5 Ways CISOs Can Use Selective Retrieval to Optimize Data Lakes

Data lakes have evolved. Once treated as passive storage archives, they're now becoming active components of enterprise risk management. The driver? Selective retrieval — the ability to park large data volumes in cold storage and later retrieve targeted slices for forensic or compliance needs.

This shift matters. According to 2025 data from Cybersecurity Insights Group, 73% of enterprises report that SIEM ingestion costs are limiting their real-time analysis capacity. At the same time, 62% of security leaders say forensic readiness is a top priority for compliance and post-incident response.

Here are five ways selective retrieval strategies are helping CISOs address both challenges without sacrificing visibility.

5 Selective Retrieval Strategies in Action

1. Focus Real-Time Analytics on High-Signal Data

Not every log needs immediate analysis. Modern pipelines now allow teams to route high-signal logs, like authentication failures or denied firewall requests, directly to SIEM tools for real-time analysis. Lower-signal data, such as successful logins or benign file accesses, can be sent directly to a central data lake.

The advantage is cost control without loss of coverage. When investigations require it, teams can retrieve dormant logs selectively, analyzing only the relevant portions.

In 2025, 58% of organizations report shifting non-critical data to cold storage to optimize real-time detection capacity. Selective retrieval operationalizes that shift without losing forensic traceability.

2. Improve Forensic Readiness While Reducing Always-On Costs

Maintaining hot access to all historical data is financially impractical. Enterprises using selective retrieval can store 12 to 24 months of logs affordably in cold storage and retrieve subsets for investigations as needed.

One European engineering firm reported a **45% reduction in SIEM licensing fees** after adopting selective retrieval, while simultaneously extending their log retention period from six months to two years.

This approach separates storage from processing, giving security teams access to necessary data without ongoing analysis overhead.

3. Balance Security Priorities with Compliance Mandates

Compliance teams often require long-term data retention, but this should not dictate analytics workflows. By tagging data on ingest, teams can store all logs for audit readiness while restricting active analytics to data supporting detection priorities.

In 2025, **69% of financial services** organizations reported using metadata tagging to manage regulatory and operational log requirements separately. This ensures compliance reporting needs are met without compromising security team efficiency.

4. Retain Full Visibility Across Noisy Data Sources

Firewall logs are a prime example of high-volume, low-actionability data. Many teams historically chose between retaining denied requests or successful connections due to storage and processing constraints.

Selective retrieval removes that tradeoff. Both denied and accepted traffic logs can be stored without analysis, preserving full visibility. When a post-incident review demands it, analysts can retrieve only the relevant data window to answer specific questions.

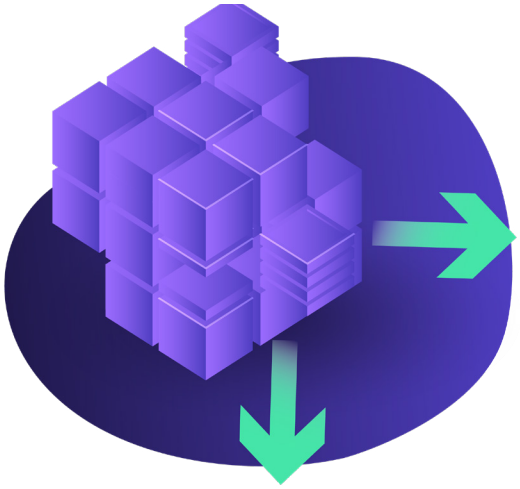
In regulated sectors like healthcare and manufacturing, **64% of security teams** now cite selective retrieval as essential for supporting internal investigations without expanding real-time infrastructure.

5. Treat the Data Lake as an Operational Asset

Data lakes are no longer passive archives. Modern architectures support preview-before-ingest capabilities and conditional retrieval triggers, transforming data lakes into active security resources.

Instead of analyzing everything or ignoring large datasets, teams can treat stored logs as a reservoir to be accessed precisely when needed — whether for compliance audits, insider threat detection, or incident response.

The result is not just cost savings. It is operational flexibility. In 2025, 71% of CISOs surveyed by TechTarget acknowledged that selective retrieval had directly improved their team's investigative efficiency.



Looking Ahead

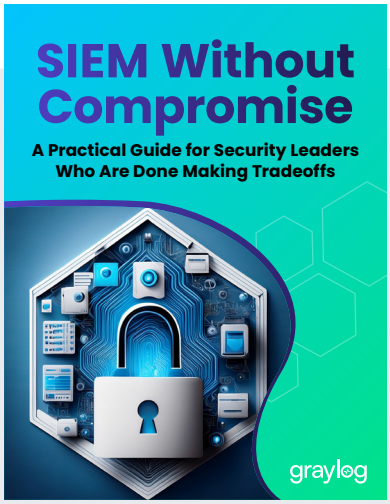
Selective retrieval represents a practical shift in how CISOs manage data growth, security visibility, and compliance obligations. By separating storage from analysis, security leaders can cover more risk scenarios without overextending budgets or infrastructure.

As data volumes continue rising, adopting selective retrieval is becoming less of a niche strategy and more of an operational standard.

Graylog supports these strategies with flexible ingestion, metadata tagging, and retrieval pipelines, enabling CISOs to protect, retain, and investigate without adding unnecessary drag to daily detection operations.

Looking for guidance on how to get started? Read: "SIEM Without Compromise — A Practical Guide for Security Leaders Who are Done Making Tradeoffs"

Learn more about Graylog Security ›



ABOUT GRAYLOG

Graylog delivers a SIEM that works the way teams actually need: full visibility, real detection, and faster investigations—without blowing the budget. Trusted by 50,000+ organizations worldwide, Graylog helps analysts skip the noise and get to what matters. Automate the heavy lifting. Stay focused on real threats. Learn more at graylog.com.