Byzantine Robustness for Future Inter-domain Routing Security through Integrated Management Plane

Vahid Heydari Famit Tafreshi, Haitham Cruickshank, Zhili Sun
Centre for Communication Systems Research (CCSR), University of Surrey, Guildford, Surrey, GU2 7XH, UK
e-mails: {v.fami; h.cruickshank; z.sun}@surrey.ac.uk

***Abstract*— Border Gateway Protocol (BGP) is the de-facto inter-domain routing protocol exploited in the Internet today. Future Internet will not serve as a trustworthy vehicle for communication without overcoming BGP security challenges. While security should be a built-in element of any good design, it seems to be an arduous add-on process for BGP. The protocol suffers from the Byzantine Failure whence a legitimate node simply misbehaves. Currently, no systematic method determines whether the received information from an Autonomous System (AS) is valid or not in a global scale. This is due to the absence of an integrated managerial plane operating upon the control plane in our minds. We propose a hybrid method by an overlay network with a global, shared view of the address space ownership performing over the highly-connected ASes merely for the veracity check of the BGP origins. Subsequently, by breaking the hop-by-hop paradigm of BGP with the aid of our introduced management plane, we reach a level of Byzantine Robustness in which the risk pertaining to BGP prefix hijacking as a severe instance of Byzantine attacks is mitigated to a large extent.**

Keywords- Byzantine Robustness/Failure; Integrated Management Plane; BGP ; Inter-domain Routing Security

I. INTRODUCTION AND BACKGROUND

A. BGP Overview

Border Gateway Protocol (BGP-v4) acts like glue to bind divergent components of the Internet together. Any of these individual components is called Autonomous System. BGP facilitates the communication of NLRI (Network Layer Reachability Information) which in turn answers the question of relative reachability address of other hosts and routers across the Internet. This routing protocol is categorized under adaptive/dynamic routing in which routing adjusts automatically to network topology changes. BGP is a hearsay network in which NLRI is whispered between ASes in a series of point to point exchanges between domains. This requires ASes to mutually have a level of trust on each other. Nevertheless, the dynamic population of routing tables via dynamic routing protocols such as BGP allows for the malicious entities to inject bogus routing information. The Byzantine Failure is an arbitrary fault that can occur during the execution of a routing algorithm. A Byzantine Robust model is the one which deal with security issues of a distributed system by tolerating the presence of corrupted nodes that trying to hinder the system to work according to its specification [1] [2]. Here, we want to prove firstly that BGP is not immune in this sense and then try to propose our design for BGP which provides the routing protocol with Byzantine Robustness model.

Communication between BGP speaker routers takes place utilizing different message types. UPDATE messages are exchanged containing NLRIs. This information is a combination of announcements of new routes-to-prefixes bindings along with its associated BGP attributes as well as the list of withdrawn routes (for prefixes which were already advertised). BGP is known to have incremental updates by which whence a new route to a prefix is appeared, the relevant UPDATE is distributed through a series of point-to-point exchanges.

On the other hand, a routing system with Byzantine Failure initially is defined as the one which behaves incorrectly when some malfunctioning nodes exist (a legitimate node misbehaves). Considering the traditional CIA security objectives (Confidentiality, Integrity and Availability), integrity under inter-domain routing security not only encompasses the correctness of NLRI in the course of transit, but also emphasizes the veracity of the information with regards to Byzantine Failure. In inter-domain context, what we are hoping is that in the existence of one or more legitimate but malicious entities exhibiting Byzantine Failure behavior, the BGP routing infrastructure can reach a consistent decision to determine the veracity of the prefix. By mitigating the risk of prefix hijacking, we claim that BGP can reach a level of Byzantine Robustness.

B. Analysis OF Byzantine Robustness Across BGP Security Solutions

Many works including [3] as the major one have explored BGP security holes and solutions so far. We sanitized and analyzed these solutions from Byzantine Robustness point of view. The result of our analysis shows that almost all the remedies in this area are vulnerable to this class of attacks due to implicit trust relationships between BGP speakers. For this gap, there is no systematic mechanism currently in place to check whether the injected information into inter-domain routing system's control plane is genuine or not (in a global scale). Amongst those solutions which slightly consider Byzantine Robustness, DNS-based [4] and sBGP [5] approaches ask for use of digital signature. This implies need to the PKI from which today's Internet is deprived. pgBGP [6] cares only about the detection (and thus ignores prevention). It also asks for revealing peering information which faces reluctance of ASes. All other remedies clearly have overlooked Byzantine Robustness from the scratch. Byzantine Failure can be thought of as the inherent deficiency with each distributed system whence every entity puts a level of trust on the others. This in fact stems from the hop-by-hop

978-3-901882-50-0

characteristics of BGP which is required to be broken. The way that we propose to break this hop-by-hop paradigm to reach a level of Byzantine Robustness emerges from studying the topological structure of today's Internet. Initial results of the simulation will better prove this analysis and reflect the required cornerstone for our method proposed later in section III.

