Hluboký web

Hluboký web (též neviditelný web, anglicky deep web nebo invisible web) se definuje jako textové stránky, soubory nebo další informace přístupné prostřednictvím WWW, které webové vyhledávače nezahrnují do svého indexu.

Důvody vedoucí ke vzniku hlubokého webu

- webové vyhledávače nedokážou indexovat dynamicky se měnící stránky
- kontextuální web - některé weby vrací různý obsah podle přístupového kontextu – různé IP adresy, URL referer, atp.
- přístup na některé stránky je chráněn heslem, CAPTCHA, privátní weby
- Non-HTML obsah - některé webové vyhledávače neindexují rámce, obrázkové mapy, linky generovaná Javascriptem apod.
- mnoho webových vyhledávačů má omezení na počet indexovaných stránek z určité domény
- většina webových vyhledávačů preferuje indexování populárních stránek
- na některých stránkách je nabízen nelegální obsah
- obsah dostupný pomocí speciálního softwaru – tzv. **darknet**
    - Tor - The Onion Routing - umožňuje uživatelům přístup k serverům s pseudodoménou .onion anonymně skrytím uživatelovy adresy
    - I2P – Invisible Internet Project – P2P síť pro anonymní komunikaci
- webové archivy

Typologie hlubokého webu

- **Nepřehledný web** - obsahuje soubory, které mohou být indexovány roboty, ale z určitých příčin indexované nejsou. Roboti je z finančních důvodů neindexují.
- **Soukromý web** - skládá se ze stránek, které by robot dokázal indexovat, ale správce stránky mu to neumožňuje.
- **Speciální nebo vlastnické weby** - jde o část webu, ke které se dostaneme jen po splnění určitých podmínek (vyplnění registračního formuláře atd.)
- **Skutečně neviditelný web** - skládá se z informací, které roboti nedokážou indexovat, protože na ně nejsou naprogramované. Většinou se jedná o formáty PDF, spustitelné programy, flash, komprimované soubory atd.

Temný web

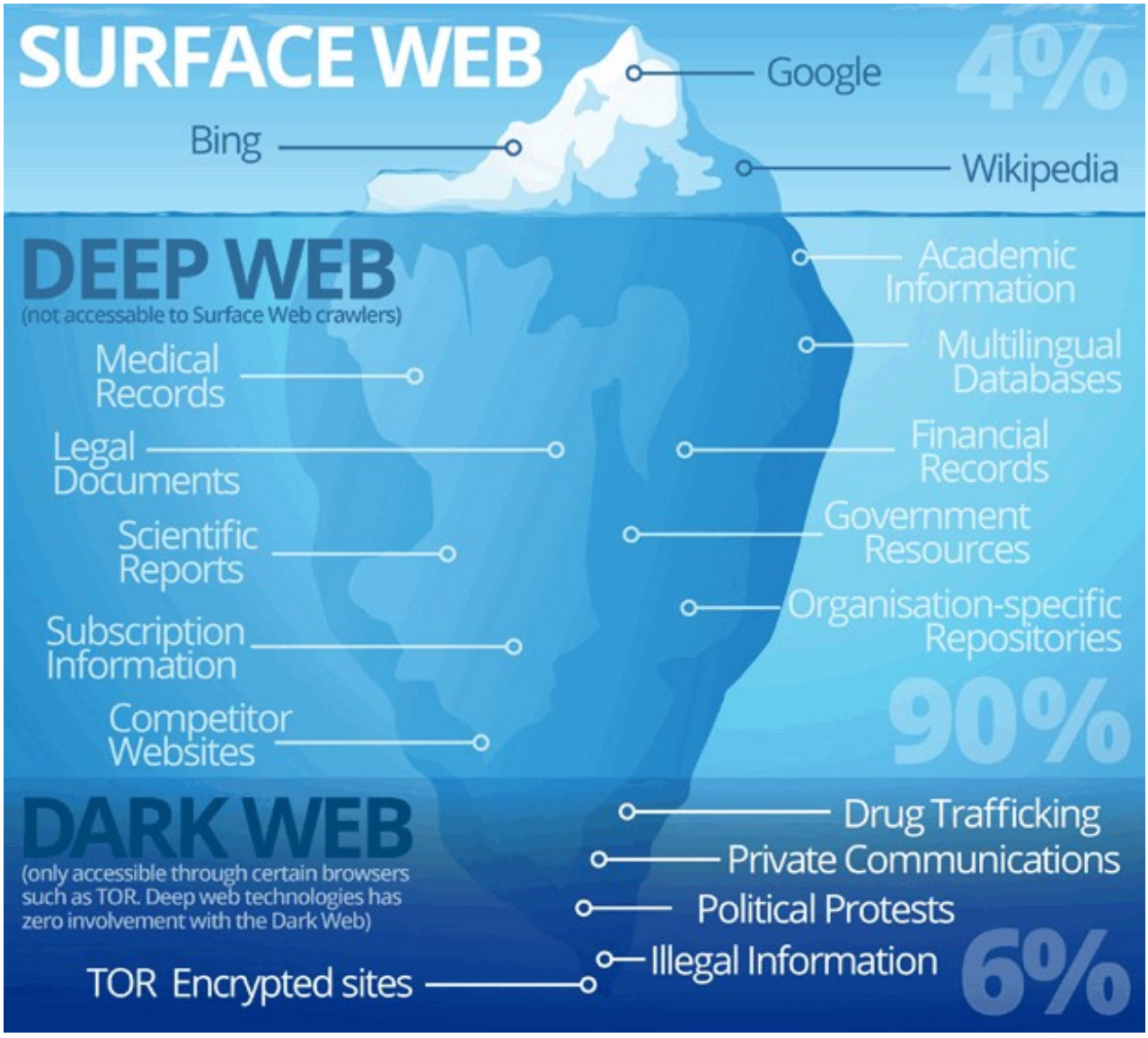
Temný web (anglicky dark web) je ta část obsahu World Wide Webu, která je umístěna na darknetech, tj. překryvných sítích, které využívají Internet, ale jsou přístupné pouze prostřednictvím speciálního softwaru nebo konfigurace. Temný web tvoří část hlubokého webu, což je oblast webu, kterou vyhledávače neindexují.

