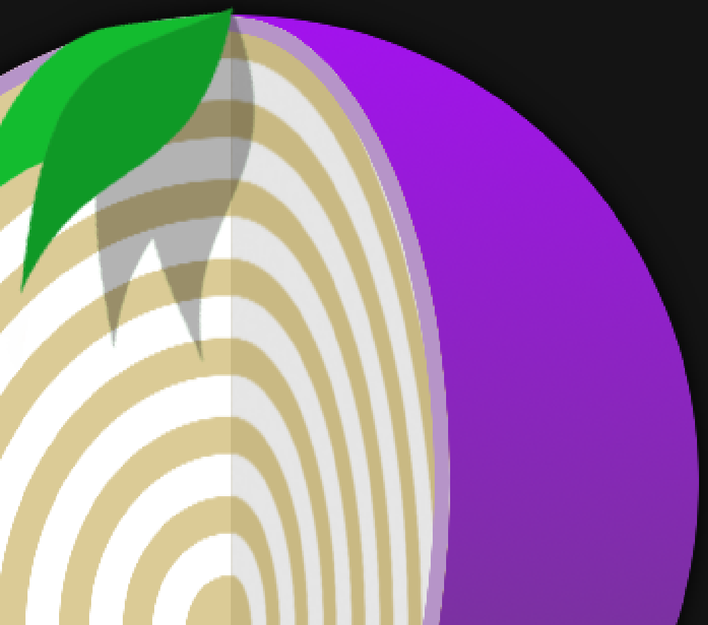
07 August 2023

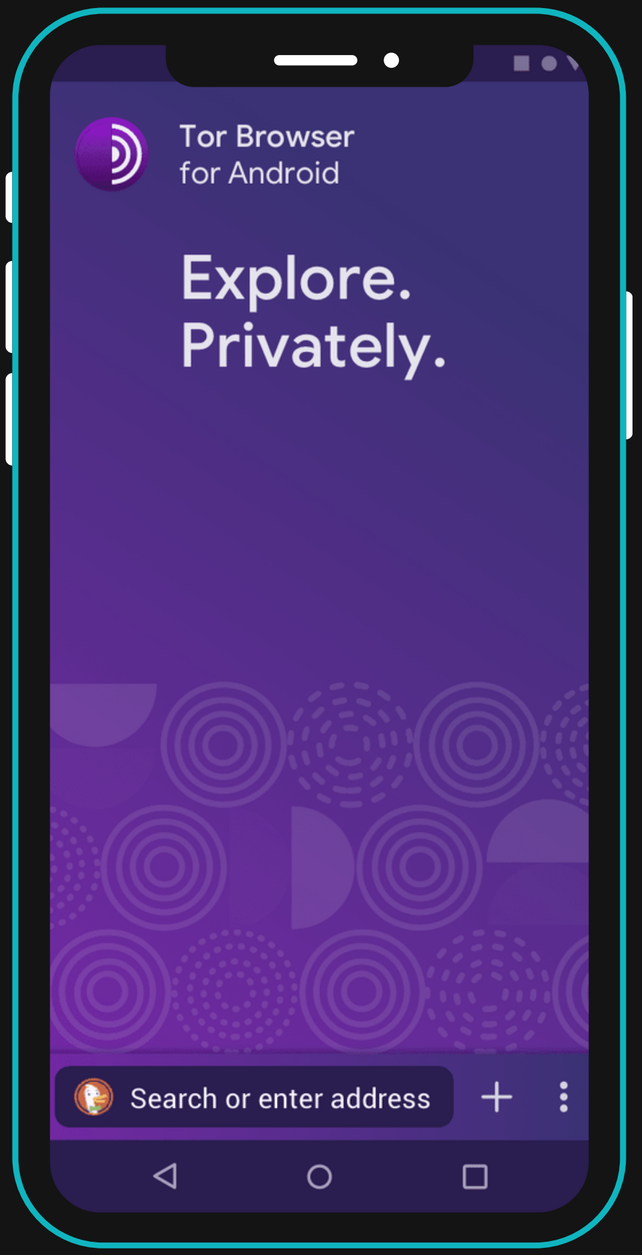
Tor/Onion Hidden Service Deanonymization Techniques.

Group 4

Index

What you need to know

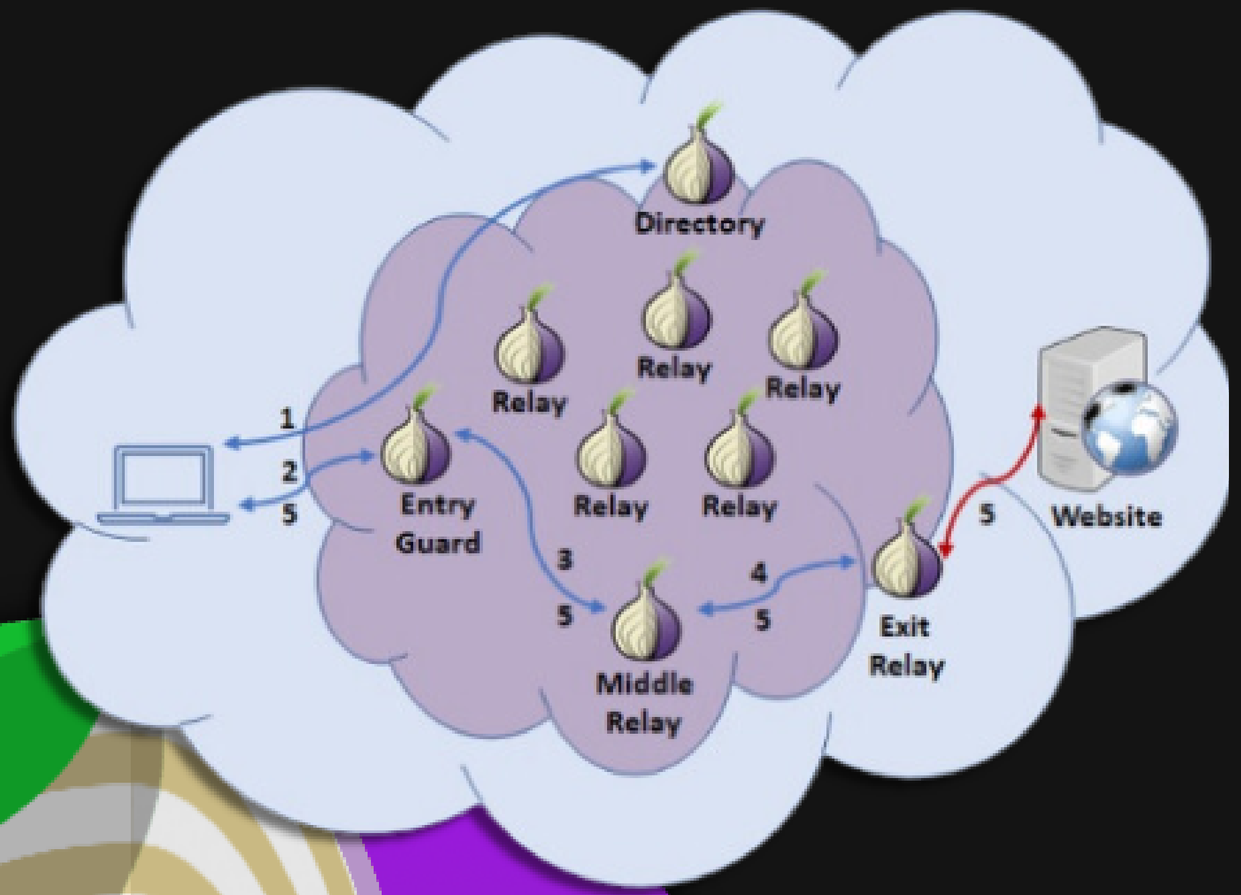




Preserving Privacy with Tor

- Widespread surveillance - increasingly challenging endeavour for internet users.
- Online threats (Spam, identity theft and DoS etc) - heightened user vulnerability and need for robust privacy protection measures.
- In response, anonymity networks have emerged as a crucial solution safeguarding user identities by providing unlinkability between their IP addresses, digital fingerprints, and online activities.
- Tor - granted significant adoption as the foremost privacy-preserving network.

Challenges and Research Landscape of Tor



- Tor's design confronts inherent challenges related to scalability and performance,
- Researchers have diligently sought to enhance the user experience by proposing various improvements,
- A comprehensive understanding of the research landscape encompassing the performance and security of low-latency anonymous communication systems, with a specific focus on Tor.

Tor Hidden Service.

Anonymous Hosting

- Host websites and services anonymously.
- Accessible only through the TOR network.
- ".onion" domain addresses conceal physical location.

End-to-End Encryption

- Layered encryption for content and communication paths.
- Ensures user and service provider anonymity.

Censorship Resistance

- Bypass censorship and access restricted content.
- Difficult for authorities to block or filter.
- Promotes open access to information.

