



Spark SQL vs Presto

In this article we've tried to lay out the comparisons of Spark SQL vs Presto. When it comes to checking out Spark Presto, there are some differences to be aware of:

Commonality:

- They are both open source, “big data” software frameworks
- They are distributed, parallel, and in-memory
- BI tools connect to them using JDBC/ODBC
- Both have been tested and deployed at petabyte-scale companies
- They can be run on-prem or in the cloud. They can be containerized

Differences:

- Presto is an ANSI SQL:2003 query engine for accessing and unifying data from many different data sources. It's deployed as a middle-layer for federation.
- Spark is a general-purpose cluster-computing framework. Core Spark does not support SQL – for SQL support you install the Spark SQL module which adds structured data processing capabilities. Spark SQL is also ANSI SQL:2003 compliant (since Spark 2.0).
- Presto is more commonly used to support interactive SQL queries. Queries are usually analytical but can perform SQL-based ETL.
- Spark is more general in its applications, often used for data transformation and Machine Learning workloads.
- Presto supports querying data in object stores like S3 by default, and has many connectors available. It also works really well with Parquet and Orc format data.
- Spark must use Hadoop file APIs to access S3 (or pay for Databricks features). Spark has limited connectors for data sources.

Many users are today are learning about Presto Spark. This lays out many of the differences on Presto vs Spark SQL and how Spark and Presto can be compared.

If you want to deploy a Presto cluster on your own, we recommend checking out how Ahana manages Presto in the cloud. We put together this [free tutorial that](https://ahana.io/learn/comparisons/spark-sql-vs-presto/)

[shows you how to create a Presto cluster.](#)

Want more Presto tips & tricks? Sign up for our [Presto community newsletter](#).

Presto community newsletter



**Ahana is excited to sponsor
this year's PrestoCon Day!**



**WHITE PAPER:
Unlocking the Business Value of
the Data Lake**



Open Source

[PrestoDB](#)[Presto Foundation](#)[PrestoDB Docs](#)[Presto Users](#)[Query Analyzer for Presto](#)

Company

[About Us](#)[Pricing](#)[Ahana Support](#)[Careers](#)[News & Press](#)[Contact Us](#)

Resources

[Blog](#)[Events & Webinars](#)[Videos & Presentations](#)[Presto Use Cases](#)[Learning Center](#)[AWS Athena vs. PrestoDB](#)

Follow Ahana



©2022 Ahana Cloud, Inc., All rights reserved. The Ahana logo is an unregistered trademark of Ahana Cloud, Inc. Presto and the Presto logo are registered trademarks of **The Linux Foundation & Presto Foundation**. The Linux Foundation has registered trademarks and uses **trademarks**. Linux is a registered trademark of Linus Torvalds. All other trademarks are property of their respective owners. [Privacy Policy](#) · [Terms of Use](#)