

CRITEO MEETS EXPANDING DIGITAL ADVERTISING DEMANDS WITH AMD

CRITEO CASE STUDY

Digital advertising firm Criteo fulfilled 16 times as many market bid requests with fewer, higher-performing servers powered by 4th Gen AMD EPYC™ CPUs.



Digital advertising has become hugely compute-intensive. Over the nearly two decades of its existence, Paris, France-based Criteo has expanded its online display ads platform enormously to meet this growing demand. The company must constantly refresh its technology to enhance infrastructure performance per watt. AMD EPYC™ processors delivered exactly what Criteo needed to achieve its goals.

“Switching to 4th Gen AMD EPYC™ CPUs will enable a replacement ratio of one to two, meaning two old machines will be replaced by one.”

Matthieu Blumberg, Senior Vice President of Engineering Infrastructure, Criteo

“We have scaled our platform from 150 machines to around 30,000 over 13 years,” says Matthieu Blumberg, Senior Vice President of Engineering Infrastructure, Criteo. “Our services help media owners, such as retailers and publishers, monetize their digital assets, while bringing their users a richer browsing experience. We also deliver value to advertisers by driving awareness, discovery, and conversions of their products and services.”

Criteo has built a global reach for its platform. “We operate in over 100 countries,” says Blumberg. “Our platform is deployed worldwide across a topology of two data centers per region. There are two in the US, two in Europe, and two in APAC. There is also one dedicated data center for what we call the offline workloads. This is a full ecosystem supporting storage, compute, and AI. It's the most powerful of our data centers, using three megawatts of capacity, while in total we have 16 megawatts across all our data centers.”

WINNING REAL-TIME ADVERTISING BIDS

For the programmatic advertising Criteo delivers, responding in real time is crucial. “When we display an ad, the space where we show the ad is an inventory that we purchase on the real-time bidding market,” says Blumberg. “There is huge competition in this practice because the real-time bidding market is what helps publishers maximize the monetization of their content.”

This live auction poses huge challenges in keeping up with the level of compute required. “The first challenge is the amount of data that we must collect to support our workloads,” says Blumberg. “We deal with millions of requests per second on the platform. On the real-time bidding market, we are given a latency envelope, where each auction should be answered in less than 100 milliseconds, including the network round trip.

INDUSTRY

Digital display advertising

CHALLENGES

Meet the continually expanding volume of real-time requests in the programmatic digital advertising market bidding process, while remaining within power consumption constraints

SOLUTION

Deploy 4th Gen AMD EPYC™
8004 Series CPUs

RESULTS

16x more real-time digital ad bids fulfilled with half the number of servers Criteo used in 2020, while remaining within data center per-cabinet power limits

AMD TECHNOLOGY AT A GLANCE

2nd Gen AMD EPYC 7502P CPUs
4th Gen AMD EPYC 8534 CPUs

TECHNOLOGY PARTNER

Dell Technologies

This leaves us with just a few milliseconds for computing the requests and generating the banner if we win the bid. To put that in perspective, we display more than 5.1 billion ads per day, which requires us to bid on several hundred billion auctions per day. Our data set is around 180 petabytes on our Hadoop distributed storage system in our offline facility.”

“Since 2020, the number of requests that we process per day has increased significantly. But we are dealing with these using fewer, higher-performing servers powered by AMD EPYC™ processors.”

Matthieu Blumberg,
Senior Vice President of Engineering Infrastructure, Criteo

Hadoop is a significant part of Criteo’s infrastructure. “We have around 3,000 nodes,” says Blumberg. “We have invested significantly in the Hadoop stack.” Criteo runs two kinds of workloads. “The online workload, hosted in the regional data centers, is where we track user data from our integration with clients’ websites, receive the bid requests, compute them, and recommend products to end users. On the other side we have the offline workload, which is the centralized storage and compute platform, where we employ machine learning. This is also where we deal with the product catalog. We have roughly six times the number of products on the Amazon Marketplace—billions of products that we must ingest from our clients.”



The Criteo system delivers online ads via a real-time bidding process

The constant need for fast compute led Criteo to try AMD EPYC processors. “We were happy to see AMD back in the race in 2020,” says Blumberg. “We were refreshing some of our hardware configurations. At that time, we chose 2nd Gen AMD EPYC 7502P processors. In the most recent selection process, we discovered that Dell was one of the very few OEMs already offering 4th Gen AMD EPYC 8004 Series CPUs. We partnered with Dell, and the relationship has been going well.”

BETTER PERFORMANCE PER WATT WITH AMD

“AMD directly helped us to get early reference platforms to test,” says Aurélien Rougemont, Staff DevOps Hardware Engineer, Criteo. Blumberg adds: “The criteria in the selection process are all about raw performance plus power consumption. We have a specific framework to benchmark CPUs, which streamlines the metrics. We calculate the watts per core, factoring the performance per core and the performance per watt.”

