CHAOSSEARCH

HubSpot

HubSpot

About HubSpot: Powerful Marketing & Sales Software Platform

HubSpot delivers SaaS applications for marketing, sales and service customers. Since 2006, the company has been on a mission to make the world more inbound. Today, over 73,400 total customers in more than 120 countries use HubSpot's software, services, and support to transform the way they attract and engage customers. HubSpot's log data ingest is in the 10s of terabytes per day - and is constantly growing.

The Challenge

HubSpot protects its customers by using Cloudflare to triage and actively prevent DDoS attacks. Since its early days, HubSpot internally managed an open-source ELK Stack to collect, route, and review its daily logs.

One of the major challenges of analyzing event data is adapting data pipelines, schema, and views when logs change. For example, the HubSpot team would need to expend effort when a field needed to be added or removed.

However, as HubSpot grew, the cost and complexity to expand and maintain their ELK Stack mounted. Even though the Elasticsearch, Logstash, and Kibana applications can be acquired at no cost, the sheer volume of existing and incoming log data from Cloudflare required HubSpot to manage a very large ELK cluster. In addition to the high maintenance costs, scaling the ELK Stack to accommodate their ever-increasing Cloudflare log data volume was difficult and costly. As a result, HubSpot was forced to sacrifice its length of data retention to just five (5) days.

The Solution

HubSpot required a scalable, cost effective, and managed log analysis solution to support the continued exponential growth of its Cloudflare log data.

Prior to onboarding ChaosSearch, HubSpot considered other solutions like Scalyr, Loggly, and AWS Athena. Ultimately, they chose ChaosSearch for its ease of onboarding new data sources, affordability, and unlimited data retention.

INDUSTRY
Sales & Marketing Platform

HEADQUARTERS
Cambridge, MA

SIZE
3,400 Employees

KEY CHALLENGES
Scalable log and event management

KEY RESULTS
· Seamless integration with data tools already in use (ie. Cloudflare CDN)

· Data retention increase from 5 days to 30 days with ChaosSearch (Unlimited capacity and unlimited data retention for search and analytics)

· Significant reduction in operations overhead

"

We are able to process and analyze 10's of terabytes a day of Cloudflare log data to identify and fend off DDoS attacks on behalf of our customers at a fraction of the cost of our previous self-hosted ELK Stack."

**- Stephen Salinas,
Engineering Lead at HubSpot**

Previously, the HubSpot team used the Kibana interface as part of their ELK Stack. Since the ChaosSearch service provides the Kibana frontend application, the transition to ChaosSearch was seamless. There was no new query language or mode of interaction that users needed to learn, and their analytics remained consistent. Specifically, the iterative query style combined with the schema-on-read model of Amazon Athena was cumbersome and slow when users tried to narrow down results. Where each Athena query would continue to scan the same amount of data for each iteration, ChaosSearch narrows the scan and the results, making it much more efficient.

In addition, HubSpot went from limited data retention while running the ELK Stack to unlimited data retention using ChaosSearch -- at a fraction of the cost.

As a result, HubSpot experienced 40% savings after making the switch from ELK Stack to ChaosSearch.

Favorite Features

- Auto discovery of schema and log format. There is no extra work or overhead if logs change or if a field is added or removed. HubSpot users can also set up new indexes of new types of logs and use cases with minimal effort.

- The Kibana UI as the frontend for ChaosSearch. Many users were already familiar with this as a frontend, so having a similar UI and query language was a huge benefit -- especially in a space where other providers often have their own custom query DSLs.

ChaosSearch Results

After implementation, HubSpot reported similar performance to the internal ELK Stack they had been using, but with the additional capability to search over a much larger data set and longer period - with more robust visualizations. This additional data retention also allowed the HubSpot security team to audit months of history to better identify bad actors and protect their customers more effectively.

In addition to these increased capabilities, HubSpot also reported an excellent experience with the ChaosSearch customer care team, saying that ChaosSearch created a quick and easy onboarding experience and was available for any questions that came up during the initial phase of use.

After several months of use, HubSpot continues to be pleased with ChaosSearch's ongoing performance, support, and cost savings.

"

Setting up our proof-of-concept and production environments were quick and easy. The ChaosSearch documentation made it easy to connect them to our source data in AWS, and auto-discovery of our data's schema allowed us to get started querying data from day one."

**- Stephen Salinas,
Engineering Lead at HubSpot**

"

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."

**- Stephen Salinas,
Engineering Lead at HubSpot**