



Introduction à AWSTATS

Olivier FONTÈS et Edouard PUTZ

Eitic

Table des matières

1.Introduction :.....	3
Qu'est-ce qu'AWSTATS :.....	3
Description sommaire des fonctionnalités :.....	3
2.Création d'un jeu de statistiques depuis alternC :.....	3
Que nous faut-il :.....	3
Création du/des compte/s :.....	3
Création du jeu de statistiques :.....	4
3.Avant d'aller plus loin :.....	5
4.Description des rubriques d'AWSTATS :.....	6
Vue générale :.....	6
Barre d'en-tête :.....	7
Résumé :.....	7
Quand.....	8
Historique mensuel:.....	8
Statistiques par jour du mois :.....	9
Statistiques moyenne par jour de la semaine :.....	10
Statistiques moyenne par heure de la journée :.....	11
Qui :.....	11
Classement des visiteurs par domaine et par pays :.....	11
Classement des visiteurs par nom de machine ou par ip :.....	12
Visites de robots/spiders :.....	12
Navigation :.....	12
Durée des visites :.....	12
Types de fichiers téléchargés :.....	13
Classement des pages les plus téléchargées :.....	13
Système d'exploitation des visiteurs :.....	13
Navigateurs des visiteurs :.....	13
Origine/Referer:.....	14
Connexion par origine :.....	14
Phrase clés et mots clés :.....	14
Autres :.....	15
Divers :.....	15
Codes status HTTP :.....	15
5.Conclusion.....	15

1.Introduction :

Qu'est-ce qu'AWSTATS :

AWSTATS est un outil d'analyse de fichiers journaux du serveur web apache : il permet une analyse complète des connexions à un site web, de façon graphique et conviviale.

Description sommaire des fonctionnalités :

AWSTATS permet un affichage mois par mois, des statistiques concernant les page visitées du site, ainsi que la configuration des clients qui s'y sont connectés en donnant un relevé des navigateurs et système d'exploitation utilisés par les clients.

Menu kimbo

Domaines - Héberger un domaine - fidafrique.eu.org
Comptes Mails - fidafrique.eu.org
Comptes FTP
Listes de discussion (avec mailman)
Gestionnaire de fichiers
Statistiques Web
Répertoires protégés
MySQL - Bases de données - Admin SQL
Afficher les quotas
Aide en ligne
Configuration Se déconnecter



2.Création d'un jeu de statistiques depuis alternC :

Que nous faut-il :

Pour pouvoir accéder à awstats sur le serveur alternC, nous devons tout d'abord créer un ou plusieurs couples « nom d'utilisateur/mot de passe » qui serviront à protéger la confidentialité de nos statistiques. Pour cette documentation nous utiliserons le compte de tests « Kimbo », sur le serveur farafina.net.

Création du/des compte/s :

Dans un premier temps, la création du compte d'utilisateur. Cela se passe dans la rubrique « Statistiques web ». Ce qui aura pour effet d'afficher sur le bureau AlternC la liste des jeux de statistiques qui pour le moment est vide. Depuis cette page nous pourrons gérer les utilisateurs accédant aux statistiques ainsi que les différents jeux de statistiques. Pour ajouter un premier utilisateur , nous allons sélectionner « Gérer les utilisateurs autorisés ».

Liste des jeux de statistiques

- [Gérer les utilisateurs autorisés](#)
- [Créer un jeu de statistiques](#)

La liste des utilisateurs autorisés étant vide, il est normal que nous n'en voyions aucun pour le moment. Nous pouvons par contre créer un nouvel utilisateur pour accéder à nos statistiques depuis ici.

Liste des utilisateurs autorisés

Nom	kimbo	–	
Mot de passe			
Créer cet utilisateur.			

Liste des utilisateurs autorisés

Nom	kimbo	–	km
Mot de passe	k99m		
Créer cet utilisateur.			

Nous devons donner un nom à notre utilisateur. Celui-ci devra commencer par le nom d'utilisateur AlternC.

Puis nous devons aussi choisir un mot de passe.

Une fois ceci fait, nous validons en actionnant le bouton

« Créer cet utilisateur ».

Liste des utilisateurs autorisés

Nom	kimbo
Mot de passe	
<input type="button" value="Créer cet utilisateur."/>	

Après l'ajout de l'utilisateur, nous voyons dans la liste des utilisateurs autorisés apparaître notre nouvel utilisateur « kimbo_km ».

Le compte utilisateur a été créé avec succès

Nom

[Changer le mot de passe](#) kimbo_km

Création du jeu de statistiques :

Nous pouvons maintenant créer le jeu de statistiques.

