



instaclustr

Apache Cassandra Project Assessment Report

Document Identifier: ICF_CPA

Document Version: 1.1

Issue Date: 15 NOV 2021

www.instaclustr.com

Table of Contents

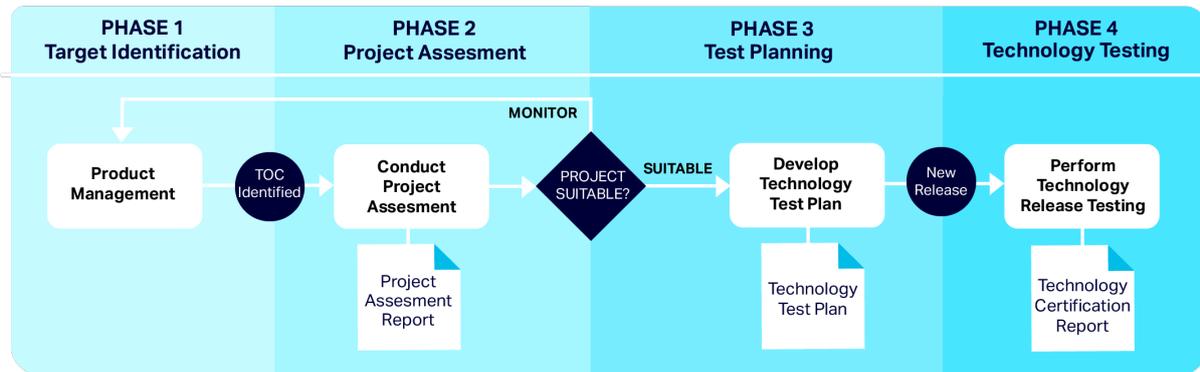
	1
1 Introduction	3
1.1 Overview	3
1.2 Documentation Identification Information	3
1.3 Copyright Statement	4
2 Assessment Approach	5
2.1 Overview	5
2.2 Scope	5
2.3 Purpose	5
2.4 Review and Updates	5
3 Assessment Results	6
3.1 Project Maturity	6
3.1.1 Observations	6
3.2 Development and Maintenance	6
3.2.1 Observations	6
3.2.2 Assessment	7
3.3 Community	7
3.3.1 Observations	7
3.3.2 Assessment	8
3.4 Licensing Model	9
3.4.1 Observations	9
3.4.2 Assessment	9
3.5 Governance	9
3.5.1 Observations	9
3.5.2 Assessment	9
4 Recommendation	10

1 Introduction

1.1 Overview

The **Instaclustr Certification Framework for Open Source Software (ICF-OSS)** provides an analytical basis for selecting and recommending various open source technologies for production usage. This is achieved by:

- a) assessing various open source projects to gain a level of assurance that the project has the foundation and capability to build and sustain production-grade software, and
- b) testing specific versions of open source technologies for function, performance, and interoperability.



1.2 Documentation Identification Information

Identifier	ICF_PAR – Apache Cassandra Project Assessment Report
Description	This document is the Project Assessment Report for Apache Cassandra as part of the Instaclustr Certification Framework for Open Source Software (ICF-OSS).
Target of Certification	Apache Cassandra http://cassandra.apache.org/

1.3 Copyright Statement

© *Instaclustr Pty Limited, 2021*

Except as permitted by the copyright law applicable to you, you may not reproduce, distribute, publish, display, communicate, or transmit any of the content of this document, in any form, but any means, without the prior written permission of Instaclustr Pty Limited.

2 Assessment Approach

2.1 Overview

This document is the Project Assessment Report (PAR) for Apache Cassandra that has been developed as part of Phase 2 of the certification process, the Instacluster Certification Framework. This document reports the result of Instacluster's assessment of the health of the Apache Cassandra project and community as a source of software for significant production usage.

2.2 Scope

This project assessment uses the following broad criteria to assess the health of the Apache Cassandra project:

- Project maturity
- Development and maintenance
- Community
- Licensing model
- Governance

2.3 Purpose

The project assessment phase of the ICF-OSS focuses on the overarching governance and management of the project as a whole. This phase assesses the non-technical characteristics of the project rather than individual technology releases.

The intent is to ensure that all open source technologies that are integrated into the Instacluster Managed Platform are being effectively managed and controlled, that risks associated with licensing and use are suitable, and that there are strong indications the project will enjoy ongoing, open community support in the foreseeable future.

