Projet de semestre
Système d'information géolocalisation IP

Ingénierie Des Technologies de l'Information

Auteur : Filmon TEWELDE

Professeur du projet : Gérald Litzistorf

mars 2017

Descriptif

La géolocalisation IP permet par exemple de connaître la provenance des requêtes adressées à un serveur web.

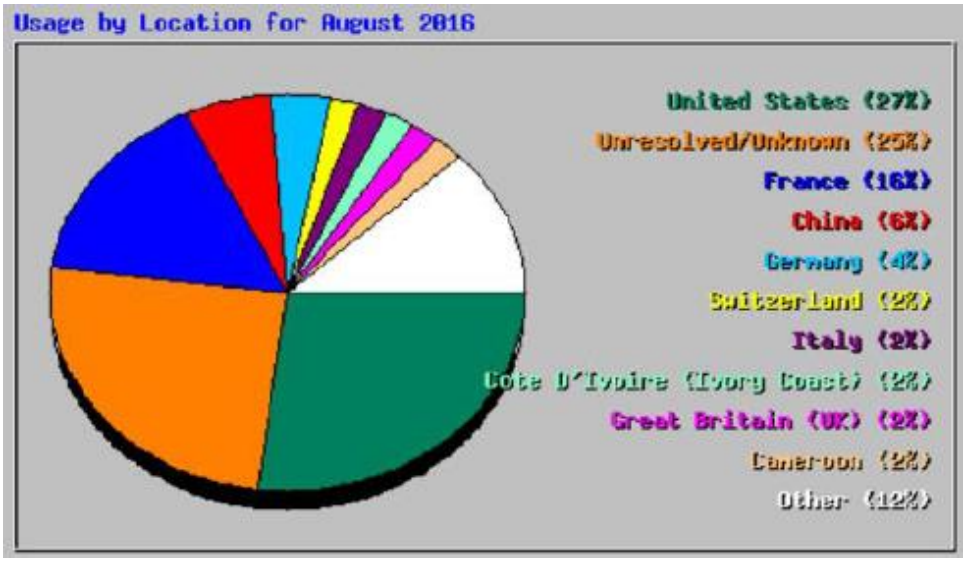


Figure 1 : Illustation avec les statisique août 2016 de www.tdeig.ch

Peut-on faire confiance à ces résultats ?
Peut-on affiner ces résultats

Travail demandé

Cette étude comprend les étapes suivantes :

1. Etudes théorique & état de l'art
   Comment fonctionne la géolocalisation IP ?
   Quels sont les principaux mécanisme utilisés (whois, BGP, …) ?
   Comment détecter les robots ?
   Quels sont les principaux risques d'erreur ?
2. Mise en œuvre sur un serveur web
   Configurer Nginx pour remplacer le serveur Apache actuel (www.tdeig.ch)
   Utiliser 2 outils différents pour corréler les résultats
3. Profiter de la durée (env. 25 semaines) du projet pour vérifier que la géolocalisation IP donne des résultats utiles et justes
   Quelles sont les adresses IP à ignorer ?

Sous réserve de modification en cours du projet de semestre

Table des matières

1 Introduction

Aujourd'hui, la géolocalisation ne fait que de prendre de plus en plus d'ampleur dans nos systèmes informatiques. Nous l'utilisons tous les jours, que nous le sachions ou pas. Mais est-elle assez précise pour des tâches importantes ?

Actuellement, nous utilisons cette technologie dans plusieurs domaines différents. *La redirection* est une des tâches principales, par exemple lorsque nous allons sur « www.google.com » nous sommes automatiquement redirigés vers le domaine « Google » de notre pays. *La publicité ciblée* utilise tout autant la localisation IP, grâce à la localisation la publicité devient tout de suite plus pertinente pour l'utilisateur et il sera plus à même de cliquer dessus. Elle est aussi grandement utilisée pour la restriction de contenu multimédia, par exemple Netflix une société de vidéo à la demande utilise la localisation IP pour modifier leur catalogue de films et séries en fonction du pays par lequel nous accédons à leur site web. Les droits de diffusion ne sont pas les mêmes dans tous les pays et grâce à des technologies telle que la localisation IP des services de vidéo à la demande a pu évoluer. Dans un registre un peu plus sérieux, la localisation IP permet de localiser des cybercriminels, depuis quelque année les sites à contenus pédopornographiques ne font que d'augmenter, actuellement les fournisseurs d'accès internet interdisent simplement ces sites web. Mais la plupart du temps, cela ne suffit pas, grâce à la localisation IP, nous pouvons localiser les processeurs et qu'ils soient jugés.

Même si aujourd'hui la précision que nous possédons nous suffit pour ce type d'action, elle ne suffira pas toujours. Nous sommes en train d'aller vers un monde ou toute la téléphonie devient IP et que va-t-il se passer lorsque nous appellerons une ambulance, les pompiers ou la police et que nous ne pourrons pas communiquer nos positions ? Comment allons-nous être redirigé vers le central le plus proche ? Ce sont des secteurs ou la précision est primordiale.

C'est donc pour tous ces différents secteurs que la localisation IP est importante et doit être étudié.

1.1 Contexte

Mais avant de rentrer dans le vif du sujet, il faut savoir qu'il n'existe pas un seul type de géolocalisation, mais une multitude comme par exemple la localisation grâce au GPS ou même par satellite. Dans ce travail, nous allons nous concentrer sur la localisation d'un hôte à travers son adresse IP.

Pour la partie pratique de ce travail le système de géolocalisation IP n'aura besoin d'une précision au pays. Mais dans la partie théorique, nous approfondirons un peu plus le sujet et nous irons jusqu'à la ville, et même la rue**.**

2 Etude théorique

2.1 Fonctionnement de la géolocalisation IP

Avant de parler de géolocalisation IP, nous devons tout d'abord voire comment est faite la gestion de l'adressage IP.

La gestion de l'internet est faite de façons très hiérarchiques. Nous avons donc au top de la gestion d'internet un organisme nommé ICANN (Internet Corporation for Assigned Names and Numbers, la société pour l'attribution des noms de domaines), leur définition de cet organisme est « *La Société pour l'attribution des noms de domaines et des numéros sur Internet est un organisme à but non-lucratif responsable de la sécurité, la stabilité et la coordination mondiale du système d'identificateurs uniques de l'Internet.* ». [1]

Dans cet organisme, le département IANA (Internet Assigned Numbers Authority) qui à lui pour but de s'occuper précisément de la répartition des adresses IP au niveau mondial. Quand ils se sont occupés de la réparation mondiale, ils passent la main aux RIR (Registre Internet Régional).

Les RIRs sont au nombre de 5 dans le monde réparti à peu près par continent.

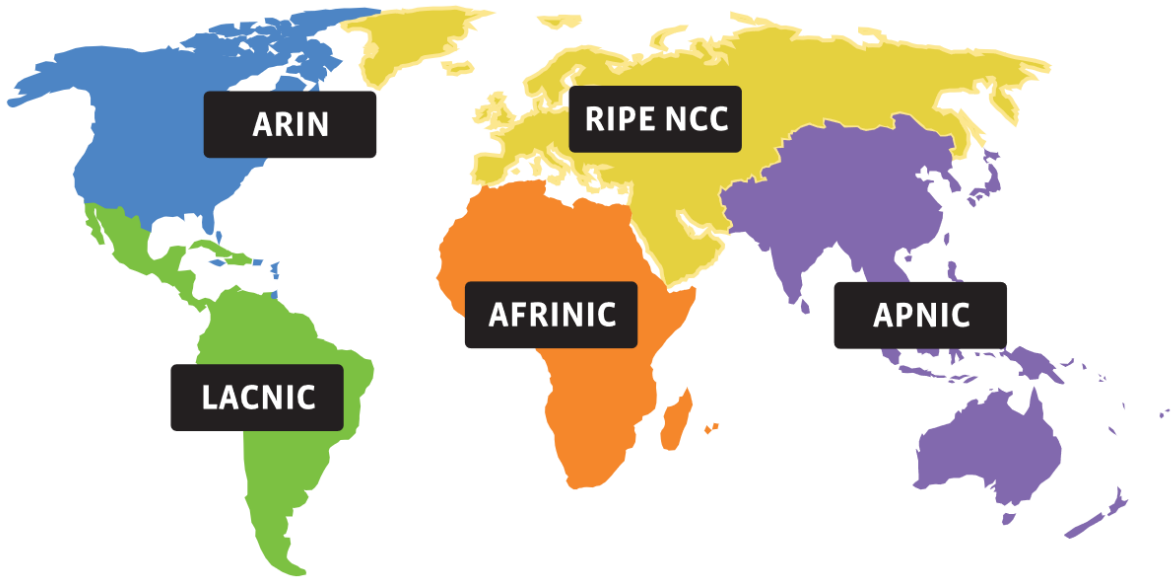


Figure 2 : Plan de répartition des RIRs

Chaque RIR on après une liste de membres à qu'ils distribuent leurs tranches adresses IP de la manière la plus intelligente possible. Les membres des RIR sont appelés LIR (Local Internet Registre), ce sont la plupart du temps des fournisseurs d'accès internet (par exemple Swisscom en suisse) qui après fournissent eux-mêmes leurs adresses IP à leur client selon un abonnement.

