

NETFLIX ARCHITECTURE

END-TO-END OVERVIEW



AGENDA

- Introduction 01
- History of Netflix 02
- Business Model & Requirements 03
- High-Level System Overview 04
- Technology Stack 05
- CDN and Edge Network 06
- Scalability & Fault Tolerance 07
- Security and Scalability 08



AGENDA

- Recommendation System (AI & ML) 09
- Challenges 10
- Solutions 11
- Conclusion 12

1. INTRODUCTION



Netflix is one of the world's largest online streaming platforms providing movies, web series, documentaries, and original content across 190+ countries.

It started as a DVD rental-by-mail service and transformed into a cloud-based entertainment platform powered by microservices and AWS Cloud.

Netflix delivers billions of viewing hours every month to over 250 million subscribers using a scalable and fault-tolerant architecture.



Key points:

- Accessible on smartphones, tablets, smart TVs, laptops, and gaming consoles.
- Offers on-demand streaming with personalized recommendations.

2. HISTORY OF NETFLIX

Founders



Reed Hastings



Marc Randolph

Timeline

1 1997 – **Netflix founded:** Originally a DVD rental-by-mail service in California.

2 **1999 – Subscription model:** Introduced monthly subscription without late fees.

3 **2007 – Streaming introduced:**

- Enabled instant video streaming on PCs.
- Began building digital content delivery infrastructure.

4 **2010s – Global Expansion:**

- Streaming expanded to Canada, Europe, Latin America, and Asia.
- Content licensing agreements became key for international availability.

5 **2016 – Cloud migration complete:**

- Fully migrated to Amazon Web Services (AWS).
- Decommissioned traditional data centers.

6 **Today – Microservices & AI-powered:**

- Hundreds of microservices power the platform.
- Personalized recommendations driven by machine learning.
- Open Connect CDN delivers content efficiently worldwide.

3.BUSINESS & REQUIREMENTS

- **Netflix business model**

  Subscription Video-On-Demand (SVOD)

  - Users pay monthly fees for unlimited streaming.

- **Technical requirements:**
  - **High availability**: Streaming should never fail.
  - **Scalability:** Millions of users streaming simultaneously.



- **Personalization:** ML-driven recommendations for engagement.
- **Global delivery:** Content available with minimal latency.
- **Reliability & fault tolerance**: System must continue even if some services fail.



Business Model Evolution

Phase 1: DVD Rental by Mail (1997–2006)

- Physical DVDs shipped to customers.
- Subscription-based; no due dates.

Phase 2: Streaming Service (2007–2010)

- Started streaming over the internet.
- Initial architecture: monolithic system hosted in data centers.
-

Phase 3: Cloud Migration & Global Expansion (2010–2016)

- Migrated entire infrastructure to AWS Cloud.
- Adopted Microservices Architecture for scalability.

Phase 4: AI-Powered Global Streaming Platform (2016–Present)

- Uses machine learning, recommendation algorithms, A/B testing, content analytics.



4. HIGH-LEVEL SYSTEM OVERVIEW

User Devices: Web browsers, mobile apps, smart TVs

API Gateway: Handles incoming requests and routes to the correct microservice.

