

AGIMUS-NG OU LES INDICATEURS D'USAGE DE NOS SERVICES NUMERIQUES

Le projet Agimus a débuté en 2010 sous la tutelle de la MINES (Mission Numérique pour l'Enseignement Supérieur) et a pour but de créer une plateforme logicielle permettant de produire des indicateurs statistiques relatifs à l'utilisation des services numériques accessibles via des applicatifs web (dans un premier temps), cette solution se basant sur l'utilisation des fichiers de traces des dites applications.

De nos jours, savoir qu'un service est plus ou moins utilisé n'est pas suffisant pour prendre des décisions d'évolutions, il est nécessaire de savoir qui l'utilise (notion de profilage), quand il est utilisé (intervalle de temps jour, mois, vacances, week-end, ...) et idéalement de quel manière le service est utilisé (périphérique de consultation, smartphones, salle libre accès, BYOD, utilisation du wifi de l'établissement, réseaux filaire en interne, depuis l'extérieur, etc.).

Fin 2014 la MINES a confié au consortium ESUP-Portail l'évolution d'Agimus (Agimus-NG) et son accompagnement au sein des établissements de l'Enseignement Supérieur et la Recherche (ESR).

Dans le domaine sensible des indicateurs, le projet porte une attention toute particulière aux contraintes suivantes :

- besoin d'anonymisation des données ;
- nécessité de pouvoir s'insérer dans des environnements informatiques hétérogènes ;
- un déploiement Agimus-NG restant le moins intrusif possible dans les Systèmes d'Information (SI) tout en les utilisant comme référentiel ;
- le développement est réalisé de manière communautaire, en « open source » sous licence libre.

LES CHOIX ARCHITECTURAUX

Agimus-NG a radicalement changé d'orientation technologique par rapport à la précédente version.

Le trio ELK : Logstash (<http://logstash.net/>), Elasticsearch (<https://www.elastic.co/products/elasticsearch>) et Kibana (<https://www.elastic.co/products/kibana>) a été choisi comme socle à cette plateforme.

Agimus-NG fonctionne sur 4 principes fondamentaux : utilisation de fichier de traces, profilage et enrichissement, constitution d'un entrepôt de données central et restitution sous forme graphique.

UTILISATION DE FICHIER DE TRACES

Les formats de fichiers peuvent avoir différentes formes et peuvent selon le cas être traités en temps réel, ou une fois par jour. L'outil Logstash est la brique qui permet l'ingestion des fichiers de traces. Sa modularité permet de définir son propre format de ligne et d'extraire tout ou partie d'un enregistrement.

Le gros enjeu en terme d'indicateurs et de pouvoir savoir « **qui** fait quoi » et pour ce faire Agimus-NG croise les fichiers de traces afin de trouver l'identité de l'utilisateur réalisant les actions. Deux cas se présentent :

- cas simple : l'application peut elle-même écrire l'identifiant de l'utilisateur dans les fichiers de traces, Agimus-NG utilisera directement cet identifiant ;
- cas le plus courant : on utilise des fichiers de traces extérieurs à l'application ne contenant pas l'identifiant : apache ou nginx par exemple (ce qui permet de limiter l'intrusion dans le SI). Afin de pouvoir associer une trace à un utilisateur, il faut légèrement modifier le système d'authentification (SSO CAS, qui est le système le plus répandu dans l'ESR) afin qu'à chaque authentification soit généré

un cookie opaque (ne portant aucune information sur l'utilisateur) qui sera enregistré au niveau du serveur CAS et au niveau des traces apache (changement de format de log). Ensuite, Agimus-NG fera le recoupement de ces fichiers afin de résoudre l'association valeur du cookie / identifiant de la personne, ce travail se fera lors du traitement des enregistrements.

PROFILAGE ET ENRICHISSEMENT

Agimus-NG permet d'enrichir les traces avec différentes informations.

Les informations relatives à l'identité de la personne sont indispensables à des fins de profilage et de regroupements statistiques. Agimus-NG permet de récupérer des attributs d'un annuaire LDAP en utilisant l'identifiant de l'utilisateur comme clé. Des attributs tels que l'affectation (composante/service de rattachement attribut supannEntiteAffectation par exemple) ou l'affiliation (type de personne student, employee, etc ... attribut eduPersonAffiliation par exemple) pourront servir aux regroupements. Agimus-NG pourra aussi être configuré pour utiliser des attributs spécifiques (hors supann) contenant un intérêt de groupement pour l'établissement.

