WHY OPTIMIZING SPLUNK IS CRITICAL FOR OIL AND GAS COMPANIES

In an incredibly competitive and volatile environment, O&G companies must adopt a more holistic approach to large scale machine generated data that not only provides new analytics use cases, but also enhances business agility, increases the value of data insights, and lowers associated processing costs.

Logging Modernization Can Help Oil and Gas Businesses Extract High Value Data Insights

One of the biggest challenges facing the Oil and Gas (O&G) industry is managing the sheer complexity and scale of machine generated data across diverse, interconnected parts of the business. The vast amounts of data that O&G companies manage comes from an array of sources in the field, including Geosciences, Land, Drilling, Reservoir Engineering, Completions, Operations, and Marketing. With unique interrelated dependencies, the ability to manage and analyze large-scale log data sets across all lines of business is critical to achieving long-term success.

Many O&G companies rely on Splunk, an industry leader in the IT Operations Management (ITOM) and Security Information and Event Management (SIEM) fields, to process their log data for analytics. However, as we move into a world of greater data complexity driven by rapidly evolving new tech like the Internet of Things (IoT) and 5G, O&G companies face a number of key challenges around extracting the highest value from large scale machine generated log data.

- Shifting towards true log forward compatibility and reducing data noise
- Creating a Splunk optimization strategy to improve data management and enhance analytics, as well as reduce costs
- Enhancing data security, identifying key logging data, and prioritizing core activities
- Driving stream processing and analytics on data-in-motion

60%

Unused critical data

Up to more than half of your data can go unused when collected and stored in single data silos, limiting operational efficiency and the sharing of information across lines of business.

Optimizing Splunk For More Nuanced Data Insights That Can Help Drive Business Success

Splunk is a powerful tool trusted by a diverse range of businesses around the world to curate and review log data from a variety of sources, but its limited data filtering and routing capabilities at the data source can create operational weaknesses for O&G companies. Splunk components include a series of Forwarders, Indexers, and Search Heads which means that all data, regardless of value, is ingested, indexed, and stored in a single data silo. Much of the data from high-frequency logs within the Oil and Gas sector are also incredibly noisy, with complexity increased by duplicates and surges caused by DNS lookups. Large scale dumps of noisy data going into and out of Splunk result in additional enterprise-wide data integration costs and provide no business critical insights.

Optimizing Splunk workflows can help achieve true log forward compatibility and reduce data noise. By introducing a combination of Apache NiFi (part of Cloudera DataFlow) and Cloudera Data Platform (CDP), firms can filter and refine logs before they are sent to Splunk, achieving a significant reduction in noise. In this scenario, NiFi first receives, compresses, filters, and transforms data based on content and/or attributes before sending it to Splunk and the Cloudera Data Platform, storing only the data required. CDP can then run deeper analysis and machine learning algorithms to provide higher value insights.

40%

Reduction in hardware costs

Achieve less noisy data faster and shared analytics with greater nuance and increased collaboration across different lines of business.

CDF is capable of processing several million transactions per second, enabling businesses to have greater control of data generated in motion and at rest without sacrificing speed and agility to secure data workflows.

About Cloudera

At Cloudera, we believe that data can make what is impossible today, possible tomorrow. We empower people to transform complex data into clear and actionable insights. Cloudera delivers an enterprise data cloud for any data, anywhere, from the Edge to AI. Powered by the relentless innovation of the open source community, Cloudera advances digital transformation for the world's largest enterprises.

Learn more at cloudera.com

Improving Data Management and Analytics With Cloudera DataFlow

One way of moving towards a more agile strategy that supports key operational outcomes that serve to significantly improve business efficiency is replacing Splunk Forwarders within the organization's digital infrastructure with MiNiFi agents, a light-weight version of Apache NiFi. Doing so means that data can be processed and filtered at the source, so complexity, load, and costs are reduced, dataflow management is improved, and data provenance recording is significantly enhanced. With the data fed into the Cloudera DataFlow (CDF) platform, data can then be managed and configured faster and more efficiently. By leveraging infrastructure that utilizes MiNiFi agents and the broader CDF platform, businesses can analyze even larger data sets quickly and easily, providing key stakeholders with actionable intelligence to drive real-time decision-making.

When businesses optimize Splunk workflows by leveraging the power of CDF they can:

- Process real-time data streaming at high volume and high scale
- Drive stream processing and analytics on log streams
- Track data provenance and lineage of streaming data
- Manage and monitor edge applications and log streaming sources

Enhancing Data Security From Edge to Enterprise

Companies within the Oil & Gas sector face a further, additional complication in managing their high value data in that cybersecurity and SecOps is a critical part of any optimization strategy. CDF supports high volume data collection at the edge, including devices using MiNiFi, and helps enhance data collection so businesses can also stream data more securely from the edge. With Apache Ranger integration built-in, CDF offers seamless security across all data from a wide range of sources including, IoT devices, enterprise applications, partner systems, and edge applications generating real-time streaming data.

Cloudera is Driving Logging Modernization for Improved Data Analytics

At Cloudera we specialize in supporting businesses with complex data flows and large volumes to acquire data more quickly and securely while prioritizing high value data for analysis with clear traceability. We understand how to manage data complexity by optimizing Splunk data workflows, so that key decision-makers get the analytics and insights needed to make real-time decisions that enhance business agility and operational efficiency.

Our Cloudera Data Platform and Cloudera DataFlow solutions are part of an intelligent, efficient logging modernization strategy to drive increased data visibility and solve even the most demanding business use cases. With real-time stream processing from the edge, data warehousing, data science, and iterative machine learning across shared data at scale, companies within the O&G sector can optimize their Splunk workflows to deliver significant cost savings and enhance their ability to extract value from data across different lines of business.

Click here for more information about how Cloudera can help improve your logging modernization strategy to gain greater data insights.

Cloudera, Inc. 395 Page Mill Road Palo Alto, CA 94306 USA cloudera.com