

Mesure d'audience des sites web

Présentation originale USMB préparée conjointement avec :
Olivier Pouydebat (Service des Systèmes Applicatifs)
et Benoit Charles (Service communication)

Chronologie des solutions



Services ciblés :

- **www ;**
- **Intranet ;**
- **Sites de services ;**
- **Sites d'évènements.**

Abandon de Google Analytics en 2017

- Ne plus envoyer de données stat. hors établissement
 - ↳ Aller vers l'abandon du traceur js
 - ↳ Décision service Com/DN
- En 2017 les recommandations de la CNIL sont :
 - Les utilisateurs doivent être informés
 - Ne tracer que le strict nécessaire
 - Durée de conservation maxi des infos collectées => 25 mois
- Solution préconisée par la CNIL (~20 solutions en 2022)
 - <https://www.cnil.fr/fr/cookies-et-autres-traceurs/regles/cookies-solutions-pour-les-outils-de-mesure-daudience>

!!
Des extensions
WP utilisent
des cookies

Sommaire doc install CNIL Matomo

Introduction	2
Deux modes d'hébergement	3
a) Matomo en version auto-hébergé (Self-hosted / On-Premise)	3
b) Matomo en version cloud : Matomo Cloud (Saas)	3
Procédure à suivre	4
1) désactiver les exports de données	4
2) permettre aux visiteurs de refuser d'être suivis (Opt-out)	4
3) vérifier que les paramètres par défaut de Matomo sont toujours en place	4
a) s'assurer que les adresses IP sont anonymes	4
b) s'assurer que les cookies tiers et le cross-domain ne sont pas utilisés	4
c) s'assurer que la mesure du "User ID" n'est pas utilisée	5
d) s'assurer que la mesure du E-Commerce n'est pas utilisée	5
e) s'assurer que les cartes de chaleurs et enregistrements de session sont désactivés	6
4) vérifier que vous ne collectez pas de données personnelles	6

Mesures via tracker js VS mesures via logs

Deux périmètres qui se recoupent

- JS : 1 évènement par page chargée
 - Richesse d'info : clics sortants ;
 - Pas de traitement batch ;
 - Nécessite un serveur d'audience toujours up ;
 - Bandeau cookie/traceur RGPD inévitable ;
 - Trafic altéré par les navigateurs modernes et adblocker.
- LOG : 1 ligne par ressource chargée
 - Plus de bruit (nécessité de filtrer l'info) : robots, attaques, API , etc. ;
 - Pas d'info sur les clics sortants ;
 - N'impacte pas les clients ;
 - Import asynchrone possible ;
 - /!\ pas de support du opt-out

🔍 Rechercher

📅 2023-02-14

NOUVELLE MISE À JOUR : MATOMO 4.13.3 📢



Tableau de bord de tous les sites web (Total : 14 109 visites, 32 756 pages vues, 34 366 actions, 0 revenu)

SITE WEB	▼ VISITES	VUES DE PAGE	REVENU	ÉVOLUTION	Visites ▼
intranet ↗	7 821	9 573	0 €	↓ -12,4 %	
www ↗	6 124	16 873	0 €	↓ -7,8 %	
ailes ↗	80	2 946	0 €	↓ -24,5 %	
btk ↗	45	3 000	0 €	↓ -35,7 %	
pilotage-statistiques ↗	39	364	0 €	↑ 62,5 %	
cockpit ↗	0	0	0 €	● 0 %	
repertoire-activites ↗	0	0	0 €	● 0 %	

