

Pentaho Big Data Integration and Analytics

DATASHEET

Within one platform — Pentaho — Hitachi Vantara provides big data tools to extract, prepare and blend your data, plus the visual analytics that will change the way you run your business. From Hadoop and Spark to NoSQL and analytic databases, Pentaho allows you to turn big data into big insights.

Broad and Adaptive Big Data Integration

Deep native connections and an adaptive big data layer accelerate access to the latest versions and capabilities of popular big data stores (see Figure 1).

- Graphical extract-transform-load (ETL) designer to simplify the creation and orchestration of big data pipelines.
- Ability to access data once and then process, combine and consume it anywhere, providing insulation from changes in the big data ecosystem.
- Support for the latest Hadoop distributions, NoSQL stores, object stores and analytic databases on premises and in the cloud, as well as Spark, AMQP, MQTT, AWS Kinesis, and Kafka.
- Automation to rapidly accelerate the ingestion and onboarding of hundreds or thousands of diverse and changing data sources into Hadoop.
- Visual design tools that empower faster development in Spark or MapReduce versus hand coding or other design tools (see Figure 2).
- Seamless switching or combining of data processing engines with in-cluster execution, load balancing and container support to maximize existing processing capacity.

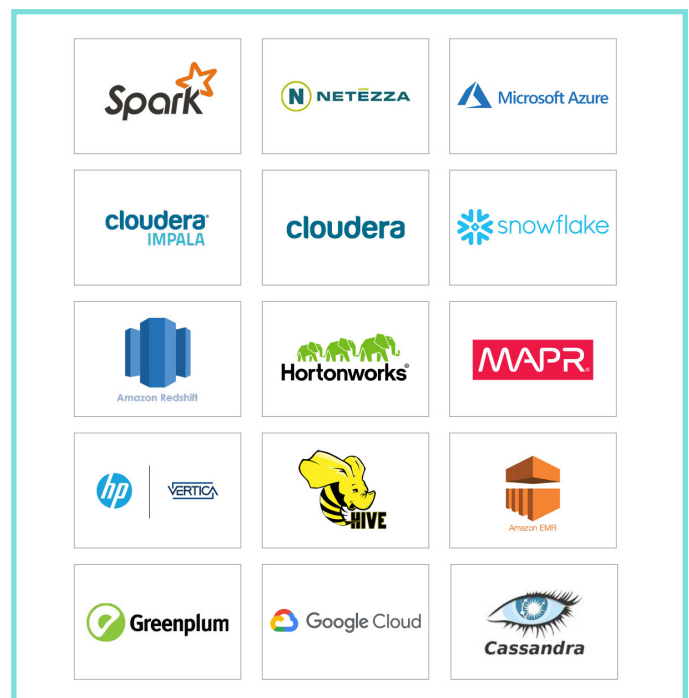


Figure 1. Pentaho integrates with a full spectrum of big data sources.

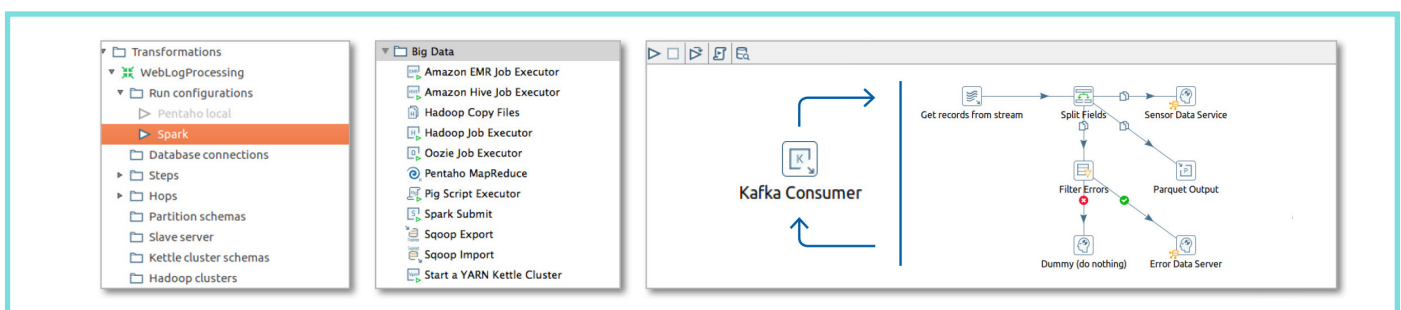


Figure 2. Visually design transformations that run in-cluster.

Power a Complete Analytic Data Pipeline

A tightly coupled data integration and business analytics platform accelerates the realization of value from blended big data.

- Provides full array of analytics: from data access and integration to data visualization and predictive analytics.
- Empowers users to architect big data blends at the source and stream them directly for more complete and accurate analytics in the form of interactive visualizations, dashboards and reports.
- Operationalizes data science models from R, Python, Scala or Weka, including ones that use Spark MLlib, scikit-learn, Tensorflow and Keras libraries, as a part of the data integration workflow, while reducing data prep time.
- Open, standards-based architecture makes it easy to integrate with or extend existing infrastructure (see Figure 3).
- Seamlessly scale data transformations with containerized architecture.