Secure Communication

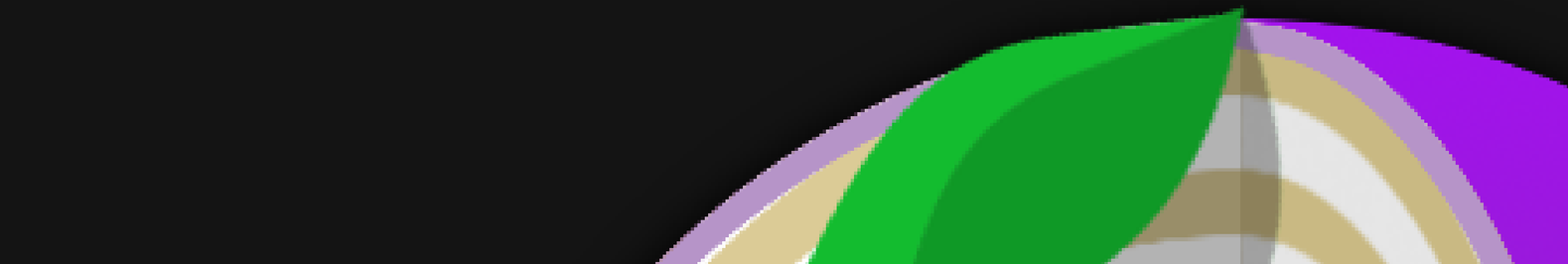
- Enables private and secure communication.
- Ideal for encrypted messaging and forums.
- Guards against surveillance and eavesdropping.

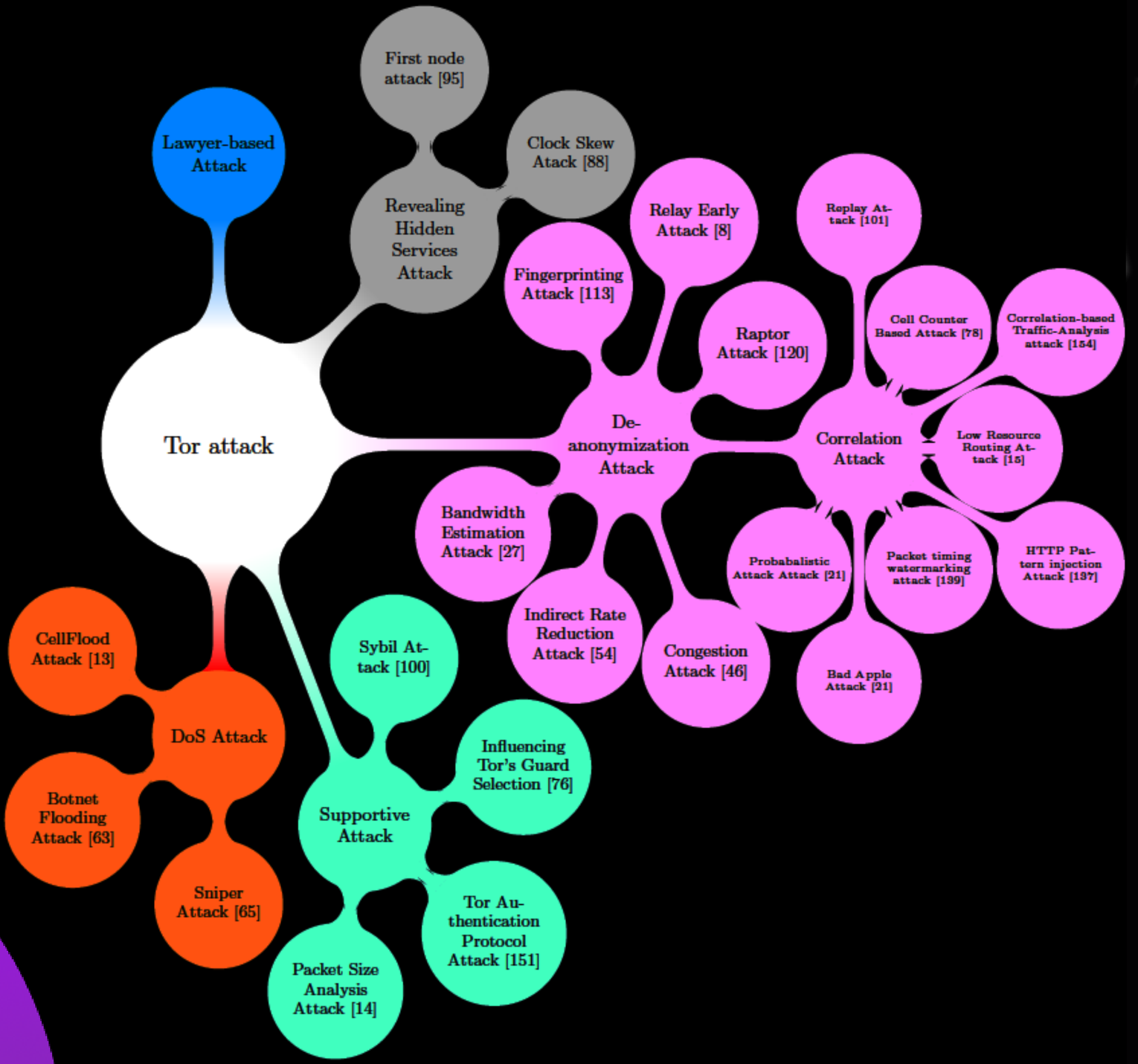
Dark Web vs. Hidden Services

- Hidden Services ≠ Illicit Dark Web.
- Many Hidden Services are legal and legitimate.
- Used for privacy, free speech, activism, and more.

Type of Attacks

Network Distribution Attack
De-anonymization Attack
Censorship Attack
Generic Attack




Tor attack
Lawyer-based Attack
First node attack [95]
Clock Skew Atack [88]
Revealing Hidden Services Attack
De-anonymization Attack
Fingerprinting Attack [113]
Relay Early Attack [8]
Raptor Attack [120]
Correlation Attack
Replay Attack [101]
Cell Counter Based Attack [78]
Correlation-based Traffic-Analysis attack [154]
Low Resource Routing Attack [15]
Bandwidth Estimation Attack [27]
Indirect Rate Reduction Attack [54]
Congestion Attack [46]
Probabalistic Attack Attack [21]
Packet timing watermarking attack [139]
HTTP Pattern injection Attack [137]
Bad Apple Attack [21]
DoS Attack
CellFlood Attack [13]
Botnet Flooding Attack [63]
Sniper Attack [65]
Supportive Attack
Sybil Attack [100]
Influencing Tor's Guard Selection [76]
Tor Authentication Protocol Attack [151]
Packet Size Analysis Attack [14]

ATTACKS	YEAR
Sybil Attack	2016
Guard Selection Attack	2015
RAPTOR Attack	2015
Torben Attack	2015
Circuit Fingerprinting	2015
The Sniper Attack	2014
BotNet Flooding Attack	2014
Relay Early Attack	2014
CellFlood Attack	2013
Hidden Service Attack	2013
Indirect Rate Reduction Attack	2012
HTTPOS Website Fingerprinting	2012
StegoTorus Attack	2012
HTTP-based application-level attack	2011
Packet Size Attack	2011
Bad Apple Attack	2011
Loop Attack	2011
Throughput Fingerprinting	2011
Traffic Analysis Attack	2010
Bandwidth Estimation Attack	2010
Passive Linking Attack	2010
Client Location Attack	2010
Adaptive Surveillance Attack	2010
Cell Counter Based Attack	2009
Protocol-level Attacks	2009
FortConsult Security Attack	2009
Bayesian Traffic Analysis Attack	

Preserving Privacy with Tor.

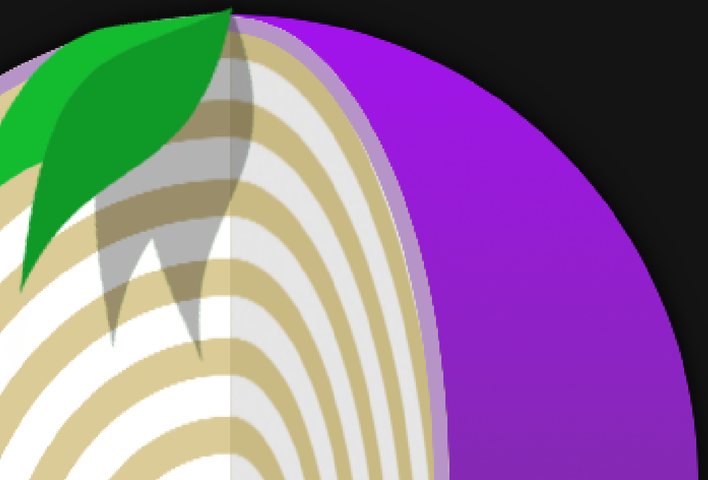
We present an overview of published attacks on the Tor network and proceed to discuss them. All the attacks are listed chronologically in the table.

ATTACKS	YEAR
Practical Congestion Attack	2009
Website Fingerprinting	2009
Bridge Deanonymization Attack	2009
Link-Based Relay Selection Attack	2009
Tor Authentication Protocol Attack	2009
AS Awareness Attack	2009
Route Fingerprinting	2008
Package Spinning Attack	2008
Replay Attack	2008
Passive logging Attack	2008
Low Resource Routing Attack	2007
Connection Start Tracking attack	2007
Packet Counting Attack	2007
Stream Correlation Attack	2007
Packet Timing Watermarking Attack	2007
Clock Skew Attack	2006
First Node Attack	2006
Congestion Attack	2005
Predecessor Attack	2004
Intersection Attack	2004
Active n - 1 Attack	2003
Website Fingerprinting	2003
Robust Watermark Correlation Attack Correlation	2003
Attack + Timing Attack	2003
Statistical Disclosure Attack	

Preserving Privacy with Tor.

We present an overview of published attacks on the Tor network and proceed to discuss them. All the attacks are listed chronologically in the table.

De-anonymization Attack on TOR.



De-anonymization Attack on Tor.

PASSIVE ATTACK

A type of attack where the attacker eavesdrops and monitors network traffic.