“Based on our results, switching to 4th Gen AMD EPYC SP6 CPUs will enable a replacement ratio of one to two, meaning two old machines will be replaced by one,” says Blumberg. “That will fit into the server cabinet power envelope. We have a design that is limited by the electricity you can draw in a single server cabinet. We also care about the blast radius of a given server.” The “blast radius” is how much computing capacity you would lose if a server goes down, which will be more for CPUs with higher core density.

“AMD EPYC CPUs allowed us to decrease the overall number of servers and improve the infrastructure economics, while keeping up with increasing traffic and the amount of data we process, with no penalty for the overall performance of the workloads.”

Matthieu Blumberg,
Senior Vice President of Engineering Infrastructure, Criteo

Criteo considered a range of 4th Gen AMD EPYC processors but wanted to keep the “blast radius” below a certain level. “Performance per watts is the main feature we are looking at,” says Xavier Noëlle, Staff SRE Lead, Criteo. “We picked AMD Zen 4c technology because it was the best for our purpose. For pure performance, Zen 4 was great, but for performance per watts, Zen 4c was better. The thermal envelope constraints made us pick 4th Gen AMD EPYC CPUs. It was the perfect match.”

HALF THE SERVERS, MORE PERFORMANCE

“We are just beginning the at-scale deployment,” says Blumberg. “AMD EPYC 8004 Series CPUs will be powering most of the configurations we have right now.” The results have exceeded Criteo’s expectations. “The jobs are running faster, with a clear impact on the response and processing times. Since 2020, the number of requests that we process per day has increased significantly. But we are dealing with these using fewer, higher-performing servers powered by AMD EPYC processors.”

AMD EPYC CPUs will also help with Criteo's sustainability goals. "We have a public Corporate Social Responsibility report," says Blumberg. "The target we have set for the company is to reduce carbon emissions by 30 percent before 2030. This also frames the future design of our infrastructure. We are actively working on a direct liquid cooling solution. Our intention is to have a higher core density and raise the power envelope we have per cabinet by using direct liquid cooling."

"AMD has been a great partner that has helped us reach our objectives."

Xavier Noëlle, Staff SRE Lead, Criteo

"On average, on the online platform, 80 percent of the workloads are now powered by AMD CPUs and on the offline platform it's 20 percent," says Blumberg. "We now have 17,500 single-socket machines running AMD EPYC CPUs." Criteo is the first customer to have a mission-critical application on 8004 Series CPUs at that level of volume. "We have also been able to reduce the number of servers we use," says Noëlle. "The current number is just over 30,000. But the AMD EPYC servers are more powerful, so they are more than half of our compute ratio."

"AMD EPYC CPUs allowed us to consolidate the overall number of servers and improve the infrastructure economics, while keeping up with increasing traffic and the amount of data we

process, with no penalty for the overall performance of the workloads," says Blumberg. Noëlle concludes: "AMD has been a great partner that has helped us reach our objectives."



Ultra-rapid compute enables Criteo to win as many digital ad bids as possible



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

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

ABOUT CRITEO

Criteo (NASDAQ: CRTO) is the global commerce media company that enables marketers and media owners to drive better commerce outcomes. Its industry leading Commerce Media Platform connects thousands of marketers and media owners to deliver richer consumer experiences from product discovery to purchase. By powering trusted and impactful advertising, Criteo supports an open internet that encourages discovery, innovation, and choice. For more information, please visit criteo.com.

ABOUT AMD

For more than 50 years AMD has driven innovation in high-performance computing, graphics, and visualization technologies. Billions of people, leading Fortune 500 businesses, and cutting-edge scientific research institutions around the world rely on AMD technology daily to improve how they live, work and play. AMD employees are focused on building leadership high-performance and adaptive products that push the boundaries of what is possible. For more information about how AMD is enabling today and inspiring tomorrow, visit the AMD (NASDAQ: AMD) [website](https://amd.com), [blog](#), [LinkedIn](#) and [X](#) pages.

DISCLAIMERS

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

The information presented in this document is for informational purposes only and may contain technical inaccuracies, omissions, and typographical errors. The information contained herein is subject to change and may be rendered inaccurate for many reasons, including but not limited to product and roadmap changes, component and motherboard version changes, new model and/or product releases, product differences between differing manufacturers, software changes, BIOS flashes, firmware upgrades, or the like. Any computer system has risks of security vulnerabilities that cannot be completely prevented or mitigated. AMD assumes no obligation to update or otherwise correct or revise this information. However, AMD reserves the right to revise this information and to make changes from time to time to the content hereof without obligation of AMD to notify any person of such revisions or changes. GD-18.

COPYRIGHT NOTICE

© 2024 Advanced Micro Devices, Inc. All rights reserved. AMD, the AMD Arrow logo, EPYC, and combinations thereof are trademarks of Advanced Micro Devices, Inc. Other product names used in this publication are for identification purposes only and may be trademarks of their respective companies.