➕ Ajouter un site

1 - 7 de 7



Rechercher

WWW

2023-02-14

TOUTES LES VISITES

TABLEAU DE BORD

NOUVELLE MISE À JOUR : MATOMO 4.13.3

- Tableau de bord
- Visiteurs
- Comportement
- Acquisition
- Objectifs
- Marché

Visites en temps-réel

DATE	VISITES	ACTIONS
Dernières 24 heures	2 608	7 371
Dernières 30 minutes	0	0

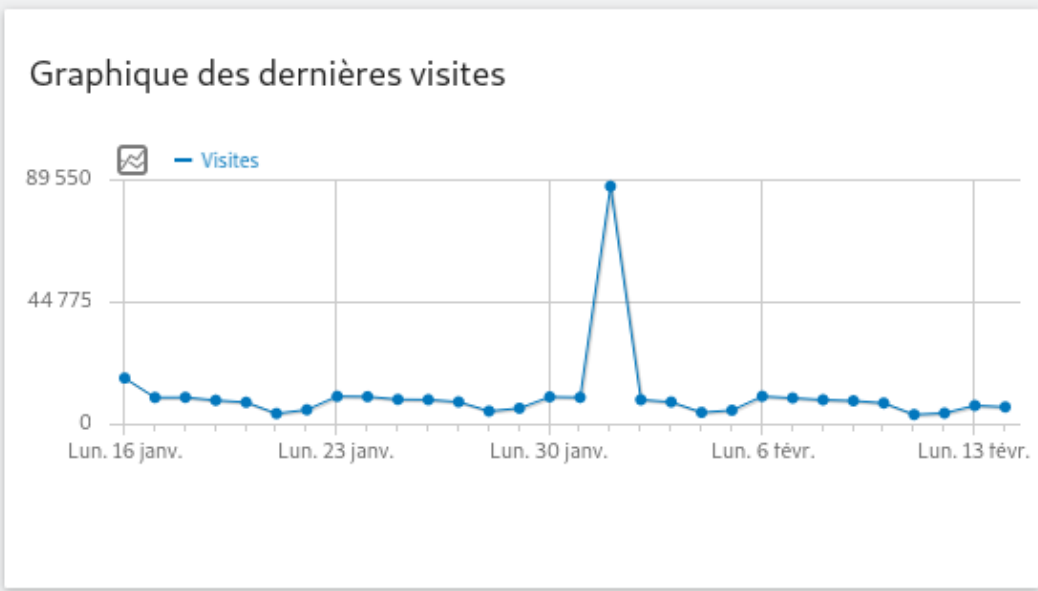
Mercredi 15 février - 00:02:01 (7s)

Entrées directes

Actions:

Mercredi 15 février - 00:02:00 (1 min 0s)

Entrées directes



Fonctionnalités et Services Premium



Rechercher

INTRANET

2023-02-14

TOUTES LES VISITES

TABLEAU DE BORD

NOUVELLE MISE À JOUR : MATOMO 4.13.3

Tableau de bord

Tableau de bord

Visiteurs

Comportement

Acquisition

Objectifs

Marché

Visites en temps-réel

DATE	VISITES	ACTIONS
Dernières 24 heures	3 365	4 055
Dernières 30 minutes	0	0

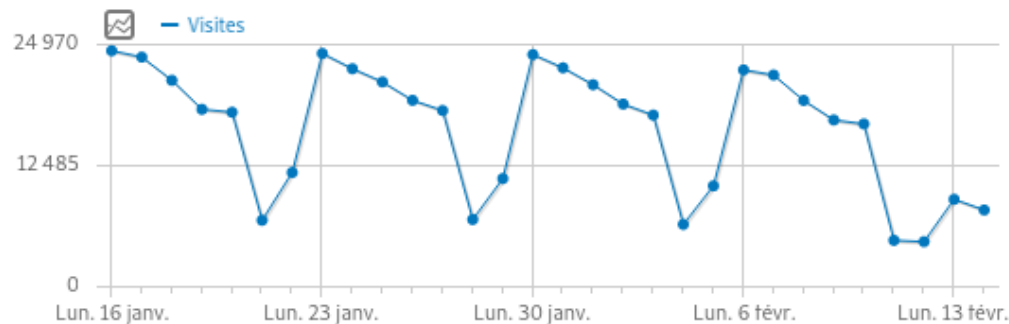
Mardi 14 février - 23:58:45 iOS

Site: cas-uds.grenet.fr

Actions:

Mardi 14 février - 23:57:50

Graphique des dernières visites



The screenshot displays the Matomo dashboard interface. At the top, the Matomo logo is visible on the left, and a search bar with the text 'Rechercher' is on the right. Below the search bar, there are several navigation buttons: 'WWW', '2023-02-14', 'TOUTES LES VISITES', and a notification for 'NOUVELLE MISE À JOUR : MATOMO 4.13.3'. The left sidebar contains a menu with the following items: 'Tableau de bord', 'Visiteurs', 'Récapitulatif', 'Journal des Visites', 'En temps réel', 'Carte en temps-réel', 'Provenances géographiques', 'Périphériques', 'Logiciel', 'Horaires', 'ID Utilisateur' (which is highlighted), 'Comportement', and 'Acquisition'. The main content area shows the 'ID Utilisateur' report, which currently displays the message: 'Il n'y a pas de données pour ce rapport.' Below this message is a link: 'Pour en savoir plus sur la façon de générer des données pour ce rapport, consultez notre guide de l'utilisateur.'

Mardi 14 février 2023 - 23:52:00

IP: 88.173.0.0

🇫🇷 Évry

Google



3 Actions - 5 min 3s

[Afficher le profil du visiteur](#)

- /2021/10/18/a-la-decouverte-des-accompagnateurs-a-la-reussite-etudiante-de-L...
- |
- /categorie/recherche/
- |
- /categorie/recherche/page/2/

Mardi 14 février 2023 - 23:56:38

IP: 89.80.0.0

🇫🇷 Lyon

Entrées directes

iOS

1 Action

[Afficher le profil du visiteur](#)

- /espace-etudiant/event/atelier-preparer-sa-candidature-en-1re-annee-de-master/

Mardi 14 février 2023 - 23:45:41

IP: 81.220.0.0

🇫🇷 Nantes

Google



5 Actions - 10 min 50s

[Afficher le profil du visiteur](#)

- /iut-annecy/
- |
- /iut-annecy/formation/reseaux-informatiques-et-telecoms/
- |
- /iut-annecy/
- |
- 🔄 2 /iut-annecy/iut/

Mardi 14 février 2023 - 23:56:22

IP: 176.166.0.0

🇫🇷 Rouen

Google



1 Action

[Afficher le profil du visiteur](#)

- /symme/members/yann-meyer/

Mardi 14 février 2023 - 22:40:53

IP: 193.48.0.0

🇫🇷 Chambéry

Entrées directes

12 Actions - 1 heures 15 min

[Afficher le profil du visiteur](#)

- /espace-etudiant/actualites/feed/
- |
- /espace-etudiant/actualites/feed/
- |
- /categorie/intranet/feed/

matomo

[Tableau de bord](#)
[Tous les sites web](#)
⚙️
ℹ️
🔗

▼ WWW

📅 2023-02-14

👤 TOUTES LES VISITES

NOUVELLE MISE À JOUR : MATOMO 4.13.3
🔔

- ☰ **Tableau de bord**
- ∞ **Visiteurs**
- Récapitulatif
- Journal des Visites
- En temps réel
- Carte en temps-réel
- Provenances géographiques
- Périphériques
- Logiciel
- Horaires
- ID Utilisateur
- 🔔 **Comportement**

Type du périphérique

▼ VISITEURS UNIQUES

TYPE	VISITEURS UNIQUES
🖥️ Bureau	4 407
📱 Téléphone intelligent (smartphone)	1 238
Inconnu	345
📱 Phablette	99
📱 Tablette	34
🎵 Lecteur de medias portable	1
🖥️ Affichage intelligent	-
📷 Caméra	-
🎮 Console	-
📞 Fonctionnalité téléphone	-

1 - 10 de 15 [Suivant >](#)

Modèle du périphérique

▼ VISITEURS UNIQUES

MODÈLE	VISITEURS UNIQUES
Générique Bureau	3 626
Apple - Générique Bureau	780
Apple - iPhone	590
Inconnu	243
Générique Téléphone intelligent (smartphone)	109
Apple - Générique Périphérique mobile	97
Apple - iPad	17
Xiaomi - Mi 11 Lite 5G	14
Samsung - Galaxy A52s 5G	12
Samsung - Galaxy S10	11