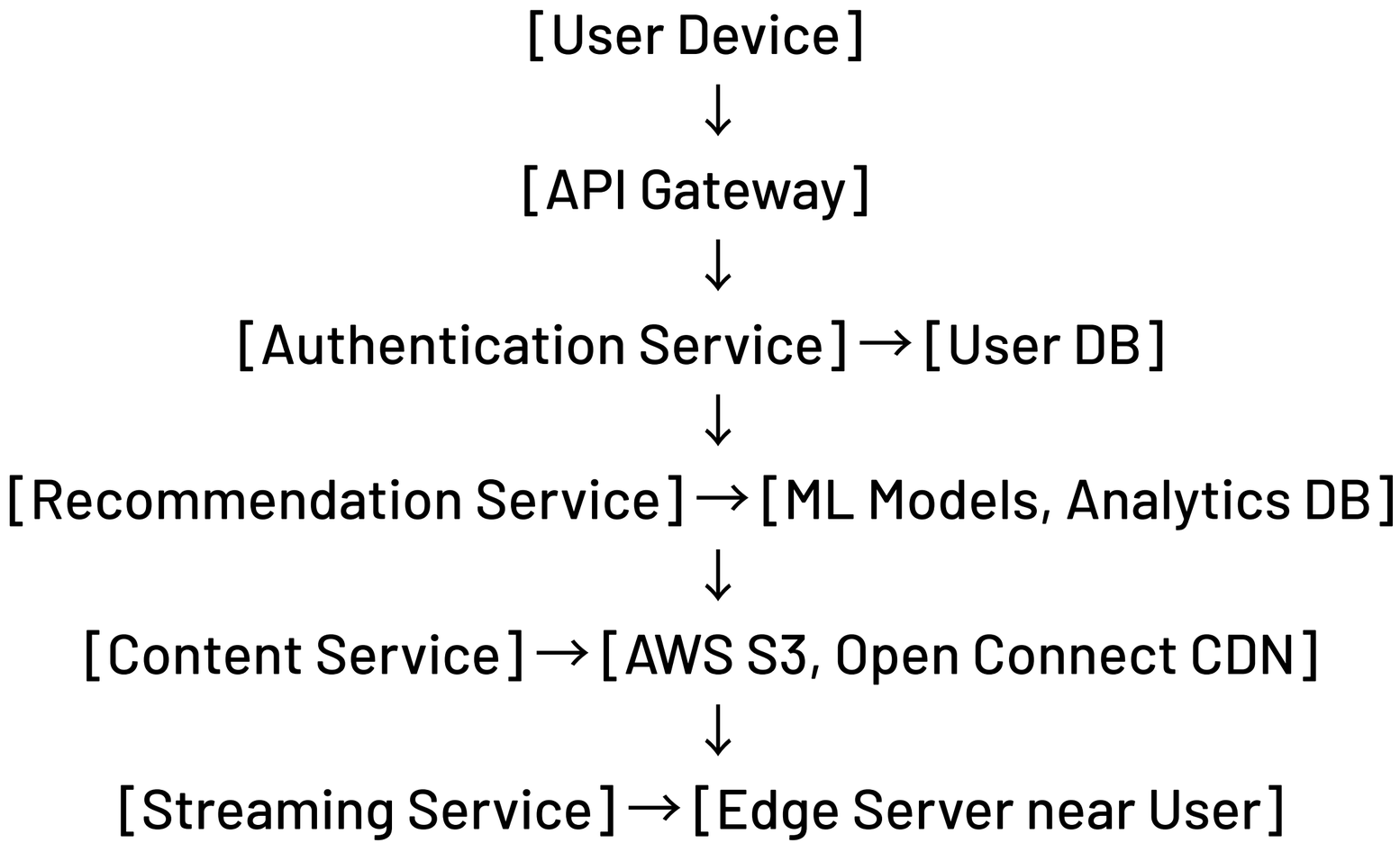
Microservices Layer: Handles business logic (user management, catalog, recommendations, playback).



Databases: Store user profiles, viewing history, content metadata.

CDN – Open Connect: Efficiently delivers video content globally.

Netflix Architecture Flow Diagram



5. TECHNOLOGY STACK

Backend	Frontend	Databases	Cloud Infrastructure	DevOps Tools
				
Java, Python, Node.js, Scala, Kotlin Frameworks: Spring Boot, RxJava	**React.js, JavaScript, Next.js**	• **Cassandra (NoSQL, for scalability)** • **MySQL (for transactions)** • **Elasticsearch (search indexing)** • **Redis (caching)** • **S3 (video storage)** • **DynamoDB (user data)**	**AWS (Amazon Web Services)**	• **Spinnaker – Continuous Deployment** • **Jenkins – CI/CD** • **Chaos Monkey – Resilience testing** • **Docker / Kubernetes – Containerization**

6. CDN AND EDGE NETWORK

Netflix uses Open Connect CDN – its own global content delivery network.

How it works:

1. Popular content (movies, shows) is stored on edge servers near users.
2. When a user requests a video:
    - Netflix redirects them to the nearest edge server.
3. Reduces latency, buffering, and improves streaming quality.

7. Scalability & Fault Tolerance

Netflix designs for failure-resilience:

- Load balancing: Distributes requests across multiple servers.
- Circuit breakers (Hystrix): Prevent cascading failures.
- Chaos Engineering: Tools like Chaos Monkey randomly kill services to ensure resilience.
- Auto-scaling: AWS EC2 instances automatically scale based on demand.



8. Security and Scalability

Security:

- OAuth 2.0 Authentication
- SSL Encryption
- Token-based access
- DRM (Digital Rights Management) for video

Scalability:

- Auto-scaling groups (AWS EC2)
- Load balancers
- Region-based redundancy (multiple data centers worldwide)



9. Recommendation System (AI & ML)

Netflix’s Recommendation Engine powers 80% of user watch activity.

Techniques used:

01



Collaborative Filtering (user behavior comparison)

02



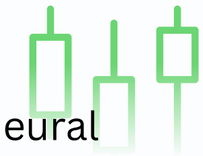
Content-Based Filtering (movie metadata)

03



Real-time analytics with Apache Kafka and Spark

04



Deep Learning models (neural networks)

10. CHALLENGES

- High traffic & scaling
- Global latency
- Deployment failures
- Resilience

11. SOLUTIONS

Challenge	Netflix Solution
High traffic & scaling	Microservices + AWS Auto Scaling
Global latency	Open Connect CDN
Deployment failures	Spinnaker + Canary Releases
Resilience	Chaos Monkey testing
Data consistency	Event-driven architecture (Kafka, Cassandra)

Conclusion

- Netflix is a cloud-native, microservices-based, highly scalable platform.
- Uses AWS, Cassandra, Kafka, Open Connect CDN, Redis, ML-based recommendations.
- Focuses on availability, personalization, fault tolerance, global delivery.
- Practices like Chaos Engineering ensure resilience and reliability.



THANK YOU

