A Due Diligence Process for

Releasing Proprietary Source Code under an Open Source License

Ibrahim Haddad, Ph.D.
Head of the Open Source Group
Samsung Research America

www.linuxfoundation.org

As companies realize the benefits of open sourcing proprietary technologies and building development communities around it, the need for a formalized process emerges to manage that transition from proprietary to open source. This paper presents an overview of a generic process that can be used to guide the legal, technical, and business considerations related to IP due diligence before making proprietary technology available under an open source license.

LINUX
FOUNDATION

Introduction

Regardless of the project, internal proprietary technology is almost always covered by varying degrees of IP. As such, when releasing code as open source, it is important to have a rigorous process in place to fully evaluate the consequences of the IP that will accompany the source code. This is to help ensure that the company has anticipated any impact the release may have on other products within their product or IP portfolio.

This paper does not offer any legal advice, and should not be viewed as such. Rather, it provides a practical approach

on due diligence process that can be deployed when evaluating the value and risks associated with the open sourcing of proprietary technology.

Setting up this process the first time is often complex, and involves significant time and effort. However, if the process is designed properly, the time and costs of each subsequent pass will decrease.

This paper describes the major steps in the process, and provides high level guidance on what inputs and outputs should be expected in each step.
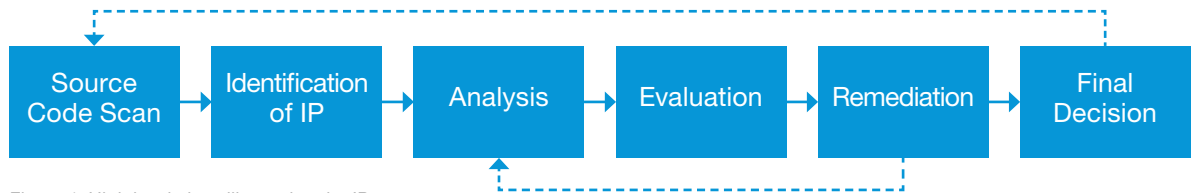
Source Code Scan	Identification of IP	Analysis	Evaluation	Remediation	Final Decision

Figure 1: High level chart illustrating the IP Due Diligence process (minimal iterations)

Overview of the Process

If you do not currently have a process in place, this paper provides you the basics to get started. Generally speaking, such a process must be methodical, well-planned, and thorough to capture all the details and nuances required to reach a decision by the time you are about to exit the process. Every company have their own tailored process for IP due diligence in relating to open sourcing code.

The first time you go through the process, it will be a time-consuming and resource-intensive exercise. You will be basically trying to make it as you go without much precedence. You will adjust the process based on the experience and most of the work will be manual. As you go through the process multiple times, you will be able to identify ways to make it more efficient and to incorporate some tooling or automation. Some companies even invest in specialized tooling that allows them to tie together their internal patent system with their open source compliance tools and some other proprietary tools.

A Due Diligence Process for Releasing Proprietary Source Code under an Open Source License

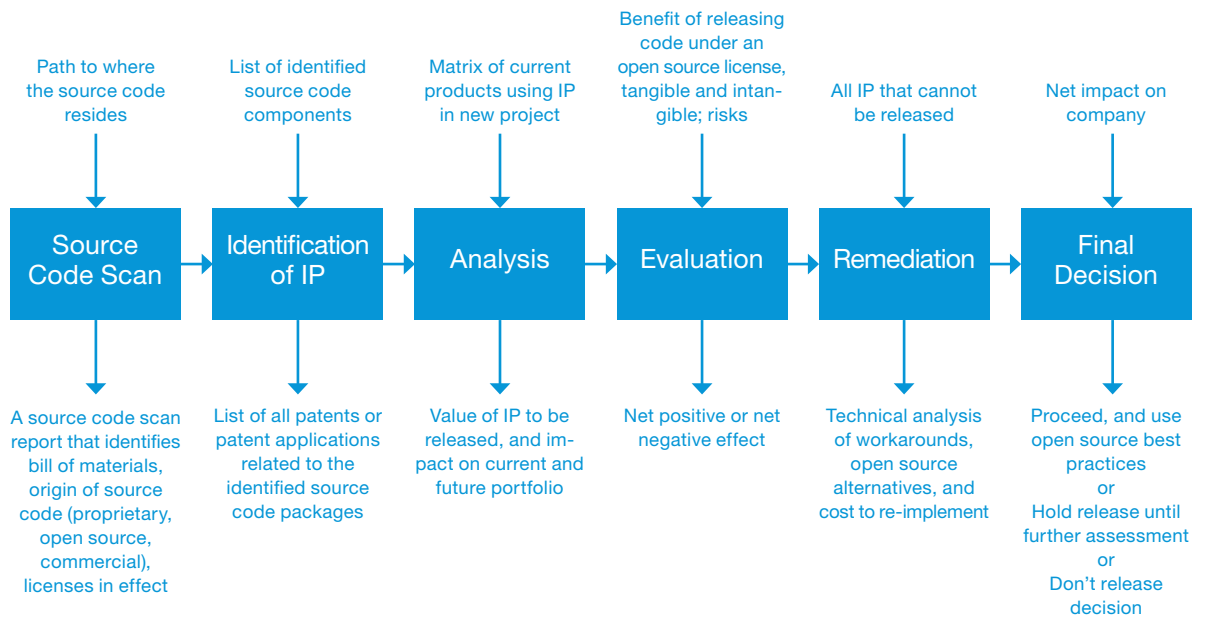| Path to where the source code resides	List of identified source code components	Matrix of current products using IP in new project	Benefit of releasing code under an open source license, tangible and intangible; risks	All IP that cannot be released	Net impact on company
Source Code Scan	**Identification of IP**	**Analysis**	**Evaluation**	**Remediation**	**Final Decision**
A source code scan report that identifies bill of materials, origin of source code (proprietary, open source, commercial), licenses in effect	List of all patents or patent applications related to the identified source code packages	Value of IP to be released, and impact on current and future portfolio	Net positive or net negative effect	Technical analysis of workarounds, open source alternatives, and cost to re-implement	Proceed, and use open source best practices or Hold release until further assessment or Don't release decision

