



Ordre du jour :

- Situation des sauvegardes

D'autres sujets ont été abordés lors de cette réunion :

- Problèmes actuel d'indexation (chargement Willey Elasticsearch)
- Expérimentation ISTEEX « ElasticSearch / GlusterFS » en mode « bare metal » sur 5 PC
- Divers

Présents : Claude Niederlender, Stanislas Perrin, Roland Lambert, Martial Luc, Sylvain Ponticelli, Patrice Ringot, Philippe Scheffer, Angel Turri, Michel Villaume.

Durée de la réunion 1h30 / Rédacteur du CR : Patrice Ringot

Situation des sauvegardes

Les VMs vpgithub, vdistex4 et vdistex8 peuvent être supprimées.

La viistex, qui sert l'ancienne API ISTEEX, et qui est utilisée par Jacques Ducloy pour ses enseignements, ne peut pas encore être supprimée. Claude dira à Jacques que cette machine sera stoppée/supprimée quand ses cours de l'année universitaire en cours seront terminés. La viistex n'a plus besoin d'être sauvegardée.

Il ne faut pas sauvegarder les machines ElasticSearch en ce moment (indexations lourdes). Quand le moment sera venu, l'équipe ISTEEX communiquera à Michel les éléments nécessaires pour une sauvegarde utilisant l'API snapshot/restore d'ElasticSearch).

Pour la suite, une session de travail (AT/MV/ISTEEX) est prévue semaine 10 pour aller dans le détail de la configuration de sauvegarde.

Problèmes actuels d'indexation

L'indexation ElasticSearch Wiley pose problème ; on n'arrive pas au bout.

IPROD propose une session commune de monitoring (Appli, OS, JVM, Infra, réseau, stockage) lors d'une prochaine indexation. Il faut prévoir une demi-journée pendant laquelle

chacun pourra contribuer avec sa perspective (à intégrer dans le prochain sprint ISTEEX). On n'aura peut-être pas des résultats lors de cette première session, mais on accumulera de l'expérience commune.

Une piste pour l'essoufflement de l'indexation peut résider dans le paramétrage de la JVM (celle-ci peut passer en mode garbage collector désespéré suite à un manque ou à une fuite mémoire d'ElasticSearch). Ce point pourra être vérifié lors de la session (à préparer à l'avance pour le paramétrage de lancement de la JVM et connexion de la jconsole).

ISTEEX va faire intervenir un consultant Zenika pour le paramétrage ElasticSearch. Des IPROD seront invités à cette occasion.

Expérimentation ISTEEX « ElasticSearch / GlusterFS » en mode « bare metal » sur 5 PC bureautique.

Indépendamment de l'infra, ISTEEX souhaite mener son expérience sur un cluster de pc bureautique, en dehors de toute virtualisation, sur les technologies ElasticSearch et GlusterFS.

De son côté, IPROD peut agir sur plusieurs leviers de l'infrastructure pour l'optimiser aux besoins du projet (ex : paramétrage VMWARE, baie de stockage) dès lors qu'elle aura une connaissance de plus en plus fine de l'architecture ISTEEX (ce qui suppose une collaboration étroite et régulière, et des métriques communes).

On peut par exemple envisager d'assigner une VM Elasticsearch par ESXi, faire de l'attachement en local de SSD sur les ESXi pour optimiser l'accès aux index ES, ne plus laisser la baie gérer la répartition SSD/non SSD, mobiliser de la mémoire pour constituer des FS « ramdisk », etc ...

Mais quelque soit l'environnement d'exécution de la plateforme ISTEEX, IPROD souligne qu'il est nécessaire de commencer à mettre en place des outils qui nous permettront d'obtenir des métriques de plus en plus solides pour évaluer les gains, les pertes d'une plateforme par rapport à une autre, les améliorations ou les régressions induites par des modifications d'architecture ou d'implémentation.

Gatling (test de perf, cf. réunion du 19/2/15) est l'un de ces outils (Graphite, Diamond, JMXtrans pourront compléter le tableau).

Un coding dojo Gatling ISTEEX/IPROD sera organisé semaine 12.

A l'issue de ce dojo, il faudra programmer une première campagne de tir commune (monitorer les aspects habituels pendant le test de perf).

A la question ISTEEX : « y-a t'il encore de l'espace disponible dans les racks ISTEEX ? », IPROD répond par l'affirmative.

Divers

- ISTEEX doit se synchroniser avec PR pour le déploiement de Java8 sur les vm de prod.
- ISTEEX signale une optimisation dans le procédé d'indexation/recherche : le JSON enrichi par son cheminement dans tous les modules LI est intégré dans les index ES ;

il n'est plus nécessaire d'aller le chercher séparément au retour d'une requête ElasticSearch.

- ISTEEX envisage d'utiliser mongodb pour stocker les résultats des modules de LoadIstex.
- Pour info, la VM vp-istex-esk (Dashboard Kibana pour une population restreinte dont l'ABES) dispose de sa propre instance Elasticsearch.
- Stanislas Perrin a effectué des relevés wget, aussi bien vers l'interne que vers l'externe, qui peuvent révéler des problèmes réseau. Il communiquera à Philippe, Sylvain et Michel son procédé et ses résultats.

Prochaine réunion (n+2) – semaine 12

- Coding dojo Gatling