

Highlights from a recent webcast on AWS and CHAOSSEARCH®

LEVERAGING AMAZON S3 FOR SERVERLESS DEVELOPMENT AND SEARCH ANALYTICS

Based on a recent webcast featuring AM Grobelny, Senior Partner Solutions Architect at Amazon Web Services (AWS), and Dave Armlin, Vice President Solution Architecture & Customer Success at CHAOSSEARCH®

From a developer perspective serverless architecture frees you to work on an app without worrying about infrastructure issues.

On the webcast, AM Grobelny, Senior Partner Solutions Architect at Amazon Web Services (AWS) provided the developer perspective on what types of workloads are good use cases for serverless, as well as how to operate with data and data transformation in that environment. He started out answering basic questions.

What is serverless?

- No infrastructure provisioning, no management
- Automatic scaling
- Pay for value
- Highly available and secure

“There’s no infrastructure provisioning,” Grobelny noted in covering the advantages of serverless architecture. “That’s one of my favorites though, since I am a developer, my background is building software. So I don’t have to worry about who’s going to go set up a server, who’s going to figure out how to keep the server on and who’s going to go patch. Those are things that just aren’t important because they’re all being handled behind the scenes.”

Automatic scaling is also a big plus for developers, he said. “I can deploy out my code and know that it’s going to be scaled in a way that is based on consumption and not me flipping dials and levers or an ops team having to come in and manage each little piece.”

But the advantages of serverless



architecture extend beyond developers to business people as well because they are only going to pay the cloud vendor for what is being used when Lambda is invoked, Grobelny explained. “You only pay when your Lambda functions are invoked for example, or if you’re using an API gateway, you only pay for when a request comes in through an API gateway. So you aren’t paying for idle time.”

Highly available and secure are the aspects of serverless that are important to the architect, the AWS architect concluded. “These pieces that you build

with are all inherently highly available and secure. You don’t have to worry about high availability. When you’re building with a service like Lambda, for example, it’s already built in to be redundant across multiple Availability Zones (AZs) in a region. You don’t have to go set that up. That’s one of my favorite pieces.”

Amazon Kinesis

There are a number of event sources that will trigger a Lambda function in an app but the webcast focused on Amazon AWS and Amazon Kinesis.

Amazon S3 has machine learning API's available to different constituencies, the developer and business user do not have to be ML experts or even know how it works to take advantage of ML, explained Grobelny, who added that he is not an ML specialist. "All I have to do is get the data into S3 and then I send it to this API. I get inference automatically without having to train any models, without having to develop any kind of machine learning. You hook up an S3 bucket to take in data and then automatically send it to one of these ML service APIs."

Kinesis is an AWS technology that handles both capture/store data, and loading data streams into AWS data stores. The capture and store component is Amazon Kinesis Data Streams. Amazon Kinesis Data Firehose loads data into stores. Data Firehose is a delivery mechanism for four different destinations within your AWS account. These databases, where data is stored for near real-time or longer term analysis, include:

- Amazon S3
- Amazon Redshift
- Amazon Elasticsearch
- Splunk

This is serverless because with Data Firehose, you're not going to have to have any code running in the server, except for what you're streaming into that firehose, the AWS architect explained. Typical sources of the data going into Firehose include devices and or applications that produce real-time data at high velocity, such as:

- Mobile devices
- Metering
- Click streams
- IoT sensors
- Logs

"CHAOSSEARCH is a search and analytics data platform that allows you to extract value from data stored directly in S3 without moving it." *—Dave Armlin, CHAOSSEARCH*

"You want to embed in your corporate website an area where every time someone clicks on something that data gets sent somewhere to get processed, Grobelny explained. "Kinesis is a really good fit for that and this really ties in to the **CHAOSSEARCH** workload really well." Kinesis Firehose is the easiest way of getting your log data from any of the disparate data sources.

Unlocking the Power of Search Analytics on Amazon S3

Dave Armlin, Vice President Solution Architecture & Customer Success at **CHAOSSEARCH**, an AWS Business Partner, explained the Boston-based company's guiding principles: "We think object storage is the home base for all data. You don't want to have to move your data out of your object storage to access it for analytics and search. All data should be accessible. And we believe that compute and storage should be separated. Engineers should do what they do best spending their time working on engineering problems, not managing an ELK cluster."

CHAOSSEARCH is a search and analytics data platform that allows you to extract value from data stored directly in S3 without moving it.

"We have our data index which is significantly smaller than the source data as it indexes the data and allows you to do search and analytics on that data," Armlin explained.

Data can come from standard log pushing mechanisms including Logstash, Kafka and fluentd as well as Amazon Fire Hose, he said. "Anything that can get data to S3 can then be utilized to do search and analytics with **CHAOSSEARCH**. In the end, you do search and analytics using an embedded Kibana inside this search platform. The database management layer allows you to discover the data that you have in your S3 bucket and catalog it and then do search and query upon it. From a user productivity perspective, we have live indexing when data enters the S3 bucket, and we will automatically index and provide access to that data."

User enhancements allow you to view query progress and also cancel inadvertent queries. It also inherits AWS security, durability and reliability.

"One of the most important parts about the **CHAOSSEARCH** platform is that there is really zero administration and you don't have to administer anything," Armlin explained. "So you're not worried about provisioning or integrating or anything like that. You just get your data to the S3 bucket and **CHAOSSEARCH** as a managed service will provide all that value for you."

SPONSORED BY:

CHAOSSEARCH

Find out more:

<https://www.chaossearch.io/>