Gate17

An Onion-Encrypting Multi-Hop Network with Intelligent Routing

An Overview

WHITE PAPER

v1.1

Daniel Hovie, Raphael Fiedler

Abstract

Packets sent through the Internet traverse many privately owned networks, which capture and sell data at will. Packets may also be redirected by an attacker via BGP hijacks who can then gather information and inject malicious payloads if unencrypted.

We propose a solution that protects packets almost all their way through the Internet using an onion-encrypting multi-hop network. To protect data from ISPs of all tiers, paths through the network are chosen in a way that exposure to the open Internet is minimized or completely eliminated. Because the path selection algorithm aims to route as near as possible to the final destination, BGP hijacks have no effect or are mitigated completely in most cases. Information gathered from the network enables clients to effectively dodge suspicious Autonomous Systems.

This network requires a vast amount of widely spread nodes and must be economically sustainable. A trusted provider acts as an escrow who receives payments from clients and supplies them with anonymous access tokens for the network. Alternatively, blockchain-based tokens or cryptocurrencies may be used to handle payments in a trickle-down like method.

To reduce the risk of successful traffic analysis, connections are multiplexed as much as possible, wire packet sizes are fixed and connections between nodes are rather static and clients only use these pre-established links. Connection establishment within the network takes zero round trips and encryption provides forward and backward secrecy with modern and proven algorithms.

Index

Your Metadata on the Internet

Metadata in this context is everything that is needed to fulfill a request for information or action on the Internet that is not the content itself, but is needed for the system to work or is a by-product of the communication itself.

The Internet has made great progress in encrypting communication in the 2010s, including massive HTTPS adoption, TLS 1.3 and encrypted DNS queries. These are tremendous changes, but all this protection mainly protects content. Metadata is still exposed: Names of websites you are accessing, the involved IP addresses, and when and where you access services.

Value of Metadata

Numerous Internet companies digitally harvest people via their metadata to increase their profits. Author and scholar Shoshana Zuboff describes this with the term Surveillance Capitalism: The growth and income focus of capitalism paired with the information gathered from huge amounts of personal data results in more and better tracking of user behavior in order to exploit them financially.

Internet Territories

The Internet is easily viewed as a somewhat mythical global network that acts in public interest and is friendly to everyone who wants to use it. But in reality, the Internet is a collection of privately owned networks called Internet Service Providers (ISPs), who all have their own agenda and try to make money in every way they can - because that is what businesses do.

Data collection

Metadata Collection

One thing we can be sure of, is that whenever an ISPs is legally allowed to sell data that crosses their network, they will. Normally, when you connect to the Internet you pay one of the smaller ISPs to give you access, and most of them are bound to protect your data. But the other dozen ISPs your data traverses may not have any obligations to protect you or your data.

The Electronic Frontier Foundation publishes a yearly report on consumer privacy rating of major US companies, called *Who Has Your Back?*. In 2017, “the four lowest performing companies are all telecoms: AT&T, Comcast, T-Mobile, and Verizon,” [1] meaning they don’t care about your privacy.

ISPs in the US have a special meaning when it comes to the Internet: five out of the 16 Tier 1 ISPs (the biggest ones) are US companies. Together, they make up 72% of the global

Internet infrastructure.[1] When you communicate over the Internet, chances are very high that your packets will go through one of these networks. Therefore, what these companies do and the laws that govern them likely affect everyone on the Internet.

Data retention

In addition to data collection that is done by companies at will, governments around the world have started to force ISPs to collect metadata and store it for a certain amount of time in order to help law enforcement. This data can then also be stolen by hackers, as an activist group demonstrated with an Australian ISP [4].

BGP Hijacks

In case someone is especially interested in the traffic of someone on the Internet, there is no need to wait until the traffic of interest occasionally crosses a controlled network. Instead, an attacker can influence how packets are routed within the Internet. The Internet uses the Border Gateway Protocol (BGP) to exchange information about who is where. Malicious actors can carefully inject information into the Internet, rerouting data wherever they want.

There are numerous high-profile examples for this, here are two recent ones:

- In April 2017 Rostelecom, a russian ISP, hijacked prefixes of financial institutions - most notably MasterCard and Visa [5]
- In July 2018 the Iran Telecommunication Company hijacked 10 prefixes of Telegram Messenger. [6]

Because of how the Internet works, detecting these hijacks is not an easy thing. A BGP monitoring company, BGPMon, offers a stream of events of the last 6 months. From January to June 2019 they detected over 6000 route changes of the Internet and classified over 1000 of them as possible hijacks, that is about 160 per month. [7]

Related Solutions

VPNs were originally created to provide secure access to a corporate network from the outside. When the first consumer VPN services started in the 2000s, they simply adapted this technology and helped customers to protect themselves from security issues within the last mile and the first ISP. Beyond that, a VPN does not protect from metadata collection or provide any other protection.

