esup portail

15 OCT. 2018 | Université PARIS Descartes

ESUP-DAYS 26

CENTRALISATION DES LOGS ET DÉTECTION D'ANOMALIES PAR APPRENTISSAGE AUTOMATIQUE

Sanghun Bang -- Université Paris 8 Vincennes-Saint-Denis
Yasmina Bensitel -- Université Paris 8 Vincennes-Saint-Denis

ESUP-DAYS 26

1. INTRODUCTION
2. ANOMALIE
3. ARCHITECTURE
4. MÉTHODOLOGIE
5. ANALYSE DES LOGS (SERVEUR CAS)
6. ANALYSE DES LOGS (SERVEUR MAIL)
7. CONCLUSION ET TRAVAUX FUTURS

APERÇU DE LA PRÉSENTATION

esup portail

ESUP-DAYS 26

Centralisation des logs et détection d'anomalies par apprentissage automatique.

INTRODUCTION

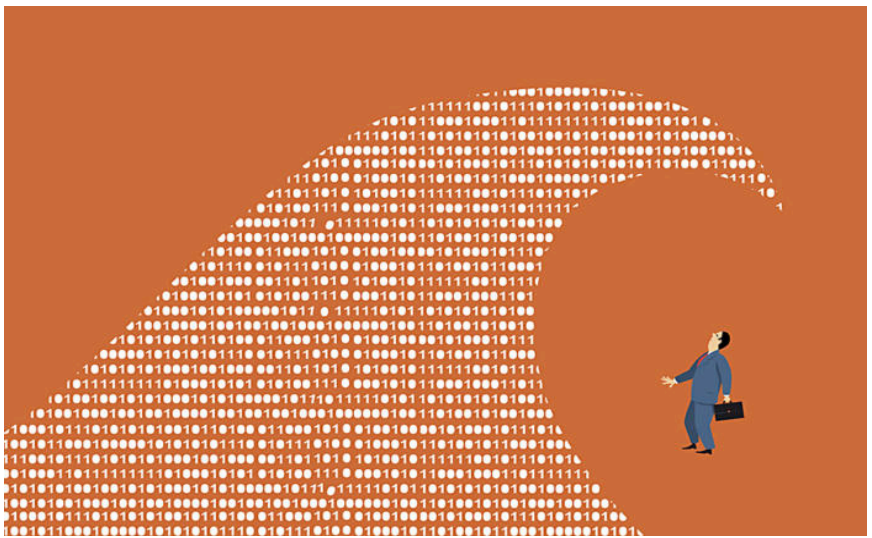
esup portail

INTRODUCTION

- UNIVERSITÉ PARIS 8 EN CHIFFRES
  - 22 045 Etudiants
  - 14 032 inscrits en licence, licence pro et DUT
  - 5 982 inscrits en master
  - 1 130 inscrits en doctorat
  - 888 inscrits dans des diplômes hors LMD
  - 899 enseignants (tous grades confondus)
  - 868 emplois de personnels BIATOSS

DÉLUGE D'INFORMATIONS

- Explosion, tsunami ou déluge de données
- Les événements anormaux se produisent relativement peu fréquemment.
- MAIS, leurs conséquences peuvent être assez dramatiques et assez souvent dans un sens négatif.



OBJECTIF

- La détection de comportements offensifs dans un système d'information :
  - L'objectif est la prévention, l'estimation des risques et la correction après avoir détecté l'anomalie.
  - Centralisation des logs vers une machine dédiée
  - Réalisation d'un système de détection d'anomalies à partir de ces logs par l'apprentissage automatique et la méthode statistique .

- Analyse des logs de différents serveurs
  - Logs de CAS et logs de mail
  - Utilisation des méthodes : Lissage exponentiel (Logs de CAS)
    - 10,110 Méga/jour, 31 860 lignes
  - Utilisation des méthodes : Apprentissage automatique (Logs de mail)
    - 242 Méga/jour, 1 548 594 lignes

ESUP-DAYS 26

Centralisation des logs et détection d'anomalies par apprentissage automatique.

ANOMALIE

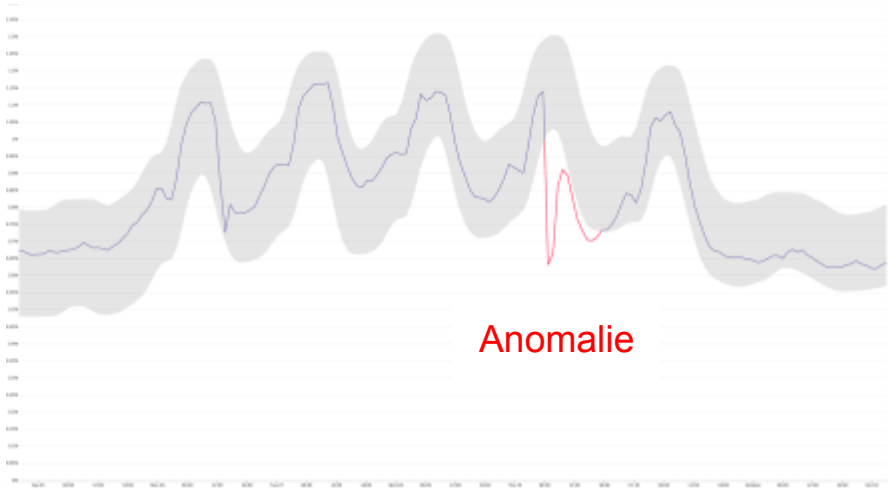
esup portail

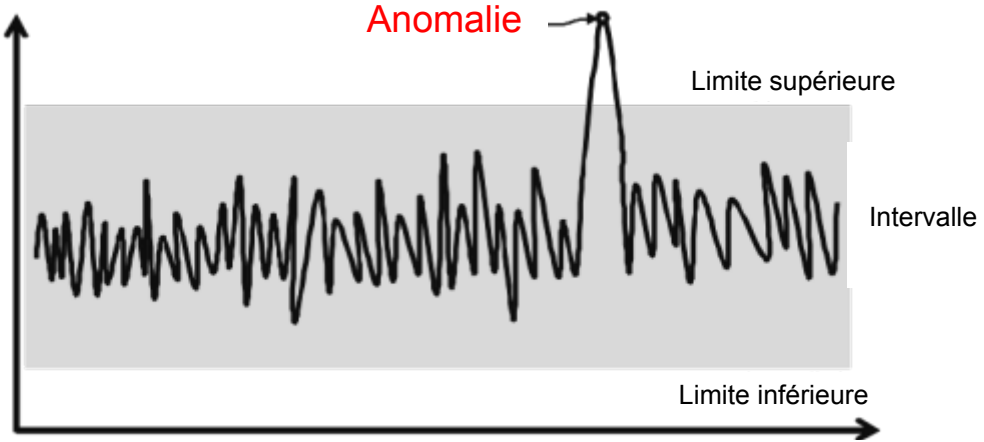
ANOMALIE: QU'EST-CE QUE C'EST ?

- La détection d'anomalie peut être utilisée dans plusieurs domaines :
  - Piratage informatique
  - Cyber-intrusion
  - Fraude de carte de crédit
  - Détection de spams
  - Etc.

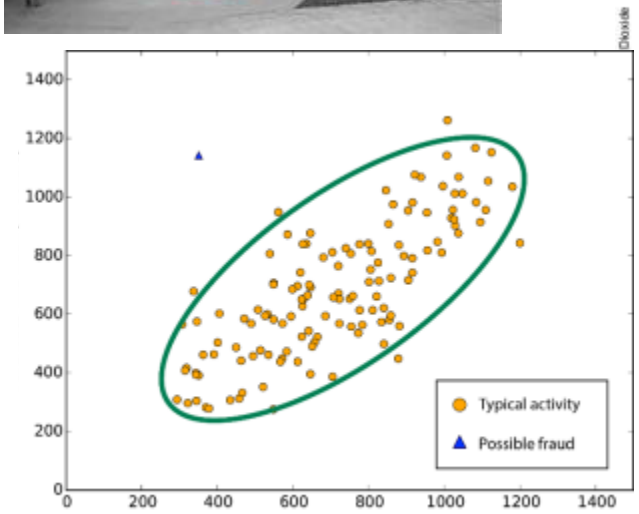
- Détecter différents types d'anomalies dans les flux d'événements.
  - Changement au niveau bidirectionnel: Une augmentation ou une diminution soutenue du niveau des valeurs, vers le haut comme vers le bas.
  - Tendance positive lente: Une lente augmentation de la tendance au fil du temps.
  - Tendance négative lente: Une lente diminution de la tendance au fil du temps.