II. Initial Simulation Results

Graphical Network Simulator version 3 (GNS3) [7] has been utilized for the simulation work. GNS3 involves Dynamips, the core program that facilitates the Cisco IOS emulation. GNS3 provides this emulation with graphical front-end. The merit here is that GNS3 bridges between simulated work and the real work environment with performing emulation under the exact conditions of the production environment rather than mere simulation. We change the topology and addressing scheme similar to [8] to form our baseline scenario illustrated in "Fig. 1". Each router is representing an individual autonomous system. AS1 as well as AS2 are emulating Tier1 ASes while AS3, AS4 and AS5 are considered Tier2 providers. Tier3 ASes consist of AS6, AS7, AS8 and AS9. The subnets which are owned by each AS as well as further address delegations are shown in the figure. Byzantine Failure as a result of implicit trust relationship vulnerability between BGP speakers can be better understood if one (R6) tries to inject malicious routing information into the control plane of inter-domain routing in the scenario. Lack of any mechanisms for BGP NLRI validation give rise to propagation of the false information generated by the attacker throughout the whole routing system. While TCP MD5 used here, ensures that the legitimate party is the one who sent the NLRI and the integrity of the message, the semantics of the NLRI is not validated in any ways and therefore can disrupt routing severely in today's Internet (Byzantine Failure).

"Fig. 2" shows the initial BGP routing table for R1 (partially) before R6 starts exhibiting Byzantine behavior. The observation is that 110.1.1.0/24 is accessible via AS4-AS5-AS9 path (indicated by *> under the first column on the left as the best path) while candidate paths to this prefix are shown (highlighted by *). R6 misbehaves by announcing 110.1.1.0/24 into BGP's control plane as an attempt to hijack this BGP prefix (the real origin is already allocated to R9). The NLRI will be propagated further through R3 to Tier1 routers. The malicious BGP entry will be injected and installed into the control plane of all the ASes on the way (R3, R1 and etc.). After the attack has been conducted, as "Fig. 3" depicts, although the AS1's routes towards the legitimate source of the prefix, that is AS9, are still installed, due to shortest BGP AS- PATH, the malicious route AS3-AS6 dominates as the best route. The substantial observation is that after the prefix has been hijacked, R2, R4 and R5 will still utilize the route towards R9 as the real source for NLRI since they have shorter AS-PATH for the same subnet via genuine route (in the absence of any other BGP policies). To conclude, the success of the Byzantine attack exploiting the implicit trust relation between BGP nodes depends on the topological structure of the network with regards to location of the misbehaving entity as well as the target and thus localization becomes substantial for achieving Byzantine Robustness.

III. Proposed Design and Discussion

Our potential solution must take into account the fact that prefix lists change significantly over the time and therefore it is indeed ICANN as the ultimate entity which knows the real owner of the prefix. As a result, we need a shared global view of valid and up-to-date address ownership integrated before being fed into inter-domain's control plane. This can be performed with the addition of management plane required in our mind to integrate the semantics of BGP (sitting on both control plane as well as data plane) across the network which is missing from today's Internet. Our architecture is considered a distributed security remedy (to avoid single point of failure) for BGP integrated at management level which imposes little overhead to the routing system due to not employing any cryptographic approaches. This, in our minds, meets the future demand of BGP security as well as is welcomed by business due to having a robust remedy with no or low cryptographic involvements for one thing (does not need any changes of infrastructure), and the piecemeal deployment option for another thing (respects current BGP's huge install base).

We argue that all the efforts so far have been dedicated to securing either BGP's control plane or data plane. Nevertheless, BGP remains vulnerable to Byzantine attacks as we saw in section II. We propose an integrated management plane sitting on top of the control plane of the inter-domain routing with the aid of an overlay network to reach a level of consistency for the semantics of BGP. The reasoning behind the integrated management plane requirement is that the malicious and inconsistent view of the network is avoided only if an ultimate entity like ICANN provides the shared global database of the address ownership. A set of anchors of trust (each anchor of trust is a server associated with the core router of the corresponding domain) which communicate with each other keeping themselves updated about the latest valid address ownership changes in a secure manner, must share this database. Only a few anchors serve the Byzantine Robustness purpose. Instead of focusing on centralized architecture and having single point of failure, we propose migrating towards a distributed architecture suggested here if Byzantine immunity is wanted.

