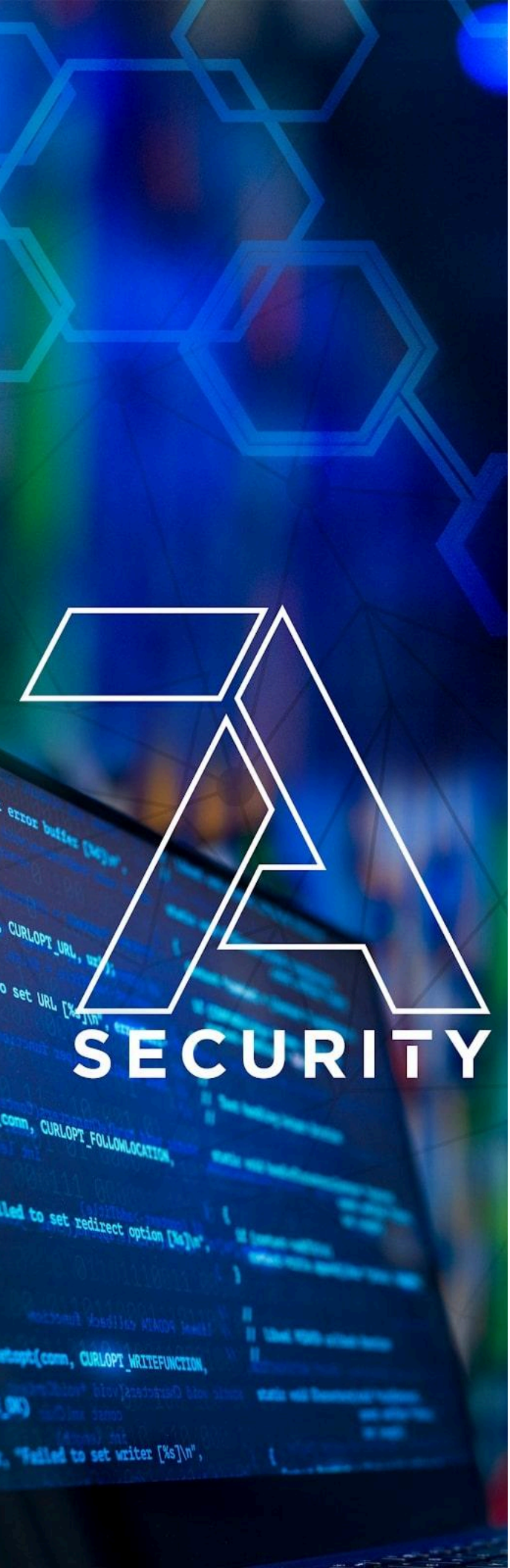

Tor Tool Test Targets:

TagTor
DescriptorParser
Margot
Exitmap
Bermuda

Privacy Report

Client:
Tor Project

7ASecurity Test Team:

- Abraham Aranguren, MSc.
- Daniel Ortiz, MSc.
- Miroslav Štampar, PhD.

7ASecurity
Protect Your Site & Apps
From Attackers
sales@7asecurity.com
7asecurity.com

INDEX

Introduction

"The Tor Project has been working on network monitoring tools to ensure that human rights defenders, journalists, activists, and other marginalized people have a safe, secure experience on the Tor network by reducing malicious relay activity and improving the health of the network. [...]
The role of each tool built in this Objective is to help humans on the network health team to better understand relay dynamics, be alerted to potential malicious activity, and track trends in the network. ***These tools are not designed to automatically remove relays or bridges, but to alert developers of potential issues in the network so they can be reviewed****."*

From *Tor Project Privacy Impact Assessment RfP*

This document outlines the results of a *whitebox* privacy impact review conducted against a number of Tor Monitoring Tools. The project was solicited by the Tor Project and executed by 7ASecurity in August 2024. The audit team dedicated 16 working days to complete this assignment. Please note that this is the first privacy impact assessment for this project. Consequently, the identification of privacy weaknesses was expected to be easier during this engagement, as more issues are identified and resolved after each testing cycle.

Please note this was strictly a privacy impact assessment; a security audit was out of scope. Any security weaknesses identified (TOR-01-Q05) were incidental to the privacy code review and not the focus of this exercise.

