



Uma Análise de Entrega de Vídeo Sob Demanda em Infraestrutura Borda-Nuvem

Daniel Ferragut Luiz

 $Luiz\ Fernando\ Bittencourt$

Relatório Técnico - IC-PFG-21-10 Projeto Final de Graduação 2021 - Julho

UNIVERSIDADE ESTADUAL DE CAMPINAS INSTITUTO DE COMPUTAÇÃO

The contents of this report are the sole responsibility of the authors. O conteúdo deste relatório é de única responsabilidade dos autores.

Uma análise de Entrega de Vídeo Sob Demanda em Infraestrutura Borda-Nuvem

Daniel Pereira Ferragut* Luiz Fernando Bittencourt[†]

Resumo

É previsto que 82% de todo tráfego de dados na Internet seja relacionado a visualizações de vídeos online como em serviços da Netflix, Youtube e Twitch. Para aumentar a qualidade de experiência (QoE) dos usuários assistindo esses vídeos, é possível utilizar técnicas como Adaptive Bitrate Streaming e DASH para sempre tentar prover o vídeo para o usuário e diminuir interrupções e tempo de espera no inicio de vídeos, a sacrifício da qualidade do vídeo sendo disponibilizado. Enquanto esse tipo de solução é feita em arquiteturas de nuvem centralizadas, também é possível utilizar Computação na Borda e trazer servidores menos poderosos mais próximos do usuários, diminuindo o congestionamento e tempo de latência dos pedidos feitos. Esse trabalho tem como objetivo analisar duas possíveis soluções para a melhora de QoE dos usuários, uma a base do uso de computação na borda com servidores cache próximos do usuários e outra sem isso, porém com velocidades maiores nos enlaces utilizados. Foram feitas diversas simulações no ns-3 em um cenário híbrido com infraestrutura borda-nuvem onde um grupo de usuários utilizava o cache na borda e o outro não, mas com velocidades mais altas. Foram testados casos para 100 e 120 usuários totais variando o aumento da velocidade de enlace para ser igual ao do grupo de cache e até 2 vezes maior. Os resultados mostram que o uso de cache na borda consegue atingir o QoE ótimo do cenário facilmente. O grupo sem cache precisou de um aumento de ao menos 50% em relação ao outro grupo para que tivesse valores similares. Além disso, o grupo sem cache teve mais casos discrepantes e usuários com experiências piores, devido a mais interrupções do vídeo pelo congestionamento da rede. Essa análise mostra que um investimento em uma infraestrutura de borda pode trazer benefícios instantâneos, como no caso analisado, assim como uma preparação para aplicações da próxima geração.

1 Introdução

Serviços de streaming de video, como Youtube, Netflix e Twitch, nunca foram tão populares. É previsto que até o final de 2021, 82% de todo trafego de dados na Internet seja relacionado a visualização de vídeos online [4]. Devido a grande competição entre esses serviços pela atenção de seus usuários, técnicas como o uso de computação na nuvem e de streaming de vídeo com taxa adaptativa de bits (Adaptive Bitrate Streaming) são usadas para que a experiência dos usuários seja a melhor possível, sendo o Dynamic Adaptive Streaming over

^{*}Instituto de Computação, Universidade Estadual de Campinas, 13081-970 Campinas, SP

[†]Instituto de Computação, Universidade Estadual de Campinas, 13081-970 Campinas, SP

HTTP (DASH) uma das maneiras mais utilizada por provedores de vídeo para implementar esse paradigma.

Em geral, arquiteturas de nuvem centralizadas em uma $Content\ Delivery\ Network(CDN)$ são utilizadas para prover vídeos. Isso é suficiente para prover uma boa experiência para vídeos sob demanda (VoDs) e transmissões ao vivo, mas para aplicações de vídeo da próxima geração essas arquiteturas centralizadas não são suficientes. Aplicações como streaming de vídeos de realidade virtual ou aumentada (VR/VA) e processamento de dados para condução autônoma de veículos precisam de um tempo de latência baixo com uma alta passagem de dados, o que não é possível quando o servidor está muito distante do usuário na nuvem.

A solução mais intuitiva desse problema seria o aumento da largura de banda dos usuários para viabilizar o uso dessas aplicações (ou melhor, a espera desse aumento). Isso, porém, é um processo lento, caro e depende muito dos provedores de internet de cada usuário, como pode ser visto na previsão da Cisco [4] em que a América Latina demoraria cerca de 5 anos para ter a mesma velocidade em Wi-Fi que a América do Norte possui atualmente. Uma solução mais imediata seria trazer o conteúdo mais próximo do usuário através de uma técnica que vem se tornando mais popular: computação na borda.

