

Quby Scales to Hundreds of Thousands of Users by Using a Unified Data Analytics Platform on AWS

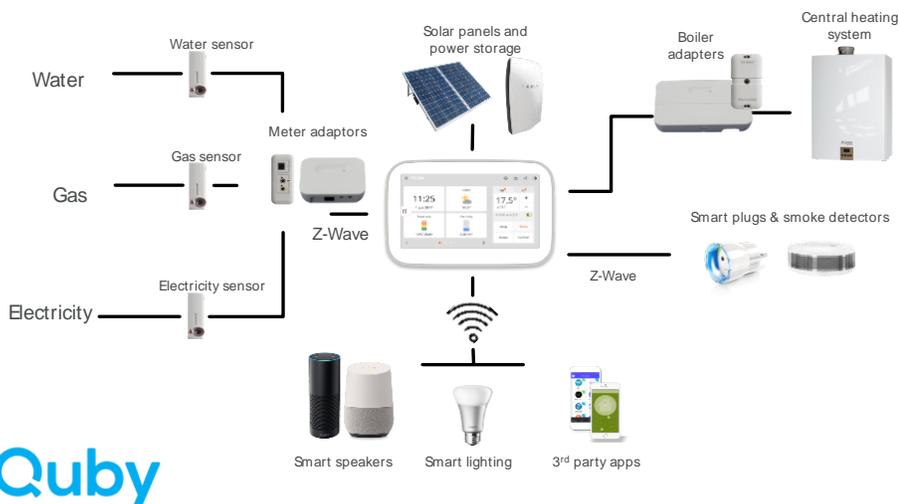


"The combination of the Databricks Unified Data Analytics Platform and AWS has allowed us to transition from a hardware-based business model to what is enabling both our partners and us to become more service orientated and ready for the future."

- Stephen Galsworthy, Chief Data Officer at Quby

Using ML and IoT Solutions for Better Smart Home Experiences

For the team at Quby, building a new generation of home energy management solutions goes beyond helping homeowners drive cost savings and efficiencies. The company's mission is to develop and evolve its Internet of Things (IoT)-connected home platform, powered by petabytes of energy usage data collected from over 400,000 households across Europe. Quby delivers insights that lead to an improved sense of home security, comfort, and connection for homeowners.



Quby delivers the Toon connected home platform powered by energy usage data to over 400k households.

Quby uses the IoT data it collects to take advantage of machine learning (ML). By using ML, Quby can detect anomalies and develop insights the company can then share with consumers and utility partners through its devices, while also driving new product and solution development.

While focusing on AI-first solution development to improve customer experiences, Quby's data engineers found themselves constrained by having to manage an on-premises Hadoop cluster infrastructure managed in separate instances of Hadoop. The company recognized that it would not be possible to scale this setup effectively while controlling costs. Quby also recognized that its teams would struggle to process streaming data reliably and efficiently in this environment. Based on its internal data science toolkit and varying technical backgrounds, Quby data scientists used different tooling to create ML models, stifling collaboration and leading to siloed development efforts.

About Quby

Quby is a European provider of smart home services. The company is the developer of **Toon**, a connected home platform and smart energy management device.

Challenge

Quby sought to take an AI-first solution development approach to evolve its IoT-connected home platform and drive better customer experiences. However, Quby's engineers were constrained by having to manage an on-premises Hadoop cluster infrastructure. The company recognized that it would not be possible to scale this setup effectively while controlling costs. Quby also realized that its teams would struggle to process streaming data reliably and efficiently in this environment.

Solution

The company began evaluating cloud-based analytics solutions to improve infrastructure and operations management while empowering data teams to drive better and more efficient collaboration. Quby selected Databricks' Unified Data Analytics Platform to migrate its infrastructure to AWS.

Benefit

The Platform has simplified end-to-end data analytics for Quby. Using the Databricks Unified Data Analytics Platform on AWS has empowered Quby to leverage ML on distributed data to build data-driven services. By transitioning to a unified data analytics approach, Quby data science teams can better leverage advanced ML algorithms and frameworks to collect better insights into energy usage and patterns to provide more personalized recommendations to customers.