Darknety zahrnují jak malé peer-to-peer sítě, tak i velké a populární sítě jako například Freenet, I2P a Tor.

Deep web vs. Dark web

Můžeme se setkat se srovnáním, že hluboký web je temný web a naopak. Toto srovnání je mylné, hluboký web jsou dokumenty a webové stránky, které mají dostupný obsah pouze po přihlášení či jiném druhu autorizace nebo jednoduše nejsou indexované žádným vyhledávačem a jsou dostupné pouze znalostí URL.

Dalo by se říct, že dark web je podmnožina deep webu, protože se k němu můžeme dostat pouze za využití specifického softwaru a i případné autorizace a není dohledatelným běžným internetovým vyhledávačem.



Dělení webů, zdroj: https://www.thedarkwebsites.com/

Je dark web legální nebo nelegální?

Dark web je síť zcela legální. V mnoha zemích, které omezují a kontrolují přístup k internetu (Čína, Rusko), umožňuje lidem přístup k internetu a ke svobodným informacím. Jiní zase využívají síť pro její bezpečnost a vyšší úroveň anonymity. Proto se temný web používá například pro sdílení citlivých dokumentů či pro zkušenější a odbornější uživatelé k tunelování připojení domácí sítě či jiné lokální sítě (LAN), která nemá přidělenou veřejnou adresu či jednoduše nemá nastavené přesměrování portů.

Ovšem vzhledem k tomu, že dark web nabízí anonymitu, mohou se zde nacházet stránky s ilegálním obsahem, jako jsou stránky s prodejem drog, falešných identit, kradených kreditních a debetních karet, čísel bankovních účtů, zbraní, nájemné vraždy a jiné. V rámci fór se vyskytují kontroverzní a nelegální témata, na úložištích se vyskytuje i nelegální pornografie.

Využití dark webu

- Bezpečnější surfování po Internetu či po samotném dark webu
- Vyšší anonymita
- Bezpečné sdílení dokumentů, souborů a jiných dat
- Obcházení cenzury Internetu
- Whistleblowing (nelegitimní, neetické nebo nezákonné praktiky na pracovišti, které se dějí se souhlasem nadřízených a jdou proti veřejnému zájmu či ohrožují veřejnost, přičemž některé verze definice jsou navíc omezeny podmínkou, že upozorňovatel jedná v dobré víře a nesleduje vlastní prospěch a že situaci nelze vyřešit interními mechanismy) a anonymní úniky informací
- Politická aktivita
- Nelegální aktivity

