DataBahn for Sumo Logic

Navigating Compute-based Pricing for Higher ROI and Volume Optimization with DataBahn's Security Data Fabric and Sumo Logic

DATABAHN

Unlocking Sumo Logic's Full Potential Starts with the Pipeline

The shift to cloud-native SIEM solutions like Sumo Logic offers significant benefits to enterprise security teams, such as enhanced visibility, UEBA baselining, and automated responses to quickly identify, investigate, and respond to threats. However, routing, moving, and managing this data across platforms can pose challenges and reduce the full value of the Sumo Logic platform.

Sumo Logic's emphasis on usage-based pricing does not eliminate the extra costs associated with log volume growth for enterprise security teams. High ingestion volumes increase the computing power needed to find relevant data, and the effort involved in preprocessing, normalizing, and routing logs to multiple destinations adds to the workload for data engineering teams. Sumo Logic may also charge more for longer log retention periods, which are necessary for compliance audits, resulting in cost overruns for teams using the platform.

By using DataBahn as an intelligent data pipeline layer, enterprises using Sumo Logic can ensure reliable and seamless log collection and aggregation, data normalization, and effortless orchestration from source to destination with full visibility and improved governance. Together, using Sumo Logic with DataBahn offers a solution that drives accelerated and optimized security operations for enterprises, while reducing cost and effort.

Challenges

While Sumo Logic's usage-based pricing can be attractive to enterprises drowning in high volumes of security data, it poses challenges and unexpected costs too:

High Usage Costs

Engaging with Sumo Logic and using its features to filter and manage raw data requires querying and data operations, increasing compute and overall cost of ownership. Higher volumes of data also lead to alert fatigue, and increase the cost, time, and effort involved in finding potential threats or vulnerabilities from security data.

Rigid Ingestion Costs

Sumo Logic provides a library of about 180 integrations for collecting telemetry data from various sources. However, managing formats, normalization, and routing to other destinations requires extra data engineering work. Telemetry health monitoring is also limited, and the platform's native tools do not offer detailed trace-level visibility, which can increase downtime.

Limited Data Shaping pre-ingest

Many logs enter Sumo Logic "as-is", and require transformation and data wrangling post-ingest, which consumes compute in Sumo Logic and increases costs and manual effort. Passing that data downstream to any other security tool also requires additional data transformation into different formats and data models.

Collecting Data from Unsupported Sources

Log collection also presents a challenge for niche applications. Advanced security teams may develop in-house custom applications and microservices or may need to ingest IoT/OT data. These use cases will require creating custom collectors or parsers and extensive engineering for data transformation to make the data usable.

