

Apache Cassandra

Apache Cassandra

Original author(s)	Avinash Lakshman, Prashant Malik
Developer(s)	Apache Software Foundation
Initial release	2008
Stable release	1.1.5 / September 10, 2012
Development status	Active
Written in	Java
Operating system	Cross-platform
Available in	English
Type	key-value store
License	Apache License 2
Website	cassandra.apache.org ^[1]

Apache Cassandra is an open source distributed database management system. It is an Apache Software Foundation top-level project^[2] designed to handle very large amounts of data spread out across many commodity servers while providing a highly available service with no single point of failure. It is a NoSQL solution that was initially developed by Facebook and powered their Inbox Search feature until late 2010.^{[3][4]} Jeff Hammerbacher, who led the Facebook Data team at the time, has described Cassandra as a BigTable data model running on an Amazon Dynamo-like infrastructure.^[5]

Cassandra provides a structured key-value store with tunable consistency.^[6] Keys map to multiple values, which are grouped into column families. The column families are fixed when a Cassandra database is created, but columns can be added to a family at any time. Furthermore, columns are added only to specified keys, so different keys can have different numbers of columns in any given family.

The values from a column family for each key are stored together. This makes Cassandra a hybrid data management system between a column-oriented DBMS and a row-oriented store.^[7] Additional features include: using the BigTable way of modeling, eventual consistency, and the Gossip protocol, a master-master way of serving read and write requests inspired by Amazon's Dynamo.^[8]

History

Apache Cassandra was developed at Facebook to power their Inbox Search feature by Avinash Lakshman (one of the authors of Amazon's Dynamo) and Prashant Malik. It was released as an open source project on Google code in July 2008.^[5] In March 2009, it became an Apache Incubator project.^[9] On February 17, 2010 it graduated to a top-level project.^[2]

Facebook abandoned Cassandra in late 2010 when they built the Facebook Messaging platform on HBase.^[4]

Licensing and support

Apache Cassandra is an Apache Software Foundation project, so it has an Apache License (version 2.0).

There are professional grade support available from a few companies. In the official wiki of Apache Cassandra's project^[10] the following ones, which collaborate with developers to the project, are mentioned

- Acunu ^[11]
- Datastax ^[12]

Main features

Decentralized

Every node in the cluster has the same role. There is **no single point of failure**. Data is distributed across the cluster (so each node contains different data), but there is no master as every node can service any request.

Supports replication and multi data center replication

Replication strategies are configurable.^[13] Cassandra is designed as a distributed system, for deployment of large numbers of nodes across multiple data centers. Key features of Cassandra's distributed architecture are specifically tailored for multiple-data center deployment, for redundancy, for failover and disaster recovery.

Scalability

Read and write throughput both increase linearly as new machines are added, with no downtime or interruption to applications.

Fault-tolerant

Data is automatically replicated to multiple nodes for fault-tolerance. Replication across multiple data centers is supported. Failed nodes can be replaced with no downtime.

Tunable consistency

Writes and reads offer a tunable level of consistency, all the way from "writes never fail" to "block for all replicas to be readable", with the quorum level in the middle.

MapReduce support

Cassandra has Hadoop integration, with MapReduce support. There is support also for Apache Pig ^[14] and Apache Hive ^[15] ^[16].

Query language

CQL (Cassandra Query Language) was introduced, an SQL-like alternative to the traditional RPC interface. Language drivers are available for **Java** (JDBC) and **Python** (DBAPI2).

Data model: keys, row keys, columns and column families

Cassandra is essentially a hybrid between a key-value and a row-oriented (or tabular) database.

Each key in Cassandra corresponds to a value which is an object. Each key has values as columns, and columns are grouped together into sets called column families. Also, each column families can be grouped in super column families.

So each key identifies a row of a variable number of elements. This column families could be considered then as tables. A table in Cassandra is a distributed multi dimensional map indexed by a key.

Furthermore, applications can specify the sort order of columns within a Super Column or Simple Column family.

