

[CASE STUDY]

NETFLIX

Now Streaming:

Why Netflix Runs a Fleet of
380+ CockroachDB Clusters



NETFLIX

Why Netflix Runs a Fleet of 380+
CockroachDB Clusters

INDUSTRY :

Media & Streaming (TV, films, documentaries, and games)

CHALLENGES :

Need for a high availability, scalable SQL database with multi-active topology.

SOLUTION :

CockroachDB-as-a-Service that's available to all Netflix developers who require a highly available, scalable, and consistent relational database.

About Netflix

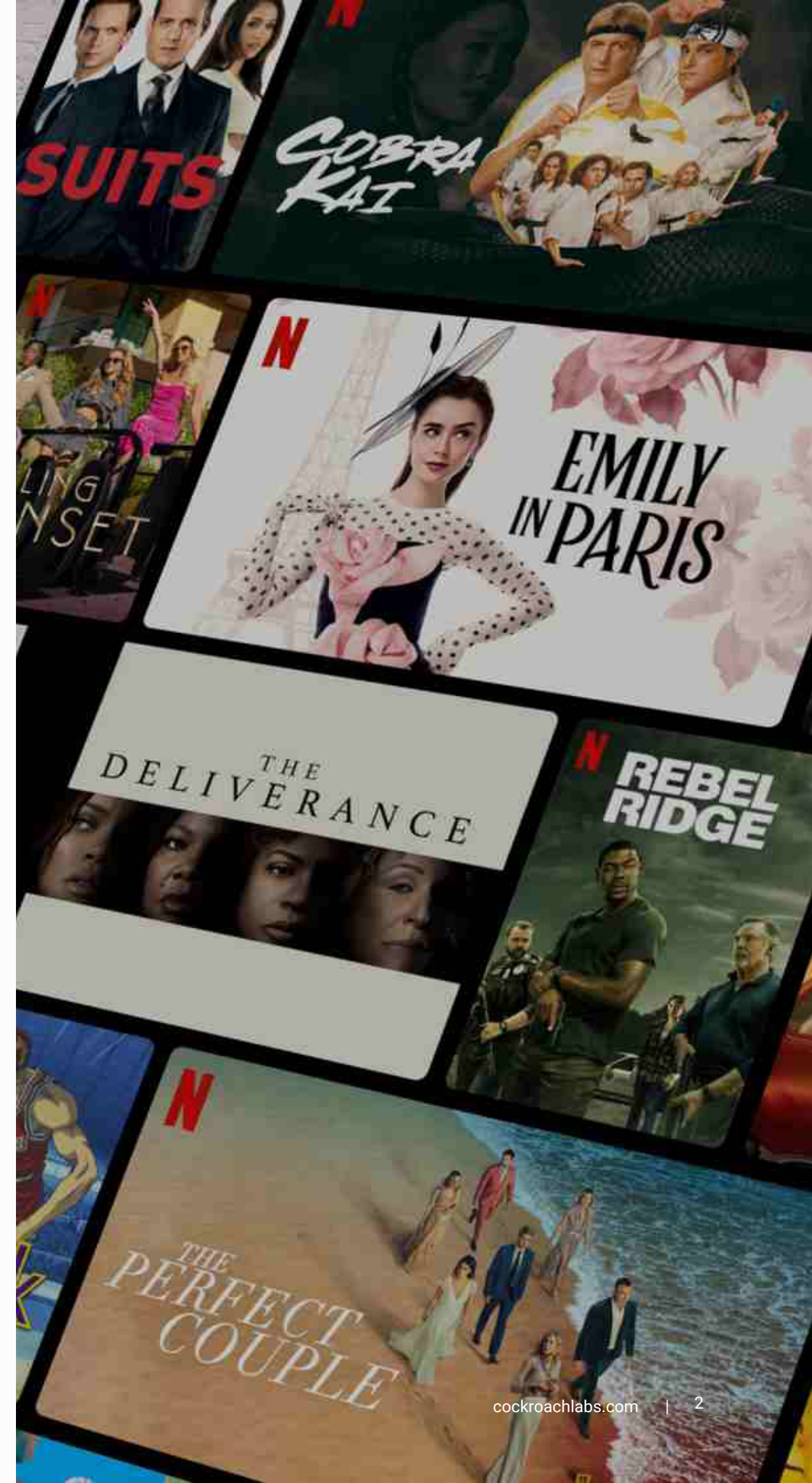
Long before [Netflix](#) became synonymous with acclaimed streaming and original TV series and films, the company first launched as a DVD rental business in the late 1990s. Since then the business and the world of media have drastically changed. Today, Netflix is a primary source of entertainment for viewers worldwide with over [278 million subscribers](#) and billions of dollars in revenue.

