Tor: The Onion Router: Part 1

Gaurav S. Kasbekar
Dept. of Electrical Engineering
IIT Bombay

References

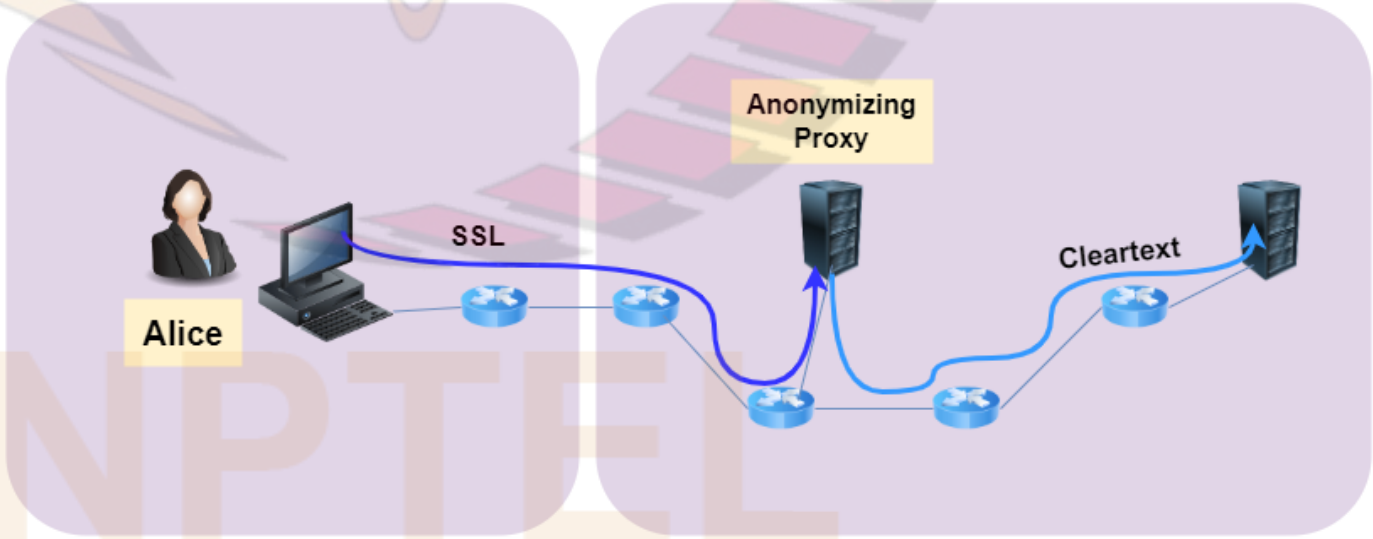
- J. Kurose, K. Ross, "*Computer Networking: A Top Down Approach*", Sixth Edition, Pearson Education, 2013
- Dingledine, R., Mathewson, N. and Syverson, P., 2004. "*Tor: The second-generation onion router"*, Naval Research Lab Washington DC
- Reed, M.G., Syverson, P.F. and Goldschlag, D.M., 1998. Anonymous connections and onion routing. *IEEE Journal on Selected areas in Communications*, *16*(4), pp.482-494.
- The Tor Project:
  - https://www.torproject.org/
- How Tor Works: Parts One, Two and Three by Jordan Wright
  - https://jordan-wright.com/blog/tags/tor/

Example

- Suppose Alice wants to visit a controversial website (e.g., a political activist site)
- Also, Alice:
  1) Does not want to reveal her IP address to the website
  2) Does not want her local ISP (e.g., her home or office ISP) to know that Alice is visiting the website
  3) Does not want her local ISP to see the data she is exchanging with the website
- How can Alice achieve such a connection to the controversial website?
- If she connects using an ordinary TCP connection without TLS:
  - ❑ she does not achieve any of the three required properties
- If she connects using TLS:
  - ❑ she achieves property 3), but does not achieve 1) and 2)
  - ❑ since her source IP address is presented to the website in every packet she sends and the destination address of every packet she sends can be sniffed by the local ISP
- Hence, we need an alternative approach

