SPONSORED CONTENT

AWS Insider.net

Highlights from a recent webcast

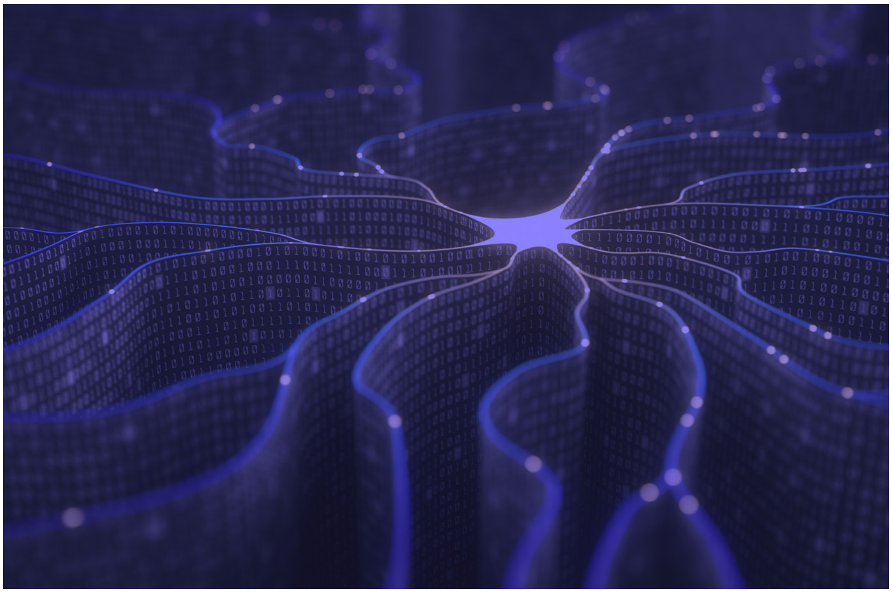
REACHING THE ELK STACK WALL

When Expanding Log and Event Data Threatens to Topple Your ELK Strategy

Based on a recent webcast featuring Betsy Bilhorn, Senior Vice President, Products at Jitterbit, Tom O'Connell, CSO at CHAOSSEARCH and Greg Schultz, Microsoft MVP, Server StorageIO.

Data is not helpful if you are drowning in it. This is especially true for the thousands of companies trying to use an ELK Stack to observe, measure and analyze their log and event data.

ELK, the acronym for three open source projects: Elasticsearch, Logstash, and Kibana, was originally touted as the fastest and most cost-efficient approach to logging challenges. But as data now grows beyond expectations, Elasticsearch is hitting walls and companies are suffering from unexpected infrastructure and human capital costs for managing this ballooning data.



ktsdesign / Shutterstock.com

The Data Explosion

With the exploding deployment of IoT devices and expanding business use of mobile computing devices, data is shattering the systems that were supposed to store, manage and analyze it. In what has been called "the data apocalypse" organizations are being swamped with more information than they can handle, especially in the cloud.

"How do we unlock the value, and find new value in that data in the cloud?" asked Greg Schultz, Microsoft MVP, Server StorageIO in a recent webcast, noting there is "unstructured data being stored in buckets, containers, objects, and blocks."

Specifically, how are you going to make use of vast amounts of data building up in your Amazon Web Services (AWS) S3 environment, especially when Elasticsearch is being stretched beyond its limits?

Current Trends & Approaches

The problems organizations are having as data multiplies beyond what Elasticsearch can handle, was explained to the webcast audience by Betsy Bilhorn, Senior Vice President, Products at Jitterbit, the API transformation company. The company needed to meet their customers' needs but database challenges were getting in the way.

"We have customers who were saying, *Hey, we want to have really robust analytics and we want to unlock the value of what you are logging for us. We expect that to be in the product. We expect you to be able to give it to us via streaming API or some other mechanism so I can take all that data from your platform and I can put it in my platforms."*

The goal was to meet these needs in a very consistent way, where things were indexed, and searches could be done very, very quickly.

As Bilhorn explained: "We wanted to have an environment where we could do search, we could do queries, we could do visualization."

She listed three challenges Jitterbit faced in reaching that goal:

1. Overall explosion of the systems that their customers are using today, which drives more business process automation/interconnectedness; and more data for synchronization, inspection, logging, analytics.

2. Mainstreaming of analytics, data lakes, etc. Customers have more data to work with and they want even more. The expectation is that, from a product perspective, vendors will supply robust analytics in-app and provide data in a format that customers can analyze outside of the app.

3. During almost a decade's experience with cloud platforms, the company had created a number of different places to log data. Meanwhile, needs and approaches have changed over time. Jitterbit had 13 different areas where it was keeping log data, which required analytics for productization and competitive differentiation.

"We had over 13 different instances across S3, Postgres, SQL Server, flat file in various places, and Loggly over six years of development," she explained. "It was impossible to do what we needed to do for analytics and usage metrics, not to mention any sort of new or differentiated offerings. It was extremely difficult to troubleshoot issues on the platform."

The company needed to consolidate on one system.

Loggly was already being used but ultimately could not be employed as part of the solution because of a combination of performance issues and HIPAA non-compliance.

SQL Server was not the solution because much of Jitterbit's data is not relational and the company ran into performance and other issues.

ELK looked like a possible solution.

"ELK seemed at the time to give us the performance we needed, visualization tooling we could embed in the product," Bilhorn said. But ELK also had limitations.

"We hit our ELK wall pretty early," she said.

For one thing, ELK turned out to have a pretty steep learning curve and the company had to hire consultants to help with it. And there were many additional costs beyond, which were not initially obvious.

"We started our initial ELK implementation and we were looking at an 18-to-24-month trend of how much data we were going to put in there and what was going to be our cost over that time," Bilhorn said. "It became very quickly obvious to us that we would pretty much almost double our entire AWS cost, which gave me pause. We were casting about and couldn't find anything and then we were introduced to **CHAOS**SEARCH. We saw that this was going to be a really groundbreaking solution for us to enable the things that we wanted to do. We worked with the **CHAOS**SEARCH team and the product delivered what they promised and we have been extremely pleased with the results."

> "We worked with the **CHAOS**SEARCH team and the product delivered what they promised and we have been extremely pleased with the results."
>
> *—Betsy Bilhorn, Senior Vice President, Products at Jitterbit*

The CHAOSSEARCH Solution

What if you could combine the cheap, scalable and secure cloud storage of Amazon S3 with search and analytic interfaces that let you extract value directly from the data without having to move it in and out of other services? Now you can.

CHAOSSEARCH is a fully managed SaaS platform that allows you to focus on search analytics directly in AWS S3 rather than moving data to expensive and complex databases for analysis. Leverage your existing AWS S3 infrastructure and let **CHAOS**SEARCH do the rest.

CHAOSSEARCH publishes Elasticsearch APIs, but we do not run any Elasticsearch software under the hood. **CHAOS**SEARCH is not an overlay or an add-on to Elasticsearch, but rather a full replacement of it. Only the API is the same. One huge benefit is that existing Elasticsearch users do not have to port their implementations. In other words, you get all of the power of **CHAOS**SEARCH without any heavy lifting.

Platform Overview:

- Fully managed service
- Unlimited data retention
- No data movement
- All on YOUR Amazon S3
- Less than half the cost of other solutions