Tor 101: How Tor Works and its Risks to the Enterprise

Executive Summary

The Tor project provides one of the most well-known tools that users can leverage to stay anonymous on the internet. People use Tor for many different reasons, both benign and malicious. However, allowing Tor traffic on enterprise networks opens the door to a variety of potential abuses and security risks.

Political activists use Tor to express their views while staying out of sight of their governments. Cybercriminals use Tor to evade defenses and hide their identity from law enforcement. Tor is famous for enabling the operation of dark web marketplaces, such as Silk Road, where customers could procure a wide range of illicit goods, including drugs, weapons and fake identification documents. Malware authors regularly use Tor for denial-of-service (DoS) attacks, hidden reconnaissance, exploitation, command and control communication and data exfiltration.

For enterprises concerned about the risks of Tor traffic, the use of Tor for malware, command and control, exfiltration, and hidden reconnaissance are some of the most important security risks. Also, employees can use Tor to bypass content blocking policies (e.g., blocking of adult or gambling sites) such as those provided by the Palo Alto Networks DNS Security and Advanced URL Filtering services. Users can also elude geographic restrictions of services or buy illicit goods unchecked using Tor. To avoid these risks, we advise the blocking of Tor in enterprise networks.

Emphasizing the importance of monitoring or blocking Tor traffic in the enterprise, we observed 6,617,473 sessions to or from 691 devices within 204 customer networks in one month.

Palo Alto Networks provides two solutions as part of Threat Prevention that are best used together to filter Tor traffic. We maintain a verified and built-in Tor Exit IP External Dynamic List that our customers can use to block connections from Tor Exit nodes. Customers can also leverage the Palo

Alto Networks traffic classification system App-ID to block incoming and outgoing Tor traffic. Additionally, customers can utilize Cortex XDR to alert on and respond to Tor-related threats on endpoints, in the network or in the cloud.

Related Unit 42 Topics	VPNs

What Is Tor?

The goal of the Tor network is to provide a tool to internet users for anonymous communication. Tor stands for The Onion Router, which is the software that enables Tor nodes to participate in onion routing. Onion routing is a technique allowing anonymous communication. Anonymity on the internet means that no one can connect a user's actions and identity. For example, anonymous communication would mean that when a user connects to a server hosting a forum and makes a public post with a made-up username, neither a global observer, the forum operators nor other users should be able to tell the user's true identity based on these actions.

To better understand how hard it is to stay anonymous, let us introduce the imaginary country of lemons ruled by the tyrant Lemonheads, where it is frowned upon to like or discuss oranges. Emilia, our heroine, is part of a small rebellious group that loves oranges. To discuss her passion with others, she would like to visit a site called `peel-the-orange[.]com`.

One of Emilia's friends, Bob, always used to connect to `peel-the-orange[.]com` without using any anonymity-enhancing technology. Bob told Emilia that his traffic is encrypted since he connects using HTTPS, so no one can read what he is doing. Unfortunately for Bob, HTTPS does not encrypt the DNS name, `peel-the-orange[.]com`, so the lemon secret police could easily tell that he was visiting a site for orange enthusiasts. Therefore, he ended up on the bad lemons list.

After that, everyone in their group became more cautious. Emilia's other friends started using a foreign VPN provider to funnel their traffic through the VPN server and hide their identities. This effort truly made it hard for the lemon police to find out who is connecting to forbidden sites.

Jess, another friend of Emilia, was high on the Lemonheads' list of suspected orange sympathizers. Observing Jess's traffic, it became clear to the secret lemon police that she is frequently connecting to a known VPN provider. While the VPN provider helped Jess protect her identity, it proved to be a single point of failure. A VPN provider can be bribed, threatened or hacked, which are the favorite methods of the Lemonheads. More advanced adversaries (such as the lemon secret police) could employ correlation attacks based on statistical methods to match incoming and outgoing traffic to and from the VPN servers.