ACTIVE ATTACK

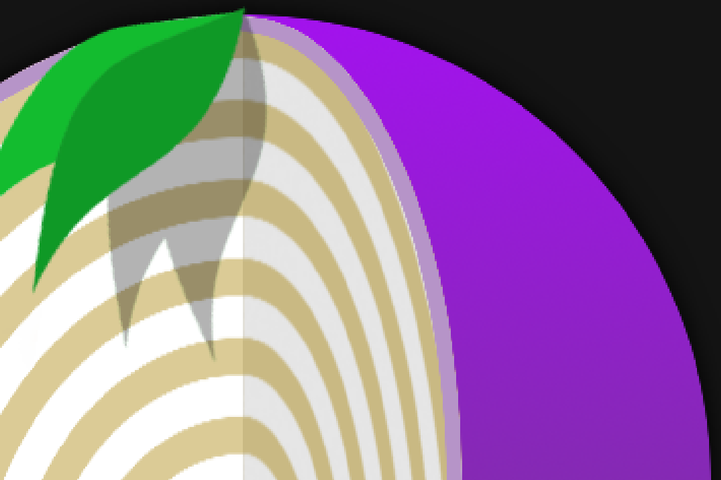
A type of attack where the attacker actively manipulates or modifies network traffic.

SINGLE-END ATTACK

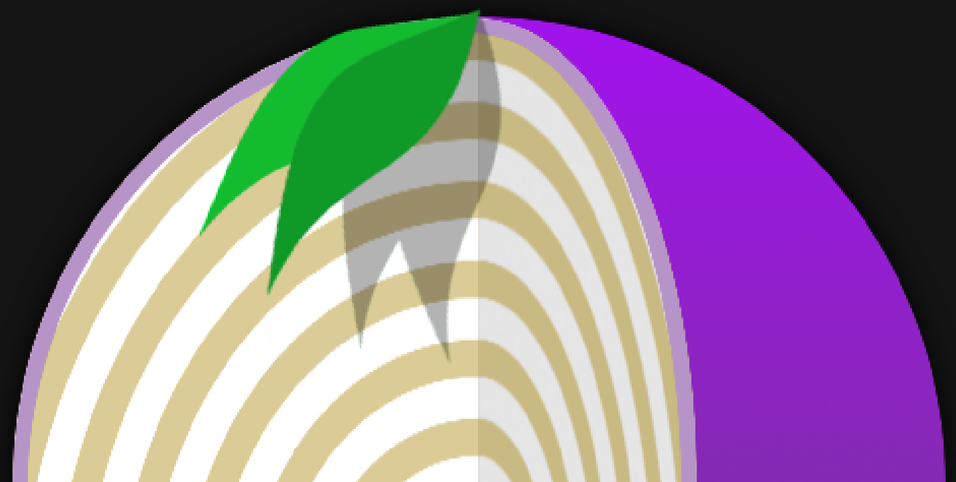
An attack that targets either the user's end or the server's end of a communication.

END-TO-END ATTACK

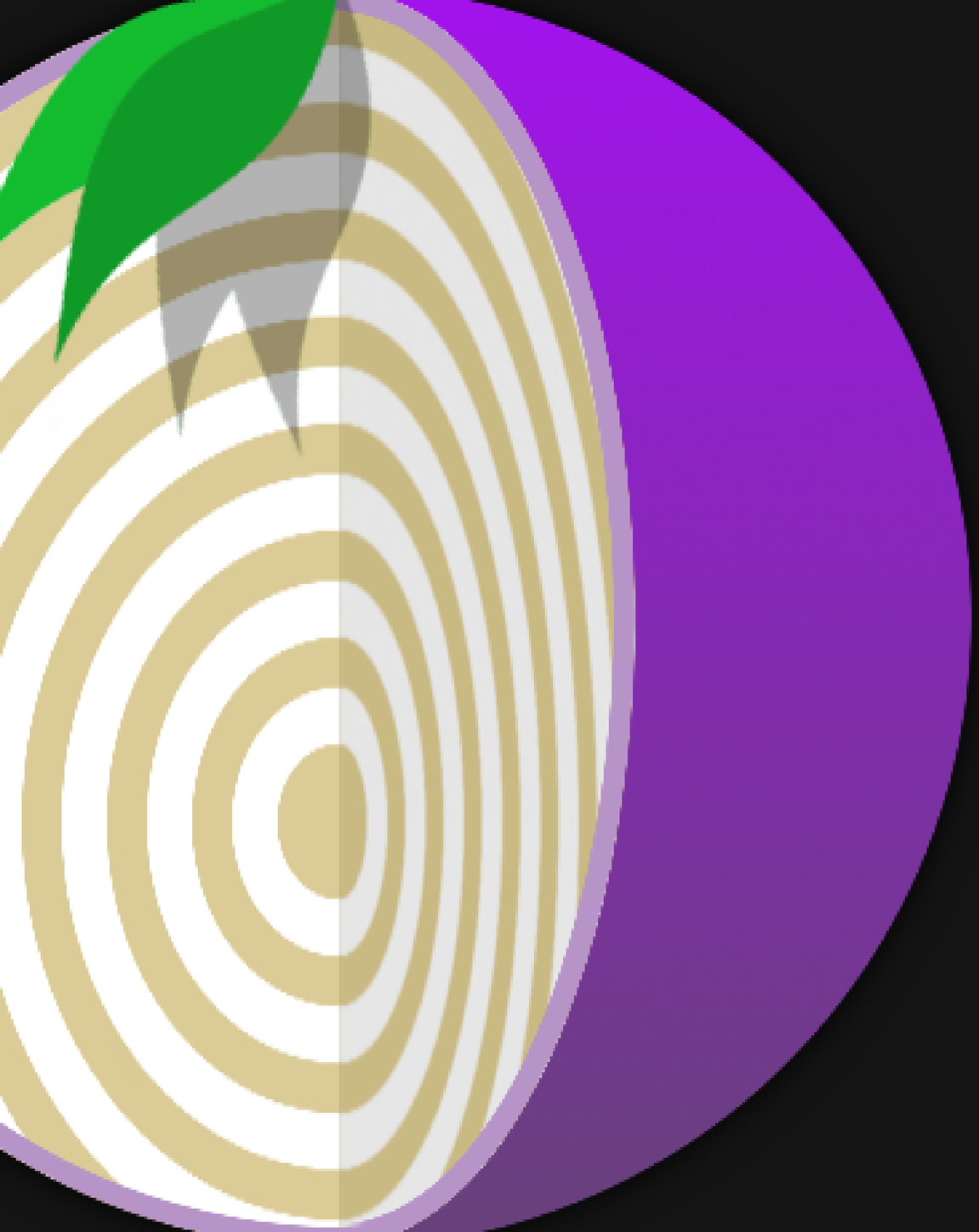
An attack that targets both ends of a communication, compromising the entire process.



7 Categories Based on Method and Goal.



1. Correlation Attack
2. Congestion Attack
3. Timing Attack
4. Fingerprint Attack
5. Denial of Service attack
6. Supportive Attack
7. Reveling Hidden services attack



Sybil Attack

Supportive

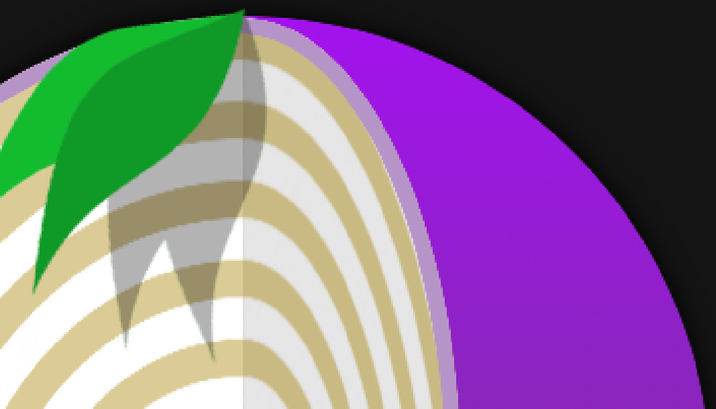
SYBIL ATTACK

OVERVIEW

- Attacks involves the creation of fake nodes.
- Attacker can cover significant area of network with fake nodes which are under his control.
- The Tor network relies on the assumption that the majority of the nodes are honest and not under the control of a single entity. However, a Sybil attacker can compromise the integrity of the Tor network by operating multiple nodes that are under their control.

TYPES

- Proliferation
- Traffic Contol
- Traffic Analysis
- Denial of Service (DoS)



SYBIL ATTACK

LIMITATIONS

- Network Diversity
- Guard Nodes
- Random Path Selection
- Difficulty in Node Placement
- Resilience to Node Failures
- High Cost and Resource Requirements
- Detection and Countermeasures

Sniper Attack

DOS Attack

SNIPER ATTACK ON PRIVACY IN TOR.

OVERVIEW

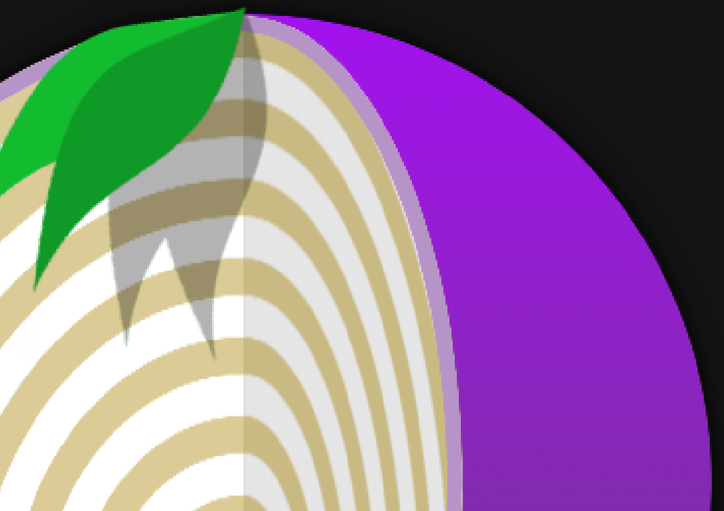
- A low-cost but highly destructive DOS attack.
- Anonymously disable TOR relays.

