CLASSIFICATION OF TOPOLOGY FOR INTERNET AUTONOMOUS SYSTEMS: AN ANALYSIS OF THE IMPLEMENTATION OF INTERNET AUTONOMOUS SYSTEMS.

Cao Yuanqing1, Midhun Chakkaravarthy1

[1] Lincoln University College, Petaling Jaya, Malaysia.

*Corresponding author: Cao Yuanqing, Lincoln University College, Petaling Jaya, Malaysia.

ABSTRACT

Data management and data routing over global networks are getting more challenging as the internet continues to expand at an exponential rate. This is the present state of affairs due to the ease of connecting to the internet. On the Internet, Autonomous Systems (AS) are decentralised networks that adhere to standard routing protocols, under the supervision of various organisations. Distributed networks, or AS, make this design possible. The system's most crucial parts are these auxiliary ones. In this study, researcher examine the impact of different topological topologies on the performance and operation of Internet AS in great detail. This study investigates the issue of topology classification inside these systems thoroughly by looking at all the many topological structures. This study will classify the currently used AS topologies based on their operational features, fault tolerance, scalability, and routing efficiency. Some of the many topologies included in this category are hybrid, ring, star, mesh, and ring, among many more. Data transfer rates, network reliability, and routing stability are the three most important performance metrics considered while building an AS. Researchers consider each and every one of these distinct aspects. Such topologies facilitate the execution of more fruitful implementation operations. With the information this project provides, network engineers and legislators will be better equipped to choose scalable and reliable alternative service deployment topologies. Incorporating this kind of classification and analysis into internet infrastructure architecture can enhance connectivity, speed, and resilience. In a world where everything is becoming digitised, this software will be invaluable. This approach aids in bettering the internet's architecture, among its many advantages.

Keywords: Topology, Internet, Autonomous Systems, Border Gateway Protocol, System.

INTRODUCTION

Studying the Internet is an intriguing topic for network engineers and scholars. The Internet's topology is not controlled by the networks that comprise it. Because of this, no one can provide us with a comprehensive map of the Internet at this time. Many people working on network protocols would find this knowledge useful; therefore, it's a major obstacle. Capturing one aspect of the Internet's architecture more easily than the rest would be a great achievement.

The Internet's AS have created this topology. The Internet makes use of these network elements while routing data across different domains. The National Laboratory for Applied Network Research has a significant amount of data on routing between different domains. Researchers offer a follow-up analysis of previous work, and this data forms the basis of other investigations. Using six instances of BGP data (from November 1997 to May 2000), researchers have computed several average AS network attributes from distributions pertaining to degree, distance, number of shortest routes, trees, etc. Some of these can be succinctly defined by power laws, according to our findings. Researchers have also derived certain empirical rules and examined the changes in these average qualities over time. Many autonomous systems, including thousands of cases managed by different administrative bodies, can be linked thanks to the Internet's connectivity. The BGP protocol regulates the behaviour of ASes during their interactions. BGP allows each AS to independently determine the routes to use for imports and exports from neighbouring systems. The restrictions are decided upon by the network administrators and are driven by AS relationships, which may be considered a form of commercial contract between ASes. There are many other kinds of peer-to-peer (P2P) connections, but two of the most prevalent are P2P and P2C partnerships. When a service provider facilitates data transmission between the Internet and other networks, the consumer bears the responsibility of paying for this service. The unfettered interchange of traffic between ASes and their clients is conceivable, but the exchange between ASes and their suppliers or other peer-respected entities is not. The prevalent economic paradigm dictates that ASs should not share their provider or peer routes with other ASs or providers. Such an arrangement is the inevitable result of the economic paradigm they're trying to change. Gaining a solid grasp of the structure, inter-domain routing dynamics, and expansion of the Internet requires an understanding of the commercial links that exist between application servers (ASes) (Gherari et al., 2023).