Interestingly, the development of Tor started (with funding from the U.S. government) before the first VPN protocol (PPTP, by Microsoft) was even released. Tor was the first system to anonymize traffic by encrypting and routing it through multiple servers. If configured and used correctly, it provides excellent network privacy, up to the point where it enters the open Internet. There, traffic is again exposed to private networks that gather

1 measured in peerings: AS degree on 02.2018. Degree is the number of neighbors that a node, an AS in our case, has in a graph. See [2], [3]

metadata. Because routes are randomly chosen in Tor, exposure will be roughly the same as with VPNs.

Gate17

Overview

Gate17 is an onion-encrypting multi-hop network with intelligent routing that aims to reduce exposure to networks that may collect and sell metadata.

Network Structure

The Gate17 network uses a static - but dynamically created - mesh network. Every node tries to optimize the network by establishing connections to other nodes, that it thinks will be useful to clients.

These connections are then published so that every client knows about every connection between nodes. This is important, as clients may only move within the established network and will never trigger a new connection between two nodes.

In order to introduce all nodes to each other, they exchange information about themselves by passing around gossip messages. When a node comes online for the first time, it fetches a list of bootstrap nodes in order to connect to the network.

Involved Parties

The network is designed for these three parties:

Network Owners
The developers of the network. They:

- maintain the project
- sign software updates
- maintain the network (kick bad nodes, ...)
- quickly react to and fix problems
- act as a trust anchor and therefore have to be trusted

The Community
People around the world, that share the owners vision. They:

- help form the network
- provide valuable feedback
- actively engage in its development
- host nodes and get paid

Clients
Whoever wants to protect their network connections. They:

- use the software and network
- pay for network resources

Nodes

Gate17 uses a system with different node types:

Trusted Nodes
are nodes that are hosted by the network owner or trusted parties. They:

- may handle unencrypted traffic as exit nodes
- need to be trusted, as they process unencrypted (eg. HTTP, DNS) requests

Community Nodes
are usually hosted by community members. They:

- diversify control over the network
- get paid for maintenance and upkeep

Tunneling

Tunneling is done on the application/session layer (#5) of the OSI network layer model. This means that the transport layer (#4, eg. TCP, UDP, ...) is terminated locally and cannot leak any information, such as the IP address. Everything above, including SSL/TLS, is tunneled through the Gate17 network seamlessly. Not tunneling the transport layer also eliminates potential bottlenecks from the congestion and flow control algorithms, as they are not optimized for this environment.

New connections are always first reviewed by the Portmaster (an application firewall) to determine if and how it should be handled by Gate17. Connections are then redirected (via DNS and DNAT) to a local port, where Gate17 awaits new connections. This tight integration with an application firewall enables greater control over data leakage and removes the need of complex system configuration.

Upon accepting a new connection, Gate17 receives information from the Portmaster and sets possible custom settings for the specific application. In principle, every app/domain pair can have its own tunnel and routing configuration.

Intelligent Routing

First, Gate17 receives all the requirements for a new tunnel, some of which are:

- Destination (Domain or IP)
- Country Override (simulate presence in another country)
- Country/AS Avoidance (blacklist of countries and AS not to go through)
- Minimum number of nodes to use

It then uses its local database of existing nodes and connections to calculate all possible paths and selects the most promising one. This selection process can be configured to lay emphasis on speed or security. Clients select a node in proximity as their entrance node. All tunnels then use this node as their first hop.

Path Selection

In general, the fastest path to the exit node is selected, while using a minimum hop count of three. Paths are always selected in a way that diversifies node ownership as much as possible, so that no single entity can effectively analyze tunnels and track users.

Additionally, the algorithm will not include nodes in the path that are located in user defined blacklisted countries or Autonomous Systems (AS). Nodes regularly investigate the network paths of their connections (eg. with tracepath) and advertise the country codes and AS numbers in order to enable clients to make these decisions.

Exit Node Selection

Exit nodes are selected in proximity to the destination server(s). This is important as it minimizes or even eliminates exposure to the open Internet, where ISPs can collect and then sell metadata. The only constraint is that unencrypted traffic will be handled by trusted nodes in order to prevent Man-in-the-Middle attacks.

Because the selected exit node is near the destination server, BGP hijacks (the malicious redirection of an IP prefix) are much less likely to affect the connection, as the exit node should be either in a neighboring or even within the destination AS. The path itself is unaffected from (targeted) BGP hijacks, as only smaller hops are taken through the Internet and the destination IP is not revealed/routed until the exit node is reached.

As every new tunnel gets its own path through the network and is routed near to the destination, clients will appear to *be everywhere at the same time*. Internet observers and trackers therefore cannot use the network location for tracking.

