Stile Education Improves Product Experience with ChaosSearch®

ANALYTICS AT STILE EDUCATION

As a science education technology company, Stile Education develops world-class curriculum resources via a digital learning platform for educators and students. With more than 1,000 digital science lessons and assessments for middle and high school students, Stile is relied upon by half a million teachers and students throughout Australia and the U.S.

Since Stile is so central to its educators' and students' day-to-day success, it is critical for the Engineering team to have fast access to application telemetry data. Analyzing log data is a crucial part of ensuring a seamless user experience, and quickly resolving any interruptions. In addition, the Engineering team wanted consistent and reliable observability into product insights, including how users adopt certain features.

Stile turned to ChaosSearch to provide a highly scalable, robust, and secure cloud data platform that centralizes data management and simplifies access to analytics. The platform is a fully managed SaaS solution connected to data within the company's existing cloud object storage environment. ChaosSearch made it easier for Stile to save log data indefinitely, without the management toil and cost of the previous Elasticsearch cluster. Now, the team can cost-effectively retain important business data for longer periods of time, which makes it easier to troubleshoot, debug, and ensure a flawless experience for users.

Stile

Industry
Education

Location
Australia and U.S.

Use Cases
- Application and infrastructure performance troubleshooting
- Customer experience improvement
- Observability into product analytics and insights

Impact
- Cost savings over previous Elasticsearch cluster
- Reduction in toil and maintenance for log analytics environment
- High availability and performance

Users
- Engineering team

THE CHALLENGE: SCALING ANALYTICS IN ELASTICSEARCH ADDED TOIL AND COST

Using its existing Elasticsearch cluster, the Stile Engineering team found it difficult to store the volume of logs they needed to analyze application performance issues that impacted the user experience. In addition, when the team shipped previews of new features to customers, they wanted to gain deeper insights into customer usage patterns over long periods of time. This data would enable the team to make changes and improvements to features in production. It became impossible for the team to maintain the cumbersome ELK Stack (Elastic, Logstash, and Kibana), while gaining access to the long-term data they needed to do their jobs effectively.

> " *Our Elasticsearch cluster was causing us pain. There were consistent maintenance issues, and it would frequently become overloaded. We had to scale up our cluster much larger than we needed it to be just to keep it stable, which was costing us too much money. And because of the way Elasticsearch works, we were only able to keep about one or two weeks of logs to optimize our cluster."*

Daniel Rodgers-Pryor
CTO at Stile

Data Management Environment

- Cloud Data Platform: ChaosSearch
- Cloud Object Storage: Amazon S3
- Analytic Interface: Kibana
- Observability Stack: Looker (BI), ChaosSearch (Logs), Grafana & Prometheus (Metrics), Jaeger (Traces)

As an education technology company, usage for Stile's product varies greatly. Peak usage occurs as teachers and students use the product in the classroom, with drop-offs during school vacation. Responding to the shift in load and keeping the appropriate amount of log data was challenging in Elasticsearch. During an incident or outage, error logs would make Elasticsearch run slowly, compromising the system when the Engineering team needed to query logs the most.

Often, by the time the Engineering team needed to troubleshoot or debug a problem, the logs in Elasticsearch were already inaccessible. The Engineering team was saving their own log backups in Amazon Simple Storage Service (Amazon S3) to prevent ballooning log retention costs in Elasticsearch. However, these logs were in the binary Elasticsearch index export format, which made them difficult to access. No one on the team had the time to load the logs back into Elasticsearch to search them, which left lingering issues unresolved.

"We looked at many traditional log analytics solutions that didn't have good usability and relied heavily on Athena, which isn't great for analytics even if you put a user interface on top of it," says Rodgers-Pryor. "Plus, we wanted an analytics tool that didn't require us to store all of our data in memory."

SOLUTION: CHAOSSEARCH CLOUD DATA PLATFORM

The team ultimately selected ChaosSearch because of its ease-of-use, low cost, and the ability to retain unlimited data directly in Amazon S3. Using a Kibana interface, the team can now search and analyze historical log data without data movement.

"One of the best things we've gotten out of ChaosSearch is the ability to keep all of our data in Amazon S3," says Rodgers-Pryor. "It's cheap and easy to keep all of our data available and indexed. We can search through it at any time to dig deeper into problems that crop up. ChaosSearch lets us ask the important questions about the root cause of issues, how long they've been happening, and other factors that might have contributed to the problem."

With ChaosSearch, the team can now retain log data indefinitely, versus saving only a week or two of data in Elasticsearch. That change has increased the team's capacity to use log data to solve business problems, and unlocked new opportunities to discover deeper product insights.