During this iteration the goal was to review the privacy as thoroughly as possible, to ensure the best possible anonymity. The methodology implemented was *whitebox*: 7ASecurity was provided with access to documentation, test users, and source code. A team of 3 senior auditors carried out all tasks required for this engagement, including preparation, delivery, documentation of findings and communication.

A number of necessary arrangements were in place by August 2024, to facilitate a straightforward commencement for 7ASecurity. In order to enable effective collaboration, information to coordinate the test was relayed through email, as well as a shared Signal Chat group. The Tor Project team was helpful and responsive throughout the audit, which ensured that 7ASecurity was provided with the necessary access and information at all times, thus avoiding unnecessary delays. 7ASecurity provided regular updates regarding the audit status and its interim findings during the engagement.

In this report, 7ASecurity directly answers 8 privacy-related questions with a confidence level ranging from *Unclear* to *Proven*.

Moving forward, the scope section elaborates on the items under review, followed by a testing methodology chapter, while a findings section documents the identified gaps with hardening recommendations.

Finally, the report culminates with a conclusion providing detailed commentary, analysis, and guidance relating to the context, preparation, and general impressions gained throughout this test, as well as a summary of the perceived privacy posture of the Tor Monitoring Tools in scope.

Scope

The following list outlines the items in scope for this project:

- **WP1: Privacy Impact Assessment of Tor Monitoring Tools**
  - https://gitlab.torproject.org/tpo/network-health/metrics/tagtor
  - https://gitlab.torproject.org/tpo/network-health/metrics/descriptorParser
  - https://gitlab.torproject.org/tpo/network-health/margot
  - https://gitlab.torproject.org/tpo/network-health/exitmap
  - https://gitlab.torproject.org/tpo/network-health/bermuda

Testing Methodology

This section outlines the testing methodology and coverage achieved during the privacy audit, focusing on various components of the Tor Monitoring Tools under scrutiny. Further details regarding the deep-dive assessments conducted on these tools, as well as the techniques employed to evaluate their privacy implications, are provided in the sections that follow.

The primary aim of the Test Methodology section is to elaborate on the assessment processes, offering context and transparency regarding all steps taken, potential privacy concerns evaluated, and the results of the analysis. Given the overlap in methodology and the similarities among the tools assessed, testing coverage is grouped into a unified section to avoid redundancy while clearly presenting the areas tested within the Tor Metric tools.

Tools Assessed

- Exitmap
- TagTor
- DescriptorParser
- Margot
- Bermuda

Summary of Findings:

- **Exitmap:** Analyzed for potential privacy leaks, focusing on how it reports and handles exit node information. No privacy issues were identified.
- **TagTor:** Assessed for data collection practices and potential leaks. No significant issues were found, but improvements in anonymization techniques (TOR-01-Q02) and security (TOR-01-Q05) were noted.
- **DescriptorParser:** Reviewed for accuracy in parsing Tor descriptors and handling sensitive data. The tool was found to be effective, with no privacy concerns.
- **Margot:** Evaluated for potential privacy weaknesses in handling Tor traffic and data anonymization. No privacy weaknesses were identified during this review.
- **Bermuda:** Investigated for privacy risks related to data storage and processing. No privacy risks were identified in the course of this audit.

Applied Methodology

The assessment began with a review of the existing Tor Monitoring Tool documentation, as well as an analysis of the prior Privacy Impact Assessment (PIA) report[1] to gain a thorough understanding of the tools within the Tor network.