The lemon secret police initiated "Operation Super Sour" by compromising several servers of the VPN provider and stealing the complete list of all lemon citizens visiting `peel-the-orange[.]com`. As a result, many of Emilia's friends, including Jess, ended up on the bad lemons list. Furthermore, the lemon government started censoring connections to known foreign VPN server IP addresses to stop the dissidents' decadent love for oranges.

Emilia became very careful after that last incident. She looked for a solution that would keep her safe, even if the lemon secret police could compromise the proxy/VPN servers she used. Additionally, she wanted to ensure that even `peel-the-orange[.]com` would not know her identity if someone infiltrated their server.

She soon found Tor, which seemed like it could help her stay anonymous. However, Emilia soon discovered that even this wasn't foolproof. She told a few acquaintances about Tor, but the lemon secret police were able to identify one of them, Gordon, because he shared his nickname Orange Cake across different sites. Luckily, Emilia only used her pseudonym Orange Girl on `peel-the-orange[.]com`.

How Tor Works on a High Level

Figure 1 helps us explain how, using Tor, Emilia can stay anonymous even if the Lemonheads would have full global visibility of network traffic or could compromise a couple of Tor nodes.

At first, we assume that Emilia already has a Tor circuit built, meaning that her computer already selected three Tor nodes (servers running Tor software) to relay messages and obtained a shared key with each of them.

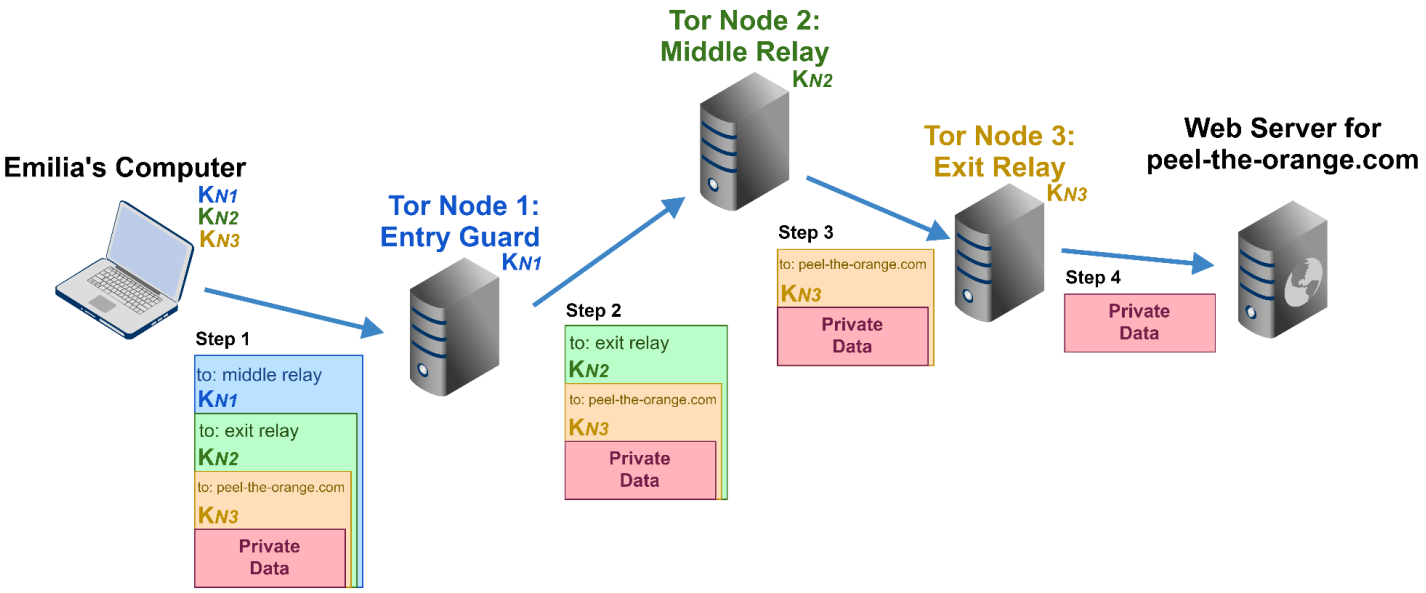

Figure 1. High-level overview of how connecting to a website looks using an already built Tor circuit.