A ideia de uma infraestrutura na borda com a nuvem é trazer os servidores mais perto do usuário final. Para conseguir tempos de latência mais baixos e uma maior taxa de transferência de dados, servidores intermediários entre a nuvem e o usuário são utilizados. Esses servidores intermediários, então, poderiam servir como um cache para segmentos de vídeos, possibilitando um resposta rápida para o usuário, deixando a *CDN* mais livre para atender outros pedidos.

Visto o problema e essas duas possíveis soluções, esse projeto tem como objetivo fazer uma análise da eficiência entre o uso de uma infraestrutura borda-nuvem focada em cache e o aumento da banda de usuários, almejando encontrar as situações em que uma solução pode prevalecer sobre a outra. Para isto, foi utilizado o *Network Simulator 3* para simular casos de interesse onde seria possível analisar dois grupos de usuários que utilizavam cada uma das soluções, medindo a experiência de cada usuário através de uma medida de *Quality of Experience*, obtida com os dados da simulação.

Esse trabalho foi dividido de tal forma em que a seção 2 explica os conceitos necessários para entender o projeto, a seção 3 contém metodologia usada para avaliar as soluções, a seção 4 apresenta os resultados e discussão e a seção 5 possui a conclusão dessa análise.

2 Conceitos Básicos

2.1 Dynamic Adaptive Streaming over HTTP

A técnica de $Adaptive\ Bitrate\ Streaming(ABR)$ se tornou um padrão para transmissão de vídeos na Internet. A fundação desse paradigma é ter a mídia que será distribuída disponível em pequenas partes codificadas com diferentes algorítimos com níveis de qualidades e tamanho diferentes. Isso faz com que seja possível monitorar o uso da CPU e de banda do usuário e alternar as partes que estão sendo pedidas pelo cliente com o objetivo de sempre conseguir fornecer segmentos e não parar o vídeo para pegar novas partes (buffering), mesmo que sua qualidade piore. O uso de ABR, portanto, resulta em um tempo de início

para o vídeo menor, pouca espera para segmentos no meio do vídeo e uma qualidade de experiência aceitável para todos os tipos de conexões, sejam elas boas ou ruins.

O Dynamic Adaptive Streaming over HTTP(DASH ou MPEG-DASH) é uma solução pública padronizada pela Moving Picture Experts Group (MPEG) para ABR em HTTP. Os segmentos do vídeo sendo providos podem ser encontrados em arquivo de manifesto chamado de Media Presentation Description (MPD) no servidor hospedeiro, com todos os segmentos agrupados por diferentes codificações, contendo informações como a resolução do segmento, URL para baixar a parte e a taxa de bits(bit rate) necessária para usar o segmento.

O cliente, então, poderia pedir dinamicamente segmentos de diferentes tipos conforme o monitoramento de diversos fatores, como por exemplo a qualidade de experiência atual do usuário, uso CPU, tráfego na rede e nível de bateria do dispositivo. Em geral essa troca de tipos de segmentos é feita pelo cliente, mas nada impede que isso seja feito pelo servidor ou por uma rede gerenciada, com cada alternativa com suas vantagens e desvantagens [1].

2.2 Quality of Experience

Um grande desafio de provedores de vídeo é corretamente mensurar a satisfação de um usuário durante e após consumir a mídia que distribuem. Para resolver isso, a medida de Quality of Experience(QoE) foi criada para quantificar a qualidade da experiência do usuário ao assistir um vídeo. Essa é uma medida subjetiva, com muitos jeitos diferentes diferentes de calcular ela na literatura, sendo descobrir as melhores maneiras de avaliar a experiência em diferentes contextos uma própria área de pesquisa, como por exemplo Mehr et Al [8] que detalha o procedimento para calcular a medida em casos de mobilidade usando DASH ou $Ge\ e\ Wang\ [6]$ que especificam um QoE para casos na borda que usam DASH.

Para esse projeto, a definição de QoE é baseada em Zhang et al [13], que modela uma equação para a qualidade para pedaços dos vídeos em redes sem fio com cache, e em Bentalab et al [1], que detalha um equação para o QoE a partir da qualidade dos segmentos de vídeo e de Teoria dos Jogos em sistemas com DASH.