- **Source Code Review**: A comprehensive examination of the source code for each tool was conducted. Static Analysis tools like *gitleaks*[2] and *semgrep*[3] were used, in conjunction with manual code review, to scan the source code for potential sensitive data exposure and privacy issues.
- **Dynamic Analysis**: Dynamic testing was performed to observe the behavior of each tool in real-time scenarios. For Exitmap, a detailed debugging process was carried out to detect any potential data leakage during execution, particularly with different modules and command-line parameters. TagTor underwent an in-depth review of its database schema and user authentication mechanisms to ensure no sensitive information was improperly exposed or stored.
- **Third-Party Dependencies**: Third-party dependencies and external services utilized by the tools, such as those listed in the *requirements.txt* file of the Exitmap tool, were scrutinized for privacy risks using tools like *Checkmarx*[4] and *Snyk*[5], as well as manually.
- **Module Interaction**: The interaction between various modules within the tools was assessed. Special attention was given to the handling of Exitmap configuration files and the note-taking feature of TagTor to identify any risks related to the exposure of personal or sensitive information.
- **Data Handling**: The handling and storage of aggregated data were evaluated. The focus was on ensuring that no privacy concerns arose from tool processing of public Tor metric data, particularly in DescriptorParser and Margot.
- **Network Interactions**: The audit reviewed network interactions, including DNS resolution processes in Exitmap and network administration commands in Margot, to verify that no sensitive information was being inadvertently exposed.
- **Tool-Specific Review**: The assessment concluded with a review of Bermuda, focusing on its role as a Tor exit scanner. The evaluation ensured that Bermuda operations do not introduce any privacy risks.

[1] https://research.torproject.org/techreports/privacy-in-memory-2017-04-28.pdf
[2] https://github.com/gitleaks/gitleaks
[3] https://github.com/semgrep/semgrep
[4] https://checkmarx.com/
[5] https://snyk.io/

Claim Verification

While not explicitly making claims about privacy protection, each tool is expected to operate without compromising privacy. The methodology aimed to verify this expectation by:

1. Confirming that no personally identifiable information is collected or stored.
2. Verifying that aggregated data cannot be easily de-anonymized, even for small user populations.
3. Ensuring that access to potentially sensitive data is properly restricted and audited.
4. Validating that the tools do not introduce new vectors for compromising anonymity.

From the perspective of 7ASecurity, based on the analysis of the tools and documentation provided, it can be concluded that the Tor Monitoring Tools generally adhere to strong privacy practices.

Privacy Analysis Findings

This section covers the privacy-related analysis results that attempt to answer 8 questions for *WP1: Privacy tests against Tor Monitoring Tools*. For this portion of the engagement, the *7ASecurity* team utilizes the following classification to specify the level of certainty regarding the documented findings. Given that this research occurred on the basis of reverse engineering, and source code analysis, it is necessary to classify the findings to address the level of confidence that can be assumed for each discovery:

- Proven: Source code and runtime activity clearly confirm the finding as fact
- Evident: Source code strongly suggests a privacy concern, but this could not be proven at runtime
- Assumed: Indications of a potential privacy concern were found but a broader context remains unknown.
- Unclear: Initial suspicion was not confirmed. No privacy concern can be assumed.

TOR-01-Q01: Information gathered or processed by the Tor Tools *(Unclear)*

This ticket summarizes the 7ASecurity attempts to answer the following question during the audit:

Q1: Does *any information, processed by the Tor Monitoring Tools in scope, have the potential to affect user privacy, and if so, how?*

7ASecurity found no evidence via analysis of the Tor Monitoring Tools, their documentation, source code review, and dynamic analysis that the Tor Monitoring Tools collect or process any information that may affect user privacy within the Tor network.

Most tools within the Tor Monitoring Tool ecosystem are used to collect and process publicly available information from the Tor network. For example, DescriptorParser[6] stores Tor network descriptors in PostgreSQL and VictoriaMetrics databases. Its architecture includes a suite of parsers, including *RouterFamilyBuilder*, *RouterStatusBuilder*, and *BandwidthParser*, among others. These parsers are designed to read descriptor files[7] from specified paths, process each descriptor, and leverage auxiliary methods to efficiently insert the processed data into the databases.

In terms of privacy, the parsers used are designed with a strong emphasis on safeguarding sensitive information. None of these parsers include methods for processing or manipulating sensitive data such as IP addresses or other personal

[6] https://gitlab.torproject.org/tpo/network-health/metrics/descriptorParser
[7] https://gitlab.torproject.org/tpo/[...]/descriptorparser/Main.java?ref_type=heads#L66

information that could raise privacy concerns. For instance, the *BridgedbMetricParser* class explicitly defines a SQL query, *INSERT_BRIDGESDB_METRIC_COUNT_SQL*[8], which is used to handle metric data without any involvement of privacy-sensitive information.