Emilia's computer first encrypts the private data in three layers (step 1 in Figure 1), hence the name onion routing. Her computer encrypts the data in reverse order: first with the key of the last exit node (Kn3), then with the middle relay node's key (Kn2) and finally with the guard node's key (Kn1). The guard node receives the data, removes the outermost layer of encryption using Kn1 and sends the decrypted message to the relay node. The middle node removes the next layer using Kn2 and relays it to the exit node. Finally, the exit node decrypts the message with Kn3 and sends the original data to the web server (in this example, `peel-the-orange[.]com`). The layered encryption enables secrecy and limits knowledge about who is involved in the communication as only the nodes that know the keys can decrypt the messages.

Nodes known	Entry Guard	Middle Relay	Exit Relay	Web Server
Emilia	knows	don't know	don't know	don't know
Entry Guard	self	knows	don't know	don't know
Middle Relay	knows	self	knows	don't know
Exit Relay	don't know	knows	self	knows
Web Server	don't know	don't know	knows	self

Table 1. This shows which nodes in the column headers know which nodes in the row headers.

Table 1 summarizes which nodes know about which other nodes. The guard node knows who Emilia is and the next node that receives Emilia's message, the middle relay node. However, the guard node does not know about the last exit node and Emilia's final destination, because decrypting with only Kn1, the message is still garbled for the entry node. The relay node knows the least. It does not know who is the original sender or the final destination and only knows the entry and exit nodes. The exit node knows about the middle relay node and the destination server, while the destination server only knows about the exit node.

Messages on the way back to Emilia are passed back in a similar fashion, each node adding a layer of encryption using the key shared with Emilia.