Primeiramente, a qualidade de cada segmento i de vídeo é calculada pela lei logarítmica sobre a taxa de bits, conforme [13] e definido na Equação 1 a seguir:

$$q(r_i) = a_1 log(\frac{a_2 r_i}{r}) \tag{1}$$

Onde r seria a taxa de bits necessária para assistir o segmento e r_i representa a taxa de bits utilizada para pegar o segmento para cada nível de codificação i. Já a_1 e a_2 são constantes positivas que dependem das categoria e funcionalidades do vídeo sendo assistido.

Com a qualidade q de cada segmento calculada, é então possível calcular a qualidade de experiência durante o vídeo de cada usuário u. Usando a equação estabelecida por [1], a partir de qualidade de cada segmento, o número de oscilações de qualidade, a quantidade de interrompimentos no meio do vídeo e do tempo espera para iniciar o vídeo, é possível calcular a qualidade de experiência. A Equação 2 descreve isso, sendo K o número de segmentos totais do vídeo, S_i a duração dos interrompimentos e ST_i o tempo de início do vídeo:

$$QoE_i = \frac{1}{K} \sum_{k=1}^{K} q(r_k) - \frac{1}{K-1} \sum_{k=1}^{K-1} |q(r_{k+1}) - q(r_k)| - \frac{1}{K} \sum_{k=1}^{K} S_k - ST_i$$
 (2)

O resultado QoE_i para o usuário i vai de 1 até 5, onde 1 seria uma experiência terrível, 2 uma ruim, 3 uma equilibrada, 4 uma boa e 5 uma ótima.

2.3 Computação na Borda e Cache

Até 2022, é esperado que o tráfego de dados mensais em dispositivos móveis suba para 77 exabytes, um aumento de 7 vezes comparado ao valor de 12 exabytes para 2017 [4]. Com mais usuários consumindo mais dados e a adoção de padrões como a Internet das Coisas (IoT) em que dispositivos de diversos tipos agora terão conexão a rede, é vital que haja uma infraestrutura que irá conseguir suportar uma essa imensa quantidade de máquinas e dados.

Para isso, o uso de computação na nuvem é essencial para garantir uma alta disponibilidade e elasticidade de recursos. Somente o uso desse paradigma, no entanto, não será suficiente. A latência causada pela distância entre a nuvem e o usuário, assim como a saturação de enlaces na rede devido ao alto tráfego de dados, fazem existem falhas nessa infraestrutura que precisam ser atendidas. Esses pontos podem ser consertados trazendo uma parte da computação em nuvem mais perto do usuário final, em uma solução chamada de: computação na borda/névoa.

A fundação dessa ideia é ter servidores intermediários mais fracos entre o usuário e a nuvem, mas ainda com as funcionalidades dela, como a disponibilidade e elasticidade de recursos. Isso, então, pode ser usado em infraestruturas de borda-nuvem, em que existem camadas de servidores entre o usuário e a nuvem, com o processamento sendo feito nos servidores intermediários sempre que possível, somente levando pedidos a nuvem quando necessário. Com o processamento mais perto do usuário, o tempo de latência é severamente reduzido e com menos pedidos para o servidor principal na nuvem, a rede fica mais balanceada, diminuindo saturações em enlaces.

Em um contexto de transmissão de vídeo, o processamento utilizado por esses servidores da borda é no cacheamento dos segmentos do *DASH*. Quando um usuário pede um novo segmento para a nuvem (uma *CDN* por exemplo), o pedido primeiramente passa no servidor da borda nas camadas inferiores, que verifica se ele já possuí o segmento, respondendo para o usuário rapidamente em caso de sucesso ou encaminhando o pedido para na camada superior em casos de ausência. Mais especificamente, nesse projeto será analisado esse cache na borda, porém ele só começará ser usado quando o enlace de um servidor com a camada superior começar a saturar devido ao alto tráfego de dados.

2.4 Network Simulator 3

O $Network\ Simulator\ 3$ (ns-3) [9] é um simulador de redes de computadores baseado em eventos discretos muito usado em pesquisas e como instrumento didático. Ele consegue simular sistemas na Internet com extrema acurácia através de eventos disparados pelos dispositivos virtuais conforme cada instante de tempo da simulação. O ns-3 tem suportes

a diversos protocolos e tipos de dispositivos, possibilitando a customização de diversos cenários, sendo que o usuário pode escrever aplicações para rodar nas máquinas simuladas ou até expandir o próprio projeto.

