

Citrix saves time and money on analytics with Matillion ETL for Amazon Redshift

ETL development and maintenance is 50% faster than hand-coding

Citrix, a cloud computing company focused on enabling mobile workstyles, is the creator of ShareFile, a secure file sync and sharing platform that helps business users collaborate internally and with their customers. ShareFile allows teams and their clients to access, sync, and securely share files from anywhere, on any device. The platform automates feedback and approval workflows to optimise operational agility and maximise productivity.

To optimise ShareFile and drive customer satisfaction, Citrix knew it needed to gather data related to how people were utilising the application. With limited resources and over 400 data events per second hitting the ShareFile pipeline on a typical day, Citrix knew that performing analytics operations to reveal the insights it needed would require an advanced solution that featured speed, power, and accuracy.

The decision to use Amazon Redshift and Matillion ETL

In 2016, Citrix began using Amazon Redshift from AWS, a fast, fully managed data warehouse that allows it to run complex analytic queries that help gain new insights into how customers use ShareFile. The application produces an event data pipeline that is written into Amazon Redshift, which delivers fast query performance by using columnar storage technology to improve I/O efficiency and by parallelising queries across multiple nodes.

When the executive team at Citrix recognised the need to bolster their use of Amazon Redshift to improve their data and reporting efforts, they turned to Matillion ETL. Built specifically for Amazon Redshift, Matillion ETL is a data integration tool that allows Citrix to aggregate data from multiple databases and provides structure to the insights and data visualisations the organisation needs. Procured from AWS Marketplace in a matter of minutes, Matillion ETL allowed Citrix to quickly move relational data sources into Amazon Redshift, and to embellish event-based data stored there with customer and location information from Microsoft SQL Server.

MATILLION ETL: ACCELERATING TIME-TO-VALUE FOR AMAZON REDSHIFT WORKLOADS

- Pushes data transformations down to Amazon Redshift, enabling you to process millions of rows in seconds, with real-time feedback.
- Features a modern browser-based environment that includes collaborative tools, version control, graphical job development, and more than 70 data read, write, join, and transform components.
- Deploys quickly and easily via Amazon Machine Image (AMI) so you can be developing ETL jobs within minutes.

CHALLENGE

Citrix needed to efficiently and accurately analyse customer data from its ShareFile data pipeline – up to 400 data events per second in a typical workday.

SOLUTION

Matillion ETL is designed specifically to streamline and accelerate data loads and transforms in Amazon RedShift.

RESULTS

Matillion ETL and Amazon Redshift empower Citrix to efficiently process their data to identify popular application features and allocate more resources to optimise their service and customer experience.

Time-to-value accelerated

Matillion ETL and Amazon Redshift have enabled Citrix to easily identify the most popular ShareFile features, as well as those that may be underperforming. With this knowledge in hand, Citrix can make informed decisions regarding where to focus product improvement efforts. The ability to load millions of rows of data into Amazon Redshift in minutes means Matillion ETL has greatly increased the efficiency and agility of Citrix's analytics initiatives.

Optimising collaboration

Citrix can now gather, collate, and analyse user data faster and on a more granular level than it ever could have when its data teams were hand-coding analytics workloads. Matillion ETL and Amazon Redshift enable the organisation to quickly process critical information regarding user location, feature use, multiple users associated with a single account, and much more. Because these kinds of details have typically been stored in scattered data warehouses, the process of extracting a meaningful story from all these disparate elements used to be a long and tedious one.

"We were in a place where we had access to a lot of really good information and data on our events platform, but not enough to tell the whole story. That's where Matillion came in."

Sheeya Gem, Staff Database Engineer, Citrix

Because of the efficiency Matillion ETL provides, Citrix saves time and money by dramatically reducing the effort required to perform complex analytics operations. Matillion ETL's graphical interface features live feedback, validation, and data preview so Citrix can see data being processed in-job, eliminating the need for time-consuming test cycles like there is with typical ETL tools. This reduces ETL development and maintenance efforts by around 50%.

ABOUT MATILLION ETL

Matillion ETL for Amazon Redshift makes loading and transforming data into Amazon Redshift fast, easy, and affordable. It takes minutes to set up and delivers results much faster than traditional ETL technologies. With just a few clicks, you can load data into Amazon Redshift from dozens of sources, including Amazon S3 and Amazon RDS; multiple databases and APIs; common systems like Google Analytics, Salesforce, Netsuite, and SAP; and even social media like Facebook and Twitter. Matillion ETL makes it easy to orchestrate and automate data load and transform, connect with other systems and AWS services, leverage scripts, and much more.

An ETL solution that drives innovation and customer satisfaction

Matillion ETL's push-down data transformation model harnesses the processing power of Amazon Redshift for transforming data. This helps deliver the performance and scalability Citrix needs to process massive amounts of event data – 50% faster than it was able to previously, when its data science team was hand-coding ETL workloads. As Citrix continues to improve ShareFile for its customers, Matillion ETL puts powerful orchestration tools at its disposal. The tool enables Citrix to choreograph data transformations and offers additional useful features like error handling – which helps the data science team verify the quality and accuracy of its query results – as well as iteration, logging, scheduling, version control, monitoring, collaboration, and environment support.

When data-driven insights are critical for making decisions around product development and improvement, customer service, and operational agility, Matillion ETL and Amazon Redshift help you reach those insights quickly and affordably.

“Matillion is a fantastic fit for us. It’s great for expert-level data technicians who write customised data queries, but also great for business users who need a tool that’s easy to configure.”

Sheeya Gem, Staff Database Engineer, Citrix



ABOUT MATILLION

Founded in 2011, Matillion is fundamentally changing data integration, enabling customers to innovate quickly and effectively with cloud-native data integration technology that solves individuals' and enterprises' top business challenges. Matillion is an AWS Advanced Technology Partner and an AWS Big Data Competency holder. [Matillion ETL for Amazon Redshift](#) is available worldwide via the AWS Marketplace.