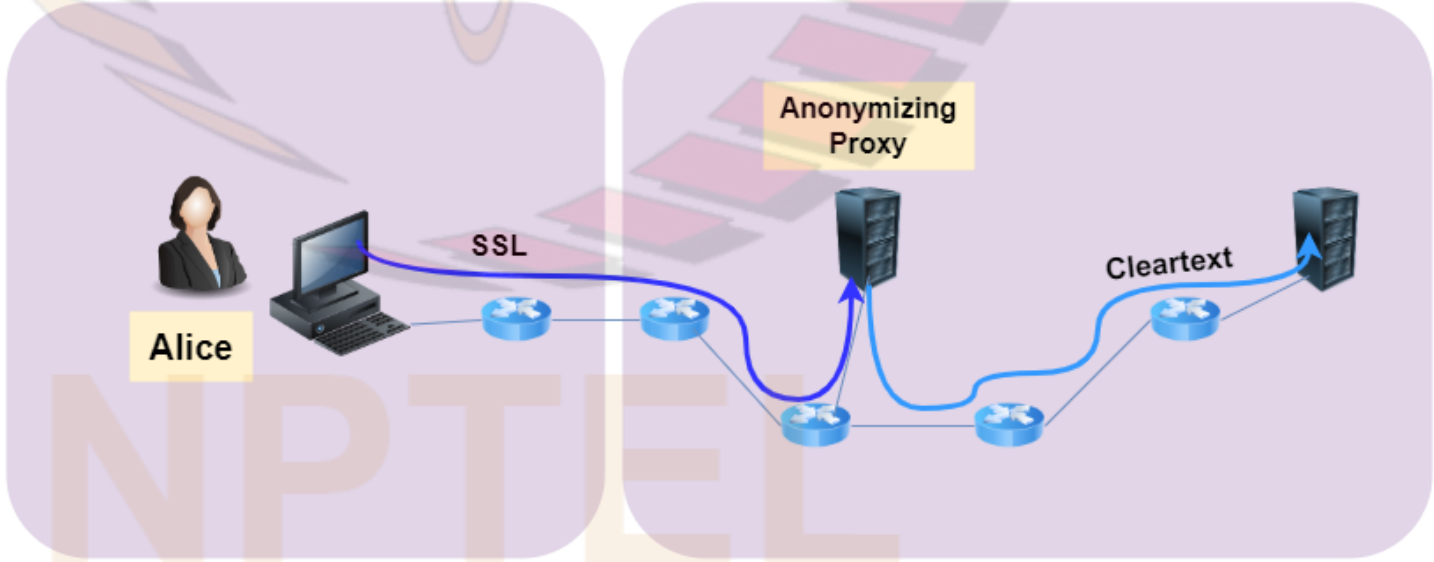
Example (contd.)

- Alice can use a combination of a *trusted proxy server* and TLS
- Alice first makes a TLS connection to the trusted proxy
- She then sends into this TLS connection an HTTP request for a webpage at the desired website
- When the proxy receives the TLS-encrypted HTTP request, it decrypts the request and forwards the cleartext HTTP request to the website
- The website responds to the proxy, which in turn forwards the response to Alice over TLS
- Property 1) is achieved because:
  - ❑ the website only sees the IP address of the proxy and not that of Alice's host
- Properties 2) and 3) are achieved because:
  - ❑ all traffic between Alice and the proxy is encrypted
  - ❑ so the local ISP cannot know which website is being visited by Alice or what data she is exchanging with it



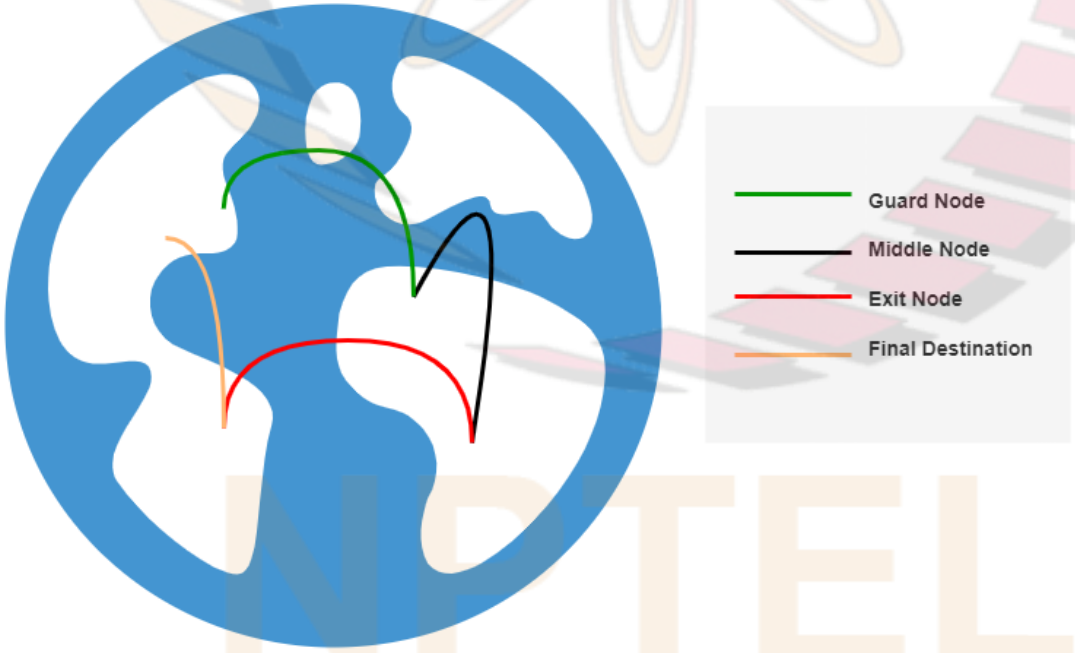
Example (contd.)

