



# ETL with DMX-H and Hortonworks Data Platform



Partner Brief



"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 Officer,  
Syncsort Data Integration

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 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 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 integrates with Hadoop and HDP directly through YARN, making it easier for users to write and maintain MapReduce jobs graphically. Additionally, through the YARN integration, the processing initiated by DMX-h within the HDP cluster will make better use of the resources and execute more efficiently.

## 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.

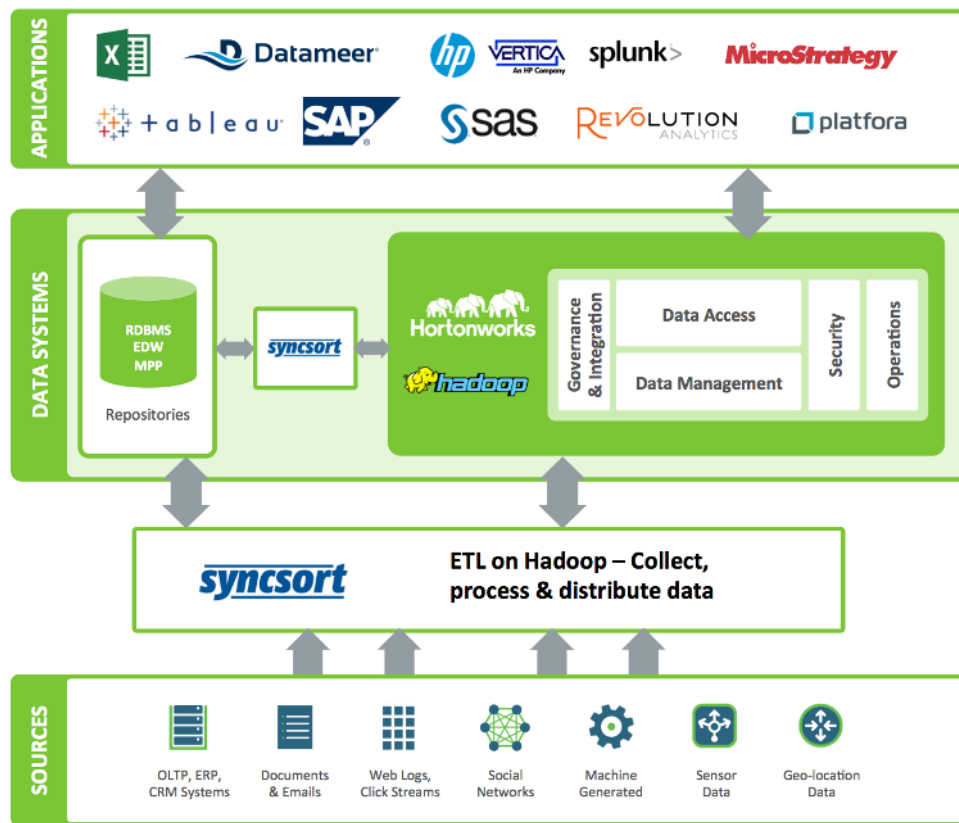
### Better Performance and Scalability

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.

### Lower OPEX and CAPEX

Organizations reduce capital and operational expenses by eliminating the need for additional compute nodes on the cluster, due to more efficient hardware utilization.

# 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 Hadoop ETL processes without writing code
- Seamlessly accelerate Hadoop performance and scalability for sort and ETL operations
- Optimize your data warehouse
- Offload heavy ETL workloads to Hortonworks Data Platform
- Deploy as part of your Hortonworks Hadoop cluster
- Fully support Kerberos & LDAP

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

## 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 and benchmarked for scale under varied workloads.



## Hortonworks. We do Hadoop.

Hortonworks is a leading commercial vendor of Apache Hadoop, the open source platform for storing, managing and analyzing Big Data. Hortonworks Data Platform, our distribution of Apache Hadoop, provides an open and stable foundation for enterprises and a growing ecosystem to build and deploy Big Data solutions.

Hortonworks is the trusted source for information on Hadoop, and together with the Apache community, Hortonworks is making Hadoop an enterprise data platform. Hortonworks provides unmatched technical support, training and certification programs for enterprises, systems integrators and technology vendors.

For additional questions, contact:

- Syncsort  
[www.syncsort.com/hadoop](http://www.syncsort.com/hadoop)  
(201) 930-8200
- Hortonworks  
[www.hortonworks.com](http://www.hortonworks.com)  
(855) 8-HORTON