The way that these anchors of trust communicate is similar to how IRV [9] is functioning and thus policy as well as peering information remain in ASes' confidence which is of their interest (local policy is respected). We then need to break the hop-by-hop paradigm of BGP through studying the topological structure of the Internet today to demonstrate which nodes are the best candidates to be included within the overlay network and to find trust anchors' locations. Bear in mind that defensive filtering is complementary to our proposal since it addresses prevention after we are able to detect Byzantine violations. The topological structure of the Internet in detail is studied in power law structure of the Internet [10]. According to this study, if d_v indicates out-degree of node v and r_v or rank is the index in the sorted list of decreasing order, then R or rank exponent is defined as the slope of the plot indicating out-degree of the ASes versus the rank of the ASes in log-to-log scale. d_v of an AS, here v, is proportional to the rank of the AS, r_v to the power of a constant, R ($d_v \propto r_v^R$). Bear in mind that R<0 in this case. The main observation here is

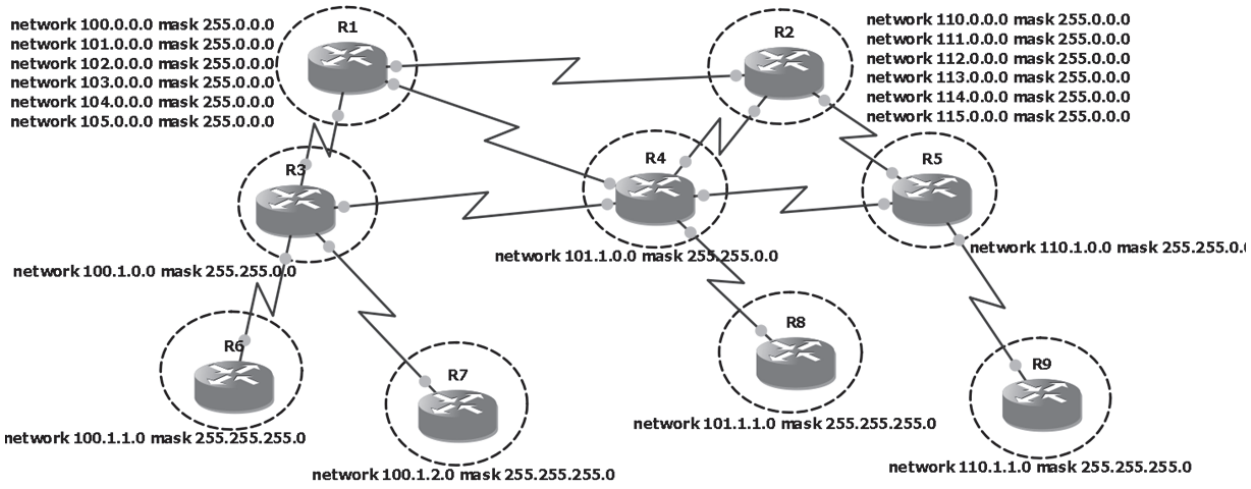

Figure 1. The baseline scenario along with the addressing scheme, each router represents one Autonomous System. TCP MD5 here, as the state-of-the-art solution fulfils integrity as well as authenticity of the BGP information exchanged across the control plane

Figure 2. AS1's initial BGP table for 110.1.1.0/24 before R6 starts exhibiting Byzantine behaviour

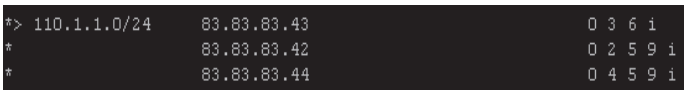

Figure 3. AS1's partial BGP table for 110.1.1.0/24 after R6 starts exhibiting Byzantine behaviour

that there are many ASes categorized as stub ASes (with $\log(d_v)$ near to 1) in the whole Internet while only a few ASes are highly-interconnected (with $\log(d_v)$ near to 1000).