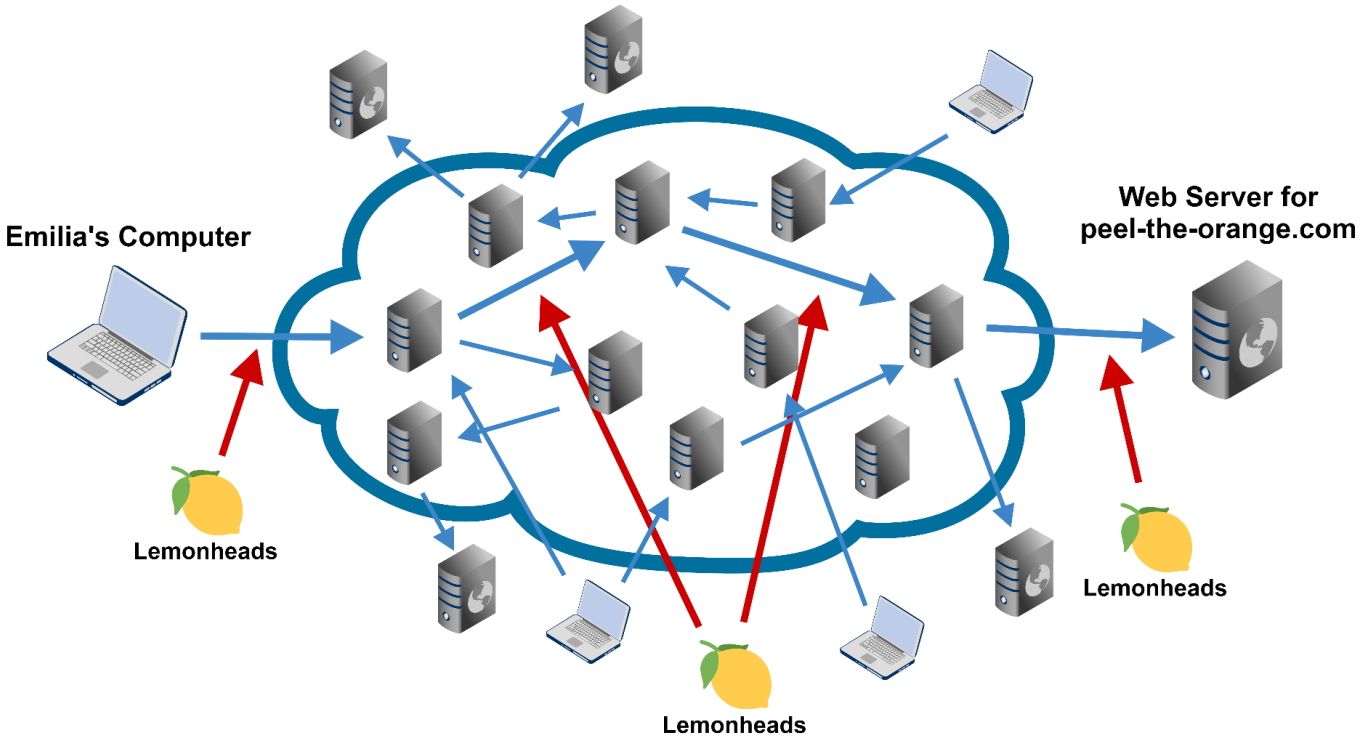

Figure 2. The global Tor Network and a global observer.

Figure 2 depicts Lemonheads as a global observer. When Emilia uses Tor, the Lemonheads can only observe her connection to the entry node. Even if they have complete global visibility of the messages passed, it becomes hard for them to track Emilia's messages as they get mixed in with all Tor users' traffic, and the layer of encryption changes every time the message is passed to another node. As discussed earlier, if Emilia were to use a single VPN provider, it would know what sites Emilia visits. Additionally, the Lemonheads could observe incoming and outgoing messages from the VPN server and possibly determine which sites Emilia is visiting.

A curious question is: How can Emilia share keys with Tor nodes without revealing her identity? Solving this problem is a two-part puzzle.

First, how can two nodes cooperate to create a key only known to them on a public network where anyone can read all communication? The answer is the Diffie-Hellman key Exchange (DHE) protocol. The idea is that first, both parties need to individually generate their own private secrets that they combine into a shared secret (Kn1) that only the two of them can compute. In practice, authenticated ECDHE based on elliptic cryptography is used to solve issues with vanilla DHE.

At this point, Emilia could go to each Tor node and establish a key with them individually, but that would reveal her identity to each of them. The second piece of the puzzle is establishing the keys using DHE. Instead of communicating directly with all three nodes, after establishing a key with the entry node, Emilia's computer encrypts all messages with Kn1 and sends the message to the relay node through the entry node. This means that the relay node only knows the entry node and not Emilia. Similarly, Emilia's computer encrypts DHE messages with Kn2 and Kn1 and sends them through the guard and relay nodes to the exit node.

Unfortunately, barely after Emilia learned about Tor and started using it, the Lemonheads started censoring connections to publicly advertised Tor node IPs. To counter such efforts, volunteers started running secret Tor bridges (private replacement of publicly advertised entry nodes) and made them available in small batches to only a few users at a time.

Making the matter worse for Emilia and her fellow orange-lovers, researchers found that they can discover Tor bridges by scanning the entire IPv4 space. We do not cover all challenges using Tor and additional issues are discussed in the DefCon 2022 Tor Presentation.

In conclusion, it turns out to be a continuous cat and mouse game for someone like Emilia to maintain anonymity online.

Malicious and Benign Use Cases for Tor

Internet users utilize Tor for many malicious and benign purposes. Political activists, similar to Emilia in the examples, want to make sure that their identity remains secret and they cannot be tied back to activities condemned by their government. Other users might want to protect their privacy and keep the sites they visit secret, even if the activity is not illicit where they live.

People might use Tor to reach geographically restricted content or to circumvent censorship by their government or content blocking by their institution. For example, if Tor traffic is not blocked, customers of Advanced URL Filtering could not stop employees using Tor from circumventing category-based filtering.

