Server Side Public License

Cloud providers and the future of open source

Marrocco Simone

February 2, 2023

Contents

0 Acknowledgments

The idea behind this document came from the video "Fork you ElasticSearch! How Open Source Works" [1] from the "Fireship" youtube channel, where he talked about the recent (for that time) change of license of the popular open source software Elasticsearch, what it means for the open source community and most importantly what open source is.

In the pages that follow we are expanding the concepts behind this video, going deeper into the problems that companies that have an open source business model face, discussing what open source is, and how to balance the idea of free software with the necessities of protecting free software developers from tech giants that use their programs to make millions but give back nothing in return.

[1] https://youtu.be/tzq4asJegKY

1 History of the Server Side Public license

1.1 Open source software abuse by cloud companies

Millions of people work every day on open source software. People volunteer their time to create or improve digital tools that everyone can use. This is done in a way that allows everyone to look at your code so that it can be free: the idea is that you, as a user, should have total control over the program, and not the other way around.

Giving the software you created for free does not mean that you cannot build a business model upon it. There are a lot of companies that managed to earn from offering consulting or their software as a service. For example, database creators may offer a cloud service with bonuses, like scalability and automatic backups, to help other businesses store their application data without the need for a complex initial setup.

But this can become difficult when a big company providing cloud services, like Google or Amazon, offers the same type of service, gaining from the fact that the software itself is free. To make things worse, the creators may use the same cloud storage services to save their data, effectively being both competitors and partners and making it difficult to compete against them. When dealing with the same product, clients may prefer the bigger company option just because of its brand or because they aggregate multiple products altogether. [2] [3]

Normally, this would not be a problem: while being a competitor, the core of Open Source is that every user contributes to the product and makes it better. But the "Cloud Giants" do not do that, preferring to just use it without helping the original creator, either by improving the software or contributing economically. "[Amazon has] taken Redis' open source project and packaged it as a service and monetized that," Manish Gupta, CMO of Redis Labs, told Business Insider. "Their contribution to the open source is minuscule, not even in the 1 percent range." [4]

[2] https://siliconangle.com/2018/12/22/stormy-weather-open-source-software-firms-limit-licenses-stop-cloud-giants/

[3] https://www.forbes.com/sites/glennsolomon/2020/10/13/finding-trillions-in-the-clouds-how-open-source-companies-can-win-in-a-world-dominated-by-amazon-microsoft-and-google/

[4] https://www.businessinsider.com/mongodb-and-redis-labs-have-a-new-plan-to-take-on-amazon-2018-10

1.2 Color.Js and Faker.Js

Some programmers, discouraged by these practices, decide to abandon their open source projects. Marak was the developer of two popular libraries, Color.js and Faker.js, for the NodeJs environment. The Color library received over 20 million weekly downloads on npm alone and had almost 19,000 projects relying on it. Whereas, Faker received over 2.8 million weekly downloads on npm, and had over 2,500 dependents. On January 9, 2022, he committed a new version of both libraries, deleting the entire code and replacing it with lines that broke every software that used them. [5]

The creator of those libraries, expressing feelings of dissatisfaction, posted on Github: "Respectfully, I am no longer going to support Fortune 500s (and other smaller-sized companies) with my free work. There isn't much else to say. Take this as an opportunity to send me a six-figure yearly contract or fork the project and have someone else work on it" [6]. Even though the open source ideology is not to make money, volunteers are not happy working freely, only to then have other businesses use the library without contribution. Various people commented on Marak's post, encouraging his actions.

1.3 MongoDB

Companies that started as open source projects and invested heavily try to fight back against Cloud Giants becoming competitors without contributing. And they usually do that by restricting their license.

MongoDB is one of the companies that fought against Amazon by creating a new license, the Server Side Public License, for its software new version. Built on top of the General Public License v3, it aims to discourage other people from offering its software as a cloud service unless they are willing to share the specifications and the code of the entire infrastructure used, in a strong copyleft way. We will discuss later in detail what this license means for the Open Source community.