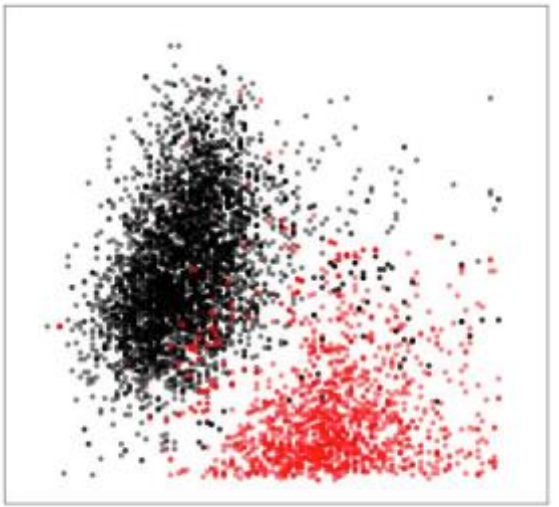
ANOMALIES COLLECTIVE DANS LES SÉRIES TEMPORELLES





DÉTECTION. CLASSIFICATION ET IDENTIFICATION D'ANOMALIES





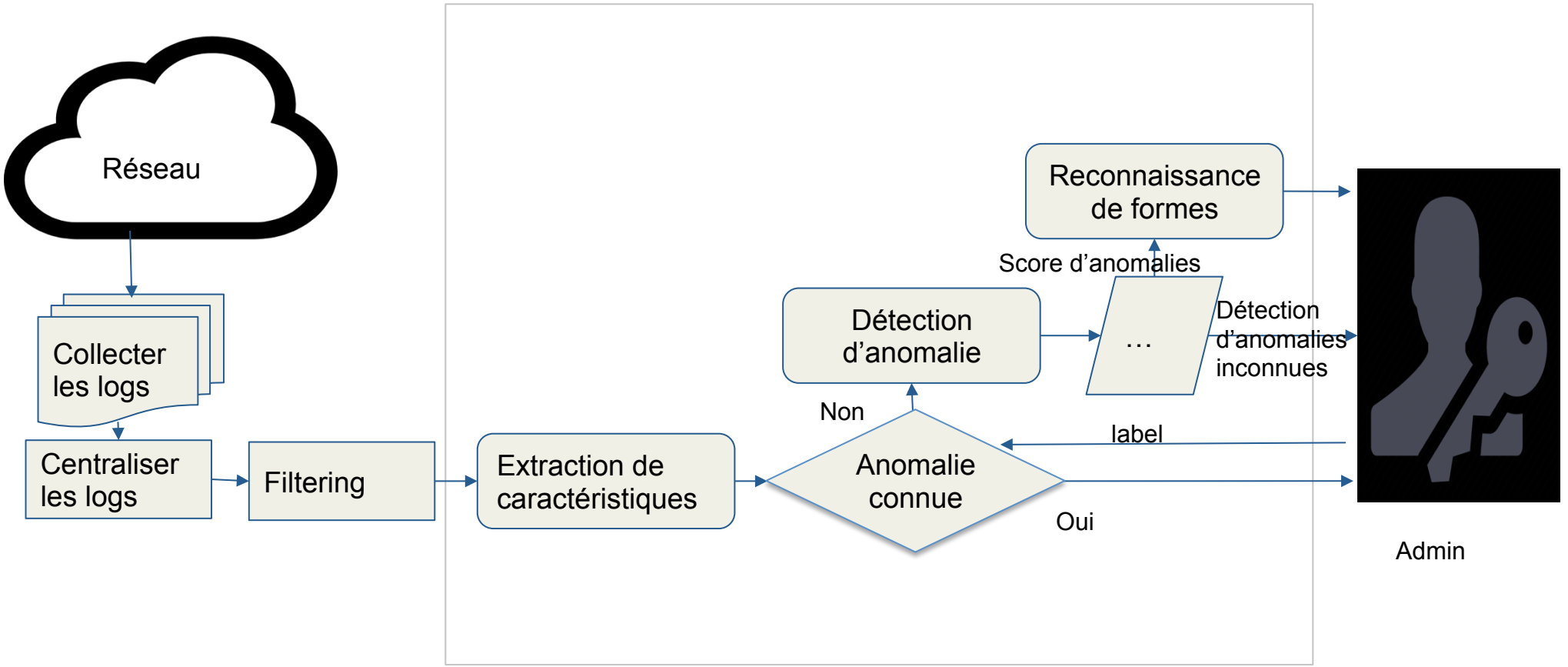
ESUP-DAYS 26

Centralisation des logs et détection d'anomalies par apprentissage automatique.

ARCHITECTURE

esup portail

ARCHITECTURE DE DÉTECTION D'ANOMALIES

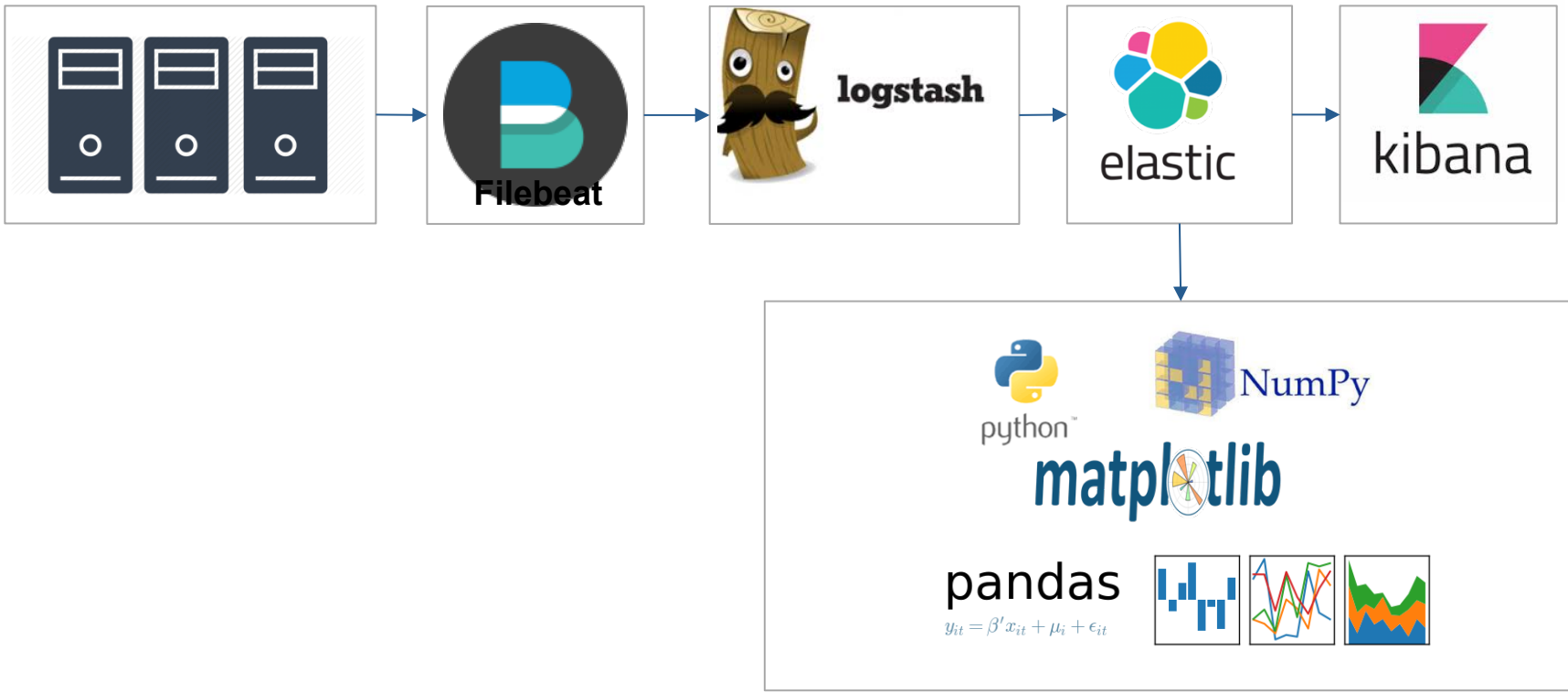


ARCHITECTURE: ANALYSE DES LOGS(CAS)

- Architecture

ARCHITECTURE: ANALYSE DES LOGS(MAIL)

- Architecture



ESUP-DAYS 26

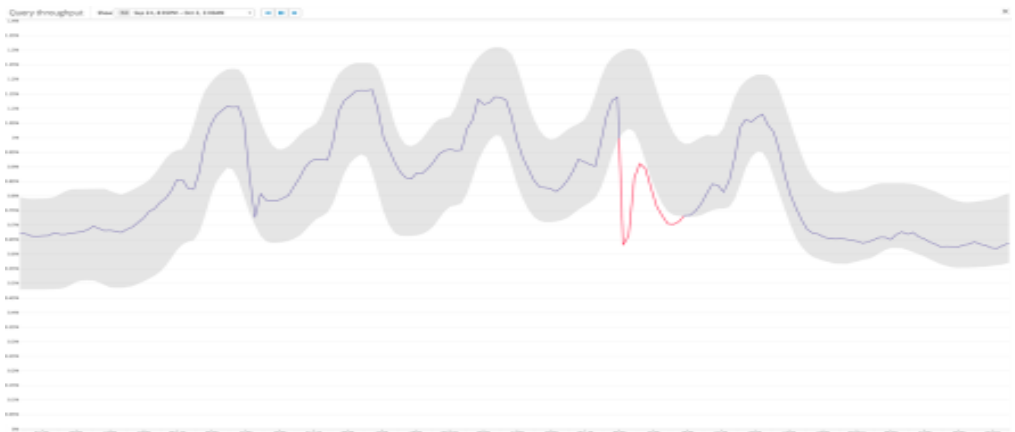
Centralisation des logs et détection d'anomalies par apprentissage automatique.

MÉTHODOLOGIE

esup portail

MÉTHODE DE LISSAGE EXPONENTIEL

- Méthode Technique de lissage exponentiel :
  - Méthode de prévision de données chronologiques : chaque donnée est lissée successivement en partant de la valeur initiale.
  - Les méthodes de lissage exponentiel sont des méthodes qui supposent que l'analyse dépend de ses valeurs passées.



