# Customer Case Study

## Viacom

# Databricks Helps Viacom Build a Real-time Analytics Platform, Improving Video Performance and Streamlining Reporting

## About Viacom

Viacom Media Networks is the parent organization to VH1, MTV, Comedy Central, Paramount Pictures, and other well-known entertainment brands. Valued at $18 billion, the company creates original content for television, motion pictures, websites, and mobile applications, with programming reaching nearly 4 billion subscribers in over 160 countries. As an extension of the company's televised content, Viacom streams out petabytes of video data through web, mobile, and streaming players. Video content includes pre-recorded episode clips, trailers, and more, in addition to live video for events such as MTV's Video Music Awards (VMAs).

## The Challenge

### DELAYED INSIGHTS AND REPORTING COMPLEXITIES

When video data is continuously streamed, it can put a strain on delivery systems and result in videos taking too long to load, endlessly buffering, or even freezing mid-stream. These failures not only affect user experience, but they can mean losses to ad revenues and even brand loyalty. To combat video failures within its own properties, Viacom's Media and Technology Services (MTS) team collects performance data from the company's proprietary, public facing video players. Data collected gives Viacom business units a window into infrastructure performance to help them identify and fix player issues.

For years, video player reports could only provide a historical snapshot of how Viacom's players were functioning hours to weeks prior. Player malfunctioning couldn't be detected until long after it began. Similarly, there was no way of knowing what videos viewers were currently watching on Viacom properties. Additionally, when reports were generated, they had to be customized by the MTS team to fit the individual needs of data scientists, data engineers, and non-technical, business users; a time-consuming process.

## The Solution

### AWS BRINGS REAL-TIME ANALYTICS TO VIACOM

With the emergence of real-time analytics technologies, the MTS team looked for ways they could leverage new capabilities to improve data timeliness and relevancy, as well as the video viewing experience. Viacom turned to AWS to help them deploy a one-month proof-of-concept (POC) that would enable one-click, real-time, user recommendations from a player window. On Viacom's side, streaming data would be ingested through Apache Kafka, analyzed within an Apache Spark cluster, routed to Amazon Simple Storage Service (S3), and presented to viewers as recommendations, in as short a timeframe as possible.

The POC was a success. At the end of one month, the user recommendation feature was implemented in players and recommendations were served to viewers in as little as 30 seconds. However, it quickly became apparent that with the massive amount of data being generated (up to 1.2 TB per day), team members were spending too much time fine-tuning clusters and managing tasks, and not enough time on application development.

### DATABRICKS' UNIFIED ANALYTICS PLATFORM

The MTS team sought out a solution to help them handle the ongoing tasks associated with managing the infrastructure behind Viacom's new, real-time efforts. After an internal recommendation, they turned to Databricks for its ability to apply managed services to Apache Spark. They were hopeful Databricks could alleviate some of the ongoing responsibilities of managing Spark clusters by proactively monitoring video feed quality and automating resource reallocation. They were also interested in Databricks' intuitive interface that could help them get up and running with their data immediately.

Though initially implemented to resolve Viacom's real-time data management issues, Databricks also helped the company resolve internal reporting challenges. Databricks' Unified Analytics Platform enabled Viacom to deliver analytics to both technical and non-technical Viacom staff using a single code base. This process simplification allows for on-demand, self-service, reporting of data at varying levels of granularity, so Viacom's data scientists can get the raw data they need to run their custom models, and product managers can get the insights they need to make improvements and launch new features.

Other benefits that Viacom gained from using Databricks on AWS include:

## Why AWS

Though the company leases datacenter space, it knew that rapidly scalable cloud infrastructure was a requirement for the volume and velocity of data that would be generated by implementing real-time analytics capabilities.

Since team members had already experimented with AWS on a separate project and were familiar with the range of services it provided, they were confident that AWS was the right choice for their POC.

**SIMPLE DEPLOYMENT**

The MTS team was able to spin up self-managing Spark clusters in minutes with Databricks serverless. They were also able to navigate the Databricks Notebook interface in just days.

**STREAMLINED REPORTING**

With Databricks, data streams for ad hoc, batch, and real-time processing can all flow through the real-time pipeline, which means the MTS team only needs to monitor one system, saving both time and money.
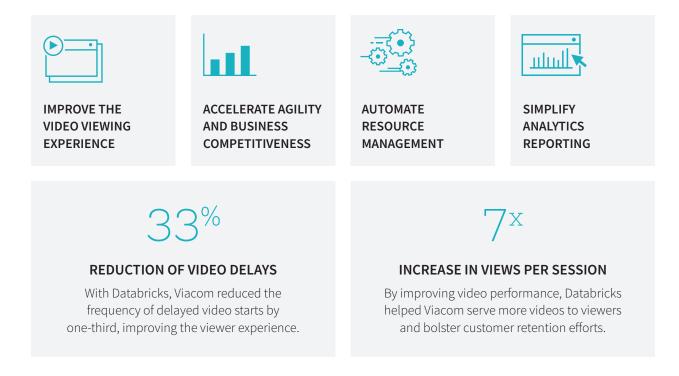
**INTEGRATION WITH OTHER BUSINESS INTELLIGENCE (BI) TOOLS**

Data can be pushed out to Tableau to accelerate business insights through visualization, and to New Relic for monitoring by Viacom Technical Support.

## Results

By working with AWS and Databricks to incorporate real-time analytics, Viacom was able to:

**IMPROVE THE VIDEO VIEWING EXPERIENCE**

**ACCELERATE AGILITY AND BUSINESS COMPETITIVENESS**

**AUTOMATE RESOURCE MANAGEMENT**

**SIMPLIFY ANALYTICS REPORTING**

## 33%

**REDUCTION OF VIDEO DELAYS**

With Databricks, Viacom reduced the frequency of delayed video starts by one-third, improving the viewer experience.

## 7x

**INCREASE IN VIEWS PER SESSION**

By improving video performance, Databricks helped Viacom serve more videos to viewers and bolster customer retention efforts.

---

*"Databricks lets us focus on business problems and makes certain processes very simple. Now it's a question of how do we bring these benefits to others in the organization who might not be aware of what they can do with this type of platform."*

- **Dan Morris**
  Sr. Director of Product Analysis, **Viacom**

---

## About Databricks

Databricks, founded by the team who created Apache Spark™, has been running on AWS since the company's beginning. The Databricks Unified Analytics Platform enables collaboration between data science teams, data engineering, and line of business analysts to get analytics results faster.

With Databricks, you can launch scalable Spark clusters in minutes that can process and analyze huge volumes of streaming data from multiple sources. Databricks can process ad hoc, batch, and streaming analytics together in a single analytics engine, delivering insights in hours and minutes instead of days and weeks. Databricks' fast deployment and fully managed cloud platform makes it easy for companies to harness the power of Spark and real-time analytics.

Learn more at databricks.com/aws

## Get started with a free trial of Databricks on AWS.