Optimizing Security Telemetry Retention

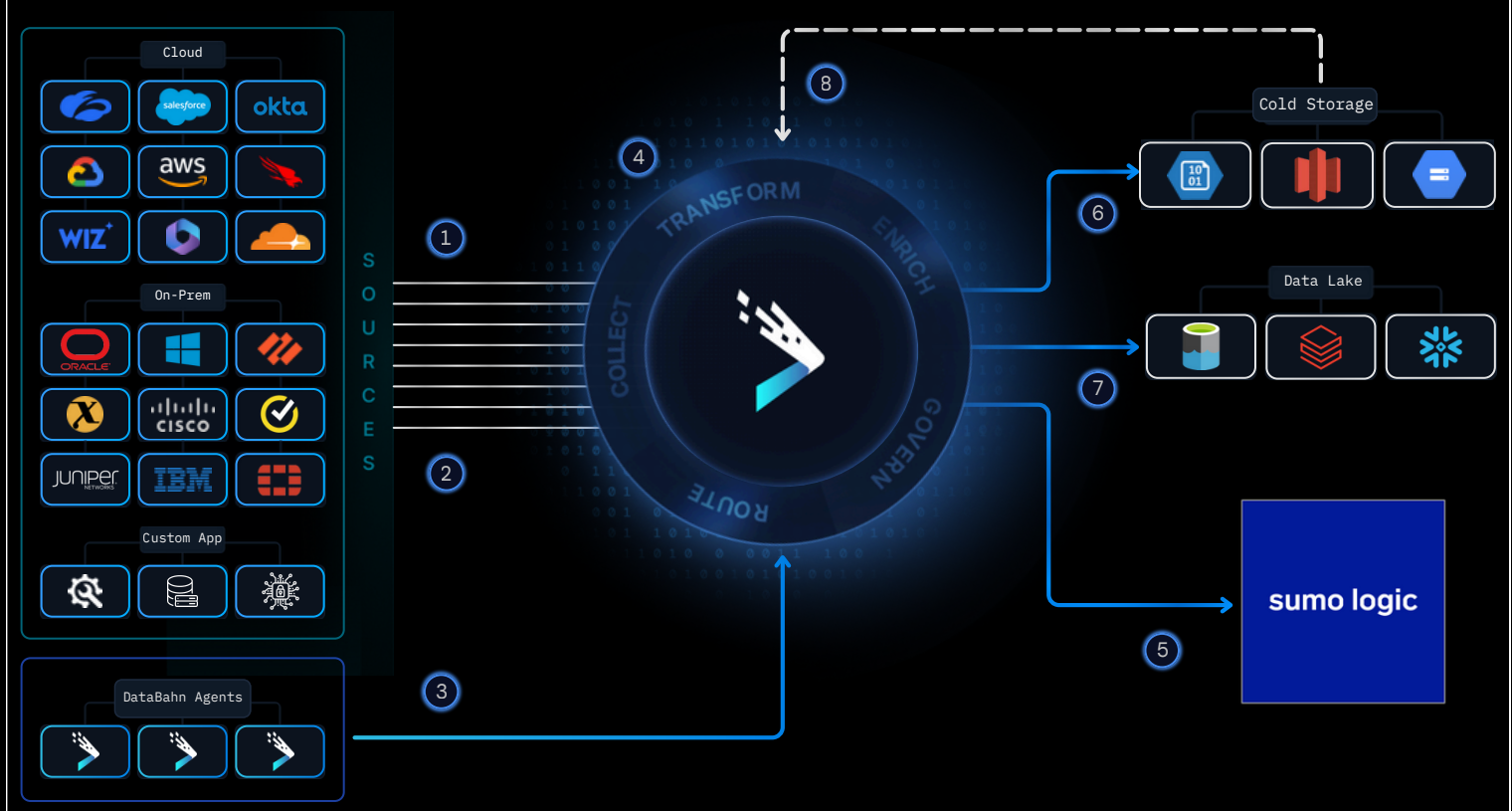
Compliance requires can impose log retention periods, and storing this in Sumo can become prohibitively expensive.

The Solution

DataBahn's purpose-built Smart Edge and Highway solutions can:

- Seamlessly ingest data from a wide range of sources such as on-premise and cloud infrastructure, security products, and applications
- Normalize and enrich this data with internal and external contextual data sources
- Orchestrate the data to extract meaningful insights
- Deliver security-relevant data and insights into your Sumo Logic platform for optimal querying, analytics, and search.

To help visualize how DataBahn supports Sumo Logic end-to-end, the following diagram maps the key stages of the pipeline:



1. **Cloud / SaaS** telemetry ingested directly using DataBahn integrations
2. **OnPrem / Application** telemetry ingested using DataBahn Smart Edge collectors
3. **Endpoint** telemetry can optionally be ingested using XDR agents. DataBahn agents can also be leveraged.
4. **Data Transformation:** Ingested logs are normalized, transformed, enriched, and reduced to retain only security-relevant events before being forwarded to Sumo Logic.
5. **Security-relevant events** are forwarded to Sumo Logic for threat detection, correlation, and analytics.
6. **Data Lake Storage:** Non-security or raw telemetry can be streamed into a data lake for cost-effective analytics, machine learning, or future enrichment workflows
7. **Cold Storage Archival:** Logs not needed for real-time analysis can be routed to cold storage for long-term archival and compliance, with the option to rehydrate them into Sumo Logic when needed.
8. **Data Replay:** Data can be replayed from storage for retrospective analytics, model rehydration, or to enhance investigations.

Smart Edge: Resilient Log Collection at Scale

Smart Edge supports both agentless and agent-based collection, built on a mesh architecture that scales horizontally and vertically. This ensures continuous ingestion, even during bursts, while simplifying deployment across on-prem, cloud, or hybrid environments.

Transform, Enrich, and Optimize in Real Time

As logs enter the pipeline, DataBahn applies a suite of enrichment and optimization steps:

- Enrich with geolocation, business context, CMDB data, or application metadata
- Suppress or reduce noise using prebuilt or custom rules
- Convert logs into lightweight summary signals or metrics, where appropriate

These operations can be performed at the edge, in transit, or centrally, giving teams maximum flexibility.

