



Università degli Studi di Camerino

SCUOLA DI SCIENZE E TECNOLOGIE

Corso di Laurea in Informatica (Classe L-31)

**Servizi Onion
Dalla teoria all'implementazione**

Laureando
Leonardo Migliorelli

Matricola 113920

Relatore
Fausto Marcantoni

Correlatore
Correlatore

Indice

1	Introduzione	11
1.1	Obiettivi	11
1.2	Struttura della Tesi	11
2	Onion Routing	13
2.1	Onion come fix alle vulnerabilità della rete internet	14
2.2	Applicazioni di utilizzo	14
2.3	Onion come Mix Network	15
3	Tor § Onion v2	17
3.1	Obiettivi	18
3.2	Network Design	18
3.2.1	Controllo di Congestione	20
3.2.2	Tor Relay	21
3.3	Directory Servers	22
3.4	Tor Browser	23
4	Implementazione	25
4.1	Servizi Onion	25
4.2	Onion V3	26
4.3	Studio delle tecnologie	27
4.4	Creazione del servizio	28
4.4.1	Gestire la comunicazione tramite i socket unix	30
4.4.2	Creazione di un url personalizzato	32
4.4.3	Far conoscere il sito	34

Listings

4.1	Connessione SSH	28
4.2	Aggiornamento del sistema	28
4.3	Installazione Nginx	29
4.4	Avvio di Nginx	29
4.5	Tor Repository	29
4.6	Installazione GPG	29
4.7	Aggiunta chiavi gpg dal repository Tor	29
4.8	Installazione Tor	29
4.9	Riavvio di Tor	30
4.10	Aggiunta/creazione socket unix	30
4.11	Rimozione connessione dalla porta 80	30
4.12	Aggiunta socket unix a torrc	30

Elenco delle figure

2.1	Vita di un pacchetto onion	13
3.1	Tor logo	17
3.2	Statistiche dei Tor Relay in Italia	22
4.1	security Group 1, SSH in entrata	28
4.2	security Group 2, TCP in uscita	28
4.3	Connessione al web server nginx	31
4.4	Comando mfp224o e output	32
4.5	Connessione con il nuovo URL	33
4.6	OnionDir, Pagina iniziale	34
4.7	OnionDir, form aggiunta sito	35
4.8	Tor Browser mostra il pulsante <i>.onion available</i> quando si visita il sito tramite Tor	36

Elenco delle tabelle

1. Introduzione

Le moderne tecnologie di rete consentono una rapida comunicazione da ogni parte del mondo, avvicinando culture altrimenti distanti migliaia di chilometri. Due dei più grandi temi del nostro secolo sono la privacy e l'anonimato, parlando di reti internet ogni connessione tra client e server passa per una moltitudine di router che conoscono esattamente l'indirizzo (e quindi l'identità) del mittente e del destinatario. Con un po' di conoscenze non è complicato scoprire questi dati e sfruttarli a proprio vantaggio, le stesse big company spesso usano l'indirizzo IP con cui ci si connette al loro sito per tracciare l'utente e fornirgli articoli e pubblicità mirata o vendere i medesimi dati a terzi. La Rete Onion è stata creata per risolvere esattamente questo problema

1.1 Obiettivi

La tesi ha come principale obiettivo la dimostrazione di come è possibile generare un servizio/web server che sfrutta la rete tor per rendere le connessioni anonime e rendere anonimo lo stesso server, creare un hostname Tor personalizzato, far sì che il servizio sia raggiungibile attraverso un motore di ricerca onion e fornire un sistema di pagamento integrato per i prodotti del servizio

1.2 Struttura della Tesi

Nel primo capitolo viene introdotto l'onion routing, i possibili utilizzi e le mix networks. Il secondo capitolo introduce la rete Tor con le relative migliorie apportate all'onion routing e gli obiettivi prefissati per l'implementazione della rete. Il terzo capitolo introduce ai servizi onion, alla rete V3 che ha apportato alcune modifiche importanti per la sicurezza, precondizioni necessarie per l'implementazione che verrà mostrata dopo uno studio delle tecnologie disponibili

2. Onion Routing



Figura 2.1: Vita di un pacchetto onion

La **rete onion** è una rete distribuita composta dall’insieme di **onion router** che agiscono come nodi di rete e collaborano per portare un pacchetto dalla sorgente alla destinazione. Il tutto avviene senza che nessun nodo possa conoscere contemporaneamente l’host sorgente e l’host di destinazione, grazie alla criptografia a strati del pacchetto, per cui ogni strato viene criptato con una chiave differente e può essere decriptato solo dal nodo con la stessa chiave simmetrica, scoprendo così le informazioni sul prossimo nodo. Il cosiddetto exit node¹ può infine decriptare la parte finale del messaggio e scoprirne il corpo. La risposta del pacchetto segue lo stesso criterio sfruttando nodi, algoritmi e chiavi differenti.

Questo meccanismo è derivato dallo studio di David Chaum riguardo alle mix networks (Sezione 2.3). Questo concede all’utilizzatore di rimanere completamente anonimo, cosa che non può essere avvenire con la semplice criptografia SSL che agisce esclusivamente sul corpo del messaggio.

Come prima operazione per utilizzare la rete onion il client deve generare un nuovo circuito specificando il percorso di nodi che ogni pacchetto dovrà seguire, iniziando dal primo nodo vengono scambiate informazioni crittografiche come l’algoritmo e le chiavi da usare, successivamente si usano le informazioni del primo nodo per generare un pacchetto indirizzato al secondo nodo per ottenere le relative informazioni crittografiche e così via fino a che non si è generato tutto il circuito, grazie a questo meccanismo neanche durante la generazione del circuito è possibile risalire al client. Dato che la criptografia asimmetrica è troppo costosa per essere utilizzata durante lo scambio dati, essa viene usata solo durante la generazione del circuito e la condivisione delle chiavi simmetriche², successivamente gli strati vengono criptati e decriptati con la stessa chiave simmetrica. Questo consente alla rete onion ad avere una bassa latenza che è incrementata solo dal numero di onion router nel percorso e non dalla tecnologia che non è distante da quella usata in HTTPS [GD99]

¹Il nodo finale del circuito che conosce l’indirizzo reale della destinazione

²In un meccanismo simile a quello dell’handshake di HTTPS/TLS

2.1 Onion come fix alle vulnerabilità della rete internet

Nella rete internet ci sono due principali vulnerabilità che gravano sull'anonimato e sulla privacy:

- **Traffic analysis**, qualsiasi utente ha la capacità di analizzare il traffico, dato che i protocolli sono standardizzati e gli indirizzi non sono criptati è facile risalire alla sorgente e destinatario del pacchetto³.
- **Eavesdropping**, consiste nell'intercettazione dei pacchetti per leggerne il contenuto⁴

La rete Onion è nata proprio per risolvere queste problematiche, implementando le giuste tecnologie per proteggere gli utenti dall'analisi del traffico e dalle intercettazioni [MGRG]

2.2 Applicazioni di utilizzo

L'onion routing può essere usato con una moltitudine di protocolli e applicazioni, tra i più comuni troviamo HTTP(S), FTP, SSH, SMTP e DNS. L'utilizzo di molti dei protocolli più comuni avviene tramite gli onion proxies, i quali sono suddivisi in tre proxy layer logici

- Un proxy che genera e gestisce le connessioni, per operare ha necessità di conoscere la topologia e i percorsi verso altri nodi, informazioni che vengono distribuite in modo sicuro all'interno della rete a ogni nuovo nodo che si connette
- Un proxy chiamato “**Application Specific Proxy**”, per ogni connessione la relativa applicazione⁵ invia al proxy il pacchetto che normalmente invierebbe al server di destinazione, il proxy si occupa di convertire lo stream di dati in un formato accettato dalla rete onion
- Un proxy opzionale chiamato “**Application Specific Privacy Filter**” che significa lo stream di dati rimuovendo informazioni dal contenuto che potrebbero identificare la sorgente

I proxy possono essere configurati in molteplici modi, tra i quali vi è la possibilità di eseguire il proxy in un server remoto e sfruttare la rete Tor senza dover installare il software in ogni dispositivo, che quindi non ne deve gestire la computazione. [GD99]

³Si prenda il considerazione il caso di comunicazioni tra aziende i cui accordi sono confidenziali, una semplice analisi del traffico potrebbe rendere disponibile a chiunque l'esistenza di accordi

⁴Onion risolve anche questa problematica come effetto collaterale soluzione apportata all'analisi del traffico, infatti insieme all'indirizzo viene criptato anche il contenuto

⁵Un mail client come un Web Browser

2.3 Onion come Mix Network

La rete Tor è una delle molteplici reti basate sullo studio di David Chaum sulle Mix network, il suo studio si concentrò sulla rete basata sulla crittografia asimmetrica e sui sistemi di risposta senza conoscere l'identità del destinatario. Tutta la rete doveva funzionare senza necessità di un ente centrale a gestire le connessioni.