HOW IS THIS DIFFERENT?

- Adversary utilizes valid protocol messages
- which consume memory in the target node.
- results in exploiting TOR's end-to-end data transport.

VULNERABILITIES EXPLOITED

- Availability: This attack exhausts the host's memory which results in termination by it's OS memory manager.



IMPLICATIONS

- Exposes onion router identity and user details.

DEFENCE AGAINST SNIPER ATTACK

- have the guard node monitor its queue length & destroy the circuit if it ever exceeds 1000 cells
- Use a adaptive out-of-memory circuit killer

HOW IT WORKS ?

CHOOSING THE ENTRY POINT

- The target relay is used as the entry point in the initial step of a client establishing a circuit.
- This relay receives information from the client and forwards it through a series of relays.

SENDING REQUESTS TO THE SERVER

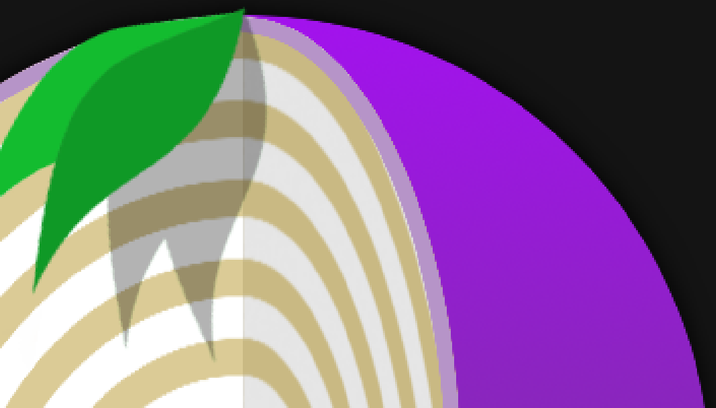
- The server then responds with the requested data after receiving a request for it from the client.

STOP READING FROM TCP CONNECTION

- The client then stops actively reading from the TCP connection to the target entry.
- But it still occasionally sends SENDME packets to the server without reading any data from it.

KILLING THE PROCESS

- The target entry stores the buffered data up until the operating system kills the Tor process, ultimately making the attack successful and deanonimizing the hidden service's anonymity.



Raptor Attack

Correlation + Supportive

ROUTING ATTACKS ON PRIVACY IN TOR.

OVERVIEW

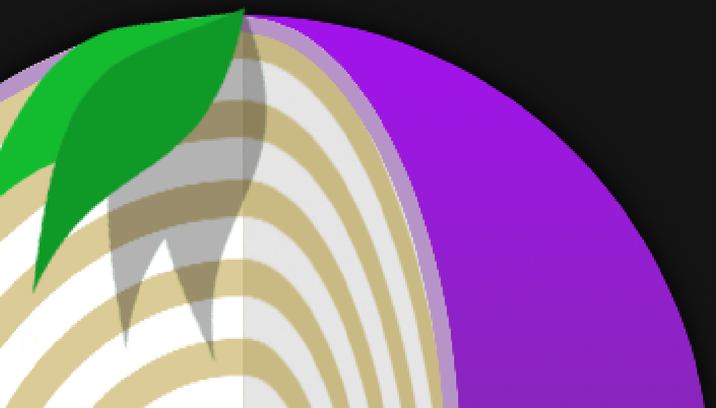
A sophisticated deanonymization technique targeting the Tor network, aiming to expose the identities of Tor users and their communication destinations.

COMPONENTS

It combines asymmetric traffic analysis, BGP path changes, and BGP hijack/interception to achieve a higher success rate in deanonymizing Tor users.

VULNERABILITIES EXPLOITED

The attack leverages Border Gateway Protocol (BGP) vulnerabilities, allowing malicious actors to interfere with Tor traffic.



IMPLICATIONS

Successful RAPTOR Attacks compromise user privacy, expose sensitive activities, and pose significant ethical concerns.

How it works.

ESTABLISHING TOR COMMUNICATION

Tor clients create layered circuits through Tor relays (entry, middle, and exit) to establish communication with a destination server.
Each relay in the circuit only knows the identity of its preceding and succeeding hop, ensuring client anonymity.

ASYMETRIC TRAFFIC ANALYSIS

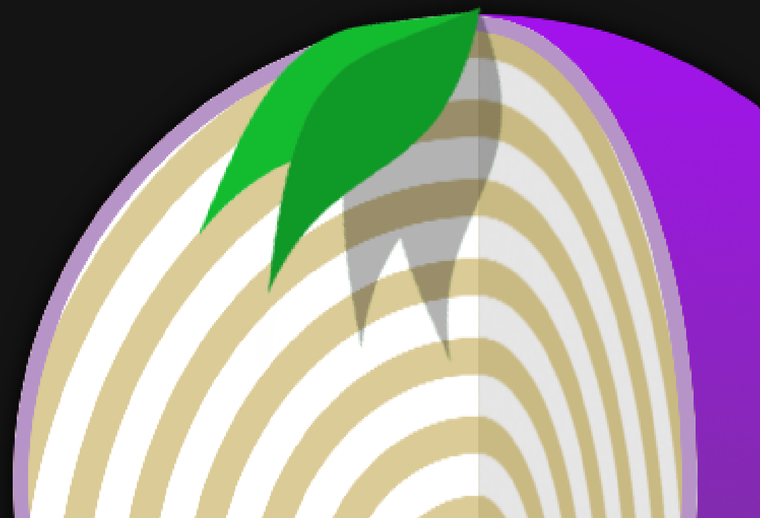
Attacker observes incoming/outgoing traffic at both ends of the communication (client-to-entry and exit-to-destination).
By correlating sequence numbers of data packets and acknowledgments, the attacker can de-anonymize the client and server.
This analysis only requires one direction of traffic at both ends and can involve different autonomous systems (ASes).

BGP PATHS

Attacker exploits the dynamic nature of BGP paths, which can change due to link or router failures.
With each change in BGP paths, the attacker introduces malicious ASes into the communication path, enhancing the chances of correlation.

BGP HIJACK/INTERCEPTION

In a BGP hijack, the malicious AS falsely advertises an IP prefix not belonging to it, capturing traffic intended for that prefix.
In a BGP interception, the malicious AS advertises a false IP prefix but intercepts and analyses the traffic before forwarding it. By intercepting the entry node's prefix, the attacker can identify all IP addresses communicating with it and use asymmetric traffic analysis to de-anonymize the client.



Drawbacks.

LIMITED REALWORLD INCIDENTS

The attack's theoretical nature makes empirical validation challenging due to the lack of documented real-world incidents.

DETECTION COMPLEXITY

Implementing detection mechanisms requires technical expertise and computational resources.

DEPENDENCE ON BGP VULNERABILITIES

Successful exploitation relies on cooperation from multiple autonomous systems.

Preventions.

SECURING INTER-DOMAIN ROUTING

Proposed protocols aim to secure BGP routing to mitigate hijack and interception attacks.

BGP PREFIX LENGTH

Tor relay operators are encouraged to use /24 prefix lengths to reduce BGP hijack vulnerabilities.

GUARD RELAY SELECTION

Tor clients can favor guard relays with shorter AS-level paths to minimize attack surface.

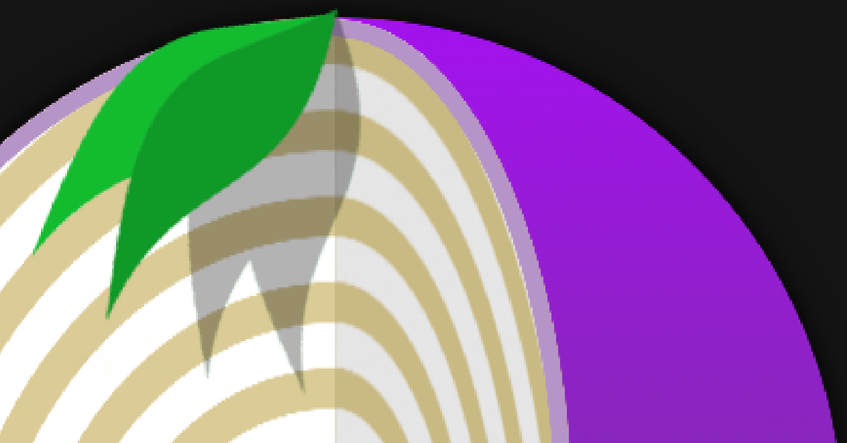
Replay Attack

Correlation

REPLAY ATTACK ON PRIVACY IN TOR.

MONITORING MANIPULATED CELLS - WITH A DIFFERENT IDEA

- Illicit replication and transmission of intercepted network data packets within the Tor network, exploiting weaknesses in encryption and authentication protocols to compromise user anonymity and undermine the integrity of communications.



YOU REPLAY THE CELLS

- Just like one of the methods in Protocol-Level attacks, we replay a cell here

- The difference lies in the approach. The idea in Protocol level attack is to use technques which makes the protocols of the TOR communication against itself.

WHAT ELSE IS THE DIFFERENCE?

- The circuit that the routers communicate with would have thousands of communications, pinpointing the correct message every time is a challenging task.

HOW DOES IT WORK?

STEP 1

- Analyse and find our requests from OP (User system). These are the requests we need to work over.

