Playing to Win

How Vertica and Pure Storage deliver Playtika's unrivalled performance in gaming

"Vertica is the best analytic platform currently available in the market. It is fast and scalable, and independent from underlying infrastructure." – Shemer Mashiach, DataInfra Solution Lead, Playtika

Overview

This is an interview based on a real-life best practice story from Playtika, a leading digital entertainment company that specializes in the development and publication of free-to-play mobile games.

To capture this story, Sivan Tziring Rosenberg, Vertica regional sales manager, interviewed Shemer Mashiach, a DataInfra Solution Lead at Playtika. Her goal was to get a sense of how Playtika's analytic database system works, how it changed over time, and the challenges Playtika faced given the company's exponential data growth. Sivan's questions are followed by answers from Shemer.

Please tell us a little bit about Playtika and your background

I have worked with databases for many years, and today I'm Playtika DataInfra Solution Lead. Since 2010, Playtika has been a pioneer in the games industry. We were among the first to offer free-to-play social games on social networks and, shortly after, on mobile platforms. We were also one of the originators of live game operations, offering users personalized, daily game experiences with new events and exciting new features 365 days a year.

Playtika has over 3.5K Employees in 19 Office locations. We process hundreds of TBs of data daily, for over 35 million monthly active users. Through our diverse and constantly growing portfolio of games, we seek to create the most engaging and beloved entertainment experiences in the world. A lot of analytics is required to support this activity.

I heard from you in our earlier discussion that you use Vertica. What does Vertica do for you?

For me Vertica is the best analytic platform currently available in the market. It is fast and scalable, and independent from underlying infrastructure.

Vertica's SQL data warehouse is trusted by the world's leading data-driven companies to deliver speed, scale, and reliability on mission-critical analytics. It combines the power of a high-performance, massively parallel processing SQL query engine with advanced analytics and machine learning so we can unlock the potential in our data with no limits.

Where did you begin your journey with Vertica, and where are you positioned today?

Playtika has been using Vertica for years. Throughout the years the concurrency requirements got higher and we found the same Vertica product more useful for this purpose in Eon Mode. In Eon Mode, the primary copy is available on shared storage. The local cache serves as the secondary copy. This means that adding or removing nodes does not redistribute the primary copy.

This shared storage model enables elasticity, meaning it is both time and cost effective to adapt the cluster resources to fit the usage pattern of the cluster. Node failure and recovery is simplified in Eon Mode. If a node goes down, other nodes are not impacted because of shared storage. Node restarts are very fast, and no recovery is needed. To implement Vertica in Eon Mode on-premises, we use Pure Storage which delivers unrivalled performance and flexibility in our fast-changing industry.

Can you share the main challenge you experienced before?

To deliver on our SLAs despite increasing data volumes, greater demand for data analysis, and a need to isolate workloads.

What were the critical success factors you found when you moved to Eon Mode?

We succeeded, first and foremost, thanks to the great support from Vertica and Pure engineers. We can now add and remove sub-clusters as needed to expand compute capacity. And we can increase performance through resource dedication and "hot" data caching.

Why did you want to separate compute and storage resources?

In two words: more flexibility. As ambitious innovators, Playtika has as one of its values "dare to challenge." And challenge it did when in 10 years it grew from just 6 employees to over 3,500. Early on, Playtika made a strategic decision for the Vertica Analytics Platform. This served the company well for many years as it grew. As a completely data-driven business, Playtika relied on Vertica to manage its Extract, Transform, Load (ETL) data processes, act as a data warehouse, and perform advanced analytics. However, when data volumes reached over 10TB every day, the hardware infrastructure groaned and we needed to adopt a new architectural approach.

With fast increasing data volumes, more demand for in-depth analysis and reports, and a need to isolate data workload by games while still delivering high performance, we needed an architecture better suited to support variable workloads.

We decided the way forward would be to separate our compute and storage capabilities. This would give us the ability to rebalance the system without sacrificing performance on any of our games. Playtika has an on-premises strategy, and to solve our storage problems (we were operating at 80% of our storage disk capacity) we purchased an S3

Shemer, Sivan and Moshe in the Vertica office

Pure Storage FlashBlade. This gave us the storage flexibility we needed, and we were delighted to discover that Vertica in Eon Mode offered a great option for dynamic resource allocation.

