

drop.chapril.org - Demande #5162

Consommation en hausse rapide sur 15 jours /\

29/01/2021 22:21 - Christian P. Momon

Statut:	Fermé	Début:	29/01/2021
Priorité:	Élevée	Echéance:	
Assigné à:	Romain H.	% réalisé:	0%
Catégorie:		Temps estimé:	0.00 heure
Version cible:	Sprint 2021 février		

Description

21:07 < PoluX> cpm_screen: on sait d'ou vient l'inflation de maine ? <https://grafana.chapril.org/d/qfRrThRMk/icinga2-influxdb-host?from=now-6M&orgId=1&to=now&var-command=disk&var-hostname=maine.cluster.chapril.org&var-service=volume%20%2Fvar%2Flib%2Flibvirt%2Fmaine&viewPanel=1>

21:11 < PoluX> arf ya drop, valise et pouet dessus

21:11 < PoluX> nos 3 gros consomateurs

21:12 < PoluX> mais même à eux 3 ils ne semble pas grossir autant

21:42 < cpm_screen> PoluX: bien vu, ça grimpe fort, en ~15 jours ~50 Go

21:43 < cpm_screen> j'ignore d'où ça vient, pas d'événement particulier il me semble

21:43 < cpm_screen> en principe, on fait un point mensuel pendant la réunion animsys, service par service, donc on devrait le détecter là

21:44 < cpm_screen> mais avec un bon dashboard Grafana, ça devrait pouvoir se comprendre avant

Historique

#1 - 29/01/2021 22:27 - Christian P. Momon

- Fichier 20210129maine.png ajouté

Effectivement, en l'espace de 15 jours, on voit bien une hausse rapide de la consommation d'espace disque sur la machine maine.

<https://grafana.chapril.org/d/qfRrThRMk/icinga2-influxdb-host?from=now-30d&orgId=1&to=now&var-command=disk&var-hostname=maine.cluster.chapril.org&var-service=volume%20%2Fvar%2Flib%2Flibvirt%2Fmaine&viewPanel=1>

20210129maine.png

#2 - 29/01/2021 22:28 - Christian P. Momon

- Fichier 20210129drop.png ajouté

Fin du suspense, ça vient de la vm drop :

<https://grafana.chapril.org/d/qfRrThRMk/icinga2-influxdb-host?viewPanel=54&orgId=1&from=now-30d&to=now&var-hostname=drop.cluster.chapril.org&var-service=All>

20210129drop.png

#3 - 29/01/2021 22:32 - Christian P. Momon

- *Priorité changé de Normale à Élevée*

Du coup, questions :

- la purge fonctionne-t-elle normalement ?
- le nombre d'utilisateur a-t-il énormément augmenté par rapport au dernier rapport d'activité ?
- un seul utilisateur est-il à l'origine de cette hausse ou est-ce une multitude disparate ?
- est-ce beaucoup de nouveaux petits fichiers ou quelques gros ?
- y-a-t-il un usage abusif du service ?

Bref, tout va bien ou faut-il mettre en œuvre une vigilance spécifique ?

#4 - 30/01/2021 02:59 - Christian P. Momon

Une première analyse rapide qui laisse perplexe quant à l'usage qui est parfois fait du service :

```
=(^-^)=root@drop:/var/log/apache2/drop.chapril.org# zgrep -h "/Jan/2021\" drop.chapril.org-access.log* | awk '{print $1}' | sort | uniq -c | sort -rn | head -20
```


330M 24 janv.
316M 29 janv.
316M 29 janv.

#6 - 30/01/2021 14:26 - Romain H.

J'ai essayé d'analyser les journaux pour comprendre.

J'ai ajouté 3 scripts non finalisés dans /root/tmp_scripts pour analyser les journaux de apache :

- log2activity.py : parse les journaux pour récupérer les téléchargements / téléversements. Les IPs sont anonymisées via un masquage qui les remplace par des uuid.
- show_top.py : liste le TOP30 des plus gros téléchargeurs, téléverseurs et fichiers avec la plus grosse activité
- list_uploaded_files.py : liste les fichiers téléversés par un utilisateur et le nombre de téléchargement pour chaque.

Vu les premiers résultats, je confirme qu'il y a des utilisateurs qui détournent l'utilisation initialement prévue par drop et qui ont automatisé le téléversement de fichiers de petite taille. Chacun de ces fichiers semble être téléchargé uniquement par un utilisateur. Je me demande si ce n'est pas pour utiliser dans des campagnes de spam ou des activités de ce genre, mais comme on a pas accès au contenu des fichiers et qu'il n'y a pas de plainte d'utilisateur on ne peut pas en être certain.