1 - 10 de 281 [Suivant >](#)

[WWW](#)

[TOUTES LES VISITES](#)
NOUVELLE MISE À JOUR : MATOMO 4.13.3

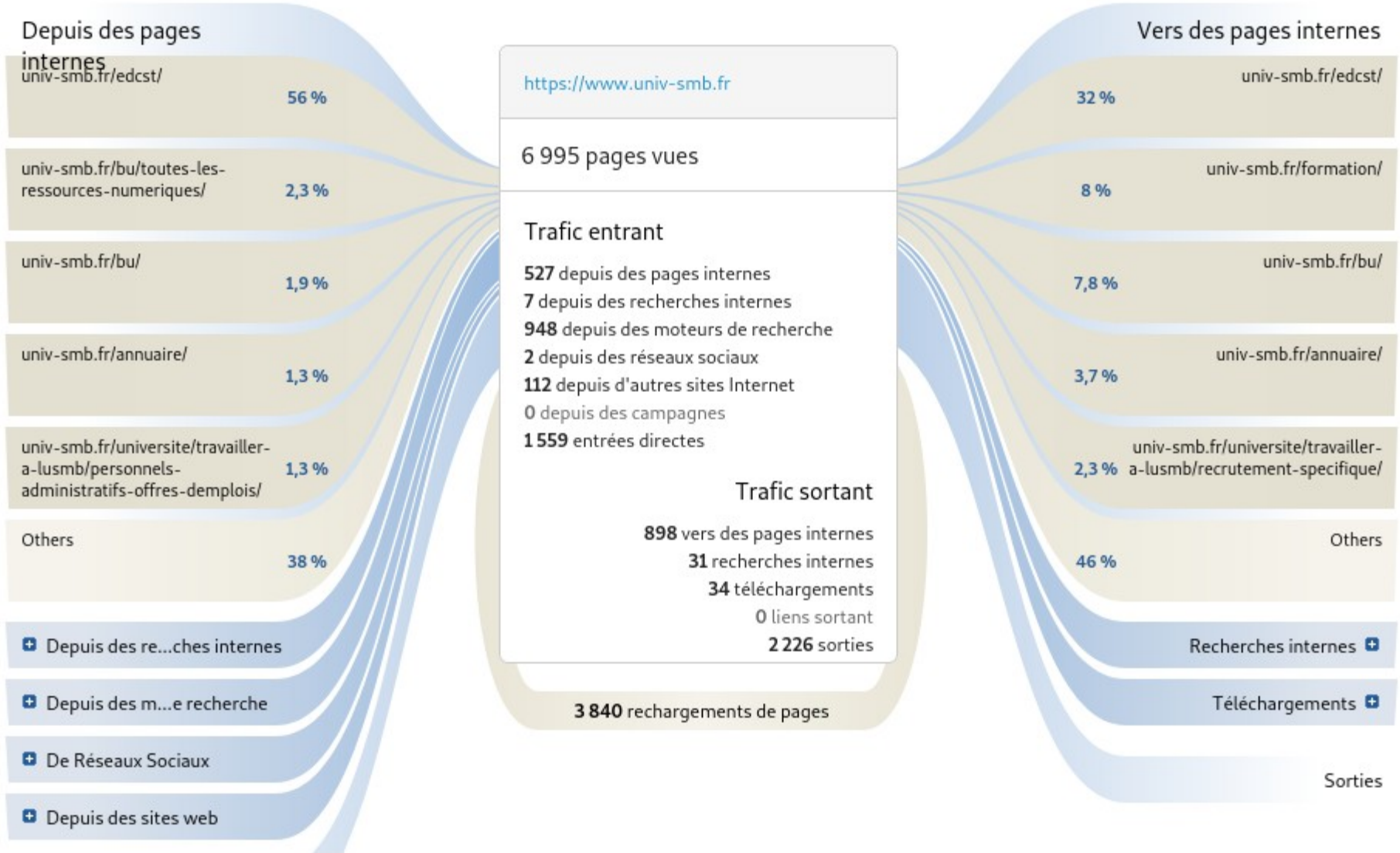

Pages

URL DE LA PAGE	VUES DE PAGE	VUES DE PAGE UNIQUES	TAUX DE REBOND	TEMPS MOYEN SUR LA PAGE	TAUX DE SORTIE	TEMPS DE CHARGEMENT MOYEN
/index	6 995	2 735	68 %	00:02:19	81 %	-
bu	1 098	879	46 %	00:02:54	57 %	-
formation	1 179	784	49 %	00:00:46	49 %	-
iut-annecy	863	625	60 %	00:02:29	56 %	-
espace-etudiant	708	603	60 %	00:01:02	49 %	-
universite	433	370	64 %	00:00:49	40 %	-



Ouvrir les transitions
Visualisez ce que les visiteurs ont fait avant et après avoir vu cette page

[Tableau de bord](#)
[Visiteurs](#)
[Comportement](#)
[Pages](#)
[Pages d'entrée](#)
[Pages de sortie](#)
[Titres des pages](#)
[Recherche sur le site](#)
[Liens sortants](#)
[Téléchargements](#)



Entrées directes

matomo

Analyse des pages web ⓘ

Mois courant ▾

Localisation: /index

Segment: Toutes les visites

Métriques principales

- 116310 Vues de page
- 48434 Visites
- 67 % Taux de rebond
- 83 % Taux de sortie
- 2 min 16s Temps moyen sur la page

Voir l'évolution de la ligne

Ouvrir les transitions

Ouvrir le journal segmenté des visites

Did you know? With [Heatmap & Session Recording](#) you understand your visitors much better. See where your visitors click, move the mouse, scroll, type, and more.

Liens directs ▾

f t YouTube i in 0.1% 0.1% 0.3% 1% 0.6%

B.U. Actualités Agenda Plans Annuaire Intranet ? Recherche

UNIVERSITÉ SAVOIE MONT BLANC 80%

UNIVERSITÉ 0.4% FORMATION 1.2% RECHERCHE 0.1% INTERNATIONAL 0.1% VIE ÉTUDIANTE 0.1% UNITA universités montium