I would like to add few words about Vertica's cost-effective high performance.

Vertica in Eon Mode for Pure Storage is the industry's first analytical database solution with an option to separate compute and storage on-premises. With excellent support from the Vertica team, including stress testing before the migration, the move from Vertica in Enterprise Mode to Vertica in Eon Mode was smooth and swift, with no need to rewrite or retune any of the existing data queries.

Each of Playtika's 14 game divisions has its own isolated compute cluster potential on the same single database. This means that research queries run by the data science team will not affect end-of-month reports being generated in another division, or executive dashboards showing data analyses across all divisions.

Playtika games send data to the backend servers where Vertica in Eon Mode digests and prepares it, and puts it to work for machine learning, data exploration, business and financial reporting. With Playtika's data volumes, the ability to cache active, or "hot," data is really helpful, too. Though the entire Vertica database is available for query on-demand, keeping this in communal storage and caching the "hot" data keeps performance high and our cost down.

While generally an on-premises strategy dictates a fixed compute capacity, Vertica in Eon Mode is designed for variable workloads. Playtika can choose to scale clusters up or down, according to variable workload demands. Node recovery is efficient, fast, and reliable, which allows easy and predictable cluster maintenance that does not interfere with production activity or performance.

Because nodes are so easily added and removed, we can choose to stand up a Quality Assurance (QA), development, or disaster recovery environment in minutes. By leveraging

Vertica in Eon Mode's sub-clustering ability, we can also give more computation power to our users as needed. This is also an option for our machine learning analysts to provide better insights to our business partners.

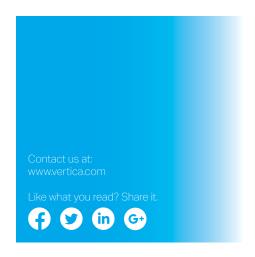### Today, what are the factors you can describe as project success?

This cost-effective model allows Playtika to assign data platform costs directly to game business units, by provisioning just the right amount of compute resources for their queries and just the right amount of storage resources for our data. As part of Playtika's architecture, Vertica is integrated with an Apache Hadoop-based data lake, Apache Spark. All these integrations are fully supported by Vertica in Eon Mode, and the additional flexibility in Eon Mode allows us to extend Vertica's analytical capabilities. The benefits are clear. With Vertica in Eon Mode and S3 Pure Storage, we have no data growth limitations. We process more than 10B events daily, executing sophisticated quantitative and advanced BI modelling to deliver actionable insights across our organization.

We have a stable, robust, flexible, and high-performance data analytics solution. We look forward to developing new capabilities, such as a Vertica and Pure Storage integrated High Availability (HA) solution for disaster recovery. We are also exploring an integration between Vertica and Micro Focus Voltage to support our GDPR compliance. I cannot stress enough how important the partnership between Vertica and Pure Storage has been in this. We received fantastic support, in my view the best I've worked with in the industry. It was great to see Vertica and Pure Storage collaborating to the benefit of all.

Can you describe the cluster environment? For example, how many nodes do you have?

One cluster includes 24 nodes. We use in each cluster 24 shards. Each node with 512 GB RAM, 88 cores, and 6 TB NVMe for the Vertica Depot. In Eon Mode, projections are not mapped to nodes directly. Data loaded into projections is instead mapped to shards (also known as segments in Vertica). During data load, nodes store the data if they serve a shard. This flexibility allows node set changes with less overhead.

The catalog in Eon Mode is broken into shards. The node maintains metadata for a subset of shards based on the mapping of which nodes have subscribed to which shards, also known as node subscriptions. These are automatically managed while running queries. At the time of database creation, you specify how many shards are created. Once this value is defined, it cannot be changed. There are two types of shards. The segment shards are for

OVER
35M
Monthly active users

segmented tables and are distributed across the cluster. The replica shards are for maintaining metadata for unsegmented tables. The replica shard lives on all nodes. If the node count is the same as the shard count, each node services a shard.

Now that you've finished the migration project, what is the status today?

With Vertica in Eon Mode and S3 Pure Storage we have no data growth limitations. We have the flexibility to add or remove nodes from the cluster without negatively impacting our games' performance. We have a stable, robust, flexible, and high-performance data analytics solution.

Thank you very much Shemer for sharing with us your experience with Vertica.

It was very interesting.

VERTICA | MICRO FOCUS®