

Uma Proposta para Gerenciamento do Tráfego do Protocolo Dropbox LAN Sync em Redes Definidas por *Software*

Gustavo Diel¹, Guilherme P. Koslovski¹

¹LabP2D – Departamento de Ciência da Computação – UDESC

Resumo. A ferramenta *Dropbox* é amplamente utilizada por usuários domésticos, empresas e universidades para sincronização de arquivos entre dispositivos locais e repositórios na nuvem. Para acelerar a transferência de arquivos, *Dropbox* utiliza o protocolo *LAN Sync*. Todavia, o protocolo é impossibilitado de atuar em diversas sub-redes. O presente trabalho utiliza *Redes Definidas por Software* para atenuar o problema.

1. Introdução

O *Dropbox*, cliente de sincronização de arquivos, oferece diversas ferramentas e recursos para seus usuários, sendo utilizado em diversos ambientes. Por exemplo, *Dropbox* é aplicado como ferramenta de suporte para o desenvolvimento de aplicações paralelas [Bretana et al. 2018]. Dentre os serviços implementados pela ferramenta, o protocolo *LAN Sync* possui destaque por permitir que a infraestrutura local de rede possa ser utilizada para a transferência de arquivos, sem efetuar o *download* de repositórios da nuvem.

LAN Sync procura possíveis clientes na rede que possuam arquivos compartilhados. Assim que encontrados, os clientes fazem uma requisição ao servidor do *Dropbox* para descobrir se algum arquivo está atualizado. Caso afirmativo, é iniciada uma conexão direta entre ambos. Porém, como o protocolo utiliza pacotes *broadcast* para descobrir quais clientes estão aptos a fazer uma conexão, este está limitado em uma única sub-rede.

Para atenuar o problema, o presente trabalho propõe a utilização de *SDN* (*Software Defined Network* [McKeown et al. 2008, Diel et al. 2017]). Por se conectar logicamente com todos os dispositivos da rede, o controlador *SDN* tem visão total do ambiente, podendo repassar informações de *broadcast* *LAN Sync*, quando necessário.

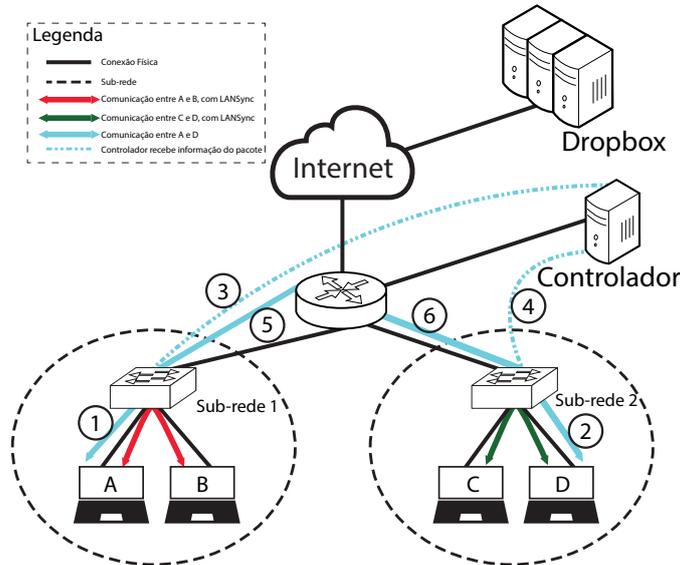
2. Propostas de Implementação

O controlador *Floodlight*, *open-source* e desenvolvido em *Java*, permite que extensões sejam instaladas em seu código diretamente [Floodlight 2011]. Essas extensões permitem modificar o comportamento de diversos recursos *SDN*, criando regras para a rede ou para o próprio controlador. O presente trabalho propõem duas formas de implementação para o ambiente do *Floodlight*. A Figura 1 é utilizada no decorrer do trabalho para facilitar as explicações das propostas.

2.1. Algoritmo geral

A sub-seção a seguir apresenta parte do algoritmo que ambas as soluções compartilham. As regras de configuração de fluxos devem ser instaladas diretamente no código do controlador, nas funções que processam a chegada de pacotes. Sempre que detectar um pacote *LAN Sync*, o controlador deve executar o módulo do código que possui a lógica principal.

