



Redis Accelerates Spark by Over 100 Times



Mountain View, February 2, 2016—[Redis Labs](#), the home of Redis, today announced its integration with Spark SQL and the release of the [Spark-Redis connector](#) package. Redis Labs' [benchmarks](#) using time-series data show that running Spark on Redis as a data store results in 135 times faster processing compared to Spark using HDFS and 45 times faster processing compared to Spark using Tachyon as an off-heap data store or Spark storing the data on-heap. The Spark-Redis connector package is open source and provides a library for writing to and reading from a Redis cluster with access to all of Redis' data structures – String, Hash, List, Set, Sorted Set, bitmaps, hyperloglogs – from Spark as RDDs. In addition, the package also ensures close cluster alignment between Spark and Redis clusters, reducing network overhead and ensuring optimal processing times. The main advantages of using Redis with Spark include:

- Acceleration of Spark performance by over 100 times, in use cases such as [spark-timeseries](#)
- Redis data structures allow elements of data to be accessed individually and rapidly, minimizing serialization/deserialization overhead and avoiding transfer of large chunks of data.

"Big data is coming of age and customers are demanding that big data insights are extracted in real-time," said Yiftach Shoolman, co-founder and CTO of Redis Labs. "This is where Redis Labs fills the gap by delivering both the right performance and optimized distributed memory infrastructure to accelerate Spark. Our goal is to make Redis the de-facto data store for any Spark deployment."

The Spark-Redis solution enables:

- Redis data structures exposed via Spark RDD and DataSet API
- Spark SQL support (via DataFrame and DataSource)
- Use of Redis Cluster as a distributed memory infrastructure

Additional planned enhancements to the solution include support for other popular use cases such as graph computation and machine learning.

Ready to use a simple, easy to deploy database to improve your team's performance?

1



"Apache Spark is becoming a default in-memory engine for high-performance data integration and analytics," said Matt Aslett, research director, data platforms and analytics at 451 Research. "The combination of Redis and Spark should enable high-performance, real-time analytics with extremely large and variable datasets."

To start using the Spark-Redis connector, visit: <http://spark-packages.org/package/RedisLabs/spark-redis>.

By continuing to use this site, you consent to our updated privacy agreement as [described here](#). You can change your cookie settings at any time as [described here](#) but parts of our site will not function correctly without them.

Continue

ABOUT REDIS LABS

Data is the lifeline of every business, and [Redis Labs](#) helps organizations reimagine how quickly they can process, analyze, make predictions with, and take action on the data they generate. As the home of [Redis](#), the most popular open source database, we provide a competitive edge to global businesses with [Redis Enterprise](#), which delivers superior performance, unmatched reliability, and the best total cost of ownership. Redis Enterprise allows teams to build performance, scalability, security, and growth into their applications. Designed for the cloud-native world, Redis Enterprise uniquely unifies data across hybrid, multi-cloud, and global applications, to maximize your business potential.

Learn how Redis Labs can give you this edge at redislabs.com.

COMPANY

About Us
Team
Careers
Newsroom
Events
Contact Us
Support

COMMUNITY

Redis Stars
Events
Why Redis
Tech Blog
Projects
Modules Hub
Redis Watch

PARTNERS

Partner Center
Find a Partner

POPULAR TOPICS

Pricing
Why Redis

TRY REDIS ENTERPRISE CLOUD FREE

Redis Enterprise Cloud provides

Ready to use a simple, easy to deploy database to improve your team's performance?

Try Now



logo are registered trademarks of Redis Labs Ltd.

Ready to use a simple, easy to deploy database to improve your team's performance?

1