



FullStory improves digital experiences with AMD EPYC™ CPUs

Lower latencies and costs with AMD EPYC CPU-powered Google Cloud Platform instances



CUSTOMER

fullstory

INDUSTRY

Digital experience analytics

CHALLENGES

Reduce latency for large customers and provide capacity for new clients

SOLUTION

Switch from Intel-based N1 to AMD EPYC CPU-powered N2D Google Cloud Platform

RESULTS

40 percent reduction in latency for large clients' queries; 13 percent cost reduction; no need for additional database cluster

AMD TECHNOLOGY AT A GLANCE

AMD EPYC CPU-powered N2D Google Cloud Platform instances

TECHNOLOGY PARTNER

Google Cloud

Providing real-time analytics to thousands of customers across e-commerce retail, SaaS and other digital services takes a considerable amount of computing power.

To keep up with demand, FullStory delivers its digital experience analytics (DXA) platform through Google Cloud Platform (GCP), running hundreds of instances across thousands of cores. FullStory wanted to improve the service it provided to its customers by reducing the compute latencies with their existing platforms and win new business with new large clients. However, to accomplish this the company was having to contemplate an expensive expansion of part of its cloud infrastructure. That is until AMD EPYC™ CPU-powered GCP instances offered a better strategy.

Delivering real-time insights from terabytes of data

"We help businesses continuously improve their digital experiences across their websites and their mobile apps for their customers," explains Jaime Yap, Director of Engineering at FullStory.

"We offer a whole suite of functionality, everything from Session Replay to Dashboards to machine-generated insights, plus we provide lots of complicated analytics, searching, complex queries, data visualizations and machine generated insights. The infrastructure that powers this is very compute- and IO-intensive." FullStory had been running its database on Intel-based GCP N1 instances but was starting to outgrow the platform's capabilities.

"We've been using Google Cloud Platform for a long time," says Yap. "We're completely in the cloud. Most of our services today run on Kubernetes, but running databases is more challenging in a Kubernetes environment and that's where we still run a significant fleet of VM instances." FullStory's database needs are considerable due to the size of its clients.

"It was pretty seamless. It [migration to EPYC instances] just worked. It was live with no downtime."

Jaime Yap, Director of Engineering at FullStory

"A large e-commerce company will have huge traffic volumes. That becomes an increment to FullStory's traffic when they sign up as a customer. That traffic volume multiplies out across thousands and thousands of other customers, so the load and the amount of data that FullStory processes is mind boggling."

"Each one of our analytics database clusters will store and maintain hundreds and hundreds of terabytes of index data," adds Yap. "That's just in a single cluster spread across hundreds of machines and thousands of CPU cores. FullStory does most of its analytics in interactive time. You can be rendering several charts and graphs as you're typing and building your queries. Delivering low latency over that amount of data is not easy to do. You run into some interesting scaling challenges along the way, the data engine is very compute-intensive. But because we want to offer our users very low latencies, we have to think hard about our system architecture."

"We have designed our systems at FullStory (including our databases) to scale horizontally. One of the constraints of the database that we are using is that in order to continue to scale horizontally, we have to share the data in binary partitions, resulting in scaling at the power of two, in some parts of the stack, which can be prohibitive," explains Yap. "Alternatively, you can enter a development cycle, optimize hot spots or make system architectural changes. We are making those essential engineering investments at FullStory, but they take time. To move at the pace digital experience demands, we needed to also consider alternatives." A potentially more cost-effective (and time-effective) option would be migrating to a faster version of the existing topology, which is where FullStory's journey towards AMD EPYC CPU-powered GCP instances began.

Discovering AMD EPYC CPU-powered instances

Yap's interest in AMD technology started at a personal level. "I've been a fan of what AMD's been doing with its Ryzen processors. I own several generations of them for my personal use," he says. "When Google announced VM instances with the new generation AMD processors, we started doing benchmarking. We used a tool we developed and open sourced, called Solr Bench, to evaluate different machine configurations." FullStory tested a workload on an AMD EPYC processor-powered GCP N2D instance against an Intel-based N1 instance like the ones currently being used for the company's digital experience database. There was very little software alteration needed because both are 64-bit x86 architectures, and results were outstanding.

"Without having to write new code, we could change the machine configurations and see what happened," says Yap. "We did some micro benchmarks of AMD EPYC CPU-powered N2D compared to N1. This showed improvements in latencies of about 17%." This was enough of a boost to make it worthwhile trying out AMD CPU-powered instances on a wider basis. FullStory did this by creating a mirror of their existing database that was running on N1 instances, with AMD EPYC server N2D instances.

