

PerfSONAR - Supervision réseau multi-domaine

Nicolas Simar
DANTE
City House, 126-130 Hills Road,
Cambridge CB2 1PQ
nicolas.simar@dante.org.uk

Frédéric Loui
GIP RENATER - Service de suivi opérationnel
151 Boulevard de l'Hôpital,
75013 PARIS
frederic.loui@renater.fr

Résumé

De nos jours, chaque opérateur déploie un certain nombre d'applications afin d'assurer la supervision de son réseau au sein de son propre domaine. Lorsqu'un problème survient entre 2 établissements interconnectés à travers un ensemble de réseaux, la résolution de celui-ci devient épineuse. Ceci est dû au fait qu'il n'existe pas de structure d'échanges des données de supervision entre ces domaines de responsabilité distincts. Il est donc très difficile pour un utilisateur de diagnostiquer ce qui se passe le long du chemin et pour un opérateur de savoir si un réseau (et alors lequel ?) est à l'origine du problème, ou si celui-ci est plutôt de nature applicative (application, système d'exploitation, matériel). PerfSONAR, dans le cadre de l'activité JRAI (« Joint Research Activity 1 ») du projet GN2 coordonné par DANTE, a été conçu pour répondre à ce besoin. Les outils élaborés dans ce projet sont capables de fournir une vision de bout en bout d'un service réseau qui s'étend sur plusieurs domaines. Ces outils sont dédiés aux services opérationnels, aux équipes d'ingénierie de chaque réseau académique pour la recherche, mais aussi à tout opérateur ayant besoin d'apporter des réponses à un problème réparti sur différents domaines. Cet article décrit l'esprit du projet, aborde ses objectifs principaux ainsi que son déploiement au sein de RENATER.

Mots clefs

Supervision, réseau, services de bout en bout, métrologie, web services, multi-domaine, RRD, SNMP, NETFLOW

1 Introduction

Les applications numériques jouent un rôle de plus en plus prépondérant au quotidien. Ces dernières années ont vu une mutation conséquente des systèmes d'information. Le réseau, dans cette redistribution des cartes, devient le maillon essentiel de cette chaîne et la disponibilité de celui-ci doit être ajustée à des niveaux toujours plus élevés.

Chaque entité, entreprise ou opérateur, supervise donc son réseau de manière à garantir le meilleur service possible avec une palette d'outils qui lui est propre et interne à son domaine. Ces plates-formes ou « NMS¹ » fournissent

l'ensemble des indicateurs nécessaires pour opérer le réseau suivant les 5 domaines du « Network Management » : FCAPS².

Ces outils sont différents d'un domaine à l'autre : ils sont dépendants de l'équipement supervisé et diffèrent selon le constructeur. Un système de supervision de routeur se distinguera ainsi d'un système de supervision d'équipement optique. De la même manière, les plates-formes de supervision WDM de deux équipementiers différents fourniront des informations propres à leur mise en œuvre des standards optiques.

Ce constat est problématique dès lors que de nouveaux types d'applications réparties voient le jour et nécessitent d'être supervisées. Celles de type « grille » par exemple, nécessitent une disponibilité importante. Les capacités importantes des grilles de calcul tirent partie d'une infrastructure géographiquement distribuée et la cohésion de l'ensemble des réseaux reliant ces éléments est la clé de voûte du système global.

Dans le cadre de ces applications, il devient difficile, en cas de dysfonctionnement, de diagnostiquer l'origine du problème et de proposer des actions curatives en un temps raisonnable. En effet, certaines configurations, où de nouvelles applications utilisent des protocoles de transport variés sur des infrastructures réseaux hétérogènes, vont présenter des signes de pannes très délicates à interpréter. D'où le besoin de disposer d'un système capable de fournir une vision synthétique du comportement de bout en bout du, ou des, réseaux.

Dans le cas d'un service restreint à un seul domaine, les points de démarcation sont bien définis et la résolution du problème reste attachée à un périmètre connu. En revanche, lorsque le service couvre plusieurs domaines, lequel d'entre eux contient l'origine du problème ? Quels sont les outils disponibles afin de qualifier rapidement le problème sans pour autant remonter la chaîne classique d'escalade et ainsi diminuer le temps de rétablissement du service réseau ?

Cet article aborde les principales raisons qui ont motivé le développement de PerfSONAR [1] et donne une

¹ Network Management System

² Fault, Configuration, Accounting, Performance, Security Management.

description de son architecture. Un point sera fait sur l'état actuel du projet ainsi que son déploiement au sein de RENATER.

1.1 La communauté de la recherche : un contexte multi-domaine

La plupart des grands projets de recherche intègrent maintenant une dimension distribuée. La notion de grille est alors au cœur de projets comme DEISA³, EGEE⁴, GRID5000⁵.