On one hand, there are many ASes categorized as stub ASes in the whole Internet while only a few ASes are highly-interconnected. On the other hand, as we saw in small scale in section II, the magnitude of the adverse effect of accepting false or malicious NLRI is reliant directly on the location of the origin and thus the localization becomes crucial. For this, we suggest to break the flat structure of the current Internet to set up management plane consisting of a set of anchors of trust (probably a hierarchy) made of a few number of highly-connected ASes. Each of these designated ASes can communicate with each other via a secure protocol and constitute the required managerial plane. One of the major defects of IRV is that the receiver would query the originating AS for the veracity of the route/origin while the other party can exhibit Byzantine behavior in a flat structure of the Internet. Nevertheless, our method asks for querying to/responding from anchors of trust merely in IRV manner instead of an alien in a semi-hierarchical structure obtained by breaking the hop-by-hop paradigm of the BGP (with the aid of ASes which are rich in the connectivity as discussed earlier). Furthermore, a subset of these anchors can also be queried (instead of one) to increase the confidence for Byzantine Robustness while to improve performance, we can cache the previous queries or employing periodic/partial queries/responses. The goal of the management plane formed with the aid of these servers is to eliminate any inconsistencies existed in the semantics of BGP before any routing information become injected and propagated into the inter-domain routing's control plane. If BGP security countermeasures are talking about guarding inter-domain routing space then Byzantine Robustness is guarding the guard and as such we claim that a separate plane established over the control and data planes of inter-domain routing is required to protect the veracity of NLRI. Due to indeterminate nature of address ownership in IP space, the added plane is of managerial type which fulfills the coordination required amongst these anchors of trust at all the time to ensure the genuineness of the repository of NLRI information.

According to this and as the results in section II suggest earlier, mitigating the threat for BGP prefix hijacking as an instance of Byzantine Failure mostly relies on the existence as well as the quality of route monitoring databases. Since we believe that Byzantine Failure is a managerial issue inherent to distributed systems like BGP, our management plane formed by an overlay network addresses this requirement with having an always-cleansed, accurate and shared view for AS-number-to-IP bindings which works globally in a systematic way. "Fig. 4" better reflects the proposed architecture. We believe that the decision made by a few number of very large ASes which are rich in connectivity have extreme effects on the routes provided to other ASes as well as the robustness of the whole routing system. To put it in another word, the vast majority of path diversity provided to other ASes is owing to these large ASes which are quantitatively few. The resulted overlay network, as shown here, benefits from piecemeal deployment and therefore requires no change in the course of the development for the underlying BGP network. Back to the scenario, R3 with 4 and R4 with 5 hands are the highly-connected ASes amongst the others which form the vantage points required to break BGP's flat structure here. Through

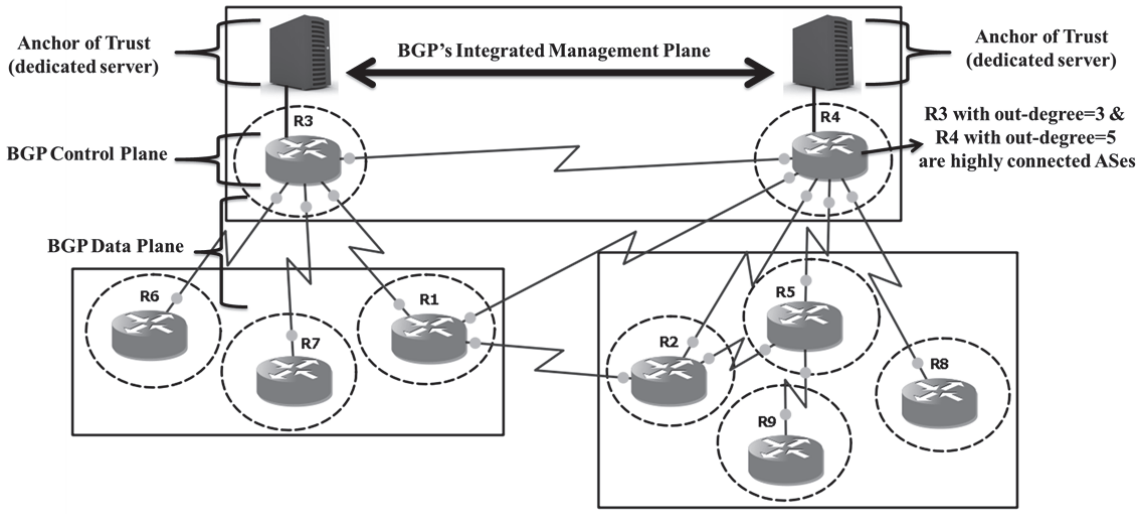

Figure 4. Proposed architecture, the addition of BGP's management plane sitting on top of the control plane implemented by anchors of trust over highly-connected ASes (merely) ensures Byzantine Robustness to a large extent with compartmentalisation of the Byzantine affected subnets