Clustering

When the cluster for Apache Cassandra is designed, an important point is to select the right partitioner. Two partitioners exist:^[17]

1. RandomPartitioner (RP): This partitioner randomly distributes the key-value pairs over the network, resulting in a good load balancing. Compared to OPP, more nodes have to be accessed to get a number of keys.
2. OrderPreservingPartitioner (OPP): This partitioner distributes the key-value pairs in a natural way so that similar keys are not far away. The advantage is that fewer nodes have to be accessed. The drawback is the uneven distribution of the key-value pairs.

Prominent users

- Talentica Software uses Cassandra as a back-end for Analytics Application with Cassandra cluster of 30 nodes and inserting around 200GB data on daily basis.^[18]
 - AppScale uses Cassandra as a back-end for Google App Engine applications^[19]
 - Cisco's WebEx uses Cassandra to store user feed and activity in near real time.^[20]
 - The CERN ATLAS experiment uses Cassandra to archive its online DAQ system's monitoring information^[21]
 - Clearspring^[22] uses Cassandra "[keep] track of how many times a URL is shared and serves over 200M view requests daily."^[23]
 - Cloudkick uses Cassandra to store the server metrics of their users.^[24]
 - Cloudtalk^[25]'s Platform contains APIs for users to create messaging apps with Cassandra as its data store.^[26]
 - connex.io^[27]'s database of user contacts is stored completely in a Cassandra cluster.^[28]
 - Constant Contact uses Cassandra in their social media marketing application.^[29]
 - Digg, a large social news website, announced on Sep 9th, 2009 that it is rolling out its use of Cassandra^[30] and confirmed this on March 8, 2010.^[31] TechCrunch has since linked Cassandra to Digg v4 reliability criticisms and recent company struggles.^[32] Lead engineers at Digg later rebuked these criticisms as red herring and blamed a lack of load testing.^[33]
 - Digital Reasoning^[34]'s Synthesys application with potential scale to over a 400-node Cassandra database was rolled out in late 2010.^[35]
 - Facebook used Cassandra to power Inbox Search, with over 200 nodes deployed.^[3] This was abandoned in late 2010 when they built Facebook Messaging platform on HBase.^[4]
 - IBM has done research in building a scalable email system based on Cassandra.^[36]
 - Isidorey^[37] is the creator of Cloudsandra: a multi-tenant platform built on Brisk (Hadoop + Cassandra).^[38]
 - Martini Media Network^[39] moved from MySQL to Cassandra.^[40]
 - Mollom^[41] uses Cassandra to track reputations from IP data^[42]
 - Netflix uses Cassandra as their back-end database for their streaming services^{[43][44]}
 - Formspring uses Cassandra to count responses, as well as store Social Graph data (followers, following, blockers, blocking) for 26 Million accounts with 10 million responses a day^[45]
 - Mahalo.com uses Cassandra to record user activity logs and topics for their Q&A website^{[46][47]}
 - Ooyala Built a scalable, flexible, real-time analytics engine using Cassandra^[48]
 - At Openwave, Cassandra acts as a distributed database and serves as a distributed storage mechanism for Openwave's next generation messaging platform^[49]
 - OpenX is running over 130 nodes on Cassandra for their OpenX Enterprise product to store and replicate advertisements and targeting data for ad delivery.^[50]
 - Outbrain^[51] uses Cassandra as a semi-persistent cache of recommendations.^[52]
 - Plaxo has "reviewed 3 billion contacts in [their] database, compared them with publicly available data sources, and identified approximately 600 million unique people with contact info."^[53]
 - PostRank uses Cassandra as their backend database^[54]
-

- Rackspace is known to use Cassandra internally.^[55]
- Reddit switched to Cassandra from memcacheDB on March 12, 2010^[56] and experienced some problems with overload handling in Cassandra in May.^[57]
- RockYou uses Cassandra to record every single click for 50 million Monthly Active Users in real-time for their online games^[58]
- ShopSavvy uses Cassandra as their storage technology for their barcode scanning application^[59]
- SimpleGeo^[60] built a scalable geospatial database on top of Apache Cassandra^[61]
- SoundCloud uses Cassandra to store user account information^[62]
- Twitter announced it is planning to use Cassandra because it can be run on large server clusters and is capable of taking in very large amounts of data at a time.^{[63][64]} Twitter continues to use it but not for Tweets themselves.^[65]
- Urban Airship uses Cassandra with the mobile service hosting for over 160 million application installs across 80 million unique devices^[66]
- Uti Labs^[67] uses Cassandra for its Low Voltage Smart System for fine grained event / reading recording and analytics^[68]
- Walmart Labs^[69] (previously Kosmix) uses Cassandra with SSD^[70]
- Yakaz uses Cassandra on a five-node cluster to store millions of images as well as its social data.^[71]
- Viocom Hosting uses Cassandra for its email and smtp gateway called (SendCluster) for large volume senders. Storing from 40 million emails a day for email marketing. This is a private platform built with cassandra for large Email Senders^[72]