BACKGROUND OF THE STUDY

It's difficult to tell it that, but domain prefix theft constantly happens on the internet. Prefix hijacking can occur for a variety of reasons, including intentional or inadvertent BGP route configuration. The latter can cause big difficulties, including service outages, privacy violations, and big money losses. If these hijackings affect even a small number of ASes in a certain location, it may be because they affect all ASes on Earth. The BGP upgrade implemented in response to PCCW Global has now spread to the global internet, per ASC 4391 (AS3491). Pakistan Communications Corporation used this technique to fool people in several countries that watched YouTube. The move made the siege last two more hours. In 2010, when China Net (AS23724) published all of its unassigned prefixes, it was almost as big as the total world routing database. China Telecom's network, called China Net, usually makes and gives out forty different prefixes. It's a good thing that these prefixes are only available to a small number of networks outside of China. In April 2010, a China Telecom AS hijacked more than 50,000 prefixes, which caused 15% of Internet traffic to go to the wrong place for fifteen minutes. Similar attacks happened while CDTDBC (Companhia de Telecommunicadores do Brasil

Central AS16735) showed its whole prefix database to its upstream providers for about five minutes. During the 2018 3ve takeover, which was a planned crime that took $29 million in phoney ad revenue, ads for actual companies, including the US Air Force, aired. Planes can be hijacked by accident, but it's more likely to be intentional. Regrettably, researchers allowed this illegal activity to continue for a year. The hijacking happened because more than 1.5 million IP addresses were taken. Errors in the BGP traffic routing protocol setup led to many of these traffic diversion difficulties. Such an incident is a big problem since researcher can't follow the routes. Cryptographic methods like Secure-BGP make these problems less serious by only allowing real routes to be broadcast. One of the greatest challenges with setting up public key infrastructure (PKI) is that it costs a lot of money to do so in terms of processing power and storage. Not doing anything about routes that look suspicious is a new technique to stop routing problems. Making plans for how to handle outliers might be a difficult job. BGP peers can communicate reachability information with each other, whether they are in the same or different ASes (Sampaio & Sousa, 2019). Researchers may use this information to create an Internet topology at the level of autonomous systems, which is a graph showing how ASs are connected to each other. A BGP peer generates a database of routing tables that shows all feasible paths and the prefixes of networks that may be accessed through any path or AS path. In this BGP table example, the "best path" is shown by the notation ">". Because BGP routers only provide their best path to other BGP peers, the way they see the Internet topology will depend on where they are. Researchers look at how to combine several BGP tables from different locations since it's important to see the Internet architecture from a wider angle. Oregon Route Views saves the BGP tables of all ASs connected to the BGP route repository every two hours. Several authors have explored a novel approach to characterising complex networks. The configuration of a sophisticated network demonstrates its operation. These papers look at three main types of topological models: the scale-free model, which shows power-law degree distributions; the small-world model, which is based on short paths between two nodes and large clustering coefficients; and the standard Erdös-Renyi model for random networks (Latif et al., 2020).

PURPOSE OF THE RESEARCH

This study aims to elucidate how various topologies influence the installation and overall performance of Internet AS, as well as to identify and classify the topologies employed within these systems. As the internet continues to grow, it is becoming more and more important for AS to be efficient, reliable, and scalable. The main goal of this research is to learn how different topologies affect routing algorithms, data transmission, fault tolerance, and the structural integrity of networks. This comprehension becomes the foundation of the research. There are many other types of topologies to choose from, such as hybrid, star, mesh, and ring. The goal of this study is to help network designers, engineers, and decision-makers choose the optimal topological frameworks for their specific operational needs. Researchers will achieve this goal by examining the advantages and disadvantages of each architecture in actual AS installations. Our ultimate objective is to make the internet's infrastructure better by promoting the use of AS in a way that is both smarter and more useful. This improvement will happen by helping to put

AS into use. As a result, connection, efficiency, and resilience will all get better throughout the world.