our overlay network, servers associated with AS3 and AS4 are communicating with each other in query/response manner to reach a shared view of the address ownerships (under ICANN). Defensive filters (currently in use) on the core routers of AS3 and AS4, is then fed via the management plane established between these ASes' servers with the latest view of the address ownership at all the time. Except these two ASes, other ASes remain untouched. The conduits for exchanges required between servers of the latest updates can use the underlying network and thus cause no change into the infrastructure. Now if R6 starts exhibiting Byzantine behavior, since AS3 had believed that 110.1.1.0/24 belongs to AS9 and has not been informed of any changes to this by the management plane provided by the overlay network, R3 discards the received malicious BGP UPDATE. In this way, R1 and all further ASes remain safe of the attack (unlike section II). The semi- hierarchical view of the network through the overlay network is now governed by an integrated management plane which limits the damage of the Byzantine Failure to direct neighborhood merely.

Compartmentalization of the affected subnet opens the door for routing around the misbehaving AS. However, topology authentication is not our concern. By focusing merely on origin authentication, we leave the flexibility for traffic engineering to address routing itself independent of security. Last but not least, securing the router management for the border routers as well as any attempts towards protecting TCP itself, adds another layer of trustworthiness to the future of inter-domain routing.

To sum up, BGP version 4 remains the de-facto protocol for inter-domain routing. The protocol's main security vulnerability is the implicit trust relationship between BGP peers. Large install base in addition to the longevity of BGP's operational life leaves no room for new protocol for near future at least. BGP has not faced with any changes so far regarding security issues and therefore bogus routing objects can be injected maliciously. While TCP MD5 for BGP session protection in addition to defensive filtering are the state-of-the-art solutions, the protocol is still suffering from Byzantine Failures. Benefitting from the Internet's topological structure, our hybrid method asks for integrated management plane to be added on top of the control plane between anchors of trust to eliminate any inconsistencies which may result in Byzantine behavior. This limits the adversary impact of prefix hijacking to a large extent.

We will carry out further work on implementing the proposed architecture in a larger scale for the next step.

REFERENCES

[1] Murilo Santos de Lima, Fabíola Greve, Luciana Arantes and Pierre Sens, "Byzantine Failure Detection for Dynamic Distributed Systems", Mar. 2010 [Online]. Available:
http://hal.inria.fr/docs/00/58/45/97/PDF/RR-7222.pdf

[2] Perlman, Radia. *Network layer protocols with Byzantine robustness.* s.l. : PhD Thesis MIT University, 1988.

[3] K. Butler, T.R. Farley, P. McDaniel and J. Rexford, "A Survey of BGP Security Issues and Solutions," *Proceedings of the IEEE*, vol. 98, pp. 100-122, Jan. 2010.

[4] T. Bates, R. Bush, T. Li and Y. Rekhter,"DNS-based NLRI Origin AS Verification in BGP", IETF Internet Draft, Jul. 1998 [Online]. Available:
http://www.tools.ietf.org/html/draft-bates-bgp4-niri-orig-verif-00

[5] S. Kent, C. Lynn and K. Seo,"Secure Border Gateway Protocol (S-BGP)," *IEEE Journals on Selected Areas in Communication*, vol. 18, pp. 582 – 592, Apr. 2000.

[6] K. Butler, P. McDaniel and W. Aiello,"Optimizing BGP security by exploiting path stability," *13th ACM Conference, Comput. Commun. Security.*, pp. 298 – 310, 2006.

[7] Graphical Network Simulator (GNS3). http://www.gns3.net, last access in May. 2012.

[8] A. Zeb, M. Farooq, "BGP Threats and Practical Security", MSc Dissertation, Department of Computer Scinence and Technology, Chalmers University of Technology, Sweden, Mar. 2011.

[9] G. Goodell, W. Aiello, T. Griffin, J. Ioannidis, P.D. McDaniel and A.D. Rubin, "Working around BGP: An incremental approach to improving security and accuracy of interdomain routing," *in Proc. NDSS*, 2003.

[10] M. Faloutsos, P. Faloutsos and C. Faloutsos, "On power-law relationships of the Internet topology," *ACM SIGCOMM ComputerCommunication Review*, vol. 29, pp. 251-262, 1999.