[1] Site officiel de l'ICANN, https://www.icann.org/fr

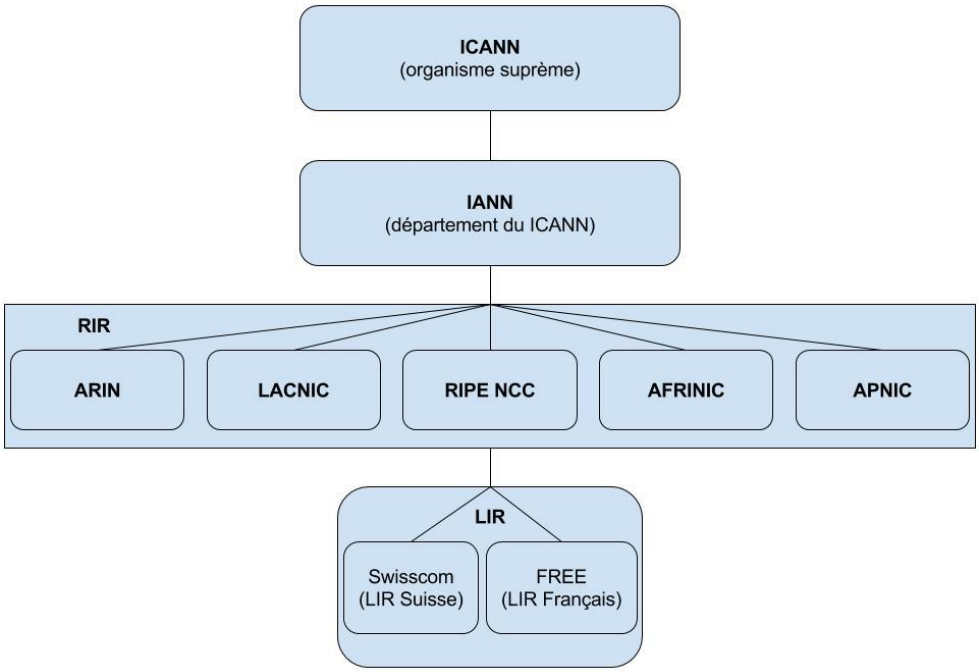


Figure 3 : Schéma de la hiérarchie internet

Maintenant que nous avons pris connaissance de toute cette hiérarchie, nous voyons que les adresses ont déjà une répartition géographique. Les RIRs nous proposent leur base de données avec un accès libre, le seul problème, c'est que leurs bases de données ne peuvent nous donner qu'au maximum le pays d'origine de l'IP. Cependant si nous voulons des localisations plus précises nous devrions voir avec les LIRs, car ils distribuent leurs adresses comme ils le désirent. Contrairement aux RIRs les fournisseurs d'accès (LIR) ne sont pas des organismes à but non-lucratif, ils ne vont donc pas nous fournir leur base de données gratuitement, et ils ne vont pas du tout fournir leurs bases de données surement pour des raisons de sécurités ou de respect de la vie privée de leurs clients.

Par conséquent si nous voulons une base de données précises, nous devons nous adresser à des entreprises tels que GeoNames ou MaxMind qui sont spécialisé dans la géolocalisation IP. Ce type d'entreprise a pour but de récolteur les informations géographiques des adresses IP partout dans le monde. Mais cette fois, elles nous proposent plusieurs types de base de données, plus nous désirions être précis plus la base de données sera chère.

Base de données	Pays	Ville
Application	Segmentez vos visiteurs par pays	Lorsque des données plus précises sont nécessaires
Continent	☑	☑
Pays	☑	☑
Subdivisions régionales		☑
Ville		☑
Code postal		☑
Latitude et longitude approximatives		☑
Fréquence de mise à jour	Hebdomadaire	Hebdomadaire
Redevances de licence de site (inclut un mois de mises à jour)	$50	$370
Facturation mensuelle	24 $ par mois	100 $ par mois
Facturation annuelle	314 $ la première année et 288 $ chacune des années suivantes.	1470 $ la première année et 1200 $ chacune des années suivantes.
	Plus d'informations	Plus d'informations

Figure 4 : Abonnement à base de données MaxMind

Bien que MaxMind dans cette image nous propose des bases de données payantes, ils ont aussi des bases de données gratuites qu'ils fournissent sur le site, mais nous affirment que ces bases de données ne seront pas aussi précises que la version payante, que cette base de données est moins souvent mise à jour et qu'officiellement et qu'ils ne fourniraient aucune aide.

2.2 Les principaux mécanismes utilisés

Il existe deux types de mesure pour la géolocalisation IP. La première est la mesure passive et la deuxième est la mesure active. Le nom de ces mesures varie entre les différents documents que l'on peut trouver. Ceux-là sont tiré de la thèse de doctorat de M. Bamba Gueye[2].

2.3 Mesure passive

La mesure passive est donc la technique qui va utiliser des bases de données comme celle de l'ICAAN et surtout les bases de données d'entreprise telles que MaxMind et GeoNames. Il existe aussi un service nommé « WHOIS » qui fournissent des informations sur les propriétaires de domaines.

[2] **Cheickh Ahmadou Bamba GUEYE, Localisation géographique des hôtes dans l'internet basée sur la multilatération,** http://edmi.ucad.sn/~gueye/articles/These_bamba.pdf (08/12/08)

2.3.1 WHOIS

Qu'est-ce que WHOIS ? C'est un service qui a été mis au point par l'ICANN pour fournir les informations des titulaires de nom de domaine. De base, il a été créé pour aider les utilisateurs de ARPANET en 1982, mais il a bien évolué depuis. Actuellement, les données WHOIS répondent au critère mise en place par le contrat RAA[3] (Registrar Accreditation Agreement, en français le contrat d'accréditation des bureaux d'enregistrement), avec des informations tel que nom du titulaire, le numéro de téléphone, un e-mail, l'adresse, etc.

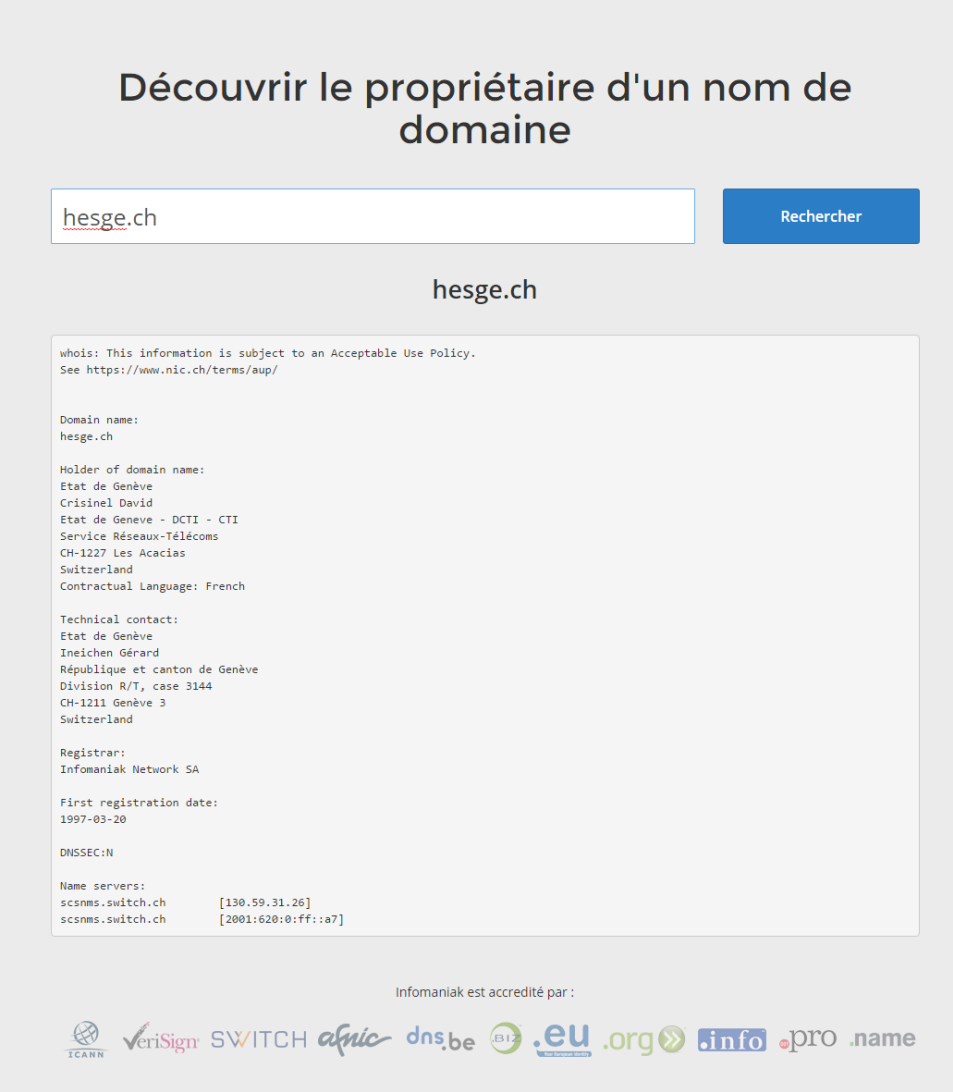


Figure 5 : Exemple de recher sur WHOIS d'Infomaniak

Comme nous pouvons le voir ici, nous avons toutes les informations sur le nom de domaine « hesge.ch » ceci est fourni par le serveur WHOIS d'Infomaniak.

[3] **Contrat RAA**, https://www.icann.org/resources/unthemed-pages/approved-with-specs-2013-10-31-fr#whois-accuracy

Malheureusement, les bases de données WHOIS ne sont pas centralisées. Chaque RIR en possède une qu'il met mise en commun par les opérateurs de registre (tels qu'Infomaniak Network SA ou Switchplus Ltd, pour la Suisse)[4] , ce qui peut par moment nous empêcher de trouver un titulaire de domaine.

Le contrat RAA a été mis à jour en 2013 pour améliorer l'exactitude des informations contenues dans les bases de données, si nous regardons l'article 1 du chapitre intitulé Spécification du programme d'exactitude du WHOIS du contrat RAA le bureau d'enregistrement à l'obligation de vérifier les données du titulaire de domaine. Pour être plus efficace ils ont aussi ajouté l'article 5 du même chapitre ou ils précisent qu'ils ont la possibilité de résilier ou mettre en attente l'enregistrement du domaine.

2.3.2 GeoCluster

Le GeoCluster est donc la méthode qui recueille les informations des BGPs (Border Gateway Protocol) et les informations de localisation depuis les formulaires de site web. Pour se faire, il faut établie l'hypothèse que deux adresses IP similaires sont proches l'une de l'autre. De plus s'il y a quelques hôtes d'où nous sommes sûrs de leur position et qui sont dans le même pool, nous pouvons donc confirmer que l'hypothèse s'avère correcte. Malheureusement, la précision de cette technique dépend beaucoup de la façon dont les ISPs (Internet Service Providers) séparent leur pool d'adresses.

Les entreprises tels que MaxMind ou GeoNames fournissent donc ce type de base de données. Malheureusement, l'exactitude de ces bases de données serait de 62% au niveau des villes[5]. Par conséquent les applications vont utiliser leurs bases de données avec lesquelles ils vont faire de simples requêtes pour trouver la localisation de l'adresse IP.