MÉTHODE DE LISSAGE EXPONENTIEL

- On appelle lissage exponentiel triple (Holt-Winters).
  Paramètres α, β et δ ∈[0,1] de $\widehat{y}_t$ cette série le processus est défini ainsi:

$$\widehat{y}_{t+h} = (l_t + hb_t)s_t$$

Prévision

$$\begin{cases} l_t = \alpha \frac{y_t}{s_{t-T}} + (1-\alpha)(l_{t-1} + b_{t-1}) \\ b_t = \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1} \\ s_t = \delta \frac{y_t}{l_t} + (1-\delta)s_{t-T} \end{cases}$$

Valeurs lissées

Tendance

Saisonnalité

APPRENTISSAGE AUTOMATIQUE

- Apprentissage supervisé
  - L'apprentissage supervisé est un type d'apprentissage automatique qui consiste à apprendre une fonction de prédiction à partir d'exemples annotés.
  - L'algorithme
    - Méthode des k plus proches voisins
    - Classification naïve bayésienne
    - L'arbre de décision
    - Machine vecteurs de support
    - Réseau de neurones (Deep Learning)

*https://fr.wikipedia.org/wiki/Apprentissage_supervis%C3%A9

APPRENTISSAGE AUTOMATIQUE

- Apprentissage non-supervisé
  - L'apprentissage non supervisé est un problème d'apprentissage automatique. Il s'agit, pour un logiciel, de trouver des structures sous-jacentes à partir de données non étiquetées.
  - Réduction de la dimensionnalité et l'extraction des caractéristiques
    - Analyse en Composantes Principales (ACP)
    - NMF Factorisation par matrices non négatives.
    - L'algorithme t-SNE (t-distributed stochastic neighbor embedding)
  - Clustering
    - Algorithme K-Moyennes
    - DBSCAN
    - Isolation Forest algorithm

* https://fr.wikipedia.org/wiki/Apprentissage_non_supervis%C3%A9

ESUP-DAYS 26

Centralisation des logs et détection d'anomalies par apprentissage automatique.

ANALYSE DES LOGS (SERVEUR CAS)

esup portail

DESCRIPTION DES DONNEES DU SERVEUR CAS

[Fri Oct 05 00:00:20 CEST 2018] [IP:86.252.**.**] [ID:aal***laigre] [TICKET:ST-367710-gzdYzTBpo9csrAyy7S1t-cas.univ-paris8.fr] [SERVICE:https://e-p8.univ-paris8.fr/uPortal/Login] [USER-AGENT:Mozilla/5.0 (Linux; Android 7.0; SM-J530F Build/NRD90M) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/69.0.3497.100 Mobile Safari/537.36]

...

[Fri Oct 05 00:00:26 CEST 2018] [IP:78.234.**.***] [ID:mpo**uvreau] [TICKET:ST-367711-5YsZsaUHPFidUm3ijGoK-cas.univ-paris8.fr] [SERVICE:https://e-p8.univ-paris8.fr/uPortal/Login] [USER-AGENT:Mozilla/5.0 (Macintosh; Intel Mac OS X 10_12_6) AppleWebKit/605.1.15 (KHTML, like Gecko) Version/12.0 Safari/605.1.15]
[Fri Oct 05 00:00:26 CEST 2018] [IP:78.250.***.**] [ID:lkyb**aliuk] [TICKET:ST-367712-4sOI2QJmsFRYqHukf7Wg-cas.univ-paris8.fr] [SERVICE:https://moodle.univ-paris8.fr/moodle/login/index.php?authCAS=CAS] [USER-AGENT:Mozilla/5.0 (Windows NT 6.2; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/69.0.3497.100 Safari/537.36]
[Fri Oct 05 00:00:29 CEST 2018] [IP:82.226.**.***] [ID:ffrix***ione] [TICKET:ST-367713-V1V5SHfIqkWpgoDj9710-cas.univ-paris8.fr] [SERVICE:https://zimbra.univ-paris8.fr/public/preauth_edu.jsp] [USER-AGENT:Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/69.0.3497.100 Safari/537.36]
[Fri Oct 05 00:00:33 CEST 2018] [IP:86.252.***.**] [ID:aal***laigre] [TICKET:ST-367714-NFLAFskbtPhzWbgfK12x-cas.univ-paris8.fr] [SERVICE:https://zimbra.univ-paris8.fr/public/preauth.jsp] [USER-AGENT:Mozilla/5.0 (Linux; Android 7.0; SM-J530F Build/NRD90M) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/69.0.3497.100 Mobile Safari/537.36]
[Fri Oct 05 00:00:35 CEST 2018] [IP:78.234.**.***] [ID:mpouvreau] [TICKET:ST-367715-BFC4kkLnDHLuobSRwG5K-cas.univ-paris8.fr] [SERVICE:https://zimbra.univ-paris8.fr/public/preauth_edu.jsp] [USER-AGENT:Mozilla/5.0 (Macintosh; Intel Mac OS X 10_12_6) AppleWebKit/605.1.15 (KHTML, like Gecko) Version/12.0 Safari/605.1.15]
[Fri Oct 05 00:00:37 CEST 2018] [IP:46.193.**.**] [ID:ama***nhica] [TICKET:ST-367716-SD4Yigced4aMito65017-cas.univ-paris8.fr] [SERVICE:https://e-p8.univ-paris8.fr/uPortal/f/u64l1s27/p/MessagerieZimbraEtudiant.u64l1n1091/max/render.uP?pCp] [USER-AGENT:Mozilla/5.0 (Macintosh; Intel Mac OS X 10_12_2) AppleWebKit/602.3.12 (KHTML, like Gecko) Version/10.0.2 Safari/602.3.12]

EXEMPLE: ANALYSE DES LOGS (SERVICESTATS)

- Visualisation des logs 'servicestats'

Timelion commence avec une fonction identifiant la source de données (.es(*))

```
$thres=5000,

.es(index='servicestats').color(#666).lines(1).label(Actual),

.es(index='servicestats').lines(1).if(eq, 0, null).holt(0.3, 0.2, 0.5,
1d).color(green).lines(2).label('Prediction'),

.es(index='servicestats').lines(1).if(eq, 0, null).holt(0.3, 0.2,
0.5,1d).subtract(.es(index='servicestats')).abs().if(lt, $thres,
null, .es(index='servicestats')).points(10,3,0).color(red).label('Anomaly').title('
Lissage exponentiel triple : CAS')
```

EXEMPLE: ANALYSE DES LOGS (SERVICESTATS)

- Visualisation des logs 'servicestats'

$thres=5000,

.es(index='servicestats').color(#666).lines(1).label(Actual),

.es(index='servicestats').lines(1).if(eq, 0, null).holt(0.3, 0.2, 0.5, 1d).color(green).lines(2).label('Prediction'),

.es(index='servicestats').lines(1).if(eq, 0, null).holt(0.3, 0.2, 0.5,1d).subtract(.es(index='servicestats')).abs().if(lt, $thres, null, .es(index='servicestats')).points(10,3,0).color(red).label('Anomaly').title('Lissage exponentiel triple : CAS')

α∈[0,1] : Poids de lissage
β∈[0,1]: Poids de tendance
δ∈[0,1]:Poids saisonnier
season (1d,1w,1s,5h,etc…)

EXEMPLE: ANALYSE DES LOGS (SERVICESTATS)

- Visualisation des logs 'servicestats'

| valeur de la prédiction - valeur actuelle | < Seuil (threshold)

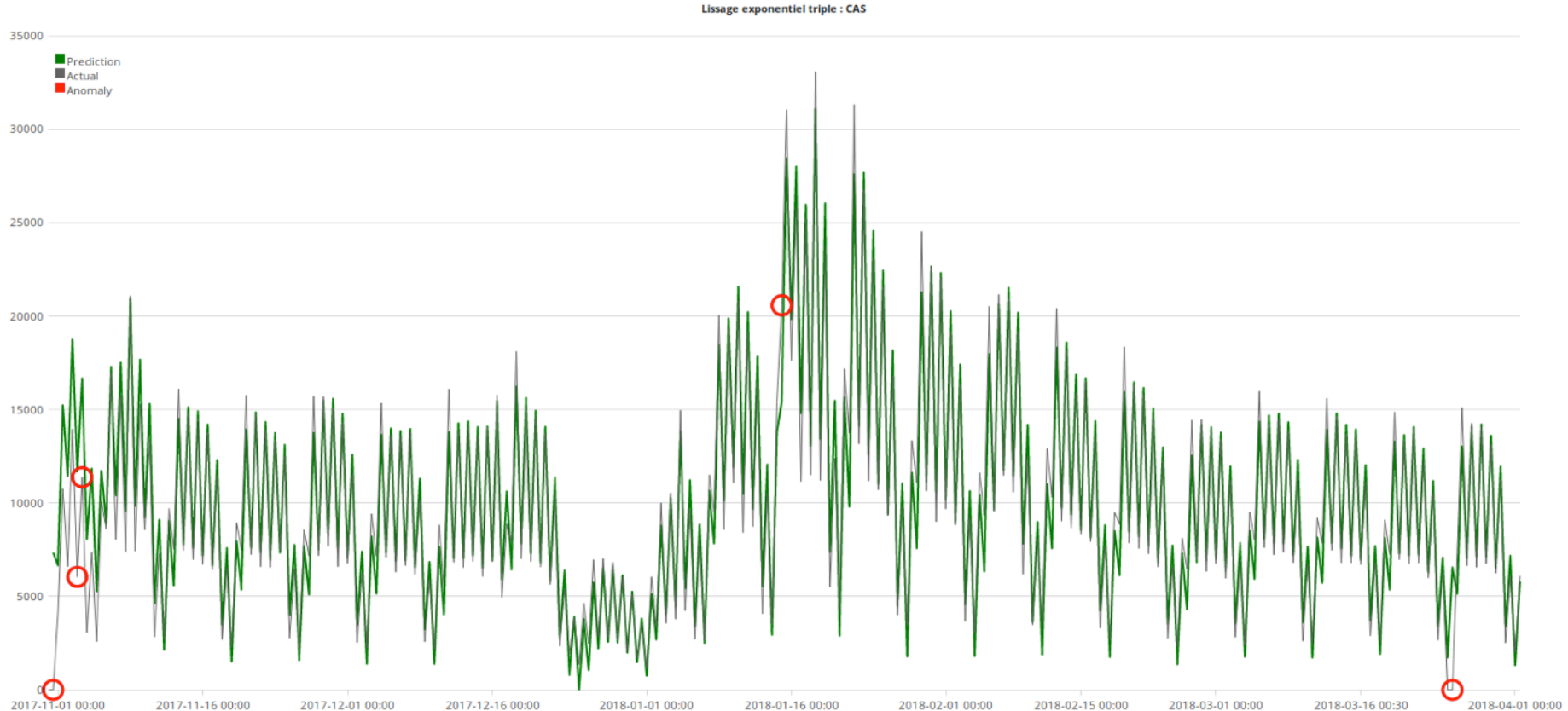
```
$thres=5000,

.es(index='servicestats').color(#666).lines(1).label(Actual),

.es(index='servicestats').lines(1).if(eq, 0, null).holt(0.3, 0.2, 0.5,
1d).color(green).lines(2).label('Prediction'),

.es(index='servicestats').lines(1).if(eq, 0, null).holt(0.3, 0.2,
0.5,1d).subtract(.es(index='servicestats')).abs().if(lt, $thres,
null, .es(index='servicestats')).points(10,3,0).color(red).label('Anomaly').title('
Lissage exponentiel triple : CAS')
```

