

cloudera[®] CONNECT



SPLUNK

INDUSTRY

Machine-generated Big Data

WEBSITE

www.splunk.com

COMPANY OVERVIEW

Splunk Inc. (NASDAQ: SPLK) provides the leading software platform for real-time Operational Intelligence. Splunk[®] software and cloud services enable organizations to search, monitor, analyze and visualize machine-generated big data coming from websites, applications, servers, networks, sensors and mobile devices.

PRODUCT OVERVIEW

More than 7,000 enterprises, government agencies, universities and service providers in over 90 countries use Splunk software to deepen business and customer understanding, mitigate cybersecurity risk, prevent fraud, improve service performance and reduce cost. Splunk products include Splunk[®] Enterprise, Hunk[™], Splunk Cloud[™] and premium Splunk Apps.

SOLUTION HIGHLIGHTS

- > Explore, analyze and visualize raw unstructured data in Cloudera Enterprise
- > Simply point Hunk at your Cloudera cluster and start exploring data immediately
- > Archive to Cloudera

Unlock the Business Value of Archived Data with Cloudera and Hunk[™]: Splunk Analytics for Hadoop

Unstructured data, much of it generated by machines or sensors, accounts for more than 90% of data today. Organizations faced with the sheer complexity and scale of this data see the benefits of Hadoop for economical long-term storage, but often struggle to manage that data in Hadoop. Without a flexible, scalable, and secure data management solution, business analysts can miss decision windows or make incomplete decisions based on limited or incomplete data—at great cost to the organization.

Leveraging the Cloudera Enterprise Data Hub and Hunk[™] for Hadoop Archive Business Analytics

The growing volume and complexity of data highlights the fault lines in conventional approaches to information management. Success in an ever-competitive data-driven market requires flexible, massively scalable data management systems that grow with your business at a reasonable cost. The enterprise data hub (EDH), delivered through Cloudera Enterprise, is a transformative active archive solution helping enterprises gain more insight across all their data to make more informed decisions. The Cloudera's enterprise data hub provides one place to economically store all historical data, in any format, at any volume, for as long as needed without costly data movement, enabling you to meet compliance management, security and governance requirements, while delivering data on demand for reporting, exploration, and analysis.

The fully integrated EDH provided by Cloudera constitutes a highly scalable storage and multi-workload processing platform, providing essential production capabilities such as security, resource management, production workload visibility, multi-file format support, and cross-workload optimizations that seamlessly integrate with specialized systems in your existing environment.

Integration of Cloudera EDH and Hunk[™]

Hunk is a full-featured platform for rapidly exploring, analyzing and visualizing data in Hadoop. Based on years of experience building big data products deployed at thousands of Splunk customers, Hunk automatically adds structure and identifies fields of interest at search time to deliver a faster, more interactive experience from the data in your EDH. In Hunk, change perspectives on-the-fly, preview results as MapReduce jobs are running, and govern access with role-based security. The result is you no longer need a science project to get business value from your data in Hadoop.

Hunk natively integrates with the Cloudera Distribution of Apache Hadoop (CDH) and the Cloudera's enterprise data hub through the Apache MapReduce framework. The combination of Hunk and Cloudera allows you to detect patterns and find anomalies across terabytes or petabytes of raw data in the EDH. Splunk's Search Processing Language (SPL[™]), Data Model and Pivot enable rapid data exploration without the need for specialized skills. With Hunk and Cloudera, unlocking the business value of data in Hadoop is faster and easier than you thought possible.

