

ENIX désigne la société ENIX SARL, dont le siège social est situé au 8 bis, rue du Centre, 94370 Sucy-en-Brie, représentée par Sébastien Wacquiez.

TDFVS désigne ... TDFVS, dont ... ??, représentée par ??.

Objet du présent devis

TDFVS souhaite corriger des problèmes de performances affectant l'utilisation de plusieurs serveurs applicatifs au sein de son réseau local.

Afin de compléter les études déjà réalisées en interne et/ou par d'autres prestataires, TDFVS mandate ENIX afin de mettre en place plusieurs systèmes de mesure réseau, afin de déterminer si certains éléments du précis impactent de manière négative les performances des applications concernées.

Ce document propose un ensemble de mesures pouvant être mises en place par ENIX. Elles se basent sur les documents transmis par TDFVS, à savoir :

- ◆ un email de Monsieur Amel Filali du 19 octobre 2007,
- ◆ un rapport d'audit Active Directory et DNS rédigé par Ludovik Dopierala,
- ◆ un rapport de mission d'optimisation Dalet rédigé par Hugues Bourissou et Norbert Gorzalka,
- ◆ une liste des services et des serveurs de BFM TV concernés.

Premier outil d'analyse : *smokeping*

Smokeping est un outil Open Source similaire au logiciel propriétaire Ping Plotter, permettant de grapher les temps de réponse d'un ensemble d'équipements IP. Ce logiciel est constitué de deux composants : une sonde collectant les données, et un générateur de graphes.

La sonde effectue à intervalle réguliers des « ping » (envoi de messages ICMP *echo-request* et mesure du temps écoulé avant réception du message *echo-reply*) vers l'ensemble des adresses IP configurées, et stocke les données résultantes dans une base de données *rrdtool*.

Le générateur de graphes permet d'afficher de manière visuelle l'évolution dans le temps de la latence (temps de réponse) et de la gigue (variation du temps de réponse autour de la valeur moyenne) de chacun des équipements.

Cet outil est fréquemment utilisé par les opérateurs IP afin d'identifier les dysfonctionnements d'équipements réseau (switches – niveau 2, ou routeurs – niveau 3), ou de liens réseaux.

La valeur ajoutée d'ENIX vis-à-vis de cet outil s'axe autour de trois points : la mise en place de l'outil ; son paramétrage pour les besoins précis de l'analyse des équipements de TDFVS ; et l'intégration d'un outil d'identification temporelle des dysfonctionnements.

La mise en place de l'outil comprend l'installation de *Smokeping* sur un serveur UNIX fourni par TDFVS (ou acheté ou loué le cas échéant) et le *provisioning* des équipements à observer dans la configuration de *smokeping*.

Le paramétrage *ad hoc* consiste en une modification du fonctionnement de *smokeping* afin que la sonde collectrice de données s'exécute en permanence (non-stop), au lieu de s'exécuter 10 secondes toutes les 5 minutes.

L'outil d'identification temporelle est une petite interface Web très simple, permettant aux utilisateurs des applications ou aux administrateurs systèmes et réseaux de TDFVS de mémoriser directement dans *smokeping* les instants précis où des ralentissements ou dysfonctionnements sont remarqués. Concrètement, lorsqu'un opérateur constate un problème applicatif imputable au réseau, il se rend sur une page Web hébergée sur le même serveur que *smokeping*, et un simple clic permet d'enregistrer la date et l'heure précises de l'incident, afin de les recouper ultérieurement avec les graphes générés par *smokeping*.

La mise en place de cet outil nécessite de la part de TDFVS la fourniture de la liste des adresses IP (ou des noms DNS) des équipements (serveurs, switches...) à observer, en les classant optionnellement par catégories (l'interface de consultation permettant de gérer une arborescence pour plus de clarté). D'autre part, il est vivement souhaité que la sonde soit connectée « au plus près » des équipements de *backbone* (si possible directement sur un switch central, ou au pire sur un équipement dont la fiabilité est reconnue suffisamment élevée).

Suivi de la charge des liens réseau grâce à une sonde SNMP haute précision

Les plate-formes de supervision classiques (HP/OV, Open NMS...) offrent généralement (entre autres) un système permettant de suivre l'occupation des liens réseau, en relevant à intervalles réguliers les compteurs des switches ou des routeurs par l'intermédiaire du protocole SNMP.