EXEMPLE: ANALYSE DES LOGS (SERVICESTATS)

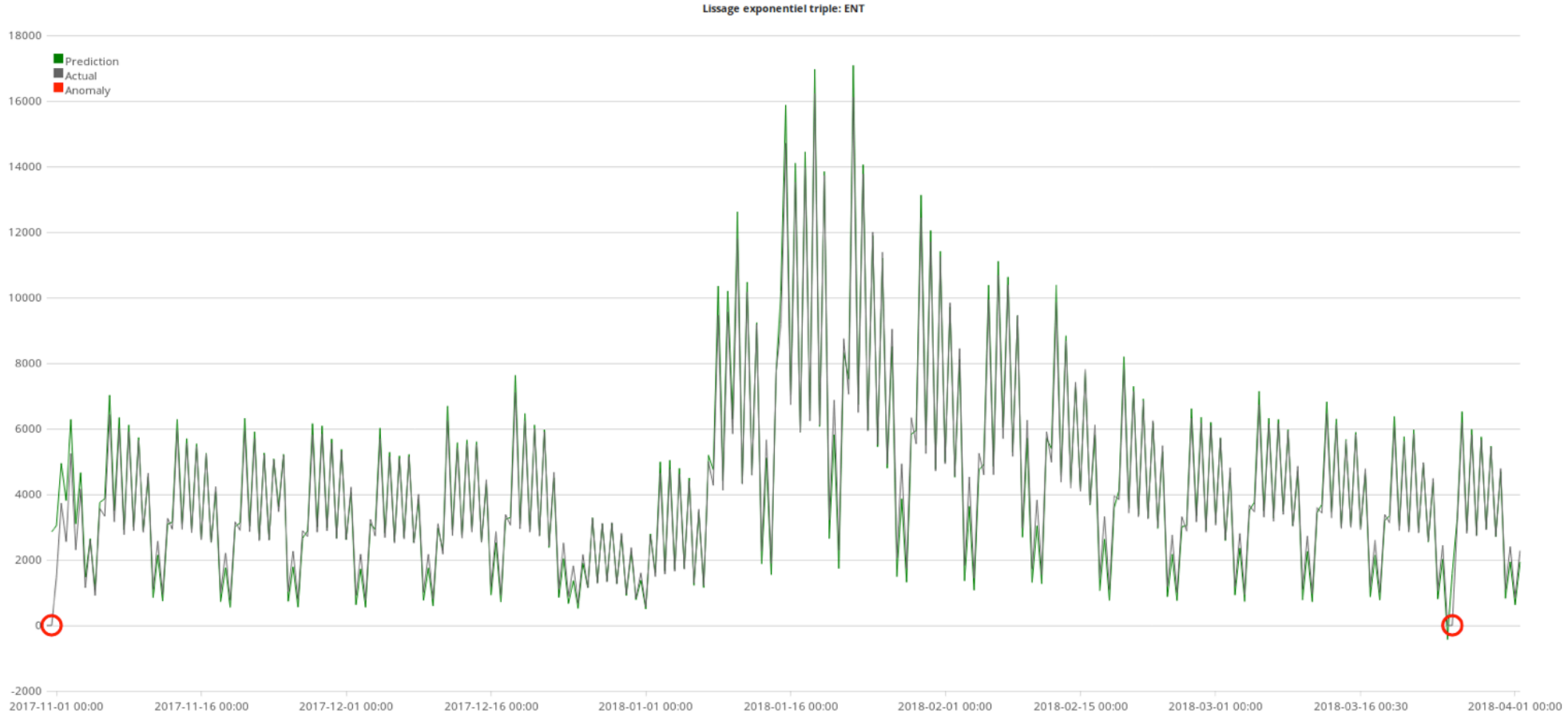


EXEMPLE: ANALYSE DES LOGS (SERVICESTATS avec nom de domaine ENT)

- Visualisation des logs 'servicestats' avec nom de domaine e-p8 (ENT)

```
$thres=2000,
.es(index='servicestats',q='message:e-p8').lines(1).if(eq, 0, null).holt(0.9, 0.2, 0.9,
1d).color(green).lines(1).label('Prediction'),
.es(index='servicestats',q='message:e-p8').color(#666).lines(1).label(Actual),
.es(index='servicestats',q='message:e-p8').lines(1).if(eq, 0, null).holt(0.3, 0.2, 0.9,
1d).subtract(.es(index='servicestats',q='message:e-p8')).abs().if(lt, $thres,
null, .es(index='servicestats',q='message:e-p8')).points(10,3,0).color(red).label('Anomaly').title('
Lissage exponentiel triple: ENT'),
```

EXEMPLE: ANALYSE DES LOGS (SERVICESTATS avec nom de domaine ENT)

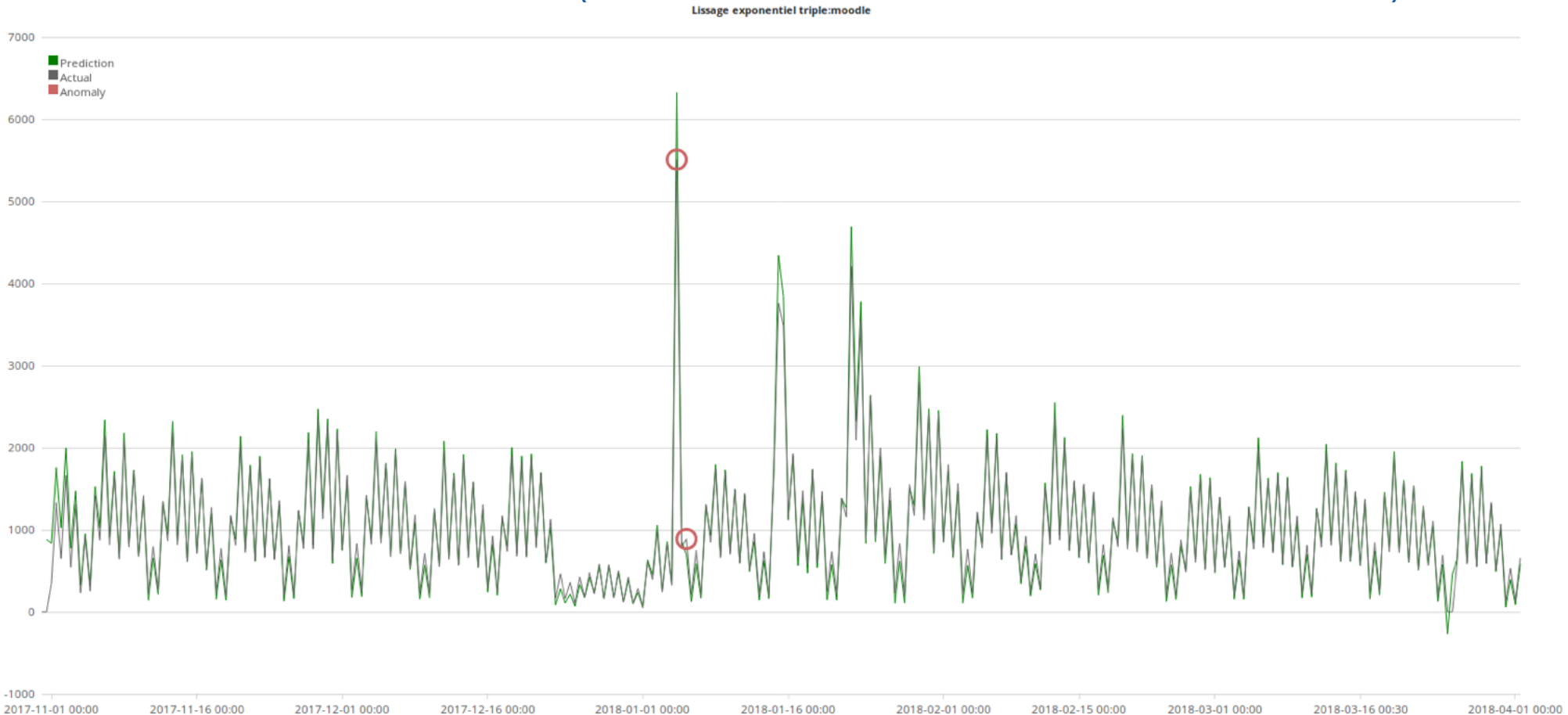


EXEMPLE: ANALYSE DES LOGS (SERVICESTATS avec nom de domaine moodle)