Tor is also famous for its onion services. For example, Tor helps to hide multiple whistleblower websites where users can report illicit and immoral activities in their organizations without having to worry about retaliation. An onion service keeps its IP address secret by allowing users to connect only using Tor. The idea is that both the user and the onion service connect through Tor, and they meet in the middle at a rendezvous point (a Tor node). While the goal of these onion services is not necessarily to enable illicit activities, past studies have found that Tor users established a large fraction or the majority of Tor hidden services for illegal purposes. However, only 6.7% of all Tor users connect to hidden services. The vast majority of users visit clear websites that are less likely to be illicit than onion services. For example, more than a million people use Tor to view Facebook's hidden service allowing access from areas where governments censor it.

Attackers can leverage Tor for their activities too. Attacks usually start with reconnaissance, where the attacker explores the target's infrastructure and searches for potential vulnerabilities, for example, by scanning for open ports and running services. Using Tor, attackers can hide their location and distribute their activity to multiple exit nodes.

Similarly, malicious actors can use Tor for later steps of an attack, such as exploiting vulnerabilities found during reconnaissance, updating malicious code on the target's machine, command and control communication, and data exfiltration. Other malicious uses of Tor include DoS attacks, fake account creation, spamming and phishing.

Miscreants have utilized Tor for ransomware attacks in a variety of ways. In the case of Ryuk and Egregor ransomware, the initial Remote Access Trojan (RAT) called SystemBC used a Tor hidden service as a backdoor for command and control communications. Using Tor hidden services for command and control is useful when building bots as this makes the command and control hard to take down and maintains its accessibility unless connections to Tor are blocked, for example, by using various Palo Alto Network products. The Gold Waterfall threat group also used Tor for backdoor communication when installing DarkSide ransomware. Tor hidden service-based leak sites also have been utilized to host stolen data related to DarkSide and Ranzy locker. Additionally, DoppelPaymer used Tor payment sites to collect ransoms. Unit 42 recently published research on Cuba Ransomware, which uses a Tor hidden service-based leak site, and BlueSky ransomware, which sends a ransom note instructing targets to download the Tor browser as part of the process of regaining access to their files.

Tor usage is not specific to malware targeting servers and personal computers. A type of Android malware also uses a Tor hidden service as a command and control server to make takedowns hard.

Additionally, researchers found that Tor is used to send various malicious spam messages, often in the form of comments and dating spam. The authors also found that emails sent through Tor can contain severe threats, including the distribution of AgentTesla RAT, Adobe-themed phishing emails and Covid-19 loan scams.

How Do Criminals Using Tor Get Caught?

While Tor provides better anonymity than many other solutions, it is not perfect.

In 2013, a Harvard student tried to avoid taking one of his final exams by sending a bomb threat. He connected through Tor to an anonymous email provider to keep his identity secret. Then he used this email provider to send the bomb threat. While he used Tor properly, the student made a big mistake when he connected to Tor from Harvard's wifi network. The student's mistake was that Tor hides what you do, but not the fact that you use Tor. Authorities found out from the email's headers that someone used Tor to send the email. From there, they checked the network logs to see if any student connected to Tor around the time the university received the email. Connecting the dots, they found the culprit, who faced criminal charges.

Ross Ulbricht, the creator of the infamous onion service Silk Road, also used Tor correctly but made a different operational mistake that led to his arrest. Silk Road was the most well-known dark web market at its time, where sellers offered goods like drugs, counterfeit cash, forged ID documents and firearms. The FBI found that early on, someone using the pseudonym "Altoid" was astroturfing to promote the Silk Road marketplace. Eight months later, Ulbricht posted a job advertisement using this pseudonym and the contact rossulbricht@gmail[.]com to hire an IT expert who can help with "a venture backed Bitcoin startup company." The FBI was reportedly then able to access both the logs of a VPN server Ulbricht used and Google's log of access to his Gmail address. Both records pointed to an internet cafe in San Francisco and led to his arrest. (Ulbricht's mistake is similar to the mistake made by Gordon in our initial example, where he was caught because of his use of the nickname Orange Cake across multiple services.)