Tools for Cassandra

Cassandra has built in tools for accessing Cassandra from the direct download such `cassandra-cli` and `node-tool`.

There are third party tools available, as the following^[73]:

Data browsers

- `chiton`^[74], a GTK data browser.
- `cassandra-gui`^[75], a Swing data browser.
- `Toad for Cloud Databases`^[76], an Eclipse plug-in data browser

Administration tools

- `OpsCenter`^[77], OpsCenter is a tool for management and monitoring of a Cassandra cluster. The Community Edition of OpsCenter is free for anyone to download and use. There is also an Enterprise Edition of OpsCenter that includes additional features.
- `Cassandra Cluster Admin`^[78], Cassandra Cluster Admin is a GUI tool to help people administrate their Apache Cassandra cluster, similar to `PHPMysqlAdmin` for MySQL administration.

Client interfaces and language Support

Cassandra has a lot of high-level client libraries for Python, Java, .Net, Ruby, PHP, Perl, C++, etc.^[79]

For a detailed list of client software go to "Client Options" article^[80] on Cassandra Wiki

Integration with other tools

There are other tools worth mentioning like **Solandra**,^[81] a Cassandra backend for Apache Solr^[82], a web application built around Lucene, for full text indexing and search.

For monitoring purposes Cassandra is well integrated with Ganglia^[83] and there are plugins for other monitoring system as, by example, Nagios.