- Visualisation des logs 'servicestats' avec nom de domaine moodle

```
$thres=1000,
.es(index='servicestats',q='message:moodle').lines(1).if(eq, 0, null).holt(0.9, 0.2, 0.9,
1d).color(green).lines(1).label('Prediction'),
.es(index='servicestats',q='message:moodle').color(#666).lines(1).label(Actual),
.es(index='servicestats',q='message:moodle').lines(1).if(eq, 0, null).holt(0.3, 0.2, 0.5,
1d).subtract(.es(index='servicestats',q='message:moodle')).abs().if(lt, $thres,
null, .es(index='servicestats',q='message:moodle')).points(10,3,0).color(#c66).label('Anomaly').title('
Lissage exponentiel triple:moodle')
```

EXEMPLE: ANALYSE DES LOGS (SERVICESTATS avec nom de domaine moodle)



ESUP-DAYS 26

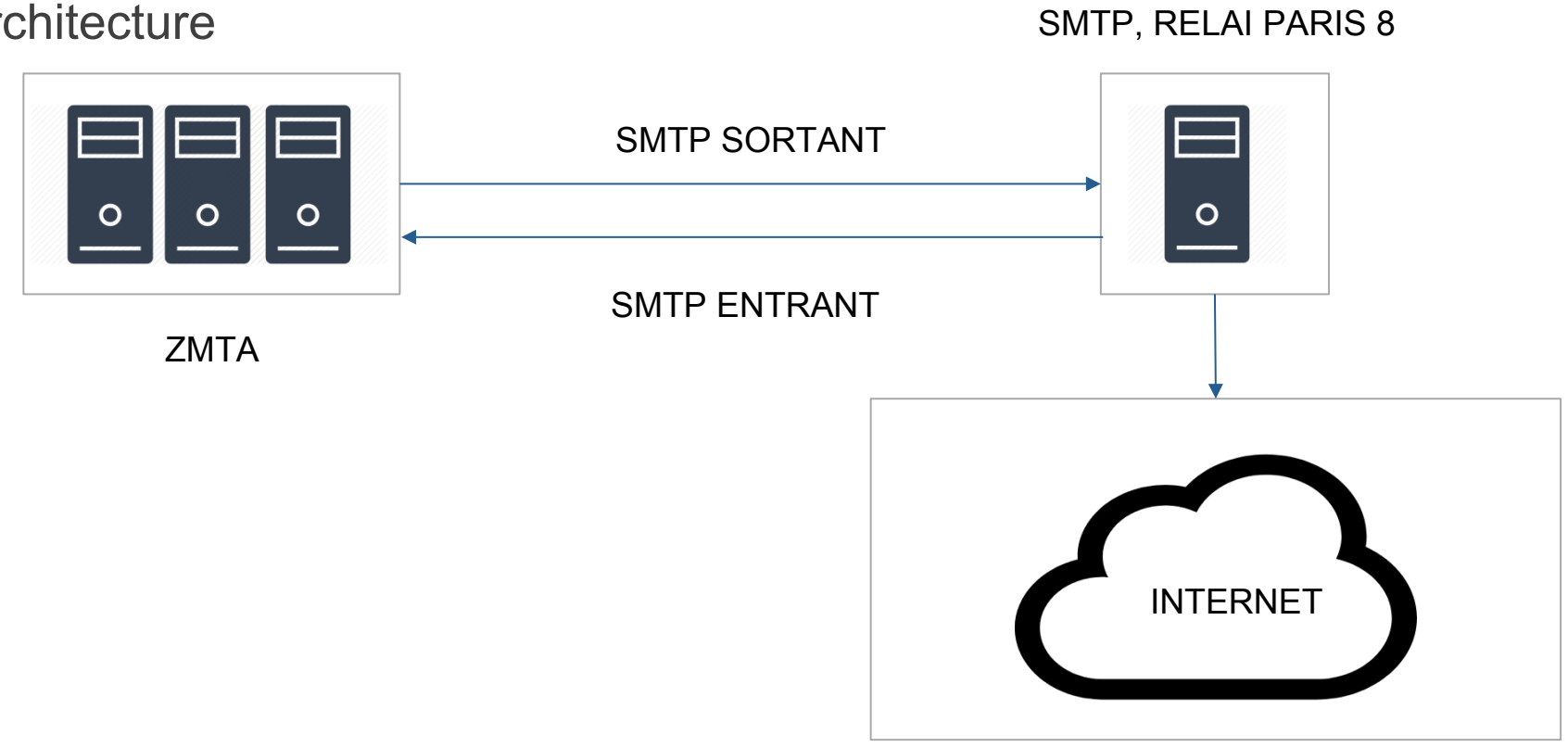
Centralisation des logs et détection d'anomalies par apprentissage automatique.

ANALYSE DES LOGS (SERVEUR MAIL)

esup portail

ARCHITECTURE: RELAI PARIS 8

- Architecture



DESCRIPTION DES DONNÉES DU SERVEUR DE MESSAGERIE

Apr 8 06:28:03 azel postfix/smtpd[22690]: warning: 14.172.**: address not listed for hostname static.vnpt.vn

Voir : L'architecture de Postfix (http://www.postfix.org/OVERVIEW.html)

Apr 8 06:28:03 azel postfix/smtpd[22690]: connect fro
Apr 8 06:28:03 azel postfix/smtpd[22595]: NOQUEUE: reject: RCPT from unknown[95.211.**]: 450 4.1.8 <bounce@promo03.thebestforyourlife.com>: Sender address rejected: Domain not found; from=<**@promo03.thebestforyourlife.com> to=<***@univ-paris8.fr> proto=ESMTP helo=<promo03.thebestforyourlife.com>
Apr 8 06:28:03 azel postfix/smtpd[22600]: NOQUEUE: reject: RCPT from unknown[95.211.**]: 450 4.1.8 <bounce@promo08.thebestforyourlife.com>: Sender address rejected: Domain not found; from=<**@promo08.thebestforyourlife.com> to=<***@univ-paris8.fr> proto=ESMTP helo=<promo08.thebestforyourlife.com>
Apr 8 06:28:03 azel postfix/smtpd[22600]: disconnect from unknown[95.211.**]
Apr 8 06:28:03 azel postfix/smtpd[22595]: NOQUEUE: reject: RCPT from unknown[95.211.**]: 450 4.1.8 <bounce@promo03.thebestforyourlife.com>: Sender address rejected: Domain not found; from=<bounce@promo03.thebestforyourlife.com> to=<***@univ-paris8.fr> proto=ESMTP helo=<promo03.thebestforyourlife.com>
Apr 8 06:28:03 azel postfix/smtpd[22595]: disconnect from unknown[95.211.**]
Apr 8 06:28:03 azel postfix/smtpd[22652]: connect from zmta2.srv.up8[192.168.**]
Apr 8 06:28:03 azel postfix/smtpd[22652]: setting up TLS connection from zmta2.srv.up8[192.168.**]
Apr 8 06:28:03 azel postfix/smtpd[22652]: Anonymous TLS connection established from zmta2.srv.up8[192.168.**]: TLSv1 with cipher ADH-AES256-SHA (256/256 bits)
Apr 8 06:28:03 azel postfix/smtpd[22652]: BFAE9B30BE: client=zmta2.srv.up8[192.168.**]
Apr 8 06:28:03 azel postfix/cleanup[22756]: BFAE9B30BE: message-id=<20180408042733.4291BB3117@azel.univ-paris8.fr>
Apr 8 06:28:03 azel postfix/smtpd[22711]: lost connection after DATA from unknown[182.187.**]

DESCRIPTION DES DONNÉES DU SERVEUR DE MESSAGERIE

- Visualisation des données sur Kibana

Time	syslog_hostname	syslog_program	syslog_pid	@timestamp	syslog_message
June 29th 2017, 10:30:00.000	azel	postfix/smtpd	11552	June 29th 2017, 10:30:00.000	warning: Connection rate limit exceeded: 102 from 66-208-242-141-ubr04a-chrstn01-pa.hfc.comcastbusiness.net[66.208.242.141] for service smtp
June 29th 2017, 10:30:00.000	azel	postfix/smtpd	10520	June 29th 2017, 10:30:00.000	connect from 66-208-242-141-ubr04a-chrstn01-pa.hfc.comcastbusiness.net[66.208.242.141]
June 29th 2017, 10:30:00.000	azel	postfix/smtpd	8902	June 29th 2017, 10:30:00.000	D8BF0C697: client=localhost[127.0.0.1]
June 29th 2017, 10:30:00.000	azel	postfix/cleanup	11661	June 29th 2017, 10:30:00.000	D8BF0C697: message-id=<S59ra2Jo.r.I1sjQxMzc6H_OraU5AAA29@hongyu.runs88.com>

ANALYSE DES LOGS (SERVEUR MAIL)

- Traitement des données du serveur de messagerie :
  - Requête sur les logs du programme smtpd
  - Sélection des attributs :
    - timestamp,
    - ip,
    - hostname,
    - status_code,
    - status_enhanced_code,
    - sasl_username

ANALYSE DES LOGS (SERVEUR MAIL)

- Traitement des données du serveur de messagerie :
  - Création de nouvelles bases de données : données groupées par périodes (30 min, 1h)
  - Calcul du résumé d'action de chaque utilisateur :
    - Nombre d'adresses IP,
    - Liste des IP,
    - Liste des noms de domaines,
    - Nombre de noms de domaines utilisés,
    - Nombre de mails envoyés,
    - Nombre d'erreur (par degré 1/2/3)
    - Nombre de messages normaux envoyés