2.4 Mesure Active

Les mesures actives sont les mesures qui en générale génèrent du trafic dans le réseau lorsqu'on veut identifier l'adresse IP d'un hôte contrairement à une mesure passive ou l'on ne fait qu'une requête pour trouver la localisation d'un hôte.

2.4.1 GeoPing

Le GeoPing se base sur l'hypothèse, que la distance géographique entre deux hôtes est mesurable en fonction du délai de transmission entre les deux. Bien sûr, elle utilise aussi un système de triangulation avec des hôtes qu'elle possède avec une localisation et des délais connus, grâce à ces hôtes, nous pouvons améliorer l'approximation de la localisation IP.

Malheureusement, l'hypothèse s'avère rarement correcte. Car la rapidité d'un réseau peu variée rapidement et donc modifier les délais. Néanmoins, comme énoncé précédemment les hôtes connus peuvent eux améliorer nos résultats et donc corriger ces petites erreurs.

[4] Liste des opérateurs de registrer agréer par l'ICANN, https://www.internic.net/origin.html
[5] Geo-location of the commune of an IP User, http://ieeexplore.ieee.org/document/4621634/

2.4.2 GeoTrack

Contrairement à la GeoPing cette technique n'émet pas d'hypothèse, mais va plutôt faire confiance aux administrateurs réseaux. Le logiciel va donc envoyer un paquet d'un point A à un point B. Il va pendant l'envoi récupérer le nom de tous les routeurs que le paquet va emprunter. Cependant, bien qu'il ait un certain nombre de personnes qui respect une certaine convention de nommage des routeurs, ce n'est pas une norme ou un standard donc personne n'a l'obligation. Ce qui fait que cette technique n'est pas optimale non plus.

2.4.3 Le projet TULIP – Trilateration Utility for Locating IP hosts

Le projet TULIP est un projet réalisé au Etats-Unis par la National University of Sciences and Techniology, School of Electrical Engineering and Computer Sciences, Standord Linear Accelerator Center et Internet End-to-end Performance Monitoring. Ce projet a pour but de géolocaliser un hôte spécifique grâce à son adresse IP.

Le principe du projet TULIP est d'utiliser le RTT pour calculer la distance de la cible. Pour ce faire, ils utilisent plusieurs points de repère (Landmark), ils en utilisent plusieurs pour pouvoir faire de la trilatération[6]. Cette technique qui utilise plus de ressource ce montre plus précise.

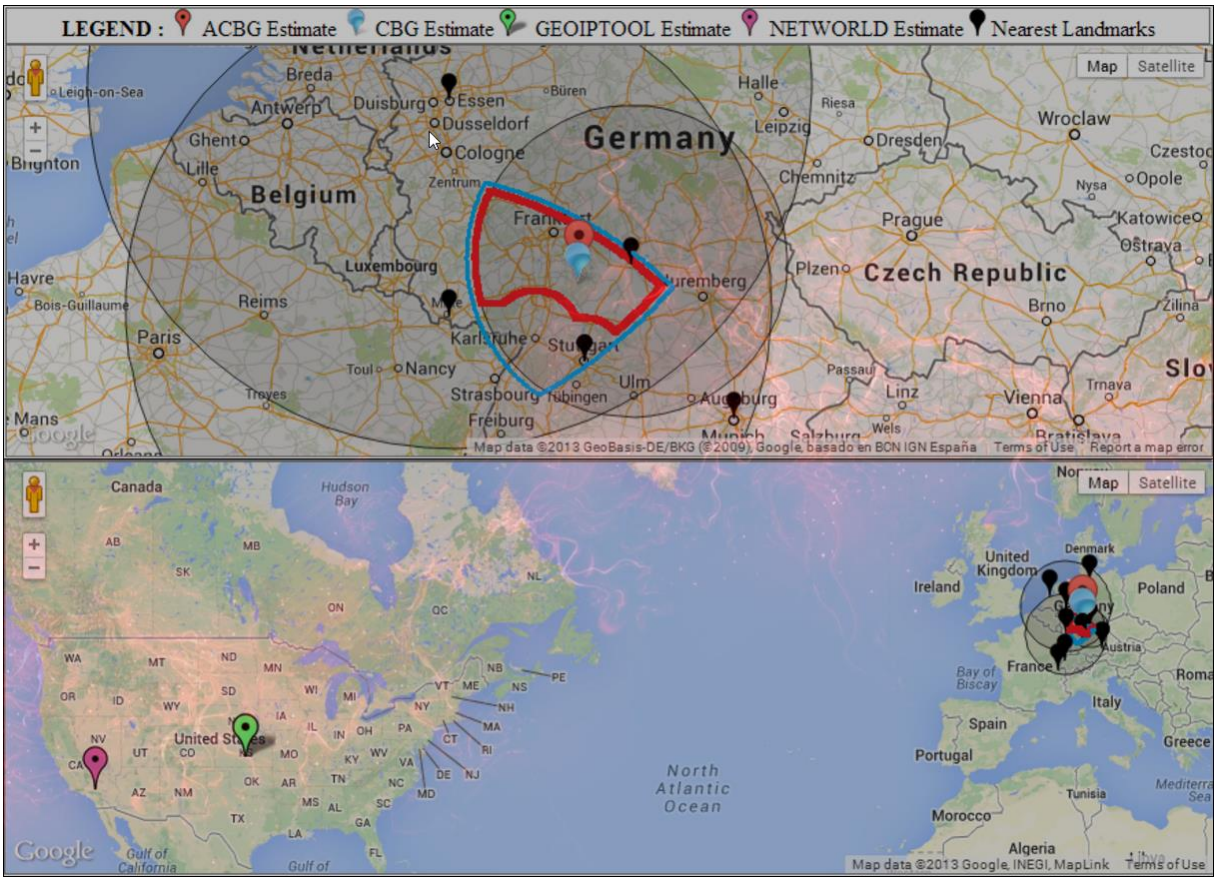


Figure 6 : Exemple fournit par le site TULIP

[6] Définition : « *La **trilatération** est une méthode mathématique permettant de déterminer la position relative d'un point en utilisant la géométrie des triangles tout comme la triangulation.* », https://fr.wikipedia.org/wiki/Trilat%C3%A9ration

Sur l'image ci-dessus, nous pouvons donc voir la différence que le projet TULIP fait. Il se base donc sur leur Landmark pour faire de la trilatération et on peut donc voir que la recherche est plus exacte qu'un logiciel de géolocalisation IP lambda.

Actuellement, le projet TULIP est utilisé pour la géolocalisation IP d'hôte, mais pas seulement. Il peut être utilisé pour vérifier le contenu des bases de données WHOIS et découvrir les hôtes qui utilisent des proxys.

2.5 Les principaux risques d'erreurs

Malheureusement, plusieurs paramètres rentrent en compte lorsque que nous cherchons à trouver la localisation de notre hôte. Une des plus commune est l'état de la base de données. Le maintien constant de la base de données est primordial, les informations de géolocalisation IP sont des informations qui évoluent constamment. Il faut donc régulièrement mettre à jour la base de données. Malheureusement, celles qui sont fournis gratuitement ne sont pas mise à jour très fréquemment, c'est pourquoi l'utilisation d'une base de données payante serait meilleure.

De plus, les informations utilisées par le service WHOIS ne sont pas toutes vérifiées, si nous regardons dans les informations données par l'ICANN plus exactement le RAA seul le nom, e-mail et numéro de téléphone sont vérifier le reste des informations sont au bon vouloir des possesseurs. Néanmoins, en 1996 la RFC 1876 a été proposée. Elle avait pour but de rajouter au DNS des informations sur la longitude la latitude et l'altitude, ce qui aurait activement améliorer la localisation d'hôte IP. Cependant, cette nouvelle RFC a été refusée.

Ensuite, l'utilisation de VPN peut aussi être un risque de plus. Malheureusement dans ce cas, nous allons trouver une localisation pour son adresse IP, mais elle sera erronée, car l'utilisation d'un VPN va créer un tunnel sécurisé qui va pourvoir à l'utilisateur de se faire passer pour un utilisateur d'un pays voulu. Tout comme le VPN, le proxy a pour but de garder son anonymat lorsque nous sommes sur internet. Par conséquent, il est lui aussi une des causes que nous ne trouvons pas la localisation de notre hôte.

2.6 Conclusions

En conclusion, aucune de ces techniques ne peut fonctionner correctement si elles travaillent chacune de leur côté. Les projets du genre TULIP utilisent la plupart de ces techniques en même temps pour qu'elles se corrigent entre elles et donc au final avoir des résultats corrects. L'entreprise MaxMind ne désire communiquer la façon d'ont-ils récupère leurs informations. Mais ils m'ont quand même dit que leur information venait deux types de sources une première qui est publique, ce qui veut donc dire des services « WHOIS ». Et un deuxième privé, je suppose que ce sont des informations achetées à dis site web qui peuvent fournir l'adresse IP et la localisation de leurs utilisateurs grâce à leurs formulaires.

Si nous prenons du recu et que nous nous mettons au niveau de la recherche IP destiné au serveur *Tdeig* du labo, ou simplement pour le site web lambda qui désire adapter la langue de son site en fonction de l'origine automatique, la précision actuellement des outils existant est suffisant. Nous pouvons avoir une base de données acceptable gratuitement ou s'en procurer payante qui va s'avérer plus précise, mais nous ne nécessitons que du pays.

Les projets comme TULIP sont actuellement les plus prometteurs dans ce secteur. Néanmoins, TULIP peut avoir des problèmes si ces repères ne sont pas actifs ou qu'il n'en possède pas assez dans certaines parties du monde.

Malheureusement, si nous voulions utiliser la géolocalisation pour la VoIP et localiser la personne qui appelle en urgence, nous ne pourrions pas risquer de se dire que cette personne utilise peut-être un VPN ou un proxy sur son téléphone ce qui rendrait donc la localisation fausse. Ce secteur manque de normes, ce qui fait que nous devons utiliser plusieurs méthodes à la fois pour trouver une simple localisation. La RFC 1876 qui aurait eu pour but de révéler la localisation des DNS aurait pu grandement améliorer les recherches. Pourquoi a-t-elle refusé, aucune idée, mais tant qu'une avancée au niveau des standards et normes ne sera pas faite la localisation IP ne sera sûr à 100%.

