Overview
The Trade Desk, Inc. is a technology company that empowers buyers of advertising. Through its self-service, cloud-based platform, ad buyers can create, manage, and optimize digital advertising campaigns across ad formats and devices. Integrations with major data, inventory, and publisher partners ensure maximum reach and decisioning capabilities, and enterprise APIs enable custom development on top of the platform.

Challenge
We live in an era of disruption and few industries have witnessed as rapid a transformation as advertising. With advertising technology evolving rapidly and getting more sophisticated than ever before, advertisers can connect with customers in many new ways. All the major stakeholders in the AdTech ecosystem, whether marketers, publishers, or consumers, want more transparency. Marketers want to know exactly where their ads are positioned, which are working best, and how much each ad contributes to new revenue. Publishers want to know how audiences interact with the ads on their websites, and how much revenue is being generated for partners. And finally, consumers want to know how their data is used.

Ron Cormier, Principal Database Engineer at The Trade Desk, explains further: “The Trade Desk has revolutionized digital and programmatic advertising. We support buyers in the RTB space, helping agencies and brands spend their advertising budget in a more targeted and smart way. But spending is only part of the challenge, transparency is also key. We don’t just say we’re spending their budget smartly; we show them, from initial impression to conversion. That’s where Vertica comes in.”

The data volumes in this game are staggering, with The Trade Desk ingesting more than 1B rows of data each hour. Other server solutions weren’t able to handle the volume or weren’t able to perform at the necessary levels of flexibility and stability required. The team looked
“When we compare Vertica to other alternatives, we consistently find that Vertica does more with less. It is more efficient at data and query processing and has a head start of years compared to other products.”

RON CORMIER
Principal Database Engineer
The Trade Desk

for an enterprise-ready data analytics platform. Cormier: “We needed support for an open, standards compliant database with a SQL interface, but most of all, we needed it to be rock solid at ingesting huge volumes of data really efficiently and allow us to easily share that data with our engineers and customers.”

Solution
Vertica’s big data vision matched The Trade Desk, and in a proof-of-concept, Vertica proved scalable, innovative, and hugely performant. Once Vertica was implemented, Cormier realized that much more could be achieved: “With our Vertica-based data warehouse, our customers now have the transparency they need. With 10 million ad auctions taking place every second, you can imagine the data volumes involved.”

The Trade Desk now manages two enterprise Vertica clusters in a traditional data center model for full contingency. Each cluster contains 50 nodes with 2PB of data. Columnar storage and Massively Parallel Processing (MPP) ensure scalability for faster performance. Vertica Flex Tables enables The Trade Desk to quickly and easily load, explore, and analyze some forms of semi-structured data. It eliminates the need for coding-intensive schemas to be defined or applied before the data is available for exploration.

The Trade Desk workload varies, as Cormier explains: “Most of our Vertica queries run between midnight and 6am, and the rest of the day the usage is much lower. This got us thinking about leveraging cloud computing for easier capacity planning and increased infrastructure cost savings.” The Trade Desk was the primary beta test site for the new Vertica deployment model, Eon Mode, that separates compute and storage. In a recent Datanami article, a spokesperson for The Trade Desk is quoted as saying: “Vertica said ‘Are you guys willing to break it for us?’ And we’re like, ‘No problem! Can do!’ We broke 25 versions of it. But they kept dialing it in and it got real good by the end of it, so we’re really happy how that turned out.”

“We now have an additional Vertica cluster in Eon Mode,” continues Cormier. “This set-up allows us to decouple the compute from the storage capacity. We have the compute power when we need it but won’t pay for it when it’s not being used, as it’s independent of the shared storage capacity. Eon Mode gives us exactly the flexibility we look for, and the set-up naturally complements our development stack. We’re really excited about expanding this concept with Vertica.”

The Eon Mode cluster has 256 nodes of varying sizes and over 3PB of raw data. Between all clusters, Vertica processes over 40,000 automated reports each day.

Results
A lean team supports the Vertica architecture: “The three Vertica clusters are managed by just three database administrators,” says Cormier. “They support over 300 Trade Desk engineers and it’s only because of the ease of use and administration that we are able to do this. Our engineers can run their own reports, leveraging all the core Vertica features of scalability, stability, and performance.”

He concludes: “When we compare Vertica to other alternatives, we consistently find that Vertica does more with less. It is more efficient at data and query processing and has a head start of years compared to other products. We also appreciate the partnership with the Vertica team, especially our close collaboration around Vertica in Eon Mode. Lots of our suggestions have made it into the solution, demonstrating that Vertica truly listens to its customers. Vertica has helped grow our business in this highly competitive space.”