References

- [1] <http://cassandra.apache.org/>
- [2] "Cassandra is an Apache top level project" (<http://www.mail-archive.com/cassandra-dev@incubator.apache.org/msg01518.html>). Mail-archive.com. 2010-02-18. Archived (<http://web.archive.org/web/20100328090322/http://www.mail-archive.com/cassandra-dev@incubator.apache.org/msg01518.html>) from the original on 28 March 2010. . Retrieved 2010-03-29.
- [3] "Niet compatibele browser" (http://www.facebook.com/note.php?note_id=24413138919&id=9445547199&index=9). Facebook. . Retrieved 2010-03-29.
- [4] Kannan Muthukkaruppan. "The Underlying Technology of Messages" (<http://www.facebook.com/notes/facebook-engineering/the-underlying-technology-of-messages/454991608919>). .
- [5] James Hamilton (July 12, 2008). "Facebook Releases Cassandra as Open Source" (<http://perspectives.mvdirona.com/2008/07/12/FacebookReleasesCassandraAsOpenSource.aspx>). . Retrieved 2009-06-04.
- [6] <http://www.cs.cornell.edu/projects/ladis2009/papers/lakshman-ladis2009.pdf>
- [7] Vishal Shinde (2011-01-19). "Apache Cassandra" (<http://www.linuxquestions.in/other-databases/apache-cassandra/?action=printpage>). <http://www.linuxquestions.in/>: Linux Questions, Help, Howtos and Tutorials. . Retrieved 2010-03-09. "The values from a column family for each key are stored together, making Cassandra a hybrid between a column-oriented DBMS and a row-oriented store"
- [8] Olivier Mallasi (2010-06-09). "Let's play with Cassandra... (Part 1/3)" (<http://blog.octo.com/en/nosql-lets-play-with-cassandra-part-13/>). <http://blog.octo.com/>: OCTO Talks. . Retrieved 2010-03-22. "Hybrid firstly because Cassandra uses a column-oriented way of modeling data (inspired by the BigTable) and permit to use Hadoop Map/Reduce jobs and secondly because it uses patterns inspired by Dynamo like Eventually Consistent, Gossip protocols, a master-master way of serving both read and write requests..."
- [9] "Is this the new hotness now?" (<http://www.mail-archive.com/cassandra-dev@incubator.apache.org/msg00004.html>). Mail-archive.com. 2009-03-02. Archived (<http://web.archive.org/web/20100425071855/http://www.mail-archive.com/cassandra-dev@incubator.apache.org/msg00004.html>) from the original on 25 April 2010. . Retrieved 2010-03-29.
- [10] "Third Party Support" (<http://wiki.apache.org/cassandra/ThirdPartySupport>) article on Apache Cassandra's wiki
- [11] <http://www.acunu.com/>
- [12] <http://datastax.com/>
- [13] "Deploying Cassandra across Multiple Data Centers" article on Datastax Cassandra Developer Center (<http://www.datastax.com/dev/blog/deploying-cassandra-across-multiple-data-centers>)
- [14] <http://pig.apache.org/>
- [15] <http://hive.apache.org/>
- [16] "Hadoop Support" (<http://wiki.apache.org/cassandra/HadoopSupport>) article on Cassandra's wiki
- [17] Dominic Williams. "Cassandra: RandomPartitioner vs OrderPreservingPartitioner" (<http://ria101.wordpress.com/2010/02/22/cassandra-randompartitioner-vs-orderpreservingpartitioner/>). <http://wordpress.com/>: WordPress.com. . Retrieved 2011-03-23. "When building a Cassandra cluster, the "key" question (sorry, that's weak) is whether to use the RandomPartitioner (RP), or the OrderPreservingPartitioner (OPP). These control how your data is distributed over your nodes. Once you have chosen your partitioner, you cannot change without wiping your data, so think carefully! The problem with OPP: If the distribution of keys used by individual column families is different, their sets of keys will not fall evenly across the ranges assigned to nodes. Thus nodes will end up storing preponderances of keys (and the associated data) corresponding to one column family or another. If as is likely column families store differing quantities of data with their keys, or store data accessed according to differing usage patterns, then some nodes will end up with disproportionately more data than others, or serving more "hot" data than others."
- [18] cite weblurl=<http://www.talentica.com>
- [19] cite weblurl=<http://appscales.cs.ucsb.edu/datastores.html#cassandra>
- [20] "Re: Cassandra users survey" (<http://www.mail-archive.com/cassandra-dev@incubator.apache.org/msg01163.html>). Mail-archive.com. 2009-11-21. Archived (<http://web.archive.org/web/20100417083733/http://www.mail-archive.com/cassandra-dev@incubator.apache.org/msg01163.html>) from the original on 17 April 2010. . Retrieved 2010-03-29.
- [21] "A Persistent Back-End for the ATLAS Online Information Service (P-BEAST)" (<https://cdsweb.cern.ch/record/1432912>). .
- [22] <http://clearspring.com/>
- [23] Matt Abrams (2011-05-05). "Clearsprings Big Data Architecture Part 1" (<http://clearspring.com/blog/2011/05/05/clearsprings-big-data-architecture-part-1>). .
- [24] https://www.cloudkick.com/blog/2010/mar/02/4_months_with_cassandra/
- [25] <http://cloudtalk.com/>
- [26] <http://cloudtalk.com/the-cloudtalk-platform.html>
- [27] <http://connex.io>
- [28] <http://blog.connex.io/why-we-replaced-syncml-with-our-own-contact-s>
- [29] Klint Finley (2011-02-18). "This Week in Consolidation: HP Buys Vertica, Constant Contact Buys Bantam Live and More" (<http://www.readwriteweb.com/enterprise/2011/02/this-week-in-consolidation-hp.php>). Read Write Enterprise. .
- [30] Ian Eure. "Looking to the future with Cassandra" (<http://blog.digg.com/?p=966>). .
- [31] John Quinn. "Saying Yes to NoSQL; Going Steady with Cassandra" (<http://about.digg.com/node/564>). .