3 Mise en œuvre

3.1 Composants

3.1.1 Fichier *log*

Pour la partie pratique, nous avons besoin de contenu à analyser. Nous avons donc récupéré le fichier log du serveur *Tdeig*. L'utilisation de ces logs nous donne une vue réelle de ce qui se passe sur ce serveur. Ces logs datent de septembre 2014, ce qui nous donne un historique assez conséquent. Nous allons donc utiliser les logiciels choisis pour analyser ce fichier.

3.2 Choix des logiciels

Nous utilisons donc un serveur web nommé Nginx comme stipule dans notre énoncé. C'est un serveur de type asynchrone, il va donc modifier son état pour s'occuper des nouvelles connexions à la place de créer un processus par client qui se connecte. Le choix de GoAccess et Atats s'est fait sur plusieurs points. Tout d'abord contrairement à Webalizer, le logiciel actuellement utilisé sur *Tdeig*, ces deux logiciels sont toujours maintenus ce qui fait qu'ils reçoivent toujours des mises à jour. De plus, ces logiciels son gratuit, contrairement à « Deep Log Analyzer » qui est 300$ pour au final un produit pas si différent.

3.2.1 AWStats

AWStats est donc un analyseur de fichier log qui génère toute sorte de statistique. Ce qui va nous intéresse, c'est son option nommée de géolocalisation, qui va donc nous fournir des informations sur la localisation des utilisateurs. Il peut analyser une multitude de format log, il peut aussi nous permettre de gérer des logs personnalisés.

3.2.1.1 Géolocalisation AWStats

Le logiciel compare les adresses IP trouvé dans le fichier log avec la base de données qu'il possède. En ce qui concerne la base de données de la géolocalisation AWStats utilise la base de données « GeoIP Legacy » fourni par MaxMind qui est gratuite. Le logiciel ne fournit pas de mise par rapport à la base de données. Si nous désirions avoir une base de données à jour, ce qui est primordial pour ce genre de travail est de posséder la base de données la plus à jour possible.

3.2.1.2 Détection des robots

Tout d'abord, il faut faire la différence avec les deux types de robots. Il y a les robots d'indexation qui sont utilisés pars les moteurs de recherche tels que Google ou Bing pour simplement mieux indexer votre site web dans leurs moteurs de recherche, ceux-là ne sont pas vraiment nocifs pour votre serveur. Ensuite, il existe les *crawlers* qui se dit en français les robots malveillants, la plupart du temps, ils servent simplement à récupérer des adresses e-mail et le plus d'information possible soit pour pouvoir générer des e-mails spam ou simplement pour la vente d'information. Ils génèrent beaucoup de trafic quand les serveurs et il est donc préférable de les rejeter.

AWStats utilise un fichier nommé « robots.pm » qui contient les noms de *crawlers* déjà identifié. Il va donc comparer les personnes qui ont accédé au site web et va nous séparer les personnes humaines des robots.

Le problème avec cette technique c'est que le monde des robots malveillant évolue très rapidement, ils sont modifiés, supprimés et créés tous les jours avec des nouveaux noms. Ce qui veut dire que notre liste devient obsolète très rapidement. Il existe des solutions créées par des utilisateurs qui adaptent le fichier « robots.pm » pour qu'il puisse détecter et ajouter les nouveaux robots.[7]

3.2.2 GoAccess

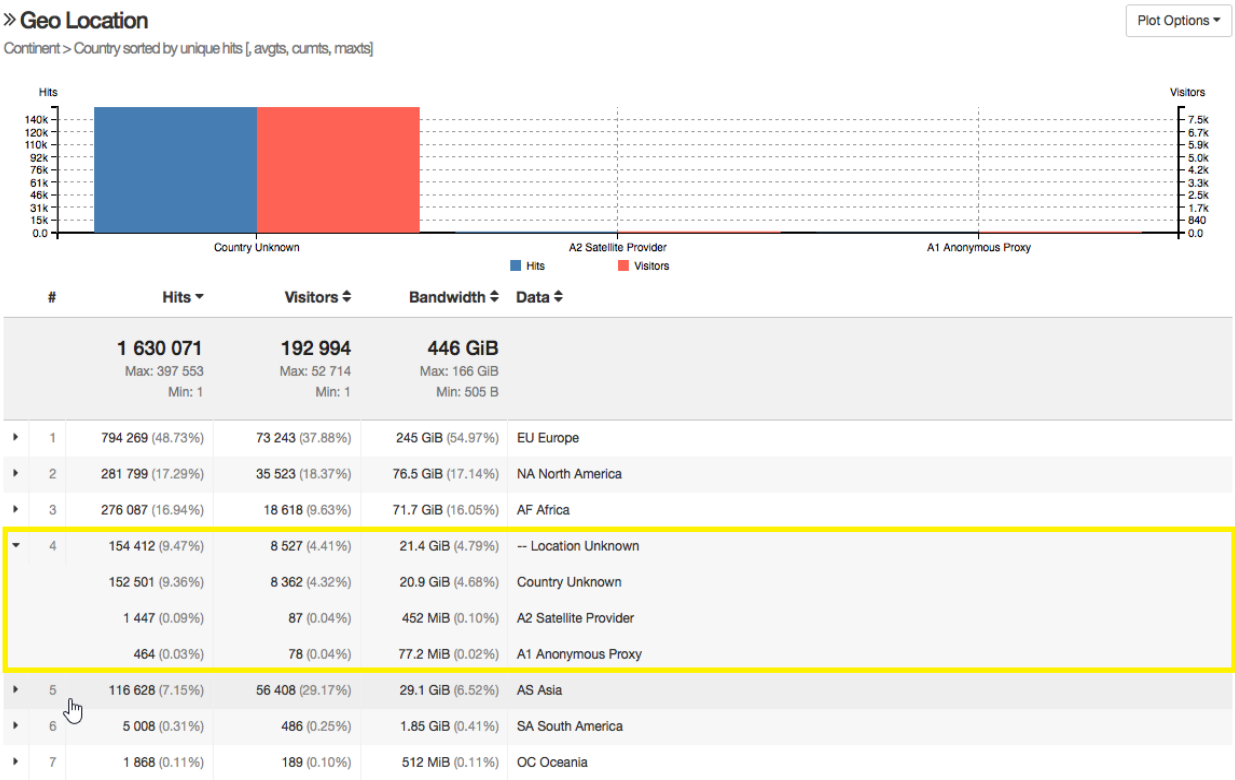
GoAccess nous permet d'accéder à toutes les informations de deux manières. La première, est comme vue précédemment grâce à une page html qui serait donc visible par nos visiteurs. La deuxième, ne génère aucun fichier html et génère une interface interactive sur le serveur directement. Dans notre cas, cela peut être très intéressant, car nous ne possédons pas d'interface sur notre serveur. Grâce à cela l'administrateur peut consulter les statistiques sans utiliser un deuxième ordinateur.

(Image en annexe)

[7] **Cornmaster (pseudo), About AWStats Robots,** http://wiki.cornempire.net/awstats/awstatsrobots (21/01/14)

3.2.2.1 Géolocalisation GoAccess

GoAccess n'est pas très différent de AWStats à ce niveau, il utilise lui aussi la base de données « GeoIP Legacy ». Si nous souhaitons fournir d'autre base de données GoAccess gère plusieurs bases à la fois.



#	Hits	Visitors	Bandwidth	Data
	1 630 071 Max: 397 553 Min: 1	192 994 Max: 52 714 Min: 1	446 GiB Max: 166 GiB Min: 505 B	
1	794 269 (48.73%)	73 243 (37.88%)	245 GiB (54.97%)	EU Europe
2	281 799 (17.29%)	35 523 (18.37%)	76.5 GiB (17.14%)	NA North America
3	276 087 (16.94%)	18 618 (9.63%)	71.7 GiB (16.05%)	AF Africa
4	154 412 (9.47%)	8 527 (4.41%)	21.4 GiB (4.79%)	-- Location Unknown
	152 501 (9.36%)	8 362 (4.32%)	20.9 GiB (4.68%)	Country Unknown
	1 447 (0.09%)	87 (0.04%)	452 MiB (0.10%)	A2 Satellite Provider
	464 (0.03%)	78 (0.04%)	77.2 MiB (0.02%)	A1 Anonymous Proxy
5	116 628 (7.15%)	56 408 (29.17%)	29.1 GiB (6.52%)	AS Asia
6	5 008 (0.31%)	486 (0.25%)	1.85 GiB (0.41%)	SA South America
7	1 868 (0.11%)	189 (0.10%)	512 MiB (0.11%)	OC Oceania

Figure 7 : Géolocalisation par GoAccess

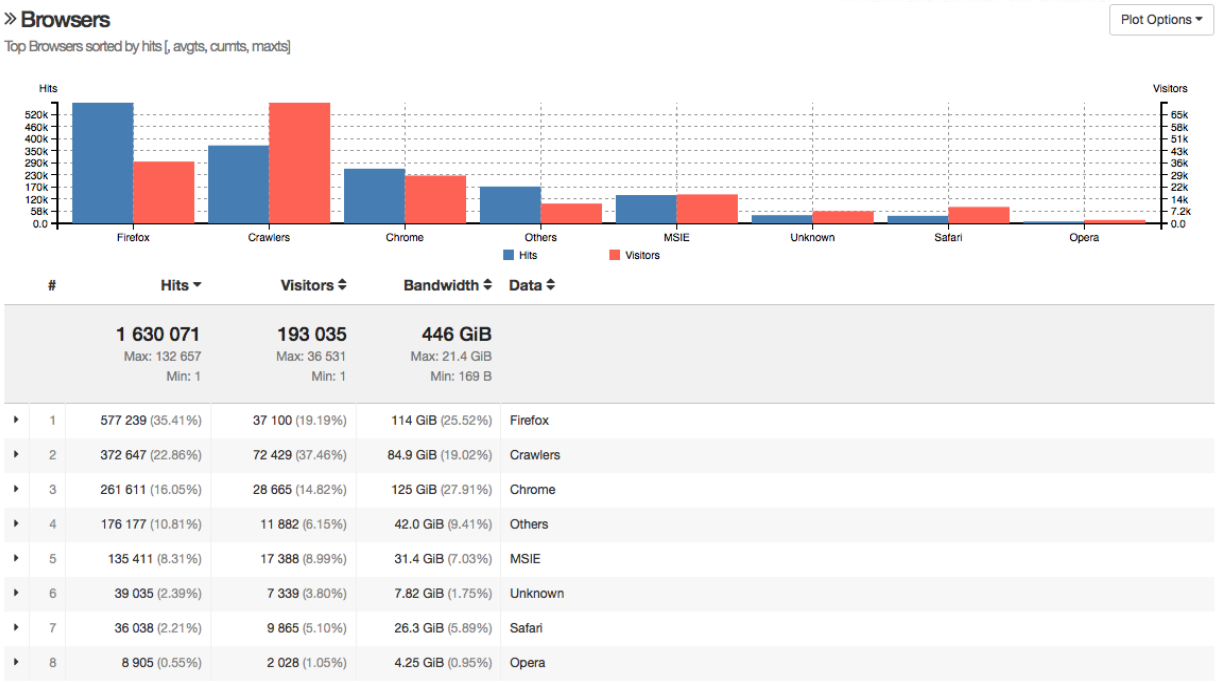
Ci-dessus, nous pouvons voir qu'il y a seulement 10% des requêtes qui ne sont pas résolues. Si nous observons la section « Location Unknown » on peut observer trois sections. La première est tout simplement les « Country Unknown » il est possible que ce problème aurait pu être évité avec une base de données plus complète. Ensuite nous avons la section « A2 Satellite Provider » qui sont des fournisseurs d'accès via satellite et les « A1 Anonymous Proxy » qui sont des utilisateurs cachés derrière des proxy pour ne pas révéler leur localisation.