Smart Routing and Replay
With built-in routing controls, only security-relevant data is sent to Sumo Logic. Additional telemetry can be streamed to alternate destinations, such as cloud-based cold storage (e.g., S3, Azure Blob), data lakes, or observability platforms. Replay functionality enables teams to pull older data back into Sumo Logic or any other tool when needed, without reprocessing through the full pipeline or duplicating ingestion.

Control Plane: Unified Visibility and Management
Fleet management, volume analytics, destination configuration, and billing are managed through a single interface. Schema drift is continuously monitored and flagged proactively, ensuring ingestion compatibility as source formats evolve. This ensures that security teams can scale confidently while maintaining control and compliance across their telemetry stack.

DataBahn: A smarter, more scalable pipeline for Sumo Logic

Smart Data Filtering & Enrichment

DataBahn's Intelligent Data Pipelines allow Security teams to pre-process and filter data before it enters the SIEM. By ensuring that only relevant, high-quality security data makes its way to Sumo Logic, DataBahn reduces the total volume of data being ingested and stored. This saves storage and retention costs, and makes it easier for teams to find and access relevant data to reduce alert fatigue and improve the time and speed of response.

Efficient Data Routing

DataBahn helps streamline the data routing process. Instead of sending raw, unstructured data to Sumo Logic, DataBahn enriches, structures, and routes the most pertinent data to the platform. This means that your SIEM platform only processes data that's directly useful, reducing the overhead of managing unnecessary logs and improving the overall efficiency of your security operations.

Optimized Querying and Analytics

DataBahn's Intelligent Data Pipelines allow Security teams to pre-process and filter data before it enters the SIEM. By ensuring that only relevant, high-quality security data makes its way to Sumo Logic, DataBahn reduces the total volume of data being ingested and stored. This saves storage and retention costs, and makes it easier for teams to find and access relevant data to reduce alert fatigue and improve the time and speed of response.

Reduced Data Management Effort and Costs

With DataBahn handling the filtering, structuring, and enriching of data upfront, you reduce the amount of data that needs to be transferred out of Sumo Logic for additional processing or analysis. This reduces the effort and cost of data wrangling, management, retention, and routing downstream to different tools and systems for a smarter and more efficient data lifecycle.

Scalability and Flexibility for Complex Environments

By using DataBahn for log aggregation and data engineering, enterprises can manage data flows from more complex environments at a significantly lower effort and cost. Enterprises looking to manage hybrid environments with on-prem footprints, complex use cases such as managing multiple SIEM instances and transactional data, etc. can now access Sumo Logic's features by decoupling data ingestion and management from Sumo Logic.

Why DataBahn + Sumo Logic Are Better Together

By integrating DataBahn with Sumo Logic, security teams can lower their total cost of ownership while maintaining full control over data management, ensuring the flexibility to adapt to new requirements and expand their security operations without being encumbered by unnecessary complexity or costs. This integrated solution ensures that organizations can maximize the ROI of their Sumo Logic SIEM deployments while benefiting from streamlined data processing and enhanced operational efficiency.

DataBahn helps with more than just cost reduction:

- **Flexible Data Routing**: DataBahn's intelligent security data pipelines enable flexible routing of security data, allowing organizations to send only relevant data to Sumo Logic to reduce alert fatigue and optimize security threat hunting and investigations
- **Simplified Data Ingestion**: With DataBahn's Smart Edge, security teams can eliminate the complexities of managing cloud log ingestion infrastructure and focus on actionable insights instead of data management
- **Streamlined Integration**: DataBahn provides easy integrations with a wide range of data sources, and automates the parsing and ingestion process for custom apps and microservices with its Agentic AI product Cruz. This ensures that Sumo Logic receives the highest quality data for security analytics and threat detection.

By combining DataBahn's intelligent security data pipelines with Sumo Logic's cloud-native SIEM platform, organizations gain a powerful, cost-effective, and scalable security operations solution that delivers actionable insights, reduces operational complexity, and ensures security teams can stay ahead of evolving threats.

DataBahn.ai is a leader in AI-driven Data Pipeline Management and Data Fabric solutions, helping organizations transform their data operations with innovative engineering and advanced analytics technologies. With its Data Fabric platform and cutting-edge AI capabilities, DataBahn.ai is committed to empowering organizations to harness the full potential of their data for a smarter, more connected future.

Learn more at databahn.ai

DATABAHN