



Partner Brief

ETL with DMX-H and Hortonworks Data Platform

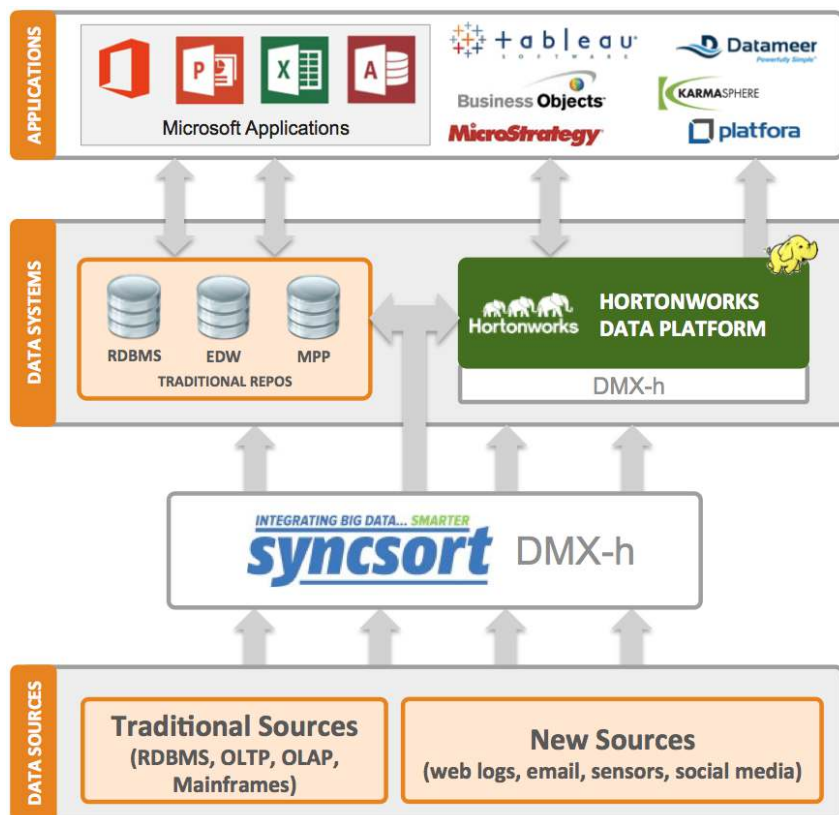


Apache Hadoop has emerged as the de facto Big Data analytics operating system to help deal with the avalanche of data coming from logs, email, sensor devices, mobile devices, social and more. While business intelligence systems are typically the last stop in extracting value from Big Data, the first

stop is commonly manipulation of the data in the process of ETL: “Extract, Transform, Load”. ETL is the process by which data is moved from source systems, manipulated into a consumable format and loaded into a target system for performing advance analytics, analysis and reporting. ETL is emerging as the one of the key use cases for Hadoop implementations. In fact, industry analyst Gartner recognizes that “most organizations will adapt their data integration strategy using Hadoop as a form of preprocessor for Big Data integration in the data warehouse.”

Syncsort, a technology partner with Hortonworks, helps organizations ease into Hadoop ETL projects with an integration tool that minimizes hurdles to pave the way for wider adoption and allows for the full potential of Hadoop to be realized.

Syncsort in the Modern Data Architecture



Features & Benefits of Combined Solution

Syncsort delivers a smarter approach to ETL with Hortonworks Data Platform and Hadoop

- Connect to virtually any data source, including mainframe and MPP databases
- Move data into and out of Hadoop up to 6x faster without the need for manual scripts
- Develop MapReduce ETL processes without writing code
- Seamlessly accelerate Hadoop performance and scalability for sort and ETL operations

One tool to connect HDP to all sources and targets

DMX-h can also extract data from HDP to other data stores, leveraging HDP as a key processing step



What Business Challenges are Solved?

- **Faster Time to Insight**
Organizations make better decisions faster based on more accurate insights, by processing more data in less time, with the same resources.
- **Lower OPEX and CAPEX Costs**
Organizations reduce capital and operational expenses by eliminating the need for additional compute nodes on the cluster, due to more efficient hardware utilization.
- **Better Performance and Scalability**
Organizations can process up to 3x more data in less time using the same, or fewer, resources. Better performance means jobs finish sooner, freeing the cluster to handle more jobs within the same processing windows, avoiding incremental capital expense.

"Syncsort and Hortonworks share a common goal of making Apache Hadoop easier to use and deploy in conjunction with proven, enterprise-class Big Data technologies. Together, Syncsort DMX-h and Hortonworks Data Platform form a powerful combination to help customers seamlessly maximize the business value and minimize the cost of their Hadoop deployments."

*---Gary Survis,
Chief Marketing Office,
Syncsort Data Integration*

Syncsort is a Certified Technology Partner

The Hortonworks Certified Technology Program reviews and certifies technologies for architectural best practices, validated against a comprehensive suite of integration test cases, benchmarked for scale under varied workloads and comprehensively documented.

For additional questions, contact:



Syncsort
+1(201) 930-8200
www.syncsort.com/hadoop
Hortonworks
USA: (855) 8-HORTON (1 for sales)
+1(408) 916-4121
www.hortonworks.com

About Hortonworks

Hortonworks develops, distributes and supports the only 100-percent open source distribution of Apache Hadoop explicitly architected, built and tested for enterprise-grade deployments.



3460 West Bayshore Road
Palo Alto, CA 94303 USA

US: 1.855.846.7866
International: 1.408.916.4121
www.hortonworks.com