On pourrait peut être ajouter un quota sur le nombre de téléversement possible par IP pour limiter cet effet. L'ajout d'un captcha pourrait fonctionner aussi mais empêcherait l'utilisation des utilitaire en ligne de commande.

#7 - 30/01/2021 15:04 - François Poulain

On pourrait peut être ajouter un quota sur le nombre de téléversement possible par IP pour limiter cet effet.

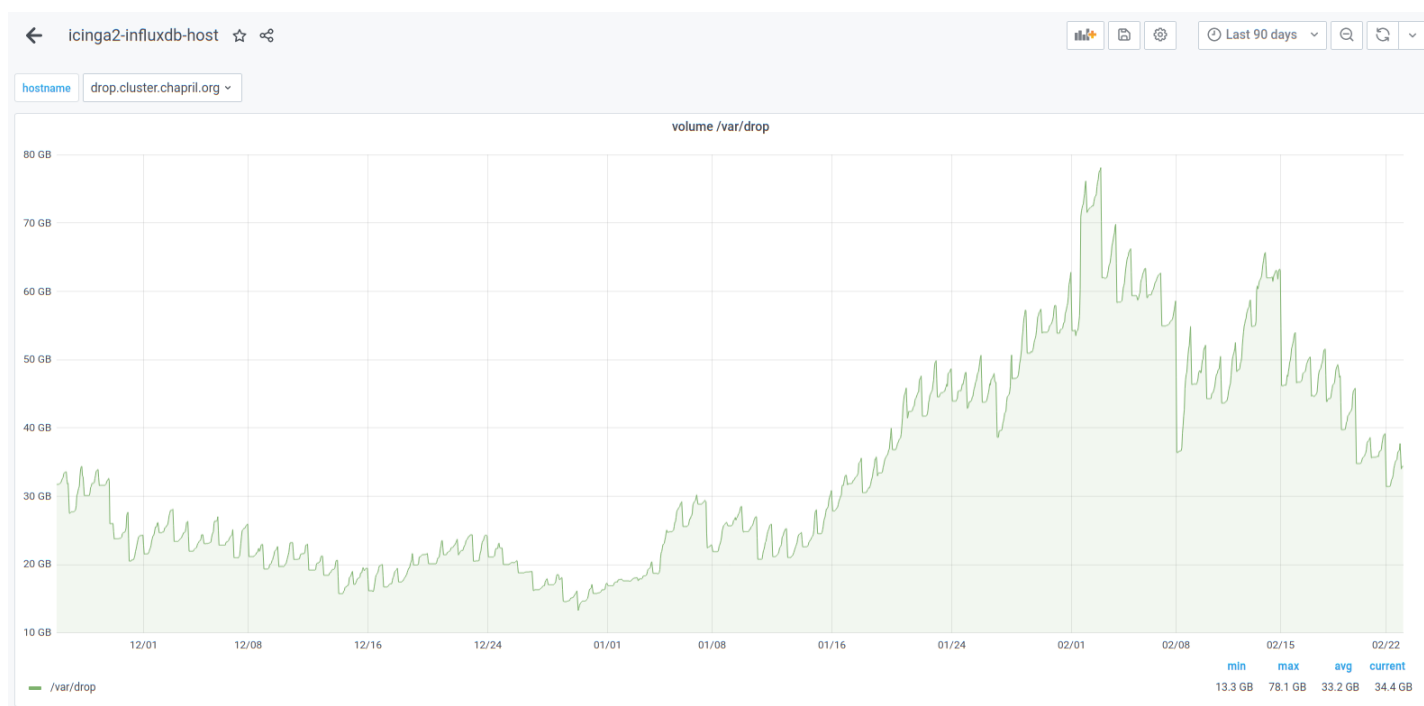
Oui si c'est possible.

#8 - 23/02/2021 04:39 - Christian P. Momon

- Fichier drop20210223var.png ajouté

- Statut changé de Nouveau à En cours de traitement

La hausse semble être retombée et le niveau être revenu à une valeur plus tranquille.



#9 - 03/03/2021 22:26 - Romain H.

- Statut changé de En cours de traitement à Résolu

Je confirme que le pic est passé.

On reste à un niveau moyen un peu plus élevé que précédemment (49 Go contre 28 Go), on envisagera d'augmenter l'espace disque si nécessaire.

#10 - 03/03/2021 23:08 - Romain H.

- Statut changé de Résolu à Fermé

#11 - 04/03/2021 01:11 - Christian P. Momon

- Version cible changé de Backlog à Sprint 2021 février

Fichiers

20210129maine.png	41,2 ko	29/01/2021	Christian P. Momon
20210129drop.png	69,3 ko	29/01/2021	Christian P. Momon
drop20210223var.png	86,2 ko	23/02/2021	Christian P. Momon