This design approach is consistent across the other tools (Exitmap[9], TagTor[10], Margot[11], Bermuda[12]), indicating a concerted effort to ensure that no privacy-related data is inadvertently processed or exposed.

7ASecurity could find no privacy concern in the current implementation, as it relates to data gathering, and hence this issue is merely informative and does not require any action. The analysis confirms that the Tor Monitoring Tools are designed and implemented with user privacy protection as a priority.

TOR-01-Q02: The Tor Tools could encourage better PII protection *(Assumed)*

This ticket summarizes the 7ASecurity attempts to answer the following question during the audit:

Q2: Is sensitive Personally Identifiable Information (PII) stored securely by the Tor Monitoring Tools in scope, and is there any risk of unauthorized retrieval?

7ASecurity did not find any evidence of the Tor Monitoring Tools storing PII. A minor exception was found in the TagTor tool, which might collect sensitive information if users manually insert it into user-generated notes:

Affected File:
https://gitlab.torproject.org/tpo/network-health/metrics/tagtor/…/tests/schema.sql#L82

Affected Code:

```
CREATE TABLE IF NOT EXISTS server_note(
  note_id                       BIGINT GENERATED ALWAYS AS IDENTITY,
  note                          TEXT                          NOT NULL,
  username                      TEXT                          NOT NULL,
  published                     TIMESTAMP WITHOUT TIME ZONE   NOT NULL,
  edited                        TIMESTAMP WITHOUT TIME ZONE   NOT NULL,
  fingerprint                   TEXT                          NOT NULL,
  status                        TEXT                          NOT NULL,
PRIMARY KEY(note_id, fingerprint)
```

[8] https://gitlab.torproject.org/tpo/[...]/parsers/BridgedbMetricsParser.java?ref_type=heads#L22
[9] https://gitlab.torproject.org/tpo/network-health/exitmap
[10] https://gitlab.torproject.org/tpo/network-health/metrics/tagtor
[11] https://gitlab.torproject.org/tpo/network-health/margot
[12] https://gitlab.torproject.org/tpo/network-health/bermuda

```
);
```

The Tor Monitoring Tool team could improve the privacy posture by educating Tor Tool users not to leak sensitive information via user-generated notes. Otherwise, there is no further action required by the Tor Monitoring Tool team to improve the privacy posture in this regard.

TOR-01-Q03: Excessive data is not gathered by the Tor Tools *(Unclear)*

This ticket summarizes the 7ASecurity attempts to answer the following question during the audit:

Q3: Do the Tor Monitoring Tools in scope collect any data beyond what is necessary for their proper functioning, and how does this impact user privacy?

7ASecurity did not find any evidence that the Tor Monitoring Tools collect additional data beyond what is necessary. Hence, no action is required by the Tor Monitoring Tool team to improve the privacy posture in this regard.

TOR-01-Q04: The Tor Tools do not appear to weaken Crypto *(Unclear)*

This ticket summarizes the 7ASecurity attempts to answer the following question during the audit:

Q4: Do the Tor Monitoring Tools in scope deliberately weaken cryptographic protocols to allow third-party decryption, and what are the privacy risks involved?

7ASecurity found no evidence to suggest that the Tor Monitoring Tools intentionally weaken cryptographic procedures to ensure third-party decryption. Based on this assessment, no action is required by the Tor Monitoring Tool team to improve the privacy posture in this regard.

TOR-01-Q05: The Tor Tools do not contain RCE Vulnerabilities *(Unclear)*

This ticket summarizes the 7ASecurity attempts to answer the following question during the audit:

Q5: Do the Tor Monitoring Tools in scope have vulnerabilities or insecure shell commands that could potentially lead to remote code execution (RCE), and what are the associated privacy risks?

7ASecurity did not identify any evidence, of intentional or unintentional vulnerabilities, that might lead to Remote Code Execution in the Tor Monitoring Tools during this audit. Furthermore, the complete lack of privacy issues identified during this audit highlights the overall privacy posture of the Tor Monitoring Tools in scope.

Please note that while no RCE issues were identified, during the code audit it was found that there is room for improvement in the current security processes, for example, the following code was found to be vulnerable to SQL injection.
Note: This was reported to the Tor Monitoring Tools team and promptly fixed during the audit.