O ns-3, no entanto, não tem suporte nativo para aplicações DASH. Para garantir tal suporte, foi usado um repositório baseado na biblioteca de DASH libdash [3] e de uma implementação dessa biblioteca com ns-3 chamada de AMuSt-ns3 [7]. Esse repositório seria o dash-cloud-edge [5], que além de garantir suporte a DASH, também contém aplicações simulando computação da borda com cache em servidores intermediários, sendo possível facilmente alterar elas para atingir cenários de interesse a esse projeto.

3 Metodologia

O principal objetivo deste projeto é a análise de casos de cache em uma infraestrutura em borda-nuvem com casos de uma infraestrutura similar sem cache porém com velocidades mais altas nos enlaces. Para isso, foram feitas diversas simulações no ns-3 com um arquitetura híbrida como o da Figura 1, com usuários consumindo um vídeo $Full\ HD\ (1920\ x\ 1080\ pixels)$ de duração de 1600 segundos separados em 800 segmentos com um buffer de 60 para segmentos futuros, usando a mesma configuração para a representação desse vídeo como grandes serviços de vídeo (Netflix) fazem.

Nessa configuração, o vídeo está sendo provido pela CDN no nó 7 na nuvem, conectado ao nó 0 que serve de acesso a borda. Esse está conectado aos nós 1 e 2, que seriam os servidores na camada mais próxima aos usuários, com cada um conectando a dois pontos de acesso(AP) de Wi-Fi(3,4 e 5,6) para grupos de usuário que utilizam a rede sem fio. Os usuários com acesso a 3 e 4 vão ter uma experiência com cache na borda, enquanto os do 5 e 6 vão ter experiência tradicional sem nenhum tipo de cache.

A partir desse momento, fica definido que os nós do grupo com cache são os nós 1, 3 e 4 e os nós do grupo sem cache são o 2, 5 e 6. Também é atribuído o termo de enlaces do grupo com cache para os links dos APs com o nó 1 (1-3 e 1-4) e o link do nó 1 com o nó 0 (0-1), sendo os enlaces do grupo sem cache os dos APs com o nó 2 (2-5 e 2-6) e o do nó 2 com nó 0 (0-2).

O número de usuários para cada AP e a velocidade de cada links são configuradas no começo da simulação. O número de usuários em cada AP é igual, mas eles não começam todos a assistir o vídeo ao mesmo tempo. Cada usuário é atribuído um tempo de início através uma estratégia de Round-robin para cada AP. Uma distribuição de probabilidade de Poisson é utilizada para atribuir os inícios de forma crescente e cada tempo é alocado para um usuário alternando o AP que recebe o tempo para o usuário igualmente.

Enquanto os nós 0 e 1 tem suporte a cache, essa funcionalidade não é ativada instantaneamente na simulação. O cache desses nós é somente ativado quando o enlace com a camada inferior é saturado, ou seja, o trafego de dados atinge o máximo suportado pelo enlace. Quando isso acontece, nesse caso quando o enlace 0-1 satura, os pedidos de 3 e 4 começam a parar no nó 1, que usa o seu cache para responder rapidamente os pedidos de segmentos de vídeo.

Para realizar as comparações entre os dois grupos sendo simulados, foi fixado um número

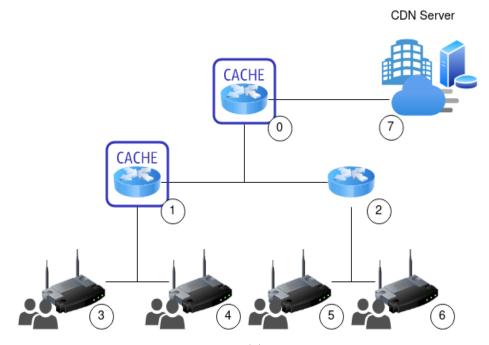


Figura 1: Arquitetura híbrida com uma CDN(7) que disponibiliza o vídeo a partir do acesso ao nó 0 que conecta um grupo com cache na borda-nuvem(1,3,4) e um grupo sem cache (2,5,6). Usuários pedem vídeo a partir dos APs (3,4,5,6)

de usuários em 25 para cada AP e também as velocidades de 0-7 (400 Mbps)e dos links do grupo de cache, com 0-1 tendo 20 Mpbs e 1-3 e 1-4 com 30 Mpbs. Então, foram testados diversos casos variando a banda dos enlaces do grupo sem cache, que inicialmente tem os mesmos valores que os nós espelhados no outro grupo e vai aumentando ela entre cada caso. Os resultados anotados seriam os dados de tráfego da mídia de todos os usuários a partir dos segmentos pedidos, com a qualidade de cada parte, assim como interrupções e tempo de espera. Após exaustar as opções de aumento com velocidade inicial, 125%, 150% e 200% dela, o número de usuários foi aumentado para 30 e os mesmo casos foram testados com um número de usuários maior por AP. Cada caso simulado para 25 e 30 usuários podem ser encontrado nas linhas da Tabela 1.