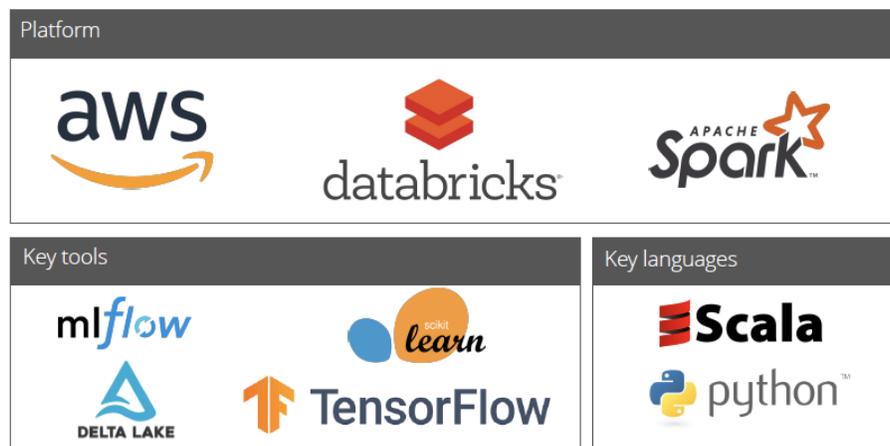
The company began evaluating cloud-based analytics solutions to improve infrastructure and operations management while empowering data teams to drive better and more efficient collaboration. "To scale up to hundreds of thousands of customers across Europe, we needed to take advantage of the cloud and adopt a more robust platform with which our data scientists could easily interface," says Stephen Galsworthy, Quby's chief data officer.

Seeking an Analytics Solution to Take Advantage of Petabytes of Data

Quby knew that migrating to Amazon Web Services (AWS) would support its data engineering and science teams and enable a company-wide, data-first transformation. Quby selected [Databricks' Unified Data Analytics Platform](#) to migrate its infrastructure to AWS. "We chose Databricks because its available tooling fulfilled our key requirement, which was to simplify our infrastructure via automated cluster management," says Galsworthy. "Over the duration of our relationship, Databricks has continued to develop and make additional tools, such as Delta and MLflow, allowing Quby to remain at the cutting edge of ML." Using the Databricks fully-managed platform on AWS with automated cluster management helps Quby simplify its infrastructure and operations as it continues to scale.

[Databricks Workspace](#), a notebook-based collaborative environment capable of running all analytic processes in one place, allows Quby teams to develop code more systematically and efficiently. The notebooks provide flexibility for engineers to seamlessly switch between programming languages, such as SQL, Scala, and Python.

The team also uses [Delta Lake](#), an open-source storage layer that brings ACID (Atomicity, Consistency, Isolation, Durability) transactions to Apache Spark and big data workloads to automatically index, compact, and cache data. Quby runs Delta Lake on Databricks which helps Quby simplify data pipelines by providing a standard API to transactionally store large historical and streaming datasets to Amazon Simple Storage Service (Amazon S3), making datasets available for high-performance analytics.



Quby

Quby's Data Science Toolkit



Quby is also using [MLflow](#), an open source platform introduced by Databricks to manage the machine learning lifecycle, to track experiments and put models into production. Using MLflow enables the data science team to more seamlessly test a variety of models on the same data set to compare results quickly. Given the complexity of Quby's machine learning models, the ease of use of the MLflow tool has encouraged the rapid adoption of MLflow amongst the Quby data science team.

Driving Collaboration and Accelerating Advanced ML Innovation

The Platform has simplified end-to-end data analytics for Quby. "The combination of the Databricks Unified Data Analytics Platform and AWS has allowed us to transition from a hardware-based business model to what is enabling both our partners and us to become more service orientated and ready for the future," says Galsworthy. Since migrating to Databricks and AWS, Quby's data engineers spend more time focusing on end-user issues and supporting data science teams to foster faster development cycles. Getting the company's first data services from proof of concept (POC) to product took over 12 months; the second release took four months, and today, Quby is able to release products to development within eight weeks.

Using the Databricks Unified Data Analytics Platform on AWS has empowered Quby to leverage ML on distributed data to build data-driven services. By transitioning to a unified data analytics approach, Quby data science teams can better leverage advanced ML algorithms and frameworks to collect better insights into energy usage and patterns to provide more personalized recommendations to customers. The company is taking advantage of flexible integrations and support from Databricks, focusing on building classification algorithms using machine and deep learning frameworks like TensorFlow, Scikit-learn, and Keras, which run on Databricks on AWS.



Databricks' mission is to accelerate innovation for its customers by unifying data science, engineering, and business. Founded by the original creators of Apache Spark, Databricks provides a Unified Analytics Platform for data science teams to collaborate with data engineering and lines of business to build data products. Users achieve faster time-to-value with Databricks by creating analytic workflows that go from ETL and interactive exploration to production. The company also makes it easier for its users to focus on their data by providing a fully managed, scalable, and secure cloud infrastructure that reduces operational complexity and total cost of ownership.



Advanced
Technology
Partner

Learn more at <https://aws.amazon.com/partners/>