STEP 2

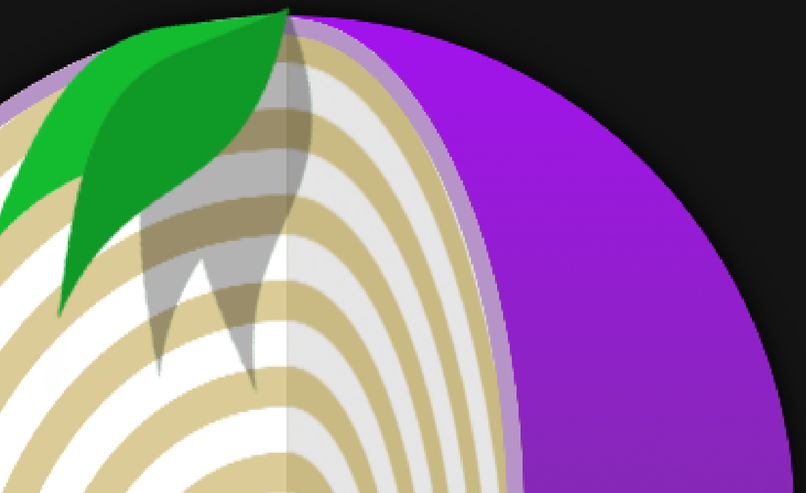
- Figure out the time to attack.
- Select specific cells for the attack

STEP 3

- You manipulate the cells to launch replay attack - duplicate the cell.

STEP 4

- See the error, confirm the relationship between sender and receiver.



DRAWBACKS

FALSE POSITIVES

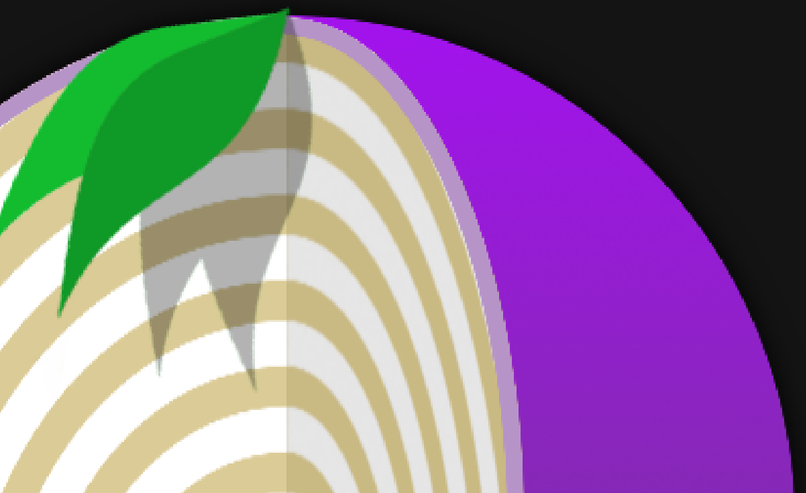
- Replay attacks can misattribute connection drops to innocent exit routers, making it challenging to identify the actual attackers.

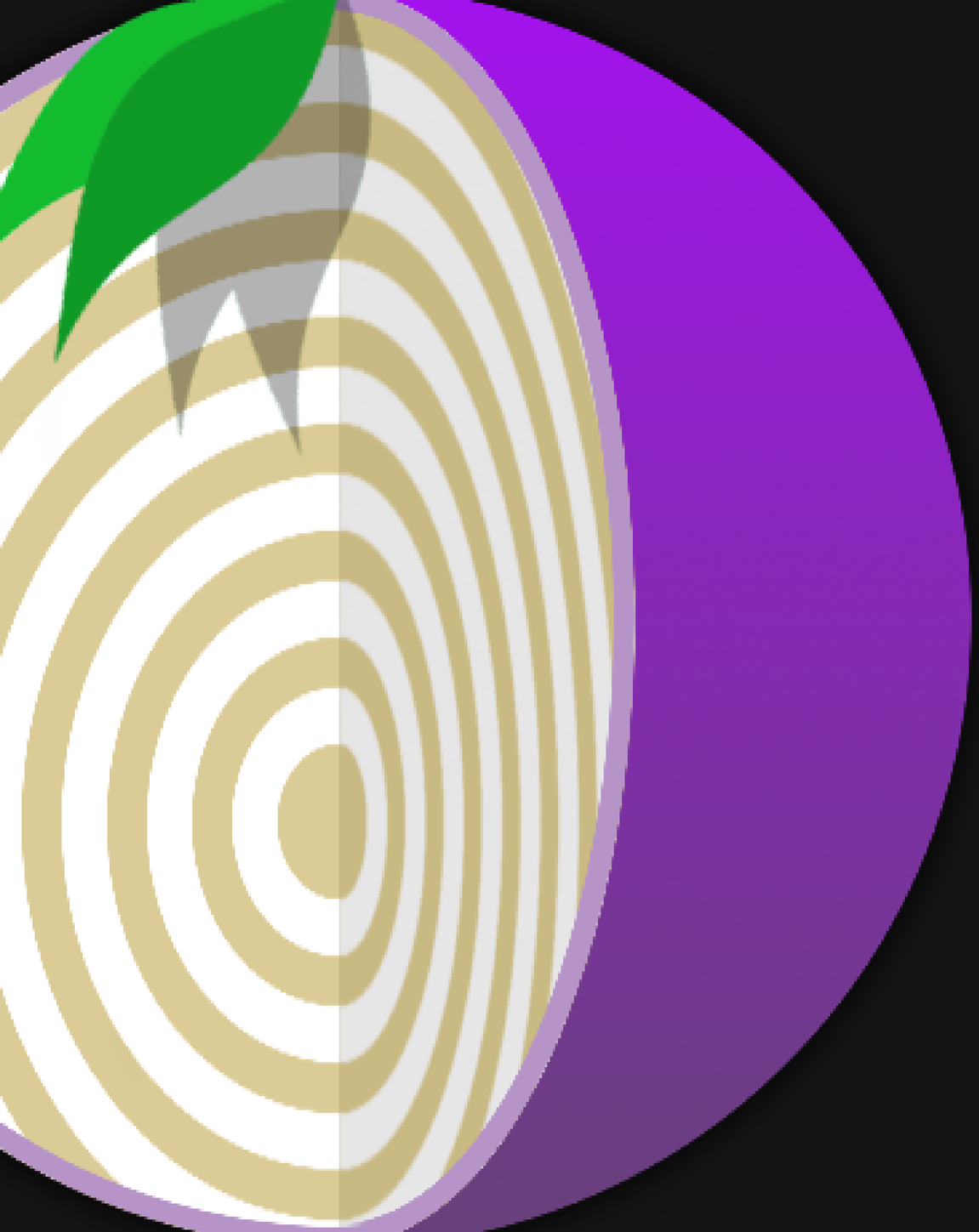
THE TECHNIQUE IS NOT EASY

- Detecting and mitigating replay attacks can be difficult due to Tor's intrinsic anonymity and encryption technologies.

RISK OF STEALTH-BREAK

- In case of the application firing an error, the attacker is at the risk of being exposed to manipulating the cell.





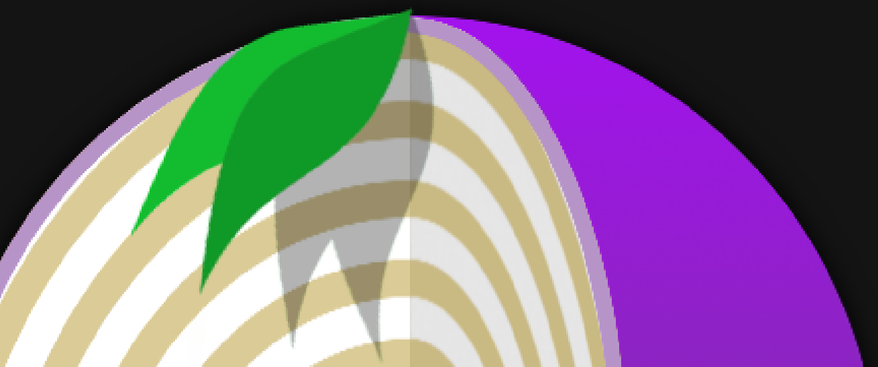
Protocol – Level Attack

Correlation + Supportive

PROTOCOL-LEVEL ATTACK ON PRIVACY IN TOR.

OVERVIEW

- Involves various techniques such as traffic analysis, timing correlation, and active probing.
- Exploit weaknesses in the Tor network's design, potentially revealing the identity and activities of users.
- By analyzing patterns and characteristics of their network traffic.



HOW IS THIS DIFFERENT?

- Goes by the simple idea - You throw a spoilt bacon and see who vomits!
- The detection involves a special algorithm which may or may not work.

IMPLICATIONS

- Exposes onion router identity and user details.

VULNERABILITIES EXPLOITED AND DOES TIME MATTER?

- Integrity: Cells are manipulated and traversed. The strength of Tor to detect integrity issues is used against it.
- Timing does not matter in Protocol-Level Attacks - but monitoring it does bring inputs.

HOW DOES IT WORK?

STEP 1

- Analyse and find our requests from OP (User system). These are the requests we need to work over.