- Several companies make such proxy services available, e.g.:
  - https://proxify.com/
- Disadvantage of this solution:
  - the proxy knows Alice's IP address, the IP address of the website she is visiting and can see all the traffic exchanged by Alice and the website
- Hence, this solution fails if the proxy is dishonest
- We want a more robust solution

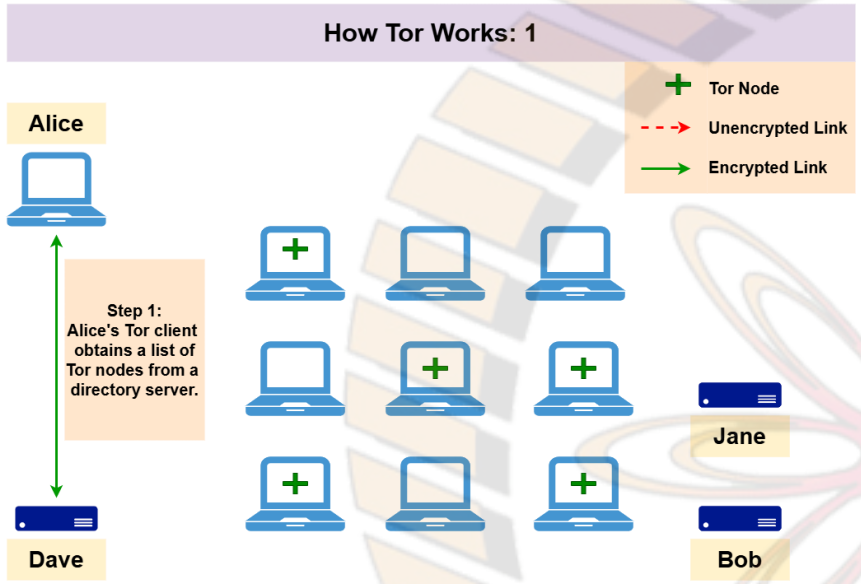


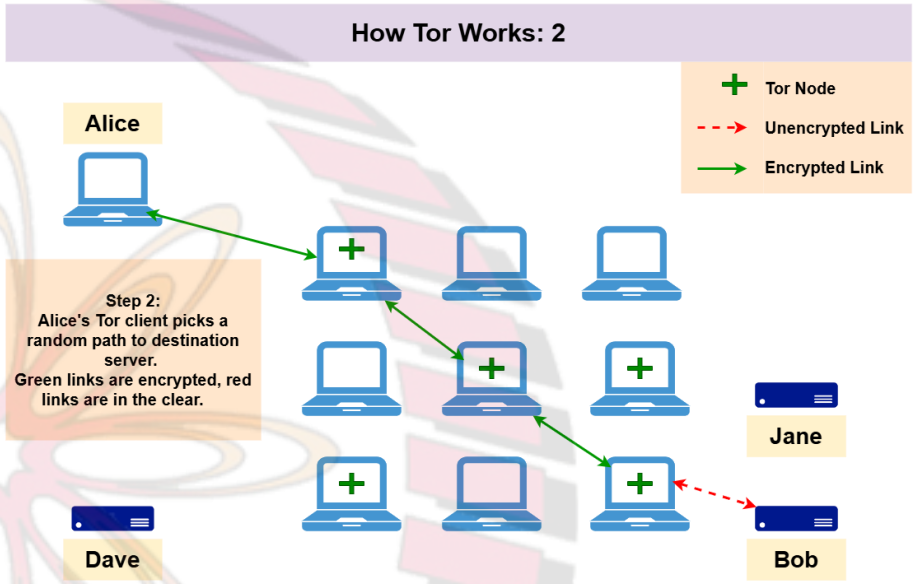
Tor: The Onion Router

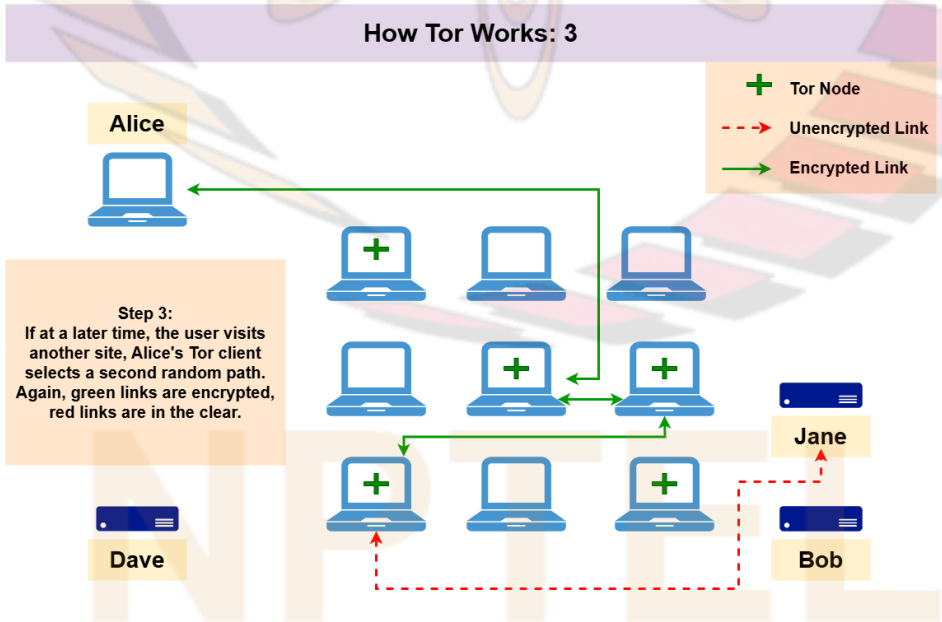
- Routes traffic through three (or more) relays
- Tor allows independent individuals (volunteers) to contribute relays to its relay pool
  - currently, there are about 6000 relays
  - relays do not need any special hardware; they only have to install the Tor software and configure it to act as a relay
  - note that volunteers need to be willing to give up some of their bandwidth
- When Alice connects to a website (destination) using Tor:
  - Tor browser on Alice's host randomly chooses, from the relay pool, a chain of three relays
  - routes all traffic between Alice and the destination over the chain
- Assuming the proxies do not collude, *no one knows that communication took place between Alice's host and the destination website*
- Also, Tor is a censorship circumvention tool:
  - allows its user, say Alice, to reach otherwise blocked destinations or content



Operation of Tor







Applications of Tor

- Individuals use Tor to prevent websites from tracking them and their family members
  - since website does not know IP address of individual browsing it
- Individuals use Tor to connect to news sites, political activist sites, sites like YouTube or Facebook, etc., when these are blocked by their local Internet providers or the countries they live in
- Journalists use Tor to communicate more safely with whistle-blowers and dissidents
- Non-governmental organizations (NGOs) use Tor to:
  - allow their workers to connect to their home website while they're in a foreign country
  - without notifying anybody nearby that they're working with that organization
- Business executives use Tor to keep their strategies confidential