ENIX propose l'installation d'une solution de supervision dédiée au suivi du trafic réseau, mais employant une résolution beaucoup plus fine, afin d'identifier d'éventuels pics de charge. Les systèmes traditionnels ont une résolution de 1 à 5 minutes. À cette résolution, un pic de charge d'un gigabit par seconde pendant dix secondes n'apparaît de manière « lissée » et n'est pas décelable. La solution développée par ENIX permet d'interroger les équipements toutes les 2 à 10 secondes, et permet d'identifier de tels pics de charge en tant que tels.

Cette solution est basée sur un collecteur de données parallélisé écrit en Python, et utilise le système *rrdtool* pour le stockage et l'agrégation des mesures. Elle a déjà été déployée avec succès sur des réseaux de plus de 300 postes, dans le cadre d'applications de type *video-on-demand* par exemple.

Ce système est plus complexe que *smokeping*, mais donne des résultats différents : il ne mesure pas les temps de réponse des équipements « vus » depuis une sonde de supervision ; il relève les compteurs de trafic d'équipements réseau SNMP, et permet donc d'identifier des surcharges de liens réseau, ou bien des problèmes de configuration de *trunks* (agrégats de ports Ethernet), par exemple.

Tout comme *smokeping*, ce logiciel s'installe sur un serveur UNIX (qui peut être le même que celui utilisé par *smokeping*, si les deux solutions sont retenues), et dispose d'un outil d'identification temporelle des dysfonctionnements permettant d'enregistrer avec précision la date et l'heure des incidents réseau pour recouper ensuite les informations avec la plus grande exactitude possible.

La mise en place de cette sonde SNMP nécessite que TDFVS fournisse la liste des équipements réseau à observer, en indiquant si nécessaire les communautés SNMP à utiliser. Si les équipements réseau ne sont pas configurés pour répondre aux requêtes SNMP, ENIX peut effectuer cette configuration en sus.

Analyse quantitative et qualitative du trafic de diffusion (broadcast)

Un des rapports d'expertise mentionnés plus haut suggère de diviser le réseau en plusieurs segments Ethernet. Sans faire entrer en ligne de compte les aspects relatifs à l'administration réseau et la gestion du parc, et en ne considérant que l'impact au niveau des performances réseau, ENIX propose d'installer une sonde enregistrant l'intégralité du trafic de diffusion sur le réseau de TDFVS. Cette sonde permettra de quantifier de manière extrêmement précise la quantité de trafic de *broadcast* transitant sur le réseau (et pouvant potentiellement le ralentir), ainsi que sa répartition selon les différents protocoles (par exemple : ARP, DHCP, annonces de contrôleur de domaine, protocoles de découverte automatique de réseau...).

Cette sonde permettra de savoir d'une part si le trafic de *broadcast* est assez important pour avoir effectivement un impact sur les performances du réseau (et justifier la séparation du réseau en plusieurs segments Ethernet), et d'autre part d'identifier les sources de ce trafic, afin de modifier le cas échéant leur configuration ou leur mode de fonctionnement pour réduire du trafic parasite.

La sonde est basée sur le logiciel Open Source *tcpdump*, modifié par ENIX afin de permettre

une capture continue et fiable du trafic réseau. L'analyse des données enregistrées se fait à l'aide d'outils développés par ENIX pour ce type d'applications ; éventuellement complétés par la suite Open Source *wireshark* pour l'analyse protocolaire.

Tout comme les deux solutions précédentes, celle-ci nécessite d'être déployée sur un serveur UNIX, qui peut là aussi être le même que celui utilisé par *smokeping* ou par le collecteur SNMP haute précision.

Installation d'un serveur syslog

Afin de consolider l'ensemble des messages émanant des équipements réseau, ENIX propose le déploiement d'un serveur *syslog* permettant de collecter l'ensemble des logs (d'information ou d'erreur) afin de les stocker de manière pérenne et de recouper plus facilement les événements affectant la topologie réseau. Par exemple, mettre côte-à-côte un message d'un commutateur X indiquant « changement de topologie STP » et un message d'un commutateur Y indiquant « perte de lien sur le port Z » au même instant, permet d'établir immédiatement une relation de cause à effet, qui n'est pas obligatoirement évidente, tout particulièrement si les *logs* des différents équipements ne sont pas accessibles simultanément et/ou que leur horloge n'est pas synchronisée de manière fiable.

