Hortonworks

Non-Stop Hadoop for the Hortonworks Data Platform

WANDISCO

Partner Brief

"Global organizations require big data solutions that can meet stringent data and availability needs across wide geographies. We believe the integration of WANdisco's active-active replication technology with the Hortonworks enterprise Apache Hadoop distribution makes sense for SAP customers looking to leverage big data."

Irfan Khan
Senior VP and GM
SAP Big Data

WANdisco has patented replication technology for Non-Stop Hadoop on the Hortonworks Data Platform (HDP) created by some of the original architects and developers of Hadoop. WANdisco delivers the first and only continuous availability solution for globally distributed multi-data center deployments. All Name Node servers in a Hadoop cluster deployed over a WAN actively support clients at each location and along with the data nodes are continuously synchronized for LAN-speed performance and access to the same data at every location. Failover and recovery are automatic both within and across data centers. WANdisco's patented technology has also been applied to deliver 100% uptime with enhanced performance and scalability for HBase – the NoSQL database for real-time interactive applications running on HDP.

Unlike other high availability solutions that only work within a single data center over a LAN, or solutions limited to a certain distance across a Metropolitan LAN, Non-Stop Hadoop for Hortonworks is optimized for use over a WAN between data centers thousands of miles apart, delivering the same capabilities regardless of distance. Enterprises can achieve continuous availability and design disaster recovery strategies that meet their global business continuity requirements.

WANdisco's patented Non-Stop technology enables continuous global availability for HBase – the NoSQL database for real-time interactive application running on Hadoop. Non-Stop Hadoop for Hortonworks eliminates the performance and availability bottlenecks in HBase's architecture and its region and master servers – by enabling each HBase instance to have multiple active master servers, as well as multiple active region servers for each region. This builds on the solid foundation laid by Non-Stop Hadoop's multiple active NameNode architecture, making HBase continuously available across data centers any distance apart.

What Business Challenges are Solved?

Global Business Continuity

Non-Stop Hadoop is optimized for use over a WAN between data centers thousands of miles apart, delivering 100% uptime, with LAN-speed read and write access to the same data at every location. Multiple active NameNodes enable performance and scalability with 100% availability.
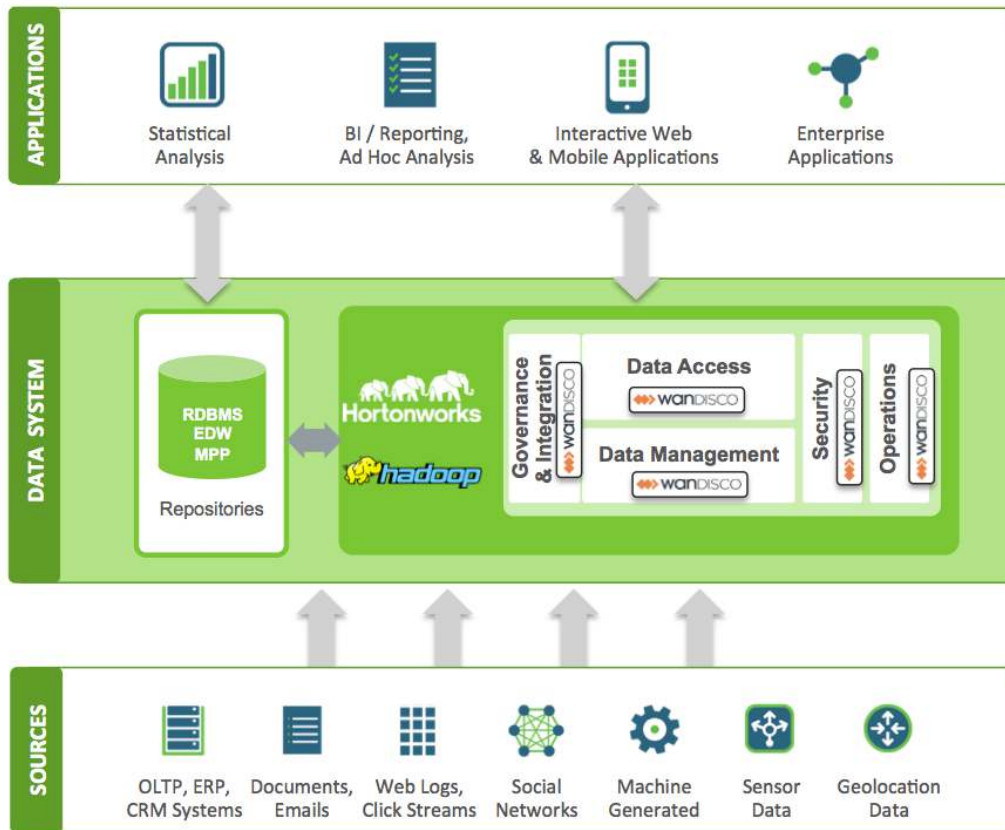
Continuous Availability

100% uptime for applications running on Hadoop even during upgrades to the Hadoop platform.

Automatic Backup, Failover and Recovery

Provides built-in hot backup with automated failover and recovery to prevent downtime and data loss regardless of whether a single server or an entire data center goes offline.

www.hortonworks.com

© 2014 Hortonworks, Inc. All rights reserved. Hadoop and the Hadoop elephant logo are trademarks of the Apache Software Foundation.

WANdisco in the Modern Data Architecture

APPLICATIONS
- Statistical Analysis
- BI / Reporting, Ad Hoc Analysis
- Interactive Web & Mobile Applications
- Enterprise Applications

DATA SYSTEM
- RDBMS EDW MPP Repositories
- Hortonworks
- Hadoop
- Governance & Integration — wandisco
- Data Access — wandisco
- Data Management — wandisco
- Security — wandisco
- Operations — wandisco

SOURCES
- OLTP, ERP, CRM Systems
- Documents, Emails
- Web Logs, Click Streams
- Social Networks
- Machine Generated
- Sensor Data
- Geolocation Data

Features and Benefits of the Combined Solution

- Certified solution on HDP
- Non-stop Hadoop for Hortonworks enables a single cluster running HBase to be deployed across multiple data centers any distance apart. HDP and HBAse are always available
- NameNodes, HBase region and master servers are continuously synchronized between globally distributed data centers
- Eliminates single point of failure
- Automated failover and recovery: if a critical server fails, their replicated peers take over
- 100% hardware utilization reduces costs and delivers maximum compute power

Continuous availability solution with WANdisco

WANdisco has a patented active-active technology for non-stop Hadoop for Hortonworks and can be applied to global multi-data center deployments of HDP.

WANdisco 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.

Hortonworks CERTIFIED TECHNOLOGY PARTNER

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:

- WANdisco
  www.wandisco.com
  (877) WANDISCO
- Hortonworks
  www.hortonworks.com
  (855) 8-HORTON