La création du jeu de statistiques se fait depuis la rubrique « Statistiques web »

Actuellement, aucun jeu de statistique n'est créé, nous n'avons donc rien dans la liste des jeux. Pour créer notre premier jeu de statistiques, nous devons choisir le lien « Créer un jeu de statistiques », qui nous emmènera au formulaire de création de jeu de statistiques suivant :

Le formulaire de création du jeu de statistiques nous propose de choisir parmi les domaines que nous avons enregistrés sous notre nom. Nous devons donc choisir sur quel nom de domaine seront réalisées les statistiques.

Créer un jeu de statistiques

Nom de domaine	www.fidafrique.eu.org
Utilisateurs autorisés	kimbo_km
<input type="button" value="Créer ce jeu de statistiques"/>	

Nous devons désigner au moins un utilisateur qui aura le droit d'accéder à ces statistiques. Nous pouvons sélectionner celui-ci en le surlignant. Si plusieurs utilisateurs doivent être sélectionnés, nous le ferons en maintenant la touche <CTRL> enfoncée.

Liste des jeux de statistiques

- [Gérer les utilisateurs autorisés](#)
- [Créer un jeu de statistiques](#)

Le jeu de statistiques a été créé avec succès

	Nom de domaine	Allowed Users	Voir
<input type="checkbox"/>	Modifier www.fidafrique.eu.org	kimbo_km	Voir
<input type="button" value="Effacer les statistiques cochées"/>			

Prêt? Alors activons le bouton « Créer ce jeu de statistiques ».

Et voilà! Notre jeu de statistiques est créé, nous le retrouvons dans la liste des jeux de statistiques.

À tout moment, nous pourrons le modifier en activant le lien modifier.

Pour voir notre jeu, nous pouvons activer le lien « Voir ». Mais avant parcourons le paragraphe suivant.

3. Avant d'aller plus loin :

Avant de poursuivre et de visualiser nos statistiques, nous devons attendre quelques heures voire quelques jours pour que les visites réalisées par les internautes commencent à produire des statistiques visibles. AWSTATS parcourt chaque nuit les fichiers journaux du serveur web apache pour mettre en forme et stocker les statistiques de l'avant-veille.

4. Description des rubriques d'AWSTATS :

Awstats comprend un grand nombre de rubriques qui renseignent l'utilisateur sur les activités des visiteurs de notre site web.

Cette partie nous permettra de détailler chacune de ces rubriques pour connaître la nature et la teneur de son contenu.


Nous pouvons accéder à nos statistiques depuis notre bureau AlternC. Le nom d'utilisateur et le mot de passe que nous avons créés nous seront demandés pour y accéder, même si nous avons déjà ouvert le bureau AlternC.

Vue générale :

AWSTATS se présente en une seule page affichant toutes les rubriques de statistiques en continu; sur la gauche se situe un menu permettant d'accéder aux différentes rubriques.

Statistiques de:
www.fidafrique.eu.org

Dernière mise à jour: 27 Nov 2005 - 04:41




Période d'analyse:

Résumé

Période d'analyse	Mois Nov 2005				
Première visite	21 Nov 2005 - 13:33				
Dernière visite	25 Nov 2005 - 06:16				
	Visiteurs différents	Visites	Pages	Hits	Bande passante
Trafic 'vu' *	6	9 <small>(1.5 visites/visiteur)</small>	20 <small>(2.22 pages/visite)</small>	28 <small>(3.11 hits/visite)</small>	65.83 Ko <small>(7.31 Ko/visite)</small>
Trafic 'non vu' *			4	4	245 Octets

* Le trafic 'non vu' est le trafic généré par les robots, vers ou réponses HTTP avec code retour spécial.

Historique mensuel



Mois	Visiteurs différents	Visites	Pages	Hits	Bande passante
Jan 2005	0	0	0	0	0
Fév 2005	0	0	0	0	0
Mar 2005	0	0	0	0	0
Avr 2005	0	0	0	0	0
Mai 2005	0	0	0	0	0
Juin 2005	0	0	0	0	0
Juil 2005	0	0	0	0	0
Août 2005	0	0	0	0	0
Sep 2005	0	0	0	0	0
Oct 2005	0	0	0	0	0

Le menu se découpe en 5 sous-rubriques:

- Quand :
- Qui :
- Navigation :
- Origine/Referer :
- Autres :

EITIC - Solutions et Services en Logiciels Libres – www.eitic.fr

Barre d'en-tête :

La barre d'en-tête permet la sélection par mois d'une année sélectionnez le mois et l'année pour lesquels vous souhaitez visualiser les statistiques puis validez ce choix en appuyant sur le bouton « OK ».