Abbiamo una chiave pubblica K e una chiave privata P, da cui derivano due funzioni:

- $K(x)$ la funzione che critta x con la chiave pubblica, può solo essere decriptato con la chiave P
- $P(x)$ la funzione che critta x con la chiave privata, può solo essere decriptato con la chiave K

Da questo è facile dedurre che $P(K(x)) = K(P(x)) = x$

Con questa tecnica però un malintenzionato potrebbe determinare il contenuto di un messaggio x creandone copie identiche y_n fino a che $K(y_n) = K(x)$ verificando l'uguaglianza $y_n = x$ e scoprendo il corpo del messaggio. Per risolvere questo problema si inserisce una stringa casuale R nella funzione crittografica del messaggio ottenendo quindi risultati sempre diversi tramite le funzioni $K(x, R)$ e $P(x, R)$, questa tecnica è chiamata sealing. Il sealing incrementa di molto la complessità, in quanto per determinare il contenuto del messaggio è inoltre necessario indovinare la stringa R

In una rete questo sistema viene implementato in maniera ridondante

$K1(R1, Kb(R0, M), B)$ ⁶

Quando il nodo 1 riceve il pacchetto lo decripta con la sua chiave privata scartando R1 e ottiene $Kb(R0, M), B$, inoltre il nuovo pacchetto⁷ a B. Questo processo può essere esteso ad N nodi (mix) incrementando la sicurezza della rete.

Il destinatario di un pacchetto deve avere la possibilità di rispondere sfruttando un indirizzo non tracciabile.

Il dispositivo A sfrutta il proprio indirizzo reale per generare un indirizzo non tracciabile, generando due chiavi pubbliche Ka e $R1$ ⁸, critta il proprio indirizzo A assieme alla chiave R1 come stringa casuale usando la chiave del mix

$K1(R1, A), Ka$

Quando B deve rispondere al pacchetto usa Ka per crittare il messaggio e $K1(R1, A)$ come indirizzo di destinazione, il mix M1⁹ che riceve il pacchetto usa la chiave privata P1 per decriptare l'indirizzo di destinazione (A) e usa la stringa casuale R1 come ulteriore chiave per crittare il messaggio

$K1(R1, A), Ka(R0, M) \Rightarrow A, R1(Ka(R0, M))$

Con questo sistema B può rispondere ad A, non conoscendo il vero indirizzo di A, il mix non conosce il contenuto del messaggio ma conosce l'indirizzo di destinazione e il dispositivo A è l'unico che può decriptare il contenuto del messaggio

Da considerare che tra M1 e B possono esserci N nodi e il messaggio può quindi essere

⁶il messaggio M viene crittato insieme a una stringa casuale R con la chiave pubblica del destinatario Kb, il tutto viene crittato assieme all'indirizzo B e a una stringa casuale R1 con la chiave pubblica K1 del nodo 1

⁷ $Kb(R0, M)$

⁸R1 in questo caso viene usato sia come stringa casuale ma in realtà esso è anche una chiave pubblica che verrà utilizzata dopo

⁹Il primo mix nel percorso da A a B, il più vicino al nodo A, viene inoltre considerato un nodo di uscita dalla rete dato che è l'unico che conosce il vero indirizzo di destinazione

criptato più volte nel percorso da B a M1

Nelle Mix Networks abbiamo un ente centrale che possiede una lista di pseudonimi creati dagli utenti. Quando un utente vuole creare il proprio pseudonimo invia la richiesta all'ente che decide se accettarla, solo allora lo aggiunge alla lista.

La richiesta contiene una chiave pubblica che agisce da pseudonimo, viene usata per verificare le firme generate dall'utente tramite la relativa chiave privata. Le richieste all'ente possono essere inviate in maniera anonima, tramite una mail non tracciabile o un altro sistema.

[Cha81]

3. Tor § Onion v2

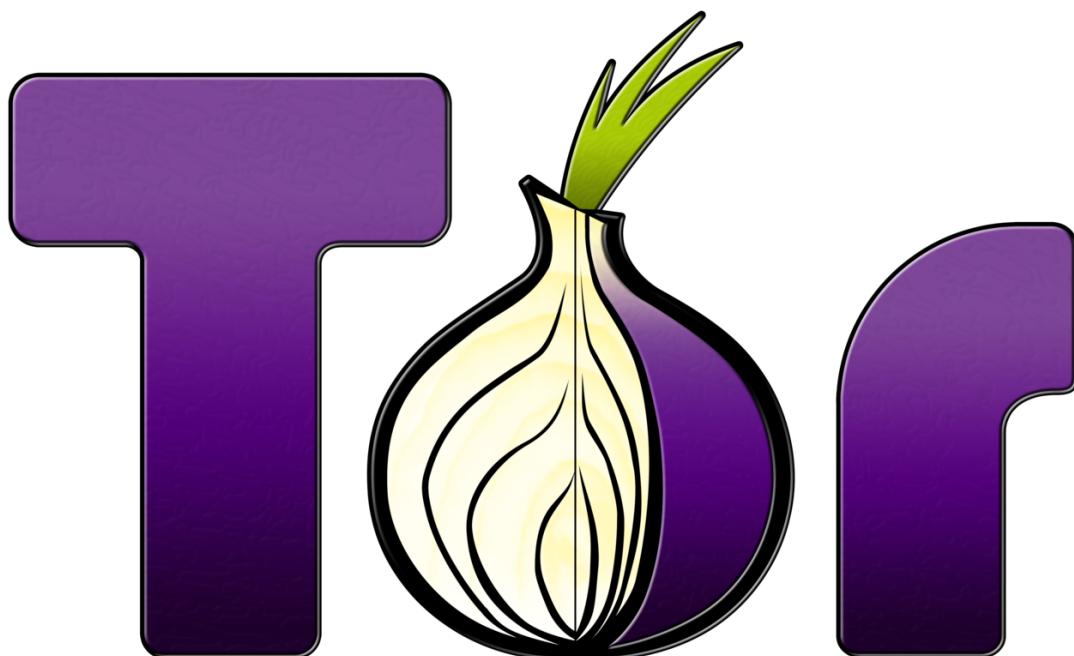


Figura 3.1: Tor logo

Nel 2002 viene presentata la rete Tor, diventata open source 2 anni dopo è l'implementazione più conosciuta di onion routing

