

« Deep web », « dark web » et darknets



Ces appellations sont d'un usage courant, que ce soit dans la presse, les conversations courantes ou même les publications scientifiques et pourtant mal comprises.

Mais de quoi parle-t-on ?

Exploring The Hidden Internet

Deep Web



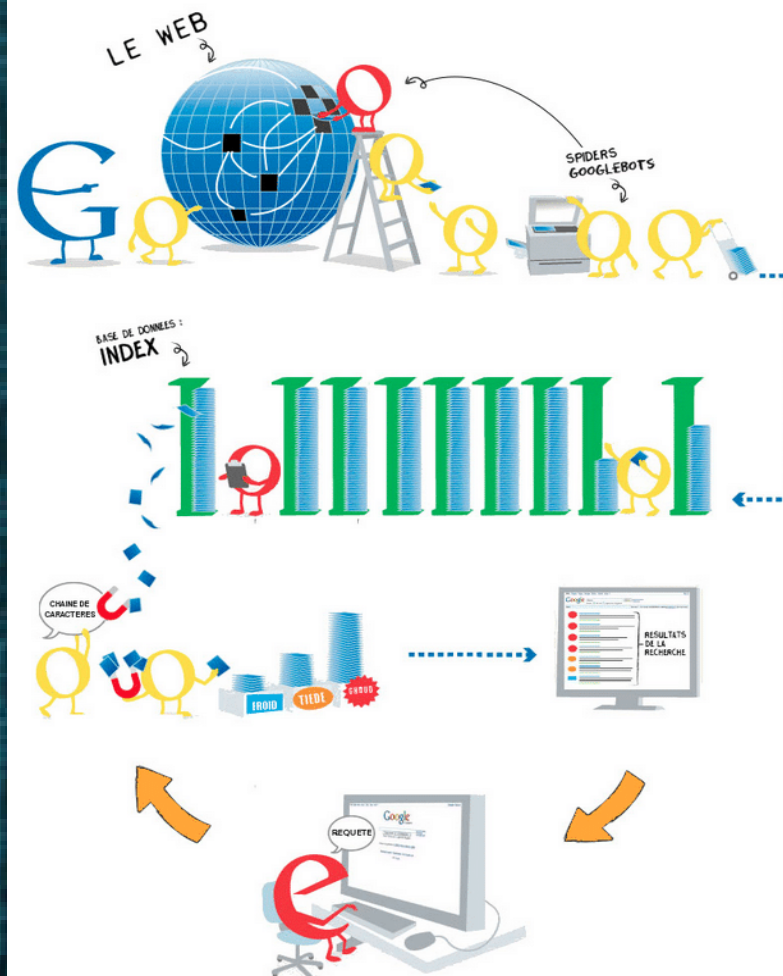
Dark Web



« Web visible » ou « web surfacique »

Le Web surfacique (aussi appelé “web visible” ou “indexable”) est la partie du World Wide Web qui est accessible en ligne par l'ensemble des internautes et indexable par les moteurs de recherche tels que Google, Yahoo, Bing, etc.

COMMENT FONCTIONNE LA RECHERCHE GOOGLE



ROBOT D'EXPLORATION

Des logiciels d'exploration, appelés "crawlers" ou "spiders", explorent régulièrement le web en parcourant les réseaux de liens. Ils scannent les pages qu'ils trouvent.

ROBOT D'INDEXATION

Le robot d'indexation extrait les données des pages scannées (URL, titre, mots-clés sous forme de chaînes de caractères), puis les classe pour former l'index de la base de données de Google

MOTEUR D'INTERROGATION

Interface entre l'internaute et l'index, le moteur d'interrogation récupère la requête tapée dans le formulaire, la compare aux données de l'index, puis recueille les résultats qu'il classe avant de les afficher. (voir l'algorithme de recherche Google)

« Deep web »

Toutes les pages web dont le contenu va être différent en fonction de l'utilisateur font partie du « deep web »

- toutes les pages non accessibles sans login ;
- toutes les pages qui vont être modifiées dynamiquement selon qui y accède ;
- toutes les pages dont l'auteur a demandé à ce qu'elles ne soient pas indexées par les moteurs de recherche

Le contenu non lié (pages sans backlinks)	Les pages sans backlinks ne peuvent pas être atteintes par les robots et sont donc dans le « Deep Web »
Le contenu de script	Des scripts JavaScript présents sur le site peuvent parfois bloquer, le plus souvent involontairement, les robots d'indexation
Les formats non indexables	formats de données incompréhensibles par les moteurs de recherche. Les moteurs de recherche s'améliorent peu à peu pour réussir à indexer le plus de formats possibles.
Le contenu trop volumineux	Base de données trop volumineuse
Les contenus privés (Web privé)	Certaines pages sont inaccessibles aux robots, car le Webmaster a inséré le fichier "robot.txt" dans le code du site afin de n'autoriser l'indexation que de certaines pages ou documents.
Le contenu à accès limité	sites Web nécessitent une authentification avec un login et un mot de passe pour pouvoir accéder aux contenus
Le contenu dynamique	L'adresse URL est générée à la demande de l'internaute en tenant compte des critères qu'il aura saisi dans un formulaire
Le contenu sous un nom de domaine non standard	Il s'agit de sites Web avec un nom de domaine dont la résolution DNS n'est pas standard avec par exemple une racine qui n'est pas enregistrée chez l'ICANN (l'Internet Corporation for Assigned Names and Numbers)

