CHAOSSEARCH

Achieving the Security Data Lake

Harnessing log data for analytics has never been more important. IT Operations, DevOps, and SecOps teams need to gain the insights within their organizations' ever-growing volumes of log data to become more robust, resilient, and responsive.

CONTENTS

INTRODUCTION

SecOps teams are in a state of continual analysis—analyzing new threats in the global landscape, analyzing specific vulnerabilities in their environment, and analyzing various alerts to determine benign or malicious intent.

A platform that centralizes data and allows for analysis, queries, visualization, and reports is crucial to a well-functioning, high-performing SecOps team. Log data is the lifeblood of the cyber defense operation.

Every transaction and event across an organization's IT landscape is captured, consolidated, and stored. These logs become a "single source of truth." SecOps teams can scan this historical data to see what has happened, understand the current situation, and hypothesize what may happen.

CRUCIAL FACTORS IN USING A LOG ANALYTICS PLATFORM FOR SECURITY

In conducting the various analyses that allow them to keep their corporate assets safe, SecOps teams always want more data rather than less. Provided they have the means to quickly and efficiently analyze it, having access to more data, from more sources (including long-term historical data) allows SecOps teams to be more efficient. They can better find and block threats, stop hacks in their tracks, mitigate the damage of threats, and investigate breaches that did occur to find and fix the root cause problems.

As SecOps teams evaluate centralized analytics platforms, they assess scalability in a few dimensions:

Scale Up: how much data can be ingested and processed each day? As data comes flooding in from multiple sources, can the platform easily be scaled up to ingest the data without creating multiple repositories and sharding databases?

Scale-Out: what are the total capacity limits of your repository? Will you need to pay for more? For cybersecurity, several years' worth of data may be necessary.

As we survey SecOps teams today, we find that while they all use analytics platforms in their overall operation, their use is suboptimal.

Working with analytics platforms limited in scale, SecOps teams are forced to ingest less data from fewer sources and retain data for shorter periods.

A common by-product of the inherent scale limitations is to create multiple islands of storage.

While this approach achieves more overall capacity, it comes with managing a sprawling environment of disparate clusters, leading to significant cost, complexity, and wasted time of the team.

WHY LEGACY SYSTEMS CAN'T SCALE EFFECTIVELY

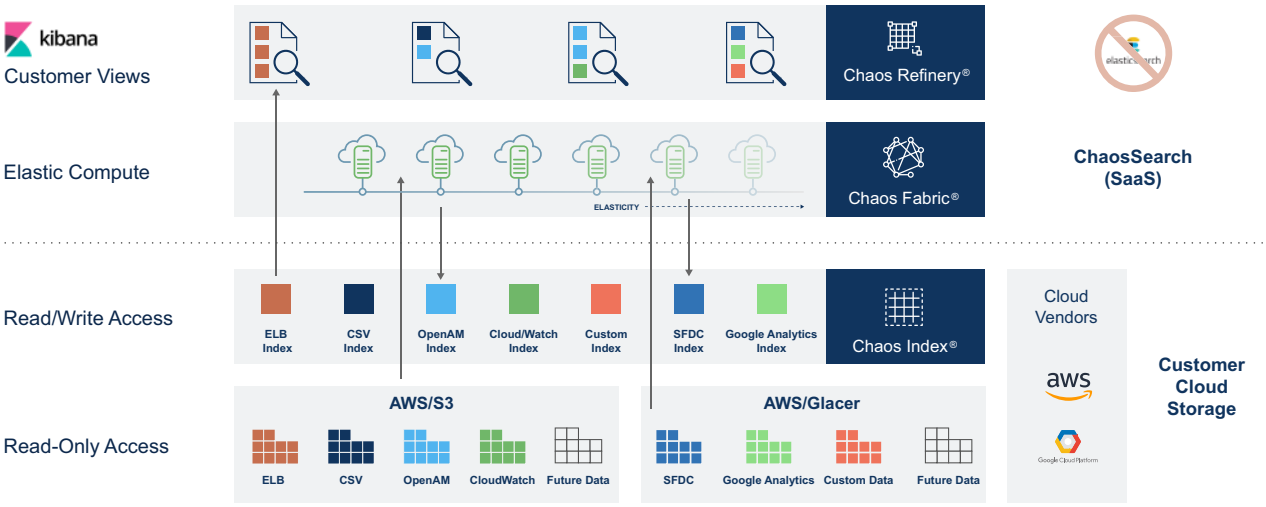
1 Legacy solutions have complex data pipelines which entail data parsing or transformations on ingest. This workflow introduces a significant bottleneck and reduces the maximum data volume processed daily.

2 Legacy solutions are "closed systems," which means after ingesting data, they take custody of the data and are responsible for it and the underlying layers of the infrastructure stack. This impacts both the daily workload and the system's overall storage capacity, limiting the customer's data retention period.