- [32] Erick Schonfeld. "As Digg Struggles, VP Of Engineering Is Shown The Door" (<http://techcrunch.com/2010/09/07/digg-struggles-vp-engineering-door/>). .
- [33] "Is Cassandra to Blame for Digg v4's Failures?" (<http://www.quora.com/Is-Cassandra-to-blame-for-Digg-v4s-technical-failures/>). .
- [34] <http://www.digitalreasoning.com/>
- [35] <http://www.datastax.com/wp-content/uploads/2011/03/CS-DigitalReasoning.pdf>
- [36] "Powered by Google Docs" (<http://docs.google.com/viewer?url=http://ewh.ieee.org/r6/scv/computer//nfc/2009/IBM-Jun-Rao.pdf>). Docs.google.com. . Retrieved 2010-03-29.
- [37] <http://isidorey.com>
- [38] <http://www.cloudsandra.com/>
- [39] <http://www.martinimedianetwork.com/>
- [40] Manicka Babu (2011-05-22). "Cassandra Part 1" (<http://manickababu.blogspot.com/2011/05/cassandra-part-1.html>). .
- [41] <http://mollom.com/>
- [42] Todd Hoff (2011-02-08). "Mollom Architecture - Killing Over 373 Million Spams at 100 Requests per Second" (<http://highscalability.com/blog/2011/2/8/mollom-architecture-killing-over-373-million-spams-at-100-re.html>). High Scalability. .
- [43] cite weblurl=<http://www.slideshare.net/adrianco/migrating-netflix-from-oracle-to-global-cassandra>
- [44] Yury Izrailevsky (2011-01-28). "NoSQL at Netflix" (<http://techblog.netflix.com/2011/01/nosql-at-netflix.html>). .
- [45] Martin Cozzi (2011-08-31). "Cassandra at Formspring" (<http://www.slideshare.net/martincozzi/cassandra-formspring>). .
- [46] "" (<http://www.datastax.com/wp-content/uploads/2011/06/DataStax-CaseStudy-Mahalo.pdf>). .
- [47] <http://blip.tv/datastax/cassandra-at-mahalo-com-4030941>
- [48] <http://www.datastax.com/wp-content/uploads/2011/04/WP-Ooyala.pdf>
- [49] <http://www.datastax.com/wp-content/uploads/2011/05/DataStax-CaseStudy-Openwave.pdf>
- [50] <http://openx.com/publisher/technology>
- [51] <http://www.outbrain.com>
- [52] Nathan Milford. "Cassandra for Sysadmins" (<http://techblog.outbrain.com/2011/08/slides-cassandra-for-sysadmins/>). .
- [53] Preston Smalley (2011-03-20). "An important milestone - and it's only the beginning!" (<http://blog.plaxo.com/2011/03/an-important-milestone-and-its-only-the-beginning/>). .
- [54] Ilya Grigorik (2011-03-29). "Webpulp TV: Scaling PostRank with Ilya Grigorik" (<http://blog.postrank.com/2011/03/webpulp-tv-scaling-postrank-with-ilya-grigorik/>). .
- [55] "Hadoop and Cassandra (at Rackspace)" (<http://www.slideshare.net/stuhood/hadoop-and-cassandra-at-rackspace>). Stu Hood. 2010-04-23. . Retrieved 2011-09-01.
- [56] Posted by david [ketrnalis] (2010-03-12). "what's new on reddit: She who entangles men" (<http://blog.reddit.com/2010/03/she-who-entangles-men.html>). blog.reddit. Archived (<http://web.archive.org/web/20100325115755/http://blog.reddit.com/2010/03/she-who-entangles-men.html>) from the original on 25 March 2010. . Retrieved 2010-03-29.
- [57] Posted by the reddit admins at (2010-05-11). "blog.reddit -- what's new on reddit: reddit's May 2010 "State of the Servers" report" (<http://blog.reddit.com/2010/05/reddits-may-2010-state-of-servers.html>). blog.reddit. Archived (<http://web.archive.org/web/20100514085008/http://blog.reddit.com/2010/05/reddits-may-2010-state-of-servers.html>) from the original on 14 May 2010. . Retrieved 2010-05-16.
- [58] Dathan Vance Pattishall (2011-03-23). "Cassandra is my NoSQL Solution but" (<http://mysqldba.blogspot.com/2010/03/cassandra-is-my-nosql-solution-but.html>). .
- [59] Alexander Muse (2011-07-18). "Shopsavvy leverages Hadoop and Cassandra" (<http://shopsavvy.mobi/2011/07/18/shopsavvy-leverages-hadoop-and-cassandra/>). .
- [60] <http://www.simplegeo.com>
- [61] Klint Finley. "How SimpleGeo Built a Scalable Geospatial Database with Apache Cassandra" (<http://www.readwriteweb.com/cloud/2011/02/video-simplegeo-cassandra.php>). Read Write Cloud. .
- [62] "Cassandra at SoundCloud" ([http://berlinbuzzwords.de/sites/berlinbuzzwords.de/files/cassandra workshop berlin buzzword 2011-Soundcloud.pdf](http://berlinbuzzwords.de/sites/berlinbuzzwords.de/files/cassandra%20workshop%20berlin%20buzzword%202011-Soundcloud.pdf)). .
- [63] Popescu, Alex. "Cassandra @ Twitter: An Interview with Ryan King" (<http://nosql.mypopescu.com/post/407159447/cassandra-twitter-an-interview-with-ryan-king>). myNoSQL. Archived (<http://web.archive.org/web/20100301151656/http://nosql.mypopescu.com/post/407159447/cassandra-twitter-an-interview-with-ryan-king>) from the original on 1 March 2010. . Retrieved 2010-03-29.
- [64] Babcock, Charles. "Twitter Drops MySQL For Cassandra - Cloud databases" (http://www.informationweek.com/news/software/open_source/showArticle.jhtml?articleID=223100894&pgno=1&queryText=&isPrev=). InformationWeek. Archived (http://web.archive.org/web/20100402075726/http://www.informationweek.com/news/software/open_source/showArticle.jhtml?articleID=223100894&pgno=1&queryText=&isPrev=) from the original on 2 April 2010. . Retrieved 2010-03-29.
- [65] "Cassandra at Twitter Today" (<http://engineering.twitter.com/2010/07/cassandra-at-twitter-today.html>). .
- [66] Erik Onnen. "From 100s to 100s of Millions" (<http://www.slideshare.net/eonnen/from-100s-to-100s-of-millions>). .
- [67] <http://www.utillabs.com/>
- [68] Hartmut Bohmer. "Low Volt Smart System" (<http://www.utillabs.com>). .
- [69] <http://www.walmartlabs.com>
- [70] Karl Mueller. "Cassandra on SSD" (<http://blog.kosmix.com/2011/01/21/cassandra-on-ssd/>). .