On their website, MongoDB said about the change: "The market is quickly moving to consume most software as a service. This is a time of incredible opportunity for open source projects, with the potential to foster a new wave of great open source server-side software. The reality, however, is that once an open source project becomes interesting, it is too easy for large cloud vendors to capture all the value but contribute nothing back to

[5] https://infotechlead.com/software/github-developer-marak-squires-corrupts-open-source-libraries-70616

[6] https://web.archive.org/web/20210704022108/https://github.com/Marak/faker.js/issues/1046

the community. The community needs a new license that builds on the spirit of the AGPL, but makes explicit the conditions for providing the software as a service." [7]

However, things did not go as planned for the company. Debian, Red Hat Enterprise Linux, and Fedora later dropped MongoDB, after deeming the license problematic. Furthermore, the change did not stop Amazon from offering a competing service: the AWS team forked the previous version, licensed under the Affero GPL, and created DocumentDB, a similar product fully compatible with its predecessor [8]: "Given that DocumentDB is designed to work with a version of MongoDB released before that license went into effect, the SSPL doesn't appear to apply to DocumentDB. But AWS thinks that it will be able to help companies that have tried and struggled to implement MongoDB on their own achieve better performance and scale than MongoDB (the corporation) can provide." [9]

1.4 Other companies following up

While the Open Source community discussed the problems that pose both this new type of license and the power of the cloud providers, other companies followed the example of MongoDB.

1.4.1 Elastic

Elastic, the owner of the two popular tools Elasticsearch and Kibana, decided to change its license from Apache 2.0 to a dual license, one of which is SSPL. The user can choose between SSPL and a proprietary license which allows the use of the software freely but prohibits hosting the software as a service without Elastic consent. In their blog, the company wrote: "So why the change? AWS and Amazon Elasticsearch Service. They have been doing things that we think are just NOT OK since 2015 and it has only gotten worse. If we don't stand up to them now, as a successful company and leader in the market, who will? We think that Amazon's behavior is inconsistent with the norms and values that are especially important in the open source ecosystem. Our hope is to take our presence in the market and use it to stand up to this now so others don't face these same issues in the future." [10], "We do not have a commercial relationship with AWS on the Amazon Elasticsearch

[7] https://www.mongodb.com/licensing/server-side-public-license/faq

[8] https://en.wikipedia.org/wiki/Server_Side_Public_License

[9] https://www.geekwire.com/2019/amazon-web-services-calls-mongodbs-licensing-bluff-documentdb-new-managed-database/

[10] https://www.elastic.co/blog/why-license-change-aws

Service. We do not actively support that service, and no longer want our investments in our software to directly benefit that service. For transparency, we also have ongoing litigation with AWS" [11]

1.4.2 Redis

Redis is another popular database that focuses on the velocity of retrieving information. Its main strength comes from six modules that allow Redis to be used as a SQL, NOSQL or other popular database structure choices. The company working on the project was one of the first to fight against cloud vendors. In August 2018 the CTO Yiftach Shoolman wrote how "Cloud providers have been taking advantage of the open source community for years by selling (for hundreds of millions of dollars) cloud services based on open source code they didn't develop (e.g. Docker, Spark, Hadoop, Redis, Elasticsearch and others). This discourages the community from investing in developing open source code" [12] and decided to change the license from Apache 2.0 to one called Common Close License, which is more restrictive. In February 2019, to clear the confusion on this license, Redis modules changed to a proprietary one called Redis Source Available License (RSAL). "Over time, other respected open source companies, like MongoDB and Confluent, created their own proposals for modern variants to open source licensing. Each company took a different approach, but all shared the same goal — stopping cloud providers from taking successful open source projects that were developed by others, packaging them into proprietary services, and using their market power to generate significant revenue streams" [13]. In November 2022 they followed Elastic by double licensing their modules under both the second version of RSAL and SSPL, calling the MongoDB one "the de facto standard for source available licenses". "RSALv2 is a permissive non-copyleft license, allowing the right to "use, copy, distribute, make available, and prepare derivative works of the software" and has only two primary limitations. Under RSALv2, you may not: Commercialize the software or provide it to others as a managed service; Remove or obscure any licensing, copyright, or other notices." [14]

[11] https://www.elastic.co/pricing/faq/licensing
[12] https://redis.com/blog/redis-license-bsd-will-remain-bsd/
[13] https://redis.com/blog/redis-labs-modules-license-changes/
[14] https://redis.com/blog/rsalv2-sspl-announcement/