2.4 Review and Updates

This assessment and report will be updated on an annual basis, or if any significant changes occur in the Apache Cassandra Project.

3 Assessment Results

3.1 Project Maturity

3.1.1 Observations

Cassandra was first open sourced by Facebook in 2008, became an Apache Foundation Incubator project in March 2009 and a top-level project in 2010.

Apache Cassandra is very widely adopted with large scale, publicly acknowledged large-scale users including:

- Apple
- Netflix
- Instagram
- Uber

Some of these companies are known to run thousands of nodes of Cassandra.

Other indicators of the adoption of Cassandra include:

- Included in 3.2k stacks on stackshare.io (compared to, for example, 87.7k for Java—a very broadly adopted technology)
- Rated number 10 on dbengines.com
- Instaclustr currently manages and supports over 3000 clusters running Apache Cassandra

3.1.2 Assessment

The Apache Cassandra project is considered to be very mature with a significant user base and deployment.

The Apache Cassandra project maturity is assessed as **EXCELLENT**.

3.2 Development and Maintenance

3.2.1 Observations

Apache Cassandra 4.0 was released on 18 July 2021. This was the first major release of Cassandra in 6 years (although there have been a steady stream of minor releases). The months leading up to the Cassandra 4.0 release saw a sustained increase in contributions from a number of parties. The project has also seen a notable increase in new committers with 11 announced over the last 12 months. Overall, the development community and effort for Cassandra appears to be healthier than ever.



Figure 2: Github activity for the Apache Cassandra project

3.2.2 Assessment

Velocity of change has picked up leading into the 4.0 released and beyond. With the community of very large scale users we do not see a significant risk that the project will not be well maintained for the foreseeable future.

The ongoing development and maintenance of the Apache Cassandra project is assessed as **GOOD**.

3.3 Community

3.3.1 Observations

Apache Cassandra has a broad and active community. Indicators include:

- Active user mailing list and community Slack channel
- Stack Overflow trends demonstrates an active flow of questions (see Figure 3 below). It is, however, fair to say that the trend of reduced questions reflects Cassandra maturing beyond the hype phase of a new technology.
- The first results page of Google search results for “Cassandra consulting” includes 10 companies with significant Cassandra focus.