Après un échange de courriel avec une personne qui travaille au service client de chez MaxMind, il m'a expliqué que les « A2 Satellite Provider » étaient souvent des utilisateurs à haut risque. Et qu'ils n'avaient pas de solutions pour trouver la localisation de ce type d'utilisateurs (A1 et A2).

3.2.2.2 Détection des robots

Ci-dessous, nous pouvons voir donc voir qu'elle type de « Browsers » a accédé au serveur, c'est aussi dans cette rubrique que les *crawlers* sont classé. Si nous regardons donc les chiffres sortis, nous voyons que 37,46% des visiteurs sont donc des *crawlers* et qu'ils génèrent un trafic très important, le deuxième le plus important. Mais malheureusement, nous ne pouvons pas assumer que c'est tout le trafic qu'ils génèrent, comme dit précédemment les fichiers que nous utilisons pour reconnaitre les *crawlers* ne sont seulement utile pour les *crawlers* déjà connu.

(La liste complet en annexe.)



#	Hits	Visitors	Bandwidth	Data
	1 630 071 Max: 132 657 Min: 1	193 035 Max: 36 531 Min: 1	446 GiB Max: 21.4 GiB Min: 169 B	
1	577 239 (35.41%)	37 100 (19.19%)	114 GiB (25.52%)	Firefox
2	372 647 (22.86%)	72 429 (37.46%)	84.9 GiB (19.02%)	Crawlers
3	261 611 (16.05%)	28 665 (14.82%)	125 GiB (27.91%)	Chrome
4	176 177 (10.81%)	11 882 (6.15%)	42.0 GiB (9.41%)	Others
5	135 411 (8.31%)	17 388 (8.99%)	31.4 GiB (7.03%)	MSIE
6	39 035 (2.39%)	7 339 (3.80%)	7.82 GiB (1.75%)	Unknown
7	36 038 (2.21%)	9 865 (5.10%)	26.3 GiB (5.89%)	Safari
8	8 905 (0.55%)	2 028 (1.05%)	4.25 GiB (0.95%)	Opera

Figure 8 : Robots Malveillants trouvés par GoAccess

Actuellement, GoAccess peut détecter les robots grâce à sa base de données qui possèdes un certain nombre de *crawlers* déjà identifié. Par conséquence, cela veut dire qu'il ne va pas de lui-même en découvrir, mais simplement nous les indiquer. Si nous le désirons, nous avons le choix de rajouter des bases de données en plus.

Ensuite, nous pouvons simplement exclure tous ces robots malveillants grâce au firewall de NGINX. Nous pouvons créer (où télécharger)[8] un fichier nommé « blacklist.conf » où nous ajouterons simplement l'adresses IP des robots malveillants que nous avons trouvé, si nous téléchargeons la version proposé il y aura quelque adresses IP en plus ce qui n'est pas mauvais. Ensuite, nous devons simplement ajouter ce fichier à la configuration de Nginx dans « nginx.conf ».

3.2.2.3 Autres options

Si nous le désirions GoAccess nous proposes trois types de la gestion de nos logs, « Default Hash Tables », « Tokyo Cabinet On-Disk B+ Tree » et « Tokoy Cabinet In-memory Hash DataBase ». La première solution est la plus rapide nous stockons simplement les logs dans notre machine et nous avons une vitesse de lecture de 87'816 lignes par seconde, La deuxième est la plus lente à 23'000 ligne seconde tous les logs sont stocké à l'extérieur de notre machine et finalement la dernière qui est un mélange des deux premières ou les données sont stocké à moitié dans la machine et le reste à l'extérieur. [9]

[8] **Hung Tran, Nginx Blacklist**, https://github.com/oohnoitz/nginx-blacklist (23.02.2014)
[9] Ces chiffres ont été donné par GoAcces et été testé avec un Intel® Core™ i7-4510U CPU @2.00GHz 8GB RAM, https://goaccess.io/faq (31/01/17)

3.2.3 Conclusion

En conclusion, le logiciel utilisé n'a pas une grande importance, car dans les deux cas, ils vont faire la même chose. Mise à part l'interface interactive de GoAccess. Ils vont simplement faire des requêtes sur les bases de données et nous donner leurs résultats. Ce qui est donc vraiment important, c'est la base de données que nous allons utiliser pour travailler.

De mon côté, je n'ai pas trouvé de logiciel avec une alternative à ce problème, il y a bien le projet TULIP qui montre des signes positifs dans la recherche active, mais cela reste une alternative couteuse et très gourmande en ressource. Et qui n'est pas implémenter pour les serveurs.

4 Déroulement du travail

4.1 Résumé

Pour se travaille, nous avions à disposition un ordinateur qui était donc le serveur où nous avons installé CentOS 7 minimal qui donc la version la plus légère et qui ne possède pas d'interface graphique. Nous avons prix Nginx 1.10.2, nous allons ensuite modifier la configuration pour avoir un serveur web utilisable. Quand nous avons installé notre machine, nous ajouterons GoAccess et copierons le fichier log du serveur *Tdeig* sur notre machine pour lui faire penser que ce sont nos logs.

4.2 Guide de sécurité NGINX

4.2.1 Activation de SELinux

SELinux nous permet de spécifier la gestion des contrôles d'accès dans notre architecture Linux. Pour ce faire il faut tout d'abord télécharger les paquets :

```
# yum install policycoreutils setroubleshoot
```

Ensuite nous allons mettre Selinx en mode permissif qui est le mode le plus agressif et de type « *targeted* » qui veut dire que nous sécurisons les dossiers visés :

```
# nano /etc/selinux/config
```

Ligne à modifier :

```
SELINUX=enforcing
SELINUX=targeted
```

Il faudra ensuite redémarrer. Et l'installation est terminé.

4.2.2 Vérification d'attaque « *Buffer Overflow* »

Une attaque de type « Buffer Overflow » peut être fatale pour une entreprise qui utilise son serveur web constamment, c'est pour cela qui faut nous défendre contre ce type d'attaque. Pour se faire nous pouvons diminuer le nombre de client simultané et géré le temps de connexion des clients inactifs. Fichier de configuration :

```
# nano /usr/local/nginx/conf/nginx.conf
```

Il faut ensuite modifier les données suivantes :

```
client_body_buffer_size  1k;
client_header_buffer_size 1k;
client_max_body_size 1k;
large_client_header_buffers 2 1k;
client_body_timeout   10;
client_header_timeout 10;
keepalive_timeout     5 5;
send_timeout       10;
```

De plus nous pouvons aussi contrôler le nombre de connexion simultané depuis la même adresse :

```
limit_zone slimits $binary_remote_addr 5m;
limit_conn slimits 5;
```

Le choix du nombre de connexion simultané et de temps de connexion a été selon https://www.cyberciti.biz/tips/linux-unix-bsd-nginx-webserver-security.html.

4.2.3 Restreindre l'accès des dossiers

4.2.3.1 Sécurisé par mot de passe

Pour améliorer la sécurité de certain contenu que l'on souhaite être accessible seulement par seulement une partie de nos utilisateurs nous devons créer le fichier des mots de passe et un utilisateur :

```
# mkdir /usr/local/nginx/conf/.htpasswd
# htpasswd -c /usr/local/nginx/conf/.htpasswdpasswd vivek
```

Maintenant nous allons modifier le fichier nginx.conf

```
Location ~/(dossier_à_protéger){
       auth_basic « Restricted » ;
       auth_basic_user_file  /usr/local/nginx/conf/.htpasswd/passwd ;
}
```

4.2.4 Bloquer certains utilisateurs nuisant (robot)

Pour ce faire nous allons éditer le fichier de configuration « nginx.conf » :

```
http   {
       …
       include /etc/nginx/blacklist.conf
}

server {
       …
       if($bad_agent){
              return 403 ;
       }
}
```

Quand ceci est fait nous devons récupérer le fichier « blacklist.conf ». Ce fichier contient tous les noms de robots déjà connu, nous n'avons plus cas le placer dans le dossier « nginx »[10].

[10] **Hung Tran, Nginx-Blacklist**, https://github.com/oohnoitz/nginx-blacklist (23/02/14)

4.3 Guide technique

4.3.1 CentOS minimal

4.3.1.1 Installation et configuration de base

Lors du début de l'installation, une interface graphique est présente. Elle n'est présente que pour l'installation, après avoir redémarré pour la première fois cette interface disparaît.

Figure 9 : Premier redémarrage

Au démarrage, le système d'exploitation n'est pas du tout configuré, la version minimale ne possède pas d'outils tels que ifconfig ou nano ne sont pas présent.