Un autre exemple d'application répartie est le projet LHC⁶ du CERN. Ce dernier va produire 15 pétaoctets de données qui seront analysées par 78 centres de calcul scientifiques répartis à travers le monde entier.

Ces projets requièrent une bande passante non négligeable (10 Gbps) et les réseaux doivent se doter d'une haute disponibilité tout le long des circuits d'un domaine à un autre.

Les grilles citées précédemment ne sont que des exemples à grande échelle de projets reposant sur une infrastructure distribuée. D'autres applications de type « peer-to-peer » sont de plus en plus répandues et couvrent, elles aussi, plusieurs domaines.

En Europe, ces projets s'appuient sur les infrastructures réseaux gérées par les réseaux nationaux pour la recherche (NRENs) et DANTE [2] pour les liaisons inter-NRENs. Certains projets ont conduit à établir des chemins de bout en bout à travers les réseaux de recherche européens, asiatiques et américains.

Au fur et à mesure des déploiements dédiés aux projets de recherche, une question s'est posée naturellement : comment superviser ces types de service bâtis sur plusieurs domaines ?

De plus, comment opérer cette infrastructure de façon simple à l'épreuve de la « mise à l'échelle » ?

1.2 GN2-JRA1 : PerfSONAR

DANTE est une organisation à but non lucratif dont les membres sont les réseaux nationaux pour la recherche en Europe (NRENs). Sa mission est de définir, déployer et opérer des réseaux pour la recherche et l'éducation. Elle déploie actuellement GEANT2, l'infrastructure de communication servant d'appui aux projets mentionnés en 1.1.

En parallèle au déploiement de GEANT2, DANTE a lancé un ensemble d'activités visant à adresser les problématiques sous-jacentes à des services réseaux couvrant plusieurs domaines.

GN2-JRA1 (« GEANT2 - Joint Research Activity 1 ») est une de ces activités et est dédiée à la supervision multi-domaine. Celle-ci est menée en collaboration avec les

réseaux académiques en Europe et aux Etats-Unis. Son ambition est d'apporter des réponses aux problématiques réseaux inter-domaine et en premier lieu de fournir aux projets de recherche des indicateurs sur les infrastructures supportant leurs applications distribuées sur plusieurs réseaux nationaux distincts.

En outre, PerfSONAR au sein de RENATER peut être utilisé pour fournir aux établissements une vision de bout en bout du service réseau. Un service point à point d'un établissement A vers un établissement B, peut être étendu sur plusieurs domaines. En effet, généralement, un établissement est raccordé à un réseau de collecte qui est lui-même connecté à un Nœud RENATER ou « NR ». Dans cet exemple, on dénombre déjà 5 domaines distincts (établissement A, réseau de collecte A, RENATER, réseau de collecte B, établissement B).

Indirectement, dans le domaine commercial, les problématiques que cherche à résoudre le projet sont les mêmes que celles des opérateurs virtuels (VNO) qui ne possèdent pas forcément d'infrastructure propre, mais qui proposent des services réseaux reposant sur plusieurs partenariats.

GN2 est un projet qui englobe une vingtaine d'organisations. L'activité JRA1⁷ relative à la supervision multi-domaine développe un ensemble de protocoles. Ces derniers s'appuient sur un jeu de services web avec des rôles spécifiques, une sémantique et une syntaxe commune standardisées au sein de l'OGF[3] NM-WG⁸. Cette standardisation rend ainsi PerfSONAR facilement extensible (les interfaces étant publiques).

Les bénéfices du déploiement d'une infrastructure telle que PerfSONAR sont :

- la possibilité d'apporter une réponse plus rapide à l'utilisateur final sur la cause d'une dégradation de performance ;
- la possibilité pour les opérateurs d'avoir accès à des données plus proches de ce que l'utilisateur perçoit ;
- la mise à disposition d'outils de diagnostic standards et prêts à l'usage.

Il est à noter que certaines applications sont développées en faisant l'hypothèse qu'un réseau WAN est considéré comme une boîte noire qui restitue les paquets en respectant l'ordre d'envoi (ne prenant pas en compte les phénomènes de partage de charge, d'asymétrie de routage, re-routage, perte de paquets, etc.). Une infrastructure de supervision de bout en bout permet d'identifier si le problème est issu du réseau ou non.