Streaming was first introduced on Netflix in 2007, and in 2008, Netflix experienced a 3-day outage due to their on-premise datacenter that would fundamentally change how they approached storing their copious amounts of data. In the ensuing years, the Netflix team would begin moving to the public cloud via AWS and Oracle. As they needed to expand to new regions, the team leveraged Cassandra for their global replication needs.

Then in 2016, Netflix Studio launched and with it came even more new considerations for the company's infrastructure. By 2019, the Netflix team was tasked with finding a new database solution that could deliver high availability, data correctness, and distributed scale as they expanded into new areas of business.

As Netflix's user base and product offerings have continued to grow and expand since they first implemented CockroachDB-as-a-Service, many teams within Netflix have come to know CockroachDB as the database solution for applications requiring data consistency, high availability, region failover, and more.

Watch  [Why Netflix Runs a Fleet of 380+ CockroachDB Clusters](#) ▶



A flexible solution that scales with the business

For the average Netflix user, we only see a polished streaming service, but teams within the global business must work together to ensure everyone can create the highest quality work. The Data Platform team at Netflix is charged with offering a suite of databases-as-a-service to other Netflix developers, who are building out various applications. Netflix was known for being an early adopter of Cassandra, and scaled to support other databases as well, such as Amazon RDS and Aurora. However, as the business grew, the team found that their options still lacked the ability to provide strongly consistent transactions or high scalability in single and multi-region deployments.

This came to a head in 2019 when the team needed a Cloud Drive to serve as a filesystem-like service for Netflix media assets. While Netflix offered other databases at the time, they were still missing a distributed database that could scale across multiple regions. For example, their content delivery network (CDN), called [Open Connect](#) needed a way to deliver content via a global control plane service for network devices.

Additionally the team wanted a solution that would survive any type of outage and grow with them in the future as their needs changed—after all, a database migration is a formidable undertaking.

In considering databases, the team evaluated several databases, including AWS Aurora and AWS DynamoDB along a variety of dimensions:

Database	CockroachDB	AWS DynamoDB	AWS Aurora
Multi-active topology	✓	✓	Single-region writer
Global consistent secondary indices	✓	Eventually consistent	✓
Global txns	✓	Single-region only	✓
Full control of the system	✓ (on-premise)	Limited	Limited
Disaster recovery	✓	No	✓

380+

CLUSTERS

60+

MULTI-REGION CLUSTERS

Up to 26TB

OF DATA FOR LARGE CLUSTERS



Ultimately Netflix decided to move forward with CockroachDB since it was the only database that met all of their requirements. The team was excited that CockroachDB had multi-active topology enabling high availability out-of-the-box since every node is a coordinator node. It would be simple to read and/or write to any database server across the globe since all nodes service both reads and writes.

Additionally, they knew they would need a highly customizable solution, and since CockroachDB is cloud-native, the team at Netflix was able to choose their own deployment method with flexibility to make changes in the future. In 2020 the team began deploying CockroachDB-as-a-Service to their clients—Netflix developers.

“If you need a high availability data store, we would recommend CockroachDB. A single node failure won’t cause a big problem, as the name suggests. In a multi-region setup, even in the event of total region failover, CockroachDB immediately transfers over leaseholders to other regions, and application owners have a totally seamless experience.”