LITERATURE REVIEW

To determine an ASN's topology, the most crucial thing to examine are the entries in the BGP routing table. An AS route between two locations may be found by looking at the routing information from a neighbouring virtual private network (VPN) and an IP address prefix block. As a means of better understanding the situation, they consulted the AS Graph. Each AS in this set is connected to the others via the nodes and edges. By connecting AS nodes along their edges in annotated AS graphs, researcher may construct new kinds of AS graphs. Findings on the P2P and P2C links provide solid evidence in favour of this conclusion. Getting BGP routes through is not too difficult. Packet Clearing House is among several organisations that advocate for IXPs. After that point, it may use this tactic to monitor IXPs globally. Internet Routing Registries (IRRs) are connected databases that include information used to examine the whole network. Compared to the links in our database, Internet Routing Registries (IRRs) do not appear very often. Several organisations and individuals, such as Merit Network and RIPE, work together to ensure that the RADb databases are always current. Multiple simultaneous connections to different IP addresses are possible for AS. To diagnose issues with your network, researcher may use traceroute and tracer. In response to an ICMP query, a router is required to provide a master list of all the IPv4 addresses to which it is connected. Created using traceroute data, IPv4 routed /24 AS connection statistics were compiled by the CAIDA Archipelago (Ark) monitoring system. Researcher may do these tasks by establishing connections to ASes and assigning them IP addresses. With Traceroute, a network may monitor a much larger volume of traffic. Researcher can examine each AS connection using one of three methods: BGP routes, Traceroute, or IRR. To research all of the area transport companies, researcher must be proficient with ASes and the internet. The number of passengers transported by an AS is a useful indicator of the number of people it serves. Customers whose trips are becoming less frequent constitute the Customer Cone (CC), according to the AS. Customers aren't making any purchases at the moment; thus, this is a perfect opportunity to reach out to them since.This data also has many potential applications (Triantopoulou et al., 2019). Administrators of internet exchange points may take the size of the client's cone into account when planning how to connect clients to other networks. During AS-to-AS communication, the consumer cone fluctuates in size. Peering networks are susceptible to this. An AS's CAIDA rating will be lower if its client cone is larger. With such a wide range of site sizes, the consumer cone fails to accurately depict the AS's actual customer connection. Level 3 is more well-liked in the United States than in Europe. Researcher could find out more about AS's condition by using the consumer cone. The two are incomparable. One way of looking at the internet is as a network of interconnected AS that coordinate their actions to accomplish certain goals. AS are autonomous and capable of routing traffic independently, but the networks that link them are crucial. Researchers and engineers in the field of networks have put in a lot of time and effort to determine the effects of different topological designs on the reliability, efficiency, and speed of data routing. Using rings, meshes, and hybrids, these models were created. Assembling AS topologies makes it simpler to comprehend the interconnections, data sharing, and communication mechanisms of networks. A potential issue with star topologies may go unnoticed because of a single vulnerable spot. Despite their notorious difficulty in implementation, mesh topologies offer unparalleled fault tolerance and redundancy. It could be challenging to grasp mesh topologies. Hybrid topologies aim to achieve a balance between the potential rate of change and the ease of adaptation. Studies suggest that the topology of large, multihomed, heavily peering networks may influence the performance of routing algorithms (Clark & Claffy, 2021).

RESEARCH QUESTION

What is the effect of network size and reach on the implementation of Internet autonomous systems?

RESEARCH METHODOLOGY

Research Design

The quantitative study used the latest version of SPSS, 25. The odds ratio and 95% confidence interval were used to assess the strength and direction of the statistical link. The researchers determined a statistically significant criterion of $p < 0.05$. An analytical examination was performed to identify the primary components of the data. Quantitative methods are often used to assess data acquired via surveys, polls, and questionnaires, along with data analysed using computational tools for statistical evaluation.

Sampling

Research participants completed questionnaires to provide data for the study. Employing the Rao-soft method, researchers selected a group of 1,260 people, yielding a total of 1,416 questions. The researchers obtained 1358 replies, removing 43 due to incompleteness, yielding a final sample size of 1315.