Tabela 1: Todos os casos testados para um número de usuários, cada linha é um cenário com velocidades diferentes para os links do grupo sem cache

Comparativo com a velocidade inicial	Link 0-2 (Mpbs)	Link 2-5 (Mpbs)	Link 2-6 (Mpbs)
100%	20	30	30
125%	25	$37,\!5$	37,5
150%	30	45	45
200%	40	60	60

4 Resultados

Os resultados gerais para os casos simulados podem ser encontrados na Figura 3, que contém os QoE médios de usuários conforme o tempo de início quando começa a consumir o vídeo, e na Figura 4, que contém o QoE médio para cada segmento do vídeo distribuído. Em ambas figuras, os gráficos a esquerda são de casos com 25 usuários e os da direta com 30 usuários, sendo que a velocidade dos enlaces do grupo sem cache vai aumentando dos gráficos de cima para os de baixo. Já as Tabelas 2 e 3 contém o QoE médio geral e para cada AP para 25 e 30 usuários, respectivamente.

Tabela 2: QoE geral e por AP para casos com 25 usuários conforme velocidade em relação a velocidade base

	AP 3	AP 4	AP 5	AP 6	Geral
100%	3.844 ± 0.312	3.838 ± 0.325	2.567 ± 0.643	2.491 ± 0.686	3.185 ± 0.836
125%	3.911 ± 0.335	3.916 ± 0.343	3.290 ± 0.605	3.208 ± 0.620	3.581 ± 0.592
150%	3.829 ± 0.292	3.827 ± 0.293	3.658 ± 0.454	3.552 ± 0.569	3.716 ± 0.492
200%	3.866 ± 0.304	3.870 ± 0.321	3.871 ± 0.296	3.849 ± 0.347	3.864 ± 0.313

Tabela 3: QoE geral e por AP para casos com 30 usuários conforme velocidade em relação a banda base

	AP 3	AP 4	AP 5	AP 6	Geral
100%	3.567 ± 0.370	3.426 ± 0.465	2.493 ± 0.687	2.169 ± 0.786	2.820 ± 0.912
125%	3.562 ± 0.375	3.429 ± 0.458	2.877 ± 0.562	2.675 ± 0.637	3.072 ± 0.660
150%	3.615 ± 0.400	3.488 ± 0.485	3.372 ± 0.529	3.230 ± 0.595	3.384 ± 0.545
200%	3.626 ± 0.381	3.482 ± 0.485	3.516 ± 0.420	3.492 ± 0.457	3.501 ± 0.444

A primeira coisa que se percebe na Figura 3 e Figura 4 seriam as duas curvas separadas representando os APs nos dois diferentes grupos. Como é possível ver na Figura 3a e na Figura 3b, essa diferença ajuda demonstrar o quanto o uso de cache na borda ajuda com o QoE dos usuários. Ambos grupos tem a mesma velocidade, mas o uso de computação da borda faz com que o grupo com cache tenha um QoE mais que um ponto maior que o grupo sem cache. Conforme o aumento de velocidade do grupo sem cache, essas curvas vão cada vez se aproximando, com a Figura 3c e 3d muito mais próximas e em 3e e 3f já não é mais facilmente perceptível a diferença entre os dois grupos. A principal diferença entre as curvas de 150% e de 200% seria a frequência dos valores discrepantes, com o de menor velocidade tendo mais valores fora do padrão e o de maior tendo valores mais próximos da curva.

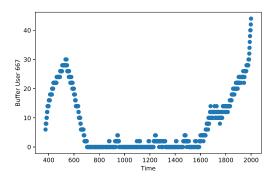
Pode-se ver também nos resultados de 150% e de 200% que o QoE do grupo sem cache não se altera muito, somente se tornando mais próximo dos valores obtidos do grupo com cache e diminuindo o desvio padrão. Isso se deve ao fato de existir um QoE máximo nessa simulação que depende também de outros valores além da banda do enlace, como a qualidade do vídeo provido e a quantidade de pedidos que um servidor consegue atender

simultaneamente.

