Instituto de Informática (INF)
Universidade Federal do Rio Grande do Sul (UFRGS)
Porto Alegre - RS
URL: https://www.inf.ufrgs.br

Orientador: Prof. Dr. Lisandro Zambenedetti Granville
Coorientador: Leandro Bertholdo
Grupo de Redes de Computadores
E-mail: granville@inf.ufrgs.br
leandro.bertholdo@ufrgs.br

Comunidades BGP: análise de adoção e criação de dataset

1 Introdução e Motivação

O Border Gateway Protocol (BGP) é um protocolo vital para o funcionamento da Internet, permitindo que informações de roteamento sejam trocadas entre sistemas autônomos. As comunidades BGP são um recurso importante deste protocolo, permitindo que os operadores de rede implementem engenharia de tráfego [7], detectem a desconexão de partes da Internet [8], localizem geograficamente uma rede [5], ou permitam descobrir quais redes estão sobre ataques DDoS [9]. Adicionalmente, as comunidades BGP também podem ser exploradas para realizar diversos ataques na Internet [11].

Embora existam padrões de comunidades BGP bem definidos (*well-known communities* [6]), estes padrões tem uma adoção desconhecida na Internet. A maioria dos sistemas autonomos (ASes) determinam comunidades privadas para essas funcionalidades [4], ou as implementam de forma incorreta [2]. Ambas as situações dificultam automação do roteamento BGP por software (Sofware Defined Networks) [1] [10]. A automação de roteamento BGP é especialmente importante para redes de distribuição de conteúdo (CDNs).

Aplicações como BGPTunner [1] e Anycast Agility[10] necessitam de uma base de dados de comunidades BGP para melhorar o roteamento BGP para redes anycast (ex. servidores de nomes globais) e reagir a ataques DDoS.

Adicionalmente, o IETF discute padronizar a implementação de algumas comunidades BGP *well-known* [3] diretamente nos roteadores. Hoje em dia, grande parte dos provedores de transito utilizam comunidades privadas implementando funções semelhantes as *well-known BGP communties*). É importante ter um método capaz de mensurar a adoção e impacto de tal mudança pela ao longo dos anos.

Este trabalho se propõe a criar um dataset de comunidades BGP utizadas por

provedores de transito. Os dados das comunidades BGP podem ser obtidos a partir daquelas divulgadas em tabelas de roteamento globais, de documentação disponibilizada na Internet (webscrapping), e outros metodos. O dataset resultante deve agrupar as comunidades BGP por funcionalidades, a partir da visão de tabelas de rotas de coletores de rotas como RIPE RIS, routeviews, e servidores de rotas em internet exchanges (IXPs).

2 Descrição das Atividades

As atividades a serem desenvolvidas durante o TG1 e TG2 são descritas abaixo:

A1. **Estudo de conceitos de roteamento BGP:** Nessa atividade o aluno deve se familiarizar com o uso do protocolo BGP, atributos e funcionalidades.

A2. **Uso de comunidades BGP para engenharia de trafego:** Aqui o aluno deve se familiarizar com o uso de comunidades BGP aplicado a engenharia de tráfego, e identificar os usos mais comuns.

A3. **Revisão bibliográfica:** Levantamento de Datasets de roteamento BGP disponiveis, Idenficação de técnicas de Ciência de dados para clusterização e visualização, técnicas de webscrapping e coleta de informações, visualização de grandes volumes de dados.

A4. **Definição e delimitação de escopo da proposta:** Esta etapa compreende a definição dos requisitos e objetivos que deverão ser atingidos. O trabalho deve considerar pesquisas recentes e propor inovação.

A5. **Design da Solução:** A solução deve ser projetada de forma flexível e genérica o suficiente para que possa ser implementada posteriormente em diferentes tecnologias. A implementação inicial será usada para demonstrar que a solução atende aos requisitos e objetivos pretendidos.

A6. **Avaliação de resultados:** A qualidade do Dataset gerado deve ser comparada a outros Datasets similares quando aos ganhos em relação a completeza, confiabilidade, e outras melhorias obtidas (ex. tempo de computação).

A7. **Escrita do TCC:** Essa atividade é continua e deverá ser realizada ao longo de todo o trabalho conforme discussao com os orientadores.

