Apache Software Foundation ▾  /  Apache Cassandra

# What is Cassandra?

The Apache Cassandra database is the right choice when you need scalability and high availability without compromising performance. Linear scalability (http://techblog.netflix.com/2011/11/benchmarking-cassandra-scalability-on.html) and proven fault-tolerance on commodity hardware or cloud infrastructure make it the perfect platform for mission-critical data.Cassandra's support for replicating across multiple datacenters is best-in-class, providing lower latency for your users and the peace of mind of knowing that you can survive regional outages.

## PROVEN

Cassandra is in use at Constant Contact (http://www.slideshare.net/daveconnors/cassandra-puppet-scaling-data-at-15-per-month), CERN (http://planetcassandra.org/blog/post/cassandra-at-cern-large-hadron-collider/), Comcast (http://www.slideshare.net/planetcassandra/nyc-tech-day-using-cassandra-for-dvr-scheduling-at-comcast), eBay (http://www.slideshare.net/jaykumarpatel/cassandra-at-ebay-13920376), GitHub (http://planetcassandra.org/blog/post/analytics-at-github-with-apache-cassandra/), GoDaddy (http://planetcassandra.org/blog/post/godaddy-worlds-largest-domain-name-registrar-and-web-host-provider-utilizes-cassandra-for-replication-and-scalability/), Hulu (http://planetcassandra.org/blog/post/cassandra-used-to-build-scalable-and-highly-available-systems-at-hulu-streaming-content-to-over-5-million-subscribers/), Instagram (http://planetcassandra.org/blog/post/instagram-making-the-switch-to-cassandra-from-redis-75-instasavings/), Intuit (http://www.slideshare.net/planetcassandra/3-mohit-anchlia), Netflix (http://www.slideshare.net/adrianco/migrating-netflix-from-oracle-to-global-cassandra), Reddit (http://planetcassandra.org/blog/post/reddit-upvotes-apache-cassandras-horizontal-scaling-managing-17000000-votes-daily/), The Weather Channel (http://planetcassandra.org/blog/post/make-it-rain-apache-cassandra-at-the-weather-channel-for-severe-weather-alerts/), and over 1500 more companies (http://planetcassandra.org/companies/) that have large, active data sets.

## 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.

## PERFORMANT

Cassandra consistently (http://vldb.org/pvldb/vol5/p1724_tilmannrabl_vldb2012.pdf) outperforms (http://www.datastax.com/resources/whitepapers/benchmarking-top-nosql-databases) popular NoSQL alternatives in benchmarks and real applications (http://blog.markedup.com/2013/02/cassandra-hive-and-hadoop-how-we-picked-our-analytics-stack/), primarily because of fundamental architectural choices (http://www.datastax.com/dev/blog/2012-in-review-performance).

## DECENTRALIZED

There are no single points of failure. There are no network bottlenecks. Every node in the cluster is identical.

## SCALABLE

Some of the largest production deployments include Apple's, with over 75,000 nodes storing over 10 PB of data, Netflix (2,500 nodes, 420 TB, over 1 trillion requests per day), Chinese search engine Easou (270 nodes, 300 TB, over 800 million requests per day), and eBay (over 100 nodes, 250 TB).

## DURABLE

Cassandra is suitable for applications that can't afford to lose data (http://wiki.apache.org/cassandra/Durability), even when an entire data center goes down.

## YOU'RE IN CONTROL

Choose between synchronous or asynchronous replication for each update. Highly available asynchronous operations are optimized with features like Hinted Handoff (http://wiki.apache.org/cassandra/HintedHandoff) and Read Repair (http://wiki.apache.org/cassandra/ReadRepair).

## ELASTIC

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

## PROFESSIONALLY SUPPORTED

Cassandra support contracts and services are available from third parties (http://wiki.apache.org/cassandra/ThirdPartySupport).