Figure 2: Sample inputs and outputs of every step in the process

Figure 2 illustrates the inputs and outputs of every phase in the process. In the following sub-sections, we examine what happens within that step of the process, who carrier the work, and the expected outcome.

SOURCE CODE SCANNING TOOLS

There are many companies that provide professional tools that can assist you in identifying the license and origin of the source code. These solutions vary by coverage and functionalities. Some offer a complete end to end solution for compliance and project management; some offer a cloud based solution; others are specialized and focused on a specific aspect. For the purpose of this article, you need the ability to identify two important aspects of the source code:

1. Its origin – Who is the author of the code? How did this code end up in our source code repository? Where did it come from?

2. Its license – Is it 100% proprietary? Is it licensed under a 3rd party commercial license? Does it include code license under an open source license?

Companies that provide commercial solutions that can assist you with this exercise include:

- Antelink
- Black Duck Software
- nexB
- OpenLogic
- Palamida
- Protecode

There is also an open source solution that can be of help. It's called **FOSSology**. Fossology allows you to identify software licenses using a heuristic pattern-matching algorithm.

LINUX FOUNDATION

Analyze the Source Code

This step involves scanning all source code packages planned to be open sourced to identify the origin and licenses of the various bodies of code, as they can be proprietary, licensed from a 3rd party, or licensed under an open source license.

The primary executor of this step is the Open Source Compliance Officer who manages the company's open source compliance program. Sometimes that officer is called Director of Open Source or Manager of the Open Source Office.

Ideally, the company maintains proper open source compliance practices and records that would allow it to generate a report quite easily without having to do a new source code scan.

The output from this step is a report that report that identifies origin and license(s) of the source code. Immediate problems can be spotted if, for example, the source code includes code from a 3rd party and licensed under a proprietary license. For such code, you will need to get the 3rd party to approve the change of license and this can be difficult in most cases, or you will need to replace the 3rd party code with your own which can be a major undertaking.

ON OPEN SOURCE COMPLIANCE

As the use of Linux and other open source software has exploded in recent years, especially in mobile and consumer electronics products, the need has arisen for a trusted, neutral, non-commercial compliance program that offers a comprehensive offering of compliance training, tools and services. To address that complexity, The Linux Foundation has developed a set of open source tools, training curricula and a new self-administered assessment checklist that will allow companies to ensure compliance in a cost-effective and efficient manner. The Open Compliance Program also includes a new data exchange standard (SPDX) so companies and their suppliers can easily report software information in a standard way.

For more details, please visit: **http://compliance.linuxfoundation.org.**

Identify IP in the Source Code

The goal of this step is to identify all IP used in the source code components intended for open sourcing. This is not easy an exercise as it involves surveying patents and patents applications, and mapping them to the identified source code components. There are some tips and tricks that you can use to minimize the amount of efforts.

There are no commercial or open source off-the-shelf tools that can facilitate this exercise. Companies that do this process often develop their own tools. Alternatively, you can use your internal tool to manage your patent portfolio as a major asset in terms of researching

LINUX FOUNDATION

patents given certain keywords or authors or even teams (if you know which team was responsible for developing a certain piece of technology).