ANALYSE DES LOGS (SERVEUR MAIL)

- Traitement des données du serveur de messagerie :
  - Enrichissement des jeux de données par les résumés d'actions des périodes précédentes :

username	nbEnvoi 'i'	nbEnvoi 'i-30min'	nbIP 'i'	nbIP 'i-30min'	liste IP 'i'	liste IP 'i-30min'	nbDomain 'i'	nbDomain 'i-30min'	liste Domain 'i'	...
459	12	0	2	0	['209.85.*.*', '209.85.*.*']	0	1	0	['google.com']	...
1283	294	314	22	25	['49.34.*.*' '47.29.*.*' '187.190.*.*' ...]	['41.13.*.*' '169.149.*.*' '223.184.*.*' ...]	3	3	['unknown' 'totalplay.net' 'co.za']	...
1323	6	0	1	0	['194.254.*.*']	0	1	0	['mshparisnord.fr']	...
1311	2	0	1	0	['92.90.*.*']	0	1	0	['sfr.net']	...
1321	2	0	1	0	['209.85.*.*']	0	1	0	['google.com']	...
1329	1	0	1	0	['109.88.*.*']	0	1	0	['voo.be']	...
1333	2	0	1	0	['117.249.*.*']	0	1	0	['unknown']	...
1343	2	0	1	0	['202.67.*.*']	0	1	0	['unknown']	...
1356	2	0	1	0	['80.200.*.*']	0	1	0	['belgacom.be']	...
1367	2	0	1	0	['193.54.*.*']	0	1	0	['unknown']	...
1392	490	232	38	18	['49.139.*.*' '27.62.*.*' '106.76.*.*' ...]	['223.184.*.*' '177.243.*.*' '42.106.*.*' ...]	2	1	['co.za' 'unknown']	...
...	...	...	...	...	...	...	...	...	...	

ANALYSE DES LOGS (SERVEUR MAIL)

- Modélisation pour la détection d'anomalies :
  - DBSCAN
    - Repose sur le concept de densité.
    - Un Cluster est une zone de l'espace où la densité d'observations est importante.

$$N_{Eps}(p) = \{q \in D | dist(p,q) \leq Eps\}$$

  - Isolation Forest
    - Construction d'un ensemble d'arbres aléatoires pour un ensemble de données.
    - Les anomalies sont des points avec le plus court chemin de la racine.

$$s(x,n) = 2^{-\frac{E(h(x))}{c(n)}}$$

ANALYSE DES LOGS (SERVEUR MAIL)

- Résultats de la détection d'anomalies :

Modèle	**Silhouette**	**Temps(s)**
DBSCAN	**0.905**	196.40
IForest	**0.975**	**3.67**

$$\text{Silhouette} = b(i) - a(i) / \max\{a(i), b(i)\}$$

ANALYSE DES LOGS (SERVEUR MAIL)

- Résultats de la détection d'anomalies :
  - Classification manuelle sur un échantillon de 2 mois.

Modèle	**F-score**	**Taux de bonne classification**
DBSCAN	**0.96**	**0.997**
IForest	0.784	0.990

$$F-score = 2.precision.rappel / precision + rappel$$

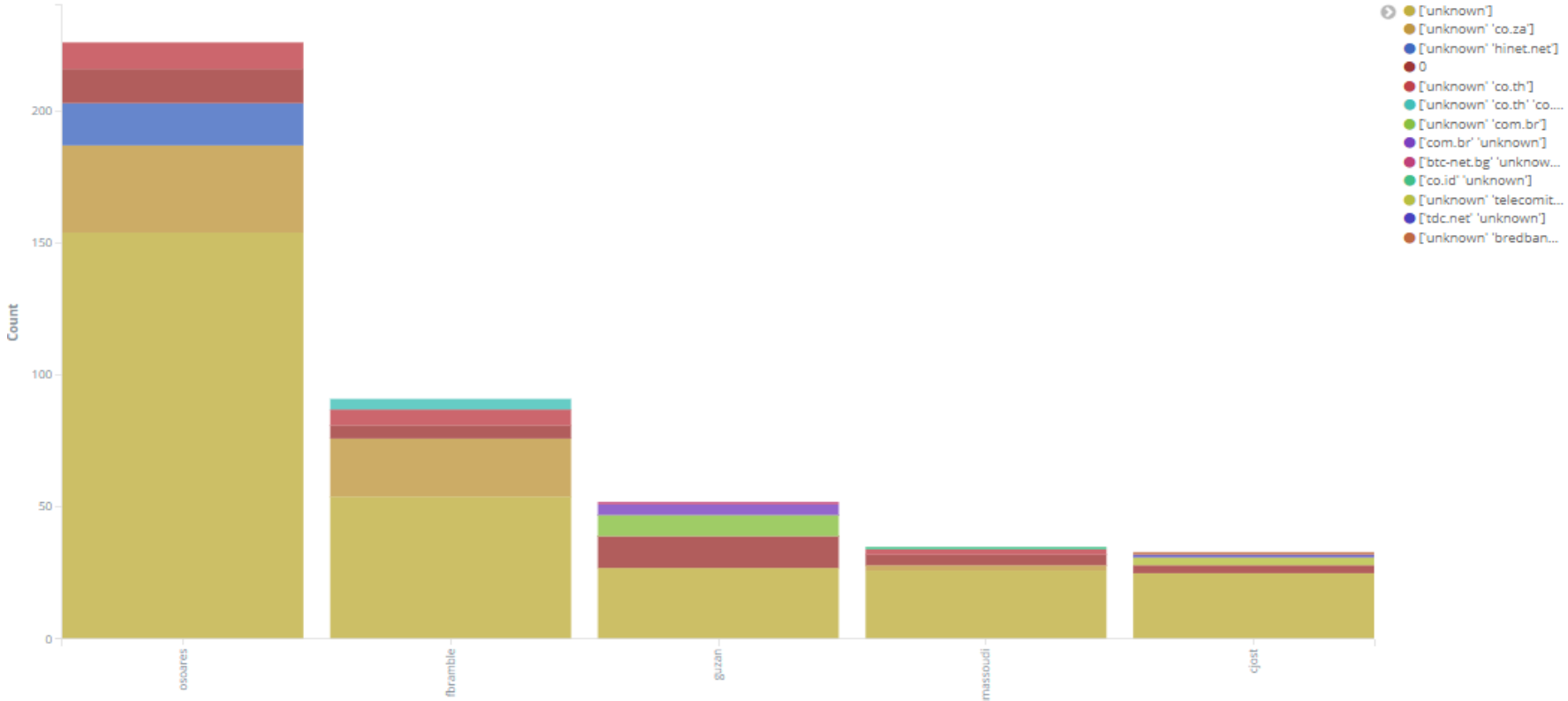
ANALYSE DES LOGS (SERVEUR MAIL)

- Résultats de la détection d'anomalies sur Kibana :

date	username	db	domains	ips
2018-03-31 22:10:46	ig	normal	['unknown']	['10.10. ']
2018-03-31 21:10:46	darc	normal	0	0
2018-03-31 20:40:46	darc	normal	['google.com']	['209.85]
2018-03-31 20:40:46	sg	normal	['unknown']	['37.16]
2018-03-31 20:40:46	ig	normal	['unknown']	['10.10 ']
2018-03-31 20:10:46	sg	normal	0	0
2018-03-31 20:10:46	rj	normal	['google.com']	['209.85. ']
2018-03-31 20:10:46	fm	normal	['wanadoo.fr']	['86.239. ']
2018-03-31 19:40:46	sg	normal	['sfr.net']	['93.6. ']
2018-03-31 19:40:46	ig	normal	['unknown']	['10.10. ' '10.10. ']
2018-03-31 19:10:46	ft	normal	0	0
2018-03-31 18:40:46	as	normal	0	0

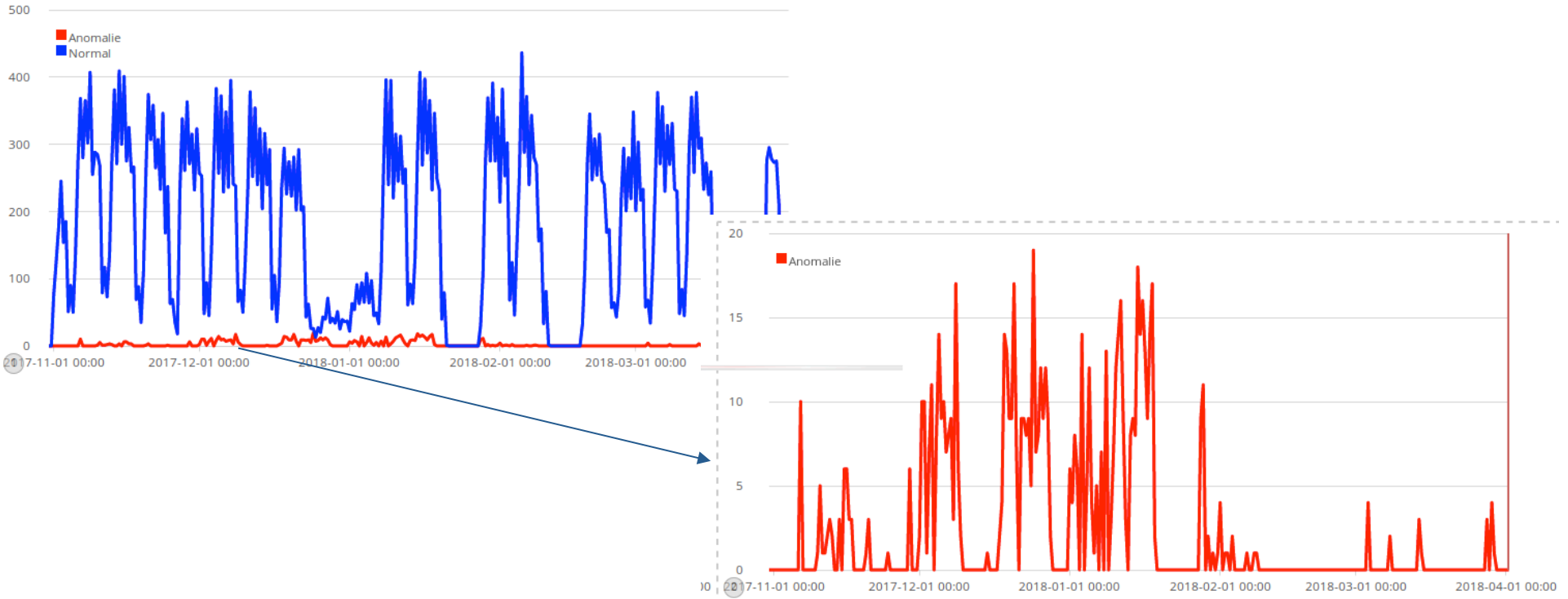
ANALYSE DES LOGS (SERVEUR MAIL) - DBSCAN

