Cloudera Enterprise Data Hub

*The modern platform for machine learning and analytics that’s optimized for the cloud.*

Many of the world’s largest companies rely on Cloudera’s multifunction and multi-environment platform. This platform provides the foundation for their critical business value drivers: growing business, protecting business, and connecting products and services.

Here’s some of what makes Cloudera Enterprise different from other data platforms:

**Enterprise grade**

The scale and performance required for today’s modern data workloads meet the security and governance demands of today’s IT. Cloudera’s modern platform makes it easy to bring thousands of users to petabytes of diverse data and provides industry-leading engines to process and query data. In addition to developing and serving models quickly, the platform also provides several layers of fine-grained security and complete auditability for companies to prevent unauthorized data access and demonstrate accountability for actions taken.

**Shared data experience (SDX)**

Eliminate costly and unnecessary application silos by bringing data warehouse, data science, data engineering and operational database workloads together on a single, integrated data platform. Cloudera SDX enables these diverse analytic processes to operate against a shared data catalog that preserves business context, such as security and governance policies and schemas. This common services framework persists even in transient cloud environments and makes it easier for IT personnel to set and enforce policies while enabling business access to self-service analytics.

**Highlights**

- One platform with applications for data science and engineering, data warehousing and operational databases.
- Deploy and manage across Amazon Web Services (AWS), Google Cloud Platform, Microsoft Azure and private networks.
- Support transient clusters and grow and shrink them as needed to avoid overpaying.
- Support permanent clusters for long-running business intelligence (BI) and operational jobs.
- Run your analytic tool of choice against cloud-native object stores like Amazon S3.
Hybrid deployment
Work where and how it’s most convenient, affordable and effective. Cloudera Enterprise can read direct from and write direct to cloud object stores, such as Amazon S3 and Azure Data Lake (ADLS) as well as on-premises storage environments, or Hadoop Distributed File System (HDFS) and Kudu on infrastructure as a service (IaaS). This advantage provides flexibility to work on the data that you want wherever it lives, with zero copies or moves. Cloudera also provides the most popular data warehouse and machine learning engines that can run on virtually any compute resource for ultimate deployment flexibility. Its hybrid control means users can self-service through a platform as a service (PaaS) offering, or opt for more configurability and management through IaaS, a private cloud or on-premises deployments.

Powerful open source
Cloudera develops and validates the best of open source innovations into one seamless and reliable platform. Key features include:

- In-memory data processing. The longest and deepest experience with Apache Spark
- Fast analytic SQL. The lowest latency and best concurrency for BI with Apache Impala
- Updatable analytic storage. The only storage for fast analytics on fast-changing data with Apache Kudu
- Open source leadership. Constant open source development and curation with the most rigorous testing

One platform. Many uses. Designed for your needs.
Cloudera Enterprise is available on a subscription basis in five editions, with each edition designed for specific needs. The Essentials edition provides superior support and advanced management for core Apache Hadoop.

Other editions are designed around potential uses for the platform, including:

- Data science and engineering for programmatic preparation and predictive modeling
- Operational database for online applications and real-time serving
- Data warehouse for BI and SQL analytics

The Enterprise Data Hub provides everything you need to become information-driven, along with complete use of the platform. All editions are available in your environment of choice, whether it’s cloud, on-premises or a hybrid deployment.

For more information
To learn more about IBM® Enterprise Data Analytics with Cloudera, visit the IBM and Cloudera webpage or contact an IBM data management expert.