The History of Apache Cassandra™

cassandra

2007

Avinash Lakshman and Prashant Malik

Facebook engineers Avinash Lakshman and Prashant Malik develop Cassandra to power Facebook's inbox search feature using large datasets across multiple servers. They name their database after the Trojan mythological prophet Cassandra - with classical allusions to a curse on an oracle. Lakshman's presentations and the team's LADIS (Large Scale Distributed Systems) paper generate excitement in the distributed systems community.

July 2008

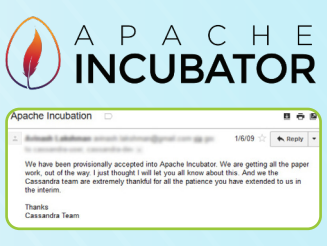
Facebook releases Cassandra as an open source project on (now defunct) Google Code.

Nov/Dec 2008

Rackspace hires distributed systems engineer Jonathan Ellis with a mandate to build a next-generation scalable database. After evaluating the extant open source projects, he forms a small Cassandra group at the company.

Jan 2009

Facebook, Lakshman, and Malik contribute Cassandra to the Apache Software Foundation, where it becomes an Apache Incubator project.



Mar 2009


Jonathan Ellis

Jonathan Ellis becomes the first new committer to Cassandra in the Apache Incubator. His blogging activity and omnipresence on IRC make him the face of Cassandra to the community.

Apr 2009

Ellis is contacted by John Vrionis of Lightspeed Ventures, who asks him if he's thought about starting a company around Apache Cassandra™. Ellis had indeed begun to have these thoughts, but wasn't quite ready to pull the trigger.


John Vrionis

May 2009


Johan Oskarsson

Twitter engineer Johan Oskarsson announces a conference for scalable databases in San Francisco. Cassandra committer Eric Evans suggests calling this a "NoSQL" conference, and the term sticks.

Nov 2009

Comcast engineers respond to a Cassandra users survey conducted by Ellis. They evaluated Cassandra positively for a data-intensive project, but corporate policy required having a company behind it that they could call for support. Ellis takes this as the sign he was looking for to start his own company based on Apache Cassandra.

Feb-Mar 2010

Jonathan Ellis and Matt Pfeil

Ellis convinces fellow Rackspace engineer Matt Pfeil to quit his job and cofound Riptano with him in Austin to commercialize Apache Cassandra.

Apr 2010

Cassandra graduates from the incubator and becomes a top-level project of the Apache Foundation, with Jonathan Ellis as project chair.

July 2010

Amid competition from multiple interested VC firms, Lightspeed Ventures leads Riptano's Series A funding round.

Jan 2011

Pfeil and Ellis decide to change Riptano's name to DataStax.

May 2011


Jonathan Ellis, Matt Pfeil and Billy Bosworth

Ellis and Pfeil hire college football player-turned-tech entrepreneur Billy Bosworth as CEO to help manage the company's hypergrowth and to get another round of investment, which was led by Crosslink Capital.

All of 2011

The NoSQL data management market explodes, with hundreds of start-ups appearing in the space.

Oct 2011

Apache Cassandra 1.0 is released, and so is Version 1 of DataStax Enterprise, the first integrated data platform with built-in analytics powered by Hadoop running on top of Apache Cassandra.

Aug 2012

University of Toronto researchers studying NoSQL systems conclude that "In terms of scalability, there is a clear winner throughout our experiments. Cassandra achieves the highest throughput for the maximum number of nodes in all experiments" although "this comes at the price of high write and read latencies." Ellis is determined to fix this.

2012-2013


Jonathan Ellis live conference

DataStax enters accelerated development and Cassandra awareness mode, significantly advancing Cassandra's capabilities and spreading the message about the power of Cassandra-based NoSQL data management throughout the tech world. They achieve adoption at many of Silicon Valley's top tech companies, including Apple, Netflix, and Twitter.

2013-2017

DataStax goes into growth mode, taking on a number of big-name customers, including Sony, eBay, Walmart, and FedEx, and going from plucky little startup to the cloud database market leader.

DataStax engineers continue to develop Cassandra, making 85% of the code commits and accelerating Cassandra's evolution through V. 3.11.

May 2018

DataStax releases DataStax Enterprise 6, which is 2 times faster than open source Apache Cassandra while eliminating significant operational complexity.