Tor

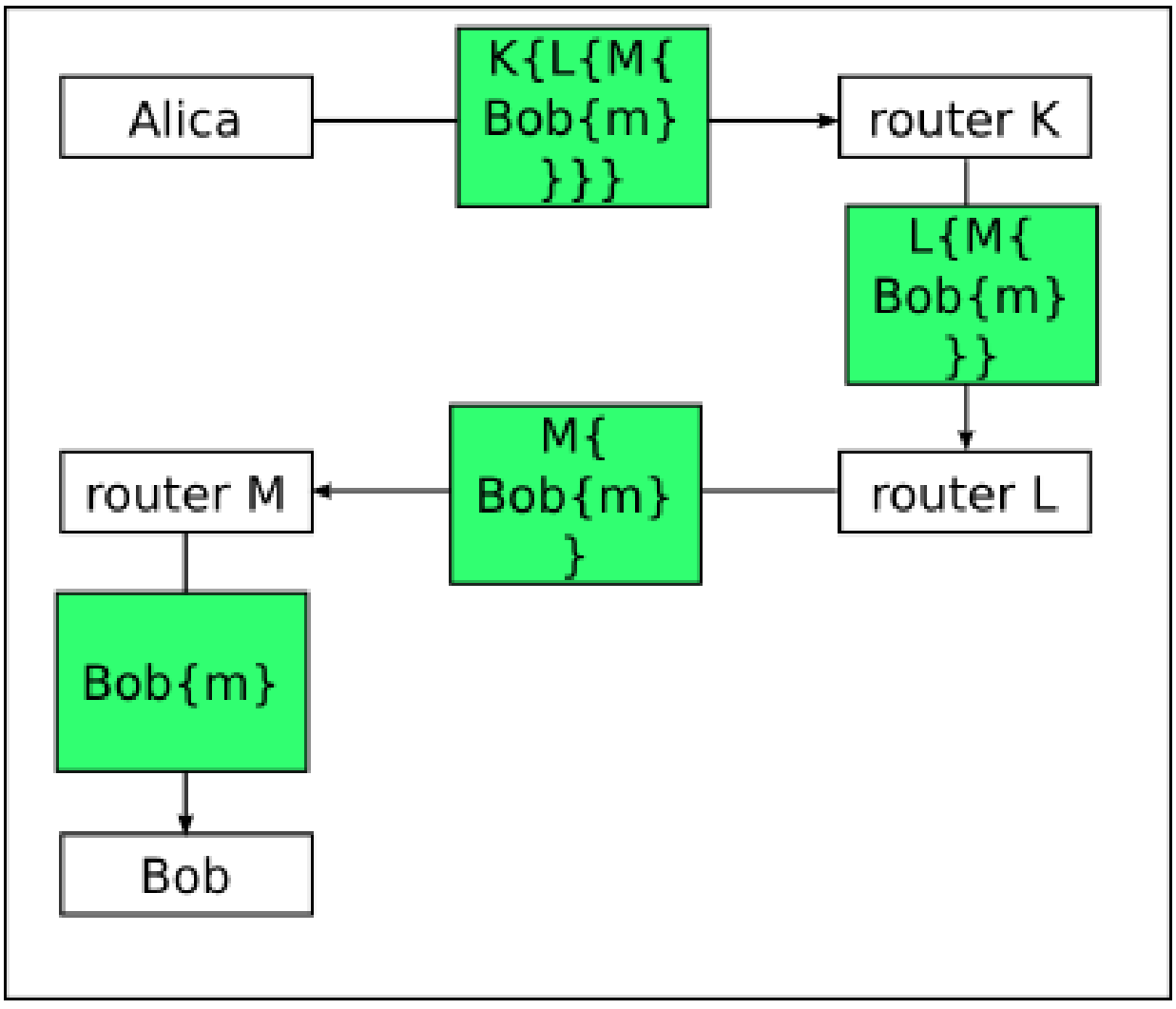
Tor je softwarový systém zajišťující anonymizaci uživatele při pohybu na Internetu, k čemuž využívá model klient-server. Uživatel využívá klientskou část a jeho datový tok prochází nejprve sítí Tor složené ze serverových částí a teprve pak k cílovému počítači. Tím je možné skrýt informace o IP adrese uživatele a dalších faktorů, které by ho mohly identifikovat. Díky používání Toru je obtížnější vysledovat stopy činnosti uživatele na Internetu včetně návštěv webových stránek, on-line příspěvků, programů pro komunikaci v reálném čase (instant messaging) a dalších forem komunikace. Je určen k ochraně osobních údajů uživatelů, jejich svobody, soukromí, a možnosti provádět důvěrné obchodování, tím že je chrání před sledováním jejich aktivit na