Dernière mise à jour:	27 Nov 2005 - 04:41	
Période d'analyse:	Nov 2005 OK	

Attention : l'analyse détaillée pour une année complète n'est pas disponible sur le serveur d'hébergement mutualisé : elle requerrait beaucoup trop de CPU, c'est à dire qu'elle ferait trop travailler le serveur et pénaliserait lourdement les connexions au serveur.

Résumé :

Résumé					
Période d'analyse	Mois Nov 2005				
Première visite	21 Nov 2005 - 13:33				
Dernière visite	25 Nov 2005 - 06:16				
	Visiteurs différents	Visites	Pages	Hits	Bande passante
Trafic 'vu' *	6	9 (1.5 visites/visiteur)	20 (2.22 pages/visite)	28 (3.11 hits/visite)	65.83 Ko (7.31 Ko/visite)
Trafic 'non vu' *			4	4	245 Octets

* Le trafic 'non vu' est le trafic généré par les robots, vers ou réponses HTTP avec code retour spécial.

Le résumé affiche un rapide abrégé des statistiques du mois sélectionné, intégrant les dates de première et dernière visite.

Le trafic « vu » est le trafic qui génère l'affichage d'une page web.

Le trafic « non vu » résulte de l'utilisation de redirection par exemple. C'est à dire qu'il ne génère pas dans le navigateur l'affichage direct d'une page web.

Les visiteurs différents sont en réalité le nombre d'ordinateurs depuis lesquels le site a été visité, si par exemple, des personnes se sont connectées depuis un réseau privé, il y a toutes les chances pour que celles-ci soient vues comme un seul et même visiteur.

Les visites précisent le nombre de fois ou le site a été consulté, une visite est une nouvelle connexion au site. Les chargements de chaque page ne sont pas enregistrés ici. Le chiffre entre parenthèses, indique le nombre moyen de visites réalisé par chaque visiteur.

Les pages donnent le nombre total de pages web téléchargées et, entre parenthèses, combien de pages en moyenne ont été chargées par visite.

Les hits indiquent le nombre de liens cliqués lors des différentes visites et, entre parenthèses, le nombre de hits réalisés en moyenne pour chaque visite. Le fait de visiter une page, puis de revenir à la page précédente = 2 hits.

Enfin la bande passante, décrit le nombre d'octets qui ont été chargés en Ko, Mo Go ou To en fonction de la quantité et, entre parenthèses, la moyenne pour chaque visite.