As atividades deverão ser realizadas ao longo de 1 ano durante as cadeiras de Trabalho de Graduação (TG) 1 e 2. As atividades de escrita deverão ocorrer em paralelo com o desenvolvimento prático do trabalho. Sugere-se que os seguintes capítulos estejam concluídos ao final de cada uma das cadeiras:

- **TG1**: Introdução e Motivação, Referencial Teórico e Trabalhos Relacionados;
- **TG2**: Abordagem, Implementações, Avaliações e Conclusões.

Porém, tais atividades e entregas poderão ser alteradas se acordadas com os orientadores para um melhor desenvolvimento do trabalho e adequação a metas/resultados do trabalho.

3 Requisitos

- Interesse em Roteamento BGP;
- Interesse por medições na Internet e ciência de dados;

References

[1] L. M. Bertholdo, J. M. Ceron, L. Z. Granville, G. C. Moura, C. Hesselman, and R. Van Rijswijk-Deij, "BGP anycast tuner: Intuitive route management for anycast services," 16th International Conference on Network and Service Management, CNSM 2020, 2nd International Workshop on Analytics for Service and Application Management, AnServApp 2020 and 1st International Workshop on the Future Evolution of Internet Protocols, IPFutu, 2020.

[2] L. M. Bertholdo, J. M. Ceron, W. B. Vries, R. D. O. Schmidt, L. Z. Granville, R. V. Rijswijk-Deij, and A. Pras, "TANGLED: A Cooperative Anycast Testbed," Proceedings of the IM 2021 - 2021 IFIP/IEEE International Symposium on Integrated Network Management, May 2021, pp. 766–771.

[3] J. Borkenhagen, R. Bush, R. Bonica, and S. Bayraktar, "Policy Behavior for Well-Known BGP Communities," Internet Requests for Comments, RFC Editor, Tech. Rep. 8642, 2019. [Online]. Available: https://www.rfc-editor.org/rfc/rfc8642.txt

[4] Brasil Peering Forum, "Lista de communities bgp wiki bpf," https://wiki.brasilpeeringforum.org/w/Lista_de_Communities_BGP, 12 2022, (Accessed On 2/5/2023 16:0).

[5] CAIDA, "Bgp community dictionary dataset caida," https://www.caida.org/catalog/datasets/bgp-communities/, 06 2021, (Accessed On 3/5/2023 10:50).

[6] R. Chandra, P. Traina, and T. Li, "BGP Communities Attribute," Internet Requests for Comments, RFC Editor, Tech. Rep. 1997, 1997. [Online]. Available: https://www.rfc-editor.org/rfc/rfc1997

[7] B. Donnet and O. Bonaventure, "On bgp communities," SIGCOMM Comput. Commun. Rev., Vol. 38, No. 2, p. 55–59, mar 2008. [Online]. Available: https://doi.org/10.1145/1355734.1355743

[8] V. Giotsas, C. Dietzel, G. Smaragdakis, A. Feldmann, A. Berger, and E. Aben, "Detecting peering infrastructure outages in the wild," SIGCOMM 2017 - Proceedings of the 2017 Conference of the ACM Special Interest Group on Data Communication, 2017, pp. 446–459.

[9] V. Giotsas, S. Zhou, M. Luckie, and K. Claffy, “Inferring multilateral peering,” CoNEXT 2013 - Proceedings of the 2013 ACM International Conference on Emerging Networking Experiments and Technologies, 2013, pp. 247–258.

[10] A. S. Rizvi, L. Bertholdo, J. Ceron, and J. Heidemann, “Anycast Agility: Network Playbooks to Fight DDoS,” Proceedings of the 31st USENIX Security Symposium, Security 2022, pp. 4201–4218, 2022.

[11] F. Streibelt, F. Lichtblau, R. Beverly, A. Feldmann, C. Pelsser, G. Smaragdakis, and R. Bush, “BGP Communities: Even More Worms in the Routing Can,” Proceedings of the Internet Measurement Conference 2018, ser. IMC ’18. New York, NY, USA: Association for Computing Machinery, 2018, pp. 279–292. [Online]. Available: https://doi.org/10.1145/3278532.3278557