Data and Measurement

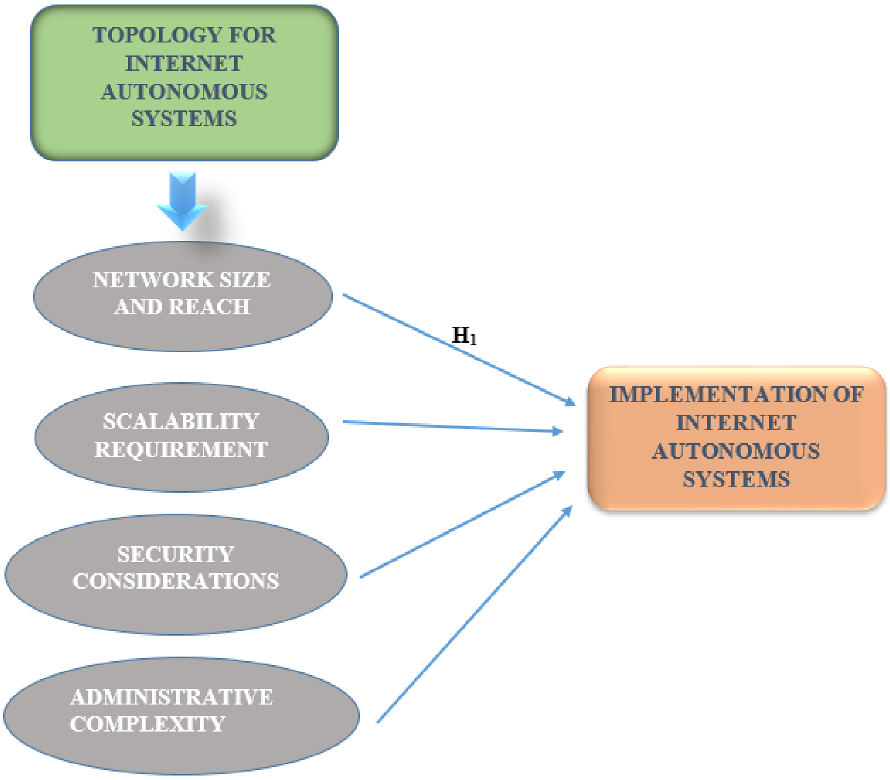
This study used a questionnaire as the main tool for data gathering. Section A of the survey requested essential demographic information, while Section B used a 5-point Likert scale to collect answers about characteristics related to online and offline channels. The secondary data was obtained from many sources, mostly online databases.

Statistical Software

The statistical analysis was performed using SPSS version 25 and Microsoft Excel.

Statistical Tools

The statistical analysis approach was used to understand the essential properties of the data being analysed. The researcher must do a data analysis using ANOVA.



CONCEPTUAL FRAMEWORK

RESULTS

Factor Analysis: Factor Analysis (FA) is often used to identify latent variables within observable data. Using regression coefficients for evaluation is a typical practice when obvious visual or diagnostic signs are lacking. Models are crucial for success in financial analysis. Modelling inherently entails mistakes, interferences, and discernible relationships. The Kaiser-Meyer-Olkin (KMO) Test may evaluate datasets generated by multiple regression analyses. Researchers assert that the model and the variables in the sample are representative. The data exhibits redundancy. Data is more intelligible when conveyed in reduced volumes. Any value between 0 and 1 may function as the KMO output. A KMO value ranging from 0.8 to 1 is considered sufficient for sample size. Kaiser believes that these are the acceptable ranges: Kaiser has delineated additional admission criteria.

An inadequate range of 0.050 to 0.059 and a subpar range of 0.60 to 0.69

The typical range for middle grades is 0.70 to 0.79.

The quality point score ranges from 0.80 to 0.89.

The range from 0.90 to 1.00 astounds them.

