

1. **Quais são os principais componentes que compõem o sistema?**

A Netflix possui uma arquitetura complexa formada por diversos componentes distribuídos. O núcleo do sistema é a Open Connect, uma CDN proprietária composta por servidores de borda instalados globalmente em parceria com provedores de internet, responsáveis por armazenar e entregar os vídeos mais acessados próximos aos usuários. O backend opera principalmente na AWS, utilizando milhares de instâncias EC2 que executam centenas de microserviços especializados para funções como autenticação, recomendações personalizadas e gestão de assinaturas. Para armazenamento de dados, a Netflix emprega bancos distribuídos como Cassandra para dados que exigem alta escalabilidade, DynamoDB para estruturas chave-valor e Elasticsearch para indexação e buscas. O tráfego entre esses serviços é gerenciado por gateways como o Zuul, que atua como balanceador de carga inteligente. No lado do cliente, aplicativos otimizados rodam em diversas plataformas como smart TVs, dispositivos móveis e navegadores web. Complementando a infraestrutura, sistemas de transcodificação convertem continuamente os vídeos para múltiplos formatos e bitrates, garantindo compatibilidade com diferentes dispositivos e condições de rede.

2. **Como a distribuição acontece neste sistema?**

A distribuição na Netflix ocorre em múltiplas camadas de forma integrada. Geograficamente, a rede Open Connect posiciona servidores de conteúdo estrategicamente próximos aos usuários finais, estabelecendo parcerias diretas com provedores de internet locais para reduzir ao máximo a latência. No nível arquitetural, o backend é fragmentado em centenas de microserviços especializados que operam de forma independente, permitindo que cada componente seja dimensionado e atualizado separadamente conforme a demanda. A distribuição de conteúdo segue uma lógica inteligente baseada em popularidade regional - os títulos mais assistidos em cada área são replicados nos servidores locais, enquanto conteúdos menos acessados ficam armazenados em datacenters centrais. Essa abordagem multi-camada permite que a Netflix atenda eficientemente a sua base global de assinantes, garantindo performance consistente mesmo durante picos de demanda em regiões específicas.

3. **Quais vantagens essa arquitetura distribuída oferece para o serviço?**

A arquitetura distribuída da Netflix proporciona benefícios estratégicos essenciais para seu modelo de negócios. A redundância inerente ao design distribuído garante alta

disponibilidade, onde falhas em componentes individuais ou mesmo em datacenters inteiros não comprometem a operação global do serviço. A proximidade física entre os servidores de conteúdo e os usuários finais resulta em baixa latência e experiência de visualização sem interrupções, com mínimos problemas de buffering. A escalabilidade elástica permite lidar com flutuações extremas de demanda, como os picos massivos durante lançamentos de séries aguardadas, através da rápida alocação de recursos adicionais. A resiliência do sistema é reforçada pela capacidade de isolar falhas e manter operações críticas mesmo durante problemas parciais. Além disso, a modularidade dos microserviços viabiliza atualizações contínuas e implantações graduais de novas funcionalidades sem a necessidade de downtime global, mantendo o serviço sempre disponível enquanto evolui.

4. Quais possíveis problemas podem surgir por ser um sistema distribuído?

A natureza distribuída da Netflix introduz desafios técnicos complexos que exigem soluções sofisticadas. Manter a consistência dos dados em um ambiente globalmente distribuído é particularmente desafiador, especialmente para informações como progresso de visualização que precisam ser sincronizadas em tempo real entre múltiplos dispositivos do usuário. A comunicação entre os diversos microserviços pode introduzir latência cumulativa, impactando o tempo de resposta geral do sistema. Falhas parciais, onde alguns componentes continuam operando enquanto outros apresentam problemas, requerem mecanismos específicos de detecção e recuperação para evitar efeitos em cascata. A complexidade do monitoramento aumenta exponencialmente com centenas de serviços distribuídos por múltiplas regiões geográficas, demandando sistemas avançados de observabilidade e análise. Problemas de concorrência podem surgir quando múltiplas requisições tentam acessar ou modificar os mesmos recursos simultaneamente, como no caso de acessos concorrentes a uma mesma conta de diferentes locais. Além disso, garantir a segurança de forma consistente em todos os componentes distribuídos representa um desafio contínuo.

5. Como o sistema parece lidar com a escalabilidade e tolerância a falhas?

A Netflix implementou estratégias inovadoras para garantir escalabilidade e tolerância a falhas em sua arquitetura distribuída. Para escalabilidade, o sistema emprega auto-scaling agressivo na AWS, onde microserviços podem automaticamente aumentar ou diminuir o número de instâncias em resposta à demanda em tempo real. Uma

hierarquia de caching sofisticada mantém os dados mais acessíveis rapidamente disponíveis em múltiplos níveis. Os bancos de dados utilizam técnicas avançadas de particionamento (sharding) para distribuir a carga de processamento de forma equilibrada. Na frente de tolerância a falhas, a Netflix é pioneira no Chaos Engineering, tendo desenvolvido ferramentas como o Chaos Monkey que testam proativamente a resiliência do sistema ao desativar componentes aleatoriamente em produção. Circuit breakers como o Hystrix previnem falhas em cascata ao isolar serviços problemáticos automaticamente. A replicação geográfica de dados críticos em múltiplas regiões assegura redundância e continuidade do serviço mesmo durante falhas significativas. Quando problemas ocorrem, o sistema é projetado para degradar graciosamente, mantendo funcionalidades essenciais operacionais enquanto serviços menos críticos podem ser temporariamente suspensos. Essa abordagem abrangente permite que a Netflix mantenha sua reputação de serviço confiável mesmo sob condições extremas.