jethro

Accelerating Big Data Performance

Data Sheet

JethroData for Qlik®: Truly Interactive BI on Hadoop

Qlik®

2015 Partner Global Award Winner

November 2015

> Jethro brings **significant performance advances** to accessing data in Hadoop, which combined with the Qlik associative experience, will enable our customers to continue discovering business value from their data—regardless of the variety or volume.
>
> — Les Bonney, COO, Qlik

Jethro for Qlik enables using SQL-on-Hadoop to gain near real-time access to big data with QlikView® and Qlik® Sense.

Jethro for Qlik provides an innovative index-based SQL engine enabling truly interactive BI on Big Data for both QlikView and Qlik Sense. It works by fully indexing select datasets on Hadoop HDFS or Amazon S3 – every single column is indexed. Queries use the indexes to access only the data they need instead of performing a full scan, resulting in a much faster response time and lower system resources utilization. Queries can leverage multiple indexes for better performance – the more you drill-down, the faster your query runs.

Jethro for Qlik Highlights

- ### Easy to set up

  Just install it on a dedicated server or two, point to your Hadoop cluster or Amazon S3 bucket, load some data and start querying. When working with Hadoop, Jethro is safe and easy to implement as it only connects remotely as an HDFS client.

- ### Performance from a unique indexing technology

  The key to Jethro's superior performance is its unique indexing technology. Jethro's indexes are sorted, multi-hierarchy, compressed bitmaps. They are created automatically for every column, and are written in an efficient, append-only fashion – avoiding expensive random writes and locking. Queries use indexes to read only the data they need, instead of performing full scans, leading to faster response time.

- ### Scalability and high-availability

  Jethro's nodes are stateless and highly elastic, allowing easy scale-out to meet concurrency requirements. Jethro's index and column files are stored as standard files on HDFS or Amazon S3 and benefit from their native scalability and high availability.

- ### Minimal Hadoop cluster load

  Other SQL-on-Hadoop solutions uses a brute force method – each node of the cluster scans and processes its local data for every query. In contrast, Jethro leverages its indexes to surgically fetch only the relevant data for each query, dramatically reducing the load on the shared Hadoop cluster – freeing it for other computations and supporting more concurrent queries.

Jethro for Qlik®: High-level Architecture

With the ability to interactively access data stored in Hadoop using Jethro for QlikView® and Qlik® Sense, customers gain high performance big data access with all of the associative analysis options of Qlik. By gaining high performance access without complicated data extracts or data silos, users maintain all the associative data model and search options of Qlik, exploring and drilling information freely in any direction rather than being confined to a predefined path of questions.

- ### SQL Interface

  Qlik connects to Jethro using its JDBC or ODBC driver and issues standard SQL queries. The driver automatically load-balances across all Jethro hosts.

- ### Query Processing

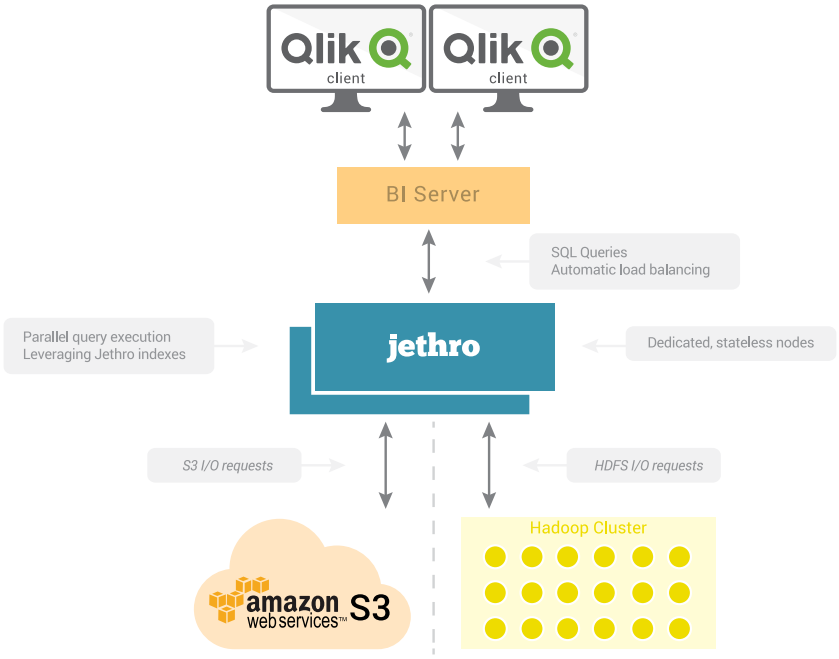
  Jethro typically runs on one or few dedicated, higher-end hosts optimized for SQL. The query hosts are stateless, and new ones can be dynamically added to support additional concurrent users.

- ### Storage Layer

  Jethro stores all its indexes and files in an existing Hadoop cluster or in an Amazon S3 bucket. With Hadoop, it uses a standard HDFS client (libhdfs) and is compatible with all common Hadoop distributions. Jethro only generates a light I/O load on HDFS – offloading SQL processing from Hadoop and enabling sharing the cluster between online users and batch processing.

- ### Data Loading & Indexes

  A loader service processes input files and creates query-optimized column and index files, which are encoded, compressed and then stored on HDFS or Amazon S3. This service can run on its own host or on one of the query processing hosts



Experience the Performance Advantage of JethroData for Qlik

View a live demo of Jethro for Qlik at qlik.jethrodata.com to see a live demo and experience the performance advantage of Jethro for Qlik.

About Jethro

Jethro was founded by a team of industry veterans committed to making big data analytics work in real time. Our passion is solving big problems, in this case building the technology that lets non-technical users interactively explore data on Hadoop and get immediate answers, using standard SQL or common BI tools.

In June 2015, Jethro closed an $8.1 million Series B financing round led by Square Peg Capital and existing investor Pitango Venture Capital.

Jethro is headquartered in New York City with an R&D office in Israel.