Tra le varie migliore abbiamo

- La segretezza del canale, nella versione originale un nodo poteva forzare altri onion router a decriptare un traffico precedentemente registrato. La rete TOR invece sfrutta una tecnica di circuiti telescopici, le chiavi generate dal client in un determinato circuito sono di breve durata, non è quindi possibile utilizzarle successivamente per decriptare il vecchio traffico
- L'implementazione del proxy di applicazione attraverso lo standard SOCKS, consente alla maggior parte del traffico TCP di funzionare senza modifiche. Precedentemente era necessario implementare un proxy per ogni applicazione, e di conseguenza generare un circuito per ogni applicazione, con conseguente duplicazione di chiavi. Il proxy implementato in TOR invece genera il circuito a livello di TCP e più processi applicativi possono utilizzarlo. Per garantire la non tracciabilità di un utente nello stesso stream dati usato da più applicazioni è stato implementato il meccanismo dei *Rotating Circuits* che genera ogni minuto un nuovo circuito se quello precedente non sta essendo usato

- Controllo di congestione, un sistema decentralizzato che sfrutta ack end-to-end per garantire l'anonimato^{3.2.1}
- Directory Server, nodi più fidati di altri che descrivono le informazioni di rete in maniera sicura e affidabile
- Politiche di uscita variabili, ogni nodo possiede delle politiche che specificano le connessioni consentite e rifiutate, fondamentale in una rete distribuita fatta da volontari
- Controllo di integrità end-to-end, viene eseguito un controllo di integrità nel momento in cui il pacchetto esce dalla rete per garantire che i contenuti non sono stati alterati

Tor però non è completamente sicura, infatti non filtra informazioni di privacy nel corpo del messaggio come invece avviene in altri sistemi come Privoxy o Anonymizer e non offre garanzie in caso di attacco end-to-end che concerne sia sorgente che destinazione [RDos]

3.1 Obiettivi

L'obiettivo principale della rete TOR è la creazione di una rete che possa rendere gli attacchi molto più complessi da portare a termine scoraggiando così un possibile hacker, da questo principio cardine sono derivati gli altri obiettivi, tra cui

- Usabilità, la rete Tor a differenza degli altri sistemi che implementano le Mix-Networks^{2.3} predilige la bassa latenza¹ e l'usabilità, un aspetto fondamentale se si vuole garantire l'anonimato². Da questo derivano altri obiettivi come
 - Basse latenze, determinate anche dal fatto che più applicazioni possono usare lo stesso circuito TCP senza doverne generare uno nuovo per ogni stream dati, questo consente di ridurre il delay causato dalla criptografia asimmetrica
 - Essere implementabile con meno configurazioni possibili, per incrementare il numero di servizi che possano attrarre utenti
 - Essere multi piattaforma, per incrementare la base di utenti
- Semplicità, la rete deve essere facile da comprendere, doveva essere usabile nel mondo reale e non doveva essere troppo costosa

[Cha81]

3.2 Network Design

A differenza di altri sistemi che usano una rete peer-to-peer, Tor è una *overlay network*³, questa scelta è derivata dalle possibili vulnerabilità di una rete basata sugli utenti, infatti un attaccante potrebbe compromettere il traffico leggendo o manipolando i dati.

¹Il che rende la rete adatta all'utilizzo tramite un web browser

²Maggiori sono i nodi più semplicemente si può garantire l'anonimato

³Una rete virtuale creata sfruttando una rete fisica preesistente

Ogni onion router è rappresentato come un processo software che mantiene due tipi di chiavi

- Chiave d'identità, una chiave di lunga durata usata per firmare i pacchetti garantendo agli altri nodi della rete l'autenticità del messaggio e delle informazioni contenute, tra cui indirizzo, bandwidth, exit policy, ecc..
- Chiave onion, una chiave di breve durata usata per decriptare le richieste all'interno dei circuiti utente

Un elemento chiave della rete TOR sono le celle, pacchetti di dimensione fissa a 512 bytes, come ogni tipo di pacchetto sono divisi in header e payload, l'header contiene l'identificativo del circuito e un comando che indica come gestire il payload. Il comando può essere

- **Padding** per mantenere viva la connessione
- **Create** per creare un nuovo circuito
- **Destroy** per eliminare un circuito

Inoltre il tipo di comando definisce il tipo della cella

- **Control**, pacchetti di controllo gestiti dal primo router che li riceve, per questo non vengono mai inoltrati
- **Relay**, trasporta stream dati per cui ha necessità di un header aggiuntivo contenente l'identificativo streamID, il checksum per il controllo di integrità, la dimensione del payload e un comando di relay. Il comando di relay può essere
 - **Begin** per iniziare uno stream
 - **End** per terminare uno stream
 - **Teardown** per terminare uno stream in modo forzato, usato per quelli "rotti"
 - **Connected** per rispondere al relay begin dell'OP, informandolo che lo stream è stato creato con successo
 - **Extend** per estendere un circuito di un router
 - **Truncate** per eseguire un teardown di una parte del circuito
 - **Sendme** usato per il controllo di congestione
 - **Drop**

L'utente per generare il circuito segue un processo di negoziazione incrementale:

1. Come primo passo l'utente, o meglio l'Onion Proxy crea e invia una richiesta *relay* al primo nodo nel percorso scelto
2. Avviene la condivisione della chiave simmetrica tramite l'handshake di Diffie-Hellman, la creazione dell'ID del circuito e di conseguenza la creazione della connessione con il primo nodo (R1)

3. L'utente invia una richiesta relay extend indicando a R1 l'indirizzo del secondo nodo nel percorso scelto (R2), R1 inizia una connessione con R2 definendo un nuovo id e associando la connessione OP-R1 con la connessione R1-R2, OP e R2 non si conoscono a vicenda e comunicano solo tramite l'intermediario R1. In fine R1 invia all'utente le informazioni del nuovo nodo tra cui la chiave simmetrica per il relativo strato
4. Questa operazione viene effettuata, richiedendo all'ultimo router corrente di creare una connessione con il suo successivo, finché il circuito non viene completato[Cha81]

Ogni applicazione, in base alle proprie necessità, invia le richieste di stream TCP all'Onion Proxy, che sceglie il circuito più recente (oppure genera un nuovo circuito) e sceglie un exit node, solitamente l'ultimo router. A questo punto l'OP invia una cella *relay begin* con un identificativo casuale all'exit node, la risposta dell'exit node conferma l'esistenza del nuovo stream TCP e l'OP può accettare i dati delle applicazioni TCP ed inoltrarli nella rete Onion

3.2.1 Controllo di Congestione

Questo design ha qualche problema derivato dal fatto che più OP potrebbero scegliere lo stesso percorso OR-OR, generando un bottleneck e saturando la rete. Le normali tecnologie di controllo di congestione non possono funzionare in una rete del genere, i pacchetti devono rispettare il proprio percorso e non possono cambiarlo, questo ne renderebbe impossibile la lettura. Per implementare un sistema di controllo di congestione sono stati sviluppati 2 meccanismi

- **Controllo di Circuito**, ogni OP possiede le informazioni di due finestre per ogni OR, mentre ogni OR possiede le informazioni delle finestre dell'OP per ogni circuito di cui fanno parte. Le 2 finestre iniziano da un valore di 1000 e decrementano a ogni cella, esse sono
 - **Packaging window**, tiene traccia del numero di celle che un OR può inviare verso l'OP, quando un router riceve abbastanza celle dati (100) invia un relay sendme all relativo OP con *streamID* = 0, quando un OR riceve un relay sendme incrementa la finestra di packaging (di 100) per il relativo OP, se la finestra di packaging raggiunge 0 il nodo che sia un Onion Router o un'Onion Proxy smette di ricevere ed inviare celle dal circuito attendendo il sendme
 - **Delivery window**, tiene traccia del numero di celle che un OR è in grado di inviare fuori dalla rete
- **Controllo di Stream**, simile al controllo di circuito ma il meccanismo è applicato ad ogni stream TCP separatamente. Ogni stream inizia con packaging = 500 con incrementi di 50 ad ogni relay sendme. Il relay sendme, a differenza del controllo a livello di circuito, viene inviato quando il numero di bytes che devono essere inviati è inferiore al threshold di 10 celle

[RDos]

