

SQL and Hadoop Moving Forward Together: RainStor Big Data Analytics for Hadoop

Abstract

In January 2012, RainStor announced the launch of its latest edition – RainStor Big Data Analytics on Hadoop. This ENTERPRISE MANAGEMENT ASSOCIATES (EMA) Impact Brief covers the announcement and identifies RainStor as a first mover in native integration of analytical capability with data storage on Hadoop infrastructure.

RainStor Announces Big Data Analytics on Hadoop

On January 18, 2012, RainStor (<http://rainstor.com>) announced the introduction of its newest product edition - RainStor Big Data Analytics on Hadoop. RainStor Big Data Analytics is an enterprise data store running natively on the Hadoop Distributed File System (HDFS™) for its storage requirements. With this configuration, RainStor performs analytics on multi-structured data without the requirement to move the data from a Hadoop HDFS to another storage environment. In addition, RainStor Big Data Analytics for Hadoop accesses this multi-structured data with both SQL and NoSQL (Pig) query languages.

With the product launch, RainStor announced partnerships with Cloudera (<http://www.cloudera.com>), Hortonworks (<http://hortonworks.com>) and MapR (<http://www.mapr.com>). These organizations are all significant contributors and distributors of the open source Hadoop code bases(s).

Founded in 2004, RainStor is a private equity firm with offices in San Francisco, California and Gloucester, United Kingdom. Sales and operations are handled from the RainStor headquarters in San Francisco. Engineering and support for end-user customers in the Communications, Financial Services, Retail, Security; Utilities; Manufacturing; and Oil and Gas industries are handled from their offices in the United Kingdom.

RainStor provides data management solutions for clients with Big Data requirements. RainStor's solutions feature patented compression algorithms and run on a variety of infrastructure choices that include Cloud and On-Premise. Technical architecture options include various editions of Windows, Unix and Linux operating systems; and Storage Area Network (SAN), Network-Attached Storage (NAS); Content Addressable Storage (CAS); and Hadoop HDFS storage systems.

Key Observations

- **Architectural Synergy:** RainStor's Big Data Analytics offering allows IT departments to exploit multi-structured information stores using its mature query interface at Big Data volumes using Hadoop HDFS storage infrastructures.
- **Leverage Multiple Learning Curves:** Employing a pair of analytical paradigms (SQL and NoSQL), allows for organizations to leverage existing business analyst knowledge base and utilize existing analytical infrastructure investment via JDBC/ODBC connectivity as well as embrace leading edge analytics like social graph and sentiment analysis.
- **Lower TCO:** Leveraging the economic advantages of the commodity hardware utilized in HDFS implementations and using the RainStor compression allows data center teams to lower their total cost of ownership (TCO) in terms of both capital and operational expense.
- **Details:** For more detailed information on the RainStor announcement, go to: <http://rainstor.com/rainstor-announces-first-enterprise-database-running-natively-on-hadoop-for-faster-analytics-at-lower-operating-cost/>

SQL and Hadoop Moving Forward Together: RainStor Big Data Analytics for Hadoop

EMA Perspective

A debate has begun with the emergence of Big Data requirements in terms of volume, variety and velocity; and the development of the Hadoop ecosystem as an alternative platform to existing analytical and storage environments like enterprise data warehouses and analytical database management platforms. The discussion revolves around Big Data analytical workloads and data storage:

- What is the best platform for Big Data workloads?
- Where is the best location to store Big Data?

With the release of Big Data Analytics on Hadoop, RainStor is making a push to answer some of these questions.

Many years ago, Henry Ford said that “If everyone is moving forward together, then success takes care of itself.” By utilizing the Hadoop HDFS platform for storage and their existing interface for queries, RainStor is aiming to create that “success” by “moving forward together” with two important components. RainStor can utilize a leading edge, Big Data storage option in Hadoop HDFS. They also offer an established and robust query interface to the Big Data information. These two components expand the options available for IT teams when faced with the challenges of multi-structured data requirements.

By taking this approach, RainStor is exploiting a pair of data access knowledge bases. With Big Data Analytics on Hadoop, IT departments can utilize both standard SQL and new NoSQL, via Hadoop Pig, interfaces to access the multi-structured data within the RainStor solution. Utilizing SQL allows RainStor customers to leverage existing SQL knowledge assets and links to business intelligence and analytics infrastructure via the JDBC/ODBC connectivity drivers. Using the Hadoop Pig (<http://pig.apache.org/>) programming platform, RainStor allows organizations to leverage the growing Hadoop development knowledge base with in corporate organizations to take advantage of NoSQL based analytical result sets like social graph and sentiment analysis.

With the adoption of the HDFS storage option and utilizing RainStor’s patented compression algorithms, IT groups can make the already economical HDFS storage option even more affordable by reducing the overall footprint the RainStor Big Data Analytics on Hadoop total cost of ownership (TCO). With the signification reduction in footprint of HDFS storage servers, data center teams can realize savings in terms of operational headcount to support the environment as well as OPEX savings in terms of electrical and environmental expenditures.

In this set of product and partnership announcements, RainStor is widening the opportunities for business teams to make analytical decisions relating to their multi-structured data. Also IT departments are not locked into a single analytical architecture whether that architecture is SQL or NoSQL based, to support business requirements.

About EMA

Founded in 1996, Enterprise Management Associates (EMA) is a leading industry analyst firm that specializes in going “beyond the surface” to provide deep insight across the full spectrum of IT management technologies. EMA analysts leverage a unique combination of practical experience, insight into industry best practices, and in-depth knowledge of current and planned vendor solutions to help its clients achieve their goals. Learn more about EMA research, analysis, and consulting services for enterprise IT professionals and IT vendors at www.enterprisemanagement.com or follow EMA on Twitter (http://twitter.com/ema_research).

0000.000000