"We did some micro benchmarks of AMD EPYC CPU-powered N2D compared to N1. This showed improvements in latencies of about 17%."

Jaime Yap, Director of Engineering at FullStory

"We spun up a new cluster and we effectively ran two clusters," says Yap. "We had our old N1 cluster and our new N2D cluster. They had the same data, they were both indexing at the same rate, they were both serving queries. We hadn't yet cut over to the new cluster, but it gave us a very direct apples to apples comparison. They were running the same workload, with the same queries, so we could see which one was better. For our largest customers, for their worst-case latencies, we saw something like 40 percent improvement."

Although this huge performance benefit wasn't universally evident, the fact that it affected the largest customers most, was exactly what FullStory needed. "FullStory's largest customers – the ones with the most data – were unfortunately experiencing the highest latency," says Yap. "N2D shifted the latency distribution back into a healthy place."

"For our largest customers, for their worst-case latencies, we saw something like 40 percent improvement [with EPYC powered instances]."

Jaime Yap, Director of Engineering at FullStory

No need for a second database cluster

The much-improved performance had numerous benefits for FullStory. Although the entire database needed to be transferred from the N1 cluster to the N2D cluster, it proved to be a very easy migration. "We leveraged our existing infrastructure to perform database migrations. Once the clusters were mirrored, things went seamlessly. It just worked. It was live with no downtime," says Yap. The N2D cluster meant that the latency concerns for the biggest customers were addressed, and FullStory could continue taking on larger clients – without the huge expense of a major instance expansion and a whole new cluster on top of the existing one.

But there were additional savings too. "The cost benefit was more of the icing on the cake," says Yap. "When we looked at the sustained use discounting that GCP offers, both N1 and N2D came up to be about the same. But if we know that we are going to be on an infrastructure for a while, we can do a commitment for a certain amount of compute resources. When we compared the costs with this commitment, the N2Ds were about 13% less expensive than N1."

Switching to AMD CPU-powered N2D GCP instances delivered all FullStory's requirements. "We needed speed, we wanted it quick, without waiting on a full development cycle," says Yap. "It really was fortuitous that N2D instances came online. They provided a way to get a significant double-digit percentage improvement quickly without having to change our code. It was an x86 machine that could run the same workloads as an Intel machine but cheaper and faster enough to make a business difference. It was 13% cheaper for the same topology, giving us a cluster doubling we didn't have to do and then business that we can say yes to."

WANT TO LEARN HOW AMD EPYC™ PROCESSORS MIGHT WORK FOR YOU?

Sign up to receive
our data center content
amd.com/epycsignup



About FullStory

FullStory's digital experience analytics platform enables businesses to continuously improve their digital customer experience across sites and apps. The platform proactively surfaces actionable insights from billions of data points, helping product teams make digital improvements that reduce costs and reclaim revenue. FullStory was founded in 2014 on the belief that everyone benefits from a more perfect digital experience; today the company has 200+ distributed employees with offices in Atlanta and London. FullStory is the go-to solution for thousands of product leaders across the world's most innovative consumer brands and Fortune 100 companies. For more information visit fullstory.com.

About Google Cloud

Google Cloud provides organizations with leading infrastructure, platform capabilities and industry solutions. We deliver enterprise-grade cloud solutions that leverage Google's cutting-edge technology to help companies operate more efficiently and adapt to changing needs, giving customers a foundation for the future. Customers in more than 150 countries turn to Google Cloud as their trusted partner to solve their most critical business problems. For more information visit cloud.google.com/cloud.

About AMD

For 50 years AMD has driven innovation in high-performance computing, graphics, and visualization technologies—the building blocks for gaming, immersive platforms, and the data center. Hundreds of millions of consumers, leading Fortune 500 businesses, and cutting-edge scientific research facilities around the world rely on AMD technology daily to improve how they live, work, and play. AMD employees around the world are focused on building great products that push the boundaries of what is possible. For more information about how AMD is enabling today and inspiring tomorrow, visit amd.com/EPYC.

All performance and cost savings claims are provided by FullStory and have not been independently verified by AMD. Performance and cost benefits are impacted by a variety of variables. Results herein are specific to FullStory and may not be typical. GD-181

©2021 Advanced Micro Devices, Inc. All rights reserved. AMD, the AMD Arrow logo, EPYC, and combinations thereof are trademarks of Advanced Micro Devices, Inc. Google is a trademark of Google LLC. Kubernetes is a registered trademark of the Linux Foundation in the United States and/or other countries. Other product names used in this publication are for identification purposes only and may be trademarks of their respective companies.