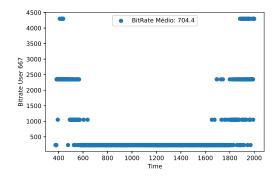
O grupo com cache consegue atingir esse QoE com facilidade, enquanto é necessário um aumento de pelo menos 50% no grupo sem cache para atingir um valor próximo ao ótimo. Pela Figura 4, é possível inferir que o aumento de velocidade que iguala a experiência dos dois grupos é seria em torno de 50% a 100 % da velocidade base, já que as curvas dos APs estão muito próximas no 150%, porém no 200% as linhas estão totalmente sincronizadas.

As diferenças de QoE iniciais para 25 e 30 usuários encontradas nas Tabelas 2 e 3 era esperada, já que com mais demanda, há um tráfego maior de dados e consequentemente mais problemas para pegar os segmentos necessários, fazendo com os casos com mais usuários tenham uma pior experiência. O crescimento de QoE entre esses dois conjuntos de casos é similar, com grupo de 30 usuários tendo um crescimento ligeiramente melhor devido a um maior espaço de melhora e também ao fato do aumento de banda afetar mais usuários. O grupo com 30 usuários, no entanto, tem mais casos discrepantes do que o grupo de 25 usuários, uma consequência de um tráfego maior na rede e espera para ter os pedidos de segmentos atendidos.

Os casos discrepantes encontrados no grupo sem cache possuem comportamentos muito parecidos. A Figura 2 mostra o buffer e taxa de bits de um usuário discrepante no AP 6 para o caso de 30 usuários por AP com 150% da velocidade base. Essa figura consegue exemplificar bem o grupo de casos discrepantes para todos os cenários observados. Em geral, esse tipo de usuário começa assistir o vídeo e recebe os segmentos normalmente, explicando a taxa de bits mais alta no começo e os segmentos armazenados. Porém, com mais usuários entrando na rede, esse usuário fica com a rede congestionada, não conseguindo ter respostas para o seus pedidos de segmento, reduzindo sua taxa de bits e, consequentemente, um buffer e QoE menor devido a um maior número de interrupções. Com mais pessoas saindo da rede, esse usuário consegue mais acesso a rede e volta a ter uma taxa de bits alta e, portanto, mais segmentos armazenados. Para usuários não discrepantes, os valores de buffer e de taxa de bits são mais altos e estáveis, o que garante um QoE maior.



(a) Segmentos armazenados para uso posterior (buffer) de um usuário assistindo o vídeo conforme o tempo



(b) Taxa de bits (bitrate) de um usuário assistindo o vídeo conforme o tempo

Figura 2: Buffer e taxa de bits de um usuário discrepante assistindo o vídeo conforme o tempo

O grupo com cache possuí bem menos casos discrepantes do que o grupo sem cache, mesmo com uma maior velocidade maior. Isso se deve as vantagens de usar servidores intermediários entre a borda e a nuvem, com menos congestionamento da rede e tempos de respostas mais baixos, causando menos usuários com espera em seus pedidos de segmentos. Já sobre o mecanismo de saturação da rede para ativação do cache, em todos os casos ele foi ativado bem nos momentos iniciais da simulação, indicando que a rede já começa a saturar nos segundos iniciais quando os usuários do grupo de cache começam a entrar na rede, portanto os resultados obtidos para esse grupo são basicamente com o cache ativado em toda a simulação.