Encryption

Packets are encrypted for every node on the path individually, only uncovering the next hop at each node. This is known as *onion routing* or *onion encryption* and has first been employed by Tor. [8]

Other than Tor, Gate17 supports zero round trip connection establishment, which brings a vital performance improvement to *onion routing*. The encryption protocol also employs forward and backward secrecy to protect encrypted data even if encryption keys are stolen, inspired by the *Double Ratchet Algorithm* by Moxie Marlinspike [9]. Encryption algorithms used are proven and modern, and can be changed or replaced on the fly.

Traffic Analysis

One major concern with all anonymity networks is passive traffic analysis. With enough network visibility, time and resources, an attacker can find out where a connection really goes to and comes from, without even interacting with the network itself.

Gordon Welchman provides excellent advice in his book *The hut six story: breaking the enigma codes* [10] on how to protect against traffic analysis.

Here we compare his (slightly simplified) principles to Gate17.

#	Principle	Gate17
1	All nodes are full nodes of the network with multiple connections (no leaf nodes).	With the exception of clients, all nodes in the network are full nodes.
2	Connections have link layer encryption.	True.
3	Communication is end-to-end encrypted.	True: end-to-end encrypted with every node.
4	All nodes *store and forward* (the longer stored the better)	As the network applications expect a more or less fast connection, this is not feasible.
5	Connections have a fixed bandwidth, which is padded to full utilization at all times.	Partly: Network packet sizes are fixed, but bandwidth is not.

Gate17 fulfills three out of these five principles on average and thus should be able withstand a fair amount of traffic analysis. Especially principles 4 and 5 could be taken into account more, but as this would slow down the network by magnitudes and would break many applications, the network would need to become a hybrid network and offer both capabilities.

Payment

To support such a complex and vast network of nodes, clients will undoubtedly have to contribute financially in order to make ongoing development and the high performance infrastructure sustainable.

In order to support standard payment methods, the network owner must act as a payment escrow and is responsible for processing these contributions. These are paid to the network owner directly, who then regularly and automatically distributes this to the participating node owners. Cryptocurrencies or blockchain-based tokens could potentially be used as an additional payment/distribution method and as an internal authorization mechanism - though this requires lots of care in order to preserve user privacy. For more details, please refer to the *"Payment via Cryptocurrencies"* section.

To ensure clients cannot be identified through their payment, authentication must be decoupled from authorization. Meaning, the system that verifies the client's identity and checks the contribution status must be separate from the system that needs to check if a client is allowed to use the network - in effect - the network nodes.

Here are three possible ways how to accomplish this:

Anonymous Tokens

First, the authentication service can simply issue a global secret token, that is identical for all users, and therefore, untrackable.

The problem with this solution is obvious: any client can abuse the network by sharing this secret token with unauthorized clients, without risking to be identified. The following measures may help mitigating this threat:

- Make sharing the secret token infeasible by regularly (eg. hourly) changing it.
- Rate-limit authentication to stop clients sharing their authentication credentials.
- Issue regular (eg. monthly) software changes that require manual patching of *cracked* software clients.

Group Signatures

Second, Group Signatures[2] provide an authorization mechanism that only proves that the client is part of a group - permitted to use the network, but network nodes cannot identify the client.

The authentication server adds clients to the signature group, who can then sign on behalf of the group to prove their status. Should a client abuse the network, a special independent trust board, consisting of highly respected community members, can then vote on unmasking a client and notify the network owner to revoke access.

While this method can also be abused by sharing the authentication secret (a private key), it may be easier to identify the malicious client and get the trust board to unveil the identity, if enough evidence is provided. For further threat mitigation, the above mentioned measures apply.

Please note that this option has not yet been fully validated to be both technically and organizationally feasible. Also, the ability for an entity to unmask users must be carefully weighed against the benefits and must be thoroughly discussed with the community.

Trickle-Down Payments

Third, instead of trying to authorize a client, the network owner can provide clients with cryptographic tokens stored on a blockchain, which are then transfered to the entry node in order to grant a certain amount of traffic. The network nodes then *trickle down* these tokens wherever their users' traffic flows. This means that the tokens trickle along the same route as the data flows. Token transactions are done only periodically to prevent any data leakage

2 brief introduction to group signatures

and will roughly match the transfered amount of data between two nodes in a relative manner.

The tricky part about this concept is that every node along the path must know how much to charge for traffic, as it will have to trickle down payments for following nodes. This is currently only achievable with approximate payments and letting every node know how many nodes are left in the chain.

This method is especially interesting, as the network owner can easily prevent abuse through keeping count of granted traffic tokens. This vastly simplifies abuse detection, while still protecting the privacy of users. In order to fully protect their privacy tough, additional procedures must be built into this system. This includes constant rotating wallet keys as well as keeping transaction amounts as low as feasible to reduce tracking potential through transaction analysis.

