

Action Vector in Hadoop

Scale the fastest SQL analytics engine on Apache Hadoop for industrial-strength deployment

Key Benefits

Analyze Hadoop Data Faster

Give SQL users turbocharged access to Hadoop data and analytics

Unleash SQL Users on Hadoop

Remove all barriers for business access to big data analytics

Improve Operational Intelligence

Enable more precise analytical predictions and decisions by enriching business analytics with Hadoop data

The World's Fastest Native RDBMS in Hadoop

Action has set the pace in the industry for fast analytics ready to provide better insights from your hybrid data environment. Action Vector in Hadoop (VectorH) is the world's fastest, most production-ready SQL database engine built natively to run on Hadoop clusters and support enterprise workloads. Running on commodity hardware and Linux open source software, VectorH makes Hadoop data easily consumable and accessible for business users, programmers, and applications using standard SQL. VectorH now supports tighter integration with Apache Spark to increase the number of supported data formats and extend the range of functionality to machine learning, graph analytics, and streaming data.

Action VectorH helps enterprises get their Hadoop data projects into production, not only with high-performance interactive SQL and analytics processing, but also with an extensive security and administrative framework to meet operational requirements. You can drill down on billions of rows of Hadoop data for deeper insights in just seconds, enabling existing applications and tools to scale to Hadoop big data volumes. You can also protect data better with row- and role-based security, handle updates from operational data sources with full ACID compliance and no performance penalty, and reconfigure cluster assets to maximize productivity.

Embracing the Spark Ecosystem

Action VectorH provides transparent access to functionality and data in the popular Apache Spark ecosystem. The Spark-Vector Connector provides the ability to join Vector data with native Hadoop data files including ORC, Parquet, and CSV. Conversely, VectorH can now directly query tables in an external format, and store results in those external table formats as well.

Beyond extending support for additional data format, VectorH allows Spark SQL to query VectorH tables and get the result in a dataframe. Whether using Spark SQL or direct SQL access to VectorH, tools and applications run faster using the VectorH query engine to process their analytic workloads. Imagine applying VectorH performance to machine learning, graph analytics, and real-time streaming.

Faster Query Processing

The core technology behind the high processing speeds of Action VectorH is vectorized processing, powered by the hardened, patented X100 query engine. It dramatically reduces the interpretation overhead typically found in other relational database systems. Additionally, it exploits performance-critical features of modern CPUs like super-scalar execution and Single Instruction, Multiple Data (SIMD) instructions. Finally, thanks to its focus on leveraging the CPU cache, main-memory traffic is reduced, which is especially important in modern multi-core systems.

SIMD applies a single operation to a vector of multiple data values in a single instruction cycle. The X100 query engine takes advantage of SIMD instructions by processing vectors of data each clock cycle using the vector instruction set available in industry-standard CPUs. The average computation for a single data value can take significantly less time than a single CPU cycle. As a result, Actian VectorH can process Hadoop data much faster than any other relational database. Much faster data processing performance accelerates time to actionable decisions.

Up to 900 Times Faster than the Competition

Actian VectorH, the world's fastest database, has been proven to run up to 900 times faster than other SQL tools on Hadoop which are based on older database technology or immature open source query engines. Give your SQL users and applications the performance and enterprise-grade SQL access they need to do their jobs.

Full Function SQL Query Engine Integrated in Hadoop

Don't be fooled by imitators who claim to offer native SQL in Hadoop support—most have immature low quality SQL support or require you to move your data out of Hadoop. Only Actian VectorH combines the most modern, scalable, high-performing database technology with the power of Hadoop, thereby enabling users to directly query data stored in HDFS. Actian VectorH supports the major distributions and versions of Hadoop, including Apache, Cloudera, and Hortonworks.

Comprehensive Enterprise-Strength Capabilities

- Accessible – standard ANSI SQL:2003 to support standard BI tools; plus advanced analytics including CUBE, ROLLUP, GROUPING SETS, and WINDOWING functions
- Secure – native DBMS security, including user and role-level authentication, data protection, and encryption of data at rest and in motion
- Reliable – fully ACID-compliant with multi-version read consistency for transactional integrity, plus system-wide failover protection
- Open – collaborative architecture for querying native Hadoop file formats without ingestion
- Concurrency – highest number of simultaneous users and tasks possible without long wait times
- Optimized – SQL, planning, and execution optimization with built-in machine learning to improve over time
- Parallelized – optimal use of every node, CPU, memory, and cache – tested to handle the largest Hadoop clusters
- Updates – fine-grained updates in Hadoop (patent pending) allow real-time insert, update and delete of individual records without impacting the performance of running queries
- Localized – data stored within HDFS for high-speed data loading and query execution
- Compression – data stored in blocks on HDFS using automatically selected smart compression schemes and tuning parameters
- Columnar – data stored in a true columnar format for faster scan rates and efficient analytical processing
- Manageable – query workloads managed automatically in Hadoop via YARN



www.actian.com | Toll Free +1 888 446 4737 | Tel +1 650 587 5500
2300 Geng Rd., Suite 150, Palo Alto, CA 94303



© 2017 Actian Corporation. Actian is a trademark of Actian Corporation and its subsidiaries. All other trademarks, trade names, service marks, and logos referenced herein belong to their respective companies. (DS14-0917)