Nous commençons par vérifier nos interfaces réseaux. La commande suivante nous permet de le faire :

```
# nmcli d
```

```
localhost login: root
Password:
[root@localhost ~]# nmcli d
PÉRIPHÉRIQUE  TYPE      ÉTAT         CONNEXION
eno16777736   ethernet  déconnecté   --
lo            loopback  non-géré     --
[root@localhost ~]# _
```

Figure 10 : Etat des périphérique réseaux

De base, la configuration du réseau est en DHCP si on désire modifier cela, nous pouvons simplement aller dans le fichier de configuration.

```
# vi /etc/sysconfig/network-scripts/ifcfg-***
```

**** Les étoiles doivent être remplacé par le nom du périphérique réseau*

Il faut aussi activer l'interface réseau pour connecter notre machin au réseau. Pour cela, on exécute la commande suivante :

```
# ifup eno***
```

Les étoiles doivent être remplacé par le nom du périphérique réseau au peut le voir sur la figure 2

4.3.1.2 Best practice

Tout d'abord modifié le hostname de notre serveur, pour vérifier le nom actuel :[11]

```
# echo $HOSTNAME
```

Ensuite pour le modifier, il faut aller changer le nom dans le fichier « *hostname* » et écrire le nom que nous désirons mettre, dans notre cas tdeigbis :

```
# nano /etc/hostname
# reboot
```

Si nous voulons améliorer la sécurité de notre serveur, il existe un firewall basic qui est conseillé d'installer dessus. Pour cela nous installer le paquet `firewalld`, ce firewall est pratique, car nous pouvons le configurer sans perdre la connexion. Après l'avoir installé le paquet, il faut simplement l'activer comme ceci :

```
# sudo systemctl start firewalld
```

Maintenant que notre firewall est lancé, nous activons les paramètres que nous désirons. Dans notre cas, nous désirons faire un serveur http donc nous devons activer le service http :

```
# sudo firewall-cmd --permanent --add-service=http
```

Quand toute notre configuration est terminée, il faut redémarrer le firewall typé simplement :

```
# sudo firewall-cmd –reload
```

Pour que notre firewall soit démarré automatiquement au démarrage, il faut :

```
# sudo systemctl enable firewalld
```

Si vous désirez voir la configuration de votre firewall :

```
# sudo firewall-cmd --permanent --list -all
```

L'horloge sur un serveur est très importante, c'est grâce à cela que si nous avons un problème et que nous avons plusieurs alertes qui surviennent sur plusieurs de notre serveur ou équipement informatique, nous pouvons les comparer. C'est pourquoi il nous faut synchroniser notre horloge avec celle d'un serveur référence[12].

[11] **Aucun nom, 30 Things to do After Minimal RHEL/CentOS 7 Installation,** http://www.tecmint.com/things-to-do-after-minimal-rhel-centos-7-installation/ (30.06.2015)

[12] **Justin Ellingwood**, ***Additional Recommended Steps fo New CentOS 7 Servers***, https://www.digitalocean.com/community/tutorials/additional-recommended-steps-for-new-centos-7-servers (05.10.2014)

Nous allons tout d'abord nous mettre sur le bon fuseau horaire, dans notre cas Zurich pour ce faire, il faut rentrer la commande suivante :

```
# timedatectl set-timezone Europe/Zurich
```

Ensuite pour le protocole NTP (protocole d'heure réseau), il nous permettre donc de posséder une heure en tout temps correct ce qui est primordial pour un serveur et surtout dans notre cas si nous désirons utiliser les logs que la machine va générer. Il faut donc installer le paquet suivant et lancé la synchronisation :

```
# yum install ntp
# systemctl start ntpd
# systemctl enable ntpd
```

Si vous voulez vérifier que tout s'est bien passé, il suffit de typer :

```
# timedatectl
```

4.3.2 Nginx Web Server

Nous utilisons Nginx comme demandé dans l'énoncé pour l'installer, nous allons télécharger les deux paquets suivant :

```
# yum install epel-release
# yum install nginx
```

Nous allons maintenant démarrer notre serveur :

```
# systemctl start nginx
```

Et la touche finale pour que le serveur démarre automatiquement :

```
# sudo systemctl enable nginx
```

4.3.2.1 Aide

Pour vérifier si vous configuration Nginx est correcte, vous pouvez simplement lancer la commande suivante :

```
# /sbin/nginx -t
```

4.3.3 GoAccess

Tout d'abord le télécharger[13] :

```
 # yum install goaccess
```

Nous allons maintenant configurer le fichier de configuration, nous devons choisir le type de log, nous allons analyser dans notre cas ça sera des logs Apache, car les logs utilisés sont ceux du serveur *Tdeig* actuel qui utilise Apache. Pour ce faire :

```
# nano /etc/goaccess.conf
```

Une fois dans ce fichier, il faut simplement chercher le paragraphe Apache et enlever les commentaires sous les lignes suivantes :

[13] **Julien Hommet, GoAccess – des logs web en temps réel,** https://computerz.solutions/goaccess-logs-web-realtime/ (01.10.2015)

```
time-format %H:%M:%S
date-format %d%b%Y
log-format %h %^[%d:%t %^] ‘’%r’’ %s%b ‘’%R’’ ‘’%u’’

log-file /var/log/apache/access.log
```

La dernière ligne, nous indique ou nos logs seront stocker

Maintenant, il ne reste plus qu'à générer le fichier html qui nous permet de voir nos statiques :

```
# goaccess -f /var/log/apache/access.log -a > /usr/share/nginx/html/report.html
```

4.3.4 Autres Installations nécessaires

Installer nano pour éditer les fichiers, c'est un éditeur un peu plus pratique que vi.

```
# yum install nano
```

Net-tools peut être intéressant à installer, il possède plein de petits outils tel que ifconfig, qui ne sont pas compris dans CentOS 7 minimal

```
# yum install net-tools
```

4.3.5 Utilisations

Nous avons donc deux modes d'affichage sur GoAccess le premier qui est donc de générer une page HTML ou générer l'interface sans GUI. Ci-dessous la version avec une page HTML :

```
# goaccess -f <fichierLog> -a -o report.html
```

Et cette fois-ci la version sans GUI :

```
# goaccess -f <fichierLog>
```

Nous pouvons si nous le désirions changer de bases données. Lorsque nous changeons de bases c'est souvent parce que nous avons pris le choix d'en payer une payante. Pour ajouter notre base de donner il nous faut donc ajouter le fichier lors de la génération des statistiques :

```
# --geoip-database <fichierGeo>
```

5 Comparaison

5.1 GoAccess Vs. Webalizer

Pour cette comparaison j'ai donc décidé de prendre un mois et de lancer une analyse avec Webalizer puis de rependre ce même mois et de lancer l'analyser avec GoAccess.

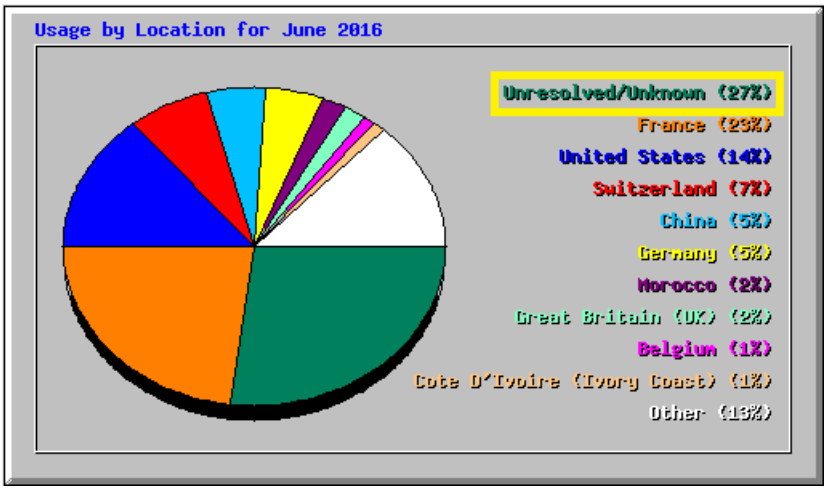
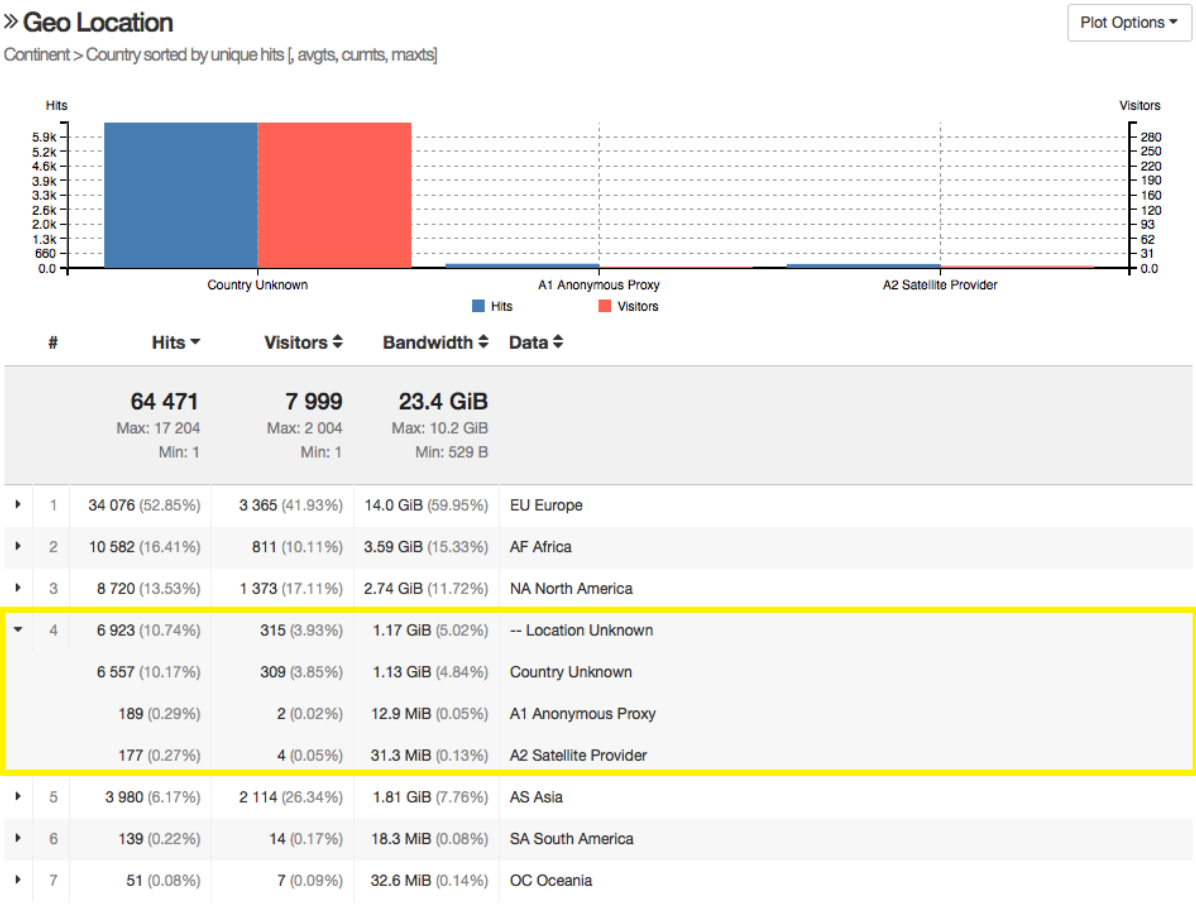


