

Action Vector in Hadoop

Turn Hadoop into a High-Performance, Industrial-Strength SQL Engine

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 is revolutionizing big data analytics with Action Vector in Hadoop, the world's fastest, most enterprise-grade massively parallel processing (MPP) SQL query engine built to run natively in Hadoop. Action Vector combines the most modern, scalable, high-performing database technology with the power of Hadoop, enabling users to directly query data stored in HDFS. Extreme performance is achieved on commodity hardware, with no database tuning. With Action Vector, Hadoop data is easily consumable and accessible by business users, programmers, and even applications that need fast access using standard SQL.

Action Vector brings Hadoop out of the batch-processing environment and closer to your business, allowing for high-performance interactive SQL and analytics processing. When you can drill down on billions of rows of Hadoop data for deeper insights in just seconds, you open up a new realm of possibilities for your SQL-enabled applications.

Vectorized Query Processing

The core technology behind the high processing speeds of Action Vector 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 storing data in the CPU cache, main-memory traffic is reduced, which is especially important in modern multi-core systems.

Single Instruction, Multiple Data

SIMD enables a single operation to be applied on a set of data at once. Action Vector takes advantage of SIMD instructions by processing vectors of data through the Streaming SIMD Extensions instruction set. Because typical data analysis queries process large volumes of data, the use of SIMD results in the average computation against a single data value taking significantly less time than a single CPU cycle.

Ultrafast Performance

Action Vector was written from the ground up to take advantage of performance features in modern CPUs, resulting in dramatically higher data processing rates compared to all other approaches. As a result, Action Vector can process Hadoop data much faster than any other relational databases. Much faster data processing performance accelerates time to actionable decisions.



Comprehensive Enterprise-Strength Capabilities

- **Performance** – quickest query response times; benchmarked up to 30 times faster than the closest competitor with broader query support
- **Accessible** – standard ANSI SQL-92 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
- **Reliable** – fully ACID-compliant with multi-version read consistency, plus system-wide failover protection
- **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 (more than 100 nodes)
- **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 analytical processing
- **Manageable** – query workloads managed automatically in Hadoop via YARN

The Only 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 and require you to move your data out of Hadoop. Only Actian Vector combines the most modern, scalable, high-performing database technology with the power of Hadoop, thereby enabling users to directly query data stored in HDFS.

Up to 30 Times Faster than the Competition

Actian Vector, the world's fastest database, has been proven to run up to 30 times faster than Cloudera Impala based on performance testing using their published benchmark environment. Give your SQL users and applications the performance and enterprise-grade SQL access they need to do their jobs.

Native Hadoop Support & YARN Ready

Actian Vector for Hadoop is optimized for use with Hortonworks HDP 2.1 and is the first SQL on Hadoop to be YARN ready. It also supports all other major distributions and versions of Hadoop, including Apache, Cloudera, and MapR.



www.actian.com

500 Arguello Street, Ste. 200, Redwood City, CA 94063
+1.888.446.4737 [Toll Free] | +1.650.587.5500 [Tel]



© 2014 Actian Corporation. Actian, Big Data for the Rest of Us, Accelerating Big Data 2.0, and Actian Analytics Platform are trademarks of Actian Corporation and its subsidiaries. All other trademarks, trade names, service marks, and logos referenced herein belong to their respective companies. (DS14-0514)