with an Apache base, also complying with the Licensing agreements. There is however a problem that the Apache HTTP server which comes from the Apache foundation is used by commercial vendors as their own product. Apache foundation highlights the problem by stating that there is no “Vendor Apache Product”.

Recent versions of the Apache Web servers[12] under the Apache HTTP server project include the Apache httpd 2.4.3, Apache httpd 2.2.23, Apache httpd 2.0.64,. These web servers can be supported by modern operating systems such as Windows NT, Unix. The versions provide improved features of authentication, proxying, OSCP support, Caching, smart filtering, security and bug fix releases. The Apache server project objective is to provide secure, extensible, services of HTTP keeping in mind the current HTTP standards.

5. Conclusions

- i. Open Source software development which provides cost benefit features among other advantaged to the software development community as can also be observed from various surveys discussed in the paper. Particularly for African countries, open sourcing has proved to be useful tool for the ICT development. The African Organizations in the Open Source environment such as FOSSFA, OSISA, AVOIR etc have been assisting in software engineering and ICT environment in the continent.
- ii. A distinction should be made between free and Open Source software based on licensing agreements which again is not the same as freeware, which can be distributed for free.
- iii. Although a commercial objective plays a role in the development of software and innovation around the world, the fact that various developers contribute to the open source software can be a case for innovation which defies the principles of economics.
- iv. It is interesting to note the motives such as “Being a good citizen” and “The problem is important to the company” play a role in the process of developers contributing to open source projects. This is unlike reasons such as “part of my job” and “I enjoy earning respect” which are assumed to be reasons for the developers.
- v. The Apache foundation was setup to provide legal setup assistance for open source software developers. The Apache Web server is good example of the success of Open source software which is being used by about 60% of web developers.

References

1. Biswas.G &Paul.D (2010), *An evaluative study on the open source digital library softwares for institutional repository: Special reference to Dspace and greenstone digital library*, International Journal of Library and Information Science Vol. 2(1) pp. 001-010, February, 2010 Available online <http://www.academicjournals.org/ijlis>. [accessed from <http://www.academicjournals.org/ijlis/PDF/pdf2010/Feb/Biswas%20and%20Paul.pdf> on 3/October 2012] .
2. Wikipedia(2012), *Free Software*, [accessed from http://en.wikipedia.org/wiki/Free_software on 5/October/2012]
3. Open Source Initiative (2012), *The Open Source Definition*, [accessed from <http://opensource.org/docs/OSD> on 9/October/2012]
4. Forbes (2012), *Why Open-Source Principles Are a Recipe For Innovation*, [accessed from <http://www.forbes.com/sites/ashoka/2012/07/25/why-open-source-principles-are-a-recipe-for-innovation/> on 9/October/2012]
5. Eric Von Hippel (2012), *Open Source Software Projects as User Innovation Networks*, [accessed from <http://mitpress.mit.edu/books/chapters/0262562278chap14.pdf> on 5 October 2012]
6. Lakhani.K.R. & Von Hippel.E(2002), *How open source software works: "free" user-to-user assistance*, Research Policy 32 (2003) 923–943 [accessed from http://www.ee.oulu.fi/~vassilis/courses/socialweb10F/reading_material/2/lakhani00-HowOpenSourceSoftwareWorks.pdf on 3/October 2012]
7. United Nations University(2011), *Free and Open Source Software in Sub Saharan Africa*, [accessed from <http://unu.edu/publications/articles/free-and-open-source-software-in-sub-saharan-africa.html> on 25/October/2012]
8. Ryan Paul(2006), *Surveys show open source popularity on the rise in industry* [accessed from <http://arstechnica.com/uncategorized/2006/01/6017-2/> on 10 October 2012]
9. Webopedia(2012), *Apache Web Server* [accessed from http://www.webopedia.com/TERM/A/Apache_Web_server.html on 17 October 2012]
10. LinuxPlanet(2012), *6 excellent Linux/Open Source Web Servers*, [accessed from <http://www.linuxplanet.com/linuxplanet/reviews/7239/1> on 16/October/ 2012]
11. The Apache Software foundation(2012), *The Apache Software foundation*, [accessed from <http://www.apache.org/foundation/faq.html> on 16/October/ 2012]
12. NTC Hosting(2012), *Apache*, [accessed from <http://www.ntchosting.com/apache-web-server.html> on 15/October/2012]

13. Apache(2012), *The Apache HTTP Server Project*, [accessed from http://httpd.apache.org/ABOUT_APACHE.html on 16/October/2012]