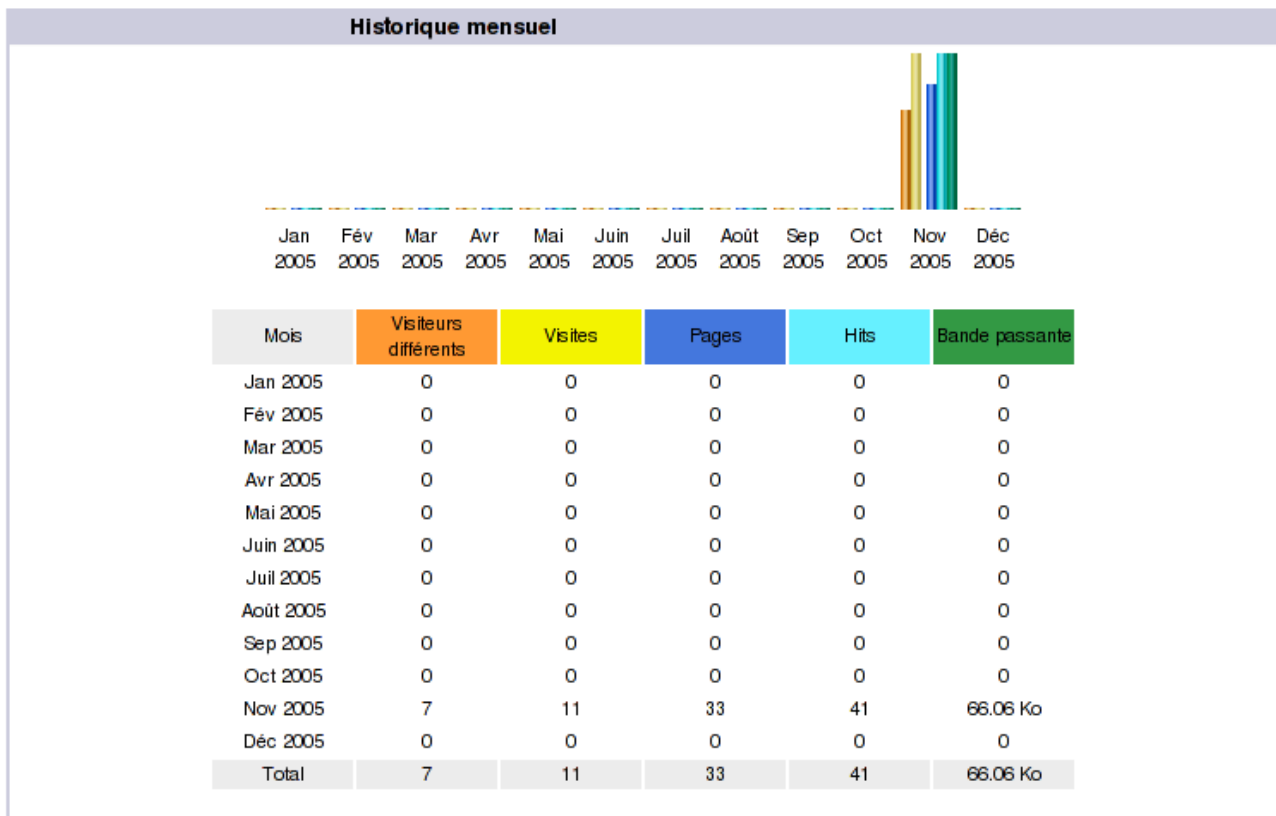
Quand

Historique mensuel :

L'historique mensuel relate mois par mois, les statistiques de consultation du site web, d'abord sous la forme d'un histogramme puis sous la forme d'un tableau permettant de visualiser précisément les chiffres pour chaque mois.

Par exemple pour novembre 2005 :

7 visiteurs ont réalisé 11 visites pendant lesquelles ils ont vu 33 pages et cliqué sur 41 liens, ce qui a généré un trafic de 66,06 Ko.