Internetu. Jedná se o open source software a síť Tor lze použít zdarma – ke stažení na: https://www.torproject.org/

Mixnet a princip Onion routingu

Pro anonymizaci komunikace na internetu se používá princip mixnetu (mix network) – každý komunikující uzel v síti je zároveň router. Úlohou každého routeru mixnetu je přeposílat zprávy, kdy každý router ví jen o routeru, od kterého zprávu obdržel a kterému routeru má dále zprávu předat. Žádný router neví, odkud se zpráva vzala a kde je její konečný cíl. Při posílání zprávy se musí použít vhodné adresování, aby odesílatel mohl adresovat konečného příjemce zprávy. Pointa je v tom, aby se nedal najít vztah mezi těmito adresami a IP adresami počítačů. Poté co zpráva projde mixnetem skrz několik routerů k zamýšlenému příjemci, až tehdy příjemce pozná, že je zpráva určená pro něho. (žádný z routerů po cestě to nevěděl). Skutečnost, že počítač je routerem mixnetu není tajemstvím – utajuje se, kdo s kým komunikuje a o čem.

Z konceptu mixnetu byl doslova odvozený koncept Onion routingu - routují se cibule. Každý router si vygeneruje pár klíčů. Pokud chce Alice poslat zprávu Bobovi skrz routery K, L, M, vyrobí Alice následující paket: Samotnou zprávu pro Boba zašifruje Bobovým veřejným klíčem. K tomu přidá pokyn, že router M má předat zprávu Bobovi. To celé zašifruje veřejným klíčem routeru M. K takovéto zprávě přidá pokyn pro router L, že má takovouto zprávu poslat routeru M. Takovouto zprávu zašifruje veřejným klíčem routeru L. K této nově vzniklé zprávě přidá pokyn pro Router K, aby zprávu předal routeru L a zprávu zašifruje veřejným klíčem routeru K.



Onion routing diagram, zdroj: https://www.root.cz/clanky/onion-routing-v-p2p-sietach-tor/

Takto připravený paket (cibuli) pošle Alice routeru K. Protože je cibule zašifrovaná jeho veřejným klíčem, umí ho router K dešifrovat. Router K odstraní vrchní vrstvu (slupku) cibule. Router K vidí pokyn, že má dále poslat cibuli routeru L a to udělá. Router L dostane paket zašifrovaný svým veřejným klíčem, tak ho rozšifruje (odstraní další vrstvu cibule) a přepošle ji dále. Takto to pokračuje, až paket dostane Bob.

Všimněte si, že router K neví, zda zpráva původně pocházela od Alice nebo ne, protože klidně mohl někdo routovat jinou cibuli skrz Alici. Stejně tak router M netuší, či po odeslání Bobovi zpráva bude putovat dále. Jednotlivé vrstvy cibule se autentizuji digitálním podpisem, aby se nedaly zprávy falšovat (protože zašifrovat jí veřejným klíčem může kdokoliv). Každý router může navíc zprávy pozdržet a přeposlat v jiném pořadí než je přijmul, a tím ztížit analýzu toku dat.

Teoretický koncept je sice funkční, ale skutečné implementace onion routingu jsou složitější, např. Šifrování veřejným klíčem je pomalá operace, což snižuje propustnost sítě a umožňuje jednoduchý denial-of-service zaplavením routeru zprávami. Cestu je vždy možné volit jinak, což pomáhá anonymitě, ale příliš častá změna může odhalit identitu komunikujících stran. Sítě implementující onion routing tak vytvářejí dočasné trasy – tunely, s omezenou životností danou bud časem (jednotky minut) nebo objemem přenesených dat.

TOR – The Onion Router

Tor slouží především jako proxy: účastník se připojí na některý z Tor routerů a nechá skrz síť poslat zprávu, která vyleze někde ze sítě ven a spojí se s cílovým serverem. Z API hlediska TOR klient funguje jako SOCKS proxy (Socket Secure – výměna dat skrz proxy server), takže je možné ho využít na proxovaní libovolných TCP spojení.