3.2.2 Tor Relay

La rete Tor si basa su nodi gestiti da volontari che forniscono la propria potenza di calcolo per consentire alla rete tor di funzionare, l'aumento di nodi porta la rete a essere più veloce, stabile e sicura per gli utenti. Questi nodi sono anche chiamati *Tor Relay*, ognuno possiede una exit policy che ne stabilisce il tipo e le caratteristiche, si suddividono quindi in

- Non-exit Relay, sono i nodi interni alla rete, vengono definiti in base alla exit-policy, essi si dividono in
 - **Guard Relay**, sono i primi nodi di ogni circuito, devono rispettare specifiche di velocità e stabilità non inferiori a 2MByte/s
 - **Middle Relay**, agiscono come tramite tra i nodi *Guard* ed *Exit*, possono diventare Guard solo se rispettano le relative specifiche
- **Exit Relay**, sono gli ultimi nodi di ogni circuito, ricevono traffico dalla rete Tor e lo inoltrano in una rete normale. Questo porta coloro che fanno hosting di questi relay a essere i più esposti a rischi legali⁴, per questo e altri motivi è sconsigliato eseguire un Tor Relay dalla propria abitazione e viene piuttosto suggerito di affidarsi ad altre istituzioni
- **Bridge Relay**, sono nodi non pubblici⁵ i cui IP non possono quindi essere bloccati da governi o ISP in quanto privati. Inoltre generalmente non ricevono denunce legali di abuso della rete internet come invece accade per gli Exit Relay.

[Tord] [Torf] Dal sito Tor [Tore] è possibile vedere le statistiche riguardanti i relay, la loro bandwidth, il tipo, gli indirizzi IP e il tempo di esecuzione. Dallo stesso sito è inoltre possibile filtrare per paese, vedendo il numero e caratteristiche dei Relay in Italia

⁴Se un utente usa la rete onion per scaricare contenuti con copyright, l'IP da cui è stato scaricato il file sarà quello del Relay e quindi le colpe ricadranno sul proprietario del server

⁵Non vengono inseriti nelle directory pubbliche Tor

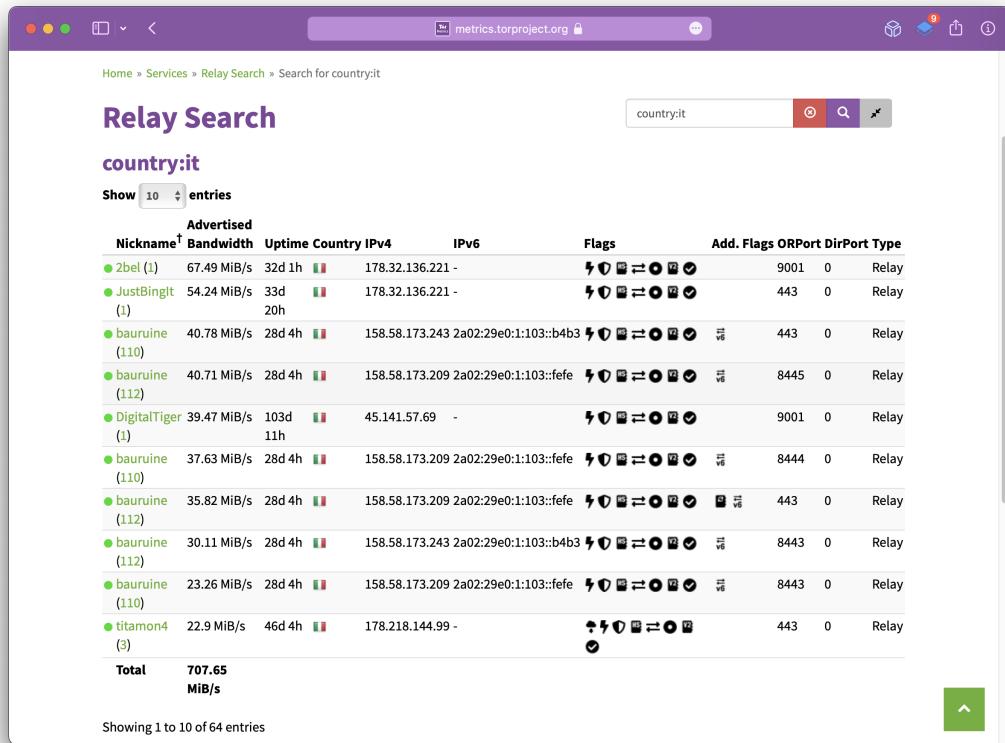


Figura 3.2: Statistiche dei Tor Relay in Italia

3.3 Directory Servers

I directory server sono un piccolo sottogruppo di onion router utilizzati per tracciare i cambiamenti nella topologia di rete. In particolare un directory server agisce come un server HTTP accessibile dai client per ottenere lo stato della rete e la lista dei router aggiornata periodicamente dagli stessi OR.

Quando il directory server riceve un aggiornamento da un OR prima di tutto controlla la chiave d'identità del router, garantendo che un attaccante non possa fingersi un OR manomettendo la rete [Cha81]. Il software di accesso alla rete onion è precaricato con le informazioni sui directory server e le relative chiavi, le informazioni di rete vengono aggiornate periodicamente dall'OP.

I directory server sono anche fondamentali per la connessione agli onion services, infatti ogni servizio creato genera un *Onion Service Descriptor* contenente una lista degli introduction points e la chiave pubblica, il pacchetto viene criptato con la chiave privata e inviato al directory server che quindi agisce come fosse un DNS server per gli onion services. Quando l'utente tenta la connessione a un sito onion richiede e riceve dal directory server il relativo *Descriptor* per l'indirizzo onion, usa la chiave pubblica, derivata dalla stringa dell'indirizzo onion a cui vuole connettersi, per decriptare il pacchetto

Nel caso il Directory Server fosse compromesso e contenesse un Descriptor malevolo generato per reindirizzare il traffico⁶, l'indirizzo onion che possediamo non sarebbe in

⁶Che però non potrà avere la stessa chiave privata

grado⁷ di decriptare le informazioni, dato che sono state criptate con una differente chiave privata [Torb]. Non c'è neanche la necessità di determinare se le informazioni sono corrette, perché a patto che la chiave privata non sia resa pubblica, nessuno potrà manomettere il Descriptor senza renderlo illeggibile e/o "*indecifrabile*"

3.4 Tor Browser

Una delle principali vulnerabilità della rete Tor è la possibilità che un servizio web crei un codice JavaScript⁸ malevolo in grado di ottenere informazioni sull'indirizzo dell'utente deanonymizzandolo. Gli sviluppatori di Tor Browser hanno inserito appositamente un sistema di sicurezza che blocca gli script, questo però di contro porta molti siti a non funzionare correttamente.

Essendo la rete Tor fortemente basata sulla criptografia la navigazione di un sito onion tramite Tor non fa comparire alcun avviso di sicurezza in caso di mancanza di HTTPS, dato che il traffico è già criptato. Da tenere inoltre in considerazione che un certificato proveniente da una Certificate Authority potrebbe generare problemi di anonimizzazione del servizio a causa della Certificate Transparency[Goo]

⁷tramite la chiave pubblica al suo interno nascosta

⁸che viene eseguito sul browser dell'utente

4. Implementazione

4.1 Servizi Onion

Nelle comuni reti internet i servizi vengono utilizzati facendo richieste basate sugli indirizzi IP pubblici, la sorgente e destinazione di un pacchetto è quindi conosciuta da tutti i dispositivi che la ricevono. Onion invece garantisce l'anonimato non solo agli utenti ma anche ai servizi, nascondendone l'indirizzo IP. Questo viene fatto tramite uno pseudonimo di lungo termine, identico in tutti i circuiti e stabile anche al un fallimento di un router

I principali obiettivi sono

- Gli attaccanti non devono riuscire a manipolare la rete sostituendosi un servizio esistente. Questo livello di affidabilità deriva dalla criptografia asimmetrica che ci garantisce che il servizio a cui cerchiamo di connetterci sia autentico, solo lui possiede la copia privata della chiave pubblica con cui tentiamo la connessione
- Sicurezza dagli attacchi DoS, viene fatto tramite l'uso di più punti d'ingresso alla rete