Figura 1. Ilustração das propostas de implementação.



A lógica consiste de duas partes, a primeira para pacotes LAN Sync de reconhecimento e a outra para pacotes LAN Sync de transmissão de dados.

Quando um pacote de reconhecimento for detectado, a ferramenta deve replicar esse pacote para as outras sub-redes presentes. Como por exemplo, originando da Sub-rede 1 e sendo replicado para a Sub-rede 2 na Figura 1. É importante notar que cada sub-rede pode possuir uma configuração de endereços IP diferente. Para contornar esse possível problema, a ferramenta deve gerar IPs falsos para cada sub-rede para cada cliente Dropbox único. Ao detectar um pacote de transmissão de dados, a ferramenta deve calcular qual será a trajetória do pacote pela rede, detectar quais IPs deverão ser substituídos e então enviar essas regras aos dispositivos de rede. Quanto a camada de enlace, para que os clientes possam preencher a tabela ARP com os devidos IPs para cada endereço MAC falso, a aplicação também deve vasculhar a rede por requisições ARP de MACs falsos, e respondê-los de acordo.

2.2. Utilizando o Controlador SDN para Transferência

É possível utilizar os fluxos que o próprio controlador gera (linhas tracejadas 3 e 4 na Figura 1) para facilitar seu funcionamento. Tais fluxos podem ser utilizados como o meio de roteamento dos pacotes, facilitando a forma de realizar o roteamento e replicação dos dados conforme necessário. É importante notar, que toda vez que algum pacote precisar utilizar esse método, todo o código deve ser executado e, conseqüentemente, todas as informações contidas no pacote vão ser repassadas ao controlador, gerando um fluxo de dados maior no enlace do mesmo.

2.3. Utilizando o Regras de Fluxo para Transferência

O método que utiliza o conceito de fluxo SDN é executado toda vez que uma nova conexão é detectada como LAN Sync (uma vez para enviar os pacotes de reconhecimento e outra para fazer a conexão direta com algum outro cliente). A ferramenta então gera

um cadastro na tabela de fluxos de cada *switch* de rede que pertence ao caminho previamente calculado. Ainda, uma regra indica que certos dados dos cabeçalhos dos pacotes sejam alterados, tais como endereços, e informações de protocolos. É possível notar que diferente do método proposto na Sub-seção 2.2, o presente algoritmo é executado apenas duas vezes para cada conexão LAN Sync, reduzindo assim o uso de rede, de recursos do sistema e não precisa de que todos os dados do pacote sejam previamente transmitidos ao controlador para aplicar as regras.

2.4. Discussão

É importante ressaltar as diferenças entre as duas técnicas. A segunda, que cria regras de fluxos para o roteamento e realiza a inserção das regras nos dispositivos é executada duas vezes por conexão; sendo a primeira vez para aplicar a regra de retransmissão dos pacotes de busca e a segunda para aplicar as regras de roteamento e mudança de IP. Uma das vantagens é que a utilização de recursos computacionais é menor, tornando a solução mais resiliente em ambientes limitados.

Por sua vez, a técnica que utiliza o controlador como meio de roteamento, necessita que todos os pacotes LAN Sync destinados a outras redes sejam retransmitidos ao controlador. Esta retransmissão dos pacotes que vai do dispositivo origem, ao controlador e então ao dispositivo destino é responsável por um consumo maior de recursos da infraestrutura, inviabilizando seu uso em larga escala. Porém uma vantagem que a técnica que utiliza o controlador tem em relação ao que utiliza fluxos é a flexibilidade de aplicar regras não disponíveis nativamente em um ambiente SDN puro. Como o controlador é executado a partir do código Java que apenas realiza chamadas aos diversos dispositivos, é possível aplicar regras externas aos pacotes da rede.