- [71] "Yakaz Technologies" (<http://www.yakaz.com/about/technologies.php>). .
- [72] "Viocom" (<https://www.viocom.co.uk/servers/clusters-sans-nosql-servers>). .
- [73] FAQ (<http://wiki.apache.org/cassandra/FAQ#gui>) on Cassandra's wiki
- [74] <http://github.com/driftx/chiton>
- [75] <http://code.google.com/p/cassandra-gui>
- [76] <http://www.quest.com/toad-for-cloud-databases/>
- [77] <http://www.datastax.com/products/opscenter>
- [78] <https://github.com/sebgiroux/Cassandra-Cluster-Admin>
- [79] "Client Options" article (<http://wiki.apache.org/cassandra/ClientOptions>) on Cassandra Wiki
- [80] <http://wiki.apache.org/cassandra/ClientOptions>
- [81] Solandra source at Github (<https://github.com/tjake/Solandra>)
- [82] <http://lucene.apache.org/solr/>
- [83] Cassandra - A Decentralized Structured Storage System (<http://www.cs.cornell.edu/projects/ladis2009/papers/lakshman-ladis2009.pdf>), a 2009 paper presenting Cassandra by their creators Avinash Lakshman and Prashant Malik

Bibliography

- Hewitt, Eben (December 15, 2010). *Cassandra: The Definitive Guide* (<http://oreilly.com/catalog/0636920010852>) (1st ed.). O'Reilly Media. pp. 300. ISBN 978-1-4493-9041-9.
- Capriolo, Edward (July 15, 2011). *Cassandra High Performance Cookbook* (<http://www.packtpub.com/cassandra-apache-high-performance-cookbook/book>) (1st ed.). Packt Publishing. pp. 324. ISBN 1-84951-512-3.