La creazione di un servizio onion passa per diversi punti prima di poter essere raggiungibile

1. Genera una coppia di chiave pubblica e privata per identificarsi
2. Definisce alcuni onion router come punti di ingresso nella rete, da cui riceverà le richieste dei clienti e invia loro la chiave pubblica
3. Crea un circuito con ogni punto di ingresso
4. Invia al servizio di lookup onion le informazioni sui punti d'ingresso e l'hash della chiave pubblica che sarà usata come hostname^{3.3}

Quando un utente tenta la connessione

1. Inizialmente esegue il lookup del dominio tramite un directory server
2. Successivamente sceglie un OR come tramite per il servizio e genera un circuito verso di esso, sfruttando uno specifico cookie per identificare il servizio
3. Viene generato uno stream verso uno dei punti d'ingresso del servizio con tutte le informazioni riguardo se stesso, il nodo tramite e il cookie. Tutto viene criptato con la chiave pubblica del servizio
4. Inizia l'handshake di Diffie-Hellman tra OP e servizio

5. Il servizio onion a sua volta genera un circuito verso il nodo tramite per poter rispondere all'OP con la seconda parte dell'handshake
6. Il nodo tramite collega i due circuiti generando un circuito dati bidirezionale

Sia il programma dell'utente che il web server non vengono modificati e non sono nemmeno a corrente che il traffico viaggia per una rete onion [Cha81]

4.2 Onion V3

La terza generazione di onion nasce per risolvere alcuni problemi di sicurezza. In particolare rispetto alla seconda generazione

- Sono stati aggiornati i sistemi di crittografia, da SHA1/DH/RSA1024 a SHA3/ed25519/curve25519
- Sono stati migliorati i directory server e il directory protocol
- È stato cambiato il sistema di hostname. Precedentemente venivano usati i primi 80 bit dell'hash (SHA1) della chiave pubblica per creare l'hostname¹, nella terza generazione vengono codificati in base32
 - La chiave pubblica, un totale di 32 byte in ed25519
 - Il checksum di 2 byte
 - Un byte di versione, di default '\x03'

In totale il nuovo hostname possiede 56 caratteri²

Inoltre l'indirizzo v3 contenendo l'intera chiave pubblica, una volta decodificato² in base32³ ci ritorna tutte le informazioni necessarie per connettersi al servizio.

[Tor13]

¹un esempio **yyhws9optuwiwsns.onion**

²esempio l'indirizzo *pg6mmjyjmcrsslykfunntlaru7p5svn6y2ymmjubnubxndf4pscryd.onion*, se proviamo a decodificarlo otteniamo la stringa 79bcc625184b05194975c28b66b66b0469f7f6556fb1ac3189a79b40dda32f1f214703, come notiamo termina con 03, il byte di versione

³Definito nell'RFC 3548 e 4648

4.3 Studio delle tecnologie

Per creare un servizio sulla rete onion si possono usare una moltitudine di tecnologie differenti, ogni singolo aspetto ha necessità di effettuare delle scelte.

- Deploy del Server, la prima scelta ricade sulla creazione del server e in particolare scegliere dove eseguire il deploy. Potremmo fare il deploy in locale, in una qualsiasi macchina intel/amd ma ci sono molteplici motivi per usare un servizio cloud, tra cui la sicurezza, la facilità d'implementazione⁴. Ci sono 3 principali competitor e molti altri secondari che per la maggior parte sfruttano i server delle 3 aziende principali:
 - Microsoft Azure, il servizio cloud di Microsoft, è il secondo più grande servizio cloud.
 - Google Cloud, il servizio cloud di Google, è il terzo più grande servizio cloud.
 - Amazon Web Service (AWS), il più importante servizio cloud, è gestito e reso disponibile da Amazon e possiede molti servizi tra cui scegliere. Per questa implementazione useremo il servizio EC2 messo a disposizione da AWS⁵ in una macchina t3.micro
- Sistema Operativo, la seconda scelta ricade sul SO della nostra macchina. I principali *"concorrenti"* sono Windows Server e Linux, anche se la scelta migliore ricade su Linux
 - Leggerezza, affidabilità e sicurezza derivate da un sistema GNU/Linux
 - Maggiore controllo sul sistema operativo e personalizzazione di ogni aspetto, è anche possibile creare un sistema operativo Linux su misura per le proprie esigenze
 - Maggiore utilizzo in ambienti server
 - Nessun costo aggiuntivo derivato dall'acquisto di una licenza d'uso come Windows server

Useremo Debian nella versione 11

- Web Server, abbiamo due web server principali
 - Apache httpd, il più vecchio dei due, rilasciato nel 1995 con licenza Apache 2.0
 - Nginx, un web server più moderno che fornisce diversi altri strumenti oltre al web server, tra cui il proxy e load balancer. È stato rilasciato nel 2004, è gratuito e come httpd è open source, ha ufficialmente superato Apache nel luglio 2021[Web]. Viene anche utilizzato in container Docker, ma noi lo useremo in una classica macchina virtuale.

⁴Per esempio il login da remoto in una macchina locale richiederebbe un IP statico e comunque non potremmo accendere e spegnere la macchina senza implementare sistemi esterni o nel migliore dei casi un Wake on Lan

⁵Gli altri servizi cloud messi a disposizione da Google o Microsoft forniscono strumenti simili, anche se la configurazione può essere differente

4.4 Creazione del servizio

Una volta creato l'account AWS ed eseguito l'accesso possiamo creare una macchina virtuale andando nella sezione EC2 -> Instances -> Launch Instances[Aws], possiamo creare la macchina in qualsiasi regione, in questo caso la andremo a creare a Milano *eu-south-1*. Durante la configurazione possiamo selezionare le opzioni che abbiamo precedentemente analizzato

1. Il nome
2. Il sistema operativo da usare, selezioniamo Debian 11 e in particolare l'architettura x86,
3. L'istanza scelta è una semplice t3.micro con 2 vCPU e 1 GiB di memoria, questo a dimostrazione che per eseguire un servizio onion non è necessario un server potente
4. Una coppia di due chiavi pubblica e privata per la connessione SSH, è anche possibile riusare una coppia precedentemente creata e quindi usare le stesse chiavi per più server
5. Successivamente è fondamentale aggiungere alla macchina i giusti security group, di default AWS non consente nessun tipo di traffico in entrata o in uscita dalla macchina
 - Consentire la connessione SSH alla porta 22

Type	Protocol	Port range	Source
SSH	TCP	22	0.0.0.0/0

Figura 4.1: security Group 1, SSH in entrata

- Consentire il traffico TCP in uscita

IP version	Type	Protocol	Port range	Destination
IPv4	All TCP	TCP	0 - 65535	0.0.0.0/0

Figura 4.2: security Group 2, TCP in uscita

Una volta creata la macchina avremo la possibilità di scaricare la chiave privata con formato *.pem che useremo per connetterci tramite SSH alla shell della macchina

```
ssh -i key.pem admin@*.compute.amazonaws.com
```

Listing 4.1: Connessione SSH

La primissima operazione sarà l'aggiornamento dei repository e del sistema

```
sudo apt update && sudo apt full-upgrade -y
```

Listing 4.2: Aggiornamento del sistema

Poi installiamo il web server nginx che abbiamo analizzato prima

```
sudo apt install nginx
```

Listing 4.3: Installazione Nginx

Una volta completata l'installazione il web server viene automaticamente avviato e messo in ascolto sulla porta 80, in caso non fosse attivo

```
sudo systemctl start nginx
```

Listing 4.4: Avvio di Nginx

Per installare tor è innanzitutto necessario installare **apt-transport-https** per utilizzare i repository tramite HTTPS

```
sudo apt install apt-transport-https
```