Attackers can deanonymize Tor users via other methods, for example, by using JavaScript or by setting up rogue Tor nodes (which might be happening in the real world now). The takeaway is that while Tor does offer a level of anonymity, users can leak their identity via operational mistakes, or they can be identified if the observer is determined and has the resources.

Methods to Block Tor Traffic

Malicious or Illicit Use Cases Blocked	Blocking Tor Exit Node IPs	Blocking Tor Guard Node IPs	Blocking Tor Traffic
Reconnaissance	Yes	-	-
Vulnerability Exploitation	Yes	-	-
C2 Communication	Part 1*	Partially	Part 2*
Data Exfiltration	Part 1*	Partially	Part 2*
DoS Attack	Yes	-	-
Evasion of content blocking	-	Partially	Yes
Evasion of Geo Restrictions	-	Partially	Yes
.onion sites	-	Partially	Yes

Table 2. How different methods can be used to stop malicious actors from leveraging Tor. The cells labeled Part 1* and 2* mean that the two solutions together need to be used to protect the enterprise.

To stop traffic to and from the Tor network, we can either block publicly advertised Tor IPs or identify and block Tor application traffic. Table 2 summarizes the use cases for each type of blocking mechanism. First, we can use the list of known exit node IPs to block attacks from Tor such as reconnaissance, exploitation, command and control communication, data exfiltration and DoS attacks. Using the list of known guard node IPs, we can stop our users and their machines from sending traffic to Tor and prevent data exfiltration, command and control communication, evasion of geo-restrictions and content blocking, and visits to `.onion` sites.

As the list of Tor bridge nodes is unknown, guard node IP-based blocking is only a partial solution. Instead, we can directly detect and block Tor traffic using App-ID, the Palo Alto Networks traffic classification system. Next to using available guard and bridge node IPs, App-ID looks at characteristics of connections – such as the cipher suite used or the size of data packets – to identify Tor traffic.

Furthermore, an attacker can initiate data exfiltration and command and control communication from the Tor network or the compromised machine. Therefore, to halt these attacks, it is best to use both exit IP-based and traffic analysis-based blocking.

Palo Alto Networks collects all publicly advertised Tor exit IPs and builds a circuit using each of them to test whether they work. The known and working Tor exits list constitutes our predefined Tor Exit IP External Dynamic List.

Using the Tor Exit IP External Dynamic List and App-ID, we observe that Tor usage is common in the enterprise, as we identified 6,617,473 sessions on 691 devices within 204 customer networks during one month.

Along with blocking Tor traffic, an enterprise can leverage endpoint protections such as Cortex XDR to provide coverage for Tor-based threats. Cortex XDR builds on user and entity behavior analytics (UEBA), endpoint detection and response (EDR), network detection and response (NDR) and cloud audit logs to detect the following activities:

- Possible network connection to a Tor relay server
- A successful VPN connection from Tor
- A successful login from Tor
- Suspicious API call from a Tor exit node
- A successful SSO sign-in from Tor

Conclusion

Tor provides anonymity to its users, which they leverage for both benign and malicious uses. On the one hand, Tor can help political activists in oppressive regimes, improve privacy and protect

whistleblower websites. On the other hand, Tor is useful for various malicious activities, including anonymous reconnaissance, data exfiltration, evasion of geo-restrictions, evasion of content blocking, and the running of illicit marketplaces on the dark web.

As cybercriminals often use Tor for malicious purposes, blocking Tor traffic in an enterprise setting is advisable. We find that attempts to use Tor are common, and we identified 6,617,473 sessions to or from 691 devices on 204 customer networks in one month. Palo Alto Networks provides two solutions for blocking Tor traffic as part of Threat Prevention that are best used together. We provide a verified and built-in Tor Exit IP External Dynamic List to our customers that they can use to block connections to Tor Exit nodes. Additionally, Tor traffic in the enterprise network can be blocked using the Palo Alto Networks traffic classification system App-ID. Furthermore, customers can leverage Cortex XDR to alert on and respond to Tor-related activities on endpoint devices, in the network or in the cloud.