This system is by far the most complex and error prone one, and needs lots of care and testing. Instead of just using cryptographic tokens, the same system can be employed with common cryptocurrencies. This would further increase the complexity due to stringer fraud prevention and trust building methods, as these tokens then have real value outside of the system.

Details of Special Interest

Trusted Provider

Why is a trusted provider necessary, why not “trustless”?

Until we have a fully encrypted Internet (we are getting there [11]), there will be information flowing from your device to some server unprotected. You definitely do not want a stranger to set up shop near a targeted destination and start attacking your connections. That is why the network owner, as the trusted provider, will handle unencrypted traffic.

Please note that while the exit node chosen for unencrypted traffic is a trusted one, the path itself could still have community nodes in it.

Payment via Cryptocurrencies

Blockchain is an emerging and *objectively over-hyped* technology that also provides seemingly anonymous payment systems. Why not use it for clients to pay node providers directly?

There are a couple projects that aim to provide a VPN-like service, where users pay nodes directly. To date, none of them have demonstrated how these transaction can be kept private (for a very long time) and how great the impact on anonymity and privacy of these payments systems are. Secondly, such a payment system would have to be both blazing fast and have a way to mitigate abuse.

If transactions cannot be kept private, one can track a client’s use of the network just by reading the live blockchain transactions to see exactly which nodes are being used by said client.

Currently, we do not think that direct payments to nodes is a feasible payment method for a privacy focused network. Nevertheless, we are closely watching these projects, in case they actually accomplish it - here are some of them:

- Orchid Protocol - the most promising, transaction privacy is out of scope for the beginning. [12]
- Substratum - pretty hyped, there are some red flags and they mixed up encryption and compression.
- Privatix - Do not mention how they plan to have privacy with transactions. [13]
- Mysterium Network - The “Payments Handling: Risk Management” section in their white paper does not mention privacy issues and even states “All the transactions [...] are accessible for everyone to read.”. [14]

Conclusion

Gate17 redefines network connection security and privacy on the Internet. By proposing a new way of thinking – away from just hiding the origin of a connection – towards intelligent destination aware routing. This enables us to thwart network level tracking and BGP hijacking attempts. Optimizing such a vast network for speed and security, as well as processing payments with full privacy will become a continuous challenge. Creating new, cutting-edge communication approaches will lead us into a new, more secure future.

References

[1] “Who Has Your Back? Government Data Requests 2017.” Available: https://www.eff.org/who-has-your-back-2017. [Accessed: 18-Feb-2018].

[2] “Tier 1 network.” Available: https://en.wikipedia.org/wiki/Tier_1_network. [Accessed: 18-Feb-2018].

[3] “CAIDA’s ranking of Autonomous Systems (AS).” Available: http://as-rank.caida.org/. [Accessed: 18-Feb-2018].

[4] “Anonymous hacks Australian ISP AAPT to demonstrate data retention problems.” Available: https://thenextweb.com/au/2012/07/26/anonymous-hacks-australian-isp-aapt-to-demonstrate-data-retention-problems/. [Accessed: 08-Jul-2019].

[5] “BGPstream and The Curious Case of AS12389.” Available: https://bgpmon.net/bgpstream-and-the-curious-case-of-as12389/. [Accessed: 18-Feb-2018].

[6] “Telegram traffic from around the world took a detour through Iran.” Available: https://www.cyberscoop.com/telegram-iran-bgp-hijacking/. [Accessed: 18-Feb-2018].

[7] “All Events for BGP Stream.” Available: https://bgpstream.com/. [Accessed: 01-Jul-2019].

[8] “Onion Routing.” Available: https://en.wikipedia.org/wiki/Onion_routing. [Accessed: 18-Feb-2018].

[9] Perrin, Trevor and Marlinspike, Moxie, “The double ratchet algorithm.” 2016.

[10] Welchman, Gordon, *The hut six story: breaking the enigma codes*. McGraw-Hill Companies, 1982.

[11] Felt, A. P., Barnes, R., King, A., Palmer, C., & Bentzel, C., “Measuring HTTPS Adoption on the Web,” 2017, Available: https://research.google.com/pubs/archive/46197.pdf.

[12] “Orchid Protocol.” Available: https://orchidprotocol.com/whitepaper.pdf. [Accessed: 18-Feb-2018].

[13] “Privatix Whitepaper.” Available: https://dxw4crzwfgmzw.cloudfront.net/whitepaper/PRIVATIX-WHITEPAPER.pdf. [Accessed: 18-Feb-2018].

[14] “Mysterium Whitepaper.” Available: https://mysterium.network/whitepaper.pdf. [Accessed: 18-Feb-2018].