External links

- Avinash Lakshman (25 August 2008). "Cassandra - A structured storage system on a P2P Network" (http://www.facebook.com/note.php?note_id=24413138919&id=9445547199&index=9). Engineering @ Facebook's Notes. Retrieved 2009-06-04.
- Project Website (<http://cassandra.apache.org/>)
- Project Wiki (<http://wiki.apache.org/cassandra/>)
- Adopting Apache Cassandra (<http://www.infoq.com/presentations/Adopting-Apache-Cassandra>) presented by Eben Hewitt on December 1, 2010
- LADIS 2009 WhitePaper by the original contributors Avinash Lakshman & Prashant Malik (<http://www.cs.cornell.edu/projects/ladis2009/papers/lakshman-ladis2009.pdf>)
- Cassandra Articles on NoSQLDatabases.com (<http://www.nosqldatabases.com/main/tag/cassandra>)
- Cassandra News and Articles on myNoSQL (<http://nosql.mypopescu.com/tagged/cassandra>)
- Cassandra @ Twitter: an Interview with Ryan King (<http://nosql.mypopescu.com/post/407159447/cassandra-twitter-an-interview-with-ryan-king>)
- Presentation on RDBMS vs. Dynamo, BigTable, and Cassandra (<http://www.slideshare.net/jbellis/what-every-developer-should-know-about-database-scalability>)
- RPM build for the apache cassandra project (<http://code.google.com/p/cassandra-rpm/>)

Article Sources and Contributors

Apache Cassandra *Source:* <http://en.wikipedia.org/w/index.php?oldid=511879100> *Contributors:* Acdx, Al3xpopescu, Alan Liefing, Andrewllavore, Andy Dingley, Anthony Appleyard, ArlebargleIV, Arthurjulian, Bearcat, Biktora, Billmantisco, Cander0000, Chris Chittleborough, ChrisGualtieri, Ciges, Cleduc, Clydewylam, Cybercobra, Dancraggs, DeansFA, Deineka, Driftx, Drmies, Dstainer, Ehn, Enchanter, Euphoria, FalconL, FrankTobia, Frap, Frecklefoot, Fyedernoggersnodden, Grafen, GreyTeardrop, Grossenhayn, Gstein, Hashar, JLaTondre, Jadave234, Jbryanscott, Jim1138, Jmhodges, Jncraton, Jweiss11, Kinglarvae, Kolyma, Krotty, Ksato9700, Kylemurph, Materialscientist, Maximgr, McSly, Midinasturazz, Mmozum, Msiebuhr, Neilc, Nemnkim, Nmiford, OrangeDog, PatrickFisher, Pmiossec, Rich Farmbrough, Ronz, Runtime, Sae1962, SamJohnston, Samarthgahire, Santiagobasulto, Slebresne, Swaroopch, Tas50, Tdmackey, TheJJJunk, Thumperward, Timendum, Timoeuy, Tommymorgan, Vanger13, Vegaswikian, Viocomnetworks, Wainstead, YPavan, Yadavjpr, YogiWanKenobi, 128 anonymous edits

License

Creative Commons Attribution-Share Alike 3.0 Unported
[//creativecommons.org/licenses/by-sa/3.0/](http://creativecommons.org/licenses/by-sa/3.0/)

Contribute to apache/cassandra development by creating an account on GitHub. Patch by Jeremy Hanna, reviewed by David Capwell and brandonwilliams for CASSANDRA-15931. Git stats. 25,381 commits. Apache Cassandra is a NoSQL multi-master database with linear scalability and no single point of failure. Install the cassandraAUR package. The package logs to /var/log/cassandra/system.log by default. To instead log to journald you will need to copy the systemd unit to /etc/systemd/system/ so the change persists. \$ cp /usr/lib/systemd/system/cassandra.service /etc/systemd/system/. Edit the unit. \$ vim /etc/systemd/system/cassandra.service. In this Apache Cassandra Tutorial, we will learn about what is Cassandra, what is NoSQL database and also a short introduction of Relational/SQL vs NoSQL database. We will also learn about different Apache Cassandra features, Cassandra History, Architecture of Cassandra. In addition, we are going to study Apache Cassandra Applications and Cassandra Data Model. As the technology advances, the data has also expanded and grown exponentially. So, there is a need for a new database.