2 L'infrastructure PerfSONAR

2.1 Les objectifs techniques

Compte tenu des objectifs du projet, PerfSONAR a été conçu de manière à intégrer les caractéristiques suivantes :

³ DEISA : Distributed European Infrastructure for Supercomputing Applications - <http://www.deisa.org/>

⁴ EGEE : Enabling Grids for E-sciencE - <http://www.eu-egee.org/>

⁵ GRID5000 : <https://www.grid5000.fr/>

⁶ LHC : Large Hadron Collider - <http://lhcb.cern.ch/>

⁷ JRA1 : Joint Research Activity 1 (Multi-domain monitoring)

⁸ OGF NM-WG : Open Grid Forum Network Management Working Group

- Modularité : La structure modulaire en composants indépendants confère à PerfSONAR une grande flexibilité et extensibilité. Celui-ci est un cadre ouvert qui autorise le développement de nouvelles fonctionnalités sans pour autant devoir restructurer son architecture.
- Dynamique : L'ajout, la modification ou la suppression de nouvelles sources de données propres à un domaine doivent être des opérations sans interruption de service durant l'exploitation au quotidien du système de supervision.
- Auto-configuration : Chaque composant doit avoir la possibilité de s'enregistrer auprès de l'ensemble du système de supervision et d'annoncer de manière flexible son existence et ses « capacités ». PerfSONAR inclut le concept de « bus applicatif » chargé de véhiculer la localisation de chaque ressource et la manière d'y accéder.
- Décentralisation : PerfSONAR a adopté un modèle décentralisé. Chaque domaine a la décision de publier ou non l'ensemble ou une partie de ses informations réseaux.
- Sécurisation : L'entité administrative décide de sa propre politique de sécurité. L'accès aux données doit au préalable être authentifié et autorisé par une infrastructure officielle dite « de confiance ». De plus, PerfSONAR ne doit pas congestionner les réseaux qu'il cherche à superviser.
- Robustesse : Dans le cas où un composant tombe en panne, l'infrastructure globale de supervision doit continuer à être opérationnelle.
- Capacité de mise à l'échelle : Il est essentiel pour un système distribué de pouvoir supporter un nombre conséquent de domaines sans qu'en soit affectée l'exploitation de l'ensemble.
- Auto-diagnostic : En cas de panne d'un composant, le système doit être capable de fournir des messages d'erreur clairs, afin d'établir un diagnostic précis et proposer des solutions curatives appropriées en un temps acceptable.
- Indépendance : Le système est distribué et ouvert. Il doit donc être indépendant de la plate-forme choisie, mais aussi du langage de développement.

2.2 Une architecture en trois niveaux

Afin de répondre aux critères listés précédemment, PerfSONAR a adopté une architecture « 3-tier ». Ce choix divise le système en trois couches distinctes. La première couche englobe les outils de supervision déployés par les opérateurs réseaux. La deuxième correspond à une couche formée de services web assurant la publication des données. La dernière couche est la couche de présentation des informations.

La figure 1 illustre ce modèle :

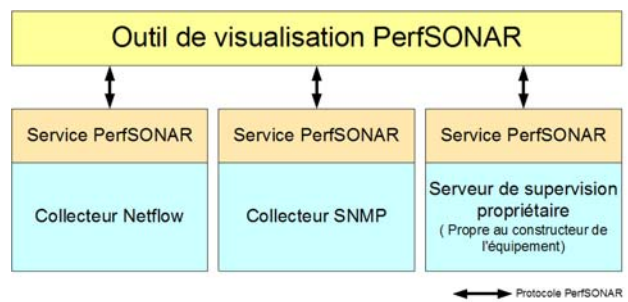


Figure 1 – Architecture « 3-tier »

La couche de plus bas niveau (« Measurement Tools »), a pour rôle de collecter les données auprès des équipements réseaux. Cette collecte est assurée par les systèmes de supervision existant déjà, utilisés par les opérateurs pour l'exploitation de leurs propres réseaux. Ces systèmes sont par exemple, un collecteur « SNMP », « CMIP » ou « NETFLOW [4] » et leur base de données associée. Ils fournissent des statistiques issues de mesures passives. Une autre catégorie d'outils inclut l'infrastructure de sondes de mesures actives. Les indicateurs provenant de ces systèmes dérivent de simulations de trafic réel dans le but de valider un niveau de service dans des conditions utilisateurs.

La couche intermédiaire (« PerfSONAR services ») a pour objet de publier les données par le biais d'une interface standardisée et les rendre ainsi disponibles aux autres domaines. Pour ce faire, une catégorie de services est chargée de fournir les données (« Measurement Archive » ou MA, « Measurement Point » ou MP, décrits plus loin dans cet article). Il arrive parfois que les données collectées nécessitent d'être corrélées avec d'autres statistiques afin de construire un indicateur commun aux domaines. Elles doivent être transformées avant d'être publiées. C'est le rôle du « Transformation Service ». D'autres services de cette couche servent d'appui logistique à la publication de données : le « Lookup Service » masque ainsi la complexité de recherche d'information à la couche de niveau supérieure. Le service d'authentification et d'autorisation assure que les règles de sécurité s'appliquent, qu'il y ait publication ou non d'information d'un domaine à un autre, et, si oui, dans quelle mesure le domaine extérieur est autorisé à accéder à tout ou partie des données.