0% Le président de l'USMB présente ses vœux pour 2023
7 février 2023

0% Féraventure, un jeu pour partir à la découverte des coréogones
6 février 2023

0% Des acteurs locaux à la rencontre des étudiants de la Faculté de droit
6 février 2023

0% Une journée dans la peau d'un étudiant de l'USMB
6 février 2023

0% Un premier appel à projets réussi pour le réseau RELIEF
6 février 2023

15 376 étudiantes et étudiants et 244 doctorantes et doctorants

7 UFR, École, Instituts

3 départements

18 laboratoires de recherche

1 Fondation universitaire et 1 Club d'entreprises

Actualisation mars 2022

MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA RECHERCHE

Liberté Égalité Fraternité

CONSEIL

ACTUALITÉ

- 0% Communiqués de presse
- 0% Dossiers de presse
- 0.1% Données et documents clés
- 0.1% Actualités
- 0% Agenda

RACCOURCIS

- 0% Paiement en ligne
- 0% Marchés publics
- 0.1% Conditions générales d'achat
- 0.3% Travailler à l'USMB
- 0.2% Plans des campus

A PROPOS

- 0% Chartre graphique de l'USMB
- 0% Glossaire
- 0.6% Réseaux et partenaires
- Aide connexion intranet

Deux durées de conservation : log réglementaires vs stats

Deux finalités distinctes

- logs : 1 an de log réglementaire
 - Pour tous les sites hébergés ;
 - A fournir en cas de réquisition judiciaire ;
 - Serveur de log dédié ;
 - Copie / déport des logs => (rsyslog, rsync, logrotate)
- Statistiques :
 - Conservation de 8j de logs bruts ;
 - Import quotidien des logs dans Matomo ;
 - 2 versions des données :
 - Détaillées en bdd => 3 mois de conservation
 - Agrégée en bdd (non individuelle) => 1 an actuellement