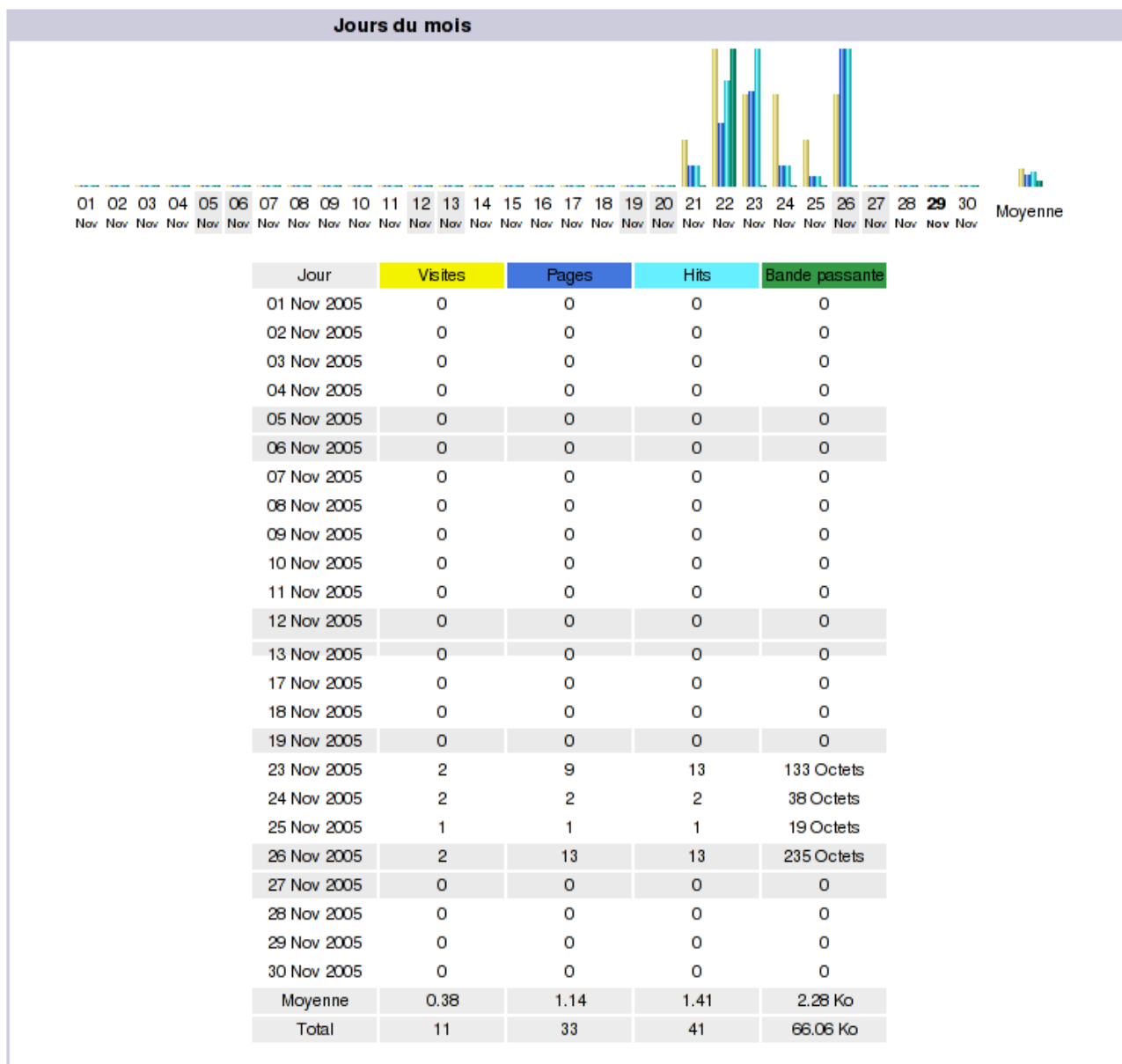


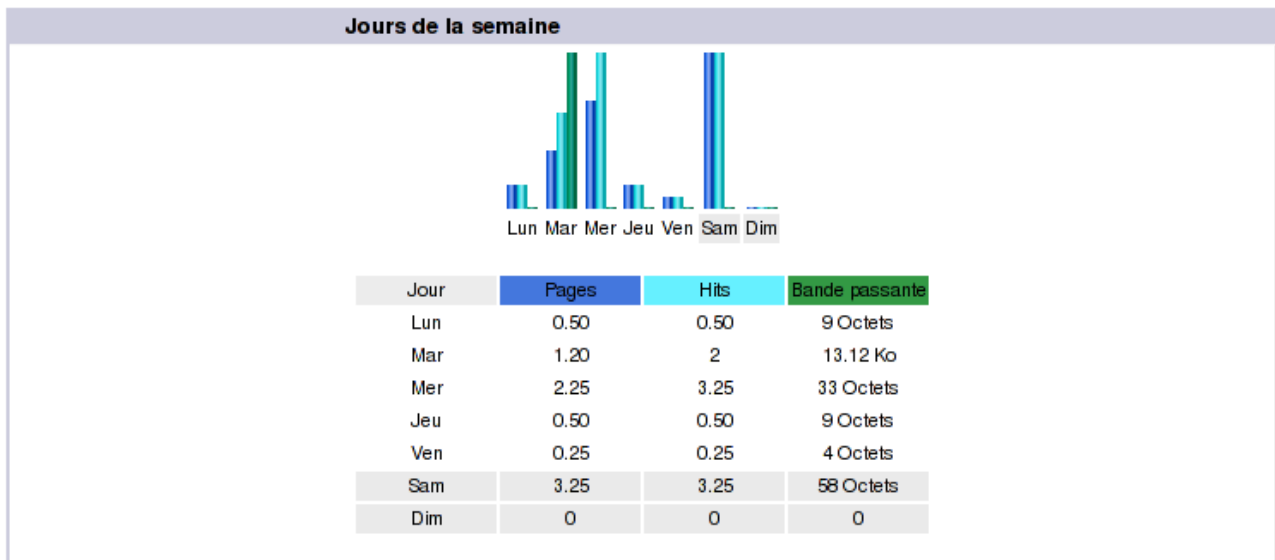
Statistiques par jour du mois :

La rubrique suivante affiche le nombre de visites et les statistiques qui s'y rapportent par jour du mois sélectionné. Les jours, sur fond grisé, représentent les samedis et les dimanches.

L'affichage est réalisé d'abord sous la forme d'un histogramme puis ensuite sous la forme d'un tableau donnant plus précisément les chiffres pour chaque jour du mois.

Par exemple, le 23 novembre, il y a eu 2 visiteurs qui ont vu 9 pages en actionnant 13 liens, ce qui a généré un trafic de 133 Octets.



Statistiques moyenne par jour de la semaine :

L'affichage des statistiques par jour de la semaine, donne le trafic moyen par jour de la semaine, en réalisant une moyenne des visites pour tous les lundis du mois, puis pour tous les mardi, etc ...