Vous naviguez sur le « deep web » quand vous :

- consultez vos emails sur Gmail par exemple;
- vous connectez à l'espace client de votre opérateur téléphonique ;
- lorsque vous consultez un doc partagé sur Google Drive ;
- vous allez sur l'Intranet de votre entreprise ;
- parlez à vos amis sur un serveur Discord

« dark web »



- Une partie du « deep web » ;
- le contenu hébergé sur les darknets (comme TOR (aussi appelé Onionland), Freenet ou i2P) ;
- Ce qui caractérise son contenu : la volonté des créateurs des sites de ne pas indexer les contenus et les spécificités techniques pour y accéder.

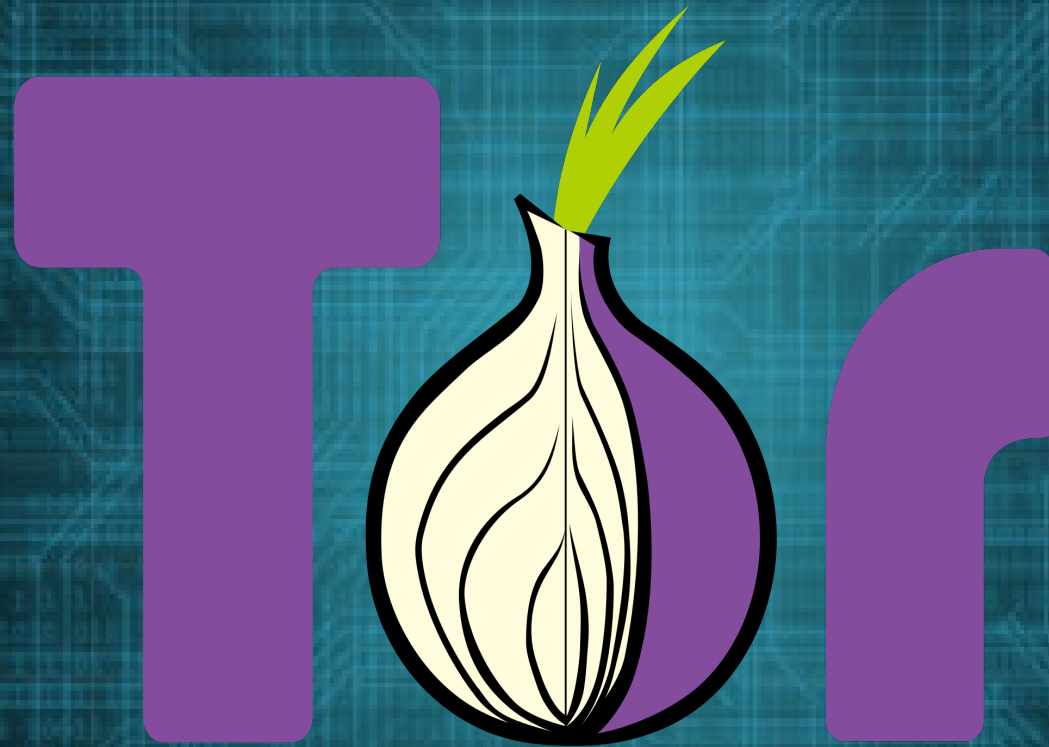
« Darknet »

Darknet est le support qui va permettre d'aller sur le « dark web »