STEP 2

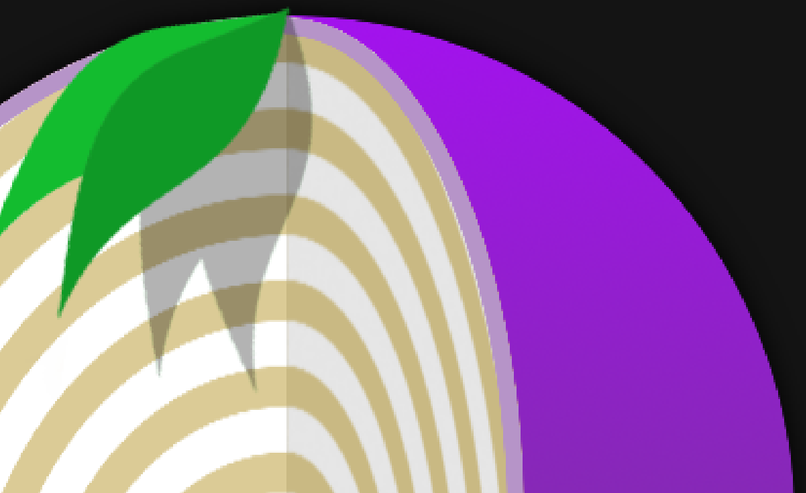
- Figure out the time to attack.
- Select specific cells for the attack

STEP 3

- You manipulate the cells to launch protocol-level attacks in four ways:
  - Replay Cell
  - Cell Modification
  - Insert Faked Cell
  - Delete a Cell

STEP 4

- See the error, confirm the relationship between sender and receiver.



DRAWBACKS

MONITORING MANIPULATED CELLS

- If these cells get detected and dropped before reaching the exit router, the effectiveness of such attacks will be reduced.

USING BRIDGE RELAYS

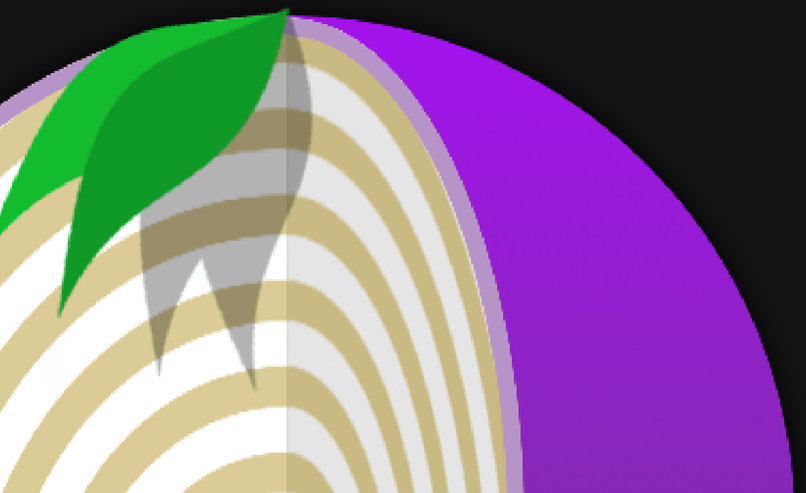
- Bridges will appear to be indistinguishable from clients if attackers use the methods mentioned above to identify clients.
- Therefore, using bridges can resist protocol-level attacks.

THE ENVIRONMENT IN THE CIRCUIT CAN BE NOISY

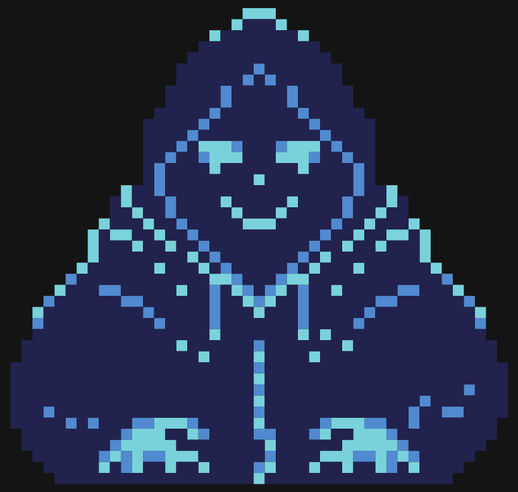
- The circuit that the routers communicate with would have thousands of communications, pinpointing the correct message every time is a challenging task.

RISK OF STEALTH-BREAK

- In case of the application firing an error, the attacker is at the risk of being exposed to manipulating the cell.

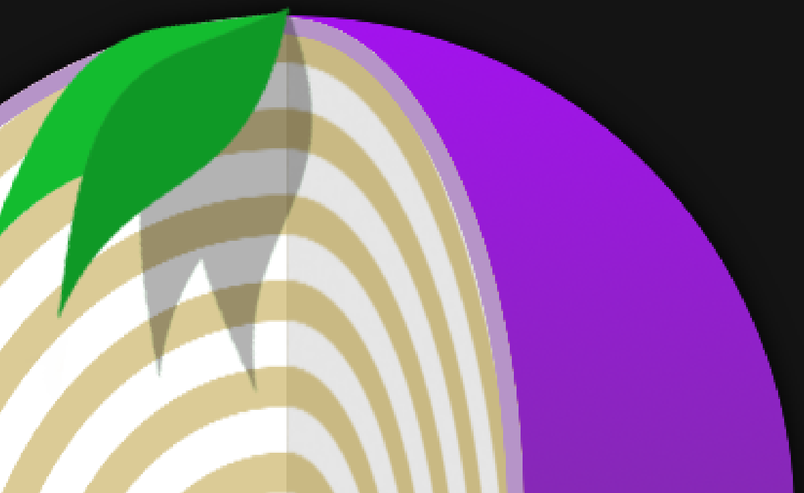


Attack Implementation on our own Hidden Service.



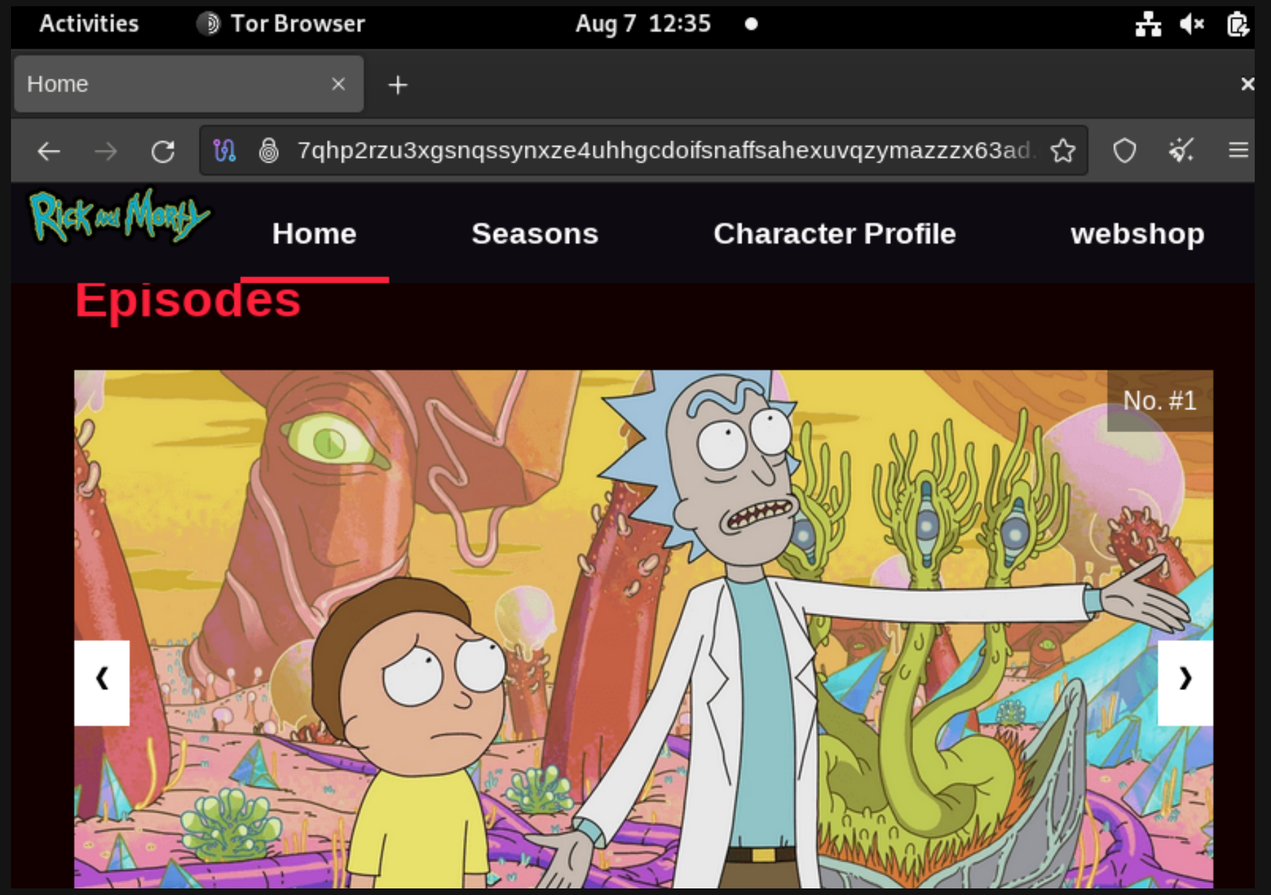
GOAL

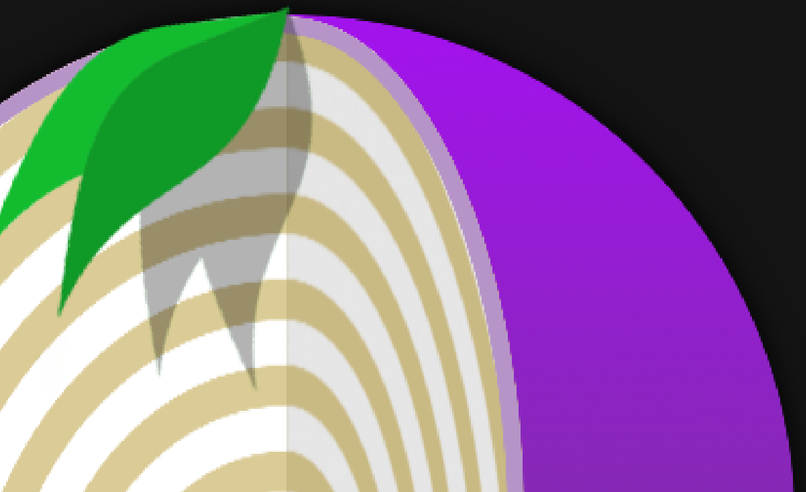
For the implementation of the attack, the main goal was somehow to leak the identity of the hidden service, like IP address, leading to it's deanonyimization.