La dernière couche correspond à une interface graphique visuelle dédiée à l'utilisateur. C'est à travers elle que les données issues de chaque domaine vont être agrégées afin de fournir une vision cohérente de bout en bout. Le tableau de bord ainsi obtenu fait abstraction des différentes techniques entre des domaines de responsabilité distincts, en présentant des caractéristiques communes le long du chemin suivi par le trafic. C'est aussi cette couche qui fait appel aux services d'authentification et d'autorisation afin de contrôler l'accès aux données.

2.3 Une architecture ouverte orientée « Services »

La section précédente présentait une hiérarchie à trois niveaux, l'un de ces niveaux étant composé de services s'interfaçant avec les outils déjà présents et les applications de visualisation. Cette section se focalise sur la structure générale d'un service PerfSONAR.

Un composant, du point de vue PerfSONAR, est un service web. Par définition, ce dernier interagit avec les autres services suivant une interface ou un protocole précis. Il aura par exemple la possibilité de répondre à des requêtes, d'en effectuer lui-même auprès d'autres services afin d'accomplir sa tâche, de s'enregistrer auprès du système afin de notifier la disponibilité de nouvelles informations.

Chaque service communique donc par le biais de messages bien définis. Ce protocole est basé sur le format SOAP XML et suit les spécifications de l' « Open Grid Forum Network Measurement Working Group (OGF NM-WG [3]) ». Par ce mécanisme, chaque service est une « boîte noire » autonome et laisse donc libre choix du langage de développement mais aussi du système d'exploitation. De cette manière, la plate-forme PerfSONAR est facilement extensible.

Cette extensibilité est renforcée par le fait que PerfSONAR, est un projet « Open Source ». Le code issu du projet est sous licence GPLv2. Toute personne désireuse de mettre en œuvre une supervision multi-domaine est ainsi libre d'utiliser PerfSONAR, voire de récupérer les sources et d'ajouter de nouvelles fonctionnalités afin de mieux répondre à ses besoins.

3 Les types de services PerfSONAR

Dans une optique de résolution de problèmes inter-domaine, le système de publication doit tenir compte du fait que :

- Les données brutes relatives au réseau sont la plupart du temps collectées par les NMS existantes. Le « Measurement Archive » agit comme une couche logicielle capable d'extraire les données de supervision de ces NMS. D'autres outils sont lancés à la demande afin d'opérer un diagnostic à un instant t dont le résultat n'est pas sauvegardé. Le « Measurement Point » s'interface avec ces outils de diagnostic.
- D'autres indicateurs reflètent la qualité d'un service en vigueur sur le réseau. Dans le cadre de la ToIP par exemple, le MOS⁹ est un indice dérivé de mesures brutes. C'est le rôle du « Transformation Service » que de transformer ces données et les rendre exploitables.
- Ces sources de données sont en nombre conséquent pour un ensemble de réseaux. Le « Lookup Service » est un mécanisme d'annuaire. Il est nécessaire pour localiser rapidement et dynamiquement l'information.

- Pour des raisons de confidentialité, l'accès aux données est soumis à une politique stricte de sécurité. Le service d'authentification et d'autorisation assure le mécanisme d'accès à l'information.

Ces familles de services ont été identifiées afin d'assurer la cohésion structurelle du système de supervision inter-domaines. Ces familles feront l'objet des sections suivantes.

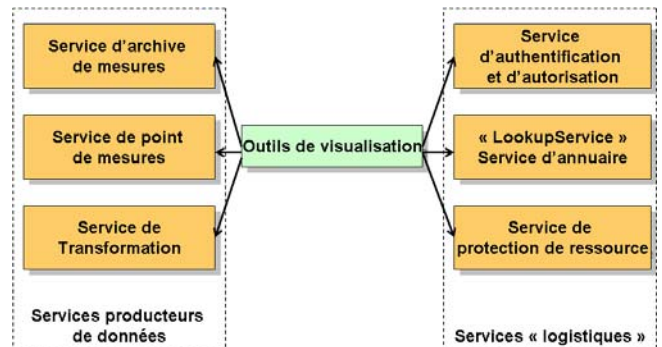


Figure 2 – Les types de service PerfSONAR

3.1 Le « Measurement Archive » : MA

Le rôle d'un « Measurement Archive » (MA) est de publier des informations provenant d'une structure de données déjà existante. C'est donc plutôt une interface entre celle-ci et les applications de la couche supérieure.

Un « MA » ne produit donc pas explicitement d'information et constitue un point d'entrée standardisé vers la source d'information. Cette source peut être une base de données RRD, une base SQL ou même un fichier plat.