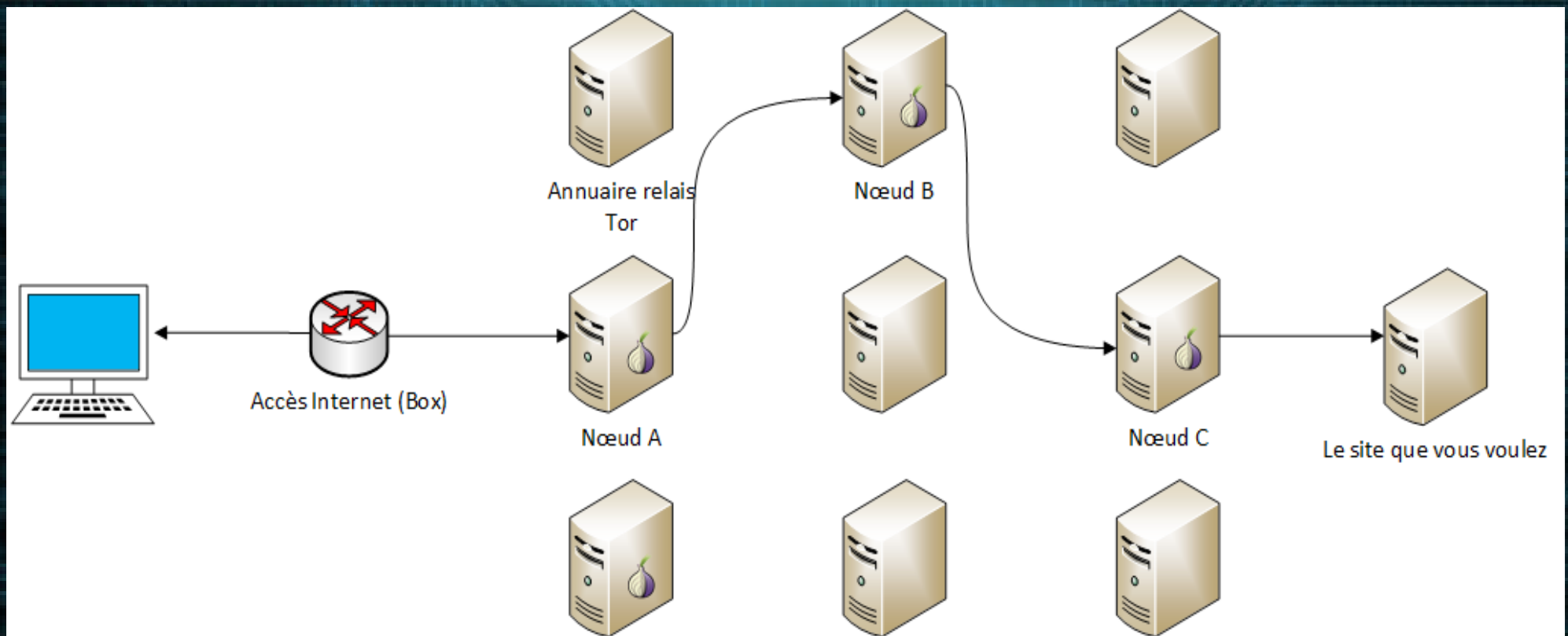
Il suppose:

- L'usage de l'infrastructure internet ;
- L'existence d'un protocole spécifique qui permet la constitution d'un sous-réseau chiffré;
- Une architecture décentralisée de type pair-à-pair.

Tor (« the oignon router »)



Plutôt que de créer une liaison directe entre le client et le serveur, il fait rebondir le trafic TCP (UDP n'étant pas supporté) sur trois serveurs relais avant d'interroger le serveur cible (site web). Ces trois serveurs sont appelés nœuds (plus de 7000) et sont mis en place par des bénévoles.



Internet = infrastructure composée d'un ensemble standardisé de protocoles de transfert de données permettant l'échange d'information numérique

web « de surface » = contenu web indexé par les moteurs de recherches populaires



WIKIPÉDIA
L'encyclopédie libre



« deep web » = contenu web simplement non indexé par les moteurs de recherches populaires, vous y allez tous les jours...



SOCIÉTÉ
GÉNÉRALE



Les darknets = des sous-réseaux d'internet utilisant des protocoles spécifiques pour permettre l'échange de contenu de manière généralement chiffrées



Bibliographie et webographie :

- Jean-François Perrat, « Un « Deep / dark web » ? Les métaphores de la profondeur et de l'ombre sur le réseau Tor », Netcom [En ligne], 32-1/2 | 2018, mis en ligne le 17 octobre 2018. URL : <http://journals.openedition.org/netcom/3134> ; DOI : <https://doi.org/10.4000/netcom.3134>
- TOR, LA FACE CHIFFRÉE D'INTERNET, entretien avec Lunar, Entretien réalisé par Camille Bosqué ; DOI 10.3917/vaca.069.0079 : <https://www.cairn.info/revue-vacarme-2014-4-page-79.htm>
- Jean-François Perrat, « Un « Deep / dark web » ? Les métaphores de la profondeur et de l'ombre sur le réseau Tor », Netcom [En ligne], 32-1/2 | 2018, mis en ligne le 17 octobre 2018. URL : <http://journals.openedition.org/netcom/3134> ; DOI : <https://doi.org/10.4000/netcom.3134>
- Jean-Paul Pinte, « Les jeunes et le Dark Web », Terminal [En ligne], 123 | 2018, mis en ligne le 31 décembre 2018, URL : <http://journals.openedition.org/terminal/3278> ; DOI : <https://doi.org/10.4000/terminal.3278>
- <https://www.lebigdata.fr/tor-tout-savoir>
- <https://www.clubic.com/antivirus-securite-informatique/vpn/dossier-361995-le-reseau-tor-est-il-infaillible-.html>
- <http://www.ibibliotech.fr/4641-a-la-decouverte-du-reseau-tor>
- <https://www.zdnet.fr/blogs/50-nuances-d-internet/le-darkweb-n-existe-pas-39858124.htm>
- <https://forums.cnetfrance.fr/tutoriels-reseaux-et-internet/653797-comment-etre-anonyme-sur-internet-et-le-web-av>
- <https://www.rennard.org/Darknet/presentation.html>

Vidéos :

<https://www.youtube.com/watch?v=nAlxgbHbDYE>

<https://www.youtube.com/watch?v=QRYzre4bf7I>

https://www.youtube.com/watch?v=9g2STPKNk_8

<https://www.youtube.com/watch?v=zjqw1VnuKLs>