1.4.3 Confluent

Some companies decided to block entirely the possibility of cloud giants abusing their software. Confluent Inc. is the primary developer of a managing system for Apache Kafka, a popular open source tool used to pipeline an enormous stream of data offered under Apache 2.0. Confluent main work is on components built upon Kafka to improve it. After Amazon announced that it was now offering the same software developed by Confluent as a competing service, the company changed its license to the Confluent Community License, which mostly allows the same liberties as an Open Source one, but prohibits offering the software as a service [15]. "What this means is that, for example, you can use KSQL however you see fit as an ingredient in your products or services, whether those products are delivered as software or as SaaS, but you cannot create a KSQL-as-a-service offering" [16]. In the same blog post explaining why the change, the CEO Jay Kreps writes, to justify his decision, that "the cloud providers have significant advantages: they control the pricing of all resources a service provider will use and can tightly integrate their own services across all their offerings. [...] [Some of those] take the open source code, bake it into the cloud offering, and put all their own investments into differentiated proprietary offerings". The co-founder then explains how this move was not done for greed or a financial problem of the company, as the project was not started to make money yet it is doing exceedingly well.

[15] https://siliconangle.com/2018/12/22/stormy-weather-open-source-software-firms-limit-licenses-stop-cloud-giants/

[16] https://www.confluent.io/blog/license-changes-confluent-platform/

2 Is SSPL an open source license?

2.1 Changes between SSPL and GPL3

"The only substantive modification is section 13, which makes clear the condition to offering MongoDB as a service. A company that offers a publicly available MongoDB as a service must release the software it uses to offer such service under the terms of the SSPL, including the management software, user interfaces, application program interfaces, automation software, monitoring software, backup software, storage software and hosting software, all such that a user could run an instance of the service using the source code made available." [17]

The SSPL is a different license that tries to solve the same problem as the Affero GPL, which says that it "is a modified version of the ordinary GNU GPL version 3. It has one added requirement: if you run a modified program on a server and let other users communicate with it there, your server must also allow them to download the source code corresponding to the modified version running there" [18]. While the AGPL was made to expand the copyleft to programs running on a server, it did not solve the problem of cloud vendors hiding their entire infrastructure behind proprietary code. Effectively, they were scaling up and profiting from open source software without sharing the advancement they made with the community that offered the same profitable code. The SSPL blocks cloud giants from serving open source software since they are not willing to share how they created their infrastructure, while, theoretically, if someone else wanted to offer the same SaaS as the original developer it could do so by sharing the full architecture of its business.

[17] https://www.mongodb.com/licensing/server-side-public-license/faq
[18] https://en.wikipedia.org/wiki/GNU_Affero_General_Public_License

2.2 Why SSPL is not an open source license

The Open Source Foundation has deemed the SSPL to be not an open source one [19]. The SSPL takes rights away from the user, in violation of the Open Source Definition 6 "No Discrimination Against Fields of Endeavor" which states: "The license must not restrict anyone from making use of the program in a specific field of endeavor" [20]. Furthermore, the OSF explains how this loss of rights is especially hard on those volunteers who contributed to the software and then cannot enjoy the fruits of their work anymore. In another article, the OSF explains the dangers of "fauxpen" licenses: "Most fauxpen source licenses plainly don't comply with the requirements of the open source definition. The source code may be available, but they are not open source" [21]. The OSF then accuses the fauxpen source vendors to disrespect the open source principles and use the name as a business stratagem to attract clients.

Other free software institutions also deemed problematic this new license in contrast with the principles that promote sharing of information and freedom to access programs. Fedora spokesman Tom Callaway tells in the Fedora Forum how "the SSPL is intentionally crafted to be aggressively discriminatory towards a specific class of users. Additionally, it seems clear that the intent of the license author is to cause Fear, Uncertainty, and Doubt towards commercial users of software under that license" [22].

"The SSPL is clearly not in the spirit of the DFSG (Debian's free software guidelines), let alone complimentary to the Debian's goals of promoting software or user freedom", mentioned Chirs Lamb, Debian Project Leader [23].