Fakt, že účastníci využívající Tor nemusí být zároveň routery, má výhodu v tom, že účastníkovi nestoupne přenos dat kvůli routování cizích zpráv. Má to ale bezpečnostní nevýhody – jednak z posledního routeru (hopu) na cestě odchází nešifrovaný plaintext na cílový server (např. u HTTP protokolu), a jednak útočník sledující skutečného odesílatele i příjemce dokáže korelovat pakety, tedy může dokázat, že spolu dvě strany komunikovaly. Tor se nesnaží bránit tomuto typu útoku (tzn. Traffic confirmation attack). Spíše se brání obecnější analýze toku dat, ze které by se útočník dozvěděl, koho má vlastně sledovat a které uzly má napadnout. Předpokládá se, že útočník dokáže sledovat část sítě (ale ne celou), že dokáže vytvářet, měnit nebo mazat pakety na TCP/IP vrstvě, kompromitovat některé onion routery a jejich klíče anebo pustit své vlastní onion routery do sítě. Tor je navržený jako systém s nízkou latencí, nevkládá žádné zdržení do posílaných paketů, ani nemění jejich pořadí. Toho sice dělá rychlejším, ale náchylnějším na časovou analýzu toku dat.

Kromě proxy funkčnosti nabízí Tor Hidden services, což jsou služby (např. Web server), které jsou skryté a dají se kontaktovat jen skrz Tor síť. Tyto služby se s uživatelského hlediska kontaktují pomocí pseudo-TLD domény .onion, která slouží k rozeznání běžných služeb od služeb skrytých.

Tor nenabízí žádný ekvivalent DNS, namísto toho jsou všechny hidden services a jejich FQDN zveřejněné jako soubor na webu, který si klientská aplikace stáhne.

Historie Toru

Alfa verze softwaru cibulového směrování byla uvedena 20. září 2002. Prezentaci sítě Tor provedli 13. srpna 2004 na 13. USENIX sympóziu věnovanému bezpečnosti Roger Dingledine, Nick Mathewson a Paul Syverson, kteří ho představili jako „The Second-Generation Onion Router“. Původně byl Tor sponzorován United States Naval Research Laboratory. V roce 2004 – 2005 byl finančně podpořen Electronic Frontier Foundation. Tor projekt je od prosince 2006 výzkumně-vzdělávací nezisková organizace se sídlem v USA, která získává rozmanitou finanční podporu.

.onion

.onion je generická pseudo-doména nejvyššího řádu označující skryté služby anonymní sítě Tor. Doména není obsažena v kořenových DNS serverech, programy, jako například webový prohlížeč, odesílají dotaz na proxy server, který ho přeposílá přes síť Tor.

Adresy s pseudo-doménou .onion bývají automaticky generovány Tor klienty spolu s veřejným PGP klíčem a skládá se z 16 znaků – malých písmen a číslic. Jsou tak většinou hůře zapamatovatelné – např. zqktlwi4fecvo6ri.onion

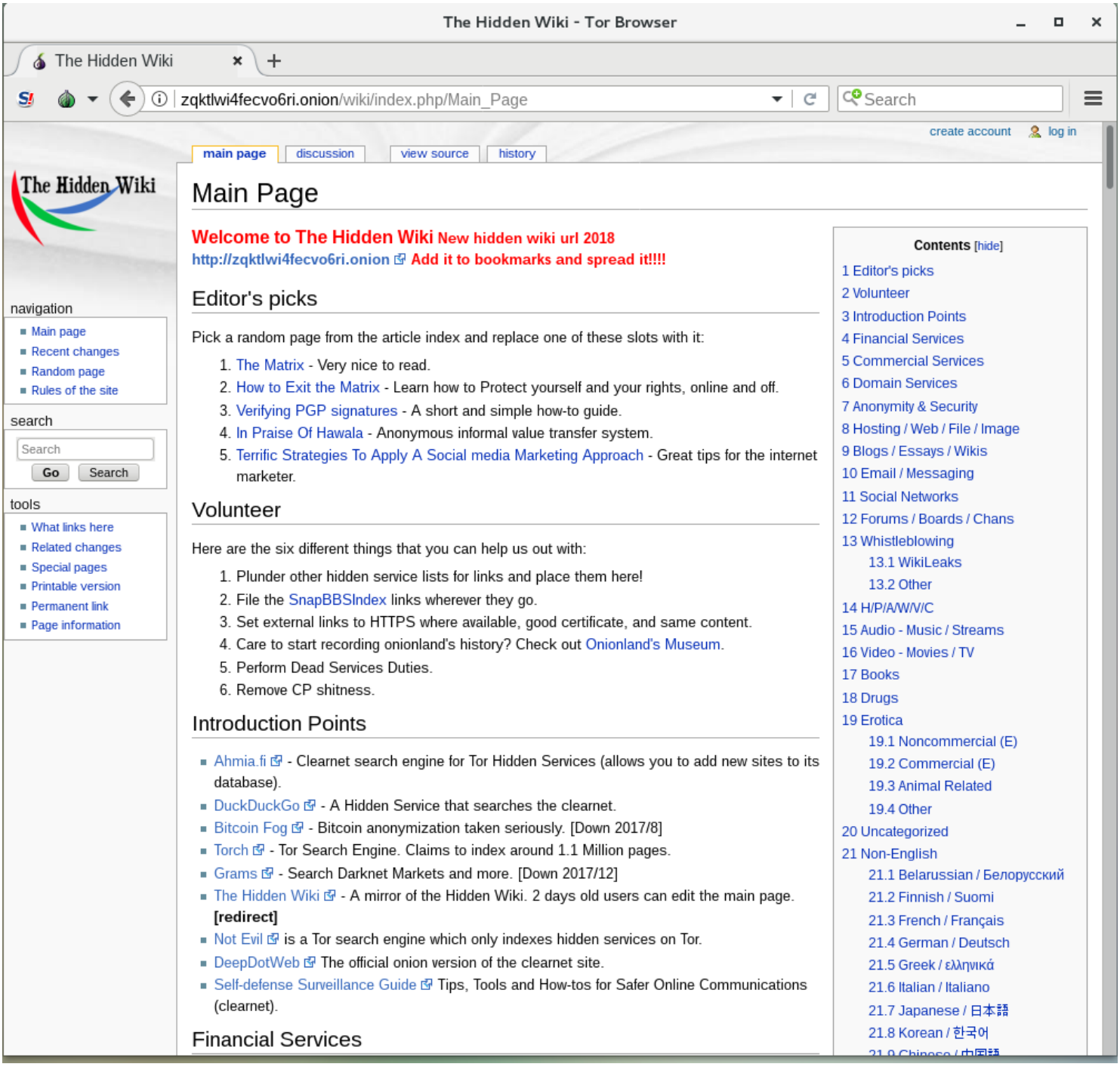
Překlad .onion adresy:

