>CASE STUDY_

Cribl Stream Powers Data Optimization for SecureCoders

HIGHLIGHTS

- Replaced difficult-to-manage Syslog servers, centralizing log management
- Achieved 64% data reduction on major log sources through serialization
- Further 30% reduction through tokenization of repetitive log data
- Developed tooling to keep lookups in sync between Cribl and Splunk software
- Implemented summarization and S3 offloading for Cisco ASA logs
- Realized approximately $700k in Splunk software license and infrastructure savings

SecureCoders

SecureCoders is a boutique cybersecurity consulting firm with more than 64 active clients. One major customer, a large and rapidly growing global pharmaceutical company, used Splunk software for years to manage its extensive and complex global data environment. They faced difficulties in configuration management, data modification, and ensuring correct host information. Cribl Stream and Cribl Search helped them overcome these challenges while realizing data optimization, capacity management, and cost-saving benefits.

Streamlining Log Management

Traditionally relying on Splunk Forwarders and Syslog network logs for data collection, the SecureCoders team found that making sense of the client's vast amounts of data from various sources was a significant challenge. Around three years ago, at the client's request, they set up Cribl Stream on-premises in a VMware cluster at the client's corporate headquarters for centralized data management.

> **"We transitioned almost all of Syslog data into the Cribl environment and immediately found that data management was far easier by consolidating previously scattered systems and simplifying configuration tasks. The transition significantly reduced the complexity of log management and enhanced our overall operational efficiency."**
>
> **—Justin Furniss**, CEO of SecureCoders

However, after realizing these benefits, SecureCoders realized they could do even more with Cribl Stream. They found that the platform paid for itself in as little as 6 months.

Serializing Data to Reduce Splunk Software Costs

SecureCoders would regularly receive requests from the client to onboard new data sources. This resulted in a constant challenge in balancing the available Splunk software license with incoming data. Previously, achieving this balance would force the SecureCoders team to delete data, but with Cribl, they could optimize it instead.

> **"We realized we could use Cribl Stream to serialize, summarize, and offload data to S3 and filter unnecessary logs before ingestion, significantly reducing the strain on Splunk software. That's when we really started to understand how much we could do with Cribl."**
>
> **—Justin Furniss**, CEO of SecureCoders

This realization inspired Secure Coders to optimize their Splunk software license utilization. As head of the project, Justin Furniss focused strategically on the client's larger log sources from external SaaS solutions like Zscaler. After analyzing these logs, he identified valuable opportunities for data refinement and decided to use Cribl to serialize them to CSV.

> **"We saw massive success overnight. We were saving hundreds of gigabytes worth of Splunk software license, which is massive when you consider not just the cost of the license but the overhead for the infrastructure that manages that data. We ultimately achieved a 64% reduction on one of the client's largest log sources."**
>
> **—Justin Furniss**, CEO of SecureCoders

By serializing other major log sources, SecureCoders realized around $700,000 of savings on their Splunk software license.

Tokenizing Inefficient Log Sources

The success of the serialization efforts then prompted a data tokenization initiative. Some of the client's largest log sources were highly repetitive. A web proxy log, for example, would contain a vast number of repetitive classifications, so SecureCoders created lookup tables shared between Cribl and Splunk software to replace repeated values with compact tokens. This resulted in an additional 30% data reduction on top of the 64% savings achieved through serialization.

> **"We developed dashboards to identify redundancy patterns across source types, calculate data savings, and prioritize tokenization efforts. Over time, we also built and are planning to open-source tools to ensure these lookups stay synchronized between Cribl and Splunk software software. For example, when a new large offender like a user-agent string emerges, our tooling programmatically updates lookups via Cribl APIs, keeping the process largely automated."**

Summarizing and Offloading Data to S3

SecureCoders then explored summarizing and offloading Cisco ASA logs using Cribl. These logs often contain repetitive data, such as repeated "allow" entries for the same source and destination IP and ports.

To optimize this data, SecureCoders used Cribl to aggregate logs by minute, only sending the first raw log to Splunk software and metadata fields that capture aggregation details. If necessary, they offload full raw logs to S3 for long-term storage so they can be replayed.

> **"Currently, this approach is used for logs we don't expect to need frequently, but it provides flexibility for handling sudden log volume increases or detailed investigations. We've documented the process and tools and plan to expand these strategies next year, shifting from proof of concept to targeting larger log sources."**
>
> **—Justin Furniss**, CEO of SecureCoders

Onboarding New Data Sources

Implementing Cribl Stream has also allowed SecureCoders to onboard new data sources they previously couldn't. The Azure AD logs they capture from Event Hub are verbose and contain duplications. Previously, they relied on Splunk's Technology Add-ons (TAs) to pull the data directly, but this approach provided minimal visibility and control, making it difficult to manage or optimize the data.

> **"With Cribl, we've built advanced pipelines to summarize and deduplicate these logs. For example, we can identify that 20% of the incoming logs are near-duplicates with minor differences, combine them, and remove the excess noise - something impossible with the Splunk TA."**
>
> **—Justin Furniss**, CEO of SecureCoders

Cribl also simplifies ingestion for SaaS products by letting SecureCoders selectively choose what data to collect rather than pulling everything. Additionally, Cribl enables SecureCoders to route new data sources to a QA environment for testing. This allows them to analyze the data, refine pipelines, and decide whether to onboard it into production.

> **"This flexibility has drastically improved our ability to evaluate and selectively onboard new data sources, making the process faster, more efficient, and cost-effective."**
>
> **—Justin Furniss**, CEO of SecureCoders

Expanding Use Cases

Looking ahead, SecureCoders has plans to implement Cribl Stream in the client's industrial environment. They hope that, by expanding Cribl into this environment, they will modernize and optimize it, thus improving scalability and efficiency while reducing the reliance on outdated infrastructure.

> **"We're really excited about this because, at the moment, the environment contains some pretty old hardware; it's still running Syslog-ng, so there's a lot of duct tape and bubblegum holding it together. For example, we distribute workloads across multiple Syslog-ng processes, but the hardware is already at its limits trying to manage the high volume of UDP logs coming into a single IP address."**
>
> **—Justin Furniss**, CEO of SecureCoders

To discover how Cribl can optimize data and reduce infrastructure and licensing costs for your organization, schedule a custom demo today.

TL;DR

- **Secure Coders implemented Cribl Stream to centralize log management and optimize data for a global pharmaceutical client**
- **Achieved 64% data reduction through serialization and an additional 30% reduction via tokenization of repetitive logs**
- **Saved approximately $700,000 on Splunk software licenses and infrastructure costs**
- **Built tools to synchronize lookups between Cribl and Splunk software for ongoing data optimization**
- **Utilized Cribl to summarize and offload Cisco ASA logs to S3, enhancing flexibility and scalability**
- **Streamlined onboarding of new data sources, improving efficiency and reducing costs**
- **Plan to expand Cribl deployment to the industrial side of the client's operations for further optimization**

ABOUT CRIBL

Cribl, the Data Engine for IT and Security, empowers organizations to transform their data strategy. Powered by a data processing engine purpose-built for IT and Security, Cribl's product suite is a vendor-agnostic data management solution capable of collecting data from any source, processing billions of events per second, automatically routing data for optimized storage, and analyzing any data, at any time, in any location. With Cribl, IT and Security teams have the choice, control, and flexibility required to adapt to their ever-changing data needs. Cribl's offerings–Stream, Edge, Search, and Lake–are available either as discrete products or as a holistic solution.

Learn more: www.cribl.io | Join us: Slack community | Follow us: LinkedIn and X (formerly Twitter)

CS-0034-EN-2-0225