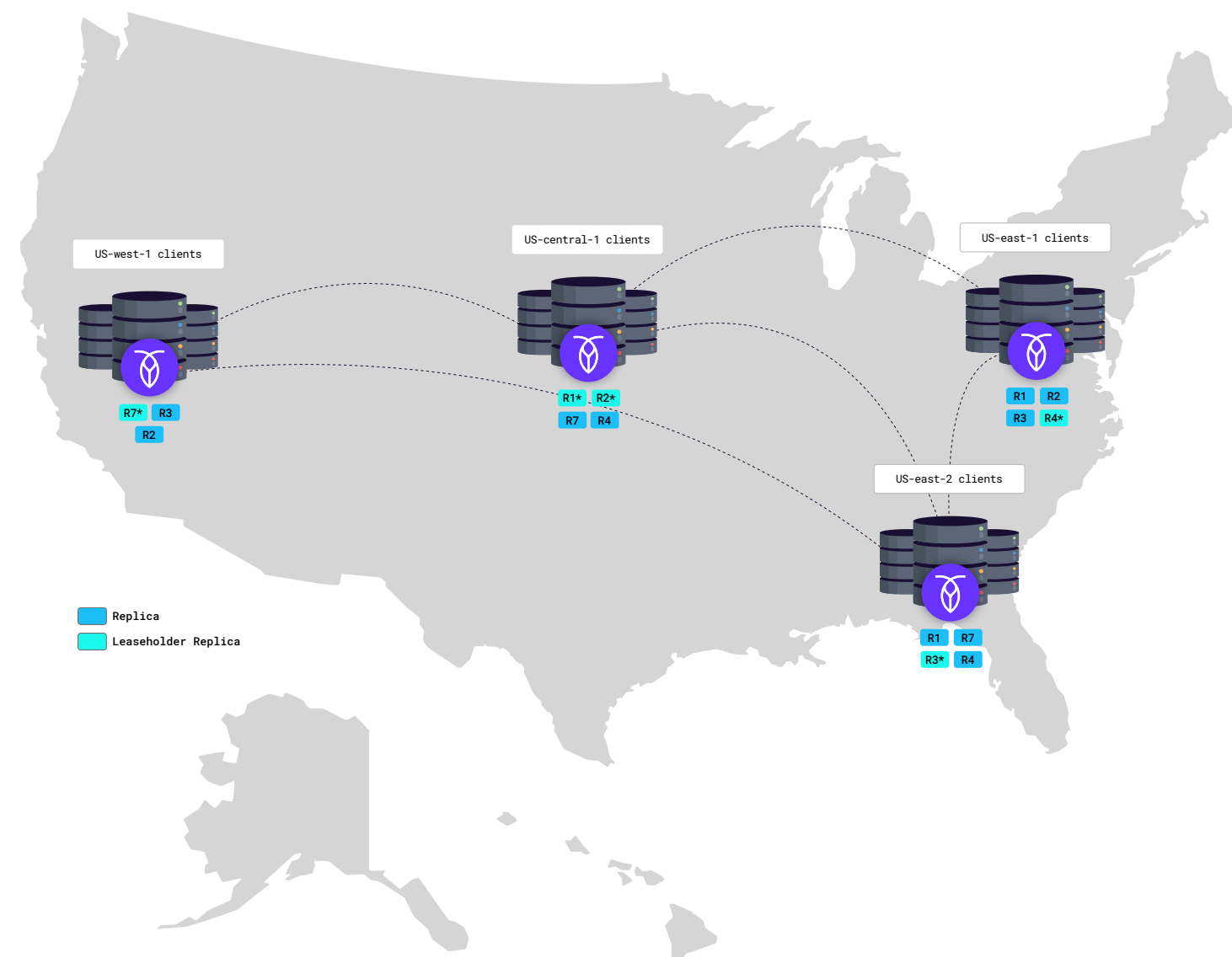
— **Shengwei Wang**, Senior Software Engineer, Netflix



Running a fleet of clusters

At Netflix, developers have the freedom to choose their own tech stack. The Online Datastores team, which Shengwei Wang is a part of, manages the various database offerings that Netflix developers can choose from. Developers don't need to determine the right technology or manage it on their own – instead Wang's team delivers the infrastructure and provides recommendations based on their use case's needs.

Since Netflix began offering CockroachDB-as-a-Service, the team has not looked back. Different business units are building different applications, and many have selected CockroachDB as their database backend, from device management and data/ML workflow orchestration to gaming.



“ ”

“As our business has scaled, as gaming has gone live, our requirements for databases have increased. We needed to scale the number of clusters over the past 2 years. Now we operate CockroachDB as a fleet.”

– **Ram Srivatsa Kannan**
Software Engineer, Netflix

In the past few years, Netflix's team has written extensively about their use of CockroachDB on their [technology blog](#). By 2022, Netflix's Data Platform team was managing over 100 production clusters, with the biggest cluster including 60 nodes in one region, storing 26.5 TB of data. Two years later, the team now deploys over 380 CockroachDB clusters, including over 160 production clusters and over 60 multi-region clusters.