The Innovative ChaosSearch Solution

Instead of ingesting, manipulating, and managing the data as a custodian, ChaosSearch simply connects to and indexes all of the data within the customer's cloud storage environment- such as Amazon AWS or Google GCP — a fundamentally different approach.

As new data floods in daily, ChaosSearch continually indexes it in real-time without any performance hit to the data ingest speeds. Since we never take custody of the underlying data, we avoid complex data pipelines as well as the need for any data parsing or data transformation. ChaosSearch indexes data as-is while auto-detecting native schemas.



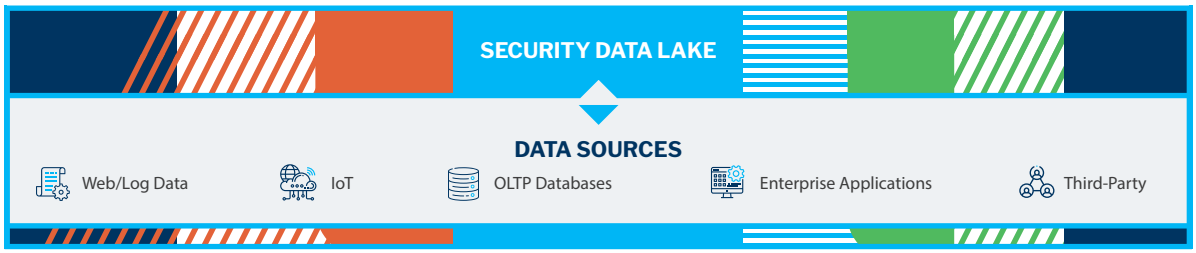
Once indexed, the ChaosSearch platform allows customers to conduct searches and review analytics using existing tools like Kibana, leveraging the open APIs of these tools.

This unique approach leverages the performance, scale, and economics of the public cloud.

HOW TO BUILD A SECURITY DATA LAKE

Given our approach, many customers select ChaosSearch to help them create a security data lake — and analysts agree this is a perfect use case!

Why? Our ability to support high-volume data storage and expanded data access for the SecOps team leads to increased data utilization, better analyses, and a superior security posture.



SecOps teams should be able to access data when and how they need it. The goal is to make the entire data lake searchable and ready for analysis.

HOW CHAOSSEARCH GIVES YOU AN ADVANTAGE OVER OTHER LOG MANAGEMENT SOLUTIONS

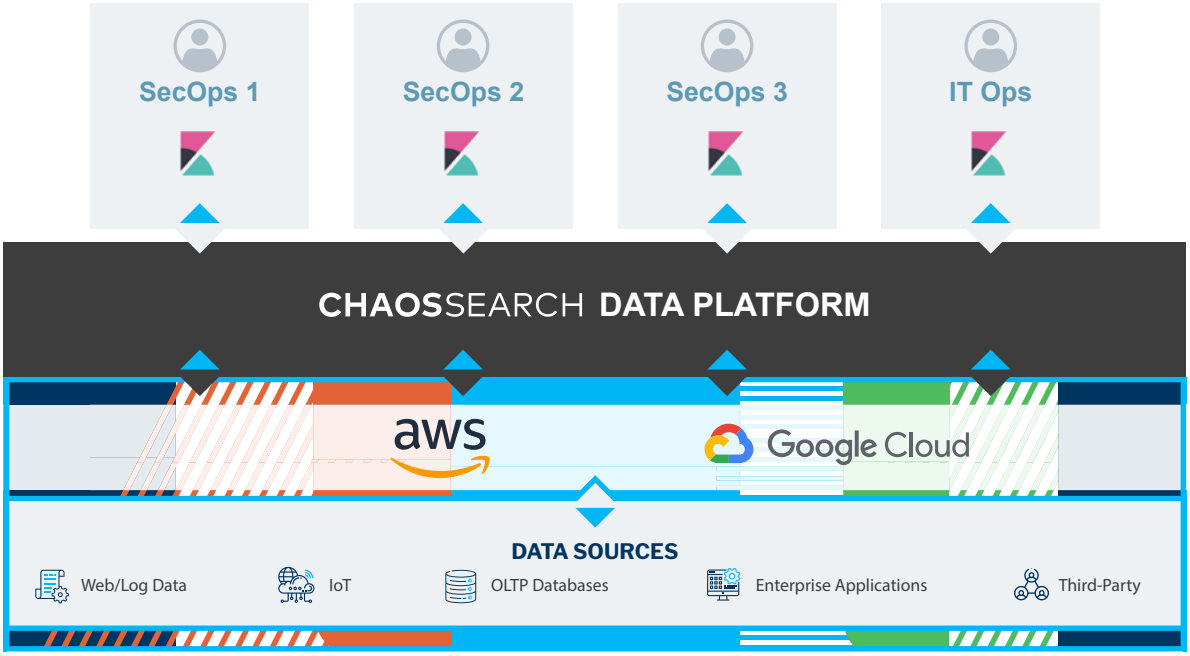
The ChaosSearch Data Platform does what the legacy platforms cannot—it allows customers to collect and analyze all of their data with virtually unlimited scalability and massive cost savings. Our customers typically see 50-80% annual cost savings compared to their legacy solution.