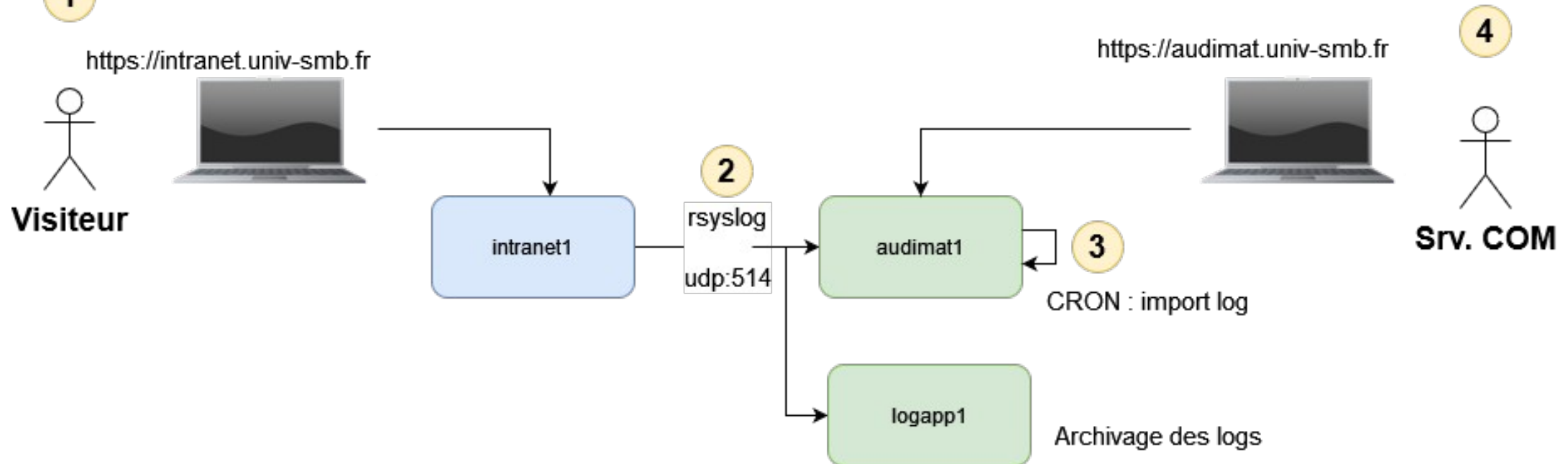
Matomo : implémentation

Installation et spécificités USMB

- Environ 10 sites web dont le `www.univ` (16k étudiants)
 - 3 mois dans la BDD = 6 millions de lignes, 2.2 Gio
`/var/lib/mysql`
- Installation On-Premise :
 - VM 4vCPU, 8 Gio de RAM ;
 - OS debian 11 ;
 - LAMP (base mariadb) ;
 - rsyslog standard UDP 514 ;

Installation et spécificités USMB

- Procédure d'installation :
 - 100% classique, cf doc officielle [install matomo "on-premise"](#)
 - + nos scripts pour les imports de logs, rotation, nettoyage, automysqlbackup
- ① • PAS de requête visiteur vers serveur audimat



Installation et spécificités USMB

- Config client rsyslog (sur les VM web, vhost apache)
 - 📄 (1^{ère} idée sans `-udp -server`, bugs espace début de `%msg%`)

```
(prod)root@intranet1:~# cat /etc/apache2/sites-enabled/vh-intranet.conf
<VirtualHost *:443>
    ServerName intranet.univ-smb.fr
    ErrorLog /var/log/apache2/intranet/error.log
    CustomLog /var/log/apache2/intranet/access.log combined
    # Forward to rsyslog server for statistics
    CustomLog "| /usr/bin/logger -t intranet -p local2.notice --udp --server 193.48.XXX.XXX" combined
    # Forward to rsyslog server for archive
    CustomLog "| /usr/bin/logger -t intranet -p local2.notice --udp --server 193.48.YYY.YYY" combined
</VirtualHost>
```