Here's a few examples of the types of use cases Netflix is running in production on CockroachDB today.

A reliable device management platform

If you have a Netflix account, you've probably seen a Netflix app on your smart TV or a Netflix button on your remote. The seamless user experience is in part thanks to the Partner Infrastructure Team, which builds and maintains their [Device Management Platform](#).

The core challenge for the platform is handling device management at scale, supporting hundreds of different device types from smart TVs to gaming consoles and streaming sticks. Beyond simple functionality, the team must ensure streaming quality and the user experience customers have come to associate with the Netflix brand.

Critical to the platform is the ability to ingest and process events in a scalable manner, especially as the platform grows to accommodate more workloads and devices. Processing these events ensures that the device information is up-to-date and that the device tests are working. In this case, the horizontal scalability, guaranteed global consistency, and SQL capabilities of CockroachDB were critical.

Supporting a data/ML orchestration platform

In order to run a billion-dollar business, data is king. Decisions must be data-driven, whether that's the layout and look of a landing page or whether or not a sequel is made. As such, Netflix needed a top-notch workflow orchestrator to schedule and manage workflows at massive scale for the data platform. For their data scientists, data engineers, analysts, and more, [Netflix chose Maestro](#) for its scalability and ability to support a myriad of use cases.

In the Maestro system, all services are stateless and can be horizontally scaled out. CockroachDB is implemented to persist workflow definitions and instance state. Here, CockroachDB's consistency guarantees, as systems scale, without much operational overhead, play a significant role.

The latest gaming platform

More recently, Netflix has expanded into gaming, with their first mobile games launching in late 2021. Currently the Gaming Team utilizes a 48-node CockroachDB cluster to store metadata across the whole control plane and runs across 4 regions. CockroachDB was a great choice as the team had a strict survival goal and knew they would be generating large volumes of competing distributed transactions. In the event of a region failover, there would still be no downtime with CockroachDB.

“ ”

“At Netflix, we provide the freedom for everyone to choose their technology. However, to reduce costs and operate efficiently, we want to offer a platform that will support a majority of the use cases. If they want white-glove support and CockroachDB fulfills their business requirement, CockroachDB is my suggestion.”

– **Shengwei Wang**

Senior Software Engineer, Netflix



Start simple, optimize later

As early adopters of CockroachDB, the Data Platform team has learned a lot about the unique capabilities of distributed SQL databases. Their first suggestion to anyone working with CockroachDB is to start with something simple, intuitive, and safe because CockroachDB is so powerful. There are so many ways to configure your clusters to suit your needs. For example, the team uses a version of geo-partitioning to control every detail at the partition level, table level, or cluster level. They are able to use various tunable features provided by CockroachDB to control the locality and read/write patterns, for example.

Since CockroachDB makes it so easy to scale, the team suggests building on CockroachDB early and then as the business expands, CockroachDB can grow with you, and it's easy to optimize parts of the deployment later.

With regard to multi-region deployments, the team recognizes that there can be pain points as you expand to a new region. They suggest starting with just a gateway node since that will not impact the latency. Later you can do more tuning, more partitioning, and make additional changes.

“ ”

“Databases are complex systems, and distributed databases even more complex. Abstracting the complexities away from the user is something that a database developer should be conscious of. In that context, CockroachDB has done a phenomenal job.”

– Ram Srivatsa Kannan
Software Engineer, Netflix



Coming Soon

Over the years, CockroachDB has been able to scale as Netflix's business ventures have evolved and advanced. The team at Netflix has begun working with CockroachDB to build higher-level abstractions for users and both parties are excited to continue this initiative. In 2020, Netflix developers had to be a distributed database expert to be able to properly leverage all of CockroachDB's benefits. But over time, features like automatic resharding have made developers' lives easier.

Looking towards the future, Netflix will continue expanding the various aspects of its business. With the gaming initiative only a few years old, the team (and the rest of the world) is excited to see what they will build next. The goal, however, remains the same—to ensure Netflix developers can use CockroachDB-as-a-service as smoothly as possible to support their use cases.

Ready to get started?

Go hands-on with 100% free CockroachDB Serverless. Spin up your first cluster in just a few clicks.