Figure 11 : Webalizer analyse du mois de juin



#	Hits	Visitors	Bandwidth	Data
	64 471 Max: 17 204 Min: 1	7 999 Max: 2 004 Min: 1	23.4 GiB Max: 10.2 GiB Min: 529 B	
1	34 076 (52.85%)	3 365 (41.93%)	14.0 GiB (59.95%)	EU Europe
2	10 582 (16.41%)	811 (10.11%)	3.59 GiB (15.33%)	AF Africa
3	8 720 (13.53%)	1 373 (17.11%)	2.74 GiB (11.72%)	NA North America
4	6 923 (10.74%)	315 (3.93%)	1.17 GiB (5.02%)	-- Location Unknown
	6 557 (10.17%)	309 (3.85%)	1.13 GiB (4.84%)	Country Unknown
	189 (0.29%)	2 (0.02%)	12.9 MiB (0.05%)	A1 Anonymous Proxy
	177 (0.27%)	4 (0.05%)	31.3 MiB (0.13%)	A2 Satellite Provider
5	3 980 (6.17%)	2 114 (26.34%)	1.81 GiB (7.76%)	AS Asia
6	139 (0.22%)	14 (0.17%)	18.3 MiB (0.08%)	SA South America
7	51 (0.08%)	7 (0.09%)	32.6 MiB (0.14%)	OC Oceania

Figure 12 : GoAccess analyse du mois de juin

Ainsi, nous voyons directement que le nombre de localisation non-résolue est passé de 27% à 10%, une amélioration de 17%. Ce que nous pouvons remarquer en nous focalisant sur une pays tel que la Suisse, c'est que nous sommes passés de 6,52% pour Webalizer à 8,41%(les listes plus complètes se trouvent en annexe). Ce sont donc toute ces adresses que nous ne trouvions pas avant qui ont été réparti parmi les pays que nous avions déjà.

Malheureusement dans ces statistiques il y a toujours ces robots malveillants qui faussent donc les données. Par exemple, les robots malveillants ont généré 12'238 visites sois 19% des visites, nous pouvons donc penser qu'une partie des 10% d'adresse non-résolus font partie de ces 10% et que certain des robots ont été localisé tel que « *Baiduspider/2.0* » qui est de Chine et qui a lui tout seul à générer 4% de visite ce qui peut expliquer les 6,17% en provenance de l'Asie. Au contraire, il n'y a pas eu beaucoup d'accès par proxy ou satellite, mais il faut assumer que dans les « Country Unknow » il y a des proxys qui ne sont pas connu par GoAccess.

En bref, GoAccess apporte de grosses améliorations en ce qui concerne la localisation des hôtes IP. Et permet d'identifier toutes les sources externes qui peuvent fausser nos statistiques tel que les robots malveillants, ce qui faudrait faire est de restreindre l'accès à ces robots et nous nous rapprocherons encore une fois d'un résultat correct.

6 Problèmes rencontrés

Pour ce travail, j'ai rencontré différents types de problèmes. Voici les problèmes que je pense nécessaire d'expliquer :

- Lors de l'installation de CentOS 7 sur mon serveur. J'ai utilisé un utilitaire pour créer une clé bootable avec CentOS 7. Malheureusement, cette application UnetBootin ne fonctionne pas avec CentOS 7, elle fait une clé défectueuse. Je n'ai remarqué que plus tard que le site www.wiki.centos.org expliquait que UnetBootin ne fonctionnerait pas. Grâce à ce site, j'ai pu après trouver un nouvel utilitaire du nom de RUFUS qui cette fois fonctionnait parfaitement.

- Lors de l'installation d'AWStats j'ai remarqué que la génération de ces statistiques sont fait sur des pages PHP ce qui veut dire que dans notre cas ce n'est pas utilisable, car notre machine ne possède pas de serveurs HTTP. J'ai quand même pris la décision d'expliquer comment installer AWStats et PHP sur ma partie pratique.

7 Conclusion

Nous avons donc pu voir dans ce travail les différents systèmes de géolocalisation IP existant et leur fonctionnement. Dans le cadre de ce projet, nous avons donc pu voir que des logiciels tels que GoAccess ou AWStats étaient suffisant pour une implémentation non commerciale et avec une précision n'allant pas plus loin que le pays. Mais nous voyons qu'à partir du moment où nous cherchons une précision allant de la ville à la rue, c'est un secteur qui est encore au début de son évolution, il a beaucoup d'amélioration possible. Il y a toute sorte de projet différent qui travaille sur son amélioration. Le projet TULIP qui pourrait être un des plus prometteur dans le secteur.

Mais la solution au problème de la précision de la géolocalisation ne serait-elle pas faisable au niveau des normes ? Comme nous avons pu le constater aucune norme n'oblige à aucun moment de stipuler la localisation de serveur ou de machine. Ne faudrait-il donc pas simplement ajouter une RFC telle que la RFC 1876, qui pour rappel n'a pas été accepté et voulait obliger l'ajout des coordonnées géographiques des serveurs DNS.

Tant que ces changements ne seront pas faits la géolocalisation IP ne sera précise et nous ne pourrons pas garantir que les appels VoIP d'urgence seront correctement redirigés vers le central le plus proche, lors d'une simple utilisation d'un proxy ou d'un VPN sur notre téléphone.

8 Table des matières des figures

9 Annexe

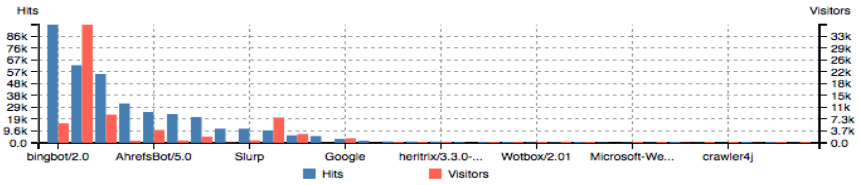
9.1 Interface GoAccess sans GUI

9.2 Liste de robots malveillants complète par GoAccess

» Browsers

Plot Options

Top Browsers sorted by hits [, avgts, cumts, maxts]



#	Hits	Visitors	Bandwidth	Data
	1 630 071 Max: 132 657 Min: 1	**193 035** Max: 36 531 Min: 1	**446 GiB** Max: 21.4 GiB Min: 169 B	
1	577 239 (35.41%)	37 100 (19.19%)	114 GiB (25.52%)	Firefox
2	372 647 (22.86%)	72 429 (37.46%)	84.9 GiB (19.02%)	Crawlers
	95 529 (5.86%)	6 045 (3.13%)	13.9 GiB (3.11%)	bingbot/2.0
	62 746 (3.85%)	36 531 (18.89%)	17.7 GiB (3.96%)	Baiduspider/2.0
	55 688 (3.42%)	8 729 (4.51%)	21.4 GiB (4.79%)	Googlebot/2.1
	31 828 (1.95%)	674 (0.35%)	309 MiB (0.07%)	DotBot/1.1
	24 946 (1.53%)	3 981 (2.06%)	240 MiB (0.05%)	AhrefsBot/5.0
	23 289 (1.43%)	740 (0.38%)	112 MiB (0.02%)	MJ12bot/v1.4.5
	20 862 (1.28%)	1 915 (0.99%)	15.3 GiB (3.42%)	YandexBot/3.0
	11 509 (0.71%)	96 (0.05%)	61.2 MiB (0.01%)	SEOkicks-Robot
	11 454 (0.70%)	768 (0.40%)	4.69 GiB (1.05%)	Slurp
	10 075 (0.62%)	7 840 (4.05%)	59.2 MiB (0.01%)	AhrefsBot/5.1
	6 056 (0.37%)	2 749 (1.42%)	9.82 GiB (2.20%)	Googlebot-Mobile/2.1
	5 477 (0.34%)	39 (0.02%)	204 MiB (0.04%)	TurnitinBot/3.0
	3 200 (0.20%)	1 413 (0.73%)	972 MiB (0.21%)	Google
	1 783 (0.11%)	7 (0.00%)	3.04 MiB (0.00%)	BacklinkCrawler
	1 134 (0.07%)	14 (0.01%)	13.1 MiB (0.00%)	netEstate
	1 035 (0.06%)	1 (0.00%)	10.3 MiB (0.00%)	heritrix/3.1.1
	1 029 (0.06%)	1 (0.00%)	9.08 MiB (0.00%)	heritrix/3.3.0-SNAPSHOT-20160309-0050
	727 (0.04%)	5 (0.00%)	5.02 MiB (0.00%)	SISTRIX
	689 (0.04%)	1 (0.00%)	17.3 MiB (0.00%)	Java/1.8.0_51
	571 (0.04%)	233 (0.12%)	5.93 MiB (0.00%)	Googlebot-Image/1.0
	441 (0.03%)	9 (0.00%)	1.24 MiB (0.00%)	Wotbox/2.01
	412 (0.03%)	369 (0.19%)	85.4 MiB (0.02%)	facebookexternalhit/1.1
	350 (0.02%)	24 (0.01%)	3.45 MiB (0.00%)	Python-urllib/2.7
	343 (0.02%)	4 (0.00%)	19.7 MiB (0.00%)	Java/1.6.0_04
	281 (0.02%)	32 (0.02%)	151 KiB (0.00%)	Microsoft-WebDAV-MiniRedir/6.1.7601
	219 (0.01%)	1 (0.00%)	11.3 MiB (0.00%)	Java/1.8.0_25
	209 (0.01%)	91 (0.05%)	25.7 MiB (0.01%)	Sogou
	196 (0.01%)	1 (0.00%)	8.87 MiB (0.00%)	Java/1.7.0_75
	178 (0.01%)	1 (0.00%)	31.6 MiB (0.01%)	crawler4j
	133 (0.01%)	54 (0.03%)	580 KiB (0.00%)	YisouSpider
	132 (0.01%)	58 (0.03%)	84.5 MiB (0.02%)	ia_archiver
	126 (0.01%)	3 (0.00%)	350 KiB (0.00%)	libwww-perl/5.833
3	261 611 (16.05%)	28 665 (14.82%)	125 GiB (27.91%)	Chrome
4	176 177 (10.81%)	11 882 (6.15%)	42.0 GiB (9.41%)	Others
5	135 411 (8.31%)	17 388 (8.99%)	31.4 GiB (7.03%)	MSIE
6	39 035 (2.39%)	7 339 (3.80%)	7.82 GiB (1.75%)	Unknown
7	36 038 (2.21%)	9 865 (5.10%)	26.3 GiB (5.89%)	Safari
8	8 905 (0.55%)	2 028 (1.05%)	4.25 GiB (0.95%)	Opera

