

Pentaho and Cloudera Solution Brief

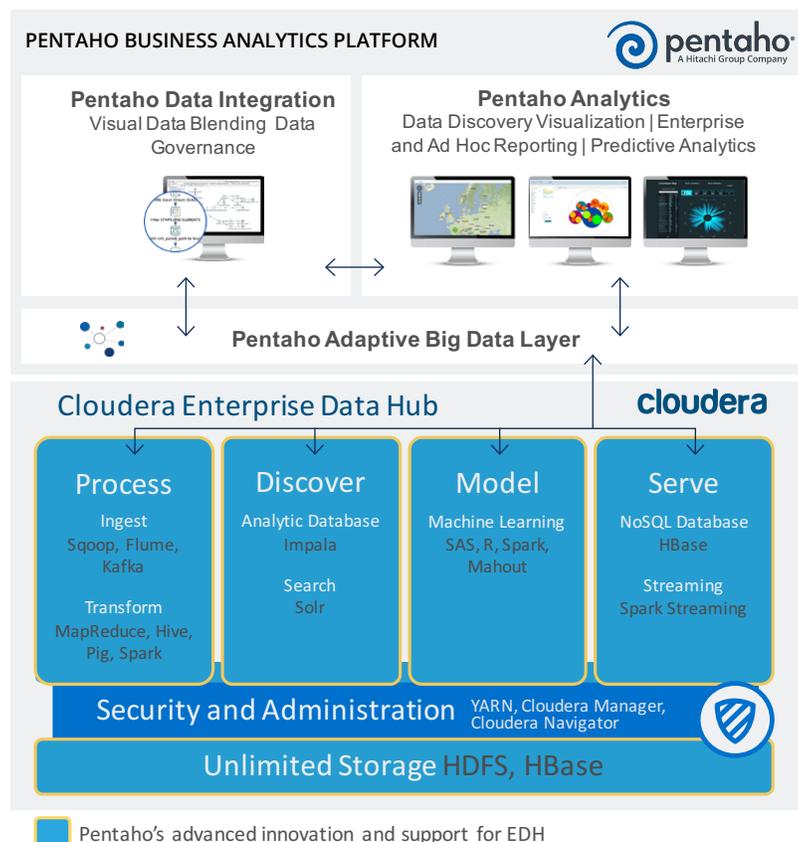
Accelerating Time to Big Data Value

Empowering the Information-Driven Enterprise

Enterprises today collect and generate more data than ever before. Business users ask more sophisticated questions and explore data at a level of detail that can tax existing data management systems. To win in a competitive marketplace by obtaining insights to improve products, understanding customer sentiment, optimizing marketing promotions, and detecting and preventing fraud, organizations must combine a variety of data from many sources to extract real, actionable insight.

Cloudera and Pentaho Realizing Insights from Hadoop

With Pentaho and Cloudera, organizations get a comprehensive view of data across their business. Cloudera's Enterprise Data Hub (EDH), powered by Apache Hadoop, delivers a complete data management platform: one place to store, process, and analyze data in full fidelity that complements existing storage strategies and analytic infrastructure investments. Pentaho's comprehensive analytics platform, data integration, and a spectrum of data visualization and analysis capabilities can bring this data to life. Deeply integrated with technologies such as Cloudera Impala, Search and YARN, Pentaho's Analytics Platform is optimized for the EDH.



ABOUT CLOUDERA

Cloudera is revolutionizing enterprise data management by offering the first unified Platform for Big Data: The Enterprise Data Hub. Cloudera offers enterprises one place to store, process and analyze all their data, empowering them to extend the value of existing investments while enabling fundamental new ways to derive value from their data.

Founded in 2008, Cloudera was the first and is still today the leading provider and supporter of Hadoop for the enterprise. Cloudera also offers software for business critical data challenges including storage, access, management, analysis, security and search. With over 40,000 individuals trained, Cloudera is a leading educator of data professionals, offering the industry's broadest array of Hadoop training and certification programs. www.cloudera.com

IT Managers

DECREASE COSTS AND REDUCE TIME IN DATA PREPARATION AND MANAGEMENT

- Pentaho's Visual MapReduce enables ETL via its intuitive drag-and-drop interface, simplifying and reducing time and complexity of data integration.
- Infrastructure costs are reduced as ETL workloads previously run on expensive systems can migrate to a Hadoop active archive in a Cloudera enterprise data hub to run at very low cost, in parallel, with increased performance and accessibility.

Developers

ACCELERATE BIG DATA ADOPTION WITH FAMILIAR TOOLS

- Pentaho's drag-and-drop ETL handles Hadoop data sources in the same way as traditional relational and file-based data sources. Instead of writing Java programs, MapReduce jobs or Pig scripts, Pentaho empowers developers to design and develop big data jobs using visual tools - resulting in greater team productivity and efficiency.
- With Pentaho's support for Cloudera Impala, the world's first open source MPP SQL query engine expressly built for Hadoop, large volumes of disparate data can be refined, blended, enriched and visualized faster than what was previously possible with Hive alone.

Business Users

UNCOVER PREVIOUSLY UNSEEN DATA PATTERNS

- With Pentaho's Analytics Platform users discover deep insight into data stored in the entire enterprise data hub to quickly spot patterns and identify insights.
- Pentaho provides highly interactive visualizations, advanced and predictive analytics. Self-service and production reporting, dashboards complete the analytics capabilities.

"What impressed me is that ETL world and big data world integrated quite well with Pentaho. The migration of moving through job flows to MapReduce jobs was easy in Pentaho Data Integrations visual interface. No one else is doing this and it saved us an immense amount of time...Using analyzer makes me a dangerous data analyst."

JEFF SIPPEL, CTO
edo

Customer Success: When Building Customer Loyalty is a Matter of Survival

edo Interactive connects brands with consumers by harnessing billions of daily data records related to customer conduct, synthesizing trends and delivering personalized offers most likely to trigger a sale.

For edo, having lightning fast data analytics systems is a matter of survival. The company has to keep innovating in a crowded social, local and mobile advertising market. But in early 2013, edo hit a wall. The team couldn't process data fast enough with their existing SQL database and as a result, couldn't get the right offers to the right people quickly enough.

Working with Cloudera and Pentaho, edo has dramatically scaled the amount of data they can process. Pentaho Business Analytics utilizes Cloudera Impala, Hive and HBase, to streamline edo's preparation and analytics processes. As a result, they cut the processing window from 29 hours to under four hours all while growing the amount of data processed 974%! This has been a key factor to customer preference and retention.

ABOUT PENTAHO

Pentaho, a Hitachi Group Company, is a leading data integration and business analytics company with an enterprise-class, open source-based platform for diverse big data deployments. Pentaho's unified data integration and analytics platform is comprehensive, completely embeddable and delivers governed data to power any analytics in any environment. Pentaho's mission is to help organizations across multiple industries harness the value from all their data, including big data and IoT, enabling them to find new revenue streams, operate more efficiently, deliver outstanding service and minimize risk. Pentaho has over 15,000 product deployments and 1,500 commercial customers. For more information visit www.pentaho.com.

Be social
with Pentaho:

