SPONSORED CONTENT

AWS Insider.net

Based on a recent webcast on leveraging customer data

NEW CLOUDFRONT BREAKTHROUGH: FROM LOG DATA TO BUSINESS INSIGHTS IN MINUTES

Accessing the wealth of business-critical insights in Amazon CloudFront logs has always been challenging… until now

Many companies are struggling to manage and leverage the growing volume of data that their systems and customers are producing. As an example, a company's Amazon CloudFront logs contain a wealth of business-critical insights (i.e. cache efficiency, cyber threat indicators, product usage and access telemetry)—but accessing them has always taken a fair amount of effort, cost, and inefficiency requiring multiple AWS services (Lambda, Kinesis, DynamoDB, Athena...) to harvest business value. Finding more efficient ways to access and analyze this information was the subject of a recent webinar: "New CloudFront Breakthrough" sponsored by **CHAOS**SEARCH.

Amazon Web Services (AWS), the leading public cloud provider, has been honing its services to collect more performance information. A.M. Grobelny, Senior Partner Solutions Architect at AWS, noted that they offer a handful of services that provide companies with information about how their infrastructure performs:

- Amazon CloudFront is a Content Delivery Network (CDN) that supports distributed applications running in data centers around the world.
- AWS Lambda@Edge is a serverless edge network compute solution that provides an organization with a way to quickly write and run code that provisions system infrastructure.
- AWS Shield Advanced protects corporations against DDoS (Distributed Denial of Service) attacks by distributing traffic across multiple points of presence (POPs) and filtering requests to ensure that only valid HTTP(S) requests are forwarded to backend hosts. The service has many benefits.
  - AWS Shield Advanced protection is always-on.
  - Its flow-based monitoring of network traffic and its active application monitoring provide near real-time notifications of suspected DDoS incidents.
  - The service employs attack mitigation and routing techniques to automatically lessen the impact of an attack.
  - Its geo restrictions feature, also known as geoblocking, helps to isolate attacks originating from a particular location.
  - AWS WAF is a web application firewall designed to prevent malicious requests, like SQL injections and cross site scripting, from invading corporate networks. CloudFront integrates with AWS WAF, so enterprises can configure rules that filter out potentially malicious requests, and to minimize latency, requests are filtered inline at each POP.

Log Collection Best Practices

Customers need to provide applications with access to compute, storage and database resources in a high performing and secure manner. CloudFront is the software layer that sits in front of each interaction with an application deployed behind CloudFront. CloudFront empowers a rich responsive and secure customer experience for applications built and deployed on AWS. Corporations configure CloudFront to periodically push logs into AWS S3 at an interval that meets their business needs.

But log data selection criteria and collection practices sometimes miss valuable transactions, and that shortcoming prevents companies from gaining insight into system performance. AWS's Grobelny offered four best practices to capture desired data:

Skorzewiak / Shutterstock.com

- Specify the longest practical (not theoretical) value for keeping cached information. The longer that time period is, the fewer misses
- When building a query, do so consistently. Developers should use the same case (uppercase or lowercase) for all instances of a parameter. Any difference may lead to a cache miss.
- To reduce latency, forward only specified cookies, instead of all cookies
- Avoid caching based on request headers that have large numbers of unique values. The more unique they are, the more likely important data will not be collected.

Once a business has the log information stored, it needs a data analytics system to analyze and extract actionable information. AWS customers will often employ Lambda to preprocess CloudFront logs and stream them using Kinesis to another S3 bucket for analysis with AWS Athena. **CHAOS**SEARCH, a SaaS search and analytics solution provider, can provide immediate access to CloudFront log data in a customer's S3 bucket in a secure and highly scalable manner, greatly reducing cost and complexity with its fully managed service. Companies search it, gain valuable insights into system security and performance, and make necessary adjustments, according to Dave Armlin, Vice President Solution Architecture and Customer Success at **CHAOS**SEARCH.

Jimmy McDermott is CTO at Transeo, an educational compliance-focused software supplier. Its solution measures items, like student service hours and readiness for college, to help schools generate reports that meet government regulations and funding guidelines.

Maintaining its service is challenging because there are many compliance guidelines and the list keeps growing. Transeo has to meet regulations when its data is created, logged, retained and retired. In some cases, they generate bulk student information and must keep it for many years, since student information needs to be housed in an accessible manner years after pupils have graduated. Transeo also needs to be able to produce that data any time it is requested.

In addition, Transeo is subject to FERPA (Family Educational Rights and Privacy Act of 1974), a federal guideline that protects students' privacy. Compounding the work, each state has slightly different data storage and reporting regulations.

Not surprisingly, meeting the policy guidelines is often a time consuming and cumbersome process. Transeo had a legacy Heroku system but was running into roadblocks with it:

- They had the data stored, cleaned, and de-anonymized, but when C-Suite officials tried to garner insights, they often had limited or no visibility. The dashboards were hard to use and the reports sometimes lacked depth.
- Poor/no retention of logs
- Ballooning log costs as the service grew
- The company was managing security reactively rather than proactively
- Also, data collection processes had grown up autonomously. Transeo created a number of workarounds but consolidating the information was time consuming.

CHAOSSEARCH, a SaaS search and analytics solution provider, can provide immediate access to CloudFront log data in a customer's S3 bucket in a secure and highly scalable manner, greatly reducing cost and complexity with its fully managed service.

Finding the Right Analytics Solution

In sum, their enterprise needed a better system, and after a competitive review process, decided to move to **CHAOS**SEARCH. The change streamlined business intelligence as well as technical processes. Among the many benefits of moving to **CHAOS**SEARCH, Transeo found:

- They now have very predictable storage costs (just S3!)
- Team members share easy-to-understand dashboards, and they now dig deep into data to glean trends
- Users can now ask any question about any student. They previously had that ability on the transactions side, but now can do it with their time-based logs as well.
- The **CHAOS**SEARCH system offers unlimited data retention. Moving forward, the vendor plans to have multiple years stacked, so that they can better understand how circumstances change over time.
- Security improved. They use data combinations to proactively alert, detect, and thwart any attempted intrusions.

Organizations have rapidly been moving to public clouds for convenience. However, growing data volumes are making it challenging to manage their system infrastructures effectively. New tools from AWS and **CHAOS**SEARCH enable companies to gain more visibility and control of their infrastructure, so it better supports the business.