Overview

For a developer or architect, making use of open-source software can be a very attractive proposition to gain access to the latest technological innovation, maximise agility and minimise cost. However, choosing software to bake in to your architecture is a long-term decision and it is important to understand all the implications of your choice.

At Instaclustr, we are 100% focused on open-source software and have spent a lot of time observing how open source projects work and thinking about the implications. A major risk that we evaluate in adopting and recommending open source software is what happens if a dominant entity associated with that software either goes out of business or takes it in a direction that is not good for the community. If your future use of the software is tied to the whim of a single commercial entity (even if you’re not paying that entity today), then many of the benefits of open-source software are greatly reduced (and may evaporate altogether in the future for that software).

To understand these risks for a particular open-source project you need to understand the licensing terms of the software you are using, the health of the ecosystem surrounding the software and the business model of any commercial entities directly involved with the software. All of these factors can have significant future implications once you are committed to a software product.
Understand the licensing terms

Firstly, let’s consider licensing terms. There are lots of good resources available explaining the different open-source licenses that are in use (for example, here and here), but what are the factors that you should consider when evaluating a license? For many people, a major concern will be whether the license risks forcing a leak for the organisation’s IP (so called “viral” or “copyleft” licenses), but issues like limitations on commercial use may also come into play with lesser known licenses.

At Instaclustr, we favour the use of open-source licenses that do not contain a copyleft provision. While IP leakage risk is one factor in this, the bigger factor is that many of the large companies that are essential for a strong independent community for an open-source project will not contribute to or use software distributed with these licenses. This increases the likelihood that maintenance of the software will be dependent on a single entity. Linux, which uses the GPL (Gnu General Public License - a popular copyleft license), is clearly an exception to this rule and shows that, where the software is sufficiently broadly applicable and organisations do not expect to customise for their own use, GPL can still be a successful community licence.

Understand the ecosystem

Secondly, consider the health of the overall ecosystem surrounding the software you are intending to use. An active and broad-based ecosystem means more organisations with a stake in the usefulness and quality of the software you are using and a better chance that it will evolve in a way that suits a broad base of users rather than a specific use case or commercial drivers of a software vendor (for example, adding features for marketing reasons rather than actual usefulness). Specific indicators to look for are:

- Is the software owned and governed by an independent body (for example the Apache Foundation or the Cloud Native Computing Foundation)? This generally ensures multiple organisations are actively contributing to the code base and provides a decision-making mechanism that guides decisions based on technical and user value rather than commercial interests.

- Is the software used by a variety of large, well-known organisations? The last thing you want is to be left as one of a small number of users of a significant piece of open-source software – with all the burden of keeping it up to date until you can migrate away. When you use software that is also widely adopted by large technology-driven organisations (for example, the FAANGs), you can expect that they will share the burden of keeping the software up to date and, if you keep your ear to the ground, you’ll have a good sense of whether they are moving away from the technology, with plenty of time to migrate.

- Does the software have a variety of other businesses depending on it? Consulting companies, managed service providers, companies with software products built on the tech all have a strong incentive to invest resources to keep the the open-source software healthy and share the load of maintenance.
Understand the commercial interests

Finally, in many cases your use of an open-source software product will be strongly associated with a single vendor, either because you are buying support or a managed service from that vendor or because that vendor is a dominant player in the development or maintenance of that software. In this case it’s important to understand the business model of the vendor or vendors as that will drive the behaviour you can expect from them in the future. There are many business models associated with open source and new models emerging all the time. However, we see a few prevalent models:

- **Free open source software (FOSS) IP builder/Open core** – These companies seek to sell software that is, at its core, free open-source software, often owned and governed by an independent body. They will typically have their own proprietary version of the software that adds features in addition to the FOSS version which are kept closed source and often licensed for a fee that reflects the value of both the FOSS software and the additional features. Many companies following this model are very well funded and spend a lot of that money developing and promoting the core software, which benefits the entire community. However, their need to maintain sufficient differentiation for their proprietary version may lead to tension about what gets contributed to the core FOSS project and the reducing value of the proprietary product over time (as FOSS versions of the proprietary features are developed). It is also common to see a tension with these companies seeing themselves as sellers of licenses to IP and customers seeing them as providers of support for the FOSS software and not getting the level of support they expect for the cost. There is also the risk of becoming unwittingly locked in to the proprietary version of the software and having high switching costs if the vendor decides to increase annual fees.

- **Open source IP owner** – These companies are similar to the FOSS IP builders but, rather than being based on FOSS that is owned by an independent foundation, they develop a code base and publicly release the source code. They may accept external contributions to the code base but at the end of the day they maintain complete control of the code and decisions about what features go in the open source version versus their proprietary version. License terms and overall strength of community are extremely important to evaluate when dealing with one of these companies. If the license terms and community strength do not make it likely that an independent community fork of the software could emerge, then your level of vendor lock-in with these providers is really no better than with closed-source software – the vendor has control of what work goes in what version and could make the open-source version unviable whenever it suits them.
- **Cloud Provider** – The big cloud providers often provide semi-managed versions of popular open-source software products. Their primary motivation is clearly to provide a service to their customers that increases the overall use of their platform. They generally won’t drive significant innovation in the products but can be expected to contribute some level of bug fixing and so forth to maintain the quality of their offerings. Community pressure is however starting to increase the level of investment of the big cloud providers back to the projects they use. We view the presence of big cloud provider offerings as a great balance against dominant players in the IP builder or open-source IP owner mode – they have the resources and interest to create and maintain a fork when necessary (for a great example see Amazon Corretto).

- **Managed service provider** – Specialised managed service providers (such as Instaclustr) have similar motivations to the big cloud providers in that they are interested in growing their own user base. While typically having smaller resources they also have more at stake in the success of the FOSS products they support. Unlike cloud providers that will be happy for you to shift between their many products to find the right fit, MSPs will be focused on making the software you’ve chosen a success for you. A good MSP will have strong capability to fix and enhance the core FOSS to meet customer needs (and unlike cloud providers, be prepared to work with requirements at an individual customer level in many cases). An MSP that contributes to a project proportionally to the profit it makes from that project should be seen as a healthy player in an open-source ecosystem. As MSPs are typically smaller players you need to consider the health of the wider ecosystem for the products they support when choosing a product.

**Conclusion**

At Instaclustr, we evaluate all of these factors before offering an open-source product on the [Instaclustr Managed Platform](https://www.instaclustr.com). Currently, this evaluation is a behind-the-scenes decision process but we will soon be releasing an Instaclustr Open Source Certification framework which will provide a structure for us to publish this analysis for specific products for the benefit of the broader community.
About Instaclustr

InStaclustr delivers reliability at scale through our integrated data platform of open source technologies such as Apache Cassandra®, Apache Kafka®, Apache Spark™ and Elasticsearch.

Our expertise stems from delivering more than 30+ million node hours under management, allowing us to run the world’s most powerful data technologies effortlessly.

We provide a range of managed, consulting and support services to help our customers develop and deploy solutions around open source technologies. Our integrated data platform, built on open source technologies, powers mission critical, highly available applications for our customers and help them achieve scalability, reliability and performance for their applications.

Build, run and scale your app with confidence

Like what you see?

If you’re looking at building a proof of concept for your application, or looking for production grade nodes, contact a member of our Sales Team to discuss your specific needs.