3. Resultados Preliminares

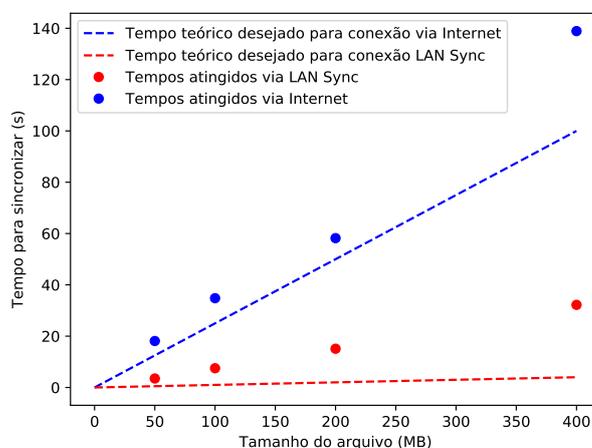
Para quantificar a diferença no uso real que utilizar o LAN Sync apresenta, testes preliminares foram realizados em uma topologia dividida em duas sub-redes. Para a realização dos testes, foi utilizada a ferramenta Mininet, capaz de simular diversas topologias de rede em um computador virtual com 2 núcleos, 4GB de RAM executando Ubuntu 16.04. O controlador SDN foi executado em uma máquina com 2 núcleos e 8GB de RAM, também com o sistema Ubuntu 16.04. Cada teste foi executado 10 vezes, e os resultados foram obtidos com uma média simples de todos os testes.

Uma sub-rede possui um par de clientes para testar a transmissão do protocolo, enquanto a outra sub-rede possui apenas um cliente para realizar a conexão com a outra sub-rede, utilizando a Internet. Os resultados podem ser observados na Figura 2. É possível notar que os desvios padrões estão pequenos demais para serem legíveis no gráfico. Para as duas curvas tracejadas, as seguintes fórmulas foram utilizadas: A primeira é $f_{internet}(x) = \frac{x}{4}$ que representa a curva de tempo para a conexão via Internet. A segunda é $f_{lansync}(x) = \frac{x}{1000}$ que representa o tempo para a conexão utilizando LAN Sync.

4. Considerações Finais

Os serviços de compartilhamento de arquivos que utilizam nuvens computacionais para realizar o armazenamento de dados ganharam espaço no cotidiano de seus usuários, sendo

Figura 2. Comparação entre sincronização via Internet e com LAN Sync.



utilizado inclusive como suporte para aplicações distribuídas. O Dropbox é um destes serviços e além de ferramentas de gerenciamento, conta com um protocolo para otimizar a transferência em redes locais. O LAN Sync oferece vantagens ao otimizar o *download* de conteúdo localmente. Por conta das limitações das redes convencionais, o protocolo está limitado à uma única sub-rede. Em decorrência desta limitação, o presente trabalho apresentou duas propostas usando SDN permitir o funcionamento do LAN Sync em uma rede segmentada. Utilizando SDN, é possível construir uma ferramenta que aprenda qualquer topologia da rede e possa desempenhar suas funções com o mínimo de configuração necessário.

Para trabalhos futuros, é interessante testar a capacidade da ferramenta de ser utilizada em outros protocolos além do LAN Sync, assim como testar a escalabilidade do sistema em execuções simultâneas. Também é interessante uma análise entre a quantidade de dados compartilhados com o uso dos enlaces de rede e a latência para descobrir possíveis gargalos no sistema.

Agradecimentos: UDESC, FAPESC e LabP2D.

Referências

- Bretana, L., Schwendler, A., and Cavalheiro, G. G. H. (2018). Computação distribuída: Desafios do uso do dropbox como suporte ao espaço de tuplas. In *Escola Regional de Alto Desempenho do Estado do Rio Grande do Sul*. SBC.
- Diel, G., Marcondes, A., and Koslovski, G. (2017). Uma ferramenta para evitar enlaces congestionados em redes definidas por software. In *Escola Regional de Alto Desempenho do Estado do Rio Grande do Sul*. SBC.
- Floodlight (2011). Project floodlight. <http://www.projectfloodlight.org/floodlight/>.
- McKeown, N., Anderson, T., Balakrishnan, H., Parulkar, G., Peterson, L., Rexford, J., Shenker, S., and Turner, J. (2008). Openflow: Enabling Innovation in Campus Networks. *ACM SIGCOMM Computer Communication Review*, 38(2):69–74.