1. A hidden service calculates its key pair (private and public key, asymmetric encryption).
2. Then the hidden service picks some relays as its *introduction points*.
3. It tells its public key to those *introduction points* over Tor circuits.
4. After that the hidden-service creates a *hidden service descriptor*, containing its public key and what its *introduction points* are.
5. The hidden service signs the *hidden service descriptor* with its private key.
6. It then uploads the *hidden service descriptor* to a *distributed hash table* (DHT).
7. Clients learn the .onion address from a hidden service out-of-band. (e.g. public website) (A $hash.onion is a 16 character name derived from the service's public key.)
8. After retrieving the .onion address the client connects to the DHT and asks for that $hash.
9. If it exists the client learns about the hidden service's public key and its *introduction points*.
10. The client picks a relay at random to build a circuit to it, to tell it a *one-time secret*. The picked relay acts as *rendezvous point*.
11. The client creates a *introduce message*, containing the address of the *rendezvous point* and the *one-time secret*, before encrypting the message with the hidden service's public key.
12. The client sends its message over a Tor circuit to one of the *introduction points*, demanding it to be forwarded to the hidden service.
13. The hidden service decrypts the *introduce message* with its private key to learn about the *rendezvous point* and the *one-time secret*.
14. The hidden service creates a *rendezvous message*, containing the *one-time secret* and sends it over a circuit to the *rendezvous point*.
15. The *rendezvous point* tells the client that a connection was established.

16. Client and hidden service talk to each other over this *rendezvous point*. All traffic is end-to-end encrypted and the *rendezvous point* just relays it back and forth. Note that each of them, client and hidden service, build a circuit to the *rendezvous point*; at three hops per circuit this makes six hops in total.

Kde začít s Hidden Services

Hidden Wiki: http://zqktlwi4fecvo6ri.onion/wiki/index.php/Main_Page



Zdroje:

https://en.wikipedia.org/wiki/Deep_web

https://cs.wikipedia.org/wiki/Temn%C3%BD_web

https://www.thedarkwebsites.com/

https://cs.wikipedia.org/wiki/Tor_(software)

https://www.root.cz/clanky/onion-routing-v-p2p-sietach-tor/

http://www.security-portal.cz/clanky/tor-onion-router-syst%C3%A9m-pro-vysoce-anonymn%C3%AD-%C5%A1ifrovan%C3%BD-p%C5%99%C3%ADstup-k-internetu

https://tor.stackexchange.com/questions/672/how-do-onion-addresses-exactly-work?rq=1

https://topnawebe.sk/top-6-najtemnejsich-zakuti-internetu/