Installation et spécificités USMB

• Config serveur rsyslog (sur le serveur matomo)

```
(prod)root@audimat1:~# cat /etc/rsyslog.d/udp.conf
# provides UDP syslog reception
module(load="imudp")
input(type="imudp" port="514")
# Template pour envoi des logs dans un rep specifique
$template DynamicFile, "/var/log/remote/%HOSTNAME%/%syslogfacility-text%.log"
$template DynamicFile2, "/var/log/remote/%HOSTNAME%/%syslogtag%::securepath%-%syslogfacility-text%.log"
# Template pour stocker le message brut sans entête rsyslog
# Per RFC 3164, anything after the colon in the syslog tag gets counted as part of the %msg% field, including any
space character
# https://stackoverflow.com/questions/30745164/empty-space-at-beginning-of-rsyslog-log-file
$template rawFormat, "%msg:2:$:%\n"
$template justMsg, "%msg%\n"
# Les logs apache2 sont en facility.prio local1.notice, les envoyer en brut sans un répertoire par serveur
local2.*    -?DynamicFile2; justMsg
local2.*    stop
```

Installation et spécificités USMB

- Logs reçus sur le serveur

```
(prod)root@audimat1:/var/log/remote# ls
serveur1  serveur2  intranet1
(prod)root@audimat1:/var/log/remote/intranet1# ls
intranet-local2.log  intranet-local2.log.2.gz  intranet-local2.log.4.gz  intranet-local2.log.6.gz
intranet-local2.log.1  intranet-local2.log.3.gz  intranet-local2.log.5.gz  intranet-local2.log.7.gz
```

Import (quotidien) dans la bdd

```
(prod)root@audimat1:/var/log/remote/intranet1# cat /etc/logrotate.d/remote-matomo
/var/log/remote/intranet1/intranet-local2.log {
    daily
    [...]
    prerotate
    if [ -r /data/sources/matomo/misc/log-analytics/import_logs.py ]; then
        chgrp www-data "$1"
        sudo -u www-data python3 ../../matomo/misc/log-analytics/import_logs.py --idsite=1 --url=https://ma....fr "$1"
    fi
    endscript
}
/var/log/remote/serveur?/local1.log {
    [...]
}
```

Installation et spécificités USMB

- Purges périodiques

```
(prod)root@audimat1:/var/log/remote/intranet1# cat /etc/cron.daily/matomo
#!/bin/sh -e
# Garder 3 mois d'activités web détaillées dans la BDD
DATE=`date +%Y-%m-%d -d "now -3 month"`

sudo -u www-data php /data/sources/matomo/console core:archive \
--force-all-websites --url=https://ma....fr >> /var/log/matomo-core-archive.log 2>&1
sudo -u www-data php /data/sources/matomo/console core:delete-logs-data \
--dates="2020-09-01,$DATE" --optimize-tables --no-interaction >> /var/log/matomo-core-delete.log 2>&1
```

Matomo : usages pratiques

Qu'utilise-t-on réellement ? que demande les services ? que manque-t-il ?

Matomo usages pratiques

Qu'utilise-t-on réellement ? que demande les services ? que manque-t-il ?

Éléments de contexte

Webmaster à la direction de la communication

Gestion des interfaces et du contenu du site de l'USMB

Création de sites internet pour différentes entités de l'USMB

"Composantes, laboratoires, services, projets"

Assistance sur les 20 sites internet déjà créés

Audience internet à la demande

Matomo usages pratiques

Qu'utilise-t-on réellement ? que demande les services ? que manque-t-il ?

Qu'utilise-t-on réellement ?