Affected File:
https://gitlab.torproject.org/tpo/network-health/metrics/.../tagtor/db.py#L80-93

Affected Code:

```
def build_query(self, limit, offset, keyword, sort, filter,
                tags, untagged, sort_by):
    query = "SELECT DISTINCT ON (fingerprint, nickname) * \
        FROM server_status "
    [...]
    sort_order = "DESC" if sort not in ('asc', 'desc') else sort.upper()
    query += " ORDER BY fingerprint, nickname, {} {} LIMIT {} OFFSET {};".format(
        sort_by, sort_order, limit, offset)
    return query, params
```

TOR-01-Q06: The Tor Tools do not contain Backdoors *(Unclear)*

This ticket summarizes the 7ASecurity attempts to answer the following question during the audit:

Q6: Do the Tor Monitoring Tools in scope have any kind of backdoor?

7ASecurity did not identify any evidence of process, or command execution calls, commonly associated with backdoors or malware in the Tor Monitoring Tools during this audit. Based on this assessment, no action is required by the Tor Monitoring Tool team to improve the privacy posture in this regard.

TOR-01-Q07: The Tor Tools do not attempt to gain Root Privileges *(Unclear)*

This ticket summarizes the 7ASecurity attempts to answer the following question during the audit:

Q7: Do the Tor Monitoring Tools in scope attempt to gain root access through public vulnerabilities or in other ways?

At the time of writing, 7ASecurity found no evidence to suggest that any of the Tor Monitoring Tool components contain code that attempts to leverage or exploit platform-specific vulnerabilities to gain elevated privileges. Based on this assessment, no action is required by the Tor Monitoring Tool team to improve the privacy posture in this regard.

TOR-01-Q08: The Tor Tools do not use Obfuscation *(Unclear)*

This ticket summarizes the 7ASecurity attempts to answer the following question during the audit:

Q8: Do the Tor Monitoring Tools in scope use obfuscation techniques to hide code and if yes for which files and directories?

7ASecurity found no obfuscation evidence across the codebase. Furthermore, the Tor Monitoring Tools are operating at a high transparency level already, as the code is publicly available online, without any closed-source components (except the Bermuda repository). Hence, no action is required by the Tor Monitoring Tool team to improve the privacy posture in this regard.

Conclusion

The assessment of the Tor Monitoring Tools indicates that they have been designed with a strong focus on privacy. The tools demonstrated several positive aspects:

- The Tor Monitoring Tools primarily operate with public Tor network data, which minimizes risks to user privacy. No evidence was found of these tools collecting or processing sensitive information (TOR-01-Q01, TOR-01-Q02, TOR-01-Q03).
- It was noted that privacy preservation is strongly emphasized in the design and implementation of these tools, and no evidence suggests any deliberate weakening of cryptographic protocols to facilitate third-party decryption (TOR-01-Q04).
- No Remote Code Execution (RCE) vulnerabilities were identified, reinforcing the security posture of the Tor Monitoring Tools (TOR-01-Q05). Additionally, no backdoors or processes commonly associated with malware were found (TOR-01-Q06).
- It was observed that the tools do not attempt to gain root privileges (TOR-01-Q07), and no obfuscation techniques were identified (TOR-01-Q08), reflecting the Tor Project commitment to transparency. Most tools are open-source, promoting community oversight.
- Manual operation by the Tor network health team ensures that access is restricted to authorized personnel only.

The privacy posture of the Tor Monitoring Tools could improve in the following areas:

- The documentation could be improved to clarify the commitment to user privacy and avoid any potential data leaks (TOR-01-Q02).
- A formal code review process ought to be implemented, focusing on privacy implications for all changes. This process should include peer reviews, automated privacy checks, and regular audits to ensure privacy considerations are consistently addressed during development.

Regular testing of these Tor Tools is suggested, at least annually or before deploying substantial changes, to ensure new features do not introduce privacy weaknesses. This approach will reduce issues and increase resilience against online attacks. Additionally, future security audits, as illustrated in TOR-01-Q05, may benefit the tools.

7ASecurity would like to take this opportunity to sincerely thank Gaba, Georg, Micah Anderson, Silvia and the rest of the Tor Project team, for their exemplary assistance and support throughout this audit.