Richard Fontana, senior commercial counsel at Red Hat Inc., tells "I have some concerns over what cloud providers are doing, but my bigger concern is how [Redis, Confluent and MongoDB] are reacting." Fontana said he's particularly concerned about contract language that uses open source terminology to define licenses that are really proprietary. [24]

[19] https://opensource.org/sspl-not-open-source

[20] https://opensource.org/osd

[21] https://opensource.com/article/19/4/fauxpen-source-bad-business

[22] https://lists.fedoraproject.org/archives/list/devel@lists.fedoraproject.org/thread/IQIOBOGWJ247JGKX2WD6N27TZNZZNM6C/

[23] https://hub.packtpub.com/red-hat-drops-mongodb-over-concerns-related-to-its-server-side-public-license-sspl/

[24] https://siliconangle.com/2018/12/22/stormy-weather-open-source-software-firms-limit-licenses-stop-cloud-giants/

2.3 "Free and Open" products

It should be noted that the open source institutions mentioned above do not suffer from cloud vendors as much as the businesses adopting SSPL. Those companies still consider themselves to be part of the community that shares the same principles of free software, freedom of usage and leaving the source code open for everyone to look at and modify: while they have written in their blogs that they should not be named as Open Source, they now refer to themselves as free and open: "While we have chosen to avoid confusion by not using the term open source to refer to these products, we will continue to use the word "Open" and "Free and Open." These are simple ways to describe the fact that the product is free to use, the source code is available, and also applies to our open and collaborative engagement model in GitHub. We remain committed to the principles of open source - transparency, collaboration, and community" writes Elastic in his FAQ explaining the change of license [25].

Heather Meeker is one of the world's foremost legal experts on open source software licensing and compliance. She's authored the go-to book on the topic, Open Source for Business, and is a General Partner at OSS Capital [26]. She's been the main helper in creating the SSPL, the RSALv2 and the Elastic License [27]. In the Commercial Open Source Software (COSS) community, she writes about open source and what these new licenses mean. In her post "SSPL Re-Takes the Stage in 2021" [28] she criticizes the OSF for its lack of transparency on the criteria of whether a license is open source or not. To justify the open-sourceness of the SSPL she starts with the "Freedom Zero": "The freedom to run the program as you wish, for any purpose" [29]. SSPL was the first license trying to be open source while limiting the power of cloud companies, unlike the other source available licenses - licenses with restrictions that do not meet the OSD, like the Confluent Community one. The discussion born from the MongoDB license created such a fierce debate that the request to the OSF was removed before a final verdict (whether SSPL is open source or not). In particular, the arguments against the SSPL were mostly against its creators, which were deemed looking for profit more than driven by a sense of morality. Another problem is the lack of definition of what "guaranteeing software freedom" actually means since the same case used for SSPL could be used against the AGPL.

[25] https://www.elastic.co/pricing/faq/licensing
[26] https://fossa.com/blog/q-a-heather-meeker-hot-topics-oss-license-compliance/
[27] https://redis.com/blog/rsalv2-sspl-announcement/
[28] https://www.coss.community/cossc/sspl-re-takes-the-stage-in-2021-2koa
[29] https://www.gnu.org/philosophy/free-sw.en.html

3 Personal Considerations on the subject

3.1 The goals of Free Software, and why SSPL respects them

The core of the Free Software ideology is that the user should have full control of the programs that run on their machines. Free, as in free speech, means that people should see the source code of the program to understand what it does, and to be sure that it does not hide trackers or malicious code. Opening the source code means that people can run and modify the software so that it can be improved. Nonexpert users know from the community that the program is safe, and there will always be a volunteer who will make the experience better. The success of this ideology stands in the fact that knowledge and ideas are shared, without bureaucracy and legal blocks, so that anyone can build on top of other people's code to advance the entire digital ecosystem. Richard Stallman, the creator of the GNU system, talks about this in a TedTalk available on YouTube called "Free software, free society: Richard Stallman at TEDxGeneva 2014" [30].

For the vast majority of users, the SSPL does not remove this power from the software licensed with it. Its goals remain the same as the free software community, and the SSPL was created with the specific objective of blocking monolitic corporations that have the power to hold for themselves what the community created together. MongoDB could have just gone the same way as other companies did, restricting a user's possibility of usage of their code, but instead, they have tried a way to save the open source community by obligating the cloud giants to participate in the open source environment, a thing they have always be reluctant. The license is by no mean perfect and it surely needs corrections, but the best place where this modification can happen is the same one that allows coding as we know it today a much easier task, the open source community.