  - e.g., an investment bank may not want competitors to be able to track what web sites their analysts are watching
- Militaries use Tor, e.g.:
  - Insurgents may monitor Internet traffic and discover all the hotels and other locations from which people are connecting to known military servers
  - So military field agents deployed away from home use Tor to mask the sites they are visiting

Limitations of TLS and VPN

- In all the above applications, note that we need to prevent intruders from finding out:
  - the IP addresses of the source and destination of some communication over the Internet
- If Alice and Bob communicate using TLS:
  - an eavesdropper can find out the IP addresses of both of them by looking at the source and destination IP addresses of packets exchanged by them
- If Alice and Bob communicate using a VPN:
  - an eavesdropper can still find out that communication is taking place between some user in Alice's network and some user in Bob's network
  - this may be unacceptable in some applications, e.g., it may reveal that two companies are collaborating

Operation of Tor

- The objective of Tor is to make it difficult for attackers to find out that a given source is communicating with a given destination
- To create a private network pathway with Tor, the client's browser incrementally builds a circuit of encrypted connections through relays on the network
- The circuit is extended one hop at a time
- Each relay along the way knows only:
  - ❑ which node gave it data
  - ❑ and which node it is giving data to
- No individual relay knows the complete path that a data packet has taken
- The first relay on the circuit knows Alice's IP address
  - ❑ so it knows that Alice's IP address is using Tor
    - however, note that Tor is not illegal anywhere in the world
  - ❑ does not know which websites Alice is visiting
- The third relay knows the destination's (Bob) IP address
  - ❑ however, it does not know who is communicating with Bob
  - ❑ can read the data being exchanged with Bob if it is not encrypted, and cannot read it if it is encrypted (e.g., if TLS is used)