Il peut toutefois publier des indicateurs provenant d'un ou plusieurs « MP » et les agréger afin d'avoir un tableau de bord commun entre domaines.

Prenons le cas d'un opérateur et d'un de ses clients désireux de mettre en commun la charge des liens réseaux. De part et d'autre, les données sont stockées avec des formats de données différents. Pour ce faire, l'opérateur et son client vont déployer, chacun, un type de MA publiant des informations similaires et accessibles de la même façon.

Ces 2 services, s'étant enregistrés au préalable auprès du bus applicatif, vont alimenter les outils de visualisation masquant ainsi toute la complexité de mise en œuvre propre à chaque domaine. La figure 3 illustre le concept de « MA » avec d'une part une base de données SQL et d'autre part une base RRD.

⁹ MOS : Mean Opinion Score. Indicateur destiné à quantifier la qualité d'une application audio.

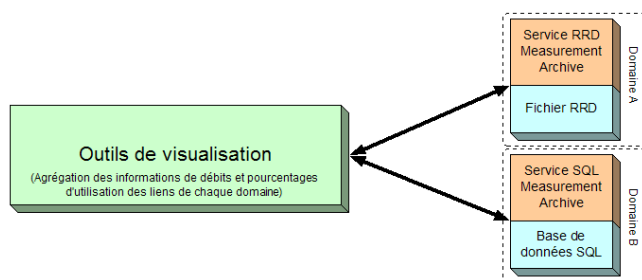


Figure 3 – Principe du "Measurement Archive"

3.2 Le « Measurement Point » : MP

Le « Measurement Point » (MP) est également un service web interagissant avec un outil fournissant des informations à la demande. Le MP permet de lancer des tests ou d'obtenir de l'information non stockée dans une archive. Si besoin est, ces informations peuvent être transmises au MA ou à la couche supérieure de visualisation.

Le « MP » communique donc avec ces outils qui sont capables de fournir des mesures actives comme les sondes Symmetricom™, HADES™ ou autres... D'autres sont aptes à fournir des indicateurs issus de mesures passives des « proxy » SNMP, NETFLOW ou TELNET/SSH.

Le « MP » va donc plutôt produire des données brutes qui ne sont pas déjà stockées ou déjà mesurées. De façon comparable au « MA », le « MP » présente la capacité de faire abstraction de l'outil de mesure.

En considérant toujours le cas d'un opérateur et de son client qui constate un problème de routage, toujours pour des raisons historiques, l'opérateur a choisi une gamme d'équipement provenant d'un fournisseur alors que son client en a choisi un autre. Si chacun des domaines installe alors son propre « TELNET/SSH MP », il est possible via un « Looking Glass » unifié d'obtenir l'information de routage sur les deux réseaux avec la même application et de distinguer comment le préfixe est routé au sein de chaque domaine de façon continue. La figure 4 illustre le principe de « MP ».

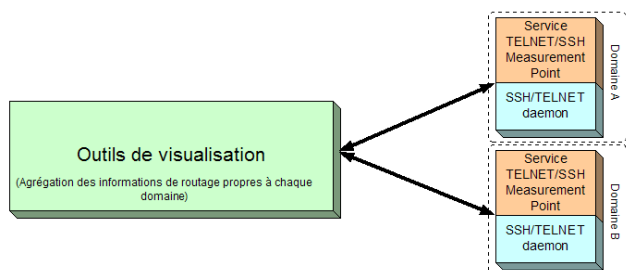


Figure 4 – Principe du "Measurement Point"

3.3 Le « Transformation Service » : TrS

Le concept de « Transformation Service » décrit la fonction permettant d'effectuer un ensemble d'opérations sur les données brutes provenant de « MA » ou de « MP » afin de dériver de nouveaux indicateurs, ou de générer des alarmes.

Ce service ne crée pas de nouvelles données brutes mais plutôt les transforme par le biais d'opérations de corrélation, de réajustement, de translation etc. Ce service est nécessaire dans le cas où un indicateur de qualité doit être associé spécifiquement à une application donnée.

Un niveau de charge processeur relatif à des équipements de même type n'a pas la même signification suivant leurs constructeurs respectifs. De la même manière, des données de base comme le délai, la gigue, le nombre de sauts peuvent être cumulés avec d'autres indicateurs afin de définir une perception de la qualité d'écoute d'une application de visioconférence. Le « TrS » répond au besoin de qualifier la qualité d'un service global réparti sur plusieurs domaines composés par des segments hétérogènes. Le « TrS » peut alors fournir des indicateurs globaux de bout en bout, que le trafic passe par un chemin « optique » ou « Ethernet » par exemple.