Security Data Lake Core Attributes:

- **Frictionless Ingest**
- **Massive Scalability**
- **Simplified Management**
- **Easy Access, Search, Query, and Analysis**
- **Cost-effective Growth**

"With our daily log and event volume exceeding tens of terabytes a day and growing, it became clear that we needed to find a new log analysis solution which could scale with Armor's growing business. We explored several options and found that ChaosSearch could deliver the reliability, scale, and expanded retention we needed."

Josh Bosquez
Chief Technology Officer at Armor, a cybersecurity software provider with over 1500 customers worldwide



The ChaosSearch Data Platform connects to a customer's existing cloud storage—in this case, either Amazon AWS or Google GCP. The SecOps team can search, query, and analyze our Data Platform using a Kibana interface.

SecOps teams use data in various security use cases that require access to data at scale. Below are a few prominent examples of how our customers use the ChaosSearch Data Platform today to improve their overall cyber defenses and threat readiness posture.

Improved ability to find advanced persistent threats (APTs):

The nature of this type of attack makes it necessary to look back further in time to identify the pattern of attack commonly used by APTs. The probing and infiltration stage can take months, which mandates the need to review data collected over a long time.

Support for threat hunting:

As threat actors become more clever and harder to spot, SecOps teams benefit from having more data to analyze and drive pattern recognition. As threats attack different aspects of infrastructure at different times, having more data and data that reflects a longer time frame improves the analytics or machine learning/AI that supports threat hunting. Increasing the retention period to go beyond average dwell time improves the odds of finding incursions that occurred weeks or months ago.

Identify DDoS Attacks in progress:

The ChaosSearch Platform's scale and resilience also enable customers to identify and neutralize threats quickly, including DDoS attacks. ChaosSearch integrates with popular content delivery networks (CDN) and security services like: Fastly, Cloudflare, AWS CloudFront, Carbon Black, Auth0, and Okta. We help customers understand application usage, traffic patterns, location of origin, and when and where their website or application is compromised. ChaosSearch's built-in alerting enables teams to set thresholds and automate responses to threats in near real-time.

The median dwell time for an incursion is 56 days.[1] If log data covers only 14 days, any incursion that happened weeks before won't be identified.

CHAOSSEARCH: YOUR FOUNDATION FOR ENTERPRISE CYBERSECURITY ARCHITECTURE

With its unique "data lake" approach, ChaosSearch delivers unmatched scalability in terms of the daily data ingest rate and the long-term data retention period.

This gives SecOps teams access to more data—precisely what they need to conduct various mission-critical operations such as forensic investigations, threat hunting, and DDoS attack mitigation.

With the growing need for superior analytics in cybersecurity, ChaosSearch is an ideal solution to deploy at the foundation of an enterprise's overall security architecture.

CYBER SECURITY EXCELLENCE AWARDS WINNER 2021

Best Security Analytics Solution

CYBER SECURITY EXCELLENCE AWARDS WINNER 2021

Best Security Log Analysis Solution

GLOBAL INFOSEC AWARDS WINNER CYBER DEFENSE MAGAZINE 2021

Cutting Edge in Cybersecurity Analytics

> ***"ChaosSearch now serves as one of our team's primary monitoring tools for identifying DDoS attacks and protecting our customers from them. The additional data retention also serves to help our security team audit issues over past months to better identify bad actors."***
>
> HubSpot
>
> **Stephen Salinas**
> Engineering Lead at HubSpot

[1] Fireeye. (2020). FireEye Mandiant M-Trends 2020 Report. https://investors.fireeye.com/news-releases/news-release-details/fireeye-mandiant-m-trends-2020-report-reveals-cyber-criminals

ABOUT CHAOSSEARCH

ChaosSearch enables customers to Know Better™, delivering data insights at scale while achieving the true promise of data lake economics. The ChaosSearch Data Platform connects to and indexes data within a customer's cloud storage environment, rendering it fully searchable and available for analysis with existing data tools – all with unlimited scale, industry-leading resiliency, and massive cost savings.

Based on these capabilities, ChaosSearch is an ideal replacement for the commonly deployed ELK stack today. With ChaosSearch, customers can perform scalable log analytics on AWS S3, using the familiar ElasticSearch API for queries, and Kibana for log analytics and visualizations while reducing costs and improving analytical capabilities.

We'd like to hear from you about your log management challenges and priorities. For any questions or requests, or to simply learn more, visit us online or send us an email.

info@chaossearch.com | **www.chaossearch.io**