The outcomes of Bartlett's sphericity test are as follows: The chi-square value is around 190, with a significance level of 0.000.

This confirms that claims made for sampling purposes are genuine. The researchers used Bartlett's Test of Sphericity to determine the significance of the correlation matrices. A Kaiser-Meyer-Olkin measure score of 0.877 indicates a satisfactory sample. Bartlett's sphericity test produces a p-value of 0.00. The association matrix does not possess a distinct value, hence satisfying Bartlett's circularity test.

Table 1. Evaluation of Sampling Adequacy using KMO and Bartlett's Test indicating Kaiser-Meyer-Olkin measure of 0.877.

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.877
Bartlett's Test of Sphericity	**Approx. Chi-Square**	3252.968
	df	190
	Sig.	.000

Bartlett's Test of Sphericity further validated the importance of the association criteria. The Kaiser-Meyer-Olkin metric of sampling adequacy is 0.877. Using Bartlett's sphericity test, researchers obtained a p-value of 0.00. The results of Bartlett's sphericity test indicated that the correlation matrix is not valid.

INDEPENDENT VARIABLE

Topology for Internet Autonomous Systems: "Internet topology" is a term that outlines how nodes and connections are set up in a data network. It might be as simple as basic or as hard as very intricate. The Internet Topology Zoo is one of several well-known projects that gathers network data for study and analysis. An AS is a set of Internet Protocol (IP) routing prefixes that work together to make sure that the Internet has a clear and consistent policy. These prefixes are managed by one or more network providers. They all belong to the same group or administrative entity. Each AS that is being looked at is assigned an autonomous system number (ASN) to make Border Gateway Protocol (BGP) routing simpler. The Internet Assigned Numbers Authority (IANA) offers regional Internet registries (RIRs) groups of Autonomous System Numbers (ASNs) so that they may be given to other RIRs. Then, these blocks are delivered to local Internet registries (LIRs) and end-user organisations. The Internet Assigned Numbers Authority (IANA) also maintains track of ASNs that are designed for private use and shouldn't be shared online. When talking about AS that are connected to the internet, the word "topology" refers to the overall shape of the links that make up the network. This is particularly essential to consider about when you think about scalability, fault tolerance, network speed, and an effective routing system. AS may have many various topologies, such as hybrid, multihomed, ring, star, and mesh. For instance, big ASes often use hybrid and mesh topologies to ensure sure their systems are constantly up and running and have backups. Smaller ASes, on the other hand, might use star topologies since they are simpler to use and need less upkeep. When building and administering an AS, choosing the correct topology is critical since the size, reach, and operational demands of the network all impact the decision-making process (Nur & Tozal, 2021).

DEPENDENT VARIABLE

Implementation of Internet Autonomous Systems: Internet AS key concerns include the development, enhancement, and operation of autonomously managed networks that aid in global traffic routing. The routing of the internet is aided in some way by these networks. A Regional Internet Registry (RIR) is a kind of Internet registry that assigns one ASN to every AS. Each AS is overseen by a separate entity. Any number of institutions, including private businesses, educational institutions, and even government agencies, could fall under this category. If the AS has this number, it has the capability to use the Border Gateway Protocol (BGP) to connect to other systems and exchange routing data with them.

Developing a thorough routing policy is the first step in successfully establishing an AS. Both incoming and outgoing traffic to and from other ASes are configured according to this policy. Transportation to and from several ASes is included in this. Configuring routers according to these requirements is the duty of skilled networking specialists who want to guarantee redundancy, fast routing selection, and good security. Setting up transit arrangements to connect to upstream providers or peering agreements with other advanced services is an additional technological step that has to be accomplished. The company model in question determines if this strategy component is applicable (Funel, 2019).