La segmentation

Indicateurs de visites brut

Indicateurs de téléchargements

Les indications annexes sur la provenance des usagers (public cible)

Les mots clés utilisés pour les recherches

Flux entrant et sortant sur un endroit du site (étude du comportement)

- **Étude des zones cliquables page par page (Traçage pour AB testing)**
- **Gestion des utilisateurs**

Matomo usages pratiques

Qu'utilise-t-on réellement ? que demande les services ? que manque-t-il ?

Que demande les services ?

Rapport ponctuel sur un fichier ou une page cible

Rapport d'indicateurs sur une période donnée

Audit d'un site web pour évolution prochaine

Matomo usages pratiques

Qu'utilise-t-on réellement ? que demande les services ? que manque-t-il ?

Que manque-t-il ?

Les indications précises sur le public cible

Données démographiques

- **Rapports graphiques automatiques**

Merci pour votre attention

(prod)root@audimat1:~# ./show-matomo-table-line-count.sh

```
====Tables de plus de 1000 lignes
4537360 usmb_log_link_visit_action
1710381 usmb_log_visit
146908 usmb_log_action
13107 usmb_archive_blob_2022_05
12862 usmb_archive_blob_2022_09
12774 usmb_archive_blob_2022_07
12558 usmb_archive_blob_2022_08
12383 usmb_archive_blob_2022_06
10243 usmb_archive_blob_2022_04
8990 usmb_archive_blob_2022_01
8133 usmb_archive_blob_2022_03
7786 usmb_archive_blob_2021_12
7760 usmb_archive_blob_2021_11
7531 usmb_archive_blob_2021_10
7124 usmb_archive_blob_2022_02
5847 usmb_archive_numeric_2022_05
5500 usmb_archive_numeric_2022_09
5466 usmb_archive_blob_2022_10
5406 usmb_archive_numeric_2022_06
5262 usmb_archive_numeric_2022_07
5160 usmb_archive_numeric_2022_08
4786 usmb_archive_blob_2021_09
4441 usmb_session
```

```
=== Poids des fichiers de plus d'1 Mio ===
760 Mio
/var/lib/mysql/matomo/usmb_log_link_visit_action.ibd
648 Mio /var/lib/mysql/matomo/usmb_log_visit.ibd
96 Mio /var/lib/mysql/ib_logfile0
76 Mio /var/lib/mysql/ibtmp1
76 Mio /var/lib/mysql/ibdata1
44 Mio
/var/lib/mysql/matomo/usmb_archive_blob_2022_01.ibd
40 Mio
/var/lib/mysql/matomo/usmb_archive_blob_2022_09.ibd
36 Mio /var/lib/mysql/matomo/usmb_log_action.ibd
36 Mio
/var/lib/mysql/matomo/usmb_archive_blob_2022_05.ibd
36 Mio
/var/lib/mysql/matomo/usmb_archive_blob_2022_04.ibd
36 Mio
/var/lib/mysql/matomo/usmb_archive_blob_2022_03.ibd
36 Mio
/var/lib/mysql/matomo/usmb_archive_blob_2021_11.ibd
32 Mio
/var/lib/mysql/matomo/usmb_archive_blob_2022_07.ibd
31 Mio
/var/lib/mysql/matomo/usmb_archive_blob_2021_10.ibd
30 Mio
/var/lib/mysql/matomo/usmb_archive_blob_2022_08.ibd
29 Mio
```

(prod)root@audimat1:~# cat ./show-matomo-table-line-count.sh

```
#!/bin/bash
echo "=== Tables de plus de 1000 lignes ==="
mysql -BN matomo <<< "SHOW TABLES;" \
| sed -ne 's/^\.*$/SELECT COUNT(*) AS count, "\0" AS name FROM \0;/p'\
| mysql -BN matomo\
| sort -rn\
| while read count table; do
if [ $count -gt 1000 ]; then
    echo $count $table
fi
done
echo
echo "=== Poids des fichiers de plus d'1 Mio ==="
find /var/lib/mysql -type f -printf '%s %p\n'\
| sort -rn\
| while read size path; do
size_mio=$((size/1024/1024))
if [ $size_mio -gt 0 ]; then
    echo $size_mio Mio $path
fi
done
```