Dal comando `lsb_release -c` possiamo vedere la distribuzione corrente e inserire il repository corretto nella configurazione, con Debian 11 siamo su di una distribuzione bullseye, creiamo un file chiamato `tor.list`⁶ nella cartella `/etc/apt/sources.list.d` dove andremmo ad aggiungere i repository TOR tramite le seguenti righe

```
deb [signed-by=/usr/share/keyrings/tor-archive-keyring.gpg] https://deb.torproject.org/torproject.org bullseye main
deb-src [signed-by=/usr/share/keyrings/tor-archive-keyring.gpg] https://deb.torproject.org/torproject.org bullseye main
```

Listing 4.5: Tor Repository

Assicuriamoci che GPG sia installato nel sistema prima di aggiungere le chiavi

```
sudo apt install gpg
```

Listing 4.6: Installazione GPG

Aggiungiamo le chiavi GPG della repository appena creata

```
wget -qO- https://deb.torproject.org/torproject.org/
A3C4F0F979CAA22CDBA8F512EE8CBC9E886DDD89.asc | gpg --dearmor | tee /usr/share/keyrings/tor-archive-keyring.gpg >/dev/null
```

Listing 4.7: Aggiunta chiavi gpg dal repository Tor

Aggiorniamo gli index e installiamo finalmente tor

```
sudo apt update && sudo apt install tor -y
```

Listing 4.8: Installazione Tor

[**Tora**]

Generiamo il file `torrc` che ci servirà per impostare tutti i parametri di configurazione

```
sudo tor -f /etc/tor/torrc
```

Apriamo il file `/etc/tor/torrc` e togliamo il commento alle righe `HiddenServicePort 80 127.0.0.1:80` e `HiddenServiceDir /var/lib/tor/hidden_service`, il primo indica che il traffico arrivato dalla porta virtuale (80) viene reindirizzato alla porta 80 del localhost, ovvero la porta in cui è in ascolto il web server nginx, il secondo ci serve per indicare

⁶Il nome è configurabile

la directory in cui salvare l'hostname del sito e le chiavi crittografiche, o eventualmente la directory in cui sono già presenti.

Riavviamo tor per aggiornare la configurazione, assicurarcì che il file torrc non contenga errori e che il sistema sia funzionante

```
sudo systemctl restart tor
```

Listing 4.9: Riavvio di Tor

4.4.1 Gestire la comunicazione tramite i socket unix

A questo punto tor ha creato gli introduction points e ha generato un circuito con ognuno di essi, ha inoltre generato le chiavi e il proprio hostname, queste informazioni sono state aggiunte nella cartella che abbiamo inserito nell'HiddenServicePort, in particolare il file chiamato hostname contiene l'indirizzo tor del nostro servizio [Torc]

Questo sistema però non è completamente sicuro, infatti Tor è connesso con il web server tramite la porta 80, essendo entrambi i processi sulla stessa macchina il sistema ottimale è utilizzare un socket unix tra i due processi. Questo impedisce a un utente **non** Tor di accedere direttamente al web server senza dover configurare un firewall che comunque aggiunge complessità nella rete.

Apriamo il file /etc/nginx/sites-enabled/default e nella sezione server aggiungiamo il socket in ascolto, così che la comunicazione possa passare anche per il socket unix

```
listen unix:/var/run/website.sock;
```

Listing 4.10: Aggiunta/creazione socket unix

Possiamo inoltre rimuovere la connessione dalla porta 80 commentando le seguenti righe

```
#listen 80 default_server;  
#listen [::]:80 default_server;
```

Listing 4.11: Rimozione connessione dalla porta 80

Dopo aver riavviato nginx ed esserci assicurati che non vi siano errori apriamo il file torrc e inseriamo lo stesso socket

```
HiddenServicePort 80 unix:/var/run/website.sock
```

Listing 4.12: Aggiunta socket unix a torrc

Infine riavviamo tor

```
sudo systemctl restart tor
```

Se tutto è andato a buon fine il web server non dovrebbe essere in grado di rispondere alle richieste HTTP dalla porta 80, ma solo a quelle provenienti dalla rete Tor tramite l'hostname .onion che abbiamo ricevuto precedentemente.

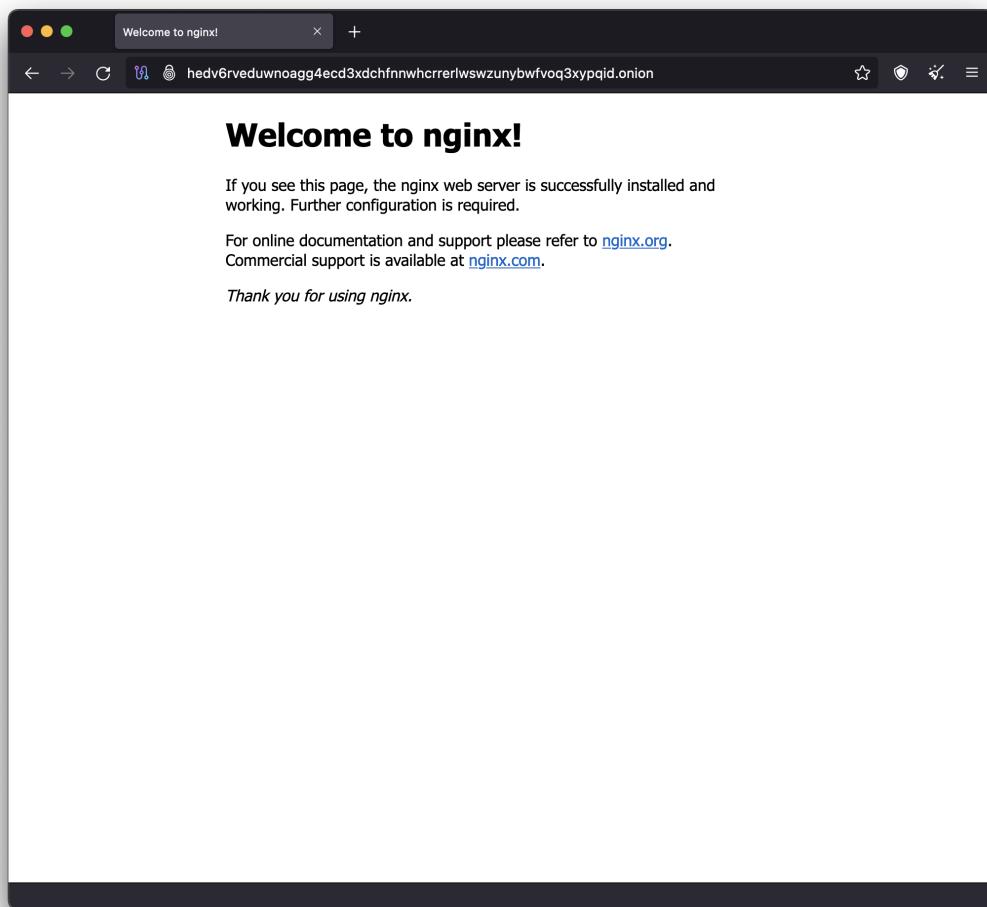


Figura 4.3: Connessione al web server nginx

4.4.2 Creazione di un url personalizzato

Ora che il servizio è attivo e disponibile possiamo iniziare a configurarlo a nostro piacimento, la prima cosa che potremmo voler fare è cambiare l'hostname impostandone uno personalizzato, almeno in parte, un esempio è il sito email di proton **proton-mailrmez3lotccipshtklegetolb73fuirgj7r4o4vfu7ozyd.onion**