- Résultats de la détection d'anomalies sur Kibana :



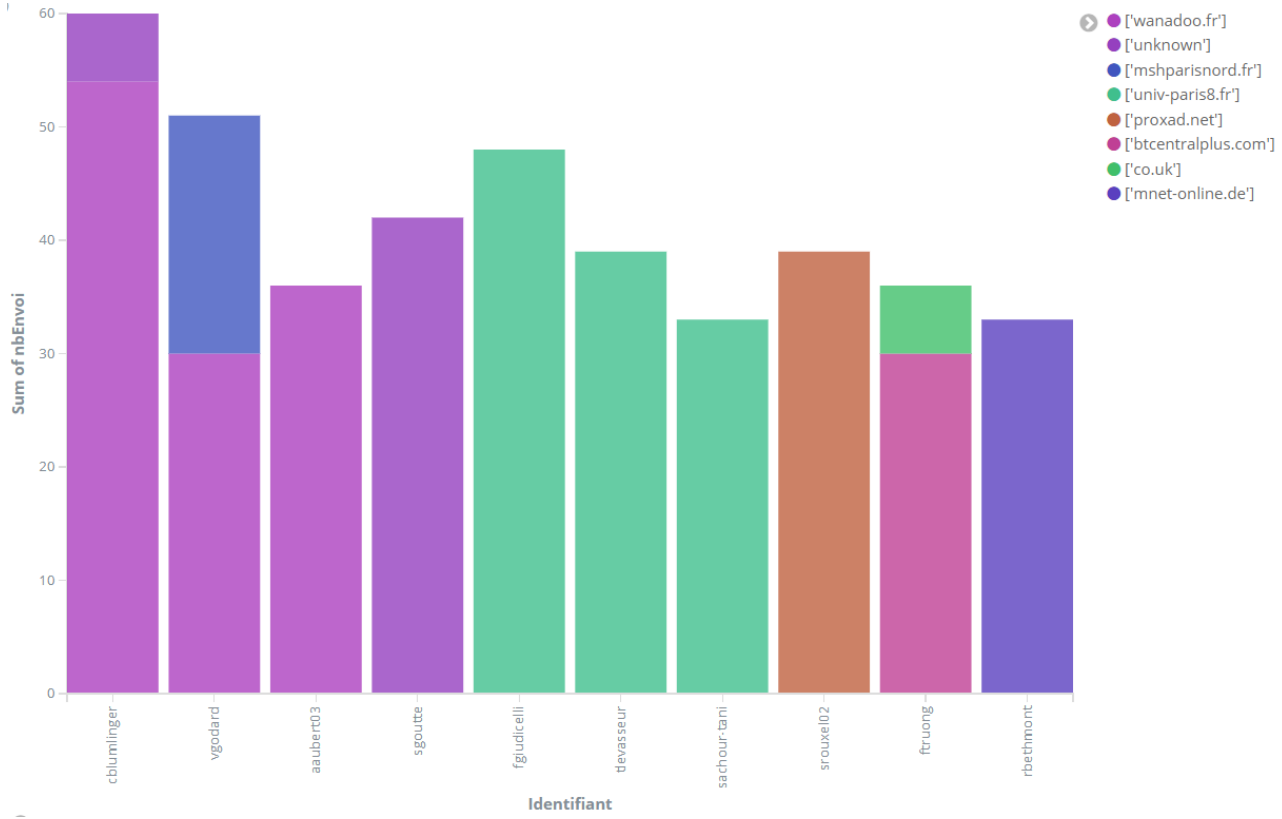
ANALYSE DES LOGS (SERVEUR MAIL) - DBSCAN

- Résultats de la détection d'anomalies sur Kibana :



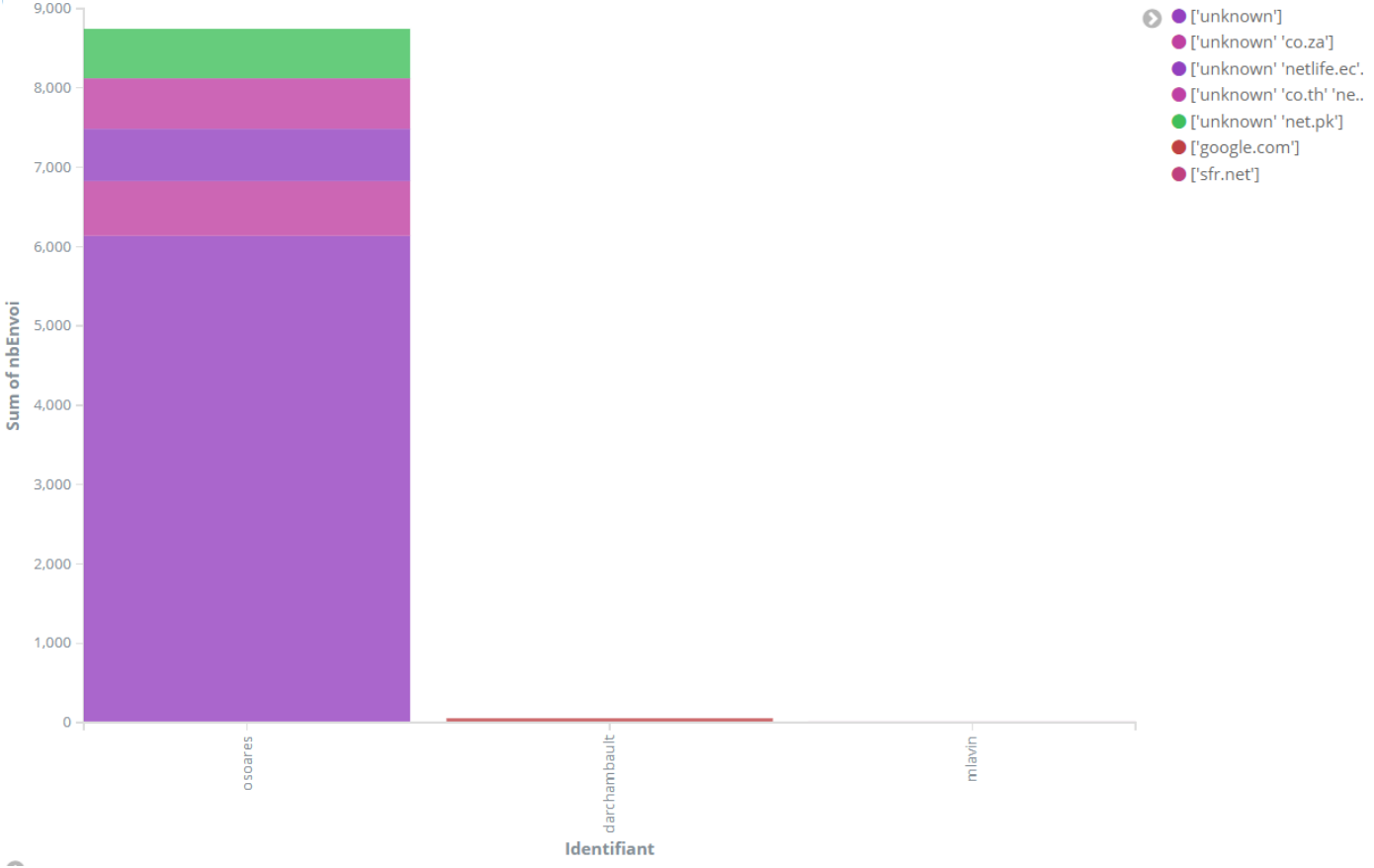
ANALYSE DES LOGS (SERVEUR MAIL) – ISOLATION FOREST

- nbEnvoi par jour (normal) :



ANALYSE DES LOGS (SERVEUR MAIL) – ISOLATION FOREST

- nbEnvoi par jour (anomalie) :



ESUP-DAYS 26

Centralisation des logs et détection d'anomalies par apprentissage automatique.