[30] https://youtu.be/Ag1AKIl_2GM

3.2 Why we should protect Open Source businesses from Cloud Companies

While the vast majority of open source software is small projects run by volunteers, some ideas need much more than a couple of hours each week. Complex projects, like big-scale databases or operating systems, require people to work on it full-time jobs, but they cannot do that if they are not able to earn money for food or rent at the end of the day. Confluent CEO Jay Kreps says, in the blog post explaining its license change: "There are thousands of libraries thriving on GitHub that don't need much investment beyond a few volunteer contributors. Distributed data systems are different. Building a successful new distributed data platform is just excruciatingly hard. [...] What I think is most stark is that the only systems that remained relevant through to today are those that, whatever their origin, managed to develop a stable commercial entity that helped sustain ongoing investment. Those that did this (MongoDB, ElasticSearch, Cassandra, Hadoop) all continue to thrive and have become part of the modern stack. Those that didn't (Voldemort, Dynomite, CouchDB, and a dozen others) have all fallen by the wayside, despite early popularity. They still exist, but most likely you have never heard of them". [31]

There are a lot of ways an open-source business can make money to pay its developers: for example, getting acquired by a big tech company, like Red Hat recently acquired by IBM [32]. But some types of software can only make money by offering their software as a service itself, like the ones we talked about here, and unfair competition that only steals and not contributes means only that products evolve slower for the greed of an external player who can exploit the morality of benevolent people. Open Source business making money is good because it means the best of both worlds: we have a strong company working actively on a product while having the security and the advantages of the open source community watching and suggesting changes on the software itself. But if we allow them to be exploited by bigger companies that gain a lot from this practice, we are telling the open source community that it is not feasible to make money from this type of software, pushing towards closed license programs.

[31] https://www.confluent.io/blog/license-changes-confluent-platform/
[32] https://www.redhat.com/en/ibm

3.3 Tangible problems of the power of Amazon

In the previous paragraphs, it was preferred a neutral stance on the subject, calling against the general "Cloud Vendors". The practical truth, however, is that the main culprit is the Amazon Web Service. Each one of the businesses we talked about fighting against AWS offering a competitive service without contributing back, and when confronted, instead of starting a fruitful collaboration, it always decided to play dirty, forking the project and then offering a compatible version (like DocumentDB or OpenSearch, the forked version of ElasticSearch). Considering that all software using SSPL has still open source code, it usually means that AWS copies the new features and stays on top, without working on them.

Of course, we need to be wary of other cloud vendors, but most open source businesses already have economic ties with them, like Confluent and its partnership with Azure [33], or Elastic: "We have built strong relationships with Google Cloud, Microsoft Azure, Alibaba Cloud, and Tencent Cloud" [34]

This may look like a war for profit, where two competing tech companies want to have as much money as they want, but the bad practices and the monopoly of AWS have real, tangible consequences in the real world, in fields far from tech and software.

In its excellent video "Is Amazon too big to fail?" [35], the YouTuber and disseminator PolyMatter explains how Amazon makes so much money thanks to AWS (nearly 60% of the total income in 2018) that it can allow the main company to compete in virtually any sector (online retail store, video and music streaming, ebooks, physical stores, and many others), at a massive scale and without the need for profit. Most of the time, it allows its other subsidiaries to stay in red to acquire an unfair monopoly in sectors where other competitors, that need to make a profit, cannot compete on the same level.

It is no surprise then how Amazon Shopping can keep the price so low, offering things like free shipping and fast delivery, while the competitors cannot: it does not make a profit from it. And this has big effects on the entire population: local shops must close because they cannot stay afloat, while postmen and warehouse workers suffer inhuman conditions and low pay.

[33] https://www.confluent.io/partner/microsoft-azure/
[34] https://www.elastic.co/pricing/faq/licensing
[35] https://youtu.be/EYPs-ya_GDA

Stories can be quickly found online, and it happens anywhere Amazon operates: in the USA [36], in Italy [37] and in Mexico [38], among many other contries.

This cycle of injustice is possible because Amazon has a large power on cloud infrastructure, a power that it got thanks to all the services it can provide on AWS, where they are integrated all together and it is possible to combine them, for a developer, with just some clicks. And this aggregation of options is made at the expense of open source software creators. Fighting against corporations, via more restrictive licenses, is much more than just considering the user right of offering a SaaS removed: if the original creator cannot compete with cloud vendors, who can?

[36] https://www.theguardian.com/technology/2020/feb/05/amazon-workers-protest-unsafe-grueling-conditions-warehouse

[37] https://www.ilpost.it/2021/10/22/sfruttamento-logistica-emilia-romagna/

[38] https://www.vice.com/en/article/pkb9qn/amazon-to-open-dollar21-million-state-of-the-art-warehouse-in-tijuana-slum