

BPF-33

Doc stack Thanos

Présentation Équipe Supervision du 30/10/2024

Objectifs des nouvelles stacks mon3/mon4

- scalabilité de la solution Thanos
 - en multipliant le nombre de Receivers
 - en doublant (voire plus) le point d'entrée "distributor"
- proposer une solution aux incidents "sapin de Noël"
- profiter de quelques fonctionnalités complémentaires
 - accepter les métriques "out of order"
 - pouvoir exécuter des expressions sur des intervalles de plus de 24h
 - ...

Composants

[https://docs-](https://docs-internal.corp.caascad.com/Architecture/Monitoring/NGOT/centra)

[internal.corp.caascad.com/Architecture/Monitoring/NGOT/centra](https://docs-internal.corp.caascad.com/Architecture/Monitoring/NGOT/centra)

- Chemin d'écriture
 - Receive "distributeur"
 - Receive
 - (Compactor)
- Chemin de lecture
 - Querier "proxy"
 - Ruler
 - Receive
 - Store Gateway
- Chemin d'alerting
 - Ruler
 - Querier
 - Receive
 - Store Gateway

MCO : où trouver la doc ?

- [Architecture/Monitoring/NGOT/central_zones/](#) (le schéma d'architecture)
- “Help_Alerts” ([exemple](#))
- [Opérations/Applications/Thanos/](#)

En cas d'incident

1. help_alerts...
2. ...qui redirige souvent vers
Opérations/Applications/Thanos/ ou
Opérations/Applications/Prometheus/

Si on veut redéployer

Certains composants ne se déploient pas avec Trackbone.

Tout est expliqué dans la documentation

[Applications/Thanos/Deployment_of_Thanos_components/](#)

Les pièges à connaître (1/2)

- Certains composants ne se déploient pas avec Trackbone
- Thanos-Receiver
 - Ne jamais changer les R/L. On ajoute un réplica à la place
 - Ne jamais augmenter la taille du PVC. On ajoute un réplica à la place
 - Ne jamais ajouter un réplica avec `kubectl edit` ou `kubectl scale`. Il faut redéployer avec la chart helm (sinon, il y a aussi un *hashRing* dans un configMap à mettre à jour)
 - Si modif sur les pods Receivers, faire un `kubectl scale replicas=0` avant
 - **Toujours suivre la doc !**
- Thanos-Receiver “distributor”
 - Le “distributor” est un Receiver avec une configuration spécifique (sans stockage local).
 - Ses “resources requests” en mémoire sont énormes par rapport à sa consommation “de croisière”. Cela est lié à son comportement au démarrage.
 - Au démarrage, à cause de ses R/L, un nouveau noeud peut être créé par l’auto-scaler. Parfois, c’est loooooong !!!
 - En cas d’OOMKill au démarrage, on peut augmenter les R/L (même à l’arrache, mais il faudra pérenniser ultérieurement).

Les pièges à connaître (2/2)

- Thanos-Query et Thanos-Query “proxy”
 - Ces deux composants sont les mêmes mais n’assurent pas le même rôle.
 - Parce que la chart helm ne permet pas de déployer ces deux composants avec un rôle différent, on déploie le “proxy” de façon séparée.
- Thanos-Ruler
 - Il est piloté par Prometheus-Operator (et donc déployé par Trackbone)
- Thanos-Compactor
 - La version du compactor diffère de celle des autres composants Thanos. Ce n’est pas grave, c’est compatible.

Dashboard Grafana

Il existe 2 dashboards de travail, qui peuvent être utilisés en cas de recherche lors d'un incident :

- **PF-2640** : ce dashboard donne une vue globale sur “mon3” et “mon4”
- **Incident sapin de Noël** : ce dashboard rassemble des métriques agrégées pour les alertes liées à la communication “remote-write” entre les Prometheus-Cluster et les stacks centrales mon1/mon2/mon3/mon4

Ces dashboard seront peut-être pérennisés ultérieurement avec des noms plus explicites.

Conclusion

GO MCO prévu en début de semaine prochaine !

Les silences sur les stacks mon3/mon4 seront retirés.

L'envoi des métriques sur mon1/mon2 sera désactivé dès que possible.

Le démantèlement des stacks mon1/mon2 sera planifié ultérieurement.