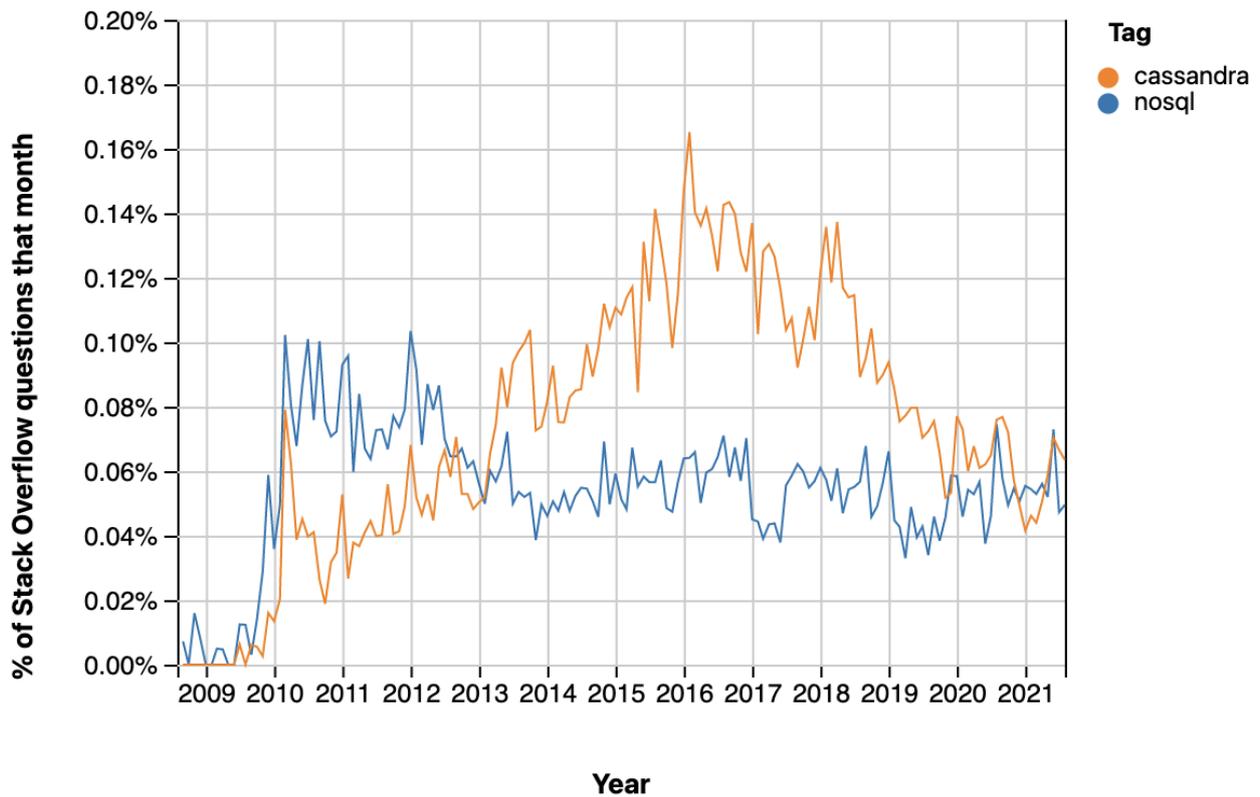


Figure 3: Stackoverflow questions related to Apache Cassandra

It is notable from these statistics that most indicators of community around Cassandra have plateaued and/or are showing slow decline. We view this as a natural part of the new technology hype cycle and a sign that developers are better understanding appropriate use cases for Cassandra rather than any long-term issue with the health of the technology.

The indicators relating to the broader ecosystem surrounding Apache Cassandra is also strong, these include:

- Multiple companies with Cassandra as a core focus of their business (DataStax and Instaclustr being well known examples).
- Several vendors offering Cassandra specific add-on products (for example <https://www.rubrik.com/solutions/cassandra/>)
- Many commercial and open source products using Cassandra as a storage engine. (If you Google “configure Cassandra storage” you will see quite a few techs in the first couple of pages of search results. We are also aware of several commercial products that use Cassandra for back-end storage but do not publish the fact.)

3.3.2 Assessment

The Apache Cassandra project has a wide and varied active community.

The community is assessed as **GOOD**.

3.4 Licensing Model

3.4.1 Observations

Apache Cassandra is licensed under the Apache 2.0 license. This license is very open and, in Instaclustr's view, strongly encouraging of use and contribution to the software.

A good summary of licence conditions can be found here:

[https://tldrlegal.com/license/apache-license-2.0-\(apache-2.0\)](https://tldrlegal.com/license/apache-license-2.0-(apache-2.0))

3.4.2 Assessment

The Apache Software Foundation's Apache 2.0 license considered the benchmark for open source technologies.

The licensing model is assessed as **EXCELLENT**.

3.5 Governance

3.5.1 Observations

The Apache Cassandra project is an Apache Foundation project and governed according to The Apache Way (<http://www.apache.org/foundation/governance/>). This governance mechanism has been proven to allow multiple parties to contribute to project in a meritocracy of ideas and contributors and has generated many extremely successful projects (for example Hadoop, Open Office, and Tomcat).

The Project Management Committee for Cassandra has representatives from multiple organizations (Apple, DataStax, and The Last Pickle having significant representation).

3.5.2 Assessment

The governance, control and openness associated with the Apache Software Foundation is considered the benchmark for open source technologies.

The governance model is assessed as **EXCELLENT**.

4 Recommendation

The Apache project is considered **HIGHLY SUITABLE** for inclusion on the Instaclustr Managed Platform.

CRITERIA	ASSESSMENT	RESULT
PROJECT MATURITY	The Apache Cassandra project is considered to be very mature with a significant user base and deployment.	EXCELLENT
DEVELOPMENT AND MAINTENANCE	While the community had a difficult few years, the push for the Cassandra 4.0 release has seen broad input further strengthened by the announcement of 11 new committers in 12 months.	EXCELLENT
COMMUNITY	The Apache Cassandra project has a wide and varied active community.	GOOD
LICENSING MODEL	The Apache Software Foundation's Apache 2.0 license considered the benchmark for open source technologies.	EXCELLENT
GOVERNANCE	The governance, control and openness associated with the Apache Software Foundation is considered the benchmark for open source technologies.	EXCELLENT