> " *Saving our logs for longer has been valuable because we can do root cause analysis on tricky problems. We can use logs for more and more types of data analysis. For example, we're looking at the data to find out how people use our new features and previews, when traffic ticks up and down, and more."*

Daniel Rodgers-Pryor
CTO at Stile

THE IMPACT: CONNECTING CHAOSSEARCH TO A BEST-OF-BREED OBSERVABILITY STRATEGY

The Stile Engineering team uses ChaosSearch as a part of a best-of-breed observability strategy, integrating it with Looker for business intelligence, Grafana and Prometheus for metrics, and Jaeger for traces. These tools interoperate well because the team generates global correlation ID tags for every request. That way, it's easy to follow the path through system events when something goes wrong. The team simply generates a link to a log query in ChaosSearch, and looks for the correlation ID in the time window when an issue occurred. That makes it simple to jump from a metric or trace into the logs, or from a message in the app into the logs to see exactly what is going on.
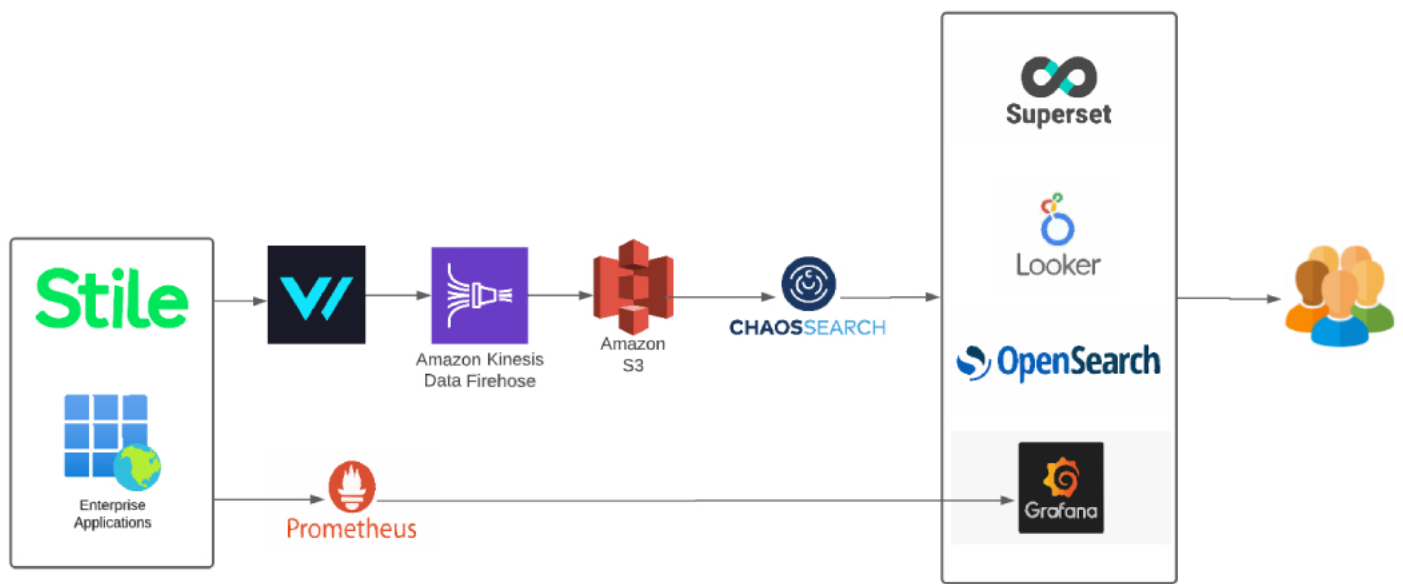
"We've just connected ChaosSearch to Looker using the SQL interface," says Rodgers-Pryor. "There's lots of potential to use ChaosSearch for more than just logs, as a general data analysis tool. For example, we'd love to build a graph of how a particular feature has been used over time, counting the number of users and plotting it by day. Without any extra code, you've got a usage graph for free. We'll also be able to measure new KPIs — for example, set a metric on the number of error logs from different services, break those down by different teams or parts of the application, or set alerts through Looker to understand certain system events like background job completions."

The Engineering team has already experienced a significant difference between the previous Elasticsearch cluster and ChaosSearch. Rodgers-Pryor says, "Rather than having to manage a cluster of components, we now have an auto-scaling analytics system with ChaosSearch and Amazon. We no longer drop log messages we need for troubleshooting, and we don't have to trade off on the amount of time we retain our log data. Now we just write our log data to Amazon S3, with no need to back it up manually."

In addition, Stile Education expanded the use of ChaosSearch to enhance security operations. "The extra capacity afforded to us by ChaosSeach has recently allowed us to start indexing our database audit logs, which gives us a new tool for investigating problems and analyzing performance, and provides additional confidence that we could respond to any potential security incident quickly by analyzing the already-indexed audit records," says Rodgers-Pryor.

The biggest measurement of success for the Stile Engineering team is the fact that their new analytics system reliably works during an incident. Troubleshooting is easier than ever before, and the team no longer has to devote valuable time to maintaining the ELK stack.

"*The most important things for us are minimizing toil by simplifying our analytics stack, and gaining capabilities to do new and interesting things with our data. ChaosSearch helps us do both of these things, while extended data retention in Amazon S3 makes it a lot easier to do our day-to-day jobs.*"

Daniel Rodgers-Pryor
CTO at Stile

To begin your free trial of ChaosSearch, visit **chaossearch.io**

ABOUT CHAOSSEARCH

ChaosSearch helps modern organizations Know Better® by activating the data lake for analytics. The ChaosSearch Cloud Data Platform indexes customers' cloud data, rendering it fully searchable and enabling analytics at scale with massive reductions of time, cost and complexity. ChaosSearch was purpose-built for cost-effective, highly scalable analytics encompassing full text search, SQL and machine learning capabilities in one unified offering. The patented ChaosSearch technology instantly transforms your cloud object storage (Amazon S3, Google Cloud Storage) into a hot, analytical data lake.

For more information, visit ChaosSearch.io.

info@chaossearch.com | www.chaossearch.io