L'affichage se présente encore une fois d'abord sous la forme d'un histogramme puis sous la forme d'un tableau affichant:

- Le nombre de pages visitées en moyenne pour chaque jour
- Le nombre moyen de hits pour chaque jour
- La bande passante consommée par jour

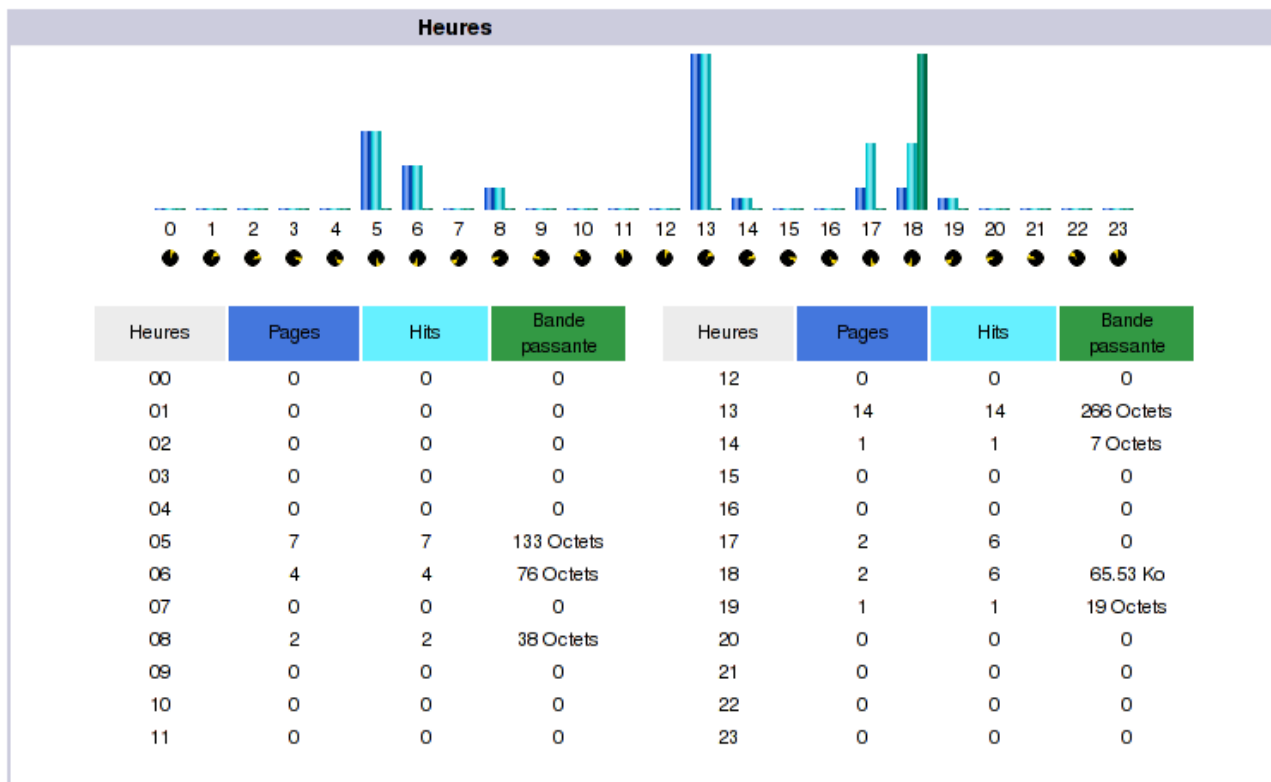
Par exemple, dans notre cas, le pic de visites -en terme de nombre de pages téléchargées- est plutôt le samedi, par contre, la consommation de bande passante est plus importante le mardi.

Statistiques moyenne par heure de la journée :

Comme pour les statistiques par jour de la semaine, les statistiques par heure de la journée sont le résultat d'une moyenne des statistiques du site pour chaque tranche horaire d'une heure.

Présentées, d'abord sous la forme d'un histogramme, puis sous la forme d'un tableau de valeurs. Elle donne, dans notre exemple, les informations suivantes :

Le pic du nombre de page téléchargées se situe dans la tranche de 13h et que par contre le trafic généré voit son apogée plutôt dans la tranche des 18 heures.



Qui :

Classement des visiteurs par domaine et par pays :

Le classement des visiteurs par domaine ou pays affiche dans l'ordre décroissant du nombre de pages téléchargées, le nombre de pages, le nombre de hits et la bande passante mobilisée par domaine et ou par pays.

L'origine est déduite à partir du nom résolu du client. Ces données ne sont pas d'une fiabilité exemplaire, puisque l'extension « .net » est considérée comme faisant partie d'un pays nommé network.

Par ailleurs, les adresses pour lesquelles une résolution de noms aurait échoué sont classées dans la catégorie Inconnu.