Acknowledgments

We want to thank Michael Giuntoli, Yue Guan, Russell Holloway, Erez Levy, Daiping Liu, Erica Naone, Jason Reverri, Siddhart Shibiraj, Zachary Weinberg, Zhibin Zhang and Jimmy Chen for their invaluable input on this blog post.

Appendix: More Details on the DHE Key Exchange

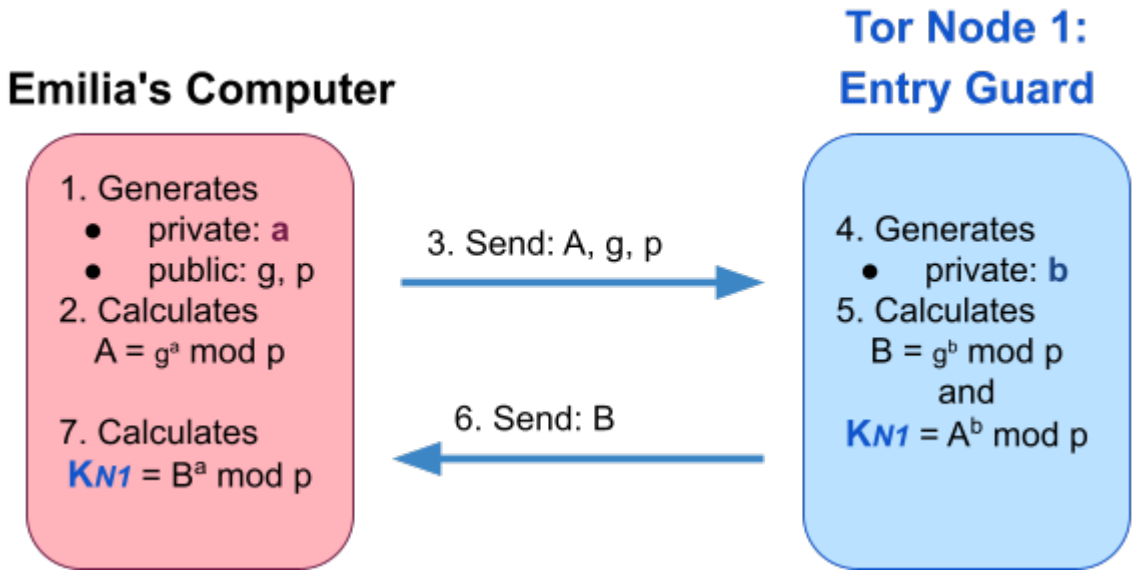

Figure 3. How the basic DHE protocol for key exchange would look between Emilia's computer and the Tor guard node. Note that in practice ECDHE (DHE with elliptic curve crypto) is used with node authentication.

We explain the DHE protocol in Figure 3. The idea is that first, both parties need to individually generate their own secret (a and b randomly selected) that they combine into a shared secret (Kn1) that only the two of them know. They achieve this by selecting large public prime numbers g and p and calculating A (= g^a mod p) and B (= g^b mod p), the public counterparts of their secrets (a and b). The trick is that only Emilia can calculate Kn1 (= B^a mod p) from public B, and only the entry guard can calculate Kn1 (= A^b mod p) from public A (leveraging that $(g^a)^b = (g^b)^a$).

There are two drawbacks to DHE. First, it is computationally expensive. Therefore, a more advanced version building on elliptic cryptography is used in practice called ECDHE. Second, Meddler-in-the-Middle (MitM) attacks are possible against plain DHE protocols as all messages are public. For example, the Lemonheads could act like they are the entry node to Emilia, generate keys with her, and at the same time pretend that they are Emilia to the entry node and share a different key with it. To counter MitM attacks, Tor uses an authenticated version of DHE. In the case of Tor networks, users authenticate Tor nodes by contacting several known directory authorities to retrieve node identities and certificates.