The output of this step is an identification of any patents that map directly into the source code in question.

This step in the process is driven by the Legal team and internal patent agents.

Analysis

There are three analyses that will happen in parallel: Legal, technical and business – lead by different teams.

Legal Analysis

The legal analysis is driven by Legal staff with the goals to:

1. Identify pending and resolved IP lawsuits related to the IP

2. Identify companies involved, and current claims that did not make it to court yet

3. Identify the need in current and possible future products for IP protection from these patents

Technical Analysis

The technical analysis is driven by senior technical staff that are very familiar with the technology in question. The goal of the technical analysis is to identify the difference between current technology and patented technology to understand the type and value of patents to be made available. For example, you may have a patent that is 5 years old but current technology is far more superior and the patent does not have much value from that perspective given aging technology.

Business Analysis

The business analysis is driven by the business P&L owner. The goals of such an analysis are to:

- Identify potential market for products / services that could come out of the patents in question
- Identify current income via royalties, if any
- Identify the business risks from making the patents in question available
- Identify how could competitors exploit the source code made available for their own good outside of the scope of the intended use (after the code is made available under an open source license)

The results of all technical, legal and business analyses is then merged and discussed by the proper committee whose job is to make the decision. There is no clear map to guide the "Go"

LINUX FOUNDATION

or "No-Go" decision on this, as every case is different and the level of risk tolerance on all of these analyses may vary as well.

IP Remediation

If there as IP identified and the result is a No-Go for making the technology / source code available under an open source license, there are possible paths to explore such as evaluating the possibility of working around the issues raised part of the analyses.

Example common issues include:

- The source code include 3rd party IP that is licensed under a commercial license
    - Possible remediation:
        - Re-negotiate terms of license to allow open sourcing.
        - Create your own replacement for the 3rd party source code and make available under an open source license (ideally the same license under which you are planning to release the remaining of the source code)
- You have identified IP as part of the technology:
    - Possible remediation:
        - Provide a clean-room implementation to work around IP
        - Locate an existing open source component that provides similar functionality and drive your contributions to that projects versus creating a new project

Final Decision

There are two possible outcomes from the process:

- "Go" - approved
- "No-Go" – denied

If the decision is a "Go" – to open source, then you proceed according to best practices of establishing a new open source project, building a development community around it, and ensuring its success via the adoption of open development methods and processes, transparent governance, etc. this will be the topic for a future planned paper.

If the decision is a "No-Go", at this conjunction, you need evaluate the extent of changes required to overcome the identified issues and proceed with a proper remediation plan (if possible), or exit the due diligence process with a No-Go decision.

Conclusion

IP due diligence is a must-do step prior to making proprietary technology available under an open source license. Such due diligence often requires a well-defined extensive cross-functional collaboration between Legal, Business and Technical teams. The success of such

a process depends on the availability of a solid compliance program as a starting point to allow identification of software components, origin and licenses, and on the availability of proper IP inventory allowing you to have a direct mapping between your IP assets and the actual source code components. This paper presented a high level process that can be used to guide you through the exercise of creating your own internal process. Future papers will examine the steps that follow a "Go" decision in terms of internal preparations that you need to go through prior to making your source publically available.

Happy "Open Sourcing"!

About the Author

Dr. Ibrahim Haddad is the Head of the Open Source Group at Samsung Research America, the North American R&D Center of Samsung Electronics. Prior to joining Samsung, Haddad was a member of the management team at The Linux Foundation responsible for technical, legal and compliance projects and initiatives. At the Linux Foundation, he worked with the largest technology companies and with the Open Source community to facilitate a vendor-neutral environment for advancing the Linux platform for next-generation computing devices.

Haddad's career started at Ericsson Research where he spent five years focusing on advanced research for system architecture of 3G wireless IP networks and on adopting Linux and Open Source software in carrier grade environments. He then joined Motorola as Director of Technology Portfolio managing the Open Source Technology Group and driving Motorola's Open Source initiatives. After Motorola, Haddad ran the Open Source Office at Palm and was responsible for all Open Source activities related to webOS, and later supported HP with open sourcing webOS to become the open webOS project.

Dr. Haddad graduated with Honors from Concordia University (Montréal, Canada) with a Ph.D. in Computer Science. He is a Contributing Editor to the Linux Journal, Co-Author of two books on Red Hat Linux and Fedora, and Technical Editor for four books on Linux System Administration, Fedora Linux and Ubuntu Linux. He is fluent in Arabic, English and French.

LINUX FOUNDATION