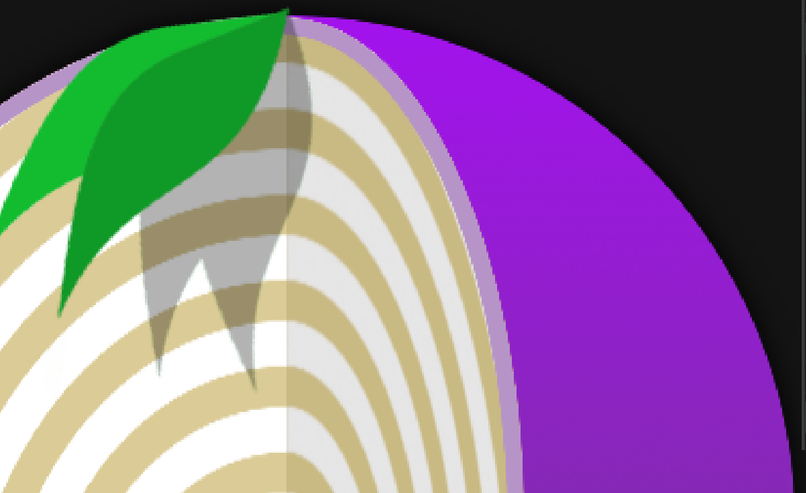
OVERVIEW

- Oracle VirtualBox with 2 machines
- Debian is running a hidden onion service on localhost（Target）
- Kali Linux is used for attacking
- The hidden service is just a static website（for testing purposes）
- URL:7qhp2rzu3xgsnqssynxze4uhhgcdoifsnaffsahexuvqzymazzzx63ad.onion





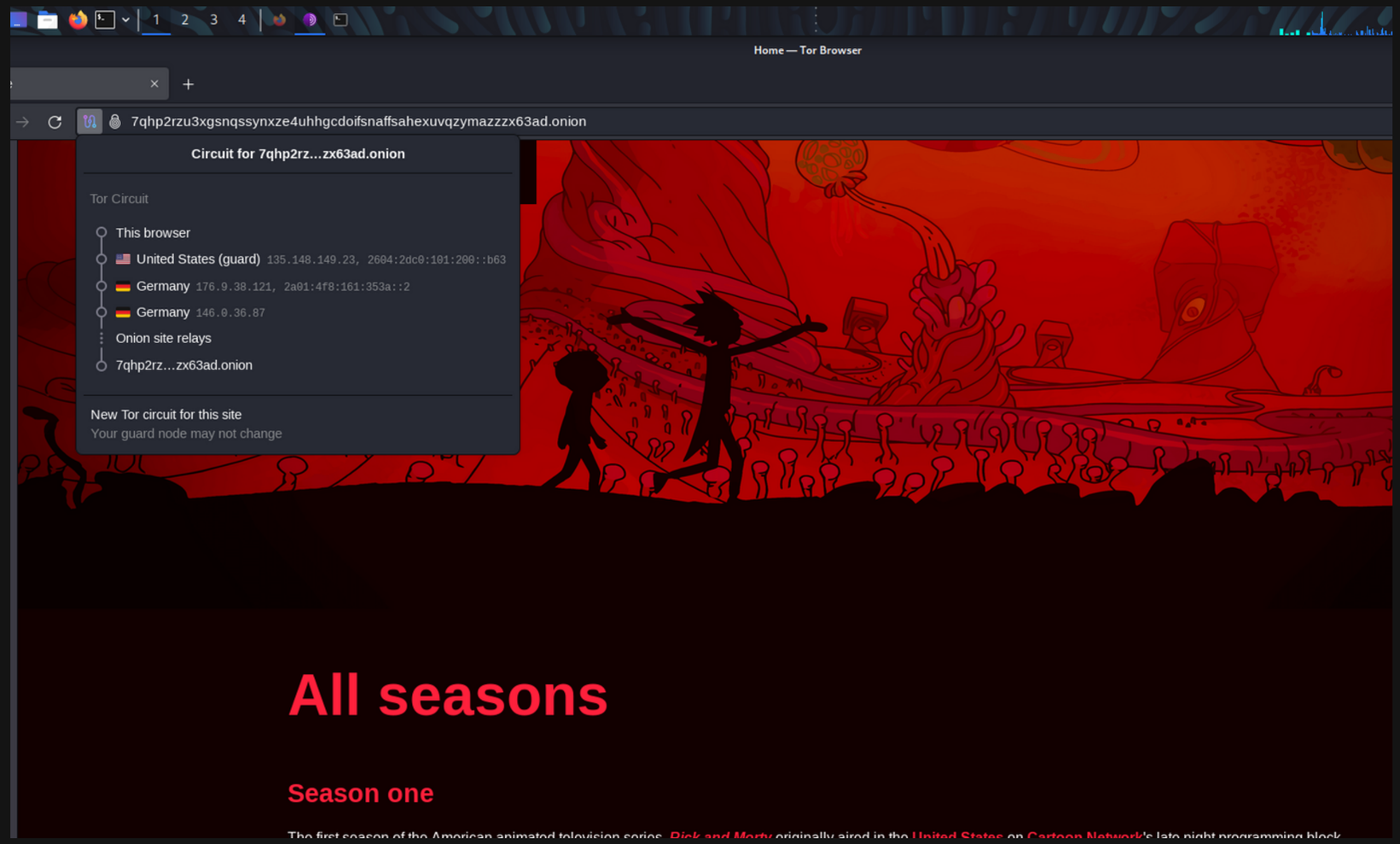
- The Debian system utilizes an Apache web server to host the website on localhost at port 80.
- Also, Tor is used to connect to the TOR networks and it helps to hide the website by making it a hidden service.