Si sfrutta la tecnica del brute force per generare coppie di chiavi casuali da cui ottenere indirizzi .onion fino a che non ne troviamo uno che inizia con la stringa che abbiamo scelto, potrebbe impiegare diverso tempo ed è consigliabile usare un computer più potente della nostra macchina virtuale. Vi sono diversi strumenti per fare quest'operazione sia v2 che v3, per questa implementazione useremo il tool mfp224o [cat] che appunto ci permette di fare quest'operazione con gli indirizzi v3, possiamo installarlo su Linux o usarlo su Windows da riga di comando semplicemente indicando il nome con cui dovrà iniziare

```
./mfp224o tesilm
```

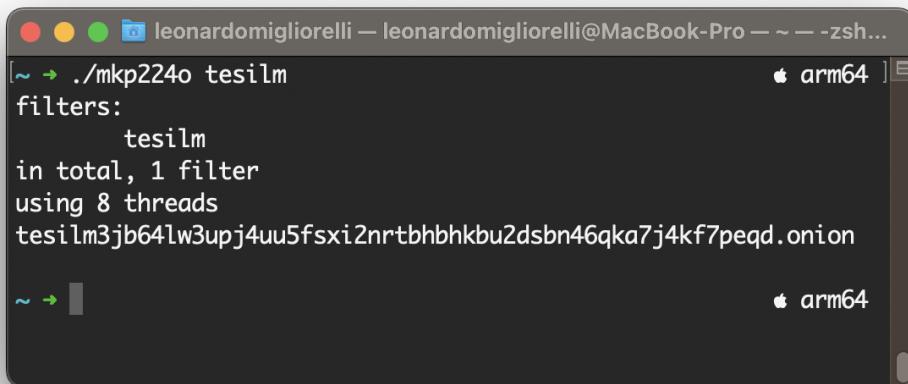
7

Da notare che più caratteri si scelgono nel matching più il software impiegherà nella generazione di un dominio. Otteniamo il nuovo indirizzo del servizio⁸ con le relative chiavi

Possiamo copiarle sul server con SCP usando l'opzione -i per definire il file d'identità, come facciamo con SSH

```
cd tesilm3jb64lw3upj4uu5fsxi2nrtbhbhkbu2dsbn46qka7j4kf7peqd.onion  
scp -i key.pem * admin@*.compute.amazonaws.com:~/var/lib/tor/  
hidden_service
```

Avviamo la connessione al server e riavviamo tor per impostare le modifiche



The screenshot shows a terminal window titled "leonardomigliorelli — leonardomigliorelli@MacBook-Pro — ~ — zsh...". The command "./mfp224o tesilm" is run, followed by the output:

```
filters:  
tesilm  
in total, 1 filter  
using 8 threads  
tesilm3jb64lw3upj4uu5fsxi2nrtbhbhkbu2dsbn46qka7j4kf7peqd.onion
```

The terminal window also displays the system status as "arm64".

Figura 4.4: Comando mfp224o e output

⁷Il file viene avviato direttamente dalla cartella del sorgente, per cui apponiamo ./

⁸tesilm3jb64lw3upj4uu5fsxi2nrtbhbhkbu2dsbn46qka7j4kf7peqd.onion

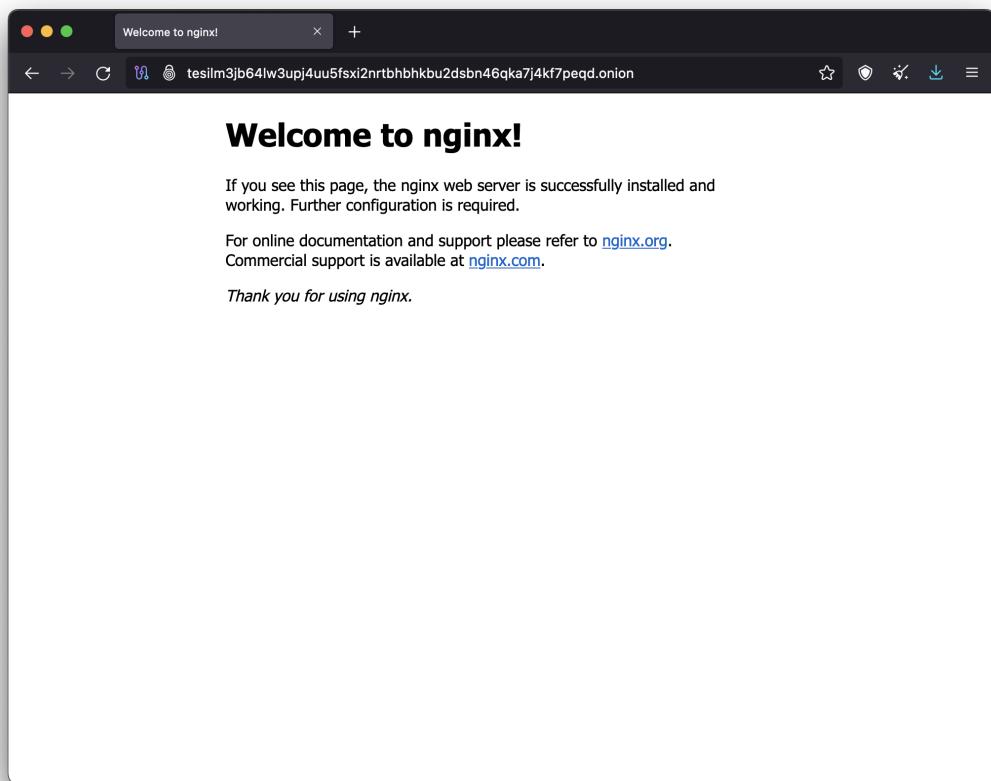


Figura 4.5: Connessione con il nuovo URL

4.4.3 Far conoscere il sito

Un'altra caratteristica importante per ogni sito è la capacità di essere trovato dall'utente finale, solitamente tramite un motore di ricerca, ci sono diversi motori che eseguono l'indexing di indirizzi onion. Tra questi c'è notevil⁹ funziona come un classico motore di ricerca, a differenza di altri motori come Torch che invece esegue la ricerca più in base al contenuto che all'url¹⁰

In NotEvil è possibile contattare gli amministratori del motore per richiedere l'aggiunta di un sito. Dalla pagina ufficiale è infatti presente il pulsante di contatto, dopo un breve controllo ci permette di inviare una richiesta anonima così che il sito possa essere analizzato e aggiunto.

Un altro sistema per far conoscere il sito agli utenti finali è l'utilizzo di siti specializzati come OnionDir¹¹ che permette di aggiungere il proprio sito in maniera completamente anonima.

In questo caso basta cliccare sul tasto Add Link nella toolbar in alto.

A questo punto possiamo inserire l'URL e una piccola descrizione, come NotEvil il

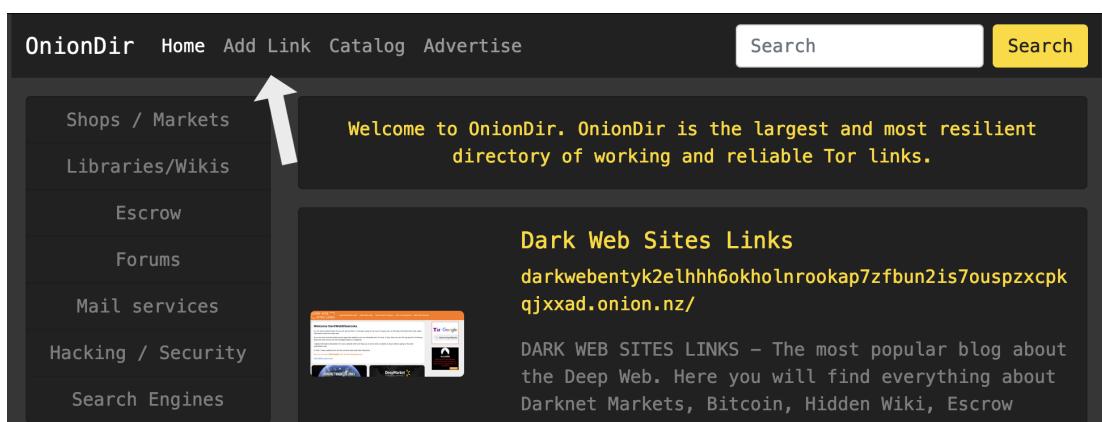


Figura 4.6: OnionDir, Pagina iniziale

sito sarà scansionato per evitare truffe o contenuti proibiti, come indicato nelle poche righe di descrizione dei termini del servizio.