Pour des raisons de fiabilité, il est recommandé d'installer le serveur *syslog* sur un serveur UNIX, qui peut être le même que celui utilisé pour les sondes citées précédemment. En revanche, il est en ce cas préconisé d'utiliser des disques durs séparés pour enregistrer les *logs*, pour garantir les performances des sondes.

Outre l'installation du serveur *syslog*, il est bien entendu nécessaire de configurer les équipements (serveurs ou routeurs) afin qu'ils envoient leurs messages vers ce serveur ; en fonction du nombre d'équipements et de leurs types, cette configuration peut nécessiter une prestation en sus.

Restructuration du backbone afin de supprimer le Spanning Tree Protocol

Les rapports d'expertise cités au début de ce document mentionnent des erreurs liés au STP (*Spanning Tree Protocol*). Ce protocole présente de nombreux défauts. Citons entre autres un temps de convergence relativement long, une inadéquation avec le protocole 802.1q utilisé pour les VLANs, ainsi qu'avec les protocoles d'agrégation de liens Ethernet (type FEC et LACP).

Un temps de convergence élevé signifie qu'en cas de micro-coupure sur un lien Ethernet, les switches situés de part et d'autre du lien vont devoir attendre jusqu'à plusieurs dizaines de secondes avant de ré-établir le trafic sur le lien concerné (ainsi, un problème de câblage ou d'interférences, au lieu de provoquer une coupure d'un dixième de seconde, peut provoquer une coupure d'une dizaine de secondes).

L'inadéquation avec les VLANs n'est pas gênante pour TDFVS tant que des VLANs ne sont pas déployés sur son réseau.

L'inadéquation avec les protocoles d'agrégation de lien, en revanche, peut s'avérer problématique ; les interactions entre le STP et les protocoles d'agrégation n'étant pas normalisées, des problèmes peuvent apparaître entre des équipements de constructeurs différents qui utilisent simultanément STP et un quelconque protocole d'agrégation.

À moins que la topologie réseau de TDFVS ne soit très particulière, ENIX propose de remplacer l'utilisation du STP par les protocoles d'agrégation de liens uniquement. Ces protocoles permettent le même niveau de redondance, mais avec une meilleure réactivité (de l'ordre de la seconde) et en permettant l'équilibrage de la charge sur les liens.

Concrètement, si deux équipements sont interconnectés par deux liens distincts et utilisent STP, un seul des deux liens sera actif à un instant T, et si ce lien est défaillant, il faudra un temps relativement long avant que les équipements ne basculent sur le lien de secours. Avec un protocole de type LACP, le trafic sera réparti entre les deux liens, et en cas de coupure d'un lien, l'autre sera immédiatement utilisé pour l'intégralité du trafic.

Cette restructuration nécessite d'établir au préalable une cartographie de la topologie

réseau, et une intervention sur la plupart des switches de coeur de réseau. L'intervention n'entraîne pas de longue coupure du réseau, mais nécessite une préparation adéquate des configurations des différents équipements.

| Récapitulatif | |
|---|---|
| Sonde <i>smokeping</i> | - déploiement de la sonde : 1 jour - configuration et <i>provisioning</i> : 1 jour - intégration de l'outil d'identification temporelle : 1 jour |
| Sonde SNMP haute précision | - déploiement de la sonde : 3 jours - configuration et <i>provisioning</i> : 1 jour - création des graphiques d'analyse : 2 à 3 jours (selon le nombre de ports et l'existence d'agrégats) |
| Sonde de trafic <i>broadcast</i> | - deployment de la sonde : 1 jour - analyse des résultats : 1 jour |
| Serveur <i>syslog</i> | - déploiement du serveur : 1 jour - configuration des équipements : 1 jour par modèle |
| Suppression du STP | - cartographie du réseau : 2 jours - planification de l'intervention + intervention : 1-4 jours |
| Installation d'un serveur UNIX Debian GNU/Linux « etch » | - installation sur votre matériel : 1 jour |

Pour toute information complémentaire sur cette proposition, contacter :

Jérôme Petazzoni – jp@enix.org - +33 6 73 23 55 40