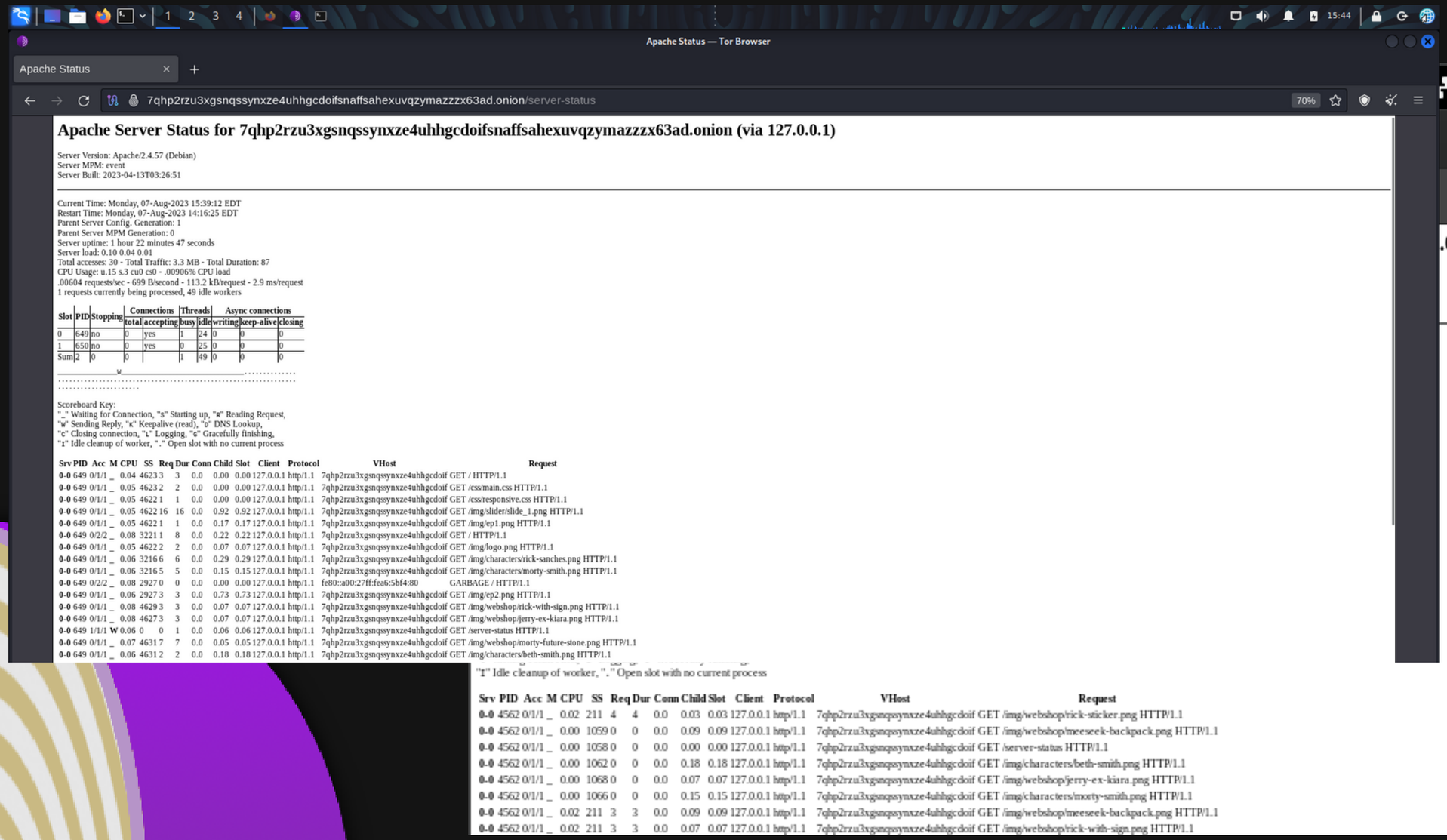
```
Activities  Terminal  Aug 7 12:37
epsilon@Epsilon: ~
epsilon@Epsilon: /var/www/tor  ×   epsilon@Epsilon: ~  ×
backups  cache  data  lib  local  lock  log  mail  opt  run  spool  tmp  www
epsilon@Epsilon:/var$ cd ~
epsilon@Epsilon:~$ sudo systemctl status apache2
[sudo] password for epsilon:
● apache2.service - The Apache HTTP Server
     Loaded: loaded (/lib/systemd/system/apache2.service; enabled; preset: enabled)
     Active: active (running) since Mon 2023-08-07 12:18:59 EDT; 17min ago
       Docs: https://httpd.apache.org/docs/2.4/
    Process: 4556 ExecStart=/usr/sbin/apachectl start (code=exited, status=0/SUCCESS)
   Main PID: 4561 (apache2)
      Tasks: 55 (limit: 4659)
     Memory: 22.3M
        CPU: 342ms
     CGroup: /system.slice/apache2.service
             ├─4561 /usr/sbin/apache2 -k start
             ├─4562 /usr/sbin/apache2 -k start
             └─4563 /usr/sbin/apache2 -k start

Aug 07 12:18:59 Epsilon systemd[1]: Starting apache2.service - The Apache HTTP Server.>
Aug 07 12:18:59 Epsilon apachectl[4560]: AH00112: Warning: DocumentRoot [/var/www/html>
Aug 07 12:18:59 Epsilon systemd[1]: Started apache2.service - The Apache HTTP Server.
lines 1-17/17 (END)
```

```
Activities  Terminal  Aug 7 12
epsilon@Ep
epsilon@Epsilon: /var/www/tor  ×
        CPU: 342ms
     CGroup: /system.slice/apache2.service
             ├─4561 /usr/sbin/apache2 -k start
             ├─4562 /usr/sbin/apache2 -k start
             └─4563 /usr/sbin/apache2 -k start

Aug 07 12:18:59 Epsilon systemd[1]: Starting apache2.service - The Apache HTTP Server.>
Aug 07 12:18:59 Epsilon apachectl[4560]: AH00112: Warning: DocumentRoot [/var/www/html>
Aug 07 12:18:59 Epsilon systemd[1]: Started apache2.service - The Apache HTTP Server.
lines 1-17/17 (END)
[3]+  Stopped                 sudo systemctl status apache2
epsilon@Epsilon:~$ sudo systemctl status tor
● tor.service - Anonymizing overlay network for TCP (multi-instance-master)
     Loaded: loaded (/lib/systemd/system/tor.service; enabled; preset: enabled)
     Active: active (exited) since Mon 2023-08-07 12:18:54 EDT; 18min ago
    Process: 4524 ExecStart=/bin/true (code=exited, status=0/SUCCESS)
   Main PID: 4524 (code=exited, status=0/SUCCESS)
        CPU: 3ms

Aug 07 12:18:54 Epsilon systemd[1]: Starting tor.service - Anonymizing overlay network>
Aug 07 12:18:54 Epsilon systemd[1]: Finished tor.service - Anonymizing overlay network>
lines 1-9/9 (END)
```

- The attacking system also has Tor installed and running.
- To be able to interact with the hidden service, Tor-Browser is also made use of.

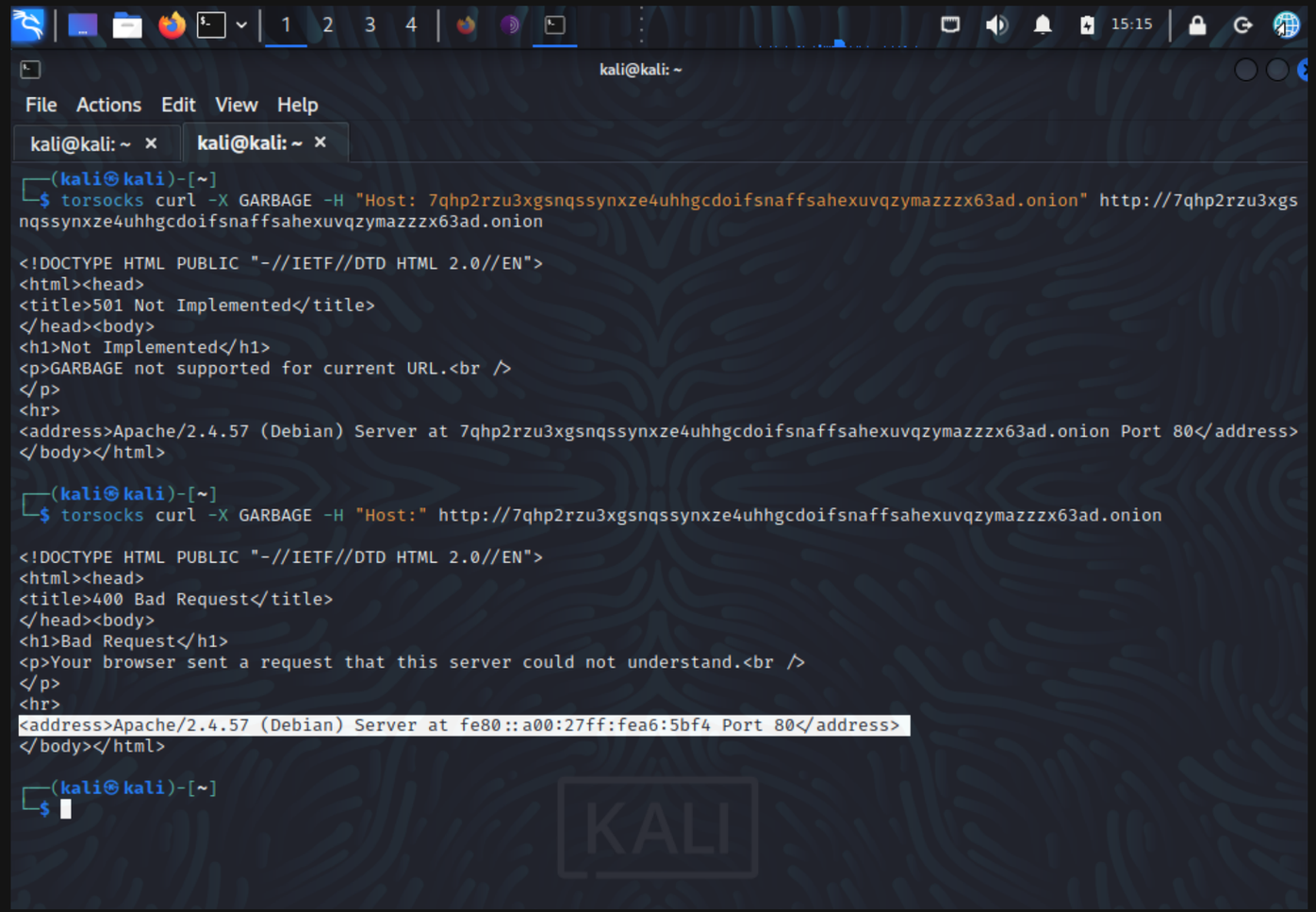


*.onion/server-status



BAD HTTP REQUEST

- It consists of sending an HTTP request to the hidden service with a blank host header
- The response returns the IP of the hidden service if it has a hosting address in the Apache website conf file.
- Here, it returns the IPv6 local address as it is being hosted on localhost



Conclusion

The implementation on attacks has always been a challenge

Sybil attack is versatile, but its very foundation of resistance against node failures is evident

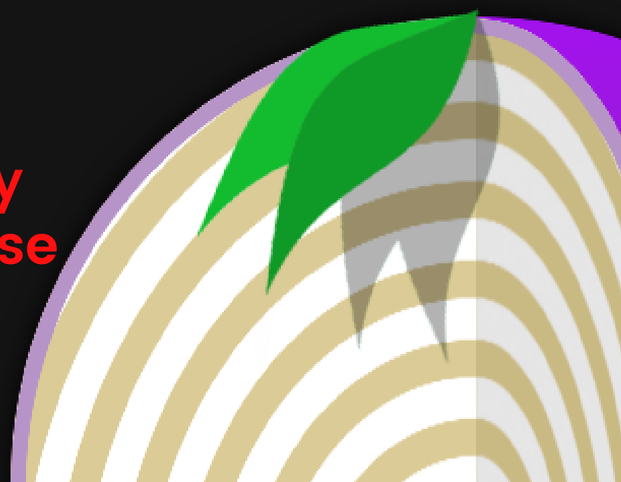
Sniper attack is very good to shoot down routers, but, it really does nothing after that

RAPTOR attacks are sophisticated, so are its detection strategies

Protocol Level attacks are relatively old, TOR now drops defected cells.

Replay attacks have the same disadvantages as Protocol-Level Attacks

Implementation of attacks in real time effectively requires control on multiple routers - say 50 in case of RAPTOR attacks, so services such as PLANET HUB are used



Do you have any questions?