CONCLUSION ET TRAVAUX FUTURS

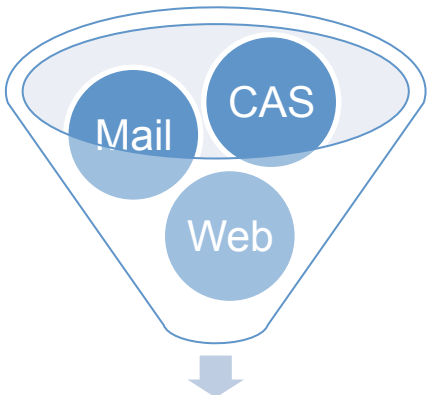
esup portail

CONCLUSION ET TRAVAUX FUTURS

- Conclusion
  - La détection d'anomalies permet de repérer des informations critiques dans les données.
  - La méthode statistique et/ou l'apprentissage automatique (le machine learning) ont permis de détecter des anomalies à partir des données récoltées.
  - Nous avons besoin de différentes approches pour résoudre un problème particulier.
    - Apprentissage supervisé et non-supervisé

TRAVAUX FUTURS

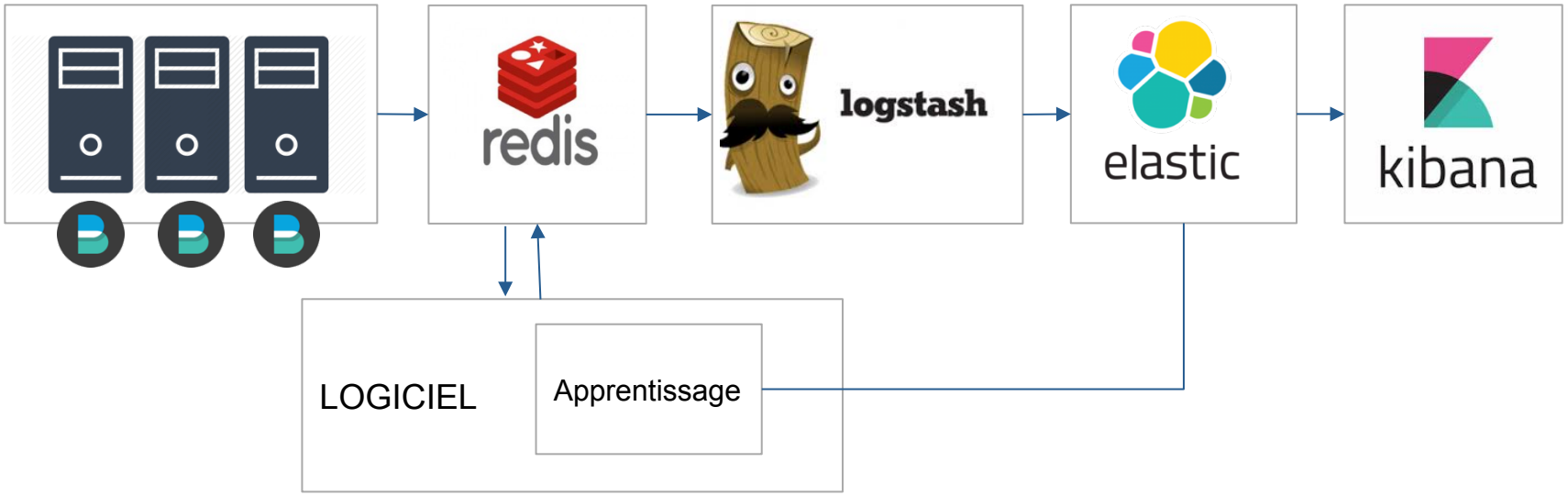
- Fusionner les données issues de plusieurs sources dans une seule base de données, pour améliorer le système de détection d'anomalies.



- Réaliser un système d'apprentissage en temps réel.

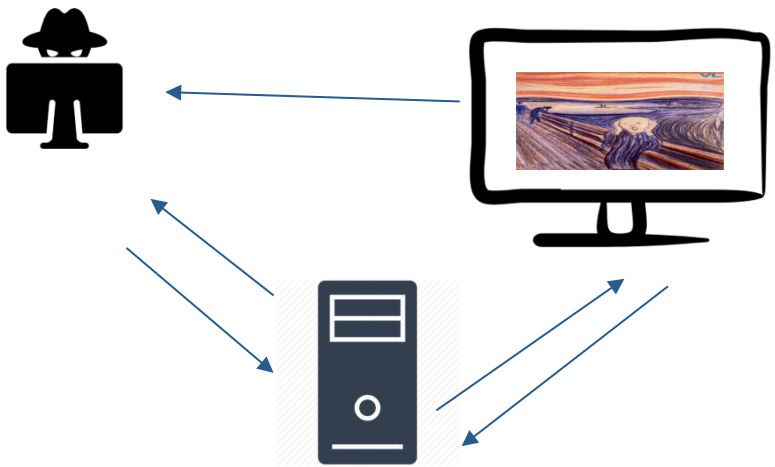
TRAVAUX FUTURS

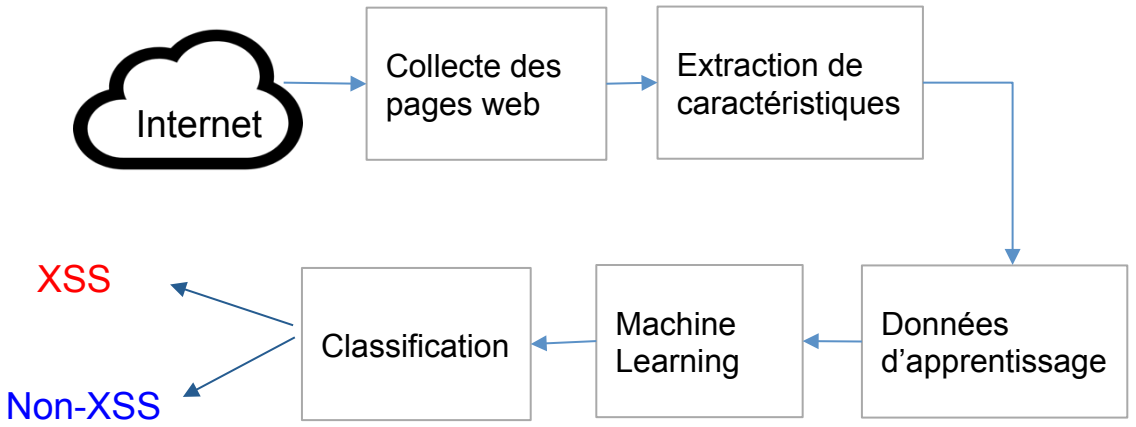
- Modification de l'architecture : Centralisation des logs et l'apprentissage automatique



TRAVAUX FUTURS

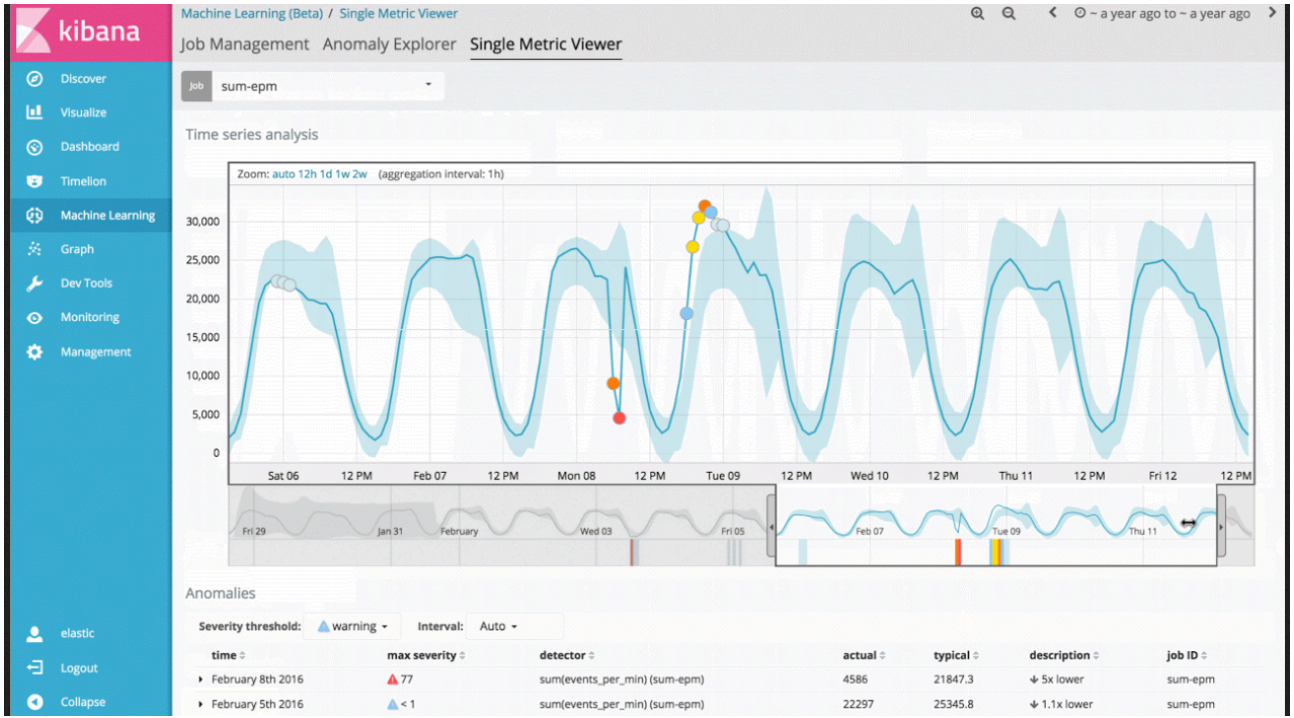
- Détection de cross-site scripting attacks(XSS) à l'aide du Machine Learning (Serveurs Web)





TRAVAUX FUTURS

- Développement d'un plugin KIBANA ?



MERCI DE VOTRE ATTENTION

Contact :
Sanghun BANG(sang-hun.bang@univ-paris8.fr)
Yasmina Bensitel(yasmina.bensitel@etud.univ-paris8.fr)