Domaines/pays visiteurs (Top 200) - [Liste complète](#)

Domaines/Pays	Pages	Hits	Bande passante
Network	26	26	482 Octets
Inconnu	5	13	65.55 Ko
France	2	2	38 Octets

Classement des visiteurs par nom de machine ou par ip :

Le classement des visiteurs par nom de machine ou par ip affiche dans l'ordre décroissant des pages téléchargées, la liste des noms de domaines (ou à défaut des adresses IP) d'où les clients se sont connectés.

Une colonne « Dernière visite » indique la date et heure de la dernière connexion pour le client correspondant.

Hôtes (Top 200) - Liste complète - Dernière visite - Adresses IP non résolues				
Hôtes : 4 Connus, 3 Inconnus (IP non résolue) - 7 Visiteurs différents	Pages	Hits	Bande passante	Dernière visite
Ins-th2-5f-81-56-229-182.adsl.proxad.net	24	24	456 Octets	26 Nov 2005 - 13:43
196.207.218.82	2	6	0	23 Nov 2005 - 17:21
196.207.234.28	2	6	65.53 Ko	22 Nov 2005 - 18:08
home.eitic.fr	2	2	38 Octets	21 Nov 2005 - 13:34
Ins-bzn-28-82-254-208-8.adsl.proxad.net	1	1	7 Octets	26 Nov 2005 - 14:12
inzu.nkurunziza.net	1	1	19 Octets	22 Nov 2005 - 08:07
84.4.58.84	1	1	19 Octets	24 Nov 2005 - 08:59

Visites de robots/spiders :

Les visites de robots sont affichées dans cette rubrique.

Visiteurs Robots/Spiders (Top 200) - Liste complète - Dernière visite			
0 robots différents*	Hits	Bande passante	Dernière visite

* Les robots présentés ici sont à l'origine de hits ou de trafic "non vus" par les visiteurs donc non représentés dans les autres tableaux.

Qu'est ce qu'un robot? Un robot est un programme mis en oeuvre généralement par un moteur de recherche pour indexer le contenu des pages d'un site. Les robots parcourent les sites web pour alimenter leur base de données et ainsi permettre une tenue à jour des moteurs de recherche.

Navigation :**Durée des visites :**





L'affichage de la durée des visites calcule d'abord la durée moyenne des visites puis présente, sous forme d'un tableau, le nombre de visites et le pourcentage que celles-ci représentent par rapport au total des visites classées par durée de 0-30 s à 1h et plus.

Le pourcentage de visites inconnues représente les visites dont le temps n'a pu être mesuré.

Durée des visites		
Visites: 11 - Moyenne: 35 s	Visites	Pourcentage
0s-30s	6	54.5 %
30s-2mn	3	27.2 %
2mn-5mn		
5mn-15mn		
15mn-30mn		
30mn-1h		
1h+		
Inconnu	2	18.1 %

Types de fichiers téléchargés :





Les types de fichiers téléchargés sont affichés dans l'ordre décroissant des téléchargement; il ne s'agit pas des fichiers proprement dits mais des types de fichiers classés en fonction de leur extension (la chaîne de caractère située après le « . » dans le nom du fichier). Les statistiques concernant le type de fichier sont comptabilisées en nombre de hit (demande de fichier par le navigateur) et en bande passante mobilisée.

Types de fichiers						
Types de fichiers			Hits	Pourcentage	Bande passante	Pourcentage
	html	HTML or XML static page	32	78 %	53.70 Ko	81.2 %
	css	Cascading Style Sheet file	4	9.7 %	4.09 Ko	6.1 %
	png	Image	4	9.7 %	8.25 Ko	12.4 %
	php	Dynamic Html page or Script file	1	2.4 %	19 Octets	0 %

Classement des pages les plus téléchargés :


Après avoir vu les types de fichiers les plus téléchargés, nous voyons maintenant les fichiers eux-même les plus téléchargés, classés par nombre de pages vues.

L'adresse de la page est indiquée après le « / », le « / » représente le nom de domaine, ici par exemple « /index.php » indique une visite de la page « www.fidafrique.eu.org/index.php ».

Pages-URL (Top 200) - Liste complète - Entrée - Sortie					
4 pages différentes	Pages vues	Taille moyenne	Entrée	Sortie	
/	28	18 Octets	8	7	
/tech-forum/maillist.html	2	13.24 Ko		2	
/tech-forum/threads.html	2	13.36 Ko	2		
/index.php	1	19 Octets	1		

Système d'exploitation des visiteurs :

L'affichage des systèmes d'exploitation des visiteurs permet de connaître la nature du système d'exploitation utilisé par l'ordinateur des visiteurs. Dans notre cas, les ordinateurs sous Linux sont majoritaires, dans la réalité, les proportions sont largement inverses.