3.4 Le « Lookup Service » : LS

La notion de « bus applicatif » a souvent été évoquée précédemment dans cet article. Bien souvent, beaucoup d'applications servant à l'exploitation du réseau sont mises en œuvre. Cependant, il arrive que les exploitants ou les utilisateurs externes ignorent comment accéder à ces outils et parfois ne savent même pas qu'ils sont disponibles. Par exemple, beaucoup d'opérateurs mettent à disposition leur « Looking Glass », mais il reste alors aux utilisateurs à découvrir leur existence.

Le « Lookup Service ou LS » peut être perçu comme un annuaire référençant tous les services du domaine. Il assure la fonction de localisation de la ressource. Grâce au LS, il est possible d'adapter l'utilisation du système de supervision multi-domaine sans pour autant explicitement développer des passerelles logicielles additionnelles à chaque fois qu'un nouveau service est disponible.

Un « LS » est déployé au sein de chaque entité et tous les autres types de services s'enregistrent auprès de celui-ci afin de se signaler et de publier leurs fonctionnalités au sein du même domaine.

De la même manière, les « LS » respectifs de chaque domaine de responsabilité peuvent être interconnectés et ainsi diffuser les informations de manière transparente et dynamique sans pour autant reconfigurer les applications de visualisation. C'est cette configuration « multi-LS » qui assure la jonction inter-domaine.

Par ce biais, PerfSONAR a défini un jeu de messages standardisés permettant à chaque domaine membre de

publier ses ressources et ainsi étendre le périmètre de supervision de l'ensemble.

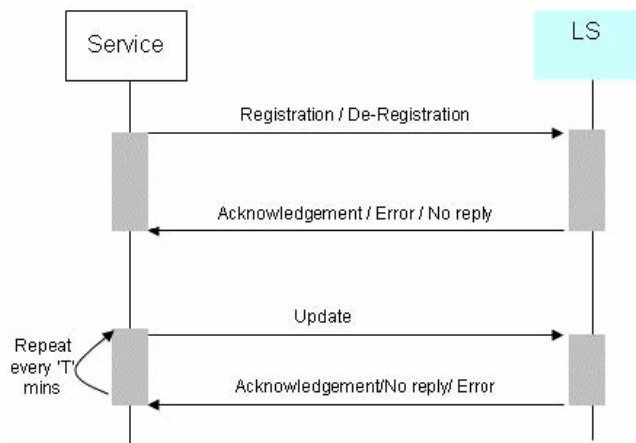


Figure 5 – Enregistrement d'un service auprès du "Lookup Service"

3.5 Le « Resource Protector Service » : RPS

Les services présentés ci-dessus n'avaient qu'un rôle de publication d'information. Le « RPS » est d'un autre type dont le rôle est de gérer la contention et la réservation de ressources limitées. Cette ressource peut correspondre à la bande passante disponible entre deux points du réseau ou à un certain nombre de requêtes simultanées qu'un équipement est capable de gérer en parallèle durant un intervalle de temps.

Un « MP » peut intrinsèquement inclure une fonction de « RPS » comme par exemple le « TELNET/SSH MP ». Ce dernier a la possibilité de limiter dans le temps le nombre de requêtes lancées sur un équipement.

Dans un service d'allocation de bande passante étendu sur plusieurs domaines, les « RPS » respectifs de chaque domaine contrôlent la charge des liens avant d'activer ou non le circuit au débit demandé.

3.6 Le « Topology Service » : TopS

Le « Topology Service » ou TopS est le service capable d'extraire du réseau les informations de topologie. Celui-ci est à même de fournir une vision de la topologie à tous les niveaux du modèle OSI. Il est donc possible par le biais du « TopS » d'avoir des informations sur l'infrastructure physique, mais aussi au niveau 3 par exemple.

Pour construire la topologie d'un ensemble de réseaux, le « TopS » peut interroger à la demande un ou plusieurs « MP ». Par la suite, les résultats sont agrégés : le « TopS » est un exemple de « TrS ».

Le « TopS » est essentiel pour comprendre la structure de du réseau de l'opérateur au quotidien. De plus, la métrologie associée à un réseau repose toujours sur une vision topologique dynamique.

3.7 Le service d'authentification et d'autorisation

Le service d'authentification et d'autorisation (AAS) est le composant commun à tous les services précédemment cités. En effet, ce service est transversal puisqu'il contrôle l'accès aux ressources d'un domaine à un autre.

Le principe de confiance entre domaines repose sur la politique de sécurité accessible par l'AAS. La politique de sécurité de PerfSONAR est totalement déléguée aux domaines. Cependant la méthode d'authentification peut être fédérée et hiérarchisée.

4 PerfSONAR : un état des lieux

4.1 L'état du projet

La version actuelle de PerfSONAR [5] est la version 2.1. Dans cette version, un ensemble de blocs élémentaires a été élaboré. Ces blocs sont composés de services web et d'outils de visualisation.