9.3 Liste de géolocalisation GoAccess du mois de Juin

34 076 (52.85%)	3 365 (41.93%)	14.0 GiB (59.95%)	EU Europe
17 204 (26.68%)	1 194 (14.88%)	10.2 GiB (43.51%)	FR France
5 421 (8.41%)	370 (4.61%)	2.06 GiB (8.83%)	CH Switzerland
2 348 (3.64%)	222 (2.77%)	427 MiB (1.78%)	DE Germany
2 018 (3.13%)	16 (0.20%)	5.28 MiB (0.02%)	NL Netherlands
1 607 (2.49%)	9 (0.11%)	23.5 MiB (0.10%)	AT Austria
1 169 (1.81%)	393 (4.90%)	55.1 MiB (0.23%)	GB United Kingdom
905 (1.40%)	56 (0.70%)	188 MiB (0.78%)	BE Belgium
704 (1.09%)	223 (2.78%)	34.4 MiB (0.14%)	UA Ukraine
699 (1.08%)	514 (6.40%)	65.8 MiB (0.27%)	IT Italy
687 (1.07%)	185 (2.31%)	442 MiB (1.85%)	RU Russian Federation
385 (0.60%)	53 (0.66%)	327 MiB (1.37%)	PL Poland
334 (0.52%)	12 (0.15%)	40.4 MiB (0.17%)	ES Spain
123 (0.19%)	33 (0.41%)	111 MiB (0.46%)	EU Europe
103 (0.16%)	5 (0.06%)	4.09 MiB (0.02%)	SE Sweden
81 (0.13%)	7 (0.09%)	13.5 MiB (0.06%)	LU Luxembourg
75 (0.12%)	13 (0.16%)	14.3 MiB (0.06%)	CZ Czech Republic
42 (0.07%)	1 (0.01%)	2.64 MiB (0.01%)	IS Iceland
42 (0.07%)	6 (0.07%)	7.85 MiB (0.03%)	NO Norway
29 (0.04%)	13 (0.16%)	31.1 MiB (0.13%)	RO Romania
21 (0.03%)	10 (0.12%)	78.0 KiB (0.00%)	MD Moldova, Republic of
13 (0.02%)	3 (0.04%)	7.69 MiB (0.03%)	RS Serbia
12 (0.02%)	9 (0.11%)	10.7 MiB (0.04%)	IE Ireland
11 (0.02%)	1 (0.01%)	7.98 KiB (0.00%)	TR Turkey
7 (0.01%)	1 (0.01%)	381 KiB (0.00%)	GR Greece
6 (0.01%)	2 (0.02%)	5.85 MiB (0.02%)	PT Portugal
5 (0.01%)	1 (0.01%)	670 KiB (0.00%)	SK Slovakia
5 (0.01%)	2 (0.02%)	1.80 MiB (0.01%)	BG Bulgaria
4 (0.01%)	3 (0.04%)	1.95 MiB (0.01%)	HU Hungary
3 (0.00%)	1 (0.01%)	1.15 MiB (0.00%)	MC Monaco
2 (0.00%)	0 (0.00%)	1.01e+3 B (0.00%)	HR Croatia
2 (0.00%)	1 (0.01%)	7.54 KiB (0.00%)	AL Albania
2 (0.00%)	2 (0.02%)	70.1 KiB (0.00%)	MK Macedonia
2 (0.00%)	1 (0.01%)	13.7 KiB (0.00%)	BY Belarus
2 (0.00%)	1 (0.01%)	13.7 KiB (0.00%)	LV Latvia
2 (0.00%)	1 (0.01%)	3.37 KiB (0.00%)	LT Lithuania
1 (0.00%)	1 (0.01%)	6.72 KiB (0.00%)	EE Estonia
10 582 (16.41%)	811 (10.11%)	3.59 GiB (15.33%)	AF Africa
2 013 (3.12%)	182 (2.27%)	841 MiB (3.51%)	MA Morocco
890 (1.38%)	55 (0.69%)	322 MiB (1.34%)	CI Cote D'Ivoire
872 (1.35%)	77 (0.96%)	239 MiB (1.00%)	SN Senegal
852 (1.32%)	83 (1.03%)	533 MiB (2.23%)	DZ Algeria
828 (1.28%)	35 (0.44%)	154 MiB (0.64%)	TG Togo
717 (1.11%)	51 (0.64%)	171 MiB (0.71%)	BJ Benin
611 (0.95%)	51 (0.64%)	224 MiB (0.93%)	CM Cameroon
566 (0.88%)	30 (0.37%)	119 MiB (0.50%)	MG Madagascar
552 (0.86%)	85 (1.06%)	309 MiB (1.29%)	TN Tunisia
476 (0.74%)	7 (0.09%)	38.8 MiB (0.16%)	GN Guinea
344 (0.53%)	14 (0.17%)	272 MiB (1.14%)	BF Burkina Faso
336 (0.52%)	7 (0.09%)	56.9 MiB (0.24%)	RE Reunion
300 (0.47%)	10 (0.12%)	54.0 MiB (0.23%)	ML Mali
255 (0.40%)	29 (0.36%)	86.8 MiB (0.36%)	CD Congo, The Democratic Republic of the
234 (0.36%)	38 (0.47%)	71.0 MiB (0.30%)	ZA South Africa
233 (0.36%)	4 (0.05%)	32.1 MiB (0.13%)	MU Mauritius
180 (0.28%)	18 (0.22%)	49.7 MiB (0.21%)	CG Congo
86 (0.13%)	4 (0.05%)	5.34 MiB (0.02%)	CF Central African Republic

9.4 Liste de géolocalisation Webalizer du mois de juin

Top 30 of 86 Total Locations

#	Hits		Files		kB F		kB In		kB Out		Location
1	17611	27.32%	13156	54.11%	4865021	19.83%	0	0.00%	0	0.00%	Unresolved/Unknown
2	14704	22.81%	13302	54.71%	10009617	40.81%	0	0.00%	0	0.00%	France
3	9139	14.18%	6859	28.21%	2748696	11.21%	0	0.00%	0	0.00%	United States
4	4203	6.52%	3434	14.12%	699952	2.85%	0	0.00%	0	0.00%	Switzerland
5	3404	5.28%	2905	11.95%	1743210	7.11%	0	0.00%	0	0.00%	China
6	3002	4.66%	2808	11.55%	812796	3.31%	0	0.00%	0	0.00%	Germany
7	1498	2.32%	1372	5.64%	519747	2.12%	0	0.00%	0	0.00%	Morocco
8	1157	1.79%	1108	4.56%	61289	0.25%	0	0.00%	0	0.00%	Great Britain (UK)
9	930	1.44%	759	3.12%	197642	0.81%	0	0.00%	0	0.00%	Belgium
10	846	1.31%	804	3.31%	309306	1.26%	0	0.00%	0	0.00%	Cote D'Ivoire (Ivory Coast)
11	800	1.24%	723	2.97%	517547	2.11%	0	0.00%	0	0.00%	Algeria
12	699	1.08%	675	2.78%	67271	0.27%	0	0.00%	0	0.00%	Italy
13	686	1.06%	663	2.73%	117740	0.48%	0	0.00%	0	0.00%	Togo
14	647	1.00%	453	1.86%	138164	0.56%	0	0.00%	0	0.00%	Russian Federation
15	588	0.91%	567	2.33%	225149	0.92%	0	0.00%	0	0.00%	Cameroon
16	524	0.81%	486	2.00%	101170	0.41%	0	0.00%	0	0.00%	Madagascar
17	448	0.69%	446	1.83%	23532	0.10%	0	0.00%	0	0.00%	Guinea
18	403	0.63%	367	1.51%	102812	0.42%	0	0.00%	0	0.00%	Canada
19	383	0.59%	302	1.24%	335228	1.37%	0	0.00%	0	0.00%	Poland
20	333	0.52%	321	1.32%	277197	1.13%	0	0.00%	0	0.00%	Burkina Faso
21	260	0.40%	244	1.00%	18021	0.07%	0	0.00%	0	0.00%	Netherlands
22	226	0.35%	208	0.86%	72628	0.30%	0	0.00%	0	0.00%	South Africa
23	186	0.29%	179	0.74%	79784	0.33%	0	0.00%	0	0.00%	Benin
24	152	0.24%	142	0.58%	33875	0.14%	0	0.00%	0	0.00%	Mali
25	146	0.23%	81	0.33%	126262	0.51%	0	0.00%	0	0.00%	Tunisia
26	133	0.21%	131	0.54%	19390	0.08%	0	0.00%	0	0.00%	Senegal
27	125	0.19%	123	0.51%	16431	0.07%	0	0.00%	0	0.00%	Reunion
28	98	0.15%	92	0.38%	3986	0.02%	0	0.00%	0	0.00%	Sweden
29	86	0.13%	84	0.35%	5467	0.02%	0	0.00%	0	0.00%	Central African Republic
30	81	0.13%	76	0.31%	13793	0.06%	0	0.00%	0	0.00%	Luxembourg