Les fichiers de traces contenant souvent les adresses IP, ces dernières peuvent être exploitées pour faire de la géolocalisation, mais également pour déterminer si l'adresse d'accès est interne ou externe à l'établissement, via réseaux filaires ou via le Wifi.

Et enfin, si on exploite des logs de serveurs web avec Agimus-NG on pourra enrichir des informations en travaillant sur le « user-agent » de l'utilisateur, ce qui permettra d'obtenir des informations sur le navigateur, le système d'exploitation (Windows, Mac OS, Linux mais aussi Android, iOS ou Windows Phone) ou encore le type de périphérique utilisé.

CONSTITUTION D'UN ENTREPOT DE DONNEES CENTRAL

Toutes ces informations ainsi collectées et enrichies doivent être stockées dans un système fiable et permettant des recherches rapides. Elasticsearch est la brique choisie pour assumer le rôle de l'entrepôt des données d'Agimus-NG. Le coté évolutif d'ElasticSearch permet d'envisager une montée en charge et en volume sans encombre.

Logstash va permettre de déverser directement les données dans Elasticsearch. Cet entrepôt contiendra donc l'intégralité des informations enrichies d'Agimus-NG, il pourra ainsi être utilisé par Agimus-NG mais aussi comme source d'information par d'autres applications dans d'autres contextes (OpenData).

RESTITUTION

La restitution des indicateurs sera assurée par l'outil Kibana. Il permet de créer des graphiques montrant des évolutions temporelles mais pas uniquement, il permet de traiter des graphiques pour tout type de groupement (affectation, type de personne, type de périphériques, etc.). Agimus-NG permet donc de créer différentes visualisations graphiques : camembert, graphique en bar, table de données, graphique en ligne, carte géographique, etc.

Une interface WYSIWYG permet de créer en temps réel des graphiques basés sur l'entrepôt. Mais toute personne créant des graphiques dans Agimus-NG, et les trouvant suffisamment génériques, pourra les exporter afin de les partager avec la communauté à travers un entrepôt Git (<https://github.com/EsupPortail/agimus-ng>).

Les graphiques seront rassemblés en tableau de bord afin de leur donner une cohérence commune et ces tableaux de bord peuvent eux même être exportés et partagés.

LES PREMIERS RESULTATS

A ce jour, le projet progresse encore : quatre universités (université de Lorraine, université de Lille 1, université Paris 1 et université François Rabelais de Tours) ont constitué leur entrepôt de données. Cet entrepôt est basé sur l'exploitation des logs de leur serveur SSO CAS, ENT (Esup-Portail version 3 et 4) et du LMS Moodle, le système permettant facilement d'ajouter de nouvelles sources d'information.

Les graphiques fournis permettent d'avoir une vision d'ensemble pour un service spécifique ou pour un ensemble de services. Il est par exemple possible d'obtenir :

- L'utilisation des services authentifiés par le SSO CAS dans le temps, groupé par type de personne (étudiant, personnel, chercheur, ...) groupé par affectation (Laboratoire X, Direction Y, UFR Z, ...)
- Répartition des connexions par catégories de personne et/ou affectation sur le LMS Moodle ou sur l'Espace Numérique de Travail (ENT)
- Nombre de personnes uniques (dé-doublonnées) s'étant connectées par jour sur le LMS ou sur l'ENT
- ...

Ce projet est le fruit du travail de plusieurs universités alliées au consortium esup-portail. Le groupe a fait le choix de mettre en place une méthode de travail agile qui permet une constante évolution et nous utilisons les outils suivants :

- Github : pour le partage des codes et configurations [<https://github.com/EsupPortail/agimus-ng>]
- Trello : afin de gérer les tâches et les « sprints » d'évolution [<https://trello.com/b/jp1aRH9q/suivi-de-projet-agimus-ng>]
- Wiki Esup-portail : afin de gérer la documentation [<https://www.esup-portail.org/wiki/x/DQCfFg>]
- Liste de diffusion : afin de gérer les échanges sur le projet [<https://listes.esup-portail.org/sympa/info/esup-utilisateurs>]