⁹<http://notevilmtxf25uw7tskqxj6njlpbebymlrerfv5hc4tuq7c7hilbyiqd.onion>

¹⁰Un test di ricerca di 'proton mail' in entrambi mostra solo notEvil restituirci il sito corretto

¹¹<https://oniondiricuc4x2y5qbucg4jyp2ael5rxy7aahy5f4fbars2jkkf7vad.onion.nz>

The screenshot shows the OnionDir website interface. At the top, there is a navigation bar with links for Home, Add Link, Catalog, and Advertise. To the right of the navigation bar is a search bar with a yellow 'Search' button. On the left side, there is a sidebar with a dark background containing links to various services: Shops / Markets, Libraries/Wikis, Escrow, Forums, Mail services, Hacking / Security, and Search Engines. The main content area has a dark background and contains the following text:

Submit a new site

Found an interesting site? You can submit your site to OnionDir catalog. But be careful, we don't SCAM or Prohibited content such as CP, Weapons, Drugs. All new sites will be checked for scam and prohibited content, after verification, the site will be published in the OnionDir.

If you agree with this short terms – fill a form below.

Onion v3 address (without https:// and slashes at the end)

Example: oniondiricuc4x2y5qbucg4jyp2ael5rxy7aahy5f4fbars2jkkf7vad.

Specify correct v3 onion link for your site.

Description

Specify a short brief about this site.

At the bottom, there is a CAPTCHA field with the text "22 + 7 =", a text input field labeled "Enter code", and a yellow "Submit" button.

Figura 4.7: OnionDir, form aggiunta sito

Un altro sistema per far conoscere il sito è tramite il cosiddetto *Onion-Location Header* che permette di indicare al browser che il sito è disponibile anche su Tor, in questo modo quando la pagina viene visitata il browser può mostrare un pulsante per aprire il sito direttamente su Tor.

Il miglior esempio è il sito ufficiale di [Tor Project](#) che mostra un pulsante per aprire il sito onion quando viene visitato tramite la rete Tor¹².

Per aggiungere questo header è sufficiente aggiungere la seguente riga al file di configurazione del server

```
add_header Onion-Location http://
    tesilm3jb64lw3upj4uu5fsxi2nrtbhbhkbu2dsbn46qka7j4kf7peqd.
    onion$request_uri;
```

Oppure possiamo aggiungere direttamente il tag <meta> nella sezione header della pagina HTML, nel caso di configurazioni particolari

```
<meta http-equiv="onion-location" content="http://
    tesilm3jb64lw3upj4uu5fsxi2nrtbhbhkbu2dsbn46qka7j4kf7peqd.onion" />
```

L'implementazione di questa seconda tecnica è stata mostrata nel sito di prova tesi.miglio.dev che mostra il pulsante per aprire il sito onion quando viene visitato tramite Tor.

¹²Da tenere in considerazione che questo metodo richiede un sito normale per funzionare

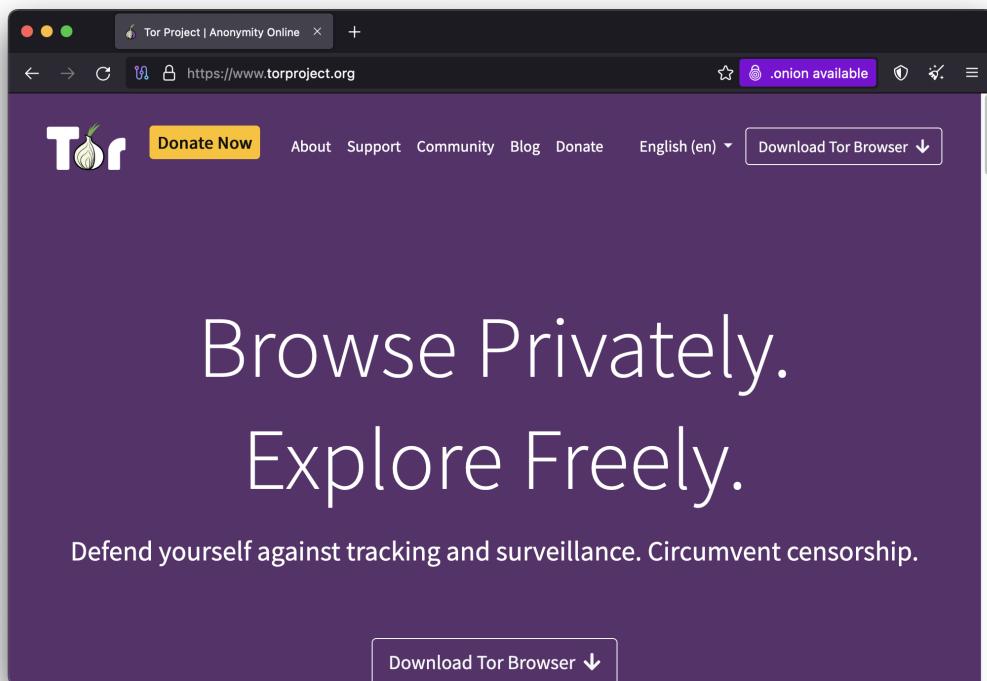


Figura 4.8: Tor Browser mostra il pulsante *.onion available* quando si visita il sito tramite Tor

Bibliografia

- [Aws] *AWS Launch Instance.* URL: <https://us-east-1.console.aws.amazon.com/ec2/home?region=us-east-1#LaunchInstances:>.
- [cat] cathugger. *mkp224o - vanity address generator for ed25519 onion services.* URL: <https://github.com/cathugger/mkp224o/>.
- [Cha81] David Chaum. «Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms». In: *ACM* (1981).
- [GD99] Syverson P. Goldschlag D. Reed M. «Onion Routing for Anonymous and Private Internet Connections». In: (1999).
- [Goo] Google. *Certificate Transparency.* URL: <https://certificate.transparency.dev/>.
- [MGRG] Paul F. Syverson Michael G. Reed e David M. Goldschlag. «Anonymous Connections and Onion Routing». In: () .
- [RDos] Paul Syverson Roger Dingledine Nick Mathewson. «Tor: The Second-Generation Onion Router». In: (agosto 2004).
- [Tora] Tor. *enable official repo.* URL: <https://support.torproject.org/apt/tor-deb-repo/>.
- [Torb] Tor. *Onion Services Overview.* URL: <https://community.torproject.org/onion-services/overview/>.
- [Torc] Tor. *Setup Onion Service.* URL: <https://community.torproject.org/onion-services/setup/>.
- [Tord] Tor. *Tor relay.* URL: <https://community.torproject.org/relay/>.
- [Tore] Tor. *Tor relay metrics.* URL: <https://metrics.torproject.org/rs.html>.
- [Torf] Tor. *Tor relay types.* URL: <https://community.torproject.org/relay/types-of-relays/>.
- [Tor13] Tor. *Rendezvous Specification - Version 3.* 2013. URL: <https://gitweb.torproject.org/torspec.git/tree/rend-spec-v3.txt>.
- [Web] WebServer History. URL: https://w3techs.com/technologies/history_overview/web_server/ms/q.

Ringraziamenti

In primo luogo vorrei ringraziare il Professore Fausto Marcantoni, non solo per avermi dato la possibilità di lavorare a questa tesi, ma anche per avermi fatto appassionare alla materia e ai suoi insegnamenti.

Ringrazio i miei genitori, per avermi sempre sostenuto ed insieme ai miei nonni per avermi permesso di studiare e di arrivare fino a questo punto.