FACTOR

Network Size and Reach: The size of a network affects how central it is. Eliminating the impact of network size allows for more precise comparisons of nodes. This change is very important since the number of connections a node has, which is what centrality means, is directly proportional to the size of the total network. Because of this, it is important to understand and contemplate the size of the network in order to do a useful analysis of the nodes' relevance and the network's structure. The total number of linked components, or nodes, in a network determines its size. The total number of physical or logical components in a network determines its reach. One way to figure out how big a network is is to count how many people, devices, or organisations are linked to it. Researcher may also think of the region that a network covers as its reach. This can be a tiny personal area network (PAN) or a massive wide area network (WAN).

In network analysis, the term "network size" refers to the overall density of a network. Nodes may be anything from people in a social network to websites on the Internet to neurones in a brain network to species in a food web. They are one of the numerous types of connected components. Nodes may also stand for many different kinds of parts. In digital marketing, "network reach" refers to the total number of unique people that an ad may reach on a certain platform or network. It is an important sign since it represents the largest number of people that marketers may leverage for their marketing activities. When marketers have a larger network, they frequently have more chances to get in touch with the people they want to reach and build a relationship with them (Witono & Yazid, 2020).

Relationship between network size, reach, and Implementation of Internet Autonomous Systems: Keeping in mind that the width and depth of the network have a significant influence on the deployment of AS over the Internet is an essential aspect to bear in mind. As a means of meeting the requirements of networks that serve a big number of users or that span a vast geographic area, it is essential for AS systems to be more durable and scalable. This will allow them to satisfy the requirements of such networks. As the size of the network continues to expand, it is becoming more important to make use of advanced routing protocols, redundant equipment, and topologies that are optimised. The reason for this is because as the network grows, the complexity of routing, traffic management, and fault tolerance gets increasingly more involved. This is the reason why this is the case. To a similar extent, large-scale networks, such as Tier-1 or worldwide Internet service providers, are required in order to establish access servers that are capable of sustaining an efficient and constant connection over a wide range of geographical locations. To fulfil this condition, it is necessary to do so. Components such as a strong network architecture, a considerable number of peering connections, and sophisticated BGP settings are often included in the configurations that are classified as belonging to this group. Smaller or more regional networks may have AS structures that are less complex, with less peering and more direct routing mechanisms. This is because smaller networks tend to be more regional. On the other hand, it is possible that bigger networks may not possess these properties. There is a correlation between the size and scope of the network, which in turn determines how the AS is deployed, and the performance, reliability, and scalability of an AS. This is due to the fact that the administration system is distributed in a way that is specified by the network (Hoeschele et al., 2021).

In considering the above debate, the researcher formulated the following hypothesis: to assess the relationship between network size, reach, and Implementation of Internet Autonomous Systems.

"H_{01}: There is no significant relationship between network size, reach and Implementation of Internet Autonomous Systems."

"H_1: There is a significant relationship between network size, reach and Implementation of Internet Autonomous Systems."

Table 2. H[1] ANOVA Test.

ANOVA					
Sum					
	Sum of Squares	**df**	**Mean Square**	**F**	**Sig.**
Between Groups	39588.620	447	5645.513	1053.660	.000
Within Groups	492.770	867	5.358		
Total	40081.390	1314			

This investigation will provide substantial outcomes. The F statistic is 1053.660, accompanied by a p-value of .000, which is below the .05 alpha threshold. The hypothesis posits that *"H_1: There is a significant relationship between network size, reach and Implementation of Internet Autonomous Systems."* The alternative hypothesis is validated, whereas the null hypothesis is rejected.