Plusieurs « Measurement Archives » capables de s'interfacer avec des bases de données SQL ou RRD ont été développés. Ces « MA » fournissent le débit et le taux de charge des liens d'un domaine. Un service d'export des indicateurs IPPM provenant des systèmes « HADES »[6] et « RIPE TTM »[7] a également été développé. Des « Measurements Points » permettant de réaliser des tests à la demande ont également été spécifiés : BWCTL MP (validation effective de bande passante), SSH/Telnet MP (exécution de commandes sur des routeurs via le protocole SSH ou TELNET), « TC MP » (capture de traces TCP à la demande).

Des annuaires de services (« Lookup Service »), de topologies (fournissant les informations de topologie réseau) et d'authentification (basé sur d'eduGAIN [8]) ont également été réalisés.

Différents outils de visualisation accédant aux données fournies par les « MA » et « MP » cités ci-dessus ont également été développés. Tout d'abord, PerfsonarUI [9] permet de visualiser la charge de tous les liens d'un ou plusieurs réseaux, d'effectuer des « traceroute », ou, à partir d'une liste d'adresses IP, d'indiquer la charge des liens entre ces adresses. Le « JRA1 Looking Glass » permet de lancer des commandes en ligne sur des équipements distribués dans plusieurs réseaux (pour autant que ceux-ci aient déployé un Telnet/SSH MP). Finalement l'outil « CNM » fournit une visualisation géographique de plusieurs réseaux ainsi que les informations produites par une trentaine de sondes « HADES » disséminées au sein de la communauté de la recherche européenne.

L'objectif des développements de PerfSONAR est de fournir un service de supervision multi-domaines à l'ensemble des « NRENs ». Les équipes cibles sont celles qui assurent les supports opérationnels, (« NOC », « PERT ») [10] mais aussi les équipes en charge des projets de grille par exemple. Ce service est nommé « Multi-Domain Monitoring » (MDM).

Une phase pilote du « MDM » s'étendra de mi-2007 à la fin 2007. Elle inclut six NRENs [11]. Ce pilote s'appuie sur des services web figurant dans la version PerfSONAR 2.0. Les objectifs du pilote sont de valider le concept du service « MDM » et la structure de support qui est mise en place. Cette structure de support vise à porter assistance aux réseaux déployant le service « MDM » ainsi qu'aux utilisateurs de ce service.

4.2 PerfSONAR et RENATER

Au sein de RENATER, l'installation de PerfSONAR 2.0 a été initiée. Ce déploiement a permis au projet GN2-JRA1 de bénéficier du retour d'expérience sur le « NREN » français.

La charge des liens de cœur de réseau est ainsi disponible via le « RRD-MA ». Chaque site ou réseau de collecte a donc la possibilité de consulter la charge des liens par lesquels son trafic transite. S'il le désire, il peut même être amené à partager ses informations en déployant son propre « MA » et avoir une supervision inter-domaines entre ses sites. La figure 6 illustre la possibilité de consulter la charge des liens de cœur du réseau RENATER via l'outil de visualisation PerfsonarUI [12].

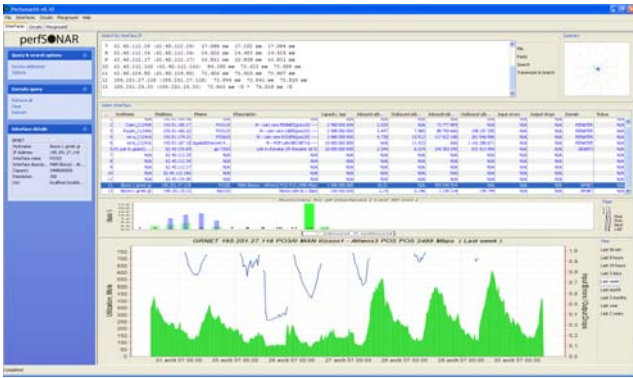


Figure 6 – Consultation de la charge de liens via PerfSONARUI

L'installation du « TELNET/SSH MP » est en cours. Son objectif est de mettre à disposition un « Looking Glass » unifié englobant l'ensemble des « NRENs ».

La mise en œuvre de PerfSONAR par RENATER est atypique dans le sens où les services sont lancés dans une machine virtuelle. Une image de celle-ci ou « VM¹⁰ » est disponible sur demande à l'ensemble de la communauté. Chaque établissement désireux de tester PerfSONAR peut récupérer cette image et mettre en œuvre rapidement sa solution de supervision inter-domaines.

4.3 Evolutions futures

Il est prévu de mettre en œuvre les deux derniers composants (le « Multi-Domain Lookup » et le service d'authentification) respectivement pour début 2008 et mi-2008. Cela achèvera le développement des blocs de base nécessaires à la mise en opération de PerfSONAR. Les outils de visualisation vont être adaptés spécialement pour

les membres des « NOCs » et des « PERTs » [10] de façon à répondre parfaitement aux besoins opérationnels.