An Expert Partner With Proven Implementations

Technology only gets you so far. People, experience and best practices are the most important drivers for project success with big data technologies.

- Big data blueprint design patterns that have allowed organizations to optimize infrastructure, refine big data for self-service analytics, create a complete view of customers, and monetize analytics as a service.
- Successful big data project across verticals, including fraud detection in financial services with Wirecard, self-service analytics in education with Ivy Tech, infrastructure optimization in hardware with NetApp, and cyber security data integration with BT.
- A dedicated team of big data services experts with offerings for every phase of the implementation life cycle, including training workshops, technical account management, solution delivery and engineering services.

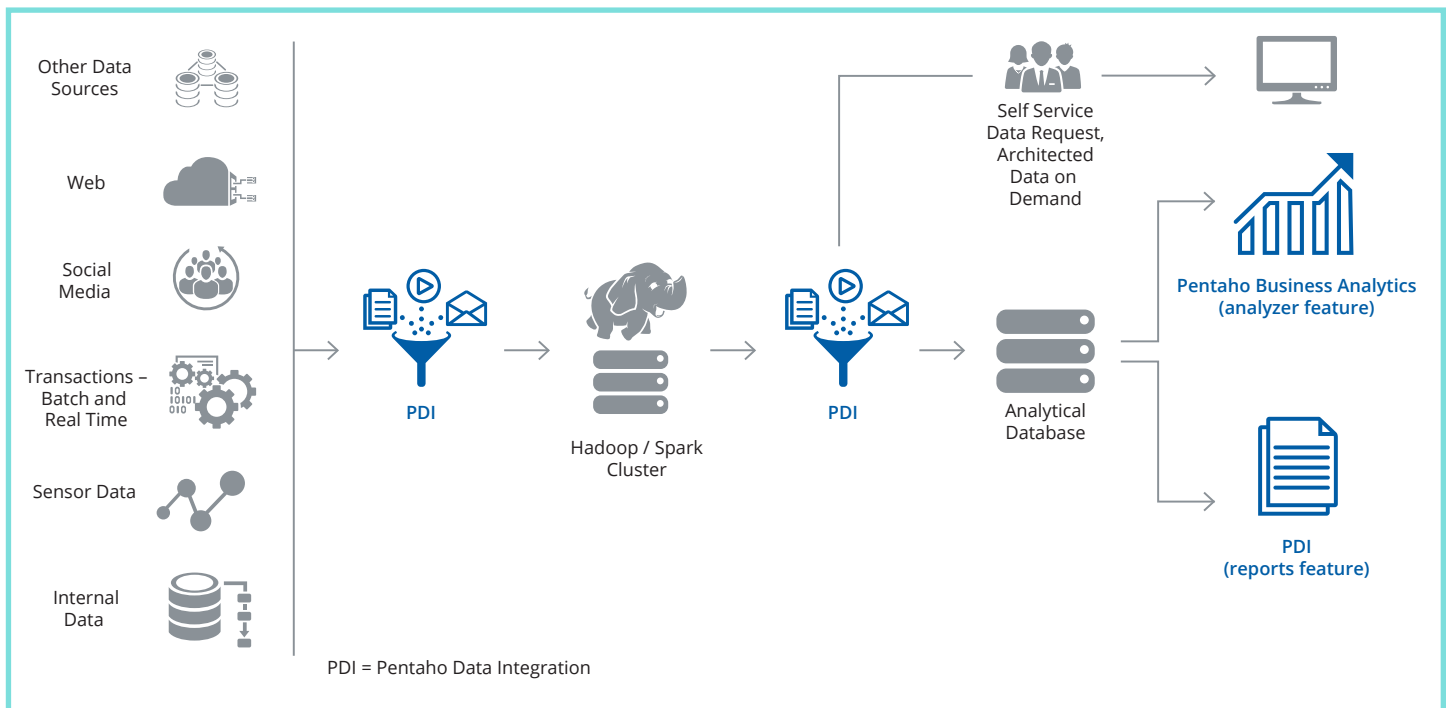


Figure 3. Sample end-to-end big data architecture with Pentaho: Integrate, blend and analyze big data for new insights.

We Are Hitachi Vantara

DataOps is the data practice for the AI era, connecting data consumers with data creators to accelerate collaboration and digital innovation. We are analytics, industrial expertise, technology and outcomes rolled into one great solution partner. Get Your DataOps Advantage.

Hitachi Vantara



Corporate Headquarters
2535 Augustine Drive
Santa Clara, CA 95054 USA
hitachivantara.com | community.hitachivantara.com

Contact Information
USA: 1-800-446-0744
Global: 1-858-547-4526
hitachivantara.com/contact

HITACHI is a registered trademark of Hitachi, Ltd. Pentaho is a trademark or registered trademark of Hitachi Vantara Corporation. All other trademarks, service marks, and company names are properties of their respective owners.