White Paper

Splunk Optimization by Imperva

thalestct.com

THALES
Building a future we can all trust

THALES

How Imperva Data Security Fabric reduces Splunk operational costs and improves data-centric security insights

Executive Summary

Imperva Data Security Fabric adds data security context into Splunk analytics and can reduce Splunk ingestion volume (and costs) by 70-95%

Many enterprises use Splunk as their primary security analytics engine. Splunk analytics give security teams a summarized view of machine data from network, data center or IT environments. Splunk is also used to retain log records for data repositories. Storing these logs enables compliance with regulatory record retention requirements. These logs can represent a significant amount of the data indexed into Splunk. Beyond log management, and the ability to store event records, they are unable to gain comprehensive visibility and an understanding of what is happening with their data.

To lower their Splunk costs, and get greater insight into what is actually happening with the data people access, many organizations implement Imperva Data Security Fabric (DSF) in concert with Splunk. Imperva Data Security Fabric acts as a data access log pre-processor that compresses the raw data, reducing the size (and cost) of what Splunk ingests. In addition, Imperva DSF enriches the records with more detail into what data was accessed, which will help the SOC staff accelerate security incident response and forensics.

Imperva Data Security Fabric also provides complimentary analytic functionality. Imperva risk-based analytics combine anomaly detection, with data sensitivity context regarding what is being accessed, and apply purpose-built active exploit detection capabilities developed by Imperva Labs. Imperva analytics look for - and find - threats that anomaly-based analytics alone do not detect. Imperva DSF then provides a simple incident report that makes incident details and severity easy to understand.

Findings

- Reduce Splunk database activity indexing and telemetry volume by 70-95%
- Enable Splunk access to database log records without Splunk ingestion costs
- Convert raw log information into actionable security risk insights
- Gain active exploit detection to quickly respond to immediate data threats

Reducing Splunk Ingestion Costs for Data Repository Logs

Imperva Data Security Fabric can reduce Splunk costs significantly. While exact cost savings will vary by organization, reviewing costs from a data telemetry volume perspective provides insight into potential savings.

Data Volume is the key

Splunk is commonly priced by the volume of data ingested into the platform (GB/day). Many organizations deploy an everything-to-Splunk strategy to get all data in one location. Imperva Data Security Fabric is designed to normalize, compress, and filter raw activity logs, resulting in 5-30% of information being sent to Splunk. The table below presents conservative (low), probable (average), and optimistic (but achievable) savings for a typical organization.

		Direct to Splunk		Pre-processed by Imperva		
Original Environment						
A	Average gigabyte (GB) per day ingestion	2,600GB				
Telemetry Savings				**Low**	**Average**	**High**
			Percent Change	-70%	-90%	-95%
			Daily logs to Splunk	780 GB	260GB	130GB
Cost Savings						
B	Total average annual Splunk ingestion cost	$835,210		$289,000	$136,000	$80,000
C	Overall Splunk analyst and developer FTE improvements		Percent Change	-10%	-20%	-30%
D	Average annual salary for Splunk analyst and developer FTEs	$90,000		$90,000	$90,000	$90,000
E	Number of analysts and developers FTEs	3		2.7	2.4	2.1
F	Average annual salary for Splunk operations FTEs	$65,000		$65,000	$65,000	$65,000
G	Number of Splunk operations FTEs	2		1.8	1.6	1.4
Annual Splunk costs (B+(D*E)+(F*G))		$1,235,210		**$397,000**	**$222,000**	**$144,000**

Ingestion cost based on annual term license and index volume of $0.88 per GB ($0.88 * 2,600 GB/day * 365 days = $835,210)

In this example Imperva DSF reduced Splunk ingestion by an average of 90% per day, from 2,600GB to 260GB. Annual Splunk costs reduced by 82%, from $1,235,210 to $222,000.

Alternatively keep the data in Imperva DSF

Imperva even provides an option to skip ingestion into Splunk altogether. A Splunk plug-in for Data Security Fabric allows Splunk native jobs to run against data sets stored in Imperva Data Security Fabric through virtual indexing. Storing the logs on Imperva rather than Splunk makes data ingestion unnecessary, and can also reduce licensing costs, yet preserves full functionality for Splunk users.