Une seconde phase de prototypage du « MDM » démarrera début 2008. Elle regroupera 12 réseaux et validera le niveau de service qui peut être proposé aux utilisateurs du service « MDM ».

Au sein de RENATER, les services PerfSONAR seront continuellement mis à jour et testés afin de fournir un retour aux équipes de développement du projet GN2-JRA1. De plus, RENATER prévoit l'export au format PerfSONAR des données IPPM issues de son infrastructure de mesures actives.

5 Conclusion

PerfSONAR est un cadre permettant l'échange contrôlé d'informations de supervision à travers plusieurs réseaux. Son objectif est de rendre disponible aux opérateurs et aux utilisateurs une vision globale sur les performances d'un service bâti sur de multiples réseaux. Les blocs de base permettant à PerfSONAR d'être utilisé en environnement opérationnel seront disponibles mi-2008. Ces mêmes blocs seront réutilisés par les « NRENs » pour fournir un service de « Multi-Domain Monitoring » à plusieurs groupes d'utilisateurs. De façon analogue, les opérateurs virtuels du marché peuvent utiliser PerfSONAR afin de vérifier que la qualité de service assurée par des réseaux partenaires est contractuellement conforme au niveau de service négocié.

En tant que projet « Open Source », les services web développés peuvent être étendus par n'importe qui afin d'intégrer de nouvelles métriques ou de nouveaux outils.

RENATER continue de suivre le projet PerfSONAR et à le mettre à disposition de la communauté de la recherche française, afin de lui donner une vision de leurs services réseaux de bout en bout.

¹⁰ VM : Virtual Machine.

Références

- [1] Site officiel PerfSONAR : <http://www.perfsonar.net>
Wiki PerfSONAR : <http://wiki.perfsonar.net>
- [2] Site officiel GRID5000 : <https://www.grid5000.fr>
- [3] OGF : <http://www.ogf.org/>
OGF NMWG : <http://nmwg.internet2.edu/>
- [4] NETFLOW :
http://www.cisco.com/en/US/products/ps6601/products_ios_protocol_group_home.html
- [5] PerfSONAR bundle download :
<http://www.perfsonar.net/download.html>
- [6] HADES :
<http://kb.pert.geant2.net/PERTKB/HadesTool>
- [7] RIPE TTM : <http://www.ripe.net/ttm/>
- [8] EduGAIN : http://www.geant2.net/upload/pdf/GN2-07-023v4-DJ5-2-3_2_Best_Practice_Guide-AAI_Cookbook-Second_Edition.pdf
- [9] Site officiel PerfSONARUI :
<http://monstera.man.poznan.pl/jra1-wiki/index.php/PerfsonarUI>
- [10] PERT : <http://www.geant2.net/pert>
- [11] MDM : http://wiki.perfsonar.net/jra1-wiki/index.php/Multi-Domain_Monitoring_Service_Pilot
- [12] Console de visualisation PerfSONARUI :
<http://perfsonar.acad.bg/>

Glossaire

AAS : Authentication and Authorization Service
BWCTL: Bandwidth Test Controller
CMIP : Common Managed Information Service Protocol
CNM : Customer Network Management
DANTE : Delivery of Advanced Network Technology to Europe
DEISA : Distributed European Infrastructure for Supercomputing Applications
EGEE : Enabling Grids for E-science
FCAPS : Fault, Configuration, Accounting, Performance and Security management
GN2-JRA1 : GEANT2 Joint Research Activity 1
GPLv2 : General Public Licence Version 2
HADES : Hades Active Delay Evaluation System
IPPM : Internet Protocol Performance Metrics
LHC : Large Hadron Collider

LS : Lookup Service
MA : Measurement Archive
MDM : Multi-Domain Monitoring
MOS : Mean Opinion Score
MP : Measurement Point
VNO : Virtual Network Operator
NMS : Network Management System
NOC : Network Operation Center
NREN : National Research and Education Network
OGF-NMWG : Open Grid Forum Network Management Working Group
PERT : Performance Enhanced Response Team
RENATER : Réseau National de Télécommunications pour la Technologie de l'Enseignement et de la Recherche
RIPE TTM : RIPE Test Traffic Measurement
RPS : Resource Protector Service
RRD : Round Robin Database
SNMP : Simple Network Management Protocol
SOAP : Simple Object Access Protocol
SQL : Structured Query Language
SSH : Secure Shell
TC MP : Trace Capture Measurement Point
TCP : Transport Control Protocol
TopS : Topology Service
TrS : Transformation Service
WAN : Wide Area Network
WDM : Wavelength Division Multiplexing
XML : Extensible Markup Language