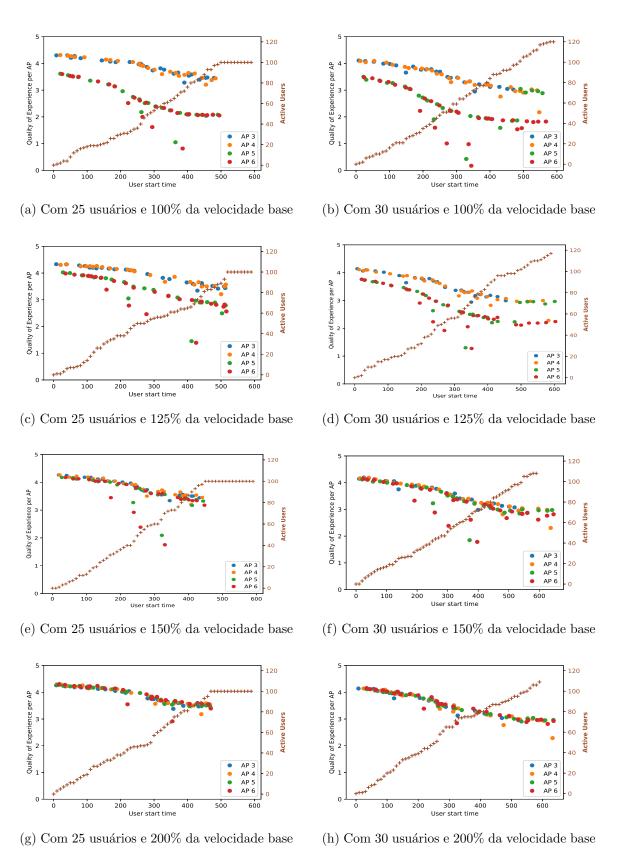


Figura 3: QoE para cada usuário separado por AP conforme data de início do vídeo

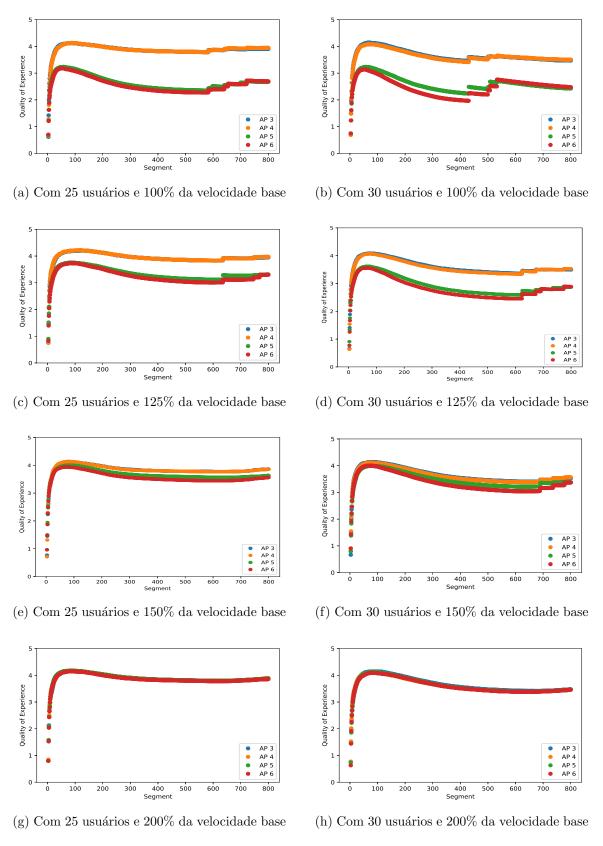


Figura 4: QoE médio para cada segmento do vídeo disponibilizado para os usuários

5 Conclusão

Através de simulações de diversos casos utilizando o ns-3 em cenários de uma infraestrutura borda-nuvem, foi possível comparar a efetividade da solução de cache na borda com a de um aumento da banda dos usuários. Seja para 25 ou 30 usuários para cada ponto de acesso, a solução com cache atingiu o QoE ótimo facilmente, enquanto o grupo sem cache precisou de um aumento de 50% da banda dos enlaces para que esse valor chegasse próximo ao do grupo com cache. Mesmo com esse aumento de banda, a solução sem cache teve mais casos discrepantes de usuários com piores qualidades de experiência devido a um maior congestionamento da rede e do servidor, causando menos segmentos sendo servidos e mais interrupções no vídeo.

Elaborando o problema analisado na América Latina, a Cisco [4] prevê que um aumento de 50% na velocidade de conexões Wi-Fi demoraria cerca de 2 anos no continente. Esse seria o tempo médio que um serviço de vídeos teria que esperar no continente para sua solução seja tão eficiente quanto o uso de cache na borda. Enquanto o investimento em uma infraestrutura na borda possa ser desafiante, há benefícios imediatos claros que essa análise demonstra, assim como um investimento a longo prazo para problemas das próximas gerações que irão precisar dos tempos de latência baixos fornecidos por esse paradigma. Olhando para o futuro com a adoção de 5G e suas velocidades muito mais altas com uma necessidade de uma maior número de antenas, é possível que serviços consigam deixar a infraestrutura de borda de lado em áreas que possuem um bom acesso a essas altas velocidades, como em grandes centros, mas para áreas com menor acesso, como nos arredores das cidades, uma solução na borda pode ser usada para manter a qualidade de experiência.