Better Together: Splunk and Imperva Analytics

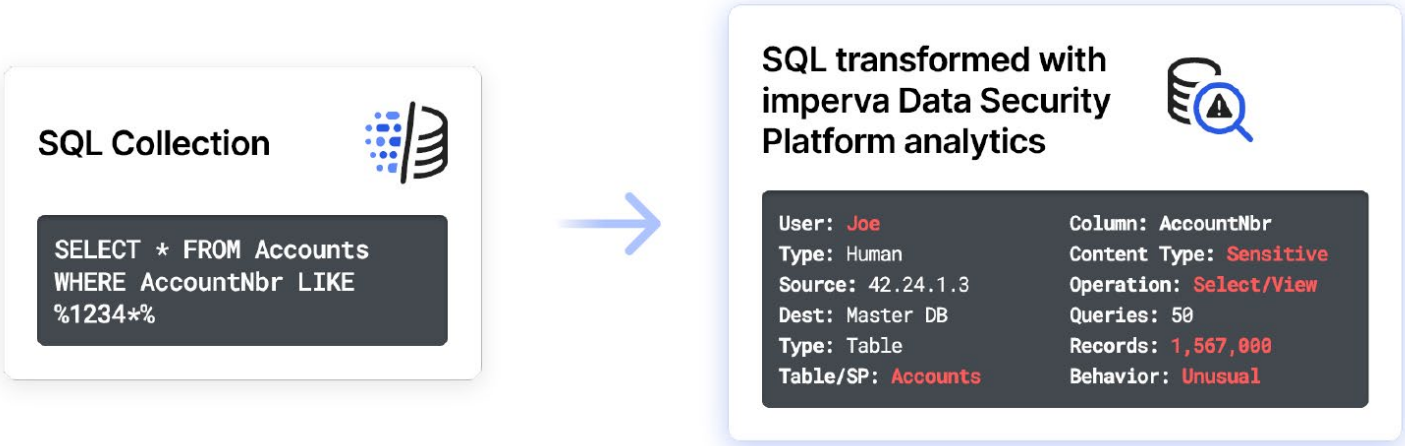
Many Splunk customers also use Imperva Data Security Fabric and gain significant benefits from the complementary analytic functionality it provides. Imperva analytics look for and find threats that Splunk analytics alone do not detect, Also, Imperva DSF provides a comprehensive summary of incident details in a simple incident report that make the incident and its severity easy to instantly understand, even when a security analyst does not have background knowledge of the data set involved.

Out of the box, Splunk consumes and correlates log telemetry from nearly every digital application. Organizations rely on Splunk for monitoring, analyzing, and searching machine-generated information. It does not, however, help security teams understand when an event might represent a critical data security issue, versus an anomaly that actually is low risk. While Splunk can consume logs from nearly every data repository its analytics lack data access context. To gain this ability requires a significant custom development effort but this is already available in Imperva DSF for dozens and dozens of data sources.

Understanding the data accessed

Whether data access is risky or not depends on the type of data. If a user logs in from a different system than normal, or at an unusual time, but does not access any sensitive data types then their activity presents no data risk. If Splunk creates an alert for this anomalous event. it would actually be a false positive.

The key to what data was accessed is in the SQL query output. Splunk analytics does not natively have the ability to parse SQL queries. Imperva DSF does decipher the SQL query so it knows exactly what data was accessed and what the user did with the data.



Advanced Threat Detection

Imperva DSF analytics continuously analyze data access activity and can automatically determine if a data access event violates compliance, a security policy, or varies from normal data access activity by peers in a user role. In addition, Imperva data risk analytics utilize purpose-built detection capabilities developed by Imperva Labs to spot attack exploits or suspicious activity even if the behavior attempts to be evasive. Imperva DSF goes beyond the anomaly detection techniques of other security products. Imperva's purpose-built algorithms are capable of identifying signs of malicious insider behavior such as privilege escalation, data exfiltration, or exhibited signs of compromised user account activity that other security tools miss.

Summary of Imperva Optimizations

Imperva filters, compresses, and enhances raw data repository log records, reducing Splunk operational costs. Imperva analytics complements Splunk analytics by providing greater incident detailed and actionable database event insights.
Splunk threat intelligence and analytics capabilities lack database security domain knowledge out of the box. Splunk DB Connect is required for custom table import and enrichment and results in a heavily manual DIY solution that can be hard to maintain and often provides inconsistent visibility.

The table below provides a before and after comparison of how Splunk and Imperva work better together.

	Events direct to Splunk	Pre-processed by Imperva
Information Access	Information flows one way	Bidirectional access enables Splunk to run reporting against a virtual index
Event Volume	Raw log data volume is high and all events stored and analyzed; can negatively impact system performance.	Imperva deduplicates and compresses content, stored on Imperva data warehouse with full access
Security	Splunk challenged with data security	Imperva is best of breed data security
Accuracy	Many false positives, SOC alert fatigue	Highly accurate alerts
Log Retention	12 months typically	12-36 months typically
Observability	Low	High
Connectivity	Splunk consumes database logs and. can use DB Connect to import tables, rows, columns for indexing, analysis, and visualization.	Imperva individually identifies different RDBMS that normalizes data retrieval, resulting in consistent visibility and no need to write complex queries or code.
Costs	High log processing and storage costs ($$$$)	Reduced log processing and storage costs ($)

Thousands of customers around the world trust Imperva to protect their data, applications, and websites. With an integrated approach combining edge, application security, and data security, Imperva protects companies through all stages of their digital journey. Imperva DSF is available for sale to the U.S. Federal Government through Thales Trusted Cyber Technologies.

About Thales Trusted Cyber Technologies

Thales Trusted Cyber Technologies, a business area of Thales Defense & Security, Inc., is a trusted, U.S. provider of cybersecurity solutions dedicated to U.S. Government. We protect the government's most vital data from the core to the cloud to the edge with a unified approach to data protection. Our solutions reduce the risks associated with the most critical attack vectors and address the government's most stringent encryption, key management, and access control requirements.

For more information, visit www.thalestct.com

Contact us - For office location and contact information, please visit thalestct.com/contact-us