DISCUSSION

The current internet is built on a network of AS. ASs are networks that function on their own and are run by organisations like internet service providers (ISPs), schools, and multinational companies to provide people with access to the internet. For a system to work, the system's topology, which arranges its nodes and connections, must be in place. This study aimed to classify AS topologies and assess their impact on the technology's practical use. Upon further scrutiny, it is clear that topology does not have universal applicability. To provide their clients excellent availability, performance, and redundancy advantages, many major ASes with a lot of global coverage adopt more complex topologies, including mesh or hybrid. If anything goes wrong, its design makes sure that the flow of data will continue without any problems from that point on. Even though smaller ASes commonly used star or stub topologies because they are easier to administer, these topologies are easier to maintain and cost less than other topologies. These designs, on the other hand, don't have any backup routes or ways to be flexible built into them. When choosing the optimum topology for a certain AS deployment, there are many important things to think about. Some of these issues include the budget, the routing plan, the network's growth, and different security needs. It is critical to adopt the right topology when networks are developing and demand is increasing. This is due to the topology's effect on routing, error handling, and overall performance. This classification may be beneficial for network engineers and lawmakers in the development of the next generation of internet infrastructure that is dependable, efficient, and resilient to future challenges.

CONCLUSION

An ecosystem of AS has emerged as a consequence of the growth of the internet, and this ecology is getting less clear and more complicated as time goes on. In issue are AS that operate autonomously, in line with their own rules, and they are designed to accommodate a diverse variety of locations, users, and objectives at the same time. In terms of its ability to foresee the routing, scalability, fault tolerance, and overall efficiency performance of an AS, the findings of this study provide evidence that demonstrates the relevance of network topology. In this article, an investigation of the implementation of a variety of distinct AS topologies in real-world scenarios is presented. Star, mesh, ring, hybrid, and multihomed environments are the classifications that are used to describe these topologies. This exemplifies the significant connection that exists between the size of the network, the geographical reach of the network, and the topology that is chosen for the network. In terms of assuring the quick transmission of data and the dependability of the data, the results suggest that bigger and more internationally linked ASes make use of more complex topologies. However, smaller and less priced ASes often choose for simpler designs that are less expensive. This is because simpler designs are available. It was shown by the hypothesis that topology is not a random design option but rather a deliberate strategic decision that has an influence on the performance and durability of the network. This was demonstrated by the fact that the hypothesis was offered.

REFERENCES

1. Clark, D., & Claffy, K. C. (2021). Trust zones: A path to a more secure internet infrastructure. Journal of Information Policy, 11, 26-62.
2. Funel, A. (2019). The graph structure of the internet at the autonomous systems level during ten years. arXiv preprint arXiv:1902.05029.
3. Gherari, M., Akbari, F. A., Habibi, S., Ali, S. O., Hmitti, Z. A., Kardjadja, Y., ... & Ajib, W. (2023). A review of the in-network computing and its role in the edge-cloud continuum. arXiv preprint arXiv:2312.00303.
4. Hoeschele, T., Dietzel, C., Kopp, D., Fitzek, F. H., & Reisslein, M. (2021). Importance of Internet Exchange Point (IXP) infrastructure for 5G: Estimating the impact of 5G use cases. Telecommunications Policy, 45(3), 102091.
5. Latif, Z., Sharif, K., Li, F., Karim, M. M., Biswas, S., & Wang, Y. (2020). A comprehensive survey of interface protocols for software defined networks. Journal of Network and Computer Applications, 156, 102563.
6. Nur, A. Y., & Tozal, M. E. (2021, April). Single packet AS traceback against DoS attacks. In 2021 IEEE International Systems Conference (SysCon) (pp. 1-8). IEEE.
7. Sampaio, A., & Sousa, P. (2019). An adaptable and ISP-friendly multicast overlay network. Peer-to-Peer Networking and Applications, 12(4), 809-829.
8. Triantopoulou, S., Papanikas, D., & Kotzanikolaou, P. (2019, July). An experimental analysis of current DDoS attacks based on a provider edge router honeynet. In 2019 10th International Conference on Information, Intelligence, Systems and Applications (IISA) (pp. 1-5). IEEE.
9. Witono, T., & Yazid, S. (2020). Portrait of indonesia's internet topology at the autonomous system level. In Computational Science and Technology: 6th ICCST 2019, Kota Kinabalu, Malaysia, 29-30 August 2019 (pp. 237-246). Singapore: Springer Singapore.