



Syncsort and Cloudera Integrate Detailed Data Lineage from All Enterprise Data for Superior Data Governance

Extends Cloudera Navigator's Reach Beyond the Hadoop Cluster for Comprehensive Insights into Where Data Came From, What Changes Were Made and Where it Lands

PEARL RIVER, N.Y., February 27, 2018 – [Syncsort](#), the global leader in Big Iron to Big Data software, today announced that it has delivered expanded integration between its industry-leading, high-performance data integration solution, DMX-h, and Cloudera Navigator. This combination uniquely enables data governance practitioners to view detailed data lineage information on enterprise-wide data as well as changes to the data both outside and within the Hadoop cluster.

In Syncsort's recently-published [Big Data survey](#), data governance was ranked among the top priorities when implementing a data lake. Nearly 60% of respondents who are testing or in production with Hadoop or Spark identified "including the data lake in data governance initiatives and meeting regulatory compliance mandates" as a significant challenge. Financial Services and Insurance professionals were twice as likely as respondents from other industries to cite enterprise-wide data governance as one of their most significant challenges. The study also found that data from legacy platforms continues to play a significant role in the data lake, and cloud repositories are gaining popularity as a data source as more organizations leverage the cloud as a deployment platform.

In keeping with these findings, and to gain comprehensive insights for data governance initiatives, organizations need visibility into what happens to data as it is collected from diverse data sources such as Mainframe, IBM i, RDBMS and other legacy sources, and where it resides in Hadoop clusters. The new solution provides detailed data lineage from all these data sources and integrates with Cloudera Navigator to provide an unmatched, detailed and comprehensive view.

"With the maturity of Hadoop as an enterprise data platform, organizations are using it to store and process significantly more data, and, in turn, more users and tools are accessing the data. The opportunity to drive greater insights is remarkable, but the volumes, diverse data sources and hybrid environments create a big governance challenge," said Tendü Yoğurtçu, CTO, Syncsort. "Cloudera recognized these challenges early on and developed Cloudera Navigator to address them. Syncsort's DMX-h data integration seamlessly integrates with Cloudera Navigator to deliver detailed data lineage information regardless of whether the data movement and transformation process was run inside or outside of Hadoop, on-premise or in the cloud. The new solution helps enterprises integrate their entire data ecosystem with Cloudera Navigator, while meeting regulatory compliance requirements."

The new integrated solution provides detailed data lineage information on:

- **Enterprise-Wide Data Coming from Outside the Hadoop Cluster:** DMX-h now leverages field-level metadata to provide detailed data lineage information on everything that happened to the data on-the-fly as it consumes data from diverse data sources, transforms it and delivers it to the Cloudera Enterprise Data Hub.
- **Data Residing Inside the Hadoop Cluster:** Cloudera Navigator tracks field-level metadata on data lineage in the cluster.
- **Changes to Data Inside the Hadoop Cluster:** DMX-h is also used for Data Integration within the cluster. ETL jobs created in the DMX-h point-and-click interface can be run on MapReduce,

Spark, or stand-alone Windows/Linux/Unix systems, on-premise or in the cloud. All data processing details are published to Cloudera Navigator.

- **Data Inside and Outside the Cluster in One Consolidated View:** The integrated solution provides a consolidated end-to-end view of all the detailed data lineage information from Syncsort DMX-h and Cloudera Navigator in the Cloudera Navigator Dashboard. The solution connects Hadoop-based governance with all enterprise data for superior audit, data lineage, metadata management and policy enforcement capabilities, on-premise or in the cloud, including complex hybrid environments.

“In support of industry-wide data governance requirements, organizations need to address the challenges of managing enterprise-level volumes of data originating from multiple data sources, including mainframe and other legacy systems, monitoring the movement of data through extensive processing chains, and collecting this data for consumption by governance technologies,” said Philippe Marinier, Vice President, Business Development, Cloudera. “Syncsort’s DMX-h delivers these requirements with its high-performance data integration capabilities, and the integration with Cloudera Navigator provides a comprehensive, end-to-end view of data lineage.”

To learn more about the integrated Syncsort DMX-h and Cloudera Navigator solution, visit booth #1022 at [Strata Data Conference](#) in San Jose from March 6-8, 2018, or watch our webinar [End-to-End, Source to Analytics, Data Lineage with Syncsort DMX-h](#).

About Syncsort

Syncsort is a trusted enterprise software provider and the global leader in Big Iron to Big Data solutions. More than 6,000 organizations, including 84 of the Fortune 100, use the company’s products to solve their most complex data management challenges, on premise and in the cloud. Syncsort helps customers optimize traditional data systems and deliver mission-critical data from these systems to next-generation analytic environments. Its Big Iron to Big Data portfolio now features the #1 high availability product for IBM i Power Systems, powerful cross-platform capacity management, best-in-class mainframe app and machine data access & integration and market-leading data quality capabilities. Rediscover Syncsort at syncsort.com.

Media Contacts:

SHIFT Communications
Sarah Babbitt, 617-779-1803
syncsort@shiftcomm.com

or

Syncsort Incorporated
Michael Kornspan, 201-930-8216
Senior Director, Corporate Communications
mkornspan@syncsort.com