Para trabalhos futuros, essa análise pode ser estendida em diferentes vertentes. Uma delas seria o uso de 5G ao invés de Wi-Fi como meio físico dos usuários, utilizando uma infraestrutura LTE-EPC para suportar isso. Também pode ser possível analisar diferentes vídeos com maiores qualidades como 4K e 8K, assim como vídeos da próxima geração, como streaming de realidade virtual. Por fim, a mobilidade de usuário é algo que precisa ser considerado, podendo ser realizada uma análise onde usuários movem de um AP para outros e as consequências do QoE para esses usuários.

6 Agradecimentos

Esse projeto não seria possível sem o apoio do grupo de pesquisa de *Adapdative Bitrate Streaming* em *DASH* com Computação na Borda da Unicamp/IC: composta também por Roger Immich, Eduardo Gama e Lucas Araujo.

Referências

[1] Abdelhak Bentaleb et al. "A Survey on Bitrate Adaptation Schemes for Streaming Media Over HTTP". Em: *IEEE Communications Surveys Tutorials* 21.1 (2019), pp. 562–585. DOI: 10.1109/COMST.2018.2862938.

- [2] Abdelhak Bentaleb et al. "Want to Play DASH?: A Game Theoretic Approach for Adaptive Streaming over HTTP". Em: *Proceedings of the 9th ACM Multimedia Systems Conference*. MMSys '18. Amsterdam, Netherlands: ACM, 2018, pp. 13–26. ISBN: 978-1-4503-5192-8. DOI: 10.1145/3204949.3204961. URL: http://doi.acm.org/10.1145/3204949.3204961.
- [3] Biblioteca libdash. Accessed 27-jun-2021. URL: https://github.com/bitmovin/libdash.
- [4] Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2017-2022. Accessed 27-jun-2021. 2017. URL: https://cyrekdigital.com/uploads/content/files/white-paper-c11-741490.pdf.
- [5] E.S Gama et al. "Video Streaming Analysis in Multi-tier Edge-Cloud Networks". Em: 2021 IEEE 8th International Conference on Future Internet of Things and Cloud (FiCloud). 2021.
- [6] Chang Ge e Ning Wang. "Real-time QoE estimation of DASH-based mobile video applications through edge computing". Em: IEEE INFOCOM 2018 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). 2018, pp. 766–771. DOI: 10.1109/INFCOMW.2018.8406935.
- [7] Christian Kreuzberger, Daniel Posch e Hermann Hellwagner. AMuSt Framework Adaptive Multimedia Streaming Simulation Framework for ns-3 and ndnSIM. 2016.
- [8] Sheyda Kiani Mehr, Prasad Jogalekar e Deep Medhi. "Moving QoE for monitoring DASH video streaming: models and a study of multiple mobile clients". Em: *Journal of Internet Services and Applications* 12.1 (abr. de 2021). DOI: 10.1186/s13174-021-00133-y. URL: https://doi.org/10.1186/s13174-021-00133-y.
- [9] Network Simulator 3 (ns-3). Accessed 27-jun-2021. URL: https://www.nsnam.org/.
- [10] P.F Prado et al. Mobile Edge Computing for Content Distribution and Mobility Support in Smart Cities. In: Anwesha Mukherjee, Debashis De, Soumya K. Ghosh, Rajkumar Buyya. (Org.). Mobile Edge Computing. Springer, 2021, pp. 1–30.
- [11] E. S. Gama, R. Immich e L. F. Bittencourt. "Towards a Multi-Tier Fog/Cloud Architecture for Video Streaming". Em: 2018 IEEE/ACM International Conference on Utility and Cloud Computing Companion (UCC Companion). Dez. de 2018, pp. 13–14. DOI: 10.1109/UCC-Companion.2018.00022.
- [12] Karima Velasquez et al. "Fog orchestration for the Internet of Everything: state-of-the-art and research challenges". Em: Journal of Internet Services and Applications 9.1 (jul. de 2018), p. 14. ISSN: 1869-0238. DOI: 10.1186/s13174-018-0086-3. URL: https://doi.org/10.1186/s13174-018-0086-3.
- [13] Weiwen Zhang et al. "QoE-Driven Cache Management for HTTP Adaptive Bit Rate Streaming Over Wireless Networks". Em: *IEEE Transactions on Multimedia* 15.6 (2013), pp. 1431–1445. DOI: 10.1109/TMM.2013.2247583.