“ I'm super excited about Hunk. Hunk is solving one of the top issues that our customers have—access to the skills and know-how to leverage the data inside of Hadoop. Splunk has a very beautiful user interface that is very easy to learn. So it bridges that gap and makes it very easy to access the data inside of Hadoop. ”

DR. AMR AWADALLAH
CTO, CLOUDERA

Benefits of Cloudera

Stores and Analyzes Any Type of Data

- > Store and analyze huge volumes of structured and unstructured data that were previously impossible or impractical
- > No need to define a data model during ingest
- > Supports multiple, flexible schemas

Massively Scalable

- > Brings compute to the data, so no need for expensive data movement prior to analysis
- > Scales linearly on industry standard x86 hardware

Industry-Leading Management and Support

- > Centralized, end-to-end management through Cloudera Manager, supporting deployment, configuration, monitoring, and issue resolution
- > Makes handling even the largest enterprise clusters simple and efficient
- > Worldwide team of Hadoop experts and project committers working for you

Benefits of Cloudera + Hunk

Unlock the Business Value of Diverse Data in Hadoop

- > **Interactive Search** – Explore data, change perspectives on-the-fly and preview results as MapReduce jobs are running
- > **Drag-and-drop Analytics** – Create data models to describe relationships in the underlying raw data, making it more meaningful and usable; Pivot to quickly generate charts, visualizations and dashboards
- > **Build Big Data Apps on Hadoop** – Quickly develop apps in familiar programming languages

Active Archiving

- > **Inexpensive Batch Storage** – Archive from Splunk Enterprise to Cloudera Enterprise

Achieve Your Active Archive Today

If your organization is seeking to retain deep histories for analysis and is struggling with access time, with complexity of datasets, or with fixed schemas that limit usability, consider the next step in data management—an EDH as an active archive for all your historical data, integrated with Hunk™: Splunk Analytics for Hadoop and NoSQL Data Stores for schema-on-the-fly business analytics, role-based access security and data visualization. The Cloudera EDH makes it technically feasible to quickly and cost-effectively store and process a large and unpredictable volume and variety of data, and Hunk™ dramatically improves the speed and simplicity of big data insights. Cloudera's enterprise data hub and Hunk provide the perfect complement to gain meaningful insight from your diverse historical data in Hadoop.

Cloudera Enterprise Data Hub

- > High-volume, cost-effective storage for structured and unstructured data
- > Repository for historical clickstream data, image data, binary formats, packet data, videos, audio, and more
- > Store all this raw data in its original fidelity
- > World's most popular open source distribution powered by Apache Hadoop
- > Cloudera Manager is the industry's first and most sophisticated management application for Apache Hadoop and EDH
- > Powerful and scalable data platform with the flexibility to run a variety of workloads
- > Distributed compute power to analyze data in cases like day/day, month/month or year/year and write results back to file

Splunk + Cloudera

- > Splunk Enterprise provides managed forwarders and high value for real-time data feeds such as network, security, app performance and business metrics
- > Splunk Enterprise customers can archive older historical data to Cloudera's EDH
- > Hunk allows you to rapidly explore, analyze and visualize data in an EDH, so you can create dashboards and share reports.
- > Preview results in Hunk and interactively search across one or more Cloudera clusters with schema-on-the-fly
- > Data Model, Pivot and role-based security provide self-service analytics
- > Download Hunk for free for a 60-day trial to connect to any size Cloudera Hadoop cluster

About Splunk Inc.

Splunk Inc. (NASDAQ: SPLK) provides the leading software platform for real-time Operational Intelligence. Splunk® software and cloud services enable organizations to search, monitor, analyze and visualize machine-generated big data coming from websites, applications, servers, networks, sensors and mobile devices. More than 7,000 enterprises, government agencies, universities and service providers in over 90 countries use Splunk software to deepen business and customer understanding, mitigate cybersecurity risk, prevent fraud, improve service performance and reduce cost. Splunk products include Splunk® Enterprise, Splunk Cloud™, Hunk™ and premium Splunk Apps. To learn more, please visit www.splunk.com/company.

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 20,000 individuals trained, Cloudera is a leading educator of data professionals, offering the industry's broadest array of Hadoop training and certification programs. Cloudera works with over 900 hardware, software and services partners to meet customers' big data goals. Leading organizations in every industry run Cloudera in production, including finance, telecommunications, retail, internet, utilities, oil and gas, healthcare, biopharmaceuticals, networking and media, plus top public sector organizations globally. www.cloudera.com