Systèmes exploitation (Top 10) - Liste complète/Versions - Inconnu				
Systèmes exploitation			Hits	Pourcentage
	Linux		21	51.2 %
	Windows		20	48.7 %

Navigateurs des visiteurs :

De même que pour le système d'exploitation des visiteurs, une autre information importante est la connaissance du navigateur qu'ils utilisent, cette indication est précieuse. Elle incite à réaliser des tests d'affichage de notre site avec ces navigateurs pour tester la compatibilité et ainsi rendre notre site plus accessible.

Navigateurs (Top 10) - Liste complète/Versions - Inconnu					
Navigateurs			Aspirateur	Hits	Pourcentage
	FireFox		Non	33	80.4 %
	MS Internet Explorer		Non	4	9.7 %
	Netscape		Non	2	4.8 %
	Opera		Non	2	4.8 %

Origine/Referer:

Connexion par origine :

L'origine d'une visite est définie comme étant la page d'ou provient le visiteur avant d'arriver sur le site. AWSTATS classe celles-ci en plusieurs catégories:

- Adresse directe/Bookmarks : le visiteur est arrivé directement sur le site soit en tapant l'adresse de celui ci soit en sélectionnant un bookmark(favoris).
- Lien depuis un NewsGroup
- Lien depuis un moteur de recherche : Le visiteur est arrivé sur notre site à la suite d'une recherche réalisée sur un moteur de recherche. Le lien liste complète permet de détailler la liste des moteurs de recherche qui ont redirigé des visiteurs vers notre site.
- Lien depuis une page externe : Le visiteur est arrivé sur notre site en cliquant un lien placé sur un autre site que le notre.
- Origine inconnue : Les visites pour lesquelles l'origine n'a pu être déterminée.

Connexions au site par				
Origine de la connexion	Pages	Pourcentage	Hits	Pourcentage
Adresse directe / Bookmarks	31	100 %	31	100 %
Lien depuis un NewsGroup				
Lien depuis un moteur de recherche Internet - Liste complète				
Lien depuis une page externe (autres sites, hors moteurs) - Liste complète				
Origine inconnue				

Phrase clés et mots clés :

Les phrases clés et mots clés indiquent, lorsque l'origine de la visite est le résultat d'un moteur de recherche, les phrases ou mots clés qui ont emmené le visiteur à trouver notre site web.

Phrases clés (Top 10)		
Liste complète		
0 phrases clé différentes	Recherche	Pourcentage

Mots clés (Top 200)		
Liste complète		
0 mots clé différents	Recherche	Pourcentage

C'est pour nous une aide précieuse : ces informations permettent de mieux référencer notre site.

Autres :**Divers :**

Les divers nous informent avec plus ou moins de fiabilité sur les capacités des navigateurs des visiteurs en particulier, si ceux ci possèdent les greffons (plugins) nécessaires à la lecture de médias flash, java, pdf ...

Divers		
Divers		
Ajout aux favoris (estimation)	21 / 7 Visiteurs	300 %
Navigateurs avec support Java actif	-	
Navigateurs avec support Macromedia Director	-	
Navigateurs avec support Flash	-	
Navigateurs avec support audio Real	-	
Navigateurs avec support audio QuickTime	-	
Navigateurs avec support audio Windows Media	-	
Navigateurs avec support PDF	-	

Codes status HTTP :

Les codes status HTTP regroupent les codes d'erreur éventuellement rencontrés par les usagers du site web, ceux ci ne sont pas comptabilisés dans les stats du site. En nous reportant au résumé, nous pouvons voir que ces codes correspondent au « Trafic non vu ».

Codes Status HTTP				
Codes Status HTTP*		Hits	Pourcentage	Bande passante
302	Moved temporarily (redirect)	3	75 %	15 Octets
404	Document Not Found	1	25 %	230 Octets

* Les codes présentées ici sont à l'origine de hits ou de trafic "non vus" par les visiteurs donc non représentés dans les autres tableaux.

5.Conclusion

AWSTATS existe en tant que module intégrable d'AlternC ce qui rend son installation et sa mise à jour extrêmement simple dans une configuration utilisant ce système d'hébergement.

Les différentes rubriques décrites ici le sont dans le cadre de son utilisation au sein d'un hébergement mutualisé AlternC. Certaines différences peuvent apparaître si vous utilisez AWSTATS différemment.

Licence de ce document :

Ce document est sous licence GNU/FDL

Copyright (c) EITIC pour Fidafrique

Permission est accordée de copier, distribuer et/ou modifier ce document en respectant les termes de la licence GNU/FDL :

<http://www.gnu.org/licenses/licenses.fr.html>