imperva

WHITEPAPER

Splunk Optimization and Cost Savings using Imperva Data Security Fabric

Executive Overview

Many enterprises use Splunk as the primary repository for native database logging and database activity monitoring (DAM) tools.

These logs can represent a significant amount of the data indexed into Splunk. Storing these logs enables adherence to compliance regulations but due to the lack of security context, are often otherwise unused.

To increase the value of raw logs, organizations implement Imperva Data Security Fabric (DSF) as a database log pre-processor to Splunk. The Imperva solution filters, compresses, stores, and indexes the raw event data and makes it available in Splunk via virtual indices. Security analysts have bidirectional access for interactive data exploration. Threat alerts include rich technical context and detection timelines.

Analytics within Imperva Data Security Fabric bridge user identity, data sensitivity, and attack detection. This approach enables organizations to increase the security value from their Splunk deployment and markedly reduces Splunk ingestion and development costs. Imperva Data Security Fabric includes broad coverage across on-premise, cloud, and hybrid data repositories.

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

Findings

- Convert raw logs into actionable security insight for use within threat event pipelines
- Gain 100% visibility into data repositories across on-premises and cloud
- Improve incident response times with pre-built and custom playbooks
- Leverage virtual indexing for bidirectional security event access without Splunk ingestion
- Reduce Splunk database activity indexing and telemetry volume (and costs) by 70-95%
- Extend data retention periods from 1 year to 3 years

Recommendations

- Evaluate IT security and compliance team use of data repository logs to understand ROI
- Explore 12 months Splunk costs and corporate growth plans to predict cost trajectory

- Evaluate database security pre-processing technologies to reduce costs and add value
- Evaluate data security architectures for areas of consolidation

Splunk and Imperva Data Security Fabric - Better Together

Out of the box, Splunk consumes and correlates log telemetry from nearly every digital application. Organizations rely on Splunk for monitoring, searching, and analyzing machine-generated information, including that from native database logging and database activity monitoring (DAM) tools.

Splunk analytics uses artificial intelligence (AI)/machine learning (ML) to relate seemingly disparate events and drive rich search capabilities across security and system monitoring use cases. As a data security tool, however, this falls short. While the system can consume logs from nearly every data repository, the lack of database context dramatically reduces the effectiveness of Splunk analytics engine. The result is large ingestion costs, added noise, and purposeless metrics.

With database logging direct to Splunk:

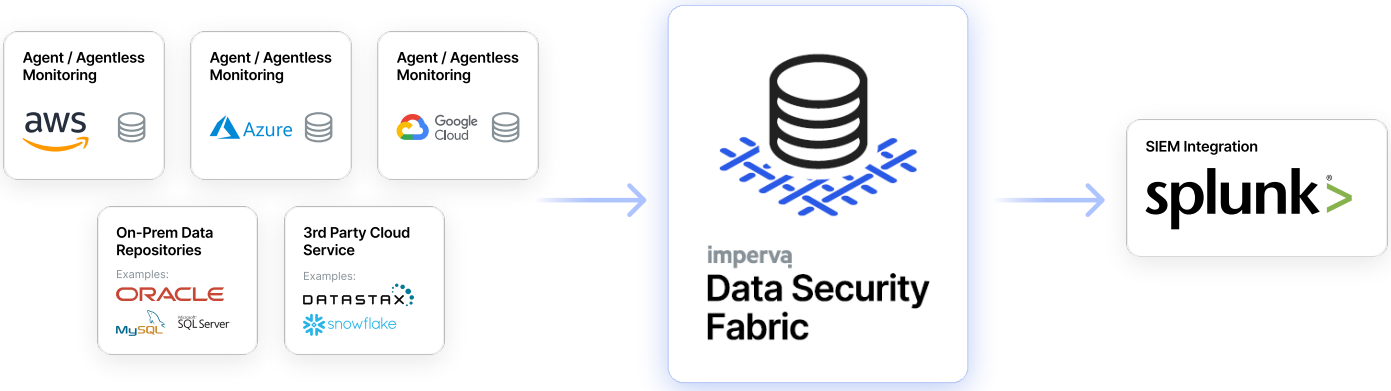
- **Database context is unavailable;** Splunk is challenged by data security and cannot decipher SQL to understand query output, nor relate this information to user behavior and access to understand risk
- **Database threat alerting is reactive**, not predictive
- **Database event storage is flat,** rather than compressed/optimized

Adding Imperva DSF as a pre-processor to Splunk provides the following value:

- **Reduction of incident remediation times.** Enriched security events sent to Splunk in an easy-to-understand format. Make analysts more effective by providing a clear lens into critical events.
- **Reduction of Splunk costs.** Native logs are optimized and stored on the Imperva DSF, dramatically reducing Splunk ingestion volume and costs.

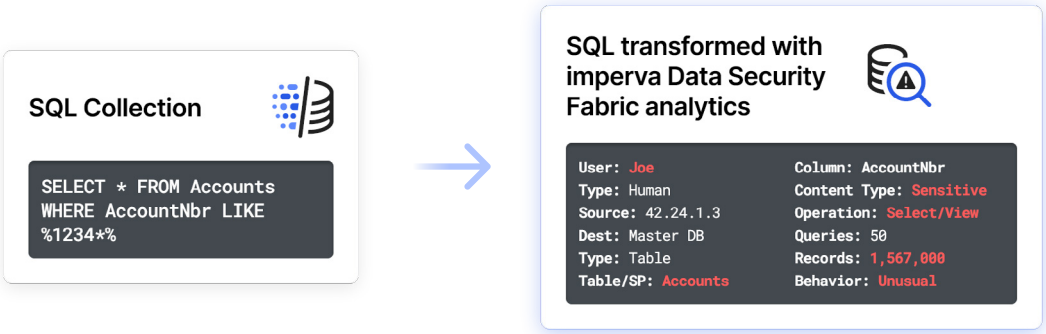
How it works

Native logs from on-premise and cloud data repositories are forwarded to Imperva Data Security Fabric for pre-processing. Imperva enriches the information with context into users, access, and threats. Only the intelligence extracted from raw data is forwarded to Splunk, allowing Splunk to remain effective at incidence response and enterprise-level event correlation.



The Imperva DSF “any database, any location, any scale” design ensures compatibility with your current and future architectures. Bidirectional integration lets Splunk users access all raw data stored in Imperva using the Splunk UI.

Splunk will normalize the traffic into a common model similar to Imperva however Imperva SQL parsing results in richer and more actionable information.



Approach Comparison

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 but limited RDBMS support results in a heavily manual DIY solution with inconsistent visibility.

Imperva DSF filters, enhances, and predicts, providing Splunk targeted and actionable database event information. Imperva’s extensive RDBMS support eliminates the need to build complex queries or code.

	Events direct to Splunk	Events to Imperva Data Security Fabric for pre-processing
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 DSF 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 ($)

Splunk Cost Reduction

Adding Imperva DSF 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.

Splunk is commonly priced by volume of data ingested into the platform (GB/day). Many organizations deploy an everything-to-Splunk strategy to get all data in one location. Database activity monitoring creates very large amounts of raw data, but only a small percentage of that traffic is relevant from a security and compliance perspective.

Imperva Data Security Fabric is designed to normalize, compress and filter raw activity logs, resulting in 5-10% 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		To Imperva DSF for pre-processing		
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)

Imperva's solution 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.

Splunk native jobs run against the data sets as a virtual index by Imperva. This reduces licensing costs, simplifies the SOC and makes events more actionable. If analysts need to dig deeper they have seamless access to the raw information that may be relevant to remediation.

"[Imperva DSF] allows us to continue to run the same searches,alerts, and dashboards while actually storing the data [in Imperva DSF]